PEPTIDES 29 (2008) 1271–1279

available at www.sciencedirect.com

ScienceDirect

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

Identification of a novel storage glycine-rich peptide from guava (*Psidium guajava*) seeds with activity against Gram-negative bacteria

Patricia B. Pelegrini [a], *André M. Murad* [a], *Luciano P. Silva* [b], *Rachel C.P. dos Santos* [a], *Fabio T. Costa* [a], *Paula D. Tagliari* [a], *Carlos Bloch Jr* [b], *Eliane F. Noronha* [a], *Robert N.G. Miller* [a], *Octavio L. Franco* [a,*]

[a] *Centro de Análises Proteômicas e Bioquímicas, Programa de Pós-Graduação em Ciências Genômicas e Biotecnologia, Universidade Católica de Brasília, Brasília-DF, Brazil*

[b] *Embrapa Recursos Genéticos e Biotecnologia, Brasília-DF, Brazil*

ARTICLE INFO

Article history:
Received 16 January 2008
Received in revised form
23 February 2008
Accepted 11 March 2008
Published on line 22 March 2008

Keywords:
Psidium guajava
Toxin
Glycine-rich proteins
Antimicrobial peptides

ABSTRACT

Bacterial pathogens cause an expressive negative impact worldwide on human health, with ever increasing treatment costs. A significant rise in resistance to commercial antibiotics has been observed in pathogenic bacteria responsible for urinary and gastro-intestinal infections. Towards the development of novel approaches to control such common infections, a number of defense peptides with antibacterial activities have been characterized. In this report, the peptide *Pg*-AMP1 was isolated from guava seeds (*Psidium guajava*) and purified using a Red-Sepharose Cl-6B affinity column followed by a reversed-phase HPLC (Vydac C18-TP). *Pg*-AMP1 showed no inhibitory activity against fungi, but resulted in a clear growth reduction in *Klebsiella* sp. and *Proteus* sp., which are the principal pathogens involved in urinary and gastro-intestinal hospital infections. SDS-PAGE and mass spectrometry (MALDI-ToF) characterized *Pg*-AMP1 a monomer with a molecular mass of 6029.34 Da and small quantities of a homodimer. Amino acid sequencing revealed clear identity to the plant glycine-rich protein family, with *Pg*-AMP1 the first such protein with activity towards Gram-negative bacteria. Furthermore, *Pg*-AMP1 showed a 3D structural homology to an enterotoxin from *Escherichia coli*, and other antibacterial proteins, revealing that it might act by formation of a dimer. *Pg*-AMP1 shows potential, in a near future, to contribute to development of novel antibiotics from natural sources.

* *Corresponding author at*: Centro de Análises Proteômicas e Bioquímicas, Programa de Pós-Graduação em Ciências Genômicas e Biotecnologia, Universidade Católica de Brasília, SGAN Quadra 916, Módulo B, Av. W5 Norte 70, 790-160 Asa Norte, Brasília-DF, Brazil. Tel.: +55 61 3448 7220; fax: +55 61 3347 4797.
E-mail address: ocfranco@pos.ucb.br (O.L. Franco).

Abbreviations: BtuB, surface receptor of colicin E3 from *E. coli*; DDR, D-alanyl-D-alanine residue; ESI-Q-ToF, electrospray-quadruple-time of flight; GlcNAc, N-acetylglucosamine; GRP, glycine-rich protein; HCl, chloridric acid; HPLC, high performance liquid chromatography; LB, Luria Bertani medium; LTP, lipid-transfer protein; MALDI-ToF, matrix-assisted laser desorption/ionization time of flight; MIC, minimum inhibitory concentration; MS, mass spectrometry; *Pg*-AMP1, antimicrobial peptide from *Psidium guajava*; Pp-TH, thionin from *Pyrularia pubera*; RIP, ribosome-inactivating protein; SDS-PAGE, sodium dodecyl sulphate polyacrylamide gel electrophoresis; *Tu*-AMP1, antimicrobial peptide from *Tulipa gesneriana*.

0196-9781/$ – see front matter
doi:10.1016/j.peptides.2008.03.013

1. Introduction

In the last two decades, a significant increase in pathogenic bacteria with enhanced resistance to commercial antimicrobials has been observed. Members of the enterobacteriaceae, such as *Escherichia coli*, *Klebsiella* sp. and *Proteus* sp., which were once susceptible to common antibiotics, are now one of the main causes of hospital infections across developing countries, with a higher mortality observed in immunosuppressed patients infected with such bacteria [27]. In response to this problem, a number of proteinaceous compounds with antimicrobial activities have been isolated from sources across plant, mammalian and microbial taxa [2,4].

Bacteria possess a number of mechanisms of resistance to antibiotics. Studies with methicillin-resistant *Staphylococcus aureus* towards vancomycin revealed that bacteria could enhance the production of *N*-acetylglucosamine (GlcNAc) in the cell wall, a murein monomer precursor (UDP-*N*-acetylmuramyl-pentapeptide), and also increase the production of penicillin-binding proteins [6], forming a thicker layer over the cell membrane. Consequently, when bacteria strengthen cell walls, antibiotic passage to the cytoplasm is inhibited, enhancing resistance [6]. In addition, some bacteria have developed resistance towards antibiotics such as vancomycin. This drug usually competes for the same substrates required by glycosyltransferase to produce peptidoglycan chains, thus leading to bacterial cell death. However, the acquirement of different D-alanyl-D-alanine residues (DDR) of lipid II precursor by these bacteria, which is important for vancomycin interaction with the cell surface, might work as false targets to vancomycin, leading to its inefficiency as an antibacterial product [6].

A number of plant proteins have demonstrated activity against bacteria and fungi, which usually play a key role in plant defense, protecting the plant host from phytopathogens [41]. These proteins have been classified into groups and families according to their function and structural similarities. Main proteins include cyclotides, defensins, lipid-transfer proteins (LPTs), lectins, pathogenesis-related proteins, chitin-binding proteins, ribosome-inactivating proteins and digestive enzyme inhibitors, amongst others [12,36]. Furthermore, five classes of R gene products have been described that specify resistance to many pathogens, such as viruses, bacteria and fungi. Resistance genes have been found across numerous plant species, with the most common family characterized by a nucleotide binding site (NBS) and a leucine-rich repeat region (LRR) [25,11]. Another R gene family contains genes such as *Cf-9* and *Cf-4* from tomato, which recognize the fungus *Cladosporium fulvum* through an interaction of a short transmembrane domain with products of *Avr9* and *Avr4* genes [17,47]. This recognition leads to the activation of multiple defense responses that impede fungal growth [17,47]. Some R gene products do not contain leucine-rich motifs, although still maintaining function, such as *RPW8.1* and *RPW8.2* from *Arabidopsis*, which demonstrated the ability to confer resistance to four different pathogens [51,52]. A group of R genes with receptor-like kinases or serine/threonine kinase domains have also been described, such as Xa21 from rice and FLS2 from *Arabidopsis*, respectively [9].

A new group of defense molecules from plant sources with activity against bacteria are the glycine-rich proteins (GRPs). They were first described as storage proteins, as they are useful in the plant as a source of essential amino acid residues [29]. GRPs can be characterized by their high percentage of glycine residues in their primary sequence, with residue number varying from one organism to another [29]. GRPs can be classified into three different groups, according to glycine content. The first group contains proteins that possess more than 70% of glycine residues in their primary structure, such as GRP from *A. thaliana* and *Brassica napus* [38]. This group is also distinct in that proteins possess conserved domains known as RNA recognition motifs (RRMs), which can be an RNP-1 (ribonucleoprotein motif) or an RNP-2 [32]. GRPs with fewer glycine residues belong to a second group, with examples in tomato and saltbush [38]. A third group of GRPs includes proteins with increased overall content of Gly, without the presence of glycine-rich domains [38]. GRP are usually more hydrophobic than hydrophilic, with tyrosine and phenylalanine residues possibly responsible for this feature [38].

Previous studies revealed that GRP genes are mainly found in vascular tissues, especially in xylem [18,19,39], although they have also been observed in hypocotyls, stems and petioles [50]. Expression of GRP genes appears to be stress-induced and influenced by environmental changes, such as wounding, hormone treatment, water and ozone stresses [1,5,8,18,21,53]. Activities have been demonstrated in cold adaptation, enhancing freezing tolerance and also acting as RNA binding factors [42]. Their involvement in changes in germination and seedling growth under stress conditions, such as salt and dehydration have also been noted [20]. Perhaps most interestingly, fungicidal activity against a number of phytopathogenic fungi has been observed with GRPs, with eight proteins isolated from *Triticum kiharae* seeds displaying fungicidal activity against *Helminthosporium sativum* and *Fusarium culmorum* [10]. Based upon these recent studies, it appears that such storage proteins might play an important role in plant defense mechanisms [10,36,49]. It seems that these proteins perhaps function as storage molecules when not required for plant defense as antimicrobial components.

This study was conducted to characterize *Pg*-AMP1 from guava seed material, a peptide with similarity to glycine-rich proteins. Inhibitory activity was observed against three Gram-negative bacterial pathogens, responsible for nosocomial, intestinal and urinary infections in immuno-suppressed patients. This is the first report of inhibitory activity in plant GRP against human bacterial pathogens.

2. Material and methods

2.1. Extraction and isolation of Pg-AMP1

Psidium guajava proteins were extracted from dry seeds, using a solution of 0.6 M NaCl and 0.1% HCl (1:3, w/v). Crude extracts were centrifuged and supernatant submitted to precipitation with ammonium sulphate (100%). Following precipitation, the rich fraction was centrifuged at 5000 rpm for 20 min at 4 °C. Precipitate was re-suspended, dialyzed against distilled water and applied onto a Red-Sepharose affinity column equilibrated

with 0.15 M Tris–HCl buffer pH 7.0 containing 5.0 mM $CaCl_2$. The non-retained fraction was displaced with equilibration buffer and the retained peak eluted with a single step of 3.0 M NaCl dissolved in equilibration buffer. Following dialyses and liophilization, 1.0 mg of retained fraction was diluted with 0.1% trifluoroacetic acid and applied onto a HPLC reversed-phase chromatography (Vydac C-18TP column). Retained proteins were eluted with a linear acetonitrile gradient (0–100%). A fraction eluted at 42% acetonitrile was re-chromatographed into a reversed-phase column (HPLC), and samples were eluted using a linear gradient of acetonitrile (0–100%).

2.2. Molecular mass analyses

Pg-AMP1 was analyzed by 15% SDS-PAGE according to Laemmli et al. [22], with minor modifications. Gels were silver stained and bromophenol blue used as tracking dye. HPLC samples were also analyzed by ESI-Q-TOF (Bruker). 2.0 $\mu g\,l^{-1}$ of peptide solution was injected in a Vydac C-18 column, using two solvents A (0.1% formic acid) and B (0.1% formic acid; 99% acetonitrile). Peptides were separated using a linear gradient of solvent B and applied on a mass spectrometer Electrospray Q-TOF, adjusted in a nanoLock-Spray source, at a speed of 250 nl min^{-1}. Mass specter was acquired to a *m/z* of 300–1600 to 1 s, followed by 3 MS/MS scans of m/s, varying from 50 to 1900 s^{-1}. Collision energy used for MS/MS varied according to the mass and charge state of eluted peptide. Leu-enk, a mass calibrator, or Lockmass, was used in a concentration range of 250 nl/min and acquired every 2 min through the run [28].

2.3. Amino acid sequencing

Samples obtained from HPLC were dried in order to remove acetonitrile and trifluoracetic acid compounds. Peptide sequences were determined using MS analysis as previously described, with minor modifications [31,48]. Protein samples were incubated with 10 $\mu g\,ml^{-1}$ bovine trypsin (Sigma) dissolved in 2.5 mM Tris–HCl pH 8.5 for 24 h at 37 °C. Tryptic fragments were then eluted in 50% acetonitrile and 0.5% trifluoracetic acid by facilitated diffusion, using ultrasonication. Tryptic protein fragments were analyzed by MALDI-TOF mass spectrometry, using 2 μl of sample, α-cyano-4-hydroxycinnamic acid as a matrix (10 mg ml^{-1} in ethanol/acetonitrile; 1:1 (v/v)), and analyzed on an Applied Biosystems Q-STAR MS/MS mass spectrometer. Spectra were obtained in a reflection-delayed extraction mode over a mass range of 1000–20,000 Da. Spectra from 64 shots at 20 different positions were continued to generate a protein mass fingerprint for the sample. The protein ions generated by autolysis of trypsin were used as internal standards for calibrating the mass spectra. Protein masses were analyzed with the MS-Fit Program using the following parameters: mass tolerance of 0.5 Da, a minimum of four peptide matches, and no missed cleavages. When isoforms were found for proteins, FindMod and PeptideMass tools in ExPaSy were used to predict possible protein modifications. Hydropathy analysis of the deduced sequences of the novel proteins was performed using TopPred2, and homology searches performed using FASTA3 and BioEdit.

2.4. In silico analyses and molecular modeling

An alignment using ClustalW [46] and BioEdit [34] was performed for the proteins in order to analyze primary sequence similarities in each group. A phylogenetic tree was visualized using TreeView [15], including protein from *P. guajava* and other glycine-rich plant proteins. Alignment against data bank was conducted to elucidate the 3D structure of *Pg*-AMP1 [30]. Bioinfo Meta Server [13] was utilized, indicating an absence of structures similar to guava peptide. Consequently, Gromacs for molecular dynamics [24] was utilized for prediction of protein structures using *ab initio* procedures. A primary structure was constructed using Deep View Swiss PDB Viewer [14] and Modeller [26] programs. This structure was then placed in a cubic box, with water molecules. A partial simulation for energy minimization was executed using 2000 steps from Steepest Descent in order to remove possible stereochemical disturbances. Complete simulation was performed using temperature and coupling pressure of 300 K and 1 atm, respectively, and Newton's movement equations (MD) as a dynamic method in 60,000 ps. The final model was visualized using Pymol [7]. All simulations were performed on a Sun AMD Opteron bioprocessor workstation.

2.5. Bacterial bioassays

Bacterial bioassays were performed in micro-plates using 1.0 ml of Luria Bertani medium (10 g l^{-1} NaCl, 5 g l^{-1} yeast extract and 45 g l^{-1} bactopeptone). Liquid cultures of *Klebsiella* sp. and *Proteus* sp. were grown in LB medium during 18–24 h at 37 °C, prior to initiating peptide evaluations, then subsequently challenged by *Pg*-AMP1 peptide from guava seeds at a concentration of 40 $\mu g\,ml^{-1}$, for 4 h at 37 °C. Sterile distilled water was used as a negative control and 40 $\mu g\,ml^{-1}$ chloramphenicol as positive control. Evaluation of bacterial growth was measured at hourly intervals at 595 nm. Each experiment was carried out in triplicate. Moreover, the minimal inhibitory concentrations MIC was also obtained according to Park et al. [35]. The microorganisms used for MIC bactericidal assay were *E. coli* ATCC 8739, *Klebsiella pneumoniae* ATCC 13883 and *Salmonella typhimurium* ATCC 14028. Bacteria were grown in tryptic soy broth (TSB) at 37 °C, growth inhibition were measured as described by Bulet et al. [3]. *Pg*-AMP1 was dissolved in sterile Milli-Q water and diluted 8-fold in TSB (OXIOD, England) broth. The standard concentrations used for the assay were 25, 50, 100 and 150 $\mu g\,ml^{-1}$. The initial inoculums were approximately 1×10^5 colony forming units (CFU ml^{-1}) and limit of detection was 10^2 CFU ml^{-1}. The final volume was 250, 25 μl of the *Pg*-AMP1 test in water, 25 μl of the inoculums in TSB, and 200 μl of TSB broth. The minimal inhibitory concentration (MIC) was measured for turbidity (OD at 595 nm) 20 h after all microorganisms were grown in stationary culture at 37 °C. The lowest concentration of the peptide in which no growth occurred was defined as the MIC.

2.6. Fungi bioassays

The inhibitory effect of *Pg*-AMP1 on growth of filamentous fungi *Trichoderma harzianum*, *Aspergillus fumigatus*, *Fusarium*

oxysporum and *Rhizoctonia solani* was assessed according to Oard et al. [33]. These microorganisms were grown in 20 ml MYG medium (0.5% malt extract, 0.25% yeast extract, 1.0% glucose, and pH 6.0) for 48 h, at 25 °C, in the presence of *Pg*-AMP1 at concentrations of 25, 50, 75, and 100 μg ml^{-1}. Sterile distilled water was used as negative control and 0.5% Capitan as positive control. Evaluation of growth inhibition in comparison with both controls was obtained by measuring fungal dry weight 48 h after incubation with test and control substances.

3. Results

3.1. *Purification and molecular characterization of guava peptide*

In order to isolate antimicrobial peptides, the rich fraction from guava seeds was dialyzed and applied onto a Red-Sepharose Cl-6B affinity column. A single retained peak was active against *Klebsiella* sp. and *Proteus* sp. while samples from non-retained peaks did not show such antimicrobial activity (Fig. 1A). The retained peak showed innumerous proteins when analyzed by SDS-PAGE, with molecular masses ranging from 5.0 to 80.0 kDa (Fig. 2). A retained fraction from guava Red-Sepharose was further analyzed via reversed-phase chromatography (HPLC), revealing a major peak eluted with 42% acetonitrile (Fig. 1B). Samples from this peak were re-chromatographed using a retard step of a 20–50% acetonitrile gradient, revealing a purified peptide 42% (Fig. 1B). After re-chromatography, the HPLC sample was submitted to molecular mass analysis by SDS-PAGE, showing a single peptide with molecular weight below 10 kDa (Fig. 2). Mass spectrometry analysis from peak 1 of HPLC showed a major monomeric peak with 6184.12 Da and small quantities of a dimeric form with 12368.25 (data not shown).

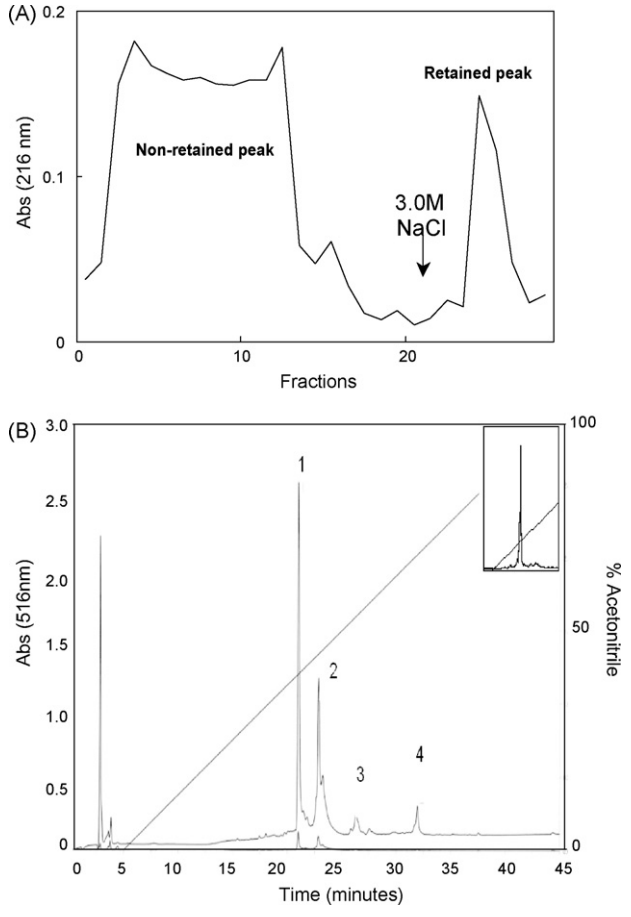

Fig. 1 – (A) Red-Sepharose Cl-6B chromatography profile of rich fraction extracted from guava seeds. Black arrows indicate initiation of protein elution using Tris–HCl buffer containing 3.0 M NaCl. (B) HPLC reversed-phase chromatogram profile (Vydac C18-TP) of Red-Sepharose retained fraction generated from guava seeds. Diagonal line indicates a linear acetonitrile gradient. Graphics on the right hand side of both chromatograms represent re-chromatography of protein fractions using the same column.

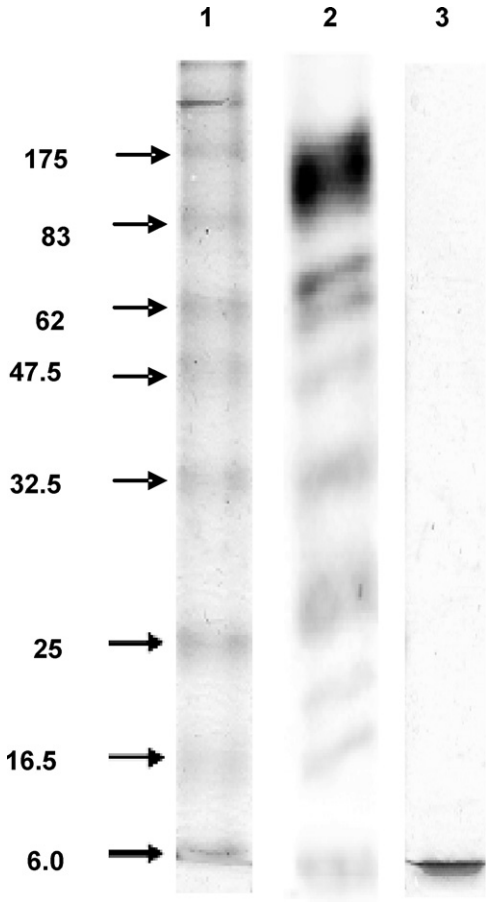

Fig. 2 – SDS-PAGE of samples from retained peak (lane 2) and purified *Pg*-AMP1 (lane 3). Lane 1 corresponds to molecular marker. The black arrow indicates the guava protein with molecular mass under 10 kDa.

3.2. *Amino acid sequencing and alignment*

Complete sequencing of the guava protein revealed a protein with a length of 55 amino acid residues (Fig. 3). A search using FASTA3 revealed 72% identity to glycine-rich proteins. A

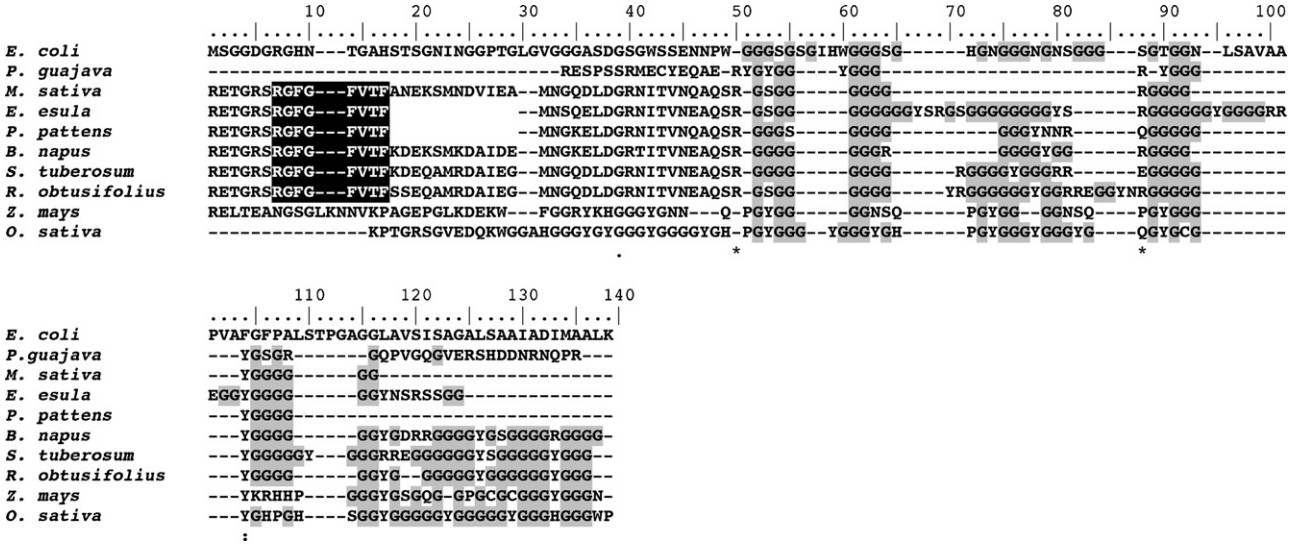

Fig. 3 – Alignment of *Pg*-AMP1 and seven other plant glycine-rich proteins and colicin E3 from *E. coli*. Gray boxed residues indicate the glycine-rich region. Black boxed residues correspond to conserved RRM motif. Asterisks indicate conserved arginine residues. Two points indicate conserved tyrosine.

comparison of *Pg*-AMP1 primary sequence with eight additional glycine-rich proteins from plant sources was performed using BioEdit [15]. Although the primary sequences between glycine-rich proteins were, in general, widely diverse, they showed a well-conserved region where glycine residues are predominant (Fig. 3). Two positively charged residues in the sequence were also conserved in most of the analyzed plant glycine-rich proteins (indicated by asterisks) (Fig. 3). Moreover, with the exception of colicin E3, a conserved tyrosine residue was also observed across the alignment (Fig. 3).

Phylogenetic groups based upon peptide data were not found to be in accordance with current taxonomic status of the species analyzed. For example, although *P. guajava*, *S. tuberosum* and *R. obtusifolius* are taxonomically close, they group separately in the tree (Fig. 4). Although glycine-rich proteins from *Z. mays* and *O. sativa* were phylogenetically close, which supports current phylogeny for these two members of the liliopsida, they also unexpectedly grouped with the guava peptide *Pg*-AMP1. Moreover, although *M. sativa* and *B. napus* are closely related taxonomically, their glycine-rich proteins were not so close in the phylogenetic tree. Colicin E3 was very close to *Pg*-AMP1, which may be more due to their functional similarities rather than sequence similarities.

3.3. *Molecular modeling*

The final model is composed of two α-helices, one at the N-terminus and the other at the C-terminus, with one loop between them (Fig. 5A). Most of the glycine residues are present within this loop, conferring greater flexibility to this structure. At the extremities of the α-helices are arginine residues, which seem to be important for providing the positive charge to this region (Figs. 5A and C). Several hydrophobic residues could also be observed along the structure, which may be important for dimerization (Fig. 3). In such a case, we could propose that, via *in silico* analyses, there is a salt bound formation between Arg^{7} from one monomer to Glu^{108} from the other monomer. Another interaction mediated by a water molecule can also occur between Arg^{45} from one of the monomers with Ser^{88} from the other. The other arginine residue seems to be important in antimicrobial activity, interacting with the pathogen cell surface (Fig. 5C). Analyses at the electrostatic layer showed two charged regions, positive at the N-terminal and negative at the C-terminal. Positively charged regions are mainly composed of arginine residues, which are located at the core of the dimer. Some of these amino acids from each monomer interact to form the dimeric molecule. The negatively charged region is formed predominantly of glutamic acid residues and can be found mainly at the external α-helices (Fig. 5B). These regions may also allow the two *Pg*-AMP1 molecules to interact with each other, forming dimers. The structure formed by two *Pg*-AMP1 molecules was analyzed by molecular dynamics, with six different positions compared. Fig. 5A illustrates the different structural possibilities for the dimerical shape of this peptide. Each color represents a unique structure. Upon exposure to a water environment, *Pg*-AMP1 structure was observed to remain stable, suggesting the possibility of homodimer formation for antibacterial activity. Analyses of *Pg*-AMP1 three-dimensional structure showed that this molecule showed high similarities with colicin E3, a protein from *E. coli* related to antibiotic proteins with bactericidal properties [43]. Both are formed by α-helices, which seems to be crucial for interaction with

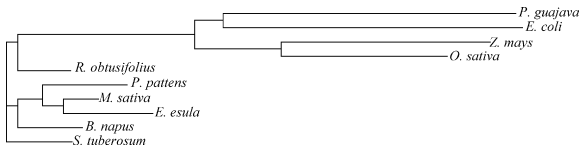

Fig. 4 – Phylogenetic tree of eight plant glycine-rich proteins.

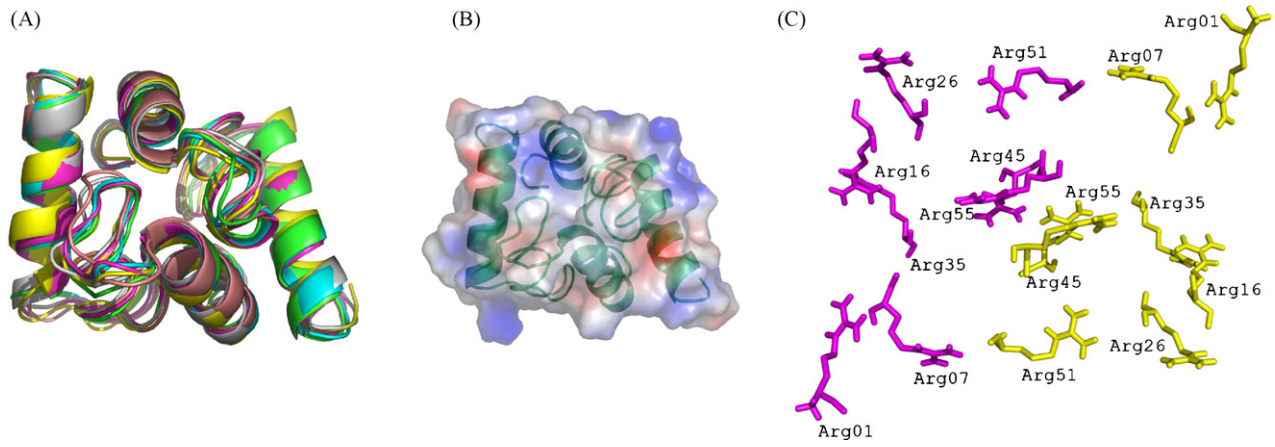

Fig. 5 – Cartoon representation of *Pg*-AMP1. (A) Front view of six different structures obtained from molecular dynamics during different time periods. (B) View of the electrostatic layer of the dimeric molecule. (C) Diagram showing arginine residues from *Pg*-AMP1 at dimeric formation. Residues in pink represent one of the monomers, while residues in yellow correspond to the other monomer.

pathogens and antimicrobial activity. The mechanism of action of colicin E3 is based on a strong interaction with receptors, such as BtuB, related to B_{12} vitamin synthesis in *E. coli*. Colicins usually act binding to these receptors at the cell membrane surface, leading to the entrance of the antimicrobial protein into the cell cytoplasm and consequently cell intoxication [43]. The three-dimensional structure of colicin E3 is composed of an A domain and two T and C globular domains. A large antiparallel helicoidal hairpin characterizes ''A'' domain, which is the region that interacts with the microorganism cell surface [43].

3.4. *Antimicrobial assays of guava seeds peptide*

Purified *Pg*-AMP1 was used for *in vitro* assays against Gram-negative bacteria. *Pg*-AMP1 revealed deleterious effects against *Klebsiella* sp. and *Proteus* sp., showing 90% and 30% inhibition, respectively, at a concentration of 40 $\mu g\ ml^{-1}$ (Fig. 6A and B). Results showed specificity to Gram-negative bacteria, while bioassays against strains of Gram-positive *S. aureus* did not reveal inhibitory activity (data not shown). With this preliminary data, MICs of *Pg*-AMP1 were obtained against *E. coli* ATCC 8739, *K. pneumoniae* ATCC 13883 e *S. typhimurium* ATCC 14028. In fact *Pg*-AMP1 showed a clear bactericidal activity, where MIC values obtained for this peptide were 72 $\mu g\ ml^{-1}$ for *E. coli* and 32 $\mu g\ ml^{-1}$ for *K. pneumoniae*. None activity was obtained towards *S. typhimurium*. This is the first report, to our knowledge, of a glycine-rich protein from plant sources showing antimicrobial activity against eubacteria pathogenic to humans. Purified *Pg*-AMP1 was also tested against human and plant pathogenic fungi: *F. oxysporum*, *A. fumigatus* and *R. solani*. Evaluation was also conducted against *T. harzianum*, a widely used model filamentous fungus for biochemical studies, once that it is an entomopathogenic organism. None of these fungi showed growth inhibition in the presence of *Pg*-AMP1 (data not shown), suggesting that its mode of action is specific to prokaryotic eubacterial cells.

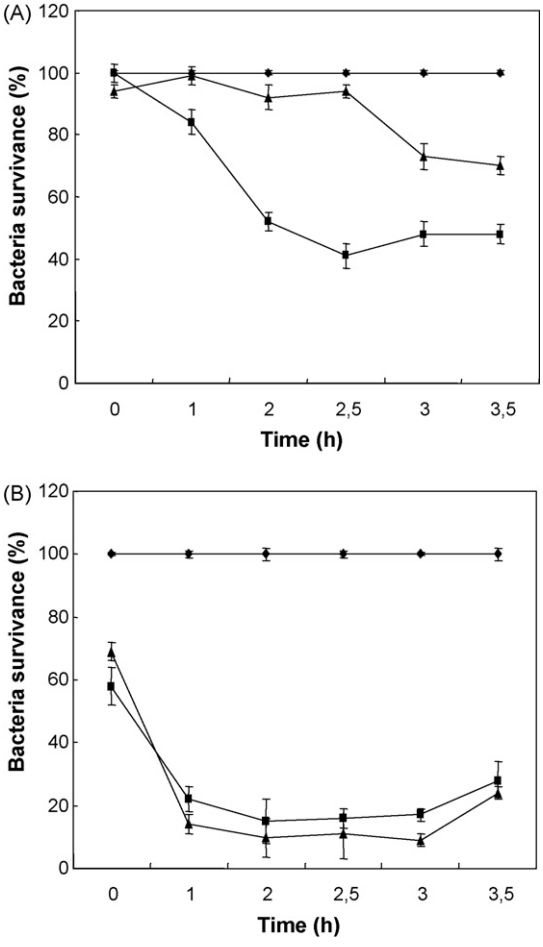

Fig. 6 – Bacterial bioassays using purified *Pg*-AMP1 against (A) *Proteus* sp. and (B) *Klebsiella* sp. (◆) indicates negative control, (■) indicates positive control and (▲) indicates *Pg*-AMP1.

4. Discussion

Sequencing of *Pg*-AMP1 revealed a fragment of 55 amino acid residues with similarity to glycine-rich plant proteins (Fig. 3). Such proteins form a wide divergent protein family with different functions and, in general, without primary sequence similarities. Their main feature is the presence of approximately 60% glycine residue content in their primary sequence, although they can be classified into three different groups according to precise number of Gly residues [29,38]. The first group shows not only a higher content of glycine in primary sequences, comprising over 70% of all residues, but also a number of conserved RRM motifs, namely RNP-1 and RNP-2. RNP-1 consists of eight amino acid residues (K/R)G(F/Y)(G/A)FVX(F/Y), while RNP-2 contains six amino acid residues with a consensus (L/I)(F/Y)(V/I)(G/K)(G/N)L. The second group in this protein family is also characterized by high glycine content and the presence of only one of these conserved domains. Finally, the third group includes proteins with no conserved motif, but with 60% of the primary sequence composed of glycine residues [32,42]. *Pg*-AMP1 seems to be classified into the third group, given that this protein does not contain RRMs present in proteins from the first two groups [29,32].

Sequence alignment with eight plant glycine-rich proteins showed that, despite primary sequence divergence, there is a considerable similarity in the glycine-rich domain (Fig. 3). A conserved pattern of glycine and tyrosine residues was observed in all sequences analyzed. Glycine and arginine residues seemed to be important for dimer formation, given that *Pg*-AMP1 forms a homodimer by interaction with these amino acids, as observed by mass spectrometry results (data not shown). A well conserved positively charged residue in the middle of the glycine-rich sequence was also observed, followed by a negatively charged amino acid residue (Fig. 3).

Similar structures have been found in other antimicrobial peptides such as distinctin, a protein isolated from *Phyllomedusa distincta*, which has shown activity against *E. coli*, *S. aureus*, *Pseudomonas aeruginosa* and *Enterococcus faecalis* [37]. The mode of action for this protein seems to be via interaction with the lipidic layer from the cell wall surface, leading to membrane permeabilization and cell death [37]. A three-dimensional structure of distinctin, obtained by NMR in water solution, is composed of two α-helices, although its' antimicrobial activity occurs by the formation of a homodimer. This shape seems to minimize the exposition of hydrophobic residues, stabilizing the structure and facilitating antimicrobial activity [37]. Furthermore, the defensin SPE10, isolated from *Pachyrrhizus erosus*, also showed similarities to *Pg*-AMP1 [44]. Its' activity against filamentous fungi *Aspergillus flavus*, *A. niger* and *Bipolaris maydis* seems to be mediated by the formation of a homodimer, where the interaction of each monomer occurs between hydrophobic residues at loop 2 (Ala^7, Phe^{10}, Pro^3, Phe^{15} and Phe^{39}) [44]. Moreover, it has been reported that dimeric conformations are important for antimicrobial activity in mammal defensins [40].

A phylogenetic tree of 20 glycine-rich proteins was generated. Proteins possessing the conserved domain (K/R)G(F/Y)(G/A)FVX(F/Y) were clustered in the tree, which included GRPs from *P. pattens*, *B. napus*, *S. tuberosum*, *M. sativa*, *R. obtusifolius* and *E. esula*, classified into the second group of GRPs. GRPs from group 2 contain, in addition to a glycine-rich region in the amino acid sequence, a single conserved domain, in contrast to the first group that possesses two conserved domains. *Pg*-AMP1 proteins from *Z. mays* and *O. sativa* do not show such conserved domains, and as such are classified into the third group of GRPs, where members present only the glycine-rich region in their sequence. *Pg*-AMP1 is closely positioned to GRPs from these monocotyledons, given that the guava peptide does not present any conserved domain. Correlations between species epithet and glycine-rich protein identity may indicate that this group of proteins evolved from a common ancestor, diverging between species during evolution. Although colicin E3 is a protein not directly related to glycine-rich proteins, it showed an unexpectedly high similarity and proximity to *Pg*-AMP1, confirming structural identity between both proteins (Fig. 4). Colicin E3 may pertain to glycine-rich protein group three, as it showed a sequence with a high content of glycine residues, without any conserved motifs or arginine contents. Given such similarity with colicin E3, the role of the peptide *Pg*-AMP1 in defense mechanisms may have been maintained during evolution from a common ancestor across both microorganisms and plants.

Through comparison of colicin E3 and *Pg*-AMP1 structures and analysis of arginine residue positions, it was possible to suggest that these residues at the external region of the dimer with the negatively charged part of bacterial membrane surface phospholipids. In order to confirm the mechanism of action of *Pg*-AMP1, nuclear magnetic resonance and molecular dynamics at the cell membrane surface are warranted as future studies. Continued discovery of novel glycine-rich proteins able to inhibit Gram-negative bacterial growth will likely provide new insights into antimicrobial mechanism of action in this protein group and also its *modus operandi*. Given that the antimicrobial peptide analyzed in this study is typically associated with storage, such plant proteins might present a secondary function in plant defense, acting against different pathogens. Numerous proteins, which have already been described in the literature, might therefore present such features, if screening for antimicrobial activities were to be conducted.

To date, there have been only few reports describing plant glycine-rich proteins with antimicrobial activities. Proteins from wheat were recently described with antimicrobial activities against filamentous fungi [10,49]. The glycine-rich protein curcumoschin from pumpkin also showed antifungal activities against *Botrytis cinerea*, *F. oxysporum* and *Mycosphaerella oxysporum* [16]. Glycine-rich proteins from oat (*Avena sativa*) and *Ginkgo biloba* have also been described, with the presence of chitin-binding domains suggested to be responsible for growth inhibition of a number of filamentous fungi [16,23]. Antiviral activity has also been described, with gene expression coding for a glycine-rich RNA-binding protein from *Nicotiana glutinosa* upregulated when the plant was infected with TMV [30]. As far as the authors are aware, prior to this current study there were no reports of proteins from this class with inhibitory activities against human bacterial pathogens.

Other members of plant storage proteins have also been described with antimicrobial activity, such as the 2S albumin groups, known typically as allergenic and storage proteins. For

example, 2S albumins from the *Brassica* family also showed inhibitory activities against *F. oxysporum*, *F. culmorum*, *Alternaria brasicola*, *B. cinerea*, *Pyrularia oryzae* and *Verticillium dahliae* [45]. Considering that glycine-rich proteins and 2S albumins contain members with antifungal and antibacterial activities, a future re-classification of these groups may be justified, as storage function might only be a secondary activity.

In conclusion, current control strategies against human pathogenic bacteria associated with urinary and gastrointestinal infections are becoming increasingly less effective as a result of selection for antibiotic resistance in pathogen populations. On the basis of our results, *Pg*-AMP1 appears to be the first glycine-rich plant protein with activity against human pathogenic Gram-negative bacteria. Discovery of this novel peptide able to inhibit Gram-negative eubacteria will facilitate, in the near future, in depth analysis of the mechanism of action of this protein group. Application in pathogen inhibition will likely include genetic engineering-derived antibiotic production or development of novel disinfectants for hospital environments.

Acknowledgments

The authors are thankful for the financial support by UCB, CAPES, CNPq and FAPEMIG.

REFERENCES

[1] Bergeron MG, Bernier M. Bactericidal activity of cefpirome (HR 810) against 513 Gram-negative bacteria isolated from blood of septicemic patients. Infection 1994;22:299–305.
[2] Bevins CL. Paneth cell defensins: key effectors molecules of innate immunity. Biochem Soc Trans 2006;34:263–6.
[3] Bulet P, Cocianich S, Dimarcq JL, Lambert J, Hoffmann D, Hofmann JA. Insect immunity, isolation from a coleopteran insect of a novel inducible antibacterial peptide and of a new member of insect defensin family. J Biol Chem 1991;236:24425–520.
[4] Campos F, Zamudio F, Covarrubias AA. Two different late embryogenesis abundant proteins from *Arabidopsis thaliana* contain specific domains that inhibit *Escherichia coli* growth. Biochem Biophys Res Commun 2006;342:406–13.
[5] Condit CM, Meagher RB. Expression of a gene encoding a glycine-rich protein in petunia. Mol Cell Biol 1987;7:4273–9.
[6] Cui L, Iwamoto A, Lian J-O, Neoh H-M, Maruyama T, Horikawa Y, et al. Novel mechanism of antibiotic resistance originating in vancomycin-intermediate *Staphylococcus aureus*. Antimicrob Agents Chemother 2006;41:428–38.
[7] DeLano WL. Unraveling host in binding interfaces: progress and challenges. Curr Opin Struct Biol 2002;12:14–20.
[8] de Oliveira DE, Seurinck J, Inze D, Van Montaque M, Botteman J. Differential expression of five Arabidopsis genes encoding glycine-rich proteins. Plant Cell 1990;437–536.
[9] Di Gaspero G, Cipriani AEG. Nucleotide binding site/leucine-rich repeats, Pto-like and receptor-like kinases related to disease resistance in grapevine. Mol Gen Genomics 2003;269:612–23.
[10] Egorov TA, Odintsova TI, Pukhalsky PA, Grishin EV. Diversity of wheat anti-microbial peptides. Peptides 2005;26:2064–73.
[11] Falk A, Feys BJ, Frost LN, Jones JDG, Daniels MJ, Parker JE. *EDS1*, an essential component of R gene-mediated disease resistance in *Arabidopsis* has homology to eukaryotic lipases. Plant Biol 1999;96:3292–7.
[12] Franco OL, Rigden DJ, Melo FR, Grossi-de-Sa MF. Plant α-amylase inhibitors and their interaction with insect α-amylase. Structure, function and potential for crop protection. Eur J Biochem 2002;269:397–412.
[13] Ginalski K, Elofsson A, Fischer D, Rychlewski L. 3D-jury: a simple approach to improve protein structure predictions. Bioinformatics 2003;19:105–18.
[14] Guex N, Peitsch MC. SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. Eletrophoresis 1997;18:2714–23.
[15] Hall TA. BioEdit: a user-friendly sequence alignment editor and analysis program for Windows 95/98/NY. Nucl Acid Symp Ser 2005;41:95–8.
[16] Huang X, Xie W, Gong Z. Characteristics and antifungal activity of a chitin binding protein from *Ginkgo biloba*. FEBS Lett 2000;478:123–6.
[17] Joosten MHAJ, De Wit PJGM. The tomato–*Cladosporium fluvum* interaction: a versatile experimental system to study plant–pathogen interactions. Annu Rev Phytopathol 1999;50:335–67.
[18] Keller B, Sauer N, Lamb CJ. Glycine-rich cell wall proteins in bean: gene structure and association of the protein with the vascular system. EMBO J 1988;7:3625–33.
[19] Keller B, Templeton MD, Lamb CJ. Specific localization of a plant cell wall glycine-rich protein in protoxylem cells of the vascular system. Proc Natl Acad Sci USA 1989;86:1529–33.
[20] Kwak KJ, Kim YO, Kang H. Characterization of transgenic Arabidopsis plants overexpressing GR-RBP4 under high salinity, dehydration, or cold stress. J Exp Bot 2005;56:3007–16.
[21] Laberge S, Castonguay Y, Vezina LP. New cold- and drought-regulated gene from *Medicago sativa*. Plant Physiol 1993;101:1411–2.
[22] Laemmli UK. Cleavage of structural proteins using assembly of the head of the bacteriophage. Nature (London) 1970;227:680–5.
[23] Lii SS, Claeson P. Cys/Gly-rich proteins with a putative single chitin-binding domain from oat (*Avena sativa*) seeds. Phytochemistry 2003;63:249–55.
[24] Lindahl E, Hess B, Spoel DVD. A package for molecular simulation and trajectory analysis. J Mol Model 2001;7:306–17.
[25] Mantovani P, van der Linden G, Maccaferri M, Sanguineti MC, Tubersoa R. Nucleotide-binding site (NBS) profiling of genetic diversity in durum wheat. Genome 2006;49:1473–80.
[26] Marti-Renom MA, Stuart A, Fiser A, Sánchez R, Melo F, Sali A. Comparative protein structure modeling of genes and genomes. Annu Rev Biophys Biomol Struct 2000;29:291–325.
[27] Mendes C, Hsiung A, Kiffer C, Oplustil C, Sinto S, Mimica I, et al. Evaluation of the *in vitro* activity of nine antimicrobials against bacterial strains isolated from patients in intensive care units in Brazil: MYSTIC antimicrobial surveillance program. Braz J Infect Dis 1997;5:236–44.
[28] Moraes CB, Lago CL. Espectrometria de massas com ionização por "Electrospray" aplicada ao de espécies inorgânicas e orgânicas. Química Nova 2003;26:556–63.
[29] Mousavi A, Hotta Y. Glycine-rich proteins: a class of novel proteins. Appl Biochem Biotechnol 2005;120:169–74.
[30] Naqvi SM, Park KS, Yi SY, Lee HW, Bok SH, Choi D. A glycine-rich RNA-binding protein gene is differentially expressed during acute hypersensitive response following tobacco mosaic virus infection in tobacco. Plant Mol Biol 1998;37:571–6.

[31] Naqvi SM, Harper A, Carter C, Ren G, Guirgis A, York WS, et al. a potent endoglucanase inhibitor secreted into the nectar of ornamental tobacco plants. Isolation, cloning, and characterization. Plant Physiol 2005;139:1389–400.
[32] Nomata T, Kabeya Y, Sato NV. Cloning and characterization of glycine-rich RNA-binding protein cDNAs in the moss *Physcomitrella patens*. Plant Cell Physiol 2004;45:48–56.
[33] Oard S, Rush MC, Oard JH. Characterization of antimicrobial peptides against a US strain of the rice pathogen *Rhizoctonia solani*. J Appl Microbiol 2004;97: 160–80.
[34] Page RDM. TREEVIEW: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 1996;12:357–8.
[35] Park CB, Kim MS, Kim SC. A novel antimicrobial peptide from *Bufo bufo* gargartzans. Biochem Biophys Res Commun 1996;218:408–13.
[36] Pelegrini PB, Franco OL. Plant gamma-thionins: novel insights on the mechanism of action of a multi-functional class of defense proteins. Int J Biochem Cell Biol 2005;37:2239–53.
[37] Raimondo D, Andreotti G, Saint N, Amodeo P, Renzone G, Sanseverino M, et al. A folding-dependent mechanism of antimicrobial peptide resistant to degradation unveiled by solution structure of distinctin. Proc Natl Acad Sci USA 2005;102:6309–14.
[38] Ringli C, Keller B, Ryser U. Glycine-rich protein as structural components of plant cell walls. Cell Mol Life Sci 2001;58:1430–41.
[39] Ryser U, Keller B. Ultrastructural localization of a bean glycine-rich protein in unlignified primary walls of protoxylem cells. Plant Cell 1992;4:773–83.
[40] Schibli DJ, Hunter HN, Aseyev V, Dtarner TD, Wiencek JM, McCray PB, et al. The solution structures of the human beta-defensins lead to a better understanding of the potent bactericidal activity of HBD3 against *Staphylococcus aureus*. J Biol Chem 2002;277:8279–89.
[41] Selitrennikoff CP. Antifungal proteins. Appl Environ Microbiol 2001;67:2883–94.
[42] Shinozuka H, Hisano EH, Yoneyama ES, Shimamoto Y, Jones EES, Forster EJW, et al. Gene expression and genetic mapping analyses of a perennial ryegrass glycine-rich RNA-binding protein gene suggest a role in cold adaptation. Mol Genet Genomics 2006;275:399–408.
[43] Soelaiman S, Jakes K, Wu N, Li C, Shoham M. Crystal structure of colicin E3: implications for cell entry and ribosome inactivation. Mol Cell 2001;8:1053–62.
[44] Song X, Wang J, Wu F, Li X, Teng M, Gong W. cDNA cloning, functional expression and antifungal activities of a dimeric plant defensin SPE10 from *Pachyrrhizus erosus* seeds. Plant Mol Biol 2005;57:13–20.
[45] Terras FRG, Torrekens S, Leuven FV, Osborn RW, Vanderleyden J, Cammue BPA, et al. A new family of basic cysteine-rich plant antifungal proteins from Brassicaceae species. FEBS Lett 1993;233–40.
[46] Thompson JD, Higgins DG, Gibson TJ. CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–80.
[47] Van der Hoorn AL, Kruijt M, Roth R, Brandwagt BF, Joosten MHAJ, De Wit PJGM. Intragenic recombination generated two distinct Cf genes that mediate AVR9 recognition in the natural population of *Lycopersicon pimpinellifolium*. Science 2001;98:10493–8.
[48] Wang Y, Sun J, Chitnis P. Proteomic study of the peripheral proteins from thylakoid membranes of the cyanobacterium *Synechocytes* sp.. Electrophoresis 2000;21:1746–54.
[49] Wang HX, Nq TB. Isolation of cucurmoschin, a novel antifungal peptide abundant in arginine, glutamate and glycine residues from black pumpkin seeds. Peptides 2003;24:969–72.
[50] Ye XY, Ng TB. Mungin, a novel cyclophilin-like antifungal protein from the mung bean. Biochem Ciophys Res Commun 1991;273:1111–5.
[51] Xiao SY, Elwood S, Calis O, Patrick E, Li TX, Coleman M, et al. Broad-spectrum mildew resistance in *Arabidopsis thaliana* mediated by RPW8. Science 2001;15:118–20.
[52] Xiao S, Brown S, Patrick B, Brearley C, Turner JG. Enhanced transcription of the Arabidopsis disease resistance genes *RPW8*. 1 and RPW8.2 via a salicylic acid-dependent amplification circuit is required for hypersensitive cell death. Plant Cell 2003;15:33–45.
[53] Zhong YS, Mita K, Shimada T, Kawasaki H. Glycine-rich protein genes, which encode a major component of the cuticle, have different developmental profiles from other cuticle protein genes in *Bombyx mori*. Insect Biochem Mol Biol 2006;36:99–110.