Journal of Global Antimicrobial Resistance 8 (2017) 13–17

Contents lists available at ScienceDirect

Journal of Global Antimicrobial Resistance

journal homepage: www.elsevier.com/locate/jgar

InverPep: A database of invertebrate antimicrobial peptides

Esteban A. Gómez*, Paula Giraldo, Sergio Orduz

Research Group Biología Funcional, Escuela de Biociencias, Facultad de Ciencias, Universidad Nacional de Colombia sede Medellín, Calle 59A No. 63-20, Medellín, Colombia

ARTICLE INFO

Article history:
Received 14 September 2016
Accepted 11 October 2016
Available online 19 November 2016

Keywords:
Invertebrate
Antimicrobial peptide
Web database
Bioinformatics

ABSTRACT

Objectives: The aim of this work was to construct InverPep, a database specialised in experimentally validated antimicrobial peptides (AMPs) from invertebrates.
Methods: AMP data contained in InverPep were manually curated from other databases and the scientific literature. MySQL was integrated with the development platform Laravel; this framework allows to integrate programming in PHP with HTML and was used to design the InverPep web page's interface. InverPep contains 18 separated fields, including InverPep code, phylum and species source, peptide name, sequence, peptide length, secondary structure, molar mass, charge, isoelectric point, hydrophobicity, Boman index, aliphatic index and percentage of hydrophobic amino acids. CALCAMPI, an algorithm to calculate the physicochemical properties of multiple peptides simultaneously, was programmed in PERL language.
Results: To date, InverPep contains 702 experimentally validated AMPs from invertebrate species. All of the peptides contain information associated with their source, physicochemical properties, secondary structure, biological activity and links to external literature. Most AMPs in InverPep have a length between 10 and 50 amino acids, a positive charge, a Boman index between 0 and 2 kcal/mol, and 30–50% hydrophobic amino acids. InverPep includes 33 AMPs not reported in other databases. Besides, CALCAMPI and statistical analysis of InverPep data is presented. The InverPep database is available in English and Spanish.
Conclusions: InverPep is a useful database to study invertebrate AMPs and its information could be used for the design of new peptides. The user-friendly interface of InverPep and its information can be freely accessed via a web-based browser at http://ciencias.medellin.unal.edu.co/gruposdeinvestigacion/prospeccionydisenobiomoleculas/InverPep/public/home_en.

1. Introduction

Since the identification of the cecropins, magainins and defensins in insects, amphibians and humans, respectively, in the 1980s [1–3], the antimicrobial peptides (AMPs) have achieved scientific and medical relevance. AMPs are effectors of the innate immune system of all organisms [4]. They are small molecules (normally 10–50 amino acids) that are easy to synthesise, with fast and effective action against different micro-organisms such as bacteria, fungi, parasites and viruses [5], and despite their structural diversity they are evolutionarily conserved [6].

During the last 30 years, AMPs have been considered as a potential source for the development of new therapeutic molecules to control infectious diseases owing to their significant specificity for micro-organisms with low toxicity for mammal cells [7,8]. For these reasons, thousands of AMPs have been isolated from plants, animals, fungi and several micro-organisms [9,10], obtained through a conventional chemistry approach [11,12] or designed using AMPs as models to produce peptides with improved selectivity and potency [13,14]. This expanse of data has been compiled in approximately 23 specialised databases that facilitate extraction of information and provide bioinformatics tools to rationalise the design of new AMPs [15].

Approximately 75% of AMPs contained in the different databases are of animal origin [16], among which invertebrates represent over 80% of the Animal Kingdom [17]. AMPs from invertebrates have been extensively studied with particular interest in marine invertebrates in the context of bioprospecting research for natural products [18] and as a result of insect immunology studies [19]. Likewise, the value of invertebrate AMP research is not only focused on their capacity to kill micro-organisms but also on their potential as insecticides, as in the case

* Corresponding author.
E-mail address: eagomezc@unal.edu.co (E.A. Gómez).

http://dx.doi.org/10.1016/j.jgar.2016.10.003
2213-7165/

of peptides found in toxins and venoms of insects and their natural enemies [20]. As an outcome of these studies, various groups of AMPs have been described, such as the cecropins and defensins in insects [21], the clavanins in tunicates [22], the penaeidins and crustins in crustaceans [18,23] and the caenopores in nematodes [24].

Nevertheless, and despite the potential of AMPs and the existence of many databases, such as CAMP [25], APD [16,26], YADAMP [27] and DRAMP [28], there is no available database specialised in AMPs from invertebrates, like there are for plants (PhytAMP) [29], bacteria (BACTIBASE) [30], amphibians (DADP) [31] or dairy origin (MilkAMP) [32]; and the penaeidins database PenBase is currently inactive [33]. Therefore, we have created InverPep, a database specialised in AMPs from invertebrates. InverPep is a manually curated database that contains information on experimentally validated peptides extracted from databases and the scientific literature. InverPep also contains an in-house developed algorithm to calculate the physicochemical properties of multiple peptides simultaneously.

2. Methods

2.1. Data collection

Data in InverPep were manually collected and curated. The information was mainly extracted from AMP databases such as APD, CAMP and YADAMP as well as from the scientific literature. A combination of different words in PubMed was used, such as 'invertebrate', 'arthropods', 'nematode', 'peptide', 'antimicrobial', 'antibacterial', 'antifungal', 'insecticide', etc. All peptides in InverPep have been experimentally validated and corroborated directly from their original scientific papers. The predicted peptides without experimental validation, peptide combinations, or articles without the peptide's amino acid sequence were excluded. InverPep was checked manually to delete redundancies and to group the peptides by source (phylum and species).

InverPep contains 18 separated fields, including InverPep code, phylum and species source, peptide name, sequence, peptide length, secondary structure, molar mass, charge, isoelectric point (pI), hydrophobicity, Boman index, aliphatic index and percentage of hydrophobic amino acids. All of the physicochemical properties were calculated using CALCAMPI, an algorithm designed for that purpose. AMPs in the InverPep database also contain their antimicrobial activity, experimental verification and links to the external literature. It is possible to perform an advanced search in InverPep by using different fields.

2.2. Database architecture

For management and operation of the database, MySQL v.5.6.17 was used. MySQL was integrated with the development platform Laravel v.4.2; this framework allows to integrate programming in PHP v.5.5.12 with HTML code, and was used to design the InverPep web page's interface.

For correct operation of Laravel, Composer v.1.0-dev was used, a manager of PHP's packages. This facilitates management of PHP dependencies and the use of necessary libraries. Git v.1.9.5 was also used to facilitate handling of the final application.

2.3. Web interface

An easy and user-friendly interface was designed (Fig. 1). A brief description of the pages of InverPep is listed is below:

- Home: brief description of the database.
- Peptides: visualisation of the AMP database. Allows basic searches (search by phylum source and antimicrobial activity) and an advanced search using other criteria.
- Tools: at present, CALCAMPI is the only available tool. Additional information of CALCAMPI is found in Section 2.4.
- Statistics: statistical information (histograms and pie charts) of the InverPep database.

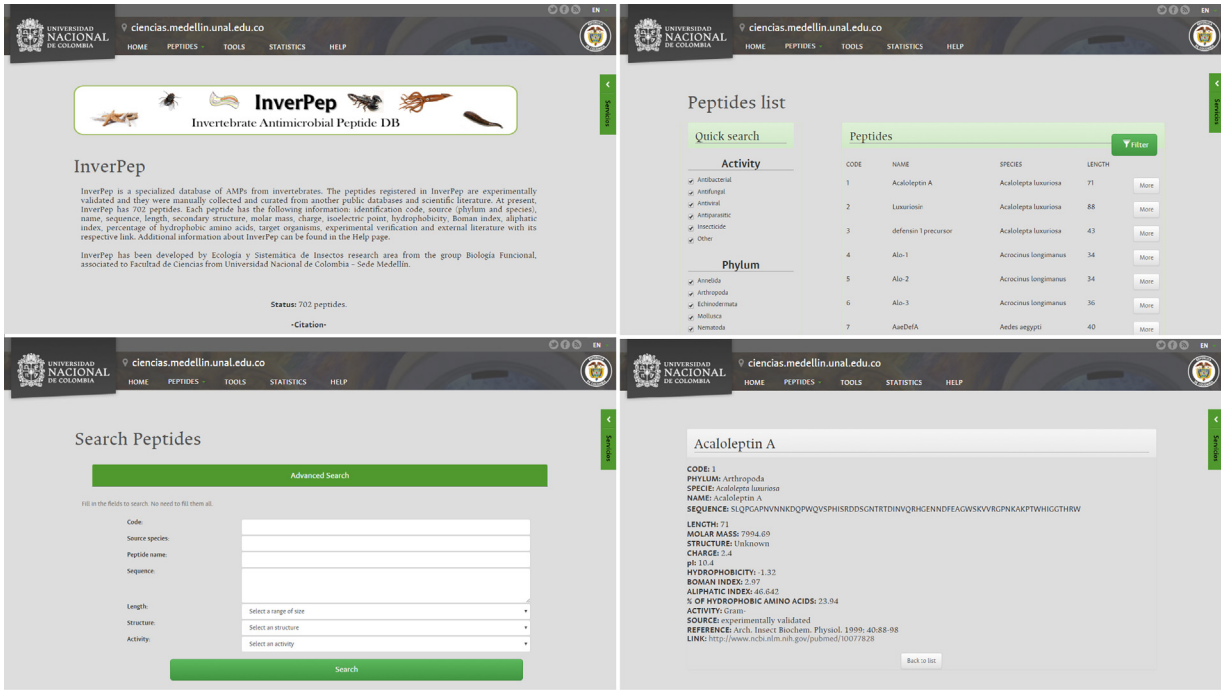

Fig. 1. Screenshots of the InverPep database.

- Help: general information on the web site, how the information was collected, CALCAMPI's description, a glossary, external links to other useful information on the web, an option for the users to submit new peptide sequences, and contact information about the creators and contributors of InverPep.

2.4. CALCAMPI design

CALCAMPI is a physicochemical properties calculator of multiple peptides at the same time. It was designed using the programming language Perl v.5.22.0.1 and was later integrated into the InverPep web interface. CALCAMPI calculates molar mass, charge, hydrophobicity, Boman index, aliphatic index, pI and percentage of hydrophobic amino acids.

Charge is calculated with the following equation:

$$\text{Charge} = \sum_i Ni\left(\frac{10^{pKai}}{10^{pH}+10^{pKai}}\right) - \sum_j Nj\left(\frac{10^{pH}}{10^{pH}+10^{pKaj}}\right)$$

where *pKai* is the pK_a value and *Ni* is the number of the N-terminus and the side chains of arginine, lysine and histidine. *Nj* and *pKaj* are the number and pK_a value, respectively, of the C-terminus and the aspartic acid, glutamic acid, cysteine and tyrosine amino acids. The pK_a values used to design the algorithm were taken from Nelson and Cox [34].

Hydrophobicity is represented in terms of the grand average of hydrophobicity. It is calculated as the sum of the hydropathic values of each amino acid (values taken from Kyte–Doolittle scale [35]) and divided by the total number of residues. Positive values indicate hydrophobic peptides and negative values indicate hydrophilic peptides.

The percentage of hydrophobic amino acids is calculated as the sum of the hydrophobic residues (Kyte–Doolittle scale) of the peptide divided by the total number of residues.

The Boman index is the sum of the free energies of the respective peptide side chains for transfer from cyclohexane to water [36] and divided by the total number of residues of the peptide. The Boman index can be seen as the potential of a peptide binding to a protein or to a membrane. Hydrophobic peptides tend to have a Boman index with negative values, whilst hydrophilic peptides have positive values.

The aliphatic index is the relative volume occupied by the aliphatic side chains of amino acids alanine, valine, isoleucine and leucine. Positive values are associated with thermostability of the peptide [37]. It is calculated using the following equation:

$$\text{Aliphatic index} = X_A + aX_V + b(X_I + X_L)$$

where X_A,X_V, X_I and X_L are the mole percent (100 per mol fraction) of alanine, valine, isoleucine and leucine; a = 2.9 and b = 3.9 are standard coefficients [38].

The pI is the pH where the peptide charge is zero. The pI was determined using an in-house developed algorithm that calculates all the possible charges when the pH is changed and the program stops when it finds the pH value with the peptide charge equal to zero.

2.5. Statistics

Statistical analysis of the InverPep data shows the variability of the peptides according to different properties such as length, percentage of hydrophobic residues, charge, pI, Boman index and aliphatic index. Besides, the percentage of peptides according to its source and antimicrobial activity was also assessed. These calculations and their respective histograms and pie charts were prepared in Microsoft Excel 2010 (Microsoft Corp., Redmond, WA).

3. Results

The InverPep database is contained in a user-friendly web-based browser available in English and Spanish, providing access to the international scientific community and to Hispanic research groups interested in this topic.

InverPep has at the moment (11 May 2016) 702 AMPs experimentally validated from different species of invertebrates of the phyla Arthropoda, Mollusca, Nematoda, Annelida, Echinodermata, Platyhelminthes, Placozoa, the Hydridae family (Cnidaria) and the subphylum Tunicata (Chordate). All of the peptides have their general information (phylum and species source, name and length) and physicochemical properties (charge, hydrophobicity, Boman index, aliphatic index, pI, percentage of hydrophobic amino acids and secondary structure); the target organism and an external link to the scientific literature are also included. There are 33 peptides in InverPep that are not reported in other actualised databases such as DRAMP [28], the last database published. Three of the peptides were discovered in Hydridae, trichoplaxin from *Trichoplax adhaerens* (Placozoa) and the rest recently discovered in Arthropoda; most of them have experimentally validated antibacterial activity.

The principal source of invertebrate AMPs are arthropods, accounting for >80% of the peptides, followed by molluscs with ca. 7%. The rest of the phyla correspond to 10% of the InverPep peptides (Fig. 2). Statistical analysis of InverPep can be seen in Fig. 3.

Peptides in the InverPep database have different lengths and physiochemical properties, which are associated with the AMP classes present in invertebrates [18,21–24]. Yet certain tendencies can be observed in the database; most of the AMPs have a length between 10 and 50 amino acids, a positive charge, a Boman index between 0 and 2 kcal/mol, 30–50% of their primary structure consists of hydrophobic residues, and their pIs are normally at alkaline pH. These properties have been already identified in cationic AMPs [39].

Invertebrates do not have an adaptive immune system but have a strong and effective innate immune system to protect themselves from pathogenic micro-organisms. The diversity of invertebrate AMPs can be attributed to the lack of adaptive immunity and to the strong innate immunity [40]. Invertebrates have a wide range of peptides to fight pathogens; this can be seen in the statistical analysis of the antimicrobial activity of InverPep (Fig. 4), where antibacterial, antifungal, antiviral and antiparasitic activities are observed and, moreover, some peptides display combinations of targets that include more than one group of pathogens.

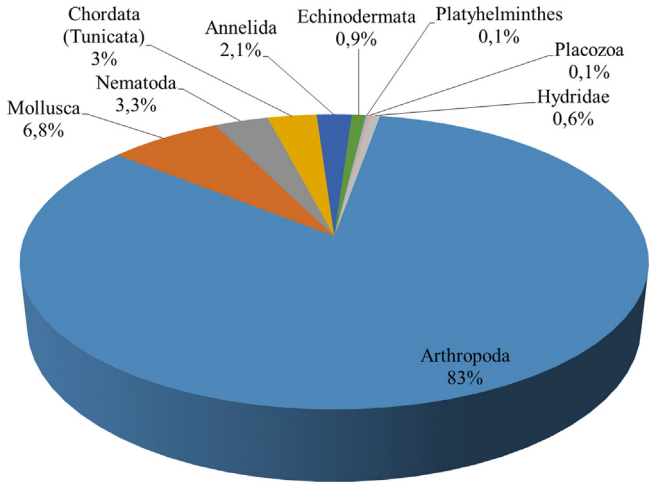

Fig. 2. Pie chart of phylum source of InverPep peptides.

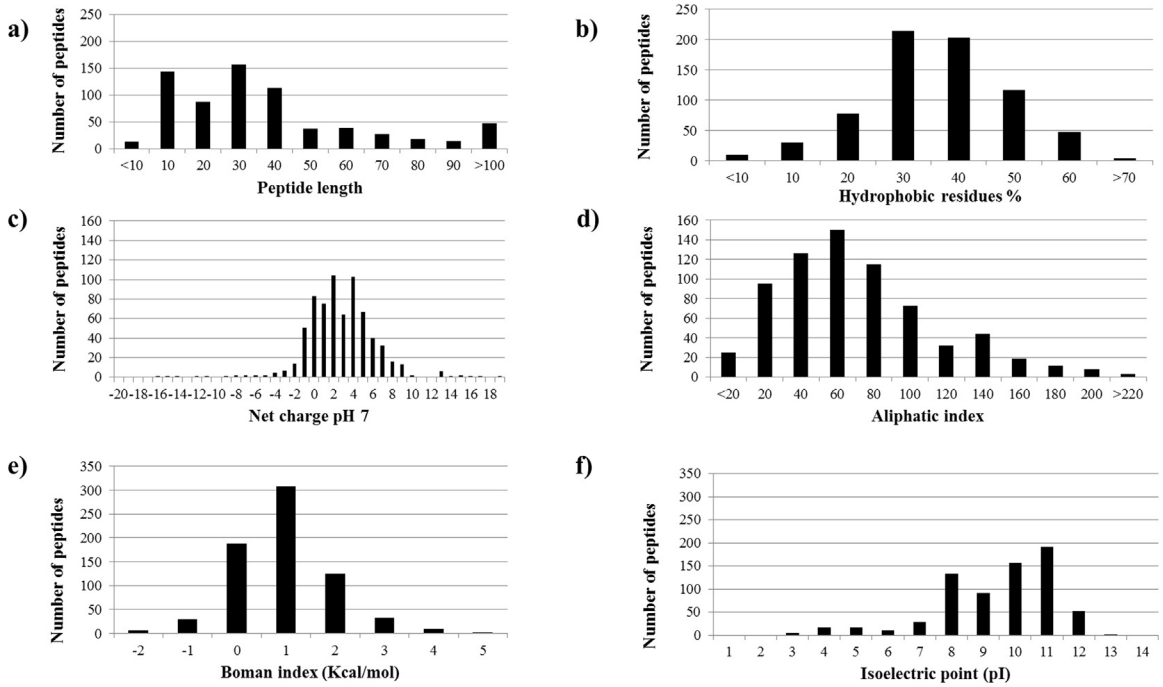

Fig. 3. Statistical analysis of InverPep. Histogram of (a) peptide length distribution, (b) percentage of hydrophobic residues, (c) charge, (d) aliphatic index, (e) Boman index and (f) isoelectric point.

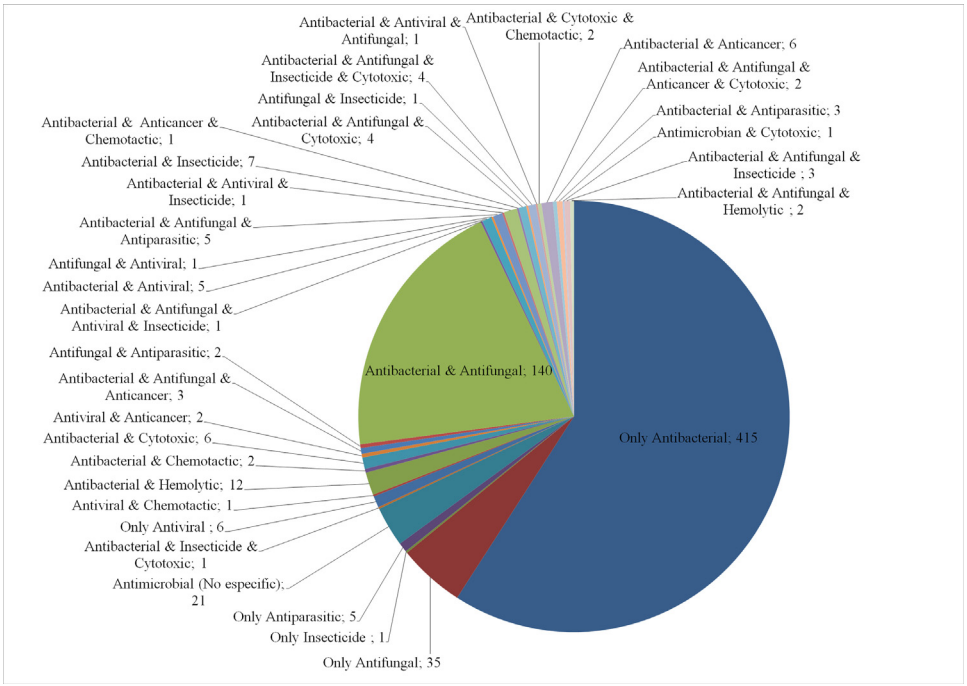

Fig. 4. Pie chart of target organisms of InverPep peptides.

4. Conclusion

AMPs are considered as effective therapeutic drug candidates. They are continually studied and new sequences are constantly discovered. Invertebrates, as the biggest group of the Animal Kingdom, are the principal source of AMPs with validated antimicrobial activity, yet there is no study focused on the physiochemical properties of these peptides. Knowledge of these properties could help in the design of new AMPs.

InverPep is a database that unites useful data of invertebrate AMPs such as source, properties and antimicrobial activity, which facilitates an understanding of their mechanisms of action; besides, InverPep provides an organised user-friendly source of information (available in English and Spanish) that can be used as a reference bibliography. This could support not only studies of AMPs but could also help in the bioprospecting of these molecules in different biotechnology fields, such as medical drug design and agricultural application.

InverPep includes CALCAMPI, an algorithm to simultaneously calculate the physicochemical properties of multiple peptides; this algorithm could be useful in identifying potential AMPs based on their primary structure.

InverPep will be updated every 6 months to provide a valuable resource, with new peptides not reported in other databases, for research in invertebrate AMPs.

InverPep is freely accessible at the following link: http://ciencias.medellin.unal.edu.co/gruposdeinvestigacion/prospeccionydisenobiomoleculas/InverPep/public/home_en.

Funding

Universidad Nacional de Colombia sede Medellín (Medellín, Colombia).

Competing interest

None declared.

Ethical approval

Not required.

Acknowledgments

The authors thank Juan David Cadavid, Ricardo Mesa and Michael Agudelo for their help in making the InverPep database available via the Internet and for creating the interface of the web page.

References

[1] Hultmark D, Steiner H, Rarmuson T, Boman HG. Insect immunity. Purification and properties of three inducible bactericidal proteins from hemolymph of immunized pupae of *Hyalophora cecropia*. Eur J Biochem 1980;106:7–16.
[2] Zasloff M. Magainins, a class of antimicrobial peptides from Xenopus skin: isolation, characterization of two active forms, and partial cDNA sequence of a precursor. Proc Natl Acad Sci U S A 1987;84:5449–53.
[3] Stanfield R, Wetbrook E, Selsted M. Characterization of two crystal forms of human defensin neutrophil cationic peptide 1, a naturally occurring antimicrobial peptide of leukocytes. J Biol Chem 1988;263:5933–5.
[4] Hancock REW, Diamond G. The role of cationic antimicrobial peptides in innate host defences. Trends Microbiol 2000;8:402–10.
[5] Van't Hof W, Veerman E, Helmerhorst E, Amerongen A. Antimicrobial peptides: properties and applicability. Biol Chem 2001;382:597–619.
[6] Schmitt P, Rosa RD, Destoumieux-Garzón D. An intimate link between antimicrobial peptide sequence diversity and binding to essential components of bacterial membranes. Biochim Biophys Acta 2016;1858:958–70.
[7] Hawrani A, Howe RA, Walsh TR, Dempsey CE. Origin of low mammalian cell toxicity in a class of highly active antimicrobial amphipathic helical peptides. J Biol Chem 2008;283:18636–45.
[8] Bobone S, Bocchinfuso G, Park Y, Palleschi A, Hahm KS, Stella L. The importance of being kinked: role of Pro residues in the selectivity of the helical antimicrobial peptide P5. J Pept Sci 2013;19:758–69.
[9] Wang G, Li X, Zasloff M. A database view of naturally occurring antimicrobial peptides: nomenclature, classification and amino acid sequence analysis. In: Wang G, editor. Antimicrobial peptides: discovery, design and novel therapeutic strategies. Wallingford, UK: CAB International; 2010. p. 1–21.
[10] Phoenix D, Dennison S, Harris F. Antimicrobial peptides. 1st ed. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co.; 2013.
[11] Wang JX, Zhao XF, Liang YL, Li L, Zhang W, Ren Q, et al. Molecular characterization and expression of the antimicrobial peptide defensin from the housefly (*Musca domestica*). Cell Mol Life Sci 2006;63:3072–82.
[12] Zhao H, Kong Y, Wang H, Yan T, Feng F, Bian J, et al. A defensin-like antimicrobial peptide from the venoms of spider, *Ornithoctonus hainana*. J Pept Sci 2011;17:540–4.
[13] Torcato IM, Huang YH, Franquelim HG, Gaspar D, Craik DJ, Castanho MA, et al. Design and characterization of novel antimicrobial peptides, R-BP100 and RW-BP100, with activity against Gram-negative and Gram-positive bacteria. Biochim Biophys Acta 2013;1828:944–55.
[14] Dong N, Zhu X, Chou S, Shan A, Li W, Jiang J. Antimicrobial potency and selectivity of simplified symmetric-end peptides. Biomaterials 2014;35:8028–39.
[15] Wang G. Database resources dedicated to antimicrobial peptides. In: Chen CY, Yan X, Jackson CR, editors. Antimicrobial resistance and food safety. San Diego, CA: Academic Press; 2015. p. 365–84.
[16] Wang Z, Wang G. APD: the antimicrobial peptide database. Nucleic Acids Res 2004;32(Database issue):D590–2.
[17] Salzet M. Neuropeptide-derived antimicrobial peptides from invertebrates for biomedical applications. Curr Med Chem 2005;12:3055–61.
[18] Otero-González AJ, Magalhães BS, Garcia-Villarino M, López-Abarrategui C, Sousa DA, Dias SC, et al. Antimicrobial peptides from marine invertebrates as a new frontier for microbial infection control. FASEB J 2010;24:1320–34.
[19] Tang T, Li X, Yang X, Yu X, Wang J, Liu F, et al. Transcriptional response of *Musca domestica* larvae to bacterial infection. PLoS One 2014;9:e104867.
[20] Vassilevski AA, Kozlov SA, Samsonova OV, Egorova NS, Karpunin DV, Pluzhnikov KA, et al. Cyto-insectotoxins, a novel class of cytolytic and insecticidal peptides from spider venom. Biochem J 2008;411:687–96.
[21] Vizioli J, Salzet M. Antimicrobial peptides from animals: focus on invertebrates. Trends Pharmacol Sci 2002;23:494–6.
[22] Lee IH, Zhao C, Cho Y, Harwig SS, Cooper EL, Lehrer RI. Clavanins, α-helical antimicrobial peptides from tunicate hemocytes. FEBS Lett 1997;400:158–62.
[23] Ghosh J, Lun CM, Majeske AJ, Sacchi S, Schrankel CS, Smith LC. Invertebrate immune diversity. Dev Comp Immunol 2011;35:959–74.
[24] Roeder T, Stanisak M, Gelhaus C, Bruchhaus I, Grötzinger J, Leippe M. Caenopores are antimicrobial peptides in the nematode *Caenorhabditis elegans* instrumental in nutrition and immunity. Dev Comp Immunol 2010;34:203–9.
[25] Waghu FH, Gopi L, Barai RS, Ramteke P, Nizami B, Idicula-Thomas S. CAMP: collection of sequences and structures of antimicrobial peptides. Nucleic Acids Res 2014;42(Database issue):D1154–8.
[26] Wang G, Li X, Wang Z, Wang G. APD3: the antimicrobial peptide database as a tool for research and education. Nucleic Acids Res 2016;37(Database issue):D590–2.
[27] Piotto SP, Sessa L, Concilio S, Iannelli P. YADAMP: yet another database of antimicrobial peptides. Int J Antimicrob Agents 2012;39:346–51.
[28] Fan L, Sun J, Zhou M, Zhou J, Lao X, Zheng H, et al. DRAMP: a comprehensive data repository of antimicrobial peptides. Sci Rep 2016;6:24482.
[29] Hammami R, Ben Hamida J, Vergoten G, Fliss I. PhytAMP: a database dedicated to antimicrobial plant peptides. Nucleic Acids Res 2009;37(Database issue):D963–8.
[30] Hammami R, Zouhir A, Ben Hamida J, Fliss I. BACTIBASE: a new web-accessible database for bacteriocin characterization. BMC Microbiol 2007;7:89.
[31] Novković M, Simunić J, Bojović V, Tossi A, Juretić D. DADP: the database of anuran defense peptides. Bioinformatics 2012;28:1406–7.
[32] Théolier J, Fliss I, Jean J, Hammami R. MilkAMP: a comprehensive database of antimicrobial peptides of dairy origin. Dairy Sci Technol 2013;94:181–93.
[33] Gueguen Y, Garnier J, Robert L, Lefranc MP, Mougenot I, de Lorgeril J, et al. PenBase, the shrimp antimicrobial peptide penaeidin database: sequence-based classification and recommended nomenclature. Dev Comp Immunol 2006;30:283–8.
[34] Nelson D, Cox M. Lehninger principles of biochemistry. 5th ed. New York, NY: W. H. Freeman; 2008.
[35] Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol 1982;157:105–32.
[36] Radzicka A, Wolfenden R. Comparing the polarities of the amino acids: side-chain distribution coefficients between the vapor phase, cyclohexane, 1-octanol, and neutral aqueous solution. Biochemistry 1988;27:1664–70.
[37] Ikai A. Thermostability and aliphatic index of globular proteins. J Biochem 1980;88:1895–8.
[38] Richards FM. Areas, volumes, packing, and protein structure. Annu Rev Biophys Bioeng 1977;6:151–76.
[39] Phoenix D, Dennison S, Harris F. Antimicrobial peptides. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA; 2013.
[40] Tassanakajon A, Somboonwiwat K, Amparyup P. Sequence diversity and evolution of antimicrobial peptides in invertebrates. Dev Comp Immunol 2015;48:324–41.