

## Encrypted antimicrobial and antitumoral peptides recovered from a protein-rich soybean (*Glycine max*) by-product

Cynthia Silva Freitas<sup>a</sup>, Mauricio Afonso Vericimo<sup>b</sup>, Manuela Leal da Silva<sup>c</sup>, Giovani Carlo Veríssimo da Costa<sup>d</sup>, Patricia Ribeiro Pereira<sup>a</sup>, Vania Margaret Flosi Paschoalin<sup>a,\*</sup>, Eduardo Mere Del Aguila<sup>a</sup>

<sup>a</sup> Instituto de Química, Universidade Federal do Rio de Janeiro, Avenida Athos da Silveira Ramos 149, Rio de Janeiro, RJ 21941-909, Brazil

<sup>b</sup> Instituto de Biologia, Universidade Federal Fluminense, Outeiro São João Batista Centro, Niterói, RJ 4020-141, Brazil

<sup>c</sup> Núcleo em Ecologia e Desenvolvimento Sócio-Ambiental de Macaé (NUPEM), Universidade Federal do Rio de Janeiro, Avenida São José do Barreto, 764, Macaé, RJ 27910-970, Brazil

<sup>d</sup> Instituto de Saude Nova Friburgo, Universidade Federal Fluminense, Nova Friburgo, Rio de Janeiro, RJ 21941-970, Brazil



### ARTICLE INFO

#### Keywords:

Soybean by-product  
Food preservation  
Functional additive  
Antimicrobial and antitumoral peptides  
Molecular modeling

### ABSTRACT

Twelve promising candidates for antimicrobial peptides encrypted in F7J075, Q948X9, Q94LX2, Q3V5S6, Q4LER6, AOAOROHYM3 and I1LMQ8 with molecular masses ranging from 718.42 to 4872.43 Da were identified in two fractions obtained from gel filtration chromatography of soybean meal aqueous extract, a defatted by-product generated by soybean oil refinery. Most peptides found in both fractions are encrypted in two regions of  $\beta$ -conglycinin alpha or alpha-prime subunits, the major protein in soybean. The pool of peptides from both fractions inhibited the growth of Gram-positive and Gram-negative foodborne pathogens and exhibited no toxicity to mouse bone marrow or fibroblast cells, but inhibited human glioblastoma proliferation. Three-dimensional structure of  $\beta$ -conglycinin domains containing the best AMP candidates, determined by molecular modeling, indicated an alpha-helix conformation. The production in large scale and use of multifunctional peptides encrypted in soybean meal proteins is an innovative application of the concept of circular bioeconomy.

### 1. Introduction

Soybean is increasingly becoming one of the most important alternative proteinaceous vegetal sources for human and animal consumption, due to its amino acid composition, low cost and high abundance, and is currently the second largest source of vegetable oil worldwide. After its refining process for oil extraction, the remaining bulk product, called soybean meal, still rich in protein and fiber, is mainly utilized for animal feeding, while a small portion is further processed into various types of soy protein-derivatives for human consumption. Although soybean meal is the most relevant by-product generated by soybean processing, with high protein content, it is not considered a high value-added food product (Day, 2013).

Protein-rich by-products generated by agro-industries have become an alternative source for obtaining encrypted bioactive peptides from protein hydrolyzation able to promote positive effects on human health (Chi, Wang, Wang, Zhang, & Deng, 2015; Lemes, Sala, Ores, Braga,

Egea, & Fernandes, 2016; Luna-Vital, Mojica, González de Mejía, Mendoza, & Loarca-Piña, 2015; Najafian & Babji, 2015; Zhu, Wang, & Guo, 2015). These peptides can display several bioactivities, able to interfere in cognitive and neurologic functions, hormonal, nutritional and metabolic activities and display pharmacological effects, such as antimicrobial or antitumoral properties (Sánchez & Vázquez, 2017). The obtainment of such peptides from by-product proteins cheapen their costs and allow for readily available health-promoting agents while adding value to these by-products (Chi et al., 2015; Luna-Vital et al., 2015; Najafian & Babji, 2015; Zhu et al., 2015). At the same time, the use of these by-products also aids in decreasing environmental impacts originating from food production (Laufenberg, Kunz, & Nystrom, 2003; Lemes et al., 2016).

Soybean meal is one of the most abundant protein-rich by-products and bioactive nitrogenous-compounds available each year (Rayaprolu, Hettiarachchy, Chen, Kannan, & Mauromostakos, 2013). Thus, the valorization of soybean meal through the extraction of two major storage

\* Corresponding author.

E-mail addresses: [cynthia.freitas@yahoo.com.br](mailto:cynthia.freitas@yahoo.com.br) (C.S. Freitas), [vericimo@vm.uff.br](mailto:vericimo@vm.uff.br) (M.A. Vericimo), [manuela@macae.ufrj.br](mailto:manuela@macae.ufrj.br) (M.L. da Silva), [giovani.verissimo@gmail.com](mailto:giovani.verissimo@gmail.com) (G.C.V. da Costa), [biopatbr@gmail.com](mailto:biopatbr@gmail.com), [patriciarp@iq.ufrj.br](mailto:patriciarp@iq.ufrj.br) (P.R. Pereira), [paschv@iq.ufrj.br](mailto:paschv@iq.ufrj.br) (V.M.F. Paschoalin), [emda@iq.ufrj.br](mailto:emda@iq.ufrj.br) (E.M. Del Aguila).

proteins, glycinin (11S) and  $\beta$ -conglycinin (7S), followed by controlled hydrolysis may generate functional ingredients with high added value, since the resulting hydrolysates frequently exhibit bioactive properties such as antioxidant, antitumoral and antimicrobial activities (Peña-Ramos & Xiong, 2002). 7S is one of the most abundant storage proteins in soybean, constituting 30–46% of the total water-extractable proteins, ranging from 180 to 210 kDa, with a compact globular structure at physiological pH, consisting of three subunits,  $\alpha$ ,  $\alpha'$  and  $\beta$  (Santiago et al., 2008). 11S is a 320 kDa-holoprotein with a quaternary structure consisting of five subunits, G1, G2, G3, G4 and G5. 11S derived-peptides present a highly hydrophobic character, which is an important feature for antimicrobial peptides (AMPs) (Kuipers, Alting, & Gruppen, 2007; Natarajan, Xu, Bae, Caperna, & Garrett, 2006; Singh, Meena, Kumar, Dubey, & Hassan, 2015).

In this context, the main purpose of this study was to recover, identify and characterize food grade water-soluble antimicrobial and antitumoral peptides obtained from soybean meal, prepared according to green chemistry methods and low-cost technologies, to be used as functional and/or preservative food additives in substitution of synthetic preservatives and antibiotics.

## 2. Material and methods

### 2.1. Organisms

Soybean meal (*Glycine max*) samples were donated by a soybean oil crushing Brazilian industry.

The toxicological tests based on cell viability assays used three cell lines, namely healthy mouse bone marrow and L929 fibroblast cell lines (ECACC Sigma-Aldrich Co, MO, USA) and human glioblastoma U-87 MG cell line (ECACC Sigma-Aldrich Co).

Antimicrobial analysis was performed by using the following microorganisms: *Acinetobacter* genospecies 3 ATCC17922 (isolated from a sludge refinery), coagulase-negative *Staphylococcus saprophyticus* KT955005 (isolated from soft cheese), *Staphylococcus aureus* strains ATCC14458 and ATCC13150, *Aeromonas hydrophila* strains FDA110-36 and ATCC7966, *Escherichia coli* strains ATCC43895, NCTC8959 and DH5alpha, *Salmonella enterica* subsp. *diarizonae* strains ATCC12325 and ATCC29934 and *Vibrio parahaemolyticus* ATCC17802 were kindly provided by the Oswaldo Cruz Institute - INCQS cell bank.

### 2.2. Preparation of the soybean meal aqueous extract

The aqueous extract was prepared according to Del Aguila, Gomes, Freitas, Pereira, and Paschoalin (2017), where 50 g of extruded soybean material was homogenized in 200 mL of distilled water followed by incubation at 50 °C, to activate endogenous proteolytic enzyme activity, for 24 h in constant agitation. The resulting suspension was centrifuged at 8000g for 10 min at room temperature, and the supernatant was then treated at 90 °C, to stop proteolysis, for 10 min under constant agitation followed by filtration through a 0.22  $\mu$ m pore membrane (Merck Millipore Co, Darmstadt, GER).

### 2.3. Peptide content determination

The peptide concentrations of the collected fractions were determined using a Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific, MA, USA), according to the manufacturer's instructions.

### 2.4. Peptide fractionation

A sample of soybean meal aqueous extract was ultra-filtered through an Amicon® Ultra-15 10 kDa Centrifugal Filter (Merck Millipore Co), according to the manufacturer's recommendations. The cutoff membrane size was chosen based on a previous study that indicated the predominance of protein-derived molecules with masses

$\leq 10$  kDa in the soybean meal aqueous extract (Del Aguila et al., 2017). Proteins and peptides  $> 10$  kDa accumulated over the membrane and those smaller than the cutoff value ( $< 10$  kDa) were collected in the filtered material. Three milliliters of the filtered suspension were fractionated on a gel filtration BIO-GEL P-30 (Bio-Rad, CA, USA) column (1.7 × 23 cm) previously equilibrated with 50 mM sodium phosphate buffer pH 7.0 at a constant flow rate of approximately 10 mL/h and at room temperature. Fractions were collected every 20 min and absorbances were monitored using a spectrophotometer (Beckman Coulter, CA, USA) at 280 and 215 nm.

### 2.5. Spectrophotometric assays

Total phenolic compounds (TPC) concentrations were determined using the Folin-Ciocalteu method (Singleton, Orthofer, & Lamuela-Raventós, 1999). Results were expressed as mg of gallic acid equivalents per g (mg GAE·g<sup>-1</sup>) of fresh weight (FW).

### 2.6. Evaluation of antimicrobial activity

Bacteria, obtained from an inoculum suspension at 10<sup>6</sup> cells/mL, were grown in the appropriate medium, following the recommended instructions provided by the guide manual from the cell bank. *A. genospecies* 3, *E. coli* strains DH5 alfa and ATCC43895 were grown in Luria-bertani (LB) media (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl) (BD™, Le Pont de Claix, FRA). *S. aureus* ATCC14458, coagulase-negative *S. saprophyticus* KT955005 and *A. hydrophila* ATCC7966 were grown in brain heart infusion (BHI) media (5 g/L beef heart infusion, calf 12.5 g/L brain infusion, 2.5 g/L disodium hydrogen phosphate, 2 g/L D (+)-glucose, 10 g/L peptone) (BD™). *V. parahaemolyticus* ATCC17802 and *Salmonella enterica* strains ATCC1225 and ATCC29934 were grown in nutrient broth (NB) media (1 g/L D (+)-glucose, 15 g/L peptone, 6 g/L sodium chloride, 3 g/L yeast extract) (Himedia, Mumbai, IND). Subsequently, cells were serially diluted (1/10) in saline solution (0.85% of NaCl, 0.2% Tween 80), plated on their respective solid media, LB, NB or BHI, incubated at 37 °C for 18 h and colony-forming units (CFU/mL) were determined.

Antimicrobial effects of samples from the crude and fractionated soybean meal aqueous extract, at concentrations ranging from 51 to 1050  $\mu$ g/mL, were tested against foodborne bacteria.

### 2.7. In vitro toxicological tests

The cytotoxicity of the soybean meal aqueous extract fractions obtained by gel filtration was assessed by *in vitro* assays against healthy murine and human tumoral cells. Bone marrow (BM) cells were collected from BALB/c mice aged 8–10 weeks (protocol approved by the Institutional Ethics Committee for Animal Research at Universidade Federal Fluminense under N° 821-16).

One hundred microliters containing healthy mouse BM cells (5.0 × 10<sup>5</sup> cells/mL), fibroblast L929 or human glioblastoma U-87 MG cells (1.5 × 10<sup>5</sup> cells/mL) were cultured in 96-well microplates for 24 h at 37 °C in a 5% CO<sub>2</sub> humidified atmosphere. One hundred microliters of the samples were added to a semi-confluent cell layer in the wells, at concentrations ranging from 222.8 to 1.74  $\mu$ g/mL. After 24 h, cell viability was evaluated by the addition of 0.05  $\mu$ g resazurin (Sigma-Aldrich Co), according to McMillian et al. (2002), with modifications, where cells were exposed to resazurin for additional 6 h following the detection of fluorescence intensity in a Victor™ X microplate reader (Perkin Elmer Inc, MA, USA) at excitation and emission wavelengths of 530 and 590 nm, respectively.

### 2.8. Peptide concentration and desalination using ZipTip® pipette tips

Pipette tips were activated by 100% acetonitrile and a 0.1% trifluoroacetic acid (TFA) aqueous solution. Peptides were attached to the

ZipTip®C-18 (Merck Millipore Co) washed by 0.1% aqueous TFA, followed by elution of the peptides adsorbed to the C-18 matrix using 4 µL of 0.1% TFA and 70% acetonitrile in water, carefully dispensed into a microfuge tube. Reagents were removed by evaporation using a Speed-Vac concentrator (Savant, MA, USA).

### 2.9. Nano-liquid chromatography, mass spectrometry and data analyses

One hundred micrograms of each sample were suspended in 400 µL of 5% acetonitrile and 0.1% formic acid pH 3.2 and then centrifuged at 12,000g for 10 min. The supernatant was recovered and a 4 µL aliquot was injected, in triplicate, into a nano-LC Ultimate 3000 (Thermo Fisher Scientific, IL, USA) coupled to a Quadrupole-Orbitrap Q-Exactive Plus mass spectrometer (Thermo Fisher Scientific, Bremen, GER). The peptides were loaded on a pre-column (2 cm length, 200 µm inner diameter), packed with ReproSil-Pur C18-AQ 5 µm resin and fractionated using a Picochip analytical column containing ReproSil-Pur C18 3 µm resin. Peptides were eluted using a gradient comprising 95% solvent A (95% H<sub>2</sub>O, 5% ACN, 0.1% formic acid) and 40% solvent B (95% ACN, 5% H<sub>2</sub>O, 0.1% formic acid) for 40 min; 40% to 85% solvent B for 10 min; and 95% solvent B for 15 min, at 300 nL/min.

The full scan and MS/MS acquisitions were obtained in positive mode applying a Data Dependent Acquisition (DD-MS<sup>2</sup>). The ion source and S-lens were optimized to spray voltage at 3.4 kV, zero flow of sheath and auxiliary gas at 250 °C and 80 S-Lens RF level. The MS full scan was acquired at a 70,000 resolution at *m/z* 400–2000 in the Orbitrap analyzer, 10<sup>6</sup> AGC, and 50 ms maximum ion injection time. The 10 most intense ions with 2–4 charges were selected for higher-energy collision dissociation fragmentation. Fragment acquisition was performed in the Orbitrap using 30 collision energies and dynamic exclusion enabled for 20 s. MS/MS scans were obtained at 17,500 resolution, 5 × 10<sup>4</sup> AGC, 50 ms maximum isolation time, microscans 1 s, range *m/z* 200–2000 and isolation window *m/z* 2.0.

The DDA raw data were processed and searched by the Proteome Discovery 2.1 software server search engine (Thermo Fisher Scientific, Bremen, GER) using the SEQUEST algorithm and Peaks 8.0 (Bioinformatics Solutions, ON, CAN) with a tolerance up to ± 0.1 Da for the precursor ions and 0.01 Da for fragment ions. Protein identification was performed by searching the mass spectrometric data against the UniProt - SwissProt protein database (available in April 2018) containing reversed sequences with a false discovery rate (FDR) < 1%.

### 2.10. Prediction of potential antimicrobial sequences

Peptide hits identified by nano-LC-MS/MS in both fractions - F1 and F2 - were screened from an initial list of 19,606 sequences by defining the selection of peptides based on two occurrences in the triplicate and establishing a cutoff value of –15 to 15 ppm.

The initial screening resulted in 308 peptide hits (Table S1), evaluated using four algorithms: support vector machines (SVM), random forests (RF) artificial neural network (ANN) and discriminant analysis (DA), to determine peptide probability to exhibit antimicrobial activity (Waghu, Barai, Gurung, & Idicula-Thomas, 2015). Peptide sequences were considered the best AMP candidates when a positive result was obtained by at least three algorithms. When the peptide sequence was classified as an AMP by two algorithms, they were considered good candidates. The four algorithms can be freely accessed at CAMP R3 website (<http://www.camp.bicnirrh.res.in/prediction.php>).

Alternatively, the β-conglycinin amino acid sequence, the main source of the bioactive peptides from fractions F1 and F2, was used to screen potential antimicrobial sequences using the AMPA algorithm available at <http://tcoffee.crg.cat/apps/ampa/do> (Torrent et al., 2011).

### 2.11. Sequence comparison of *Glycine max* proteins and peptides

A multiple sequence alignment (MSA) between β-conglycinin

subunits (P11827, P13916, Q4LER6, Q3V5S6 and F7J075) and peptides with high potential for AMP as determined by the algorithms (Section 2.10) was performed. The CLUSTAL Omega 1.2.4 (Sievers et al., 2011) parameter inputs were used, namely no dealing input sequence, 1 max guide tree iterations, no combined iterations, order aligned, MBED-like clustering iterations and MBED-like clustering guide-tree.

### 2.12. Structural modeling of β-conglycinin protein and location of AMP regions

Secondary structures (SS) were evaluated using MSA (Section 2.11) as input by Ali2D (Zimmermann et al., 2018), PSIPRED and QUICK2D (Buchan, Minneci, Nugent, Bryson, & Jones, 2013). The pGenTHREADER method (Lobley, Sadowski, & Jones, 2009) was used to predict a profile based on fold recognition and I-TASSER (Yang, Yan, Roy, Xu, Poisson, & Zhang, 2015) to generate a three-dimensional structure (3D) model for Q3V5S6 and F7J075. The Modeller software (Webb & Sali, 2016) was used to optimize the secondary structure predicted for peptides regions. The 3D analyses were performed using the PyMOL software (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

### 2.13. Statistical analyses

All experiments were performed in triplicate, including antimicrobial activity, peptide identification and cell toxicology tests. Multiple comparison analyses were performed by the ANOVA test followed by the Tukey post-test (Zar, 1984) and significance was considered at *p* < 0.05, as determined by the GraphPad Software v.7 (GraphPad Software Inc., CA, USA).

## 3. Results

### 3.1. Fractionation of bioactive peptides and evaluation of antimicrobial activity

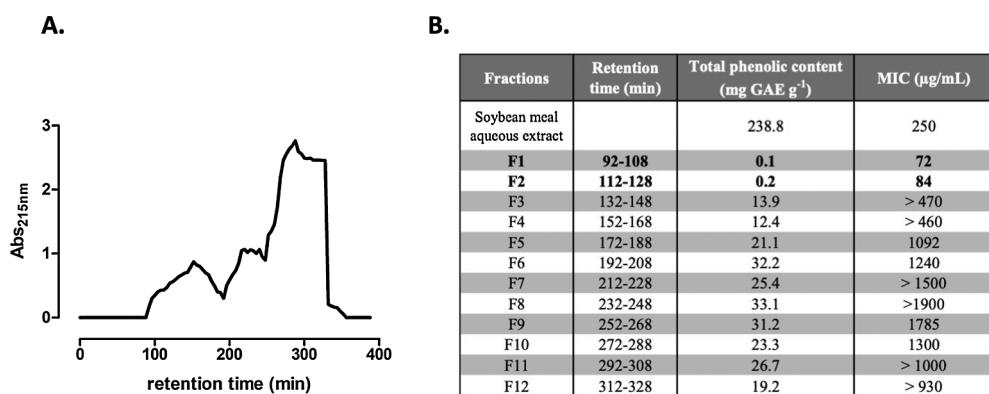
Twelve fractions containing peptides with molecular masses < 10 kDa were obtained, tested against the model microorganism *A. genomospecies* 3 and their phenolic compound content was quantified (Fig. 1A). Although fractions F1, F2, F5, F6, F9, F10 were able to completely inhibit the growth of *A. genomospecies* 3, F1 and F2, exhibited lower MIC values (72 µg/mL and 84 µg/mL, respectively) combined with the lowest phenolic compound contents (0.1 mg GAE·g<sup>-1</sup> and 0.2 mg GAE·g<sup>-1</sup>, respectively), minimizing the contribution of these compounds to the overall bioactivity of these fractions, but leaving antibacterial activity to the peptide pool (Fig. 1B).

Additionally, the antimicrobial activities of F1 and F2 fractions revealed that the peptide pools were able to inhibit the growth of other bacteria belonging to different species and genus, such as *Escherichia coli* DH5 alfa ATCC43895, *Staphylococcus aureus* ATCC14458, coagulase-negative *Staphylococcus saprophyticus* KT955005, *Aeromonas hydrophila* ATCC7966, *Vibrio parahaemolyticus* ATCC17802 and *Salmonella enterica* strains ATCC1225 and ATCC29934. MIC values were species-dependent and ranged from 750 to 1050 µg/mL and from 51 to 160 µg/mL for Gram-positive and Gram-negative bacteria, respectively (Table 1).

### 3.2. Antitumoral effect of phenolic-reduced fractions on human tumoral cells

Toxicological tests on fractions F1 and F2 were performed on healthy murine BM and L929 fibroblast cells.

Decreasing concentrations of F1, ranging from 220.2 to 1.72 µg/mL, showed no toxicity to healthy mouse BM and L929 cell lines, except for the highest concentration of 220.2 µg/mL, which caused a 70–60% decrease in cell viability. The F1 fraction had an IC<sub>50</sub> of 199.3 µg/mL



**Fig. 1.** Fractionation of the ultra-filtered soybean meal aqueous extract, anti-microbial activity and phenolic content of the resultant fractions. (A) Fractionation of the < 10 kDa-rich filtered suspension of soybean meal aqueous extract on a BioGel P-30 column equilibrated with 50 mM phosphate buffer pH 7.0. Peptide content of the collected fractions was monitored by absorbance at 215 nm. (B) Antimicrobial activity against the model microorganism *A. genomospecies* 3, cultivated in Luria-bertani media, and total phenolic concentration (TPC) of each fraction determined by the Folin-Ciocalteu method were evaluated.

and 134.4 µg/mL for BM and L929 cells, respectively (Fig. 2, top panel).

Despite its non-toxic effect on murine cells, F1 caused at least a 30% decrease in viability (Fig. 2, top panel) at 110.1–172 µg/mL and ended the viability of U-87 MG tumor cells when added at a concentration of 220.2 µg/mL ( $IC_{50} = 6.74 \mu\text{g/mL}$ ).

Similarly, F2 fraction did not affect healthy BM and L929 cells after a 24 h incubation, except at the highest concentration, of 222.8 µg/mL, that caused 78% and 55% increases, respectively, with  $IC_{50}$  of 158.6 µg/mL and 196.5 µg/mL (Fig. 2, bottom panel). On the other hand, the viability of U-87 MG tumor cell line was reduced in at least 50% at all tested concentrations, reaching 100% inhibition at 222.8 µg/mL ( $IC_{50} = 2.57 \mu\text{g/mL}$ ) (Fig. 2, bottom panel).

### 3.3. Identification of peptides in F1 and F2 by mass spectrometry

The nano-LC-MS/MS analysis of F1 and F2 fractions revealed a variety of peptides with molecular masses ranging from 718.42 to 4872.43 Da after the ultrafiltration process of the whole soybean meal aqueous extract through a 10 kDa molecular weight cutoff membrane (Table 2). A total of 308 non-identical peptide hits were identified in these fractions (Table S1). Some of these peptides showed truncated and/or overlapping sequences, revealed by comparison with data available at the UniProt protein database and Peaks 8.0 software analysis (Fig. S1). These peptides were encrypted in soybean proteins, mainly in β-conglycinin alpha and alpha-prime subunits, followed by glycinin, uncharacterized proteins, proteins involved in maturation process, oleosin, Kunitz trypsin inhibitor and seed biotinylated protein 68 kDa isoform (Table 2).

**Table 1**

Minimal inhibitory concentrations (MICs) estimated for F1 and F2 fractions from soybean meal aqueous extract against Gram-positive and Gram-negative bacteria.

Microorganisms tested		MIC – Fraction F1 (µg/mL)	MIC – Fraction F2 (µg/mL)
Gram positive	<i>Staphylococcus aureus</i> ATCC14458	1000	1050
	coagulase-negative <i>S. saprophyticus</i> KT955005 <sup>a</sup>	750	1050
	<i>S. aureus</i> ATCC13150	1010	900
Gram negative	<i>Acinetobacter</i> genomospecies 3 <sup>b</sup>	72	84
	<i>Aeromonas hydrophila</i> FDA110-36	90	100
	<i>A. hydrophila</i> ATCC7966	69	51
	<i>Escherichia coli</i> DH5alpha	105	98
	<i>E. coli</i> ATCC43895	105	98
	<i>E. coli</i> NCTC8959	105	98
	<i>Salmonella enterica</i> ATCC12325	140	150
	<i>S. enterica</i> ATCC29934	140	150
	<i>Vibrio parahaemolyticus</i> ATCC17802	160	110

F1 and F2 minimal inhibitory concentrations to stop the growth of  $10^6$  cells/mL at 37 °C for 18 h. Remaining bacterial cells were serially diluted (1/10), plated, incubated at 37 °C for 18 h and colony-forming units were counted. The growth inhibition test was performed in triplicate for each F1 and F2 concentration. *A. genomospecies* 3 and *E. coli* were grown in LB; *S. aureus*, coagulase-negative *S. saprophyticus* and *A. hydrophila* were grown in BHI; *V. parahaemolyticus* and *Salmonella enterica* were grown in NB media.

<sup>a</sup> Strain isolated from soft cheese (Nunes, de Souza, Pereira, Del Aguila, & Paschoalini, 2016).

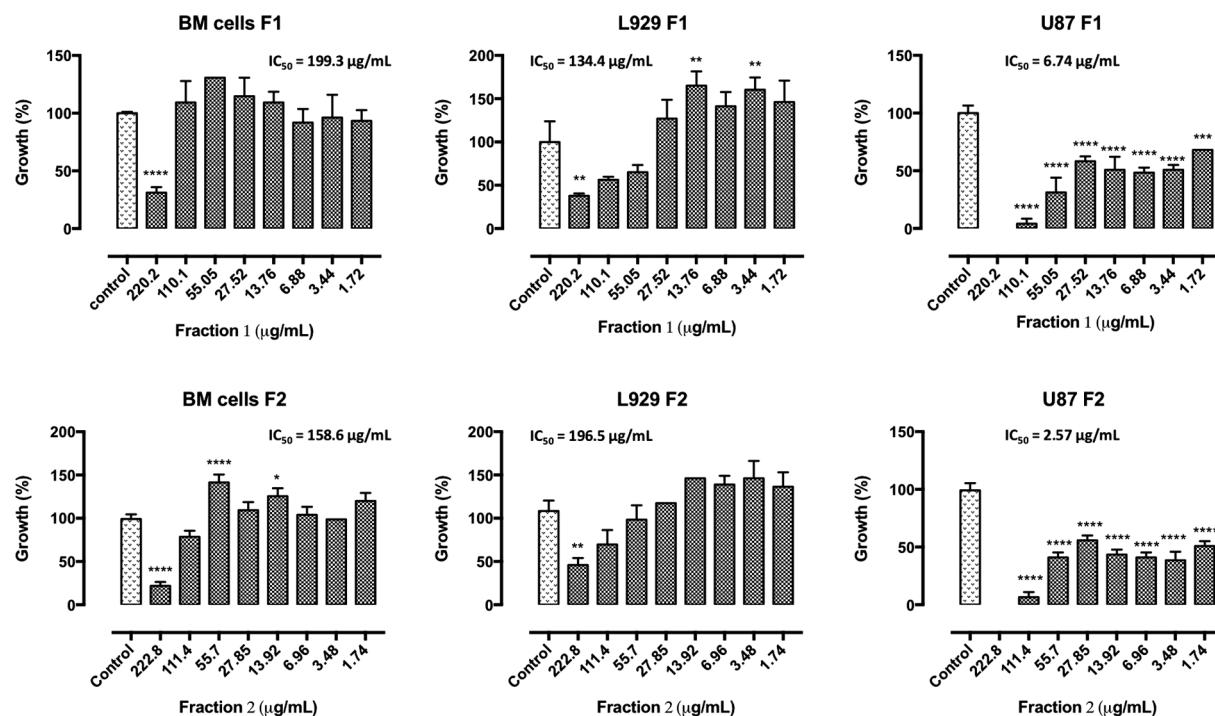
<sup>b</sup> Strain isolated from refinery sludge (Pinhati et al., 2014).

### 3.4. Antimicrobial peptide prediction evaluation

The potential to act as antimicrobial peptides (AMPs) were determined using four algorithms, support vector machines (SVM), random forests (RF) artificial neural network (ANN) and discriminant analysis (DA) available at the CAMP<sub>R3</sub> (Collection anti-microbial peptide) website (Waghlu et al., 2015). A total of 83 peptide sequences were classified as AMP candidates (Table 3). Peptides were considered the best AMP candidates when matched at least 3 algorithms. Ten peptides were selected, and 70% originated from the β-conglycinin alpha and alpha-prime subunits, while the remaining 30% originated from uncharacterized proteins. Alternatively, the AMPA algorithm was used to discover potential AMPs within the primary sequence of the anti-microbial protein β-conglycinin. The analysis revealed two potential antimicrobial peptides, found in both fractions, F1 and F2 (Table 3). A mass spectrometry analysis generated reliable and highly precise results represented by the spectrum of an AMP candidate (PRPIPFPQP), displayed in Fig. 3A.

### 3.5. Comparative analysis of *Glycine max* sequence proteins

A preliminary multiple sequence alignment (CLUSTAL Omega1.2.4) was performed between β-conglycinin subunits Q3V5S6, F7J075, Q4LER6, P11827 and P13916, followed by the location of AMP candidates within the protein sequences. AMP sequences showed high identity with the protein sequences found in the database as shown in Table S2. The location of the 10 best AMPs within β-conglycinin subunits revealed two distinct short domains (region 1 and region 2)



**Fig. 2.** Toxicological screening of F1 and F2 against healthy and tumoral cells. F1 (top panel) and F2 (bottom panel) were tested against mouse healthy bone marrow (BM) and fibroblast L929 cells and human glioblastoma U-87 MG cell line survival after 24 h exposure. Results were expressed as means  $\pm$  standard deviation ( $n = 3$ ) of bacteria growth versus fraction concentration plot. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$  and \*\*\*\* $p < 0.0001$  indicates difference between the untreated cells (control) and those incubated with F1 or F2 concentrations from 222.8  $\mu$ g/mL to 1.72  $\mu$ g/mL evaluated by the One-way ANOVA test with Tukey post-test.

(Fig. 3B). However, the location of the best AMP candidates within the tertiary structure of  $\beta$ -conglycinin subunits was not directly obtained, since the 3D structure of this specific region is not available in public databases.

### 3.6. Structural modeling of $\beta$ -conglycinin protein and location of AMP regions

To determine the three-dimensional structure of the two regions containing the 10 best AMP candidates within  $\beta$ -conglycinin, a consensus was constructed using the SS results from I-TASSER, Ali2D, PSIPRED and PSSPRED. Fig. 4A shows part of MSA for the target

regions in P11827, P13916, Q4LER6, Q3V5S6 and F7J075. I-TASSER simulations generated a large ensemble of structural conformations for the Q3V5S6 and F7J075 proteins. The top 5 3D models were predicted and ranked and the final model analyzes the confidence of each model and the peptide regions compared with other secondary predictions. The most relevant 3D model for Q3V5S6 was obtained with a C-score value of  $-2.86$  and estimated TM-score of  $0.39 \pm 0.13$ , and the most relevant 3D model for F7J075 was obtained with a C-score value of  $-2.48$  and estimated TM-score of  $0.43 \pm 0.14$ . I-TASSER uses the TM-align structural alignment program to match the first I-TASSER model to all structures in the PDB library. Proteins displaying the closest structural similarities with the Q3V5S6 and F7J075 3D models from the

**Table 2**

Quantitative analysis of peptide hits from F1 and F2 fractions evaluated by nano-LC-MS/MS.

Protein source	Access No.	# peptides		% #		Theo. MH + [Da]	
		F1	F2	F1	F2		
Beta-conglycinin	<b>F7J075</b> ; Q948X9; Q94LX2; Q4LER6	Q3V5S6; F7J075; Q4LER6	150	118	63.03	67.82	857.47–4872.43
Uncharacterized	K7K5Q8; I1JKB3; K7LDT9; AOA0R0LAD9; I1MUI0; K7N2A2; I1KID4; I1M222; <b>AOA0ROHYM3</b> ; AOA0R0JWF4; I1M0W1; I1LHP6; I1K7E6; AOA0R0FBK4; C6T1V2; C6SWV3; K7LEQ5; I1L957; I1K1B0; I1JY66; I1KV72; AOA0R0KKD6; K7MID0.	AOA0R0HYM3; I1M222; I1M0W1; AOA0R0JWF4; I1K7E6; K7LEQ5; I1LHP6; <b>I1LMQ8</b> ; I1L957; I1JMQ3; I1M5Y9	55	22	23.11	12.64	718.42–3808.66
Glycinin	P93707; Q9SB11	P93707; Q9SB11	22	26	9.24	14.94	1411.72–3560.53
Maturation protein	Q42447	Q42447	3	1	1.26	0.57	1436.82–1669.74
Seed maturation proteins	Q541U1; Q9S7N8;	Q9ZTY1; Q541U1; Q9XES8	3	4	1.26	2.30	1128.56–2173.97
Oleosin	I1N747	ND	3	ND	1.26	0.00	1308.68–1645.84
Kunitz trypsin inhibitor	Q39898	Q39898	1	3	0.42	1.72	1211.61
Seed biotinylated protein	C6K8D1	ND	1	ND	0.42	0.00	1090.62
68 kDa isoform							

Peptide hits from fractions F1 and F2 identified by nano-LC-MS/MS were screened by ppm cutoff value of 15 to  $-15$  and occurrence in at least two of three replicates. The quantity of peptides generated by each protein source was expressed in percentage (%). Access No. refers to the protein source code deposited at the protein database UniProt available at <https://www.uniprot.org> (accessed on April 2018). Accesses numbers in bold correspond to the protein source that originated the peptides classified as the best AMP candidates by the CAMPR3 analysis. ND – not detected.

**Table 3**  
Candidate AMP sequences assorted after specific algorithm evaluation.

Peptides	CAMPR3* analysis	Physicochemical characteristics			Occurrence	
		Charge	Hydrophobic ratio	Proline ratio	F1	F2
GEQQHEEEEREREHPPQPHPPHERG	+++ -	-6	0	16%	x	
GEQQHEEEEREREHPPQPHPPHE	+++ -	-6	0	18%	x	
ERQQHGEKEEDEGEQPRPFPFPRPRQPHQEE	+++ -	-4	6%	19%	x	
EQEQRPRPFPFPRP	+++ -	0	15%	38%	x	x
GEOQRPRFPFPFRP	+++ -	1	16%	41%	x	x
PRPIPFPFRPQP	+++ -	2	18%	54%	x	x
IPRPRPRPQHPEREPQ	+++ -	2	6%	37%		x
EKSKRILRGLKTLFFLITMVISLLL	+++ -	4	56%	0		x
EQDEREHRPHQPHQKEEKKH	+++ -	-3	0%	14%	x	
FPFPRPPHQK	+++ -	2	20%	40%	x	
DEREHPRPHQPHQKEEKKH	++ - -	-2	0%	15%	x	
EEQDEREHRPHQPHQKEEKKH	++ - -	-4	0%	13%	x	
GEQPRPFPFP	++ - -	0	20%	40%	x	
EQEQRPRPFPFPFP	++ - -	-1	18%	36%	x	x
EQEQRPRPFPFPFRPQP	++ - -	1	12%	37%	x	
DEGEQPRPFPFPFRPQP	++ - -	0	11%	35%	x	x
GEKEEDEDEQPRPIPFPFRPQPQRQE	++ - -	-4	8%	25%	x	
EKEEDEDEQPRPIPFPFRPQP	++ - -	-4	10%	30%	x	x
EEKRGEKGSEEEDEDEEEQDERQFPFPRPHQKE	++ - -	-10	5%	11%	x	
EKRGEKGSEEEDEDEEEQDERQFPFPRPHQKE	++ - -	-9	5%	11%	x	
RPRPQHPEREPQQPGEKEE	++ - -	-1	0%	26%	x	x
ERKQEEDEDEEQQRESEESED	++ - -	-10	0%	0%	x	
QPGEKEEDEDEQPRPIPFPFRPQPQRQ	++ - -	-3	8%	28%	x	
GEIPRPRPRPQHPEREPQ	++ - -	1	5%	33%	x	
GEIPRPRPRPQHPEREPQQPG	++ - -	1	4%	33%	x	
EDEQPRPIPFPFRPQP	++ - -	-1	13%	40%	x	x
DEEEDQPRPDHPPQRPS	++ - -	-4	0%	29%	x	x
EEEDQPRPDHPPQRPS	++ - -	-3	0%	33%	x	
EDEEEDQPRPDHPPQRPS	++ - -	-5	0%	27%	x	x
EEEDQPRPDHPPQRPSRPE	++ - -	-3	0%	31%	x	
EEEDQPRPDHPPQRPS	++ - -	-3	0%	31%	x	x
EEEEREREREHPPQPHPPHERG	++ - -	-5	0%	20%	x	x
EEEREREREHPPQPHPPHERG	++ - -	-4	0%	21%	x	x
EEREREHPPQPHPPHERG	++ - -	-3	0%	22%	x	
DEDEDEDKPRPSRPSQGK	++ - -	-3	0%	16%	x	x
GRGHERRKEEEETARR	++ - -	1	6%	0%	x	x
FPFPRPP	++ - -	1	28%	57%	x	
DEEQDERQFPFPRPP	++ - -	-3	13%	26%	x	x

(continued on next page)

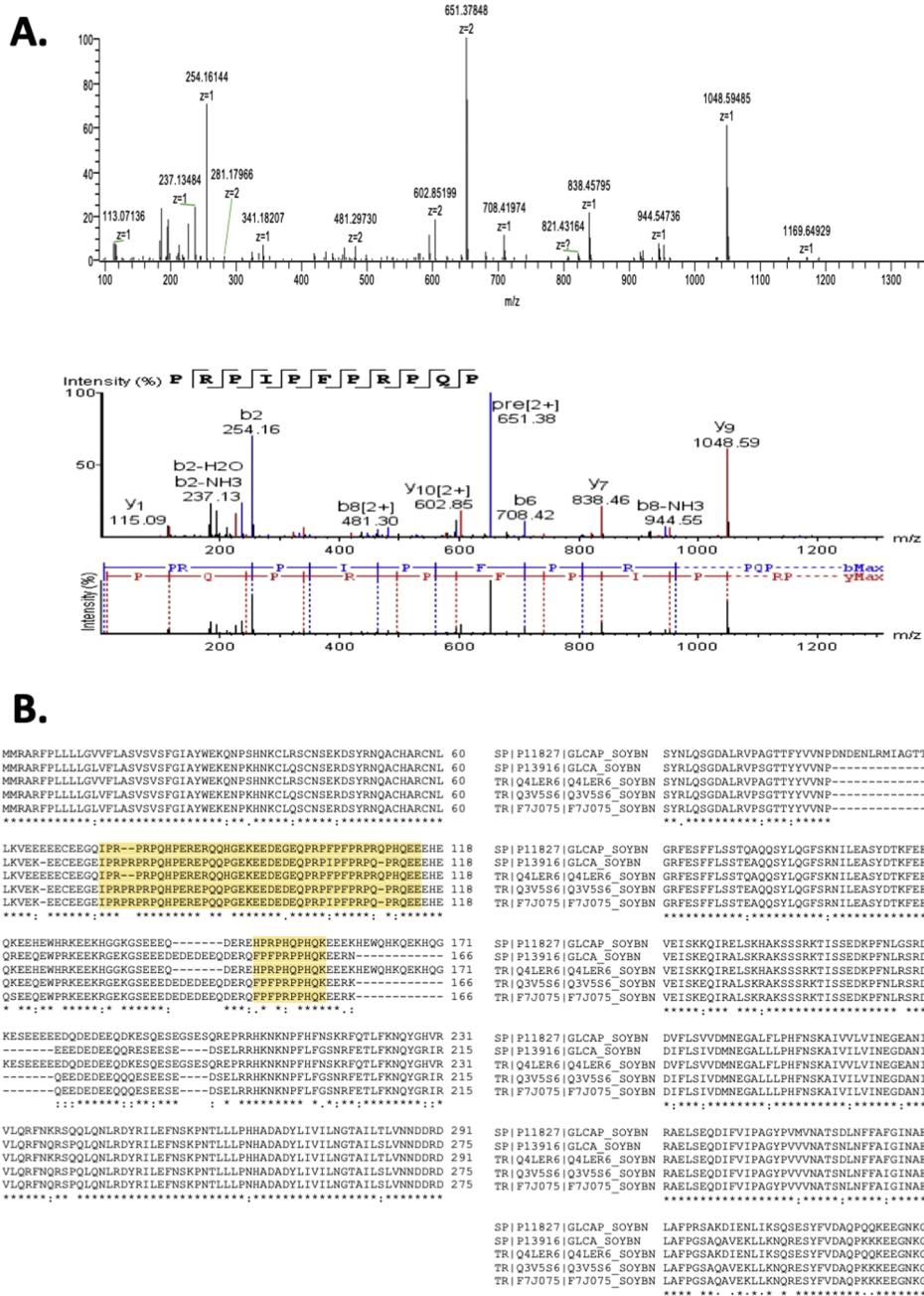
**Table 3 (continued)**

EDEQPRPIPFP	++-	-2	18%	36%	x	x
EDEDEEQDERQFPFPFRPP	++-	-6	11%	22%	x	x
EEQDERQFPFPFRPP	++-	-2	14%	28%	x	x
DEEDEEQDERQFPFPFRPP	++-	-5	11%	23%	x	x
EDEDEEQPRPIPFP	++-	-4	15%	30%	x	x
EDEDEEQPRPIPFP	++-	-5	14%	28%	x	x
DEQPRPIPFP	++-	-1	20%	40%	x	x
EDEEQDERQFPFPFRPP	++-	-4	12%	25%	x	x
EDEDEEQDERQFPFPFRPP	++-	-8	10%	20%	x	
DEDEDEEQDERQFPFPFRPP	++-	-7	10%	21%	x	x
DEEQDERQFPFP	++-	-4	16%	16%	x	x
EDEDEEEQDERQFPFP	++-	-9	11%	11%	x	
DEDEDEEQDERQFPFP	++-	-8	12%	12%	x	x
DEEDEEQDERQFPFP	++-	-6	14%	14%	x	x
QPGEKEEDEDEQPRPIPFPFRPQP	++-	-4	8%	30%	x	
GEKEEDEDEQPRPIPFPFRPQP	++-	-4	9%	28%	x	x
DEQPRPIPFPFRPQP	++-	0	14%	42%	x	x
EEDEDEEDEEQDERQFPFPFRPHQ	++-	-9	8%	17%	x	
DEDEQPRPIPFPFRPQP	++-	-2	12%	37%	x	x
DEEQDERQFPFPFRPHQ	++-	-3	11%	23%	x	
EDEDEEEQDERQFPFPFRPHQ	++-	-8	9%	18%	x	x
EDEDEQPRPIPFPFRPQP	++-	-3	11%	35%	x	x
EEDEDEEDEEQDERQFPFPFRPHQ	++-	-10	8%	16%	x	
DEDEDEEQDERQFPFPFRPHQ	++-	-7	9%	19%	x	x
DEDEDEEQDERQFPFPFRPHI	++-	-7	10%	20%	x	
DEDEQPRPIPFPFRP	++-	-2	14%	35%	x	x
EHPRPHQPHDEDEEQDERQFPFPFRPHQ	++-	-5	7%	25%	x	
FSRAFPFPPR	++-	2	40%	30%	x	
DEDEQPRPIPFP	++-	-3	16%	33%		x
DEQPRPIPFPFRP	++-	0	16%	41%		x
EDEDEEQDERQFPFP	++-	-7	13%	13%		x
EDEDEEQDERQFPFPFRPHQ	++-	-6	10%	20%		x
EEDEDEQPRPIPFPFRP	++-	-4	12%	31%		x
EEDEDEQPRPIPFPFRPQP	++-	-4	11%	33%		x
KEEDEDEQPRPIPFPFRPQP	++-	-3	10%	31%		x
DEDEQPRPFPRPQP	++-	-2	12%	37%		x
EEDEDEQPRPFPRPQP	++-	-4	11%	33%		x
GSEEEQDEREHPRPHQPH	++-	-4	0%	16%		x
REHPRPHQPHQKEEKH	++-	0	0%	17%	x	x
DEGEQPRPFPRPQRPH	++-	0	11%	33%		x
DEEEDQPRPDHPPQRP	++-	-4	0%	31%		x
RPEQQEPRG	++-	0	0%	22%		x
GEQQHEEEEREREHPQPHPPHEE	++-	-7	0%	17%		x
HPEREPQQPG**	NA	-2	0	27%	x	x
EQDERQFPFP **	NA	-2	20%	20%	x	x

Peptide hits were evaluated by the \*CAMP R3 method using 4 algorithms, namely support vector machines (SVM), random forests (RF) artificial neural network (ANN) and discriminant analysis (DA) (Waghu et al., 2015). NA – Not applied.

The first ten peptide sequences were considered the **best candidates**, since they were positively classified as AMP by at least 3 algorithms (+ + + -). Positive sequences in only two of four algorithms (+ + - -) were considered **good candidates** and are shaded in gray.

\*\*Sequences within  $\beta$ -conglycinin identified by the AMPA algorithm analysis as potential active regions (Torrent et al., 2011). The two peptides were included in the best candidates group.



**Fig. 3.** Spectra and localization of the best AMPs candidates within  $\beta$ -conglycinin primary structure. (A) Representative spectra of the original and processed peptide (PRPIPFPRPQP) identified by nano-LC-MS/MS proposed as one of the best candidates for AMPs in Table 3 (ppm < 6.7 and 651.38 m/z represent the precursor ion). (B) The predictive functional sequences of five  $\beta$ -conglycinin subunits Q3V5S6, F7J075, Q4LER, P11827 and P13916, deposited at UniProt database, were aligned by Clustal Omega and the 10 AMP best candidates were localized within the sequences. Regions shaded in orange correspond to antimicrobial domains containing the complete or partial amino acid sequences of the best AMP candidates exhibited in Table 3.

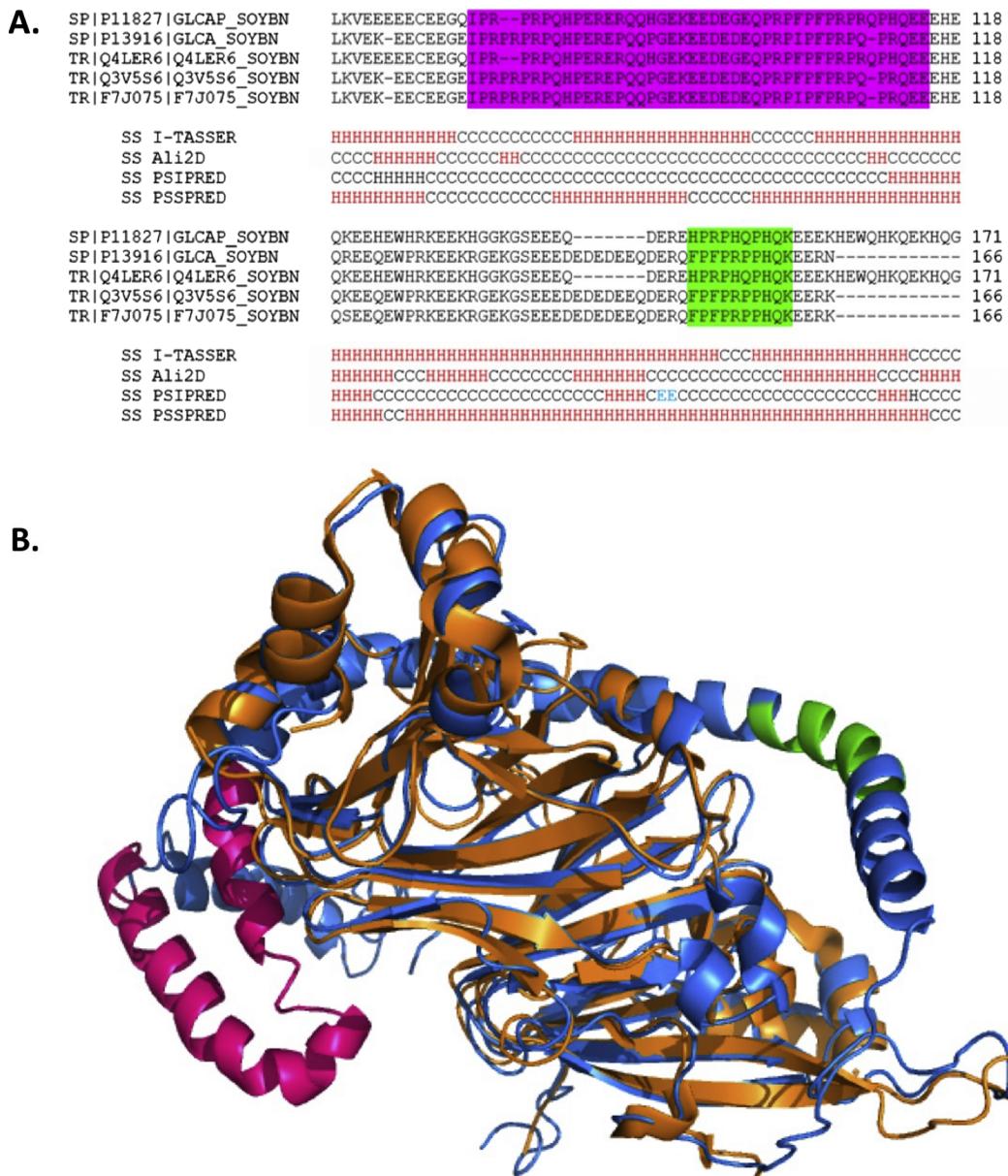
PDB were 7S globulin-1 from *Phaseolus angularis* (PDBid 2EA7\_A) and  $\beta$ -conglycinin from *Glycine max* (PDBid 1IPK\_C), respectively. In Fig. 4B, the 3D model built for F7J075 protein (in blue) is presented after applying secondary structure restraints. SS prediction was used to guide this module using the Modeller package.

#### 4. Discussion

The emergence of multi-resistant bacteria as a result of the excessive use of antibiotics has become a problem not only for the public health system but also for the food industries. Foodborne pathogen resistance has increased in the last decades, disseminating through products of

animal origin due to the indiscriminate use of antibiotics in the veterinary clinical practice. *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Salmonella* spp., coagulase-negative *Staphylococcus* spp., *Shigella* spp., *Enterococcus* spp. and *Escherichia coli* are among the main bacteria presenting multidrug resistance, all included in the category of community and hospital-acquired pathogens, in response to the strong demand by consumers for novel and broad-spectrum antibiotics against several pathogens (Fisher & Phillips, 2008).

The search for alternative agents like AMPs, which exhibit broad antimicrobial spectrum but low ability to trigger bacterial resistance mechanisms, has boosted the development of these novel antimicrobial compounds (Shagaghi, Palombo, Clayton, & Bhave, 2018).



**Fig. 4.** Three-dimensional structure of  $\beta$ -conglycinin regions containing AMPs candidate sequences. (A) Partial alignment between P11827, P13916, Q4LER6, Q3V5S6 and F7J075 with the region 1 sequence shaded in magenta and the region 2 in green. The prediction of the secondary structure is represented by red H for alpha-helix and black C for beta-sheet. (B) Structural analysis by overlap of the F7J075 3D model in blue and 7S globulin-1 from *Phaseolus angularis* in orange (PDB ID 2EA7\_A). Region1 is in magenta and region 2, in green, containing the 10 best AMPs.

Antimicrobial peptides can be naturally present or encrypted into a larger protein sequence where, in its encrypted form, they do not display any functionality. To exert their biological effects, encrypted peptides must be released by different ways, including enzymatic gastrointestinal digestion, microbial proteolytic enzymes, fermentation processes or the action of endogenous proteolytic enzymes on the matrix where peptides are encrypted in (Agyei, 2015; Guaadaoui, Benicha, Elmajdoub, Bellaoui, & Hamal, 2014).

Soybean proteins and/or peptides are generally extracted by water or an aqueous buffer containing phosphate salts (Amnuaycheewa & de Mejia, 2010). Herein, the peptide-rich soybean meal extract was prepared by aqueous extraction and the hydrolysis was carried out not by the addition of exogenous enzymes, but by thermal activation of those naturally present in the soybean meal, under mild temperature condition ( $50^{\circ}\text{C}$ ). The proteolysis step was ended by a severe heat shock at  $90^{\circ}\text{C}$ . A previous study by our research group has demonstrated that this soybean meal aqueous extract is rich in protein-derived molecules

with masses predominantly  $\leq 10\text{ kDa}$ , able to inhibit the growth of foodborne bacteria (Del Aguila et al., 2017). The antimicrobial effect of the soybean meal aqueous extract can be attributed not only to the antimicrobial peptide pool with molecular masses lower than  $10\text{ kDa}$  but also to phenolic compounds, such as isoflavones (Freitas et al., 2019). Previous reports have shown that soybean isoflavones can display antibacterial activity per se and the ability to act as potentiators for other antimicrobial agents, such as  $\alpha$ -linoleic acid, and synthetic and natural antibiotics, as norfloxacin and berberine (Hong, Landauer, Foriska, & Ledney, 2006; Laodheerasiri & Pathirage, 2017). According to Morel, Stermitz, Tegos, and Lewis (2003), the use of isoflavones from *L. angustus* in combination with the antibiotic berberine enhanced the uptake of berberine by naturally resistant *Staphylococcus aureus* and improved antimicrobial effectiveness.

Considering that the soybean meal aqueous extract is rich in peptides and isoflavones, separation of the antibacterial compounds was performed by ultrafiltration through a  $10\text{ kDa}$  cutoff membrane

followed by gel filtration chromatography fractionation, to eliminate or minimize isoflavone interference and facilitate the identification of antimicrobial peptides. Additionally, vegetable extracts contain several interfering substances in their glycoside forms, including flavonoids, where the sugar content decreases the effectiveness of antibacterial activity, resulting in poor reproducibility, one of the main obstacles in this context. In addition, qualitative and quantitative variations in bioactive phytochemical content in plant extracts result in inconstant effectiveness (Negi, 2012).

The gel permeation process of the soybean meal aqueous extract resulted in 12 fractions with distinct antimicrobial potential and peptide and phenolic compound content. To guarantee that phenolic compounds, especially isoflavones, would not interfere in the antimicrobial assays, the fractions with only traces of phenolics, F1 and F2 ( $\leq 0.2 \text{ mgGAEg}^{-1}$ ), were selected for further analysis. Both fractions were able to stop foodborne bacteria growth, with MIC (minimal inhibitory concentrations) ranging from 51 to 1050  $\mu\text{g/mL}$ , depending on the bacteria species. It is interesting to note that Gram-negative bacteria exhibited higher sensitivity to the antimicrobial peptides pool from soybean meal aqueous extract, requiring concentrations around 10-fold lower than those required to abolish the growth of Gram-positive bacteria. The physicochemical characteristics of these peptides are essential to their antimicrobial action, where the selectivity concerning Gram-negative bacteria could be the result of positively charged peptides interaction with negatively charged membrane constituents, such as phospholipids and lipopolysaccharides, that promote membrane rupture or inhibit internal cell constituents. Hydrophobicity also plays an essential role in peptide ability to interact with bacterial membranes (Del Aguila et al., 2017). No difference in F1 and F2 antimicrobial efficiencies were observed, since they stopped bacterial growth at similar concentrations, indicating that they probably share many antimicrobial peptides, evidenced by the absence of sharp peaks released by Biogel P-30 permeation, resulting in poor peptide separation in F1 and F2 (Fig. 1A).

Neither fraction exhibited toxic effects to mouse healthy bone marrow and fibroblast L929 cells at peptide concentrations ranging from 111.4 to 1.74  $\mu\text{g/mL}$  (Fig. 2). These results suggest that AMPs in F1 and F2 may be promising antimicrobial agents to be safely applied to food products without causing toxicity to consumers. Further studies are still necessary in order to evaluate the effects of such AMPs on healthy human cells.

F1 and F2 also exhibited similar antitumoral effects against the glioblastoma U-87 MG cell line, reinforcing the possibility that they probably share many peptides (Fig. 2). The cells had their viability reduced by at least 90% at 111.4  $\mu\text{g/mL}$ , maintaining reduced viability up to 50% when exposed to the lowest concentration of 1.74  $\mu\text{g/mL}$ . These results indicate a higher specificity of F1 and F2 peptide pools for the glioblastoma U-87 MG cell line, a malignant brain tumor of invasive nature and rapid progression, resulting in low survival rates after surgical resection and specific treatment (Stupp et al., 2005). Based on this, the peptides recovered from the soybean meal aqueous extracts should have their individual effectiveness investigated concerning such a serious tumor cell lineage, aiming for their use as functional additives in food products to aid in the fight against glioblastoma and possibly contribute to improve the quality of life or extend patient survival.

It has been reported that many antimicrobial peptides, including anionic or cationic and cell-penetrating peptides, can also act as anticancer peptides (ACPs), sharing similar mechanisms of action and displaying antimicrobial functionality. The use of AMPs as ACPs presents the same advantages of AMPs, namely low propensity to resistance and low toxicity to healthy host cell lines. The peptides described herein could also be considered valuable anticancer candidates, due to their toxicity to U-87 MG tumor cells (Deslouches & Di, 2017; Felício, Silva, Gonçalves, Santos, & Franco, 2017; Gaspar, Veiga, & Castanho, 2013; Prabhu et al., 2013; Prabhu, Dennison, Mura, Lea, Snape, & Harris, 2014).

The identification of peptides in fractions F1 and F2 by mass spectrometry confirmed the presence of 308 non-identical peptide hits and, as suspected, many are shared by both fractions, supporting the similar effects observed in the antimicrobial, antitumoral and toxicological tests (Table S1).

Not surprisingly, the majority of these peptides were encrypted in  $\beta$ -conglycinin, followed by glycycinin, which, together, account for  $> 80\%$  of total soybean protein content (Wang, Qin, Sun, & Zhao, 2014). These major storage globulins are known to exhibit antimicrobial properties (Sithohy, Mahgoub, & Osman, 2012; Vasconcellos, Woiciechowski, Soccol, Mantovani, & Soccol, 2014) and, therefore, are potential sources of antimicrobial and/or anticancer peptides. In fact, after the analysis of each identified peptide, using five algorithms to predict their potential to exert antimicrobial activities, 12 peptide sequences were selected as the best AMP candidates, 9 from  $\beta$ -conglycinin and 3 from uncharacterized proteins (Table 3). AMP prediction tools are based on typical antimicrobial characteristics, including composition, physicochemical characteristics and structural amino acids features (Liu et al., 2017; Thomas, Karnik, Barai, Jayaraman, & Idicula-Thomas, 2009). These peptides shared certain characteristics, such as a high proline (P) frequency, in most of the sequences, positive or negative net charge and, in some cases, high ratio of hydrophobic residues, all essential features required by AMPs to efficiently exert their function (Li et al., 2014; Tam, Wang, Wong, & Tan, 2015). Although the candidate peptides identified herein exhibit the required antimicrobial properties, experimental validation is an essential step to verify each particular effectiveness. Since antimicrobial peptides can act in synergy, potentiating antimicrobial activity, further studies should include the synthesis of AMP candidates followed by antimicrobial tests, to evaluate their individual and/or combined contribution to inhibit bacteria growth (Yu, Baeder, Rego, & Rolff, 2016).

Alignment analysis between  $\beta$ -conglycinin alpha and alpha-prime subunits and the best AMP candidates revealed that the peptides are coincidentally encrypted in two domain regions. Uniprot database information revealed that the  $\beta$ -conglycinin subunit (F7J05) is composed of 605 amino acid residues, where signal peptide sequence is represented by amino acids residues from 1 to 22, while residues from 23 to 605 are inferred to as the functional region of the protein (available at <https://www.uniprot.org/uniprot/F7J075>). However, the available information at the PDB structure database and other databanks comprised only part of the structure of the molecule, beginning at 195–598 amino acid residues (available at <https://swissmodel.expasy.org/repository/uniprot/F7J075?csm=B6198E8E4610AD25>). Surprisingly, the portion of the sequence in which the new AMP candidates were identified by mass spectrometry analysis did not comprise the available structure stretch at the databank repository. According to Qi, Wilson and Tan-Wilson (1992) and Qi, Chen, Wilson and Tan-Wilson (1994), alpha and alpha-prime subunits are used during germination as a source of amino acids until the plant becomes fully autotrophic. Therefore, both subunits are subjected to sequential proteolysis, initially by protease C1, which acts in the region between residues 22–195, where our AMP candidates are found. For this reason, the total three-dimensional structure of 22–605 residues of  $\beta$ -conglycinin was determined in this study, revealing an alpha-helix structure where the predicted peptides described as antimicrobial candidates are located. To date, no information on the antimicrobial activity of this predicted region had been described, but it is known that this is a domain region comprising the Cupin-like superfamily (available at <https://www.ncbi.nlm.nih.gov/Structure/mmdb/mmdbsrv.cgi?dps=1&uid=1UIJ>).

It is widely known that AMPs can assume diversified structural features giving rise to 4 families, alpha-helical ( $\alpha$ ) AMPs; beta-sheet ( $\beta$ ) AMPs;  $\alpha\beta$ AMPs and non- $\alpha\beta$  (extended) AMPs (Shagaghif et al., 2018). The primary sequences presented herein suggested an extended structure due to the lack of Cys (C) and, although the preliminary results of structural modeling could be an indicative for the presence of  $\alpha$ -helix, most AMP candidates described herein are composed by 11–57%

proline (P). The propensity of this residue to assume the  $\alpha$ -helix structure decreases as proline content increases, giving rise to random conformations in aqueous solutions, except when in the membrane environment (Li, Goto, Williams, & Deber, 1996). Therefore, the individual secondary structure of the AMP candidates should be carefully evaluated, especially those displaying proline occurrence.

Based on this, studies are in progress in order to evaluate the contribution of individual or combined AMP candidates to bacteria growth inhibition and to determine which one(s) are the most potent. Structural characterization of the selected peptide(s) will, then, be carried out in order to evaluate its/their propensity to assume an  $\alpha$ -helix conformation in aqueous or micellar solutions.

## 5. Conclusions

Antimicrobial peptides are a novel class of biopharmaceuticals that possess significant therapeutic potential. The results of this study indicate that soybean meal peptides pool containing GEQQHEEEERERE-HPQPHPPHEERG; GEQQHEEEEREREHPQPHPPHE; ERQQHGEKEEDE-GEQPRPFPRPRQPHQE; EGEQPRPFPRQPHQE; GEQPRPFPRP; PRPIPFPQ; IPRPRPRPQHPEREPQ; EKSKRILRGLKTLFFLTMVIS-LLL; EQDEREHPRPHQPHQKEEKH; FPFPFRPHQK; HPEREPQPG and EQDERQFPF could possibly be used in the development of biocompatible and biodegradable alternative antimicrobial or antitumoral agents for use in the food industry or as adjuvants in chemotherapy against cancer. There is a dire need to develop and exploit high throughput techniques for drug analysis and antimicrobial/antibiotic screening in the pharmaceutical industry. Encrypted antimicrobial peptides obtained by a sustainable processing from an agroindustry protein-rich by-product without generating hazardous sub products should encourage the efficient industrial-scale production of those antimicrobial peptides. The large scale production and use of multi-functional peptides encrypted in soybean meal proteins is an innovative application of the concept of circular bioeconomy.

## Conflicts of interest

The authors declare no conflicts of interest and the funding sponsors had no role in the study design, data collection, analysis or interpretation, neither in the manuscript preparation, writing or in the decision to publish the results.

## Ethics statements

Protocol approved by the Institutional Ethics Committee for Animal Research at Universidade Federal Fluminense under N° 821-16.

## Acknowledgments

The authors are thankful for the donation of bacteria strains from the ATCC, FDA and CDC collections, kindly provided by FIOCRUZ-INCQS, Rio de Janeiro, Brazil, from the Coleção de Microrganismos de Referência em Vigilância Sanitária-CMRVS. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 [Grant No. 1383744] and by Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) [Grant No. E-26/203.039/2015 and E-26/202.860/2016].

## Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jff.2019.01.024>.

## References

- Agyei, D. (2015). Bioactive proteins and peptides from soybeans. *Recent Patents on Food, Nutrition & Agriculture*, 7, 100–107.
- Amnuaycheewa, P., & de Mejia, E. G. (2010). Purification, characterization, and quantification of the soy allergen profilin (Gly m 3) in soy products. *Food Chemistry*, 119, 1671–1680.
- Buchan, D. W., Minneci, F., Nugent, T. C., Bryson, K., & Jones, D. T. (2013). Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Research*, 41, W349–W357.
- Chi, C.-F., Wang, B., Wang, Y.-M., Zhang, B., & Deng, S.-G. (2015). Isolation and characterization of three antioxidant peptides from protein hydrolysate of bluefin leather jacket (*Navodon septentrionalis*) heads. *Journal of Functional Foods*, 12, 1–10.
- Day, L. (2013). Proteins from land plants – Potential resources for human nutrition and food security. *Trends in Food Science & Technology*, 32, 25–42.
- Del Aguila, E. M., Gomes, L. P., Freitas, C. S., Pereira, P. R., & Paschoalin, V. M. F. (2017). Natural antimicrobials in food processing: Bacteriocins, peptides and chitooligosaccharides. In Atta-ur-Rahman, & M. Ibbal Choldhary (Eds.). *Frontiers in anti-infective drug discovery* (pp. 55–108). Sharjah: Bentham Science Publishers.
- Deslouches, B., & Di, Y. P. (2017). Antimicrobial peptides with selective antimutator mechanisms: Prospect for anticancer applications. *Oncotarget*, 8, 46635–46651.
- Felício, M. R., Silva, O. N., Gonçalves, S., Santos, N. C., & Franco, O. L. (2017). Peptides with dual antimicrobial and anticancer activities. *Frontiers in Chemistry*, 5, 1–9.
- Fisher, K., & Phillips, C. (2008). Potential antimicrobial uses of essential oils in food: Is citrus the answer? *Trends in Food Science & Technology*, 19, 156–164.
- Freitas, C. S., da Silva, G. A., Perrone, D., Vericimo, M. A., Baião, D. S., Pereira, P. R., Paschoalin, V. M. F., & Del Aguila, E. M. (2019). Recovery of antimicrobials and bioaccessible isoflavones and phenolics from soybean (*Glycine max*) meal by aqueous extraction. *Molecules*, 24, 74–92.
- Gaspar, D., Veiga, A. S., & Castanho, M. A. R. B. (2013). From antimicrobial to anticancer peptides A review. *Frontiers in Microbiology*, 4, 1–16.
- Guaadaoui, A., Benaicha, S., Elmajdoub, N., Bellaoui, M., & Hamal, A. (2014). What is a bioactive compound? A combined definition for a preliminary consensus. *International Journal of Nutrition and Food Sciences*, 3, 174–179.
- Hong, H., Landauer, M. R., Foriska, M. A., & Ledney, G. D. (2006). Antibacterial activity of the soy isoflavone genistein. *Journal of Basic Microbiology*, 46, 329–335.
- Kuipers, B. J., Alting, A. C., & Gruppen, H. (2007). Comparison of the aggregation behavior of soy and bovine whey protein hydrolysates. *Biotechnology Advances*, 25, 606–610.
- Laodheerasiri, S., & Pathirage, N. H. (2017). Antimicrobial activity of raw soybean, soybean flour and roasted soybean extracted by ethanol-hexane method. *British Food Journal*, 119, 2277–2286.
- Laufenberg, G., Kunz, B., & Nyström, M. (2003). Transformation of vegetable waste into value added products: (A) the upgrading concept; (B) practical implementation. *Bioresource Technology*, 87, 167–198.
- Lemes, A. C., Sala, L., Ores, J. C., Braga, A. R. C., Egea, M. B., & Fernandes, K. F. A. (2016). Review of the latest advances in encrypted bioactive peptides from protein-rich waste. *International Journal of Molecular Sciences*, 17, 950–962.
- Li, S. C., Goto, N. K., Williams, K. A., & Deber, C. M. (1996). Alpha-helical, but not beta-sheet, propensity of proline is determined by peptide environment. *Proceedings of the National Academy of Sciences*, 93, 6676–6681.
- Li, W., Tailhades, J., O'Brien-Simpson, N. M., Separovic, F., Otvos, L., Hossain, M. A., & Wade, J. D. (2014). Proline-rich antimicrobial peptides: Potential therapeutics against antibiotic-resistant bacteria. *Amino Acids*, 46, 2287–2294.
- Liu, S., Fan, L., Sun, J., Lao, X., & Zheng, H. (2017). Computational resources and tools for antimicrobial peptides. *Journal of Peptide Science*, 23, 4–12.
- Lobley, A., Sadowski, M. I., & Jones, D. T. (2009). pGenTHREADER and pDomTHREADER: New methods for improved protein fold recognition and superfamily discrimination. *Bioinformatics*, 25, 1761–1767.
- Luna-Vital, D. A., Mojica, L., González de Mejía, E., Mendoza, S., & Loarca-Piña, G. (2015). Biological potential of protein hydrolysates and peptides from common bean (*Phaseolus angularis* L.): A review. *Food Research International*, 76, 39–50.
- McMillian, M. K., Li, L., Parker, J. B., Patel, L., Zhong, Z., Gunnell, J. W., ... Johnson, M. D. (2002). An improved resazurin-based cytotoxicity assay for hepatic cells. *Cell Biology and Toxicology*, 18, 157–173.
- Morel, C., Stermitz, F. R., Tegos, G., & Lewis, K. (2003). Isoflavones as potentiators of antibacterial activity. *Journal of Agricultural and Food Chemistry*, 51, 5677–5679.
- Najafian, L., & Babji, A. S. (2015). Isolation, purification and identification of three novel antioxidative peptides from patin (*Pangasiususutchi*) myofibrillar protein hydrolysates. *LWT-Food Science and Technology*, 60, 452–461.
- Natarajan, S. S., Xu, C., Bae, H., Caperna, T. J., & Garrett, W. M. (2006). Characterization of storage proteins in wild (*Glycine soja*) and cultivated (*Glycine max*) soybean seeds using proteomic analysis. *Journal of Agricultural and Food Chemistry*, 54, 3114–3120.
- Negi, P. S. (2012). Plant extracts for the control of bacterial growth: Efficacy, stability and safety issues for food application. *International Journal of Food Microbiology*, 156, 7–17.
- Nunes, R. S. C., de Souza, C. P., Pereira, K. S., Del Aguila, E. M., & Paschoalin, V. M. F. (2016). Identification and molecular phylogeny of coagulase-negative staphylococci isolates from Minas Frescal cheese in southeastern Brazil: Superantigenic toxin production and antibiotic resistance. *Journal of Dairy Science*, 99, 2641–2653.
- Peña-Ramos, E. A., & Xiong, Y. L. (2002). Antioxidant activity of soy protein hydrolysates in a liposomal system. *Journal of Food Science*, 67, 2952–2956.
- Pinhati, F. R., Del Aguila, E. M., Tôrres, A. P. R., Sousa, M. P. D., Santiago, V. M. J., Silva, J. T., & Paschoalin, V. M. F. (2014). Evaluation of the efficiency of deterioration of aromatic hydrocarbons by bacteria from wastewater treatment plant of oil refinery.

- Química Nova*, 37, 1269–1274.
- Prabhu, S., Dennison, S. R., Lea, B., Snape, T. J., Nicholl, I. D., Radecka, I., & Harris, F. (2013). Anionic antimicrobial and anticancer peptides from plants. *Critical Reviews in Plant Sciences*, 32, 303–320.
- Prabhu, S., Dennison, S. R., Mura, M., Lea, R. W., Snape, T. J., & Harris, F. (2014). Cn-AMP2 from green coconut water is an anionic anticancer peptide. *Journal of Peptide Science*, 20, 909–915.
- Qi, X., Chen, R., Wilson, K. A., & Tan-Wilson, A. L. (1994). Characterization of a soybean [beta]-conglycinin-degrading protease cleavage site. *Plant Physiology*, 104, 127–133.
- Qi, X., Wilson, K. A., & Tan-Wilson, A. L. (1992). Characterization of the major protease involved in the soybean  $\beta$ -conglycinin storage protein mobilization. *Plant Physiology*, 99, 725–733.
- Rayaprolu, S. J., Hettiarachchy, N. S., Chen, P., Kannan, A., & Mauromostakos, A. (2013). Peptides derived from high oleic acid soybean meals inhibit colon, liver and lung cancer cell growth. *Food Research International*, 50, 282–288.
- Sánchez, A., & Vázquez, A. (2017). Bioactive peptides: A review. *Food Quality and Safety*, 1, 29–46.
- Santiago, L. G., Maldonado-Valderrama, J., Martín-Molina, A., Haro-Pérez, C., García-Martínez, J., Martín-Rodríguez, A., & Gálvez-Ruiz, M. J. (2008). Adsorption of soy protein isolate at air–water and oil–water interfaces. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 323, 155–162.
- Shagaghi, N., Palombo, E. A., Clayton, A. H., & Bhave, M. (2018). Antimicrobial peptides: Biochemical determinants of activity and biophysical techniques of elucidating their functionality. *World Journal of Microbiology & Biotechnology*, 34, 1–13.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., ... Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7, 539.
- Singh, A., Meena, M., Kumar, a. D., Dubey, A. K., & Hassan, I. (2015). Structural and functional analysis of various globulin proteins from soy seed. *Critical Reviews in Food Science and Nutrition*, 55, 1491–1502.
- Singleton, V. L., Orthofer, R., & Lamuela-Raventós, R. M. (1999). Analysis of total phenols and other oxidation substrates and antioxidants by means of Folin-Ciocalteau reagent. In Methods of enzymology (pp. 152–178). New York, NY.
- Sitohy, M. Z., Mahgoub, S. A., & Osman, A. O. (2012). *In vitro* and *in situ* antimicrobial action and mechanism of glycinin and its basic subunit. *International Journal of Food Microbiology*, 154, 19–29.
- Stupp, R., Mason, W. P., Van Den Bent, M. J., Weller, M., Fisher, B., Taphoorn, M. J., & Curschmann, J. (2005). Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *New England Journal of Medicine*, 352, 987–996.
- Tam, J. P., Wang, S., Wong, K. H., & Tan, W. L. (2015). Antimicrobial peptides from plants. *Pharmaceuticals*, 8, 711–757.
- Thomas, S., Karnik, S., Barai, R. S., Jayaraman, V. K., & Idicula-Thomas, S. (2009). CAMP: A useful resource for research on antimicrobial peptides. *Nucleic Acids Research*, 38, D774–D780.
- Torrent, M., Di Tommaso, P., Pulido, D., Nogués, M. V., Notredame, C., Boix, E., & Andreu, D. (2011). AMPA: An automated web server for prediction of protein antimicrobial regions. *Bioinformatics*, 28, 130–131.
- Vasconcellos, F. C. S., Woiciechowski, A. L., Soccol, V. T., Mantovani, D., & Soccol, C. R. (2014). Antimicrobial and antioxidant properties of-conglycinin and glycinin from soy protein isolate. *International Journal of Current Microbiology and Applied Sciences*, 3, 144–157.
- Wagh, F. H., Barai, R. S., Gurung, P., & Idicula-Thomas, S. (2015). CAMPR3: A database on sequences, structures and signatures of antimicrobial peptides. *Nucleic Acids Research*, 44, 1094–1097.
- Wang, T., Qin, G. X., Sun, Z. W., & Zhao, Y. (2014). Advances of research on glycinin and  $\beta$ -conglycinin: A review of two major soybean allergenic proteins. *Critical Reviews in Food Science and Nutrition*, 54, 850–862.
- Webb, B., & Sali, A. (2016). Comparative protein structure modeling using MODELLER. *Current Protocols in Bioinformatics*, 47, 5–6.
- Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2015). The I-TASSER Suite: Protein structure and function prediction. *Nature Methods*, 12, 7–8.
- Yu, G., Baeder, D. Y., Regoes, R. R., & Rolff, J. (2016). The more the better? Combination effects of antimicrobial peptides. *Antimicrobial Agents and Chemotherapy*, 60, 1717–1724.
- Zar, J. H. (1984). *Biostatistical Analysis* (2nd ed.). Englewood Cliffs: Prentice-Hall Inc718.
- Zhu, K.-X., Wang, X.-P., & Guo, X.-N. (2015). Isolation and characterization of zinc-chelating peptides from wheat germ protein hydrolysates. *Journal of Functional Foods*, 12, 23–32.
- Zimmermann, L., Stephens, A., Nam, S. Z., Rau, D., Kübler, J., Lozajic, M., ... Alva, V. (2018). A completely Reimplemented MPI bioinformatics toolkit with a new HHpred server at its Core. *Journal of molecular biology*, 430, 2237–2243.