

Short communication

Theoretical structural insights into the snakin/GASA family

William F. Porto, Octavio L. Franco*

Centro de Análises Proteômicas e Bioquímicas, Pós-Graduação em Ciências Genômicas e Biotecnologia, Universidade Católica de Brasília, Brasília-DF, Brazil¹

ARTICLE INFO

Article history:

Received 25 January 2013

Received in revised form 14 March 2013

Accepted 14 March 2013

Available online 8 April 2013

ABSTRACT

Among the main classes of cysteine-stabilized antimicrobial peptides, the snakin/GASA family has not yet had any structural characterization. Through the combination of *ab initio* and comparative modeling with a disulfide bond predictor, the three-dimensional structure prediction of snakin-1 is reported here. The structure was composed of two long α -helices with a disulfide pattern of Cys^I-Cys^{IX}, Cys^{II}-Cys^{VII}, Cys^{III}-Cys^{IV}, Cys^V-Cys^{XI}, Cys^{VI}-Cys^{XII} and Cys^{VIII}-Cys^X. The overall structure was maintained throughout molecular dynamics simulation. Snakin-1 showed a small degree of structural similarity with thionins and α -helical hairpins. This is the first report of snakin-1 structural characterization, shedding some light on the snakin/GASA family.

© 2013 Elsevier Inc. Open access under the [Elsevier OA license](http://creativecommons.org/licenses/by-nc-sa/4.0/).

1. Introduction

The first plant antimicrobial peptides (AMPs) were reported in 1942 in a manuscript describing the purothionins isolated from wheat (*Triticum aestivum*) [6]. After more than a half century, over 200 plant AMPs have been described [6]. These compounds have been recognized as playing a pivotal role in plant defense mechanisms against microorganisms [6,22]. Thus, numerous studies about their structure–activity relationship have been carried out [6,22].

The majority of plant AMPs are cysteine-rich [6,22,31], with few examples of plant disulfide-free AMPs [17,18,23,30,32]. The disulfide-free peptides are composed mainly of α -helical and unstructured folding; while the cysteine-stabilized AMPs are composed of several classes, which are divided according to their structural scaffolds and disulfide patterns [26]. The main plant cysteine-stabilized AMP classes are thionins [11,28], defensins [7,36], cyclotides [24,25], hevein-like peptides [4,27], α -helical hairpins [20,21] and snakins [3,29].

Among plant cysteine-stabilized AMP classes, the snakin has not had any structural characterization so far. The first snakin, snakin-1, was isolated from potato plants (*Solanum tuberosum*) [29]. This peptide class shows clear similarity with members of the GAST (giberellic acid stimulated transcript) and GASA (giberellic acid stimulated in *Arabidopsis*) protein families from *Arabidopsis*. In this conjuncture, both have been classified as members of the snakin/GASA family [3,22].

Mature snakin-1, from potatoes, is composed of 63 amino acid residues including 12 cysteine ones, which are involved in the formation of six disulfide bonds [29]. Nevertheless, no information about the three-dimensional structure or their cysteine bonding pattern has been provided until now. The lack of structural confirmation of plant bactericidal peptides prevents more detailed classification of plant AMPs [6,22]. Furthermore, this structural knowledge can help us to avoid errors in AMP classification as was observed for plant defensins, which were classified as a subclass of thionins before their structural characterization [6,22]. Bearing this in mind, this paper describes the prediction of the three-dimensional structure of snakin-1 through the combination of *ab initio* and comparative molecular modeling together with a disulfide bond predictor.

2. Materials and methods

2.1. Sequence clustering and alignment

The snakin-1 sequence was taken from the UniProt database (UniProt: Q948Z4) and the mature sequence was extracted according to the annotation (residues 26–88). The mature sequence was used as a seed for searching against UniProt, through PHI-BLAST [1] and the pattern “CX{3}CX{3}CX{7,11}CX{3}CX{2}CCX{2}CX{1,3}CX{11}CX{1,2}CX{11,14}KCP” [31], where ‘X’ indicates a wild card, which can be filled up by any of 20 natural amino acid residues, and the numbers between brackets indicate the number of repetitions of the prior character (i.e. ‘X{7,11}’ means that ‘X’ can be repeated seven to eleven times). The mature sequences from retrieved sequences were taken according to the annotation. The multiple sequence alignment was done in ClustalW 2 [33].

* Corresponding author. Tel.: +55 61 34487167/34487220; fax: +55 61 33474797.
E-mail address: ocfranco@gmail.com (O.L. Franco).

¹ <http://www.capb.com.br>.

2.2. Molecular modeling

The snakin-1 mature sequence was submitted to the QUARK *ab initio* molecular modeling server [35] in order to create an initial model. Then the cysteine connectivity was predicted as follows: the cysteine residues involved in disulfide bonds in the initial model were replaced by serine residues and then this modified sequence was submitted to the DiANNA 1.1 server [10], in order to predict the remaining cysteine pairs. The final model was constructed with MODELLER 9.10 [9]. The *ab initio* model was used as a template and the disulfide bonds were included using the method patch from the automodel class. Thus, 100 molecular models were constructed, and the final model was selected according to the discrete optimized protein energy (DOPE) scores. This score assesses the energy of the model and indicates the most probable structures. Subsequently, an energy minimization with 2000 steps of the steepest descent using the GROMOS96 implementation of Swiss PDB Viewer (SPDBV) [12] was done in order to adjust the distances between the sulfur atoms to 2 Å. The minimized model was evaluated through Verify 3D [16], ProSA II [34] and PROCHECK [15]. PROCHECK checks the stereochemical quality of a protein structure, through the Ramachandran plot, where reliable models are expected to have more than 90% of the amino acid residues in the most favored and allowed regions, while ProSA II indicates the fold quality; additionally, Verify 3D analyzed the compatibility of an atomic model (3D) with its own amino acid sequence (1D). Structure visualization was done in PyMOL (The PyMOL Molecular Graphics System, Version 1.4.1, Schrödinger, LLC).

2.3. Molecular dynamics

The molecular dynamics simulation (MD) was carried out in a water environment, using the Single Point Charge water model [2]. The analyses were performed by using the GROMOS96 43A1 force field and the computational package GROMACS 4 [14]. The dynamics used the three-dimensional model of snakin-1 as initial structure, immersed in water in a cubic box with a minimum distance of 0.5 nm between the complexes and the edges of the box. Chlorine ions were added in order to neutralize the system charge. The geometry of water molecules was constrained by using the SETTLE algorithm [19]. All atom bond lengths were linked by using the LINCS algorithm [13]. Electrostatic corrections were made by Particle Mesh Ewald algorithm [8], with a cut-off radius of 1.4 nm in order to minimize the computational time. The same cut-off radius was also used for van der Waals interactions. The list of neighbors of each atom was updated every 10 simulation steps of 2 fs. The system underwent an energy minimization using 50,000 steps of the steepest descent algorithm. After that, the system temperature was normalized to 300 K for 100 ps, using the velocity rescaling thermostat (NVT ensemble). Next, the system pressure was normalized to 1 bar for 100 ps, using the Parrinello–Rahman barostat (NPT ensemble). The systems with minimized energy, balanced temperature and pressure were simulated for 50 ns by using the leap-frog algorithm. The trajectories were evaluated through RMSD and DSSP. The initial and the final structures were compared through the TM-Score [37], where structures with TM-Scores above 0.5 indicate that the structures share the same fold.

3. Results

The peptide snakin-1 was selected as a prototype for the snakin/GASA family (Fig. 1). The prediction of snakin-1 three-dimensional structure and disulfide bonding pattern was performed using the combination of *ab initio* and comparative modeling techniques with a disulfide bond predictor.

Initially, there were 66 possible combinations of disulfide bonds for snakins, since they have 12 cysteine residues involved in six disulfide bonds. Through QUARK modeling, four disulfide bonds were formed, reducing the possibilities of disulfide bond pairs to six combinations, since only two disulfide bonds were missing in the model. Therefore, a modified snakin-1 sequence was generated through the replacement of cysteine residues by serine residues. This replacement was done in order to avoid a wrong disulfide connectivity prediction in the DiANNA server. The final disulfide bonding pattern of snakin-1 was represented by Cys^I–Cys^{IX}, Cys^{II}–Cys^{VII}, Cys^{III}–Cys^{IV}, Cys^V–Cys^{XI}, Cys^{VI}–Cys^{XII} and Cys^{VIII}–Cys^X. This disulfide pattern could be extrapolated to other members of the snakin/GASA family through sequence alignment (Fig. 1A).

After the inclusion of remaining disulfide bonds through MODELLER, the best model showed a DOPE score of –5036.17432. The final model was obtained after energy minimization on SPDBV. The final model shows a minimum and a maximum 1D–3D average score of 0.3 and 0.55 in Verify 3D. In the Ramachandran plot, 72.2% of the residues are in favored regions; 14.8% are in additional allowed regions and 11.1% in generously allowed regions; and an overall G-factor of –0.23. The Z-score on ProSA was –5.85. The three-dimensional model was composed of two long α -helices composed of residues ²SSFCDCKKLRCSKA¹⁶ and ²⁰DRCLKYCGICCEE³² and one short 3_{10} -helix composed of ⁴³NKH⁴⁵, in addition, the structure was stabilized by six disulfide bonds (Fig. 1B). Furthermore, the same structural scaffold could be observed for other members from this family through secondary structure predictions algorithms (data not shown). The snakin-1 final model is available as supplementary file 1.

In the molecular dynamics simulation, a large displacement of two C-terminal segments, ³⁷PSGTYGNK⁴⁴ and ⁵⁰YRDKKNSKGKS⁶⁰, was observed. The root mean square deviation (RMSD) stabilization occurs after 30 ns of simulation with a variation of about 4.5 Å (Fig. 2A); removing the two C-terminal segments from RMSD calculation, a variation of about 3.5 Å was observed (Fig. 2A), reinforcing the supposition that the C-terminal segments are mainly responsible for the variation of 4.5 Å in the complete structure. The DSSP analyses indicated that the short 3_{10} -helix underwent a coil transition (Fig. 2B). However, the overall structure is maintained, since it is knotted by six disulfide bonds (Fig. 2C). In addition, the root mean square deviation by residue also indicated that the C-terminal segments were responsible for the structural modification (Fig. 2D). The TM-Score with a value of 0.5023 indicated that the initial and final structures share the same fold.

4. Discussion

The snakin/GASA family has enormous biotechnological potential since it plays a defensive role against several plant pathogens, especially against the pathogens that attack potato, one of the most important crops worldwide [22]. However, the omission of structural information about this family hinders a deeper understanding about the class.

The prediction of the disulfide bonding pattern of snakins is an enormous challenge, since there are 66 possible combinations of disulfide bonds. In this report, this puzzle was handled using the QUARK *ab initio* molecular modeling server, the DiANNA disulfide bond predictor and the MODELLER package to yield a three-dimensional model, joining the data generated by QUARK and DiANNA. In a previous work, QUARK and MODELLER were used together for predicting the structure of another plant AMP, Pg-AMP1, and also for its recombinant analog [32]. Here, once more, these two methods were used together. However, in this report MODELLER was used to include the remaining disulfide bonds, while for Pg-AMP1 and its recombinant analog, MODELLER was

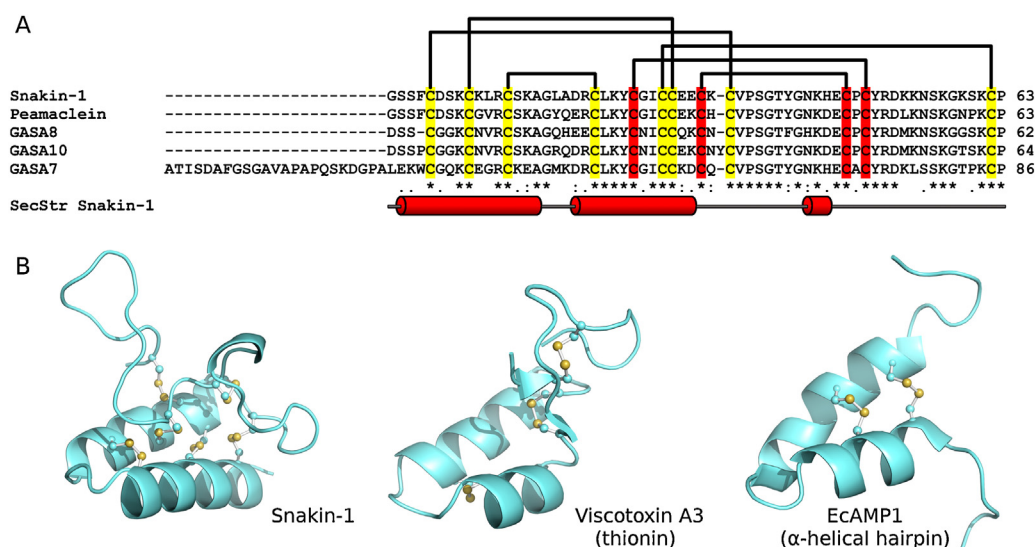


Fig. 1. The predicted structure and disulfide bonding pattern of snakin-1. (A) The multiple sequence alignment of snakin/GASA sequences: snakin-1 (UniProt: Q948Z4), peamaclein (UniProt: P86888), GASA8 (UniProt: O80641), GASA10 (UniProt: Q8LFM2) and GASA7 (UniProt: O82328); the disulfide bonds with cysteine residues highlighted in yellow were predicted by QUARK and those highlighted in red were predicted by DiANNA server; the lines connecting the cysteine residues represent the disulfide bonds; the α-helices are represented by the red cylinders. (B) Comparison between the final molecular model of snakin-1 (left) and the structures of viscotoxin A3 (thionin, middle) [10] and EcAMP1 (α-helical hairpin, right) [18]. The disulfide bonds are displayed as balls and sticks. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

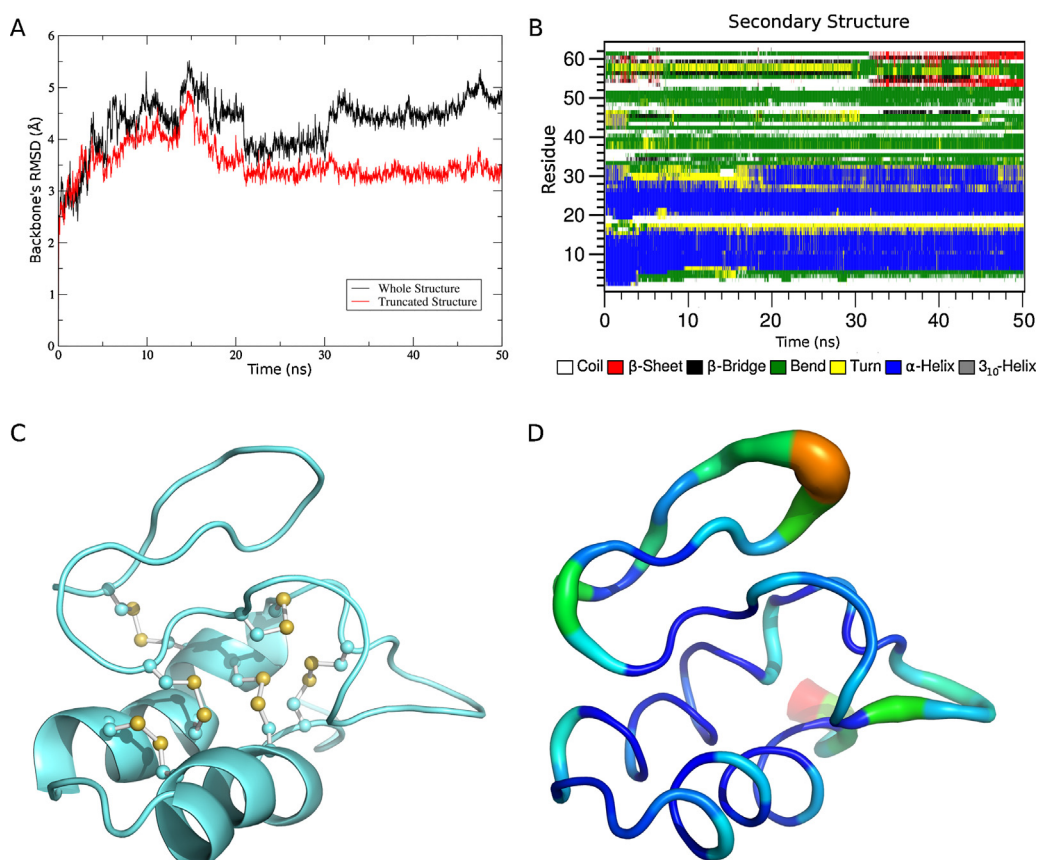


Fig. 2. Molecular dynamics simulation results. (A) The backbone's root mean square deviation throughout the 50 ns of simulation in relation to the minimized structure at 0 ns. (B) The DSSP analyses throughout the simulation: residues involved in coils are in white; β-sheets in red; β-bridge in black; bend in green; turn in yellow; α-helix in blue and 3₁₀-helix in gray. (C) The cartoon representation of the final three-dimensional snakin-1 structure after 50 ns of simulation. (D) The B-factor representation of root mean square fluctuation. The thicker the region, the more unstable the region was in the simulation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

used for refining loop conformations, generating several possible poses [32].

By using this method, a structure composed of one short 3_{10} -helix and two long α -helices, connected by loops, was generated. Among the plant AMPs, there are two classes with a structure composed of two long α -helices, the thionins [11,28] and the recently established α -helical hairpins [20,21] (Fig. 1B). Indeed, this degree of structural similarity with thionins reinforces the proposition of Silverstein et al. [31], who posited that some classes of plant cysteine-rich peptides could have a common ancestor, since they had observed internal duplications and cysteine rearrangements in diverse plant cysteine-rich sequences, including sequences for both GASA/GAST and thionin classes. Although the cysteine residues may be conserved in sequences, the disulfide bonds may not be structurally conserved. In this case, different disulfide bonding patterns could be observed, i.e. Cys^I-Cys^{IV}, Cys^{II}-Cys^V and Cys^{III}-Cys^{VI} (typical for cyclotides) or Cys^I-Cys^{VI}, Cys^{II}-Cys^V and Cys^{III}-Cys^{IV} (typical for thionins) [6,22].

Despite the structural similarity with thionins, the snakins' mechanism of action is still unclear. Thionins seem to be able to aggregate and induce leakage in negatively charged vesicles [5], while the snakins are also able to aggregate similar vesicles, but were unable to cause cytoplasmic leakage [5]. Similarly, the peptide EcAMP1, pertaining to the α -helical hairpins class, was unable to cause cell membrane disruption, but it has the ability to internalize into fungal cells [20]. The cell membrane was the only target tested so far, but there are a number of targets, such as cell wall, ribosomes, DNA or even a combination of these targets. In fact, more studies are needed to identify the mechanism of action of this AMP class.

5. Conclusion

This is the first report of the structural characterization of the peptide snakin-1, which belongs to the snakin/GASA family. Through the method applied here, combining *ab initio* and comparative modeling together with disulfide bond prediction, we hope that other peptides and proteins may be successfully modeled. The predicted snakin-1 structure presented here could be a step forward in the understanding of the missing biological information on snakins in plant biology. In addition, the predicted snakin-1 structure indicates that the snakin/GASA family could be closely related to the thionin family. In the near future, the method described here can be applied to predict the structure of other snakins and allow the engineering of more potent snakins for crop protection.

Acknowledgments

This work was supported by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico); CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior); FAPDF (Fundação de Amparo a Pesquisa do Distrito Federal) and UCB (Universidade Católica de Brasília).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.peptides.2013.03.014>.

References

- [1] Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl Acid Res* 1997;25(17):3389–402.
- [2] Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J. Interaction models for water in relation to protein hydration. In: Pullman B, editor. *Intermolecular force*. Dordrecht: Reidel; 1981. p. 331–42.
- [3] Berrocal-Lobo M, Segura A, Moreno M, López G, García-Olmedo F, Molina A. Snakin-2, an antimicrobial peptide from potato whose gene is locally induced by wounding and responds to pathogen infection. *Plant Physiol* 2002;128(3):951–61.
- [4] Broekaert WF, Mariën W, Terras FR, De Bolle MF, Proost P, Van Damme J, et al. Antimicrobial peptides from *Amaranthus caudatus* seeds with sequence homology to the cysteine/glycine-rich domain of chitin-binding proteins. *Biochemistry* 1992;31(17):4308–14.
- [5] Caaveiro JMM, Molina A, González-Manas JM, Rodríguez-Pelencuela P, García-Olmedo F, Goñi FM. Differential effects of five types of antipathogenic plant peptides on model membranes. *FEBS Lett* 1997;410:338–42.
- [6] Cândido ES, Porto WF, Amaro DS, Viana JC, Dias SC, Franco OL. Structural and functional insights into plant bactericidal peptides. In: Méndez-Vilas A, editor. *Science against microbial pathogens: communicating current research and technological advances*. Formatex; 2011. p. 951–60.
- [7] Chen KC, Lin CY, Chung MC, Kuan CC, Sung HY, Tsou SCS, et al. Cloning and characterization of a cDNA encoding an antimicrobial protein from mung bean seeds. *Bot Bull Acad Sin* 2002;43:251–9.
- [8] Darden T, York D, Pedersen L. Particle mesh ewald: an n long(n) method for ewald sums in large systems. *J Chem Phys* 1993;98:10089–92.
- [9] Eswar N, Marti-Renom MA, Webb B, Madhusudhan MS, Eramian D, Shen MY, et al. Comparative protein structure modeling with MODELLER. *Curr Protoc Bioinform* 2007;(Suppl. 15):5.6.1–30.
- [10] Ferrè F, Clote P, DiANNA 1.1: an extension of the DiANNA web server for ternary cysteine classification. *Nucl Acid Res* 2006;34:W182–5.
- [11] Fujimura M, Ideguchi M, Minami Y, Watanabe K, Tadera K. Purification, characterization, and sequencing of novel antimicrobial peptides, Tu-AMP 1 and Tu-AMP 2, from bulbs of tulip (*Tulipa gesneriana* L.). *Biosci Biotechnol Biochem* 2004;68(3):571–7.
- [12] Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 1997;18(15):2714–23.
- [13] Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS. A linear constant solver for molecular simulations. *J Comp Chem* 1997;18(12):1463–72.
- [14] Hess B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* 2008;4:435–47.
- [15] Laskowski RA, Macartur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 1993;26:283–91.
- [16] Lüthy R, Bowie JU, Eisenberg D. Assessment of protein models with three-dimensional profiles. *Nature* 1992;356(6364):83–5.
- [17] Mandal SM, Dey S, Mandal M, Sarkar S, Maria-Neto S, Franco OL. Identification and structural insights of three novel antimicrobial peptides isolated from green coconut water. *Peptides* 2009;30(4):633–7.
- [18] Mandal SM, Migliolo L, Das S, Mandal M, Franco OL, Hazra TK. Identification and characterization of a bactericidal and proapoptotic peptide from *Cycas revoluta* seeds with DNA binding properties. *J Cell Biochem* 2012;133:184–93.
- [19] Miyamoto S, Kollman PA. SETTLE. An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J Comp Chem* 1992;13(8):2134–44.
- [20] Nolde SB, Vassilevski AA, Rogozhin EA, Barinov NA, Balashova TA, Samsonova OV, et al. Disulfide-stabilized helical hairpin structure and activity of a novel antifungal peptide EcAMP1 from seeds of barnyard grass (*Echinochloa crus-galli*). *J Biol Chem* 2011;286(28):25145–53.
- [21] Oparin PB, Mineev KS, Dunaevsky YE, Arseniev AS, Belozersky MA, Grishin EV, et al. Buckwheat trypsin inhibitor with helical hairpin structure belongs to a new family of plant defence peptides. *Biochem J* 2012;446:69–77.
- [22] Padovan L, Scocchi M, Tossi A. Structural aspects of plant antimicrobial peptides. *Curr Protein Pept Sci* 2010;11(3):210–9.
- [23] Pelegrini PB, Murad AM, Silva LP, Dos Santos RC, Costa FT, Tagliari PD, et al. Identification of a novel storage glycine-rich peptide from guava (*Psidium guajava*) seeds with activity against Gram-negative bacteria. *Peptides* 2006;29(8):1271–9.
- [24] Pinto MF, Fensterseifer IC, Migliolo L, Sousa DA, de Capdville G, Arboleda-Valencia JW, et al. Identification and structural characterization of novel cyclotide with activity against an insect pest of sugar cane. *J Biol Chem* 2012;287:134–47.
- [25] Pinto MFS, Almeida RG, Porto WF, Fensterseifer ICM, Lima LA, Dias SC, et al. Cyclotides: from gene structure to promiscuous multifunctionality. *J Evidence-Based Comp Alt Med* 2012;17:40–53.
- [26] Porto WF, Pires AS, Franco OL. CS-AMPPred. An updated SVM model for antimicrobial activity prediction in cysteine-stabilized peptides. *PLoS ONE* 2012;7:e51444.
- [27] Porto WF, Souza VA, Nolasco DO, Franco OL. In silico identification of novel hevein-like peptide precursors. *Peptides* 2012;38:127–36.
- [28] Romagnoli S, Ugolini R, Fogolari F, Schaller G, Urech K, Giannattasio M, et al. NMR structural determination of viscotoxin A3 from *Viscum album* L. *Biochem J* 2000;350:569–77.
- [29] Segura A, Moreno M, Madueño F, Molina A, García-Olmedo F. Snakin-1, a peptide from potato that is active against plant pathogens. *Mol Plant Microbe Interact* 1999;12:16–23.

- [30] Silva ON, Porto WF, Migliolo L, Mandal SM, Gomes DG, Holanda HH, et al. Cn-AMP-1: a new promiscuous peptide with potential for microbial infections treatment. *Biopolymers* 2012;98(4):322–31.
- [31] Silverstein KAT, Moskal WA, Wu HC, Underwood BA, Graham MA, Town CD, et al. Small cysteine-rich peptides resembling antimicrobial peptides have been under-predicted in plants. *Plant J* 2007;51(2):262–80.
- [32] Tavares LS, Rettore JV, Freitas RM, Porto WF, Duque AP, Singulani de L, et al. Antimicrobial activity of recombinant Pg-AMP1, a glycine-rich peptide from guava seeds. *Peptides* 2012;37(2):294–300.
- [33] Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl Acid Res* 1994;22(22):4673–80.
- [34] Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucl Acids Res* 2007;35:W407–10.
- [35] Xu D, Zhang Y. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins* 2012;80:1715–35.
- [36] Yount NY, Yeaman MR. Multidimensional signatures in antimicrobial peptides. *Proc Natl Acad Sci U S A* 2004;101(19):7363–8.
- [37] Zhang Y, Skolnick J. Scoring function for automated assessment of protein structure template quality. *Proteins* 2004;57(4):702–10.