Amino Acids
https://doi.org/10.1007/s00726-018-2598-3

ORIGINAL ARTICLE

Antimicrobial and structural insights of a new snakin-like peptide isolated from *Peltophorum dubium* (Fabaceae)

Susana Rodríguez-Decuadro[1] · Mariana Barraco-Vega[2] · Pablo D. Dans[3,4] · Valesca Pandolfi[5] · Ana Maria Benko-Iseppon[5] · Gianna Cecchetto[2,6]

Received: 21 March 2018 / Accepted: 31 May 2018

Abstract

Snakins are antimicrobial peptides (AMPs) found, so far, exclusively in plants, and known to be important in the defense against a wide range of pathogens. Like other plant AMPs, they contain several positively charged amino acids, and an even number of cysteine residues forming disulfide bridges which are considered important for their usual function. Despite its importance, studies on snakin tertiary structure and mode of action are still scarce. In this study, a new snakin-like gene was isolated from the native plant *Peltophorum dubium*, and its expression was verified in seedlings and adult leaves. The deduced peptide (PdSN1) shows 84% sequence identity with potato snakin-1 mature peptide, with the 12 cysteines characteristic from this peptide family at the GASA domain. The mature PdSN1 coding sequence was successfully expressed in *Escherichia coli*. The purified recombinant peptide inhibits the growth of important plant and human pathogens, like the economically relevant potato pathogen *Streptomyces scabies* and the opportunistic fungi *Candida albicans* and *Aspergillus niger*. Finally, homology and ab initio modeling techniques coupled to extensive molecular dynamics simulations were used to gain insight on the 3D structure of PdSN1, which exhibited a helix–turn–helix motif conserved in both native and recombinant peptides. We found this motif to be strongly coded in the sequence of PdSN1, as it is stable under different patterns of disulfide bonds connectivity, and even when the 12 cysteines are considered in their reduced form, explaining the previous experimental evidences.

Keywords Pathogenesis-related peptide · Heterologous expression · *Escherichia coli* · Ab initio and homology modeling · Molecular dynamics simulations

Introduction

Plants and animals own a diverse group of small proteins with antimicrobial activity, so-called antimicrobial peptides (AMPs). These peptides are evolutionarily ancient, being important in the defense against a wide range of

Handling Editor: M. S. Palma.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00726-018-2598-3) contains supplementary material, which is available to authorized users.

✉ Gianna Cecchetto
gianna.cecchetto@gmail.com

[1] Departamento de Biología Vegetal, Facultad de Agronomía, Universidad de la República, Garzón 780, 12900 Montevideo, Uruguay

[2] Departamento de Biociencias, Facultad de Química, Universidad de la República, General Flores 2124, 11800 Montevideo, Uruguay

[3] Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Baldiri Reixac 10-12, 08028 Barcelona, Spain

[4] Joint BSC-IRB Research Program in Computational Biology, Baldiri Reixac 10-12, 08028 Barcelona, Spain

[5] Universidade Federal de Pernambuco, Centro de Biociências, Av. Prof. Moraes Rego, 1235, Recife, PE CEP 50.670-420, Brazil

[6] Instituto de Química Biológica, Facultad de Ciencias, Facultad de Química, Universidad de la República, General Flores 2124, 11800 Montevideo, Uruguay

Published online: 08 June 2018

pathogens, including bacteria, fungi, viruses, and protozoa (Zasloff 2002). Known AMPs differ in size, composition, and molecular structure, although they share some features like a small size (< 10 kDa), an amphipathic structure, and a net positive charge at physiological pH (Yeaman and Yount 2003; Brogden 2005). This high structural diversity allows AMPs to act against a wide range of microbial agents in diverse physiological environments (Tossi and Sandri 2002). In plants, these peptides often contain a high number of cysteine residues that stabilize protein structure through disulfide bond formation (Lay and Anderson 2005). Moreover, disulfide bonds may act catalytically and can be reversibly reduced and oxidized, playing a role in redox regulation (Nahirñak et al. 2012b). The main groups of plant AMPs are defensins, thionins, snakins, lipid transfer proteins, and cyclotides (Benko-Iseppon et al. 2010).

Snakins constitute a family of AMPs described from potato Snakin-1 (StSN1, Segura et al. 1999) and Snakin-2 (StSN2, Berrocal-Lobo et al. 2002), the first peptides isolated of this family. These peptides belong to a group of proteins encoded by Snakin/GASA genes, that are characterized by having a GASA (Gibberellic Acid Stimulated in Arabidopsis) domain of approximately 60 amino acids, with 12 highly conserved cysteine residues that may be involved in the formation of up to six disulfide bonds (Nahirñak et al. 2012b). Mass spectrometry was first applied to decipher the disulfide connectivity, with the hypothesis that they may be required for the generation or stabilization of a particular 3D motif responsible for the function. Accordingly, a partial connectivity pattern with three disulfide bonds was identified for StSN2 (Harris et al. 2014). Nearly, at the same time, Porto and Franco (2013) predicted the StSN1 three-dimensional structure through an ab initio and comparative modeling in combination with a predictor of disulfide bridges. Although the arrangement of cysteines along the sequence is conserved in the snakin family, the StSN1 model only shared one disulfide bond with the mass spectrometry data obtained for StSN2. The disagreement between experiments and modeling became evident with the X-ray determination of StSN1 (Yeung et al. 2016) that fully supported the mass spectrometry data, reporting a structure with a helix–turn–helix (HTH) motif stabilized by six disulfide bonds between specific cysteine residues. Surprisingly, the StSN1 model of Porto and Franco that only shared two disulfide bonds with the X-ray structure, also presented an HTH motif, suggesting that a precise connection between specific cysteine pairs may not be necessary for the peptide to fold to the native and functional state. Furthermore, some evidence exist pointing out that the reduced form of the StSN1 and StSN2 peptides (where all cysteines are free, and there are no disulfide bonds) were equally active against pathogens as the native forms (Harris et al. 2014).

The role assignment for snakins has been based on expression profiling analysis (Segura et al. 1999; Berrocal-Lobo et al. 2002; Meiyalaghan et al. 2014; Nahirñak et al. 2016; Herbel et al. 2017) and transgenic expression of these genes (Almasia et al. 2008; Balaji and Smart 2012; Rong et al. 2013; Mohan et al. 2014). Overexpression of snakin-1 or 2 in transgenic potato and tomato plants increased resistance to pathogens (Almasia et al. 2008; Balaji and Smart 2012; Mohan et al. 2014). In vitro, StSN1 and StSN2 displayed a broad spectrum of antimicrobial activity (Segura et al. 1999; Berrocal-Lobo et al. 2002). On the other hand, silencing of snakins-2 homologous genes in *Nicotiana benthamiana* increases susceptibility to *Clavibacter michiganensis* ssp. *michiganensis* (Balaji et al. 2011). In addition, potato snakin-1 gene silencing seems to influence cell wall composition, cell division, and leaf primary metabolism in potato plants, suggesting that StSN1 may also be involved in several cellular processes (Nahirñak et al. 2012a).

The biological properties of snakins make them attractive biotechnological targets, especially for the development of novel disease control agents (Oliveira-Lima et al. 2017). Heterologous expression is the most widely used method for medium and large-scale production. Several systems (producing organism/vector) have been developed to achieve a cost-effective and large-scale production of several proteins, in hosts such as bacteria, yeasts, fungi, and plants (Thevissen et al. 2007; Padovan et al. 2010; Silva et al. 2011). Specifically, recombinant snakins have been successfully produced in *E. coli*, *Pichia pastoris,* and more recently in baculovirus/insect cells (Almasia et al. 2017), displaying the expected antibacterial and antifungal activity (Kovalskaya and Hammond 2009; Mao et al. 2011; Herbel et al. 2015; Kuddus et al. 2016).

In the last decade, the bioinformatics tools have identified AMP-coding genes in several plant species. Due to the increasing availability of their omics, cultivated plants and model species have been the main target. Native species that stand out for their importance in folk medicine and exhibit considerable plant biodiversity remain underexplored, mainly because their molecular data are generally unavailable (Pestana-Calsa et al. 2010). In this work, we report the first snakin-like gene *PdSN1* isolated from *Peltophorum dubium*, a non-cultivated South American tree with medicinal properties. Our expression analysis showed that *PdSN1* is strongly expressed during seedling development and could be a constitutive element of the defense mechanisms of the storage organs. The snakin peptide was successfully produced through recombinant expression in *E. coli*, purified and characterized for its activity against a set of microorganisms. In addition, we report the first extensive analysis on the importance of the disulfide bridges on the stability of key structural marks present in the snakin family, like the HTH motif found in both the native and recombinant PdSN1

peptides. Supported by our structural analyses, we found that disulfide bridges might not be essential for the activity of the snakin family. In addition, we hypothesize that these peptides possibly could, in addition to the previously proposed lipid-membrane targeting mechanism of action, alter the microbial gene expression by binding to DNA/RNA.

Materials and methods

Biological material

Peltophorum dubium seeds and leaves were obtained from Facultad de Agronomía Garden (Montevideo, Uruguay). Seeds were immersed in concentrate grade H_2SO_4 during 15 min for scarification. Surface sterilized seeds were germinated on Whatman paper soaked with distilled water in Petri dishes at 28 °C. Leaves and 5-day-old seedlings (ca. 6-cm long) were frozen in liquid nitrogen and stored at − 70 °C until further usage.

Bacteria and fungi strains used for bioactivity assays were obtained from the Laboratory of Microbiología Facultad de Química collection. Fungal cultures (*Candida albicans* CCMG13, *Aspergillus niger* CCMG17, *Botrytis cinerea* CCMG14 g, *Alternaria alternata* CBS916.96, and *Penicillium expansum* CCMG14s) and *Streptomyces scabies* DSM41658 were grown on potato dextrose agar (PDA) at 28 °C. *Staphylococcus aureus* ATCC6538P and *Escherichia coli* CCMG50 were grown on Tryptone soya agar (TSA) at 37 °C. *C. michiganensis* ssp. *michiganensis* MAI1008 and *Xanthomonas versicatoria* MAI2020 were grown on Nutrient Agar at 28 °C. *E. coli* TOP10 (Invitrogen, Carlsbad, USA), Shuffle (New England Biolabs, Ipswich, USA) and Rosetta-gami(DE3)pLysS (Novagen, Madison, USA) were used as cloning and expression host, respectively, being grown on Luria–Bertani Agar.

Isolation of DNA and RNA from *P. dubium*

Genomic DNA was extracted from leaves using the standard cetyl-trimethylammonium bromide (CTAB) method (Doyle 1991). Total RNA was extracted using Qiagen RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) and treated with RNase-free DNase (Invitrogen, Carlsbad, USA) to eliminate any residual DNA. Genomic and plasmid DNA were visualized under UV illumination after electrophoresis on agarose gels in 1× TBE buffer, stained with GoodView (Ecoli s.r.o., Bratislava, Slovak Republic).

PdSN1 gene cloning

PdSN1 gene was PCR amplified from ATG to stop codon, from *P. dubium* genomic DNA using degenerate primers SN_dgF 5′-ATGAAGCCAGCATTTKCARCTMT-3′ and SN_dgR 5′-TTAAGGGCATTTWGSCTTKCC-3′. Primers design was based on the alignment from plant snakin encoding sequences that were identified in the UniProt database (http://www.uniprot.org/) and the GenBank database of the National Center for Biotechnological Information (NCBI) (https://www.ncbi.nlm.nih.gov/genbank/). The PCR reaction was performed in a 20-µl reaction containing: 1× buffer with 2-mM $MgCl_2$, 0.2-mM dNTPs, 2 µM of each primer, 50-ng template DNA, and 0.5 U Taq DNA polymerase (Invitrogen, Carlsbad, USA). The PCR program was as follows: 94 °C for 3 min, followed by 35 cycles of 94 °C for 30 s, 54 °C for 40 s and 72 °C for 40 s. PCR products were purified according to Richero et al. (2013) and cloned into a pGEM-T easy vector (Promega Corporation, Madison, USA). *PdSN1* gene sequence was confirmed by sequencing three clones at Macrogen Inc. (Seoul, Korea) and analyzed using BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

PdSN1-coding sequence (termed *PdSN1*c) was verified by the synthesis of 5′ and 3′ cDNA regions from total RNA extracted from *P. dubium* seedlings. In the 5′-RACE experiment (Rapid Amplification cDNA Ends), the cDNA was obtained using the 5′ RACE kit (Invitrogen) and the specific oligonucleotides PdSNc_R1 5′-AGTTCTTGAGGTCTCTGT AGCA-3′ and PdSNc_R2 5′-CTCTGTAGCAAGGGCACT CGT-3′ (Fig. 1). The 3′ region was obtained by 3′-RACE according to Frohman et al. (1988) using the specific oligonucleotides PdSNc_F1 5′-GTGCCTCTTTCGTTGAGG TC-3′ and PdSNc_F2 5′-GCCTTGTACTCAGTGCCTCTT-3′. PCR products were cloned in the pGEM-T easy vector (Promega Corporation, Madison, USA) and sequenced using the Sanger method. The resulting plasmid (pG-*PdSN1*c) contained the complete coding sequence of the mature peptide. The cDNA sequence was deposited to GenBank with accession number MG229644.

Bioinformatics analysis of the PdSN1 sequence

The deduced amino acid sequences of PdSN1 were created in BioEdit (Hall 1999) and analyzed with the Expasy-ProtParam tool (http://web.expasy.org/protparam/) to obtain the different peptide parameters. The peptide structure was evaluated for the presence of a signal peptide sequence with SignalP (http://www.cbs.dtu.dk/services/SignalP). The deduced amino acid sequences encoding for the mature *P. dubium* peptide were aligned against a set of mature plant proteins with GASA domain. All sequences used were taken from the UniProt database (http://www.uniprot.org/) (Table S1). Alignment with the newly isolated snakin was performed in ClustalW from BioEdit program. An unrooted tree was generated using the Neighbor-Joining method of the MEGA package version 5 (Tamura et al. 2011), with 5000 bootstrap replicates.

Fig. 1 Nucleotide genomic sequence of *PdSN1* and amino acidic sequence of PdSN1. The lower case letters indicate the 5′ UTR, intron, and 3′ UTR regions. Amino acid sequence of PdSN1 mature peptide indicated by gray-shaded is preceded by a signal peptide sequence predicted using the SignalP (http://www.cbs.dtu.dk/services/SignalP). Oligonucleotide sequences are indicated by horizontal arrows

Gene expression analysis

Quantitative reverse transcription-PCR (RT-qPCR) was employed to determine the relative expression levels of *PdSN1* gene in *P. dubium*'s adult leaves and seedlings (inoculated with *A. niger* and mock-not inoculated control). Five-day-old seedlings were inoculated at five points with 20 µl per point of a spore suspension of *A. niger* (1×10^6 conidia ml^{-1}). Seedling samples were collected 48 h after inoculation and frozen in liquid nitrogen. Mock inoculations were done with Tween 10^{-4}. RT synthesis was performed using MMLV reverse transcriptase (Invitrogen, Carlsbad, USA) and random primers (QIAGEN, Hilden, Germany). Specific primers (PdSN1_qF1 ACTCAGTGC CTCTTTCGTTG and PdSN1_R1 CATTCCTGCTTTCGA GCATCT) were used to amplify the cDNA fragments. PCR reactions were carried out using SensiMixPlus SYBR PCR Master Mix (2×) (Quantace, London, UK) in a Corbett Rotor Gene™ 6000 as follows: 15-min pre-denaturing at 95 °C, followed by 30 cycles of 94 °C for 15 s, 59 °C for 1 min. Three biological and two technical replicates were performed for each sample. The specificity of each reaction was confirmed by inspection of the melting curve profiles.

To choose the reference genes (RGs), the expression of actin (*PdACT*), elongation factor 1-α (*PdEF1*), F-box/kelch-repeat protein (*PdF-box*), phosphoenolpyruvate carboxylase-related kinase 1 (*PdPEPKR*), protein phosphatase 2A (*PdPP2A*), and tubulin A (*PdTUA*) were evaluated for all conditions. The expression stability of RGs was assessed using geNorm algorithm (Vandesompele et al. 2002). Previously, the candidate genes were partially cloned by 5′- and 3′-RACE using reported primers from *Caragana korshinskii* (Zhu et al. 2013; Table S2). qPCR specific primers of RGs of *P. dubium* were then located in the first cloned region of each one (Table S3).

Partial cDNA sequences of RGs were deposited in GenBank with accession numbers MG397035, MG397036, MG397037, MG397038, MG397039, and MG397040.

Construction of the expression plasmid

The pET102/D-TOPO vector (Invitrogen, Carlsbad, USA) was used for expression of *PdSN1* gene. This vector allows the expression of PdSN1 mature peptide fused to a thioredoxin protein (N-ter) and a His-tag (C-ter) to facilitate solubility and downstream purification using affinity chromatography. The mature coding sequence of *PdSN1* was PCR amplified from pG-*PdSN1*c using the following primers: PdSN1_pEF 5′-CACCGGTTCTGAGTTCTGT GACTCCAAGTGCGCG-3′ (TOPO cloning nucleotides CACC are underlined) and PdSN1_pER 5′-CCTTCCCTC GAT GGGGCATTTGGGCTTGCC-3′. Reverse primer has 12 nucleotides which code for the factor Xa recognition site. Blunt-end PCR product was obtained using a Phusion High-Fidelity DNA Polymerase (Thermo-Scientific, Waltham, USA). The PCR conditions were 95 °C for 1 min followed by 30 cycles (95 °C for 30 s, 60 °C for 30 s and 72 °C for 1 min) and finally 72 °C for a 5-min extension. The TOPO cloning reaction and the transformation of *E. coli* were performed according to the manufacturer's instructions. The resulting plasmid pE-PdSN1 was confirmed by DNA sequencing. In parallel, a pET102/D without PdSN1 sequence (pE-Trx*) was obtained.

Expression and purification of recombinant PdSN1

The *PdSN1* carrying plasmid pE-PdSN1 was introduced into *E. coli* Rosetta-gami (DE3)pLysS and Shuffle strains. A transformation with pE-Trx* was performed as a control. Cells were grown at 37 °C and 150 rpm in the presence of 100-µg ml^{-1} ampicillin (Amresco, Solon, USA) until the culture density reached an OD600 of 0.5–0.7. Induction was carried on with 1-mM IPTG (Sigma-Aldrich, Saint Louis, USA) varying time (2, 4, and 24 h) and temperature (20, 28, and 30 °C). After incubation, the bacterial cells were harvested by centrifugation at 4000×*g* for 15 min at 4 °C, washed twice in 50-mM potassium phosphate buffer, pH 7.8, and stored at − 70 °C until protein extraction.

The cell pellet was resuspended in 2 ml of ice-cold lysis buffer (50-mM potassium phosphate buffer, pH 7.5, 500-mM NaCl, 10% glycerol, 10-mM imidazole). The cells were lysed by sonication 4× pulse at 4 °C (30% amplitude; 6 cycles of 6-s pulse-on, 9-s pulse-off; 15 s) (Cole-Parmer, Vernon Hills, USA). The lysate was centrifuged 15 min at 4 °C and 10,000 rpm, and the supernatant was filtered through a 0.22-µm filter (Millipore). The recombinant protein was purified using a HisTrap-FF (GE Healthcare, Little Chalfont, UK) Ni-affinity chromatography, according to the supplier's recommendations. Unbound proteins were eluted by adding 20- and 30-mM imidazole to the buffer, while 300-mM imidazole was used to elute the bound proteins. To remove the thioredoxin (Trx) fusion fragment, the immobilized Trx-PdSN1protein was subjected to EkMax Enterokinase (Invitrogen) digestion for 15 h at room temperature. Eluted fractions were analyzed by 15% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) using Tris–Tricine buffer (Schägger and von Jagow 1987). Fractions containing recombinant peptide were pooled, dialyzed against distilled water, and lyophilized. Subsequently, the peptide was resuspended in water and the protein concentration was measured by Bradford assay (Bradford 1976) using bovine serum albumin as a standard.

Identification of recombinant PdSN1 peptide

The peptide sequence was confirmed by peptide mass fingerprinting and subsequent fragmentation of selected peptides (MS/MS). Disulfide bonds were reduced by incubation with DTT (100 mM) for 1 h at 56 °C and Cys were alkylated with iodoacetamide (300 mM). Tryptic digestion was performed overnight at 37 °C using sequencing grade trypsin (Promega). Peptide mixtures were mixed with matrix solution (-cyano-4-hydroxycinnamic acid in 60% acetonitrile, 0.1% TFA) directly onto the sample plate. Spectra were acquired in positive reflector mode and were externally calibrated using a standard mixture of peptides (Applied Biosystems). Peptide sequences were assigned using an in-house Mascot v.2.3 version (Matrix Science) for searching a local database that includes the sequence of PdSN1. The following parameters were used with variable modifications: oxidation (M), carbamidomethyl (C), ammonia loss (N-term C); mass values: monoisotopic; peptide mass tolerance: ± 0.05 Da; and fragment mass tolerance: ± 0.45 Da. Selected *m*/*z* values were further fragmented, and the assignment to PdSN1 amino acid sequence was validated by manual inspection of MS/MS spectra.

Antimicrobial activity assay

Antimicrobial activity was estimated microspectrophotometrically (Broekaert et al. 1990), determining the IC50 value (peptide concentration at which 50% inhibitions was reached). Fungal spore suspension (1×10^4 spores ml^{-1}) or bacterial suspension (1×10^5 bacteria ml^{-1}) was cultivated in 100-µl reaction, containing Potato Dextrose Broth (for fungi and *S. scabies*), Nutrient Broth or Tryptone Soya Broth (for bacteria), and serial dilutions of recombinant PdSN1. The maximum peptide concentration used in all assays was 1.8 µM. Control wells contained no peptide. The optical density of fungal or bacterial suspension was measured in a 96-well microtiter plate under 595 nm, after 48–72 h of incubation (or 24 h for *E. coli* and *S. aureus*) in the dark. Antimicrobial activity of PdSN1 was expressed as a percentage of growth inhibition. Serial dilutions of Gentamicin were used as positive control for bacteria and propiconazole for fungi. The maximum concentrations used were 30 and 100 µM, respectively. All experiments were performed in three replicates.

3D structure prediction and molecular dynamics simulations

The prediction of the 3D structure of the PdSN1 peptide was achieved using three different approaches: (1) ab initio or de novo 3D modeling; (2) homology modeling; and (3) multi-microsecond long, unbiased molecular dynamics. In all cases, the histidine residues were considered as protonated in ε.

Ab initio predictions

For the ab initio 3D determination, we followed the work of Porto and Franco (2013), who used a combination of tools, including the QUARK ab initio molecular modeling server (Xu and Zhang 2012), the DIANNA disulfide bond predictor (Ferrè and Clote 2006), the MODELLER 9.14 modeling tool (Webb and Sali 2014), and short unbiased molecular dynamics simulations. The mature form of the native *PdSN1* sequence was submitted to the QUARK server, which delivered an initial guess of the 3D structure. Then,

the disulfide bond pattern was determined using five different predictors: DIANNA, DISULFIND (Ceroni et al. 2006), CYSCON (Yang et al. 2015), DIPRO (Cheng et al. 2006), and DISLOCATE (Savojardo et al. 2011). These predictors gave different connectivity patterns between the six pairs of cysteines; hence, an individual model was built according to each prediction (see Table S4). Each 3D structure was then subjected to microsecond long unbiased MD simulation with options as described below.

Homology modeling

Using the recently determined X-ray structure of the potato snakin-1 peptide (PDB id: 5E5Q) as a template, we build a 3D model by homology using MODELLER (Webb and Sali 2014). This peptide shares 84% similarity in sequence with the mature form of the native *PdSN1* (63 residues). The connectivity pattern of disulfide bonds was taken from 5E5Q, and the final structure was subject to unbiased MD simulations (labeled $PdSN1_{5E5Q}$). A model of the full recombinant peptide (104 residues) was also built, where the GASA domain was obtained from 5E5Q, and the remaining amino acids were modeled using the QUARK server ($PdSN1_{5E5Q\text{-}ALL}$).

Long unbiased MD simulations

To assess the importance of the disulfide bonds on the stability and conformation of the 3D and secondary structure, we took the model generated by homology modeling and run a multi-microsecond long MD simulation considering all the cysteines in their reduced form (no disulfide bonds, $PdSN1_{5E5Q\text{-}RED}$).

MD simulation protocol

All models were minimized in vacuo, neutralized with 8 Cl^-, solvated (with explicit water and 0.15 M of K^+Cl^-), and minimized in solution with positional restraints on the peptide using our well-established multi-step protocol (Pérez et al. 2007; Dans et al. 2016). To produce the final models, the minimized structures were thermalized to 298 °C at NVT, and then simulated during at least 1 µs using molecular dynamics simulations at NPT ($P = 1$ atm). The first 100 ns of simulations were considered as an equilibration step and were discarded for further analysis. For representing the snakin peptide, we used the state-of-the-art ff14SB force field (Maier et al. 2015), surrounded by a truncated octahedral box of ~ 5000 TIP3P water molecules (Jorgensen et al. 1983), and Dang parameters for ions (Smith and Dang 1994). Ions were initially placed randomly, at a minimum distance of 5 Å from the solute and 3.5 Å from one another. All systems were simulated using the Berendsen algorithm (Berendsen et al. 1984) to control the temperature and the pressure, with a coupling constant of 5 ps. Center of mass motion was removed every 10 ps to limit the build-up of the translational kinetic energy of the solute. SHAKE (Ryckaert et al. 1977) was used to keep all bonds involving hydrogen at their equilibrium values, allowing the use a 2-fs step for the integration of Newton equations of motion. Long-range electrostatic interactions were accounted for using the particle mesh Ewald method (Darden et al. 1993) with standard defaults and a real-space cutoff of 9 Å. All simulations were carried out using the PMEMD CUDA code module (Salomon-Ferrer et al. 2013) of AMBER 16 (Case et al. 2017) and analyzed with CPPTRAJ (Roe and Cheatham III 2013). Structures and MD trajectories were visually analyzed using VMD 1.9 (Humphrey et al. 1996).

Results

Isolation and sequence analysis of *PdSN1*

PdSN1 genomic sequence isolated by PCR presents 424 nt, from de ATG start codon to stop codon, with one intron of 157 nt (Fig. 1). The coding *PdSN1* sequence, isolated from seedlings cDNA, allowed for the confirmation of the sequence and localization of the intron. Analysis of the deduced amino acid sequence (88 residues) showed a signal peptide sequence of 25 amino acids, followed by a mature peptide with 63 residues. The mature peptide had a calculated mass of 6978.1 Da and a basic isoelectric point of 8.99.

PdSN1 BLAST search showed high similarity to Snakin/GASA proteins, especially with snakin type 1 (i.e., 76% sequence identity to StSN1 from potato and 74% to MsSN1 from *Medicago sativa*). It has a GASA domain (84% sequence identity to StSN1), characteristic from this protein family, with 12 cysteines in conserved positions ($XCX_3CX_3CX_8CX_3CX_2CCX_2CXCX_{11}CXCX_{12}CX$). Alignment analysis of PdSN1 mature sequence and plant mature proteins with GASA domain revealed that the newly isolated gene shared a higher homology with both StSN1 and MsSN1 snakins-1, with a peamaclein (PMLN) from *Prunus persica* and with Gibberellin regulated proteins 7, 8, and 10 from *Arabidopsis thaliana* (GASA7; GASA8 and GASAA). All these peptides belong to Snakin/GASA subfamily I, according to the classification of Berrocal-Lobo et al. (2002) (Fig. 2). This analysis was restricted to sequences with reported experimental evidence such as antimicrobial activity, including a snakin-2 from tomato (|E5KBY0|E5KBY0_SOLLC; Herbel et al. 2015), a snakin-1 from alfalfa (|H9D2D5|H9D2D5_MEDSA; García et al. 2014), a snakin-1 from pepper (|B2ZAW4|B2ZAW4_CAPAN; Mao et al. 2011), a snakin-3 from potato

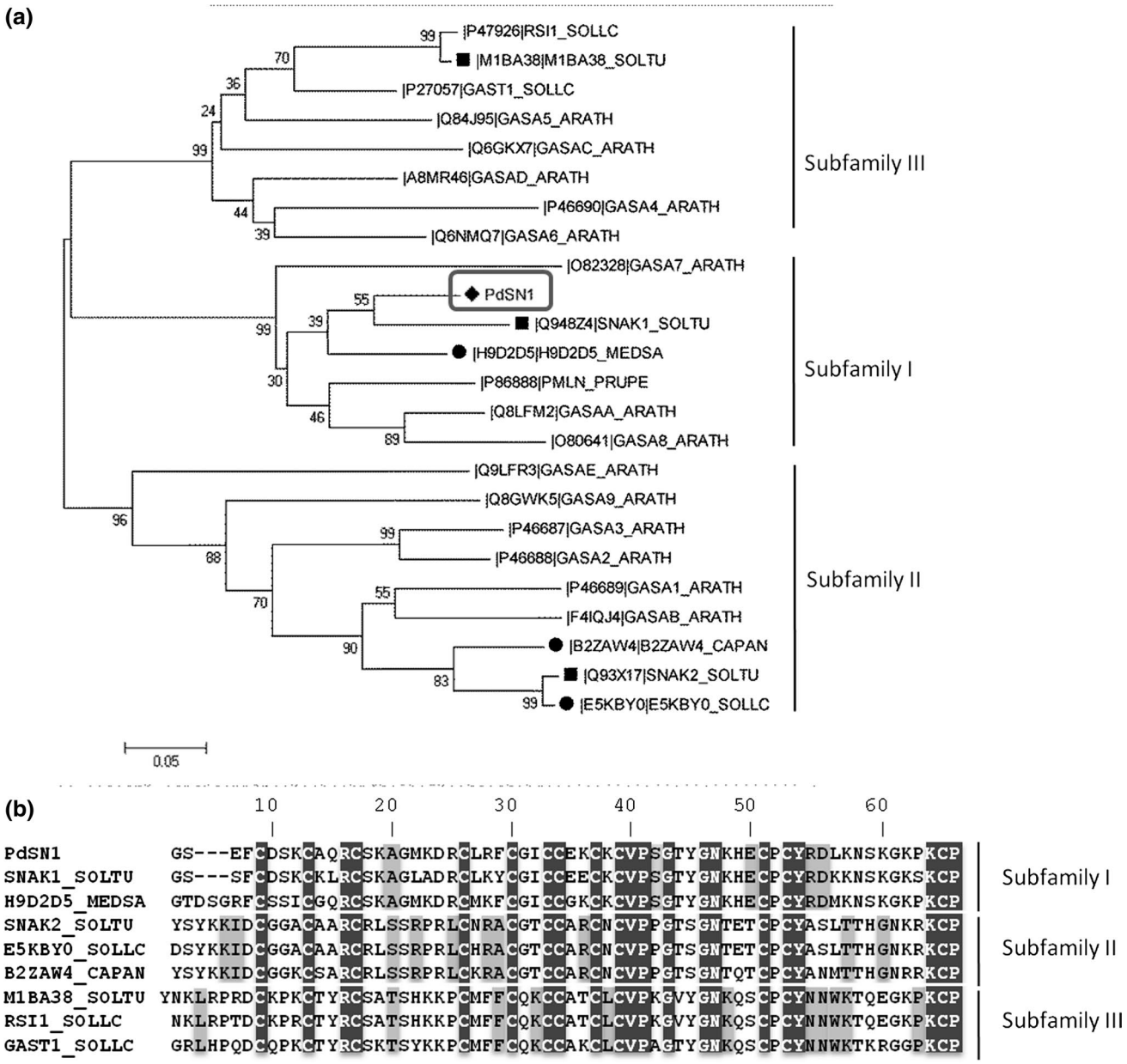

Fig. 2 Sequence alignments and distance matrix of PdSN1 peptide and members of the plant Snakin/GASA protein family. The deduced amino acid sequence of the newly isolated gene (*PdSN1*) was aligned with a set of mature plant proteins with GASA domain. Sequences used were manually annotated and reviewed from the UniProt database (http://www.uniprot.org/). Four snakins with reported experimental evidence were also included (|E5KBY0|E5KBY0_SOLLC, |H9D2D5|H9D2D5_MEDSA, |B2ZAW4|B2ZAW4_CAPAN, |M1BA38|M1BA38_SOLTU). Snakin/GASA proteins grouped following their classification into three subfamilies (I–III) as proposed by Berrocal-Lobo et al. (2002). **a** Neighbor-Joining unrooted tree. Values in the nodes regard bootstrap values (5000 replicates). Each sequence was named according to its UniProt Entry identifier, followed by a Mnemonic identifier of a UniProtKB entry. ARATH: *A. thaliana*, SOLTU: *S. tuberosum*, SOLLC: *Solanum lycopersicum*, PRUPE: *P. persica*. MEDSA: *M. sativa*, CAPAN: *Capsicum annum*. Black diamond indicates PdSN1; black squares indicate snakin-1, snakin-2, and snakin-3 (M1BA38) from *S. tuberosum* and black circles indicate snakin-1 (H9D2D5) from *M. sativa* and snakin-2 from *S. lycopersicum* (E5KBY0) and *C. annum* (B2ZAW4). **b** Alignment of mature sequences of representative snakin/GASA proteins, following their classification in subfamilies I, II, and III (indicated at the right side of the alignment). Amino acids conserved across all the family members are black shaded. Most conserved residues which are relevant for subfamily classification (according to Berrocal-Lobo et al. 2002) are shaded in gray

(|M1BA38|M1BA38_SOLTU; Nahirñak et al. 2016), and sequences obtained from the curated UniProt database (http://www.uniprot.org/).

PdSN1 gene expression in plant

PdSN1 mRNA steady states from leaves and seedlings (inoculated with *A. niger* and uninoculated) were measured by RT-qPCR using three reference genes. The geNorm tool identified *PdACT*, *PdEF1,* and *PdPP2A* as the most stable genes (expression stability value [*M*]: 0.705, 0.903, and 1.069, respectively), among six candidates analyzed in all conditions (methodology described in "Gene expression analysis"). The geometric mean of these three genes was used for normalization of *PdSN1* expression, as recommended by Vandesompele et al. (2002). The expression of *PdSN1* was 40 fold higher in seedling than in adult leaves (Fig. 3). The presence of *A. niger* did not produce significant variations of seedlings' mRNA steady-state levels 48 h after inoculation.

Expression and purification of recombinant PdSN1 in *E. coli*

To reach the highest yield of the soluble Trx-PdSN1 fusion protein, its expression was tested by varying temperatures (30, 28, and 21 °C) and times after induction (2, 4, and 24 h), using *E. coli* strains Rosetta-gami pLysS and Shuffle. Both *E. coli* strains were transformed with pE-PdSN1 and pE-Trx* (control without PdSN1). A schematic representation of the peptides resulting from these vectors, Trx-PdSN1 and Trx*, respectively, is shown in Fig. 4a. The SDS-PAGE analysis of total lysates from induced cells revealed the presence of recombinant proteins Trx-PdSN1 and Trx* with an expected molecular weight of ~ 25 and ~ 17 kDa, respectively (Fig. 4b, lanes 1–4). No significant peptide bands corresponding to the recombinant protein were observed in non-induced cultures (Fig. 4b, lane 2). Some amount of the recombinant protein was present in inclusion bodies, in all conditions. We found that the highest yield (0.9 mg l^{-1} bacterial culture) was obtained with Rosetta-gami(DE3)pLysS at 28 °C and 24 h of induction, while with Shuffle strain, we found a lower percentage of recombinant protein in the soluble fraction (data not shown). The recovery of the 25-kDa band at the end of the purification by Ni-affinity chromatography confirmed that this band contains the recombinant peptide (Fig. 4b, lane 6).

For practical reasons a cleavage on-column, Trx was conducted before peptide elution. A previous pilot test, carried out with several concentrations of enterokinase, showed that three µg of the fusion protein was digested entirely with one unit of the enzyme overnight at room temperature (data not shown). The developed protocol resulted in an efficient separation of the recombinant PdSN1 peptide from Trx (Fig. 4b, lanes 7–8).

The band of the PdSN1 peptide seemed to be larger (ca. 14 kDa) than the expected (11 kDa), possibly due to an abnormal migration which is sometimes observed with cationic peptides on the SDS-PAGE (Shi et al. 2012; Herbel et al. 2015). The results obtained by the mass fingerprinting of the peptide confirmed that its sequence corresponded to PdSN1 (100% coverage, Table S5; Fig. S1).

Antimicrobial activity of the recombinant peptide PdSN1

The antimicrobial activity of the recombinant peptide was tested in vitro against several bacterial and fungal species, as detailed in "Materials and methods". To the maximum concentration (1.8 µM) of PdSN1 tested, significant inhibition of growth (99.7%) was observed for *S. scabies,* while for *S. aureus*, *C. michiganensis* ssp. *michiganensis*, *C. albicans*, *B. cinerea*, *A. niger*, and *A. alternata*, the inhibition was intermediate (60.0, 56.3, 65.5, 53.6, 56.7, and 58.0%, respectively) (Fig. 5). The percentage of inhibition decays at lower concentrations of the peptide for all organisms. *S. aureus* was excluded from Fig. 5, because it only displayed inhibition at the maximum concentration tested (1.80 µM, 56.3% inhibition). No activity was observed for *E. coli*, *X. versicatoria,* and the fungus *P. expansum* at this concentration. IC50 values of recombinant PdSN1 for *C. michiganensis* ssp. *michiganensis, S. scabies, C. albicans, B. cinerea, A. alternata,* and *A. niger* were estimated to be 1.7, 0.3, 1.2, 0.4, 0.4, and 1.4 µM, respectively. IC_{50} value of propiconazole, which was used against *A. niger* as a positive control, was 3.0 µM. IC_{50} value of Gentamicin, used against *C. michiganensis* ssp. m*ichiganensis* was 0.5 µM.

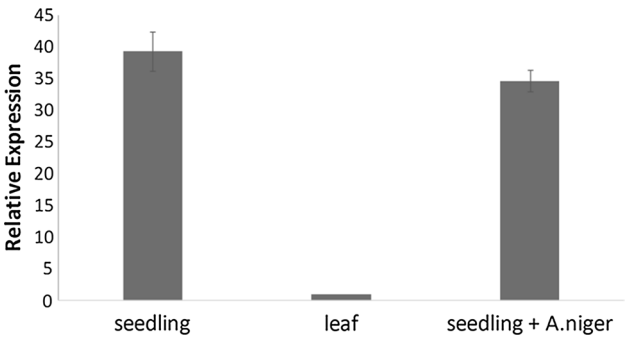

Fig. 3 Expression of *PdSN1* gene in *P. dubium* by RT-qPCR. The geometric mean of *PdACT, PdEF1,* and *PdPP2A* genes was used for normalization of *PdSN1* expression from leaves, mock seedlings, and seedling inoculated with *A. niger*. The final data were obtained by rescaled normalized expression: $(Q_{sample}/NF_{sample})/Min(Q_{sample}/NF_{sample})$. Each bar graph represents the mean relative fold change ± SD of three independent biological replicates

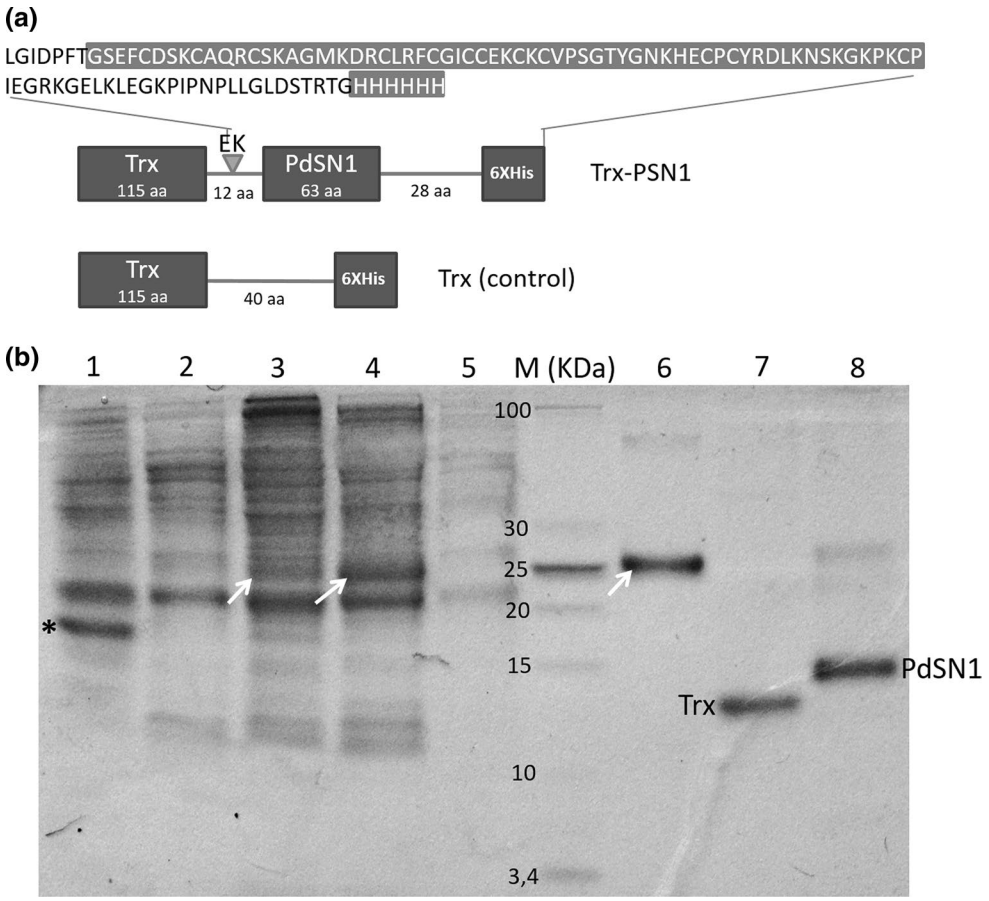

Fig. 4 Affinity purification of the recombinant PdSN1 peptide. **a** Schematic representation of recombinant Trx-PdSN1 and Trx* (control) proteins. **b** Tris–Tricine SDS-PAGE (15%) analysis of the purification steps of the Trx–PdSN1fusion protein and PdSN1 peptide. Protein fractions from *E. coli* Rosetta-gami (DE3) cells transformed with pE-Trx expression plasmid (used as a control) (1), pE-PdSN1 expression plasmid incubated in the absence (2) or the presence of IPTG (3: crude extract; 4: soluble fraction); Unbound proteins removed by the first washing step in which the fusion protein is bound to the column (5); final elution step, where the purified Trx-PdSN1is released from the Ni-affinity column (6); Trx protein after cleavage with enteroquinase (7); elution of PdSN1 peptide (8). *M* PageRuler Low Range Unstained Protein Ladder (ThermoFisher). Asterisk indicates expression of Trx* protein; white arrows show Trx-PdSN1fusion protein

3D structure determination

Using the recently determined X-ray structure of the potato snakin-1 peptide, the homology modeling of the 3D structure of the native PdSN1 was solved. Connection pattern of the disulfide bonds was also taken from the experimental structure, leading after extensive unbiased molecular dynamics simulations to a final model exhibiting an HTH motif (two α-helixes: H1 and H2, linked by a small turn), and two other small helixes (H3 and H4) in the region of the C-terminal flexible loops (Fig. 6a). The extra amino acids present in the recombinant PdSN1 peptide did not alter the HTH domain (present in the GASA domain), nor the presence of the third and fourth helixes, although H4 appeared in a different region nearest to the C-terminal end (Fig. 6b). 3D predictions coming from pure ab initio methods (a combination of QUARK and MODELLER) and different disulfide bond predictors (see Table S4) also conserved the HTH motif (Fig. 6c), although the connectivity pattern of disulfide bonds was substantially different across the models (Table S4). Furthermore, the model built without disulfide bonds, where all cysteines were represented in their reduced form, also showed a stable HTH motif in the microsecond timescale (Fig. 6c). The time evolution of the root mean squared displacements of the HTH backbone in respect to the X-ray conformation showed that all structures converged to a stable HTH motif (after 650 ns) which was less than 2.75 Å apart from the crystal (Fig. 6d). The superposition of the final structures obtained from MD clearly highlights the similarities within the HTH motifs of all models, independent of the presence of disulfide bridges or the specific disulfide bond connectivity.

While the HTH was present in all cases given to all models a very similar 3D conformation on the N-terminal side,

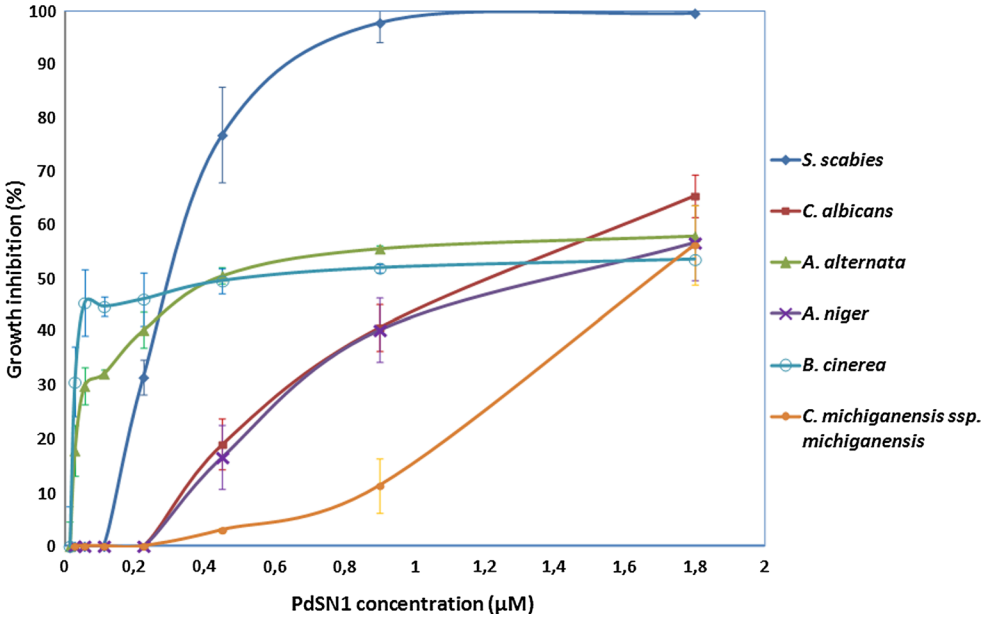

Fig. 5 In vitro antimicrobial activity of PdSN1. Dose-dependent growth inhibition curves of the following bacterial and fungal pathogens: *C. michiganensis* ssp. *michiganensis* (black circles), *S. scabies* (diamonds), *A. niger* (crosses), *B. cinerea* (open circles), *A. alternata* (triangles), and *C. albicans* (squares) were determined by measuring the cultures at OD595, in the presence of different amounts of recombinant PdSN1. The data are mean ± SD ($n = 3$)

the C-terminal region was quite variable (see the red arrows in Fig. 6a–c). The model built by homology modeling was similar to the X-ray structure, showing two small helixes embedded in between flexible loops (Fig. 6a), while the extra amino acids present in the recombinant PdSN1 sequence increased the separation between H3 and H4 with more disordered loops (Fig. 6b). The ab initio models have different C-terminal regions, one of which showed no H3 and H4 helixes (see $PdSN1_{Q\text{-}CYS}$ in Fig. 6c). Interestingly, the model with reduced cysteines exhibited a helix in the C-terminal region which seems to be a condensation of the 3_{10} helix (H3) and α-helix (H4) present in the X-ray structure of the potato snakin-1 peptide. Finally, the computed molecular electrostatic potential of the HTH motifs and extra helixes present in the recombinant PdSN1 peptide also showed a positively charged cleft (see the blue arrow in Fig. 6e), as in the X-ray structure, and a tail of histidines (only present in our recombinant peptide) that could be protonated, enhancing the electrophilic tendency of the peptide (Fig. 6e).

Discussion

In this study, a new snakin gene from *P. dubium* was isolated, aiming to explore the potential of this South American native plant as a producer of novel AMPs with antimicrobial activity. For this purpose, a PCR-based isolation strategy was used to amplify putative snakins, using degenerate primers designed from known snakin sequences of other plants. With the isolated sequence, the potential antimicrobial activity of the new snakin was tested in vitro, using a heterologous expression system for its production. In addition, an introductory analysis was performed to study the expression of this new snakin gene in different plant tissues. Finally, using comparative and ab initio modeling together with molecular dynamics simulations, we explored the structure–activity relationship of this peptide.

Alignment of the snakin genes from the NCBI and UniProt databases revealed a low level of similarity in the sequence that encodes the signal peptide. For this reason, we choose to use degenerate primers, which allowed for the obtention of a sequence, named *PdSN1* (snakin gene from *P. dubium*). This new snakin exhibits an ORF that codes a peptide of 88 residues including a GASA domain with 12 cysteines in conserved positions, sharing 76% identity with potato StSN1. After comparing PdSN1 against UniProt curated sequences bearing a GASA domain, the most significant similarity was obtained with StSN1 from potato, MsSN1 from *M. sativa*, GASA7, GASA8, and GASA10 (GASAA) from Arabidopsis and PMLN from *P. persica*. Of these proteins, only StSN1 (Segura et al. 1999; Kovalskaya and Hammond 2009; Kuddus et al. 2016) and MsSN1 (García et al. 2014) were reported to have antimicrobial activity. GASA10, expressed in *E. coli*, showed toxicity to the host cell, while no significant effect on several microorganisms growth was observed when expressed in *P. pastoris* (Trapalis et al. 2017). From its side, Peamaclein (PMLN), found in protein extracts from peach peel and pulp, was characterized as a new allergen (Tuppo et al. 2013). An NJ analysis done with MsSN1 and additional type 1 snakins reported that MsSN1 could be ortholog to snakin-1 from *Solanum tuberosum* and, ortholog to the GASA7 protein from *A. thaliana* (García et al. 2014), suggesting that they can share with PdSN1 a similar function in plants.

The expression of *PdSN1* gene was analyzed in seedlings and leaves with expression in both, but steady-state

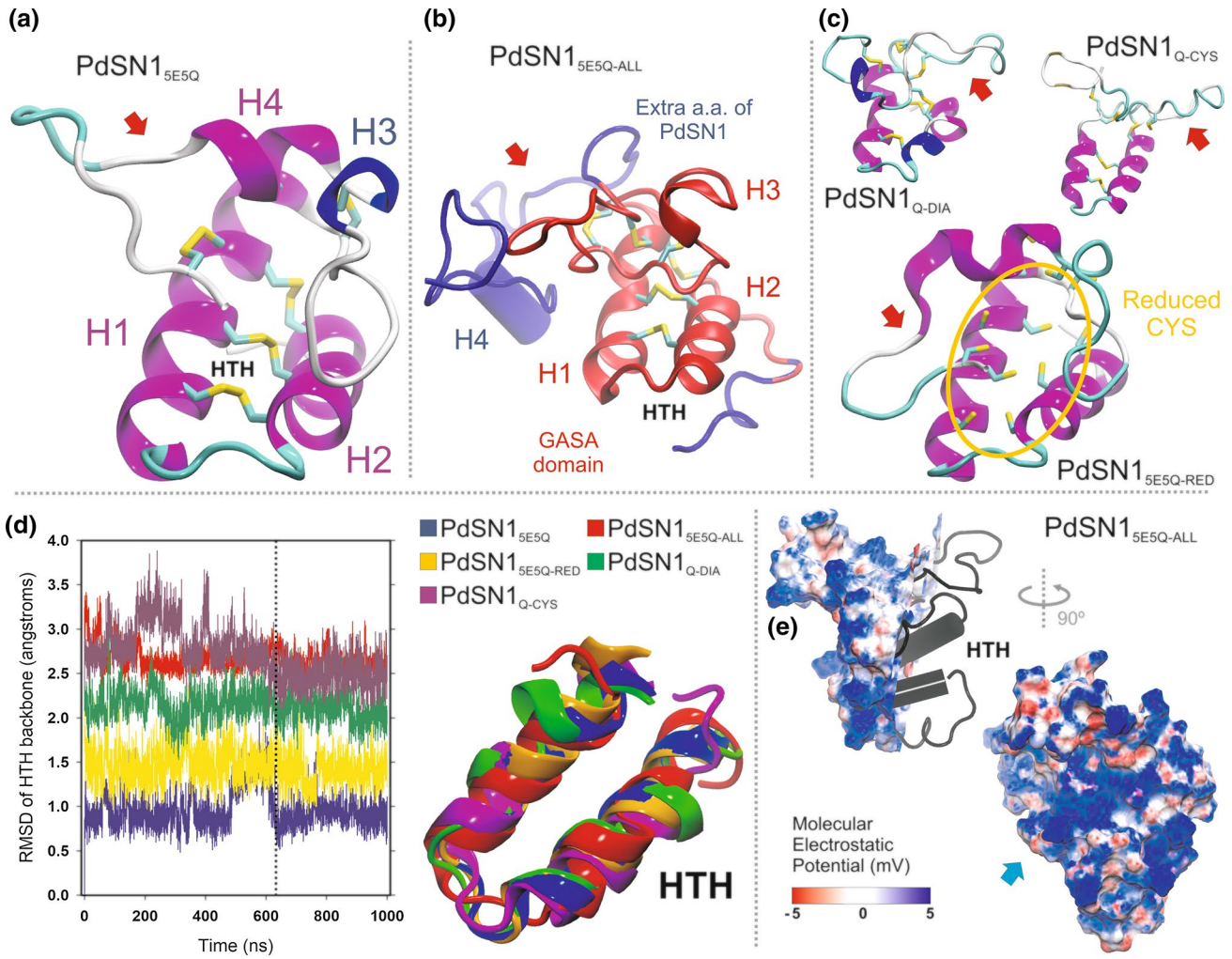

Fig. 6 Homology modeling and 3D structure prediction of the native and recombinant PdSN1 peptide. **a** Structural prediction of the native PdSN1 with his GASA domain, obtained using as template the X-ray structure with PDB id 5E5Q, showing two long helixes (H1 and H2) that define the helix–turn–helix motif (HTH), and two short helixes (H3 and H4) located in the flexible loops of the C-terminal region. **b** Model of the full PdSN1 peptide. The GASA domain is depicted in red while remaining amino acids are shown in blue. **c** 3D structure predictions combining QUARK, DIANNA and/or CYSCON, servers (top); and, bottom, the last structure of the long MD simulation (1.5 µs) of the PdSN1 GASA domain, where all cysteines were considered in their reduced form with no disulfide bonds (see "Materials and methods"). **d** Root mean square deviations in Å of the backbone atoms of the HTH motif respect to the X-ray structure (5E5Q) along time. Structural superposition of the HTH motifs obtained from the last frame of the corresponding MD simulations. **e** Molecular electrostatic potential (MEP) of the full recombinant PdSN1 peptide (front and side view). Note that a clipping plane was used in the top structure to cut the MEP allowing to visually locating the HTH motif (color figure online)

levels of *PdSN1* mRNA were higher in seedlings than in adult leaves. The results obtained for leaves are in agreement with recently reported data from *StSN1* (Meiyalaghan et al. 2014; Nahirñak et al. 2016). However, in the past, Segura et al. (1999) and Berrocal-Lobo et al. (2002) did not detect *StSN1* mRNA in potato leaves. The high levels of mRNA detected in seedlings suggest that expression of AMPs genes are strong during seedling development, justified by their susceptibility to the attacked by a large number of pathogens present in the soil. Strikingly, there was no significant increase in expression 48 h after inoculation with *A. niger*, a pathogen commonly found in soil. Plant AMPs are found constitutively in storage organs like seeds and generative tissues (reproductive organs, fruits, and flowers; revised by Benko-Iseppon et al. 2010), so our data suggest that PdSN1 could be a component of the constitutive defense barriers organs and tissues during their early development, besides fulfilling a physiological role in seedlings (Nahirñak et al. 2012b). A more exhaustive analysis should be done, including a follow-up in time after inoculation with the fungus, something beyond the scope of the present contribution.

The mature PdSN1 peptide was successfully produced in *E. coli*. Owing to the codon bias and the inability of *E. coli* to form disulfide bridges within its cytoplasm (Lobstein

et al. 2012), we decided to work with Rosetta-gami and Shuffle strains that allow for the expression of eukaryotic proteins that contain codons rarely used in *E. coli*, while enhancing disulfide bond formation. Kovalskaya and Hammond (2009) reported that the use of a recombinant expression system like *E. coli* for the production of potato defensin PTH1 and snakin-1 yielded inactive protein aggregates which needed to be denatured and refolded for gaining activity. To minimize protein aggregates into inclusion bodies, a polypeptide fusion partner such as the hydrophilic tag thioredoxin (Trx) (Terpe 2003) was used. Such tags could also serve to avoid the possible inherent sensitivity of *E. coli* towards the expression of peptides with antibacterial activity. We used pET102/D which included a N-terminal Trx protein, a C-terminal 6xHis tag, and a protease recognition site for cleavage of the fusion Trx from the recombinant peptide PdSN1. This last feature is important, since some fused proteins may not possess antimicrobial activity. For example, a scots pine defensin expressed in *E. coli*, fused to glutathione S-transferase (GST), was biologically inactive until GST was removed (Kovaleva et al. 2011). Using Rosetta-gami/pET102/D, we obtained the snakin peptide mostly in a soluble form, which after a single chromatographic step, resulted in the purification of a biologically active peptide.

The purified recombinant PdSN1 showed variable in vitro antibacterial and antifungal activities, depending on the microorganism analyzed. The Gram-negative bacteria *X. versicatoria* and *E. coli*, besides the fungus *P. expansum,* were not sensitive to PdSN1 at the maximum concentration tested (1.8 µM). Higher concentrations should be used to affirm that PdSN1 do not affect their growth. In fact, the recombinant StSN1 showed activity against the Gram-negative bacteria *Salmonella enterica* and *E. coli* at concentrations of 5 and 10 µM (Kuddus et al. 2016). The PdSN1 IC_{50} values are comparable to others, either native or recombinant snakin peptides. In our experiments, IC_{50} for the fungus *B. cinerea* was 0.4 µM, while for native StSN1 and StSN2, the effective concentration for 50% inhibition was 0.8 and 2.0 µM, respectively (Berrocal-Lobo et al. 2002). When StSN1 was produced in *E. coli,* IC_{50} for *B. cinera* was 9.0 µM (Kovalskaya and Hammond 2009); this peptide was obtained exclusively in the form of inclusion bodies and, therefore, needed to be solubilized and refolded for activity. Kuddus et al. (2016) reported a recombinant StSN1 expressed in *P. pastoris* with minimum fungicidal concentration for *Candida parapsilosis* and *P. pastoris* yeasts, of 5.0 and 10.0 mM, respectively. Our PdSN1 had an IC_{50} value for the yeast *C. albicans* in the concentration of 1.2 µM. A snakin-2 from tomato showed activity against all tested microorganisms (Gram-negative bacteria, Gram-positive bacteria, and fungi) with IC_{50} values between 0.1 µM for the Gram-positive bacteria *Micrococcus luteus* and 1.6 µM for the mold *Fusarium solani* (Herbel et al. 2015). Interestingly, our recombinant PdSN1 showed potent activity (99.7% inhibition with 1.8 µM of the peptide) against the Gram-positive bacteria *S. scabies*, a plant pathogen that causes the economically relevant potato disease named 'common scab'. PdSN1 also displayed intermediate activity against other critical plant pathogens like *C. michiganensis* ssp. *michiganensis, A. alternata,* and *B. cinerea* while inhibiting the growth of opportunistic pathogens such as *C. albicans* and *A. niger*.

The 3D structure obtained from the PdSN1 sequence using the X-ray structure of the potato snakin-1 as a template showed a clear HTH motif in the GASA domain that was preserved in the microsecond long MD simulation of the peptide that was performed in near physiological conditions. As previously observed (Yeung et al. 2016), the HTH motif is also present in the plant thionins and in the α-helical hairpin protein classes of cysteine-rich antimicrobial peptides. These other classes exhibited not only the same two helixes of the HTH motif in the same relative orientation but also shared some of the disulfide bond connections (Vila-Perelló et al. 2005). It has been suggested, based on the antimicrobial activity of truncated thionin containing only the HTH motif (Vila-Perelló et al. 2005), that these two helixes linked by a short turn seem to be essential for the activity. Our results show that the HTH motif is present independently of the specific connectivity pattern of the disulfide bonds. Moreover, the complete absence of those disulfide bonds, after modeling the PdSN1 peptide with all the cysteines in their reduced form, also lead to an HTH motif stable in the microsecond timescale (this simulation was extended to 1.5 µs with the same result).

Altogether, these results suggest that the HTH motif is strongly determined by the sequence, and do not primarily rely on the disulfide bonds, explaining the observed antimicrobial activity of unfolded synthetic potato snakin-1 and -2, where the cysteines were in their reduced form (Harris et al. 2014). Furthermore, recombinant peptides produced in BL21(DE3) *E. coli*, were still active (Mao et al. 2011; Herbel et al. 2015). In this strain, the folding could not be guided by the formation of disulfide bridges, and reinforcing the hypothesis that disulfide bonds may be not essential for antimicrobial activity. The HTH is an evident evolutionary conserved motif of plant defense peptides, and the disulfide bonds could have appeared later in the evolution, collaborating to further stabilize the structural motif and providing and entropic advantage when considering the free energy of binding of snaking-1 to its biological targets.

As in the case of the potato snakin-1, PdSN1 also displayed a large positive electrostatic surface, with a pronounced electrophile cleft. It has been proposed that this may be particularity important in targeting these peptides to their site of action (Yeung et al. 2016). Since many antimicrobial peptides exert their function by disrupting (or interacting with) the negatively charged microbial membrane surfaces,

a lipid-membrane-targeting mechanism is the currently most accepted hypothesis about their mode of action. Nevertheless, the other natural target for positively charged peptides is DNA, which is the most negatively charged polymer found in nature (Cuervo et al. 2014). Moreover, the HTH motif is a well-established motif found in many proteins that regulate gene expression, like transcription factors, or the famous λ-repressor (Wintjens and Rooman 1996), where the motif is known to bind to the major groove of DNA.

The positive molecular electrostatic potential (MEP), added to the apparent requirement of the HTH motif for antimicrobial activity, leads us to hypothesize about a possible additional mechanism of action based on the deregulation of the microbial gene expression, which could be acting together with the lipid-membrane-targeting mechanism hypothesis. Indeed, some evidence already exists suggesting that snakin peptides could disrupt the microbial membrane (Herbel et al. 2015; Kuddus et al. 2016), besides some other types of AMPs that are known to interact with intracellular targets like DNA, RNA, or proteins (Brogden 2005; Lay and Anderson 2005). Considering defensins, for example, uncertainties have arisen on which are the main antimicrobial strategies: membrane permeabilization and leakage of cytoplasmic contents, or intracellular interaction with DNA/RNA and proteins from the synthesis machinery (Lay and Anderson, 2005). All these hypotheses about the mode of action clearly deserve future studies for confirmation, which are beyond the present contribution.

Conclusions

Given that snakins with antimicrobial activity are strong candidates for the development of novel biotechnological products meant for the control of diseases, we focused our work on the isolation of a new snakin gene from an unexplored genome such as a native, not cultivated tree with ethnobotanical tradition. Using a PCR-based strategy with degenerate primers, we successfully isolated a new snakin-like gene, named *PdSN1*. This gene was expressed in *E. coli*, allowing for the obtention of small amounts of a purified recombinant peptide with promising antimicrobial activity against relevant plant pathogens and opportunistic human fungi. Our 3D structural analysis brought some insights about possible relationships between the HTH motif present in the GASA domain and PdSN1 activity. Additional studies are ongoing in our laboratories to further characterize the antimicrobial activity of PdSN1 and to verify its possible membrane and/or DNA interactions.

Acknowledgements The authors thank CSIC (Comisión Sectorial de Investigación Científica, Universidad de la República, Uruguay), PEDECIBA (Programa de Desarrollo de las Ciencias Básicas, Uruguay), ANII (Agencia Nacional de Investigación e Innovación, Uruguay), CNPq (National Council for Scientific and Technological Development, Brazil), and CAPES (Coordination for the Improvement of Higher Education Personnel, Brazil) for financial support and fellowships. We also thank Dr. Rosario Durán and Madelón Portela for MALDI-TOFF analysis. SRD was supported by a doctoral fellowship from CSIC. IRB Barcelona is the recipient of a Severo Ochoa Award of Excellence from MINECO (Ministerio de Economía, Industria y Competitividad. Government of Spain). PDD and GC are PEDECIBA and SNI (Sistema Nacional de Investigadores, Uruguay) researchers.

Compliance with ethical standards

Conflict of interest The authors declare that no competing interests exist.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

Almasia NI, Bazzini AA, Hopp HE, Vazquez-Rovere C (2008) Overexpression of snakin-1 gene enhances resistance to *Rhizoctonia solani* and *Erwinia carotovora* in transgenic potato plants. Mol Plant Pathol 9:329–338. https://doi.org/10.1111/J.1364-3703.2008.00469.X

Almasia NI, Molinari MP, Maroniche GA, Nahirñak V, Barrios Barón MP, Taboga OA, Vazquez-Rovere C (2017) Successful production of the potato antimicrobial peptide Snakin-1 in baculovirus-infected insect cells and development of specific antibodies. BMC Biotechnol 17:1–11. https://doi.org/10.1186/s12896-017-0401-2

Balaji V, Smart CD (2012) Over-expression of snakin-2 and extensin-like protein genes restricts pathogen invasiveness and enhances tolerance to *Clavibacter michiganensis* subsp. *michiganensis* in transgenic tomato (*Solanum lycopersicum*). Transgenic Res 21:23–37. https://doi.org/10.1007/s11248-011-9506-x

Balaji V, Sessa G, Smart CD (2011) Silencing of host basal defense response-related gene expression increases susceptibility of *Nicotiana benthamiana* to *Clavibacter michiganensis* subsp. *michiganensis*. Phytopathology 101:349–357. https://doi.org/10.1094/PHYTO-05-10-0132

Benko-Iseppon AM, Galdino SL, Calsa T, Kido EA, Tossi A, Belarmino LC, Crovella S (2010) Overview on plant antimicrobial peptides. Curr Protein Pept Sci 11:181–188. https://doi.org/10.2174/138920310791112075

Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81:3684–3688. https://doi.org/10.1063/1.448118

Berrocal-Lobo M, Segura A, Moreno M, López G, García-Olmedo F, Molina A (2002) Snakin-2, an antimicrobial peptide from potato whose gene is locally induced by wounding and responds to pathogen infection. Plant Physiol 128:951–961. https://doi.org/10.1104/pp.010685.1

Bradford MM (1976) A rapid and sensitive method for the quantitation microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254. https://doi.org/10.1016/0003-2697(76)90527-3

Broekaert WF, Terras F, Cammue BP, Vandedeyden J (1990) An automated quantitative assay for fungal growth inhibition. FEMS Microbiol Lett 69:55–60. https://doi.org/10.1016/0378-1097(90)90412-J

Brogden KA (2005) Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? Nat Rev Microbiol 3:238–250. https://doi.org/10.1038/nrmicro1098

Case DA, Cerutti DS, Cheatham TE III, Darden TA, Duke RE, Giese TJ, Gohlke H, Goetz AW, Greene D, Homeyer N, Izadi S, Kovalenko A, Lee TS, LeGrand S, Li P, Lin C, Liu J, Luchko T, Luo R, Mermelstein D, Merz KM, Monard G, Nguyen H, Omelyan I, Onufriev A, Pan F, Qi R, Roe DR, Roitberg A, Sagui C, Simmerling CL, Botello-Smith WM, Swails J, Walker RC, Wang J, Wolf RM, Wu X, Xiao L, York DM, Kollman PA (2017) AMBER 2017. University of California, San Francisco

Ceroni A, Passerini A, Vullo A, Frasconi P (2006) DISULFIND: a disulfide bonding state and cysteine connectivity prediction server. Nucleic Acids Res 34:W177–W181. https://doi.org/10.1093/nar/gkl266

Cheng J, Saigo H, Baldi P (2006) Large-scale prediction of disulphide bridges using kernel methods, two-dimensional recursive neural networks, and weighted graph matching. Proteins: structure, function. Bioinformatics 62:617–629. https://doi.org/10.1002/prot.20787

Cuervo A, Dans PD, Carrascosa JL, Orozco M, Gomila G, Fumagalli L (2014) Direct measurement of the dielectric polarization properties of DNA. Proc Natl Acad Sci 111:E3624–E3630. https://doi.org/10.1073/pnas.1405702111

Dans PD, Danilāne L, Ivani I, Dršata T, Lankaš F, Walther J, Illa Pujagut R, Battistini F, Gelpí JL, Lavery R, Orozco M (2016) Long-timescale dynamics of the Drew–Dickerson dodecamer. Nucleic Acids Res 44:4052–4066. https://doi.org/10.1093/nar/gkw264

Darden T, York D, Pedersen L (1993) Particle mesh Ewald: an N-log(N) method for Ewald sums in large systems. J Chem Phys 98:10089–10095. https://doi.org/10.1063/1.464397

Doyle J (1991) DNA protocols for plants-CTAB total DNA isolation. In: Hewitt GM, Johnston A (eds) Molecular techniques in taxonomy. Springer, Berlin, pp 283–293

Ferrè F, Clote P (2006) DiANNA 1.1: an extension of the DiANNA web server for ternary cysteine classification. Nucleic Acids Res 34:W182–W185. https://doi.org/10.1093/nar/gkl189

Frohman MA, Dush MK, Martin GR (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc Natl Acad Sci 85:8998–9002

García AN, Ayub ND, Fox AR, Gómez MC, Diéguez MJ, Pagano EM, Berini CA, Muschietti JP, Soto G (2014) Alfalfa snakin-1 prevents fungal colonization and probably coevolved with rhizobia. BMC Plant Biol 14:248. https://doi.org/10.1186/s12870-014-0248-9

Hall T (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98

Harris PWR, Yang SH, Molina A, Lopez G, Middleditch M, Brimble MA (2014) Plant antimicrobial peptides snakin-1 and snakin-2: chemical synthesis and insights into the disulfide connectivity. Chemistry 20:5102–5110. https://doi.org/10.1002/chem.201303207

Herbel V, Schäfer H, Wink M (2015) Recombinant production of snakin-2 (an antimicrobial peptide from tomato) in *E. coli* and analysis of its bioactivity. Molecules 20:14889–14901. https://doi.org/10.3390/molecules200814889

Herbel V, Sieber-frank J, Wink M (2017) The antimicrobial peptide snakin-2 is upregulated in the defense response of tomatoes (*Solanum lycopersicum*) as part of the jasmonate-dependent signaling pathway. J Plant Physiol 208:1–6. https://doi.org/10.1016/j.jplph.2016.10.006

Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. J Mol Graph 14:33–38. https://doi.org/10.1016/0263-7855(96)00018-5

Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79:926–935. https://doi.org/10.1063/1.445869

Kovaleva V, Krynytskyy H, Gout I, Gout R (2011) Recombinant expression, affinity purification and functional characterization of Scots pine defensin 1. Appl Microbiol Biotechnol 89:1093–1101. https://doi.org/10.1007/s00253-010-2935-2

Kovalskaya N, Hammond RW (2009) Expression and functional characterization of the plant antimicrobial snakin-1 and defensin recombinant proteins. Protein Expr Purif 63:12–17. https://doi.org/10.1016/j.pep.2008.08.013

Kuddus MR, Rumi F, Tsutsumi M, Takahashi R, Yamano M, Kamiya M, Kikukawa T, Demura M, Aizawa T (2016) Expression, purification and characterization of the recombinant cysteine-rich antimicrobial peptide snakin-1 in *Pichia pastoris*. Protein Expr Purif 122:15–22. https://doi.org/10.1016/j.pep.2016.02.002

Lay FT, Anderson MA (2005) Defensins—components of the innate immune system in plants. Curr Protein Pept Sci 6:85–101. https://doi.org/10.2174/1389203053027575

Lobstein J, Emrich CA, Jeans C, Faulkner M, Riggs P, Berkmen M (2012) SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. Microb Cell Fact 11:753. https://doi.org/10.1186/1475-2859-11-56

Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C (2015) ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99sb. J Chem Theor Comput 11:3696–3713. https://doi.org/10.1021/acs.jctc.5b00255

Mao Z, Zheng J, Wang Y, Chen G, Yang Y, Feng D, Xie B (2011) The new CaSn gene belonging to the snakin family induces resistance against root-knot nematode infection in pepper. Phytoparasitica 39:151–164. https://doi.org/10.1007/s12600-011-0149-5

Meiyalaghan S, Thomson SJ, Fiers MW, Barrell PJ, Latimer JM, Mohan S, Jones EE, Conner AJ, Jacobs JM (2014) Structure and expression of GSL1 and GSL2 genes encoding gibberellin stimulated-like proteins in diploid and highly heterozygous tetraploid potato reveals their highly conserved and essential status. BMC Genom 15:2. https://doi.org/10.1186/1471-2164-15-2

Mohan S, Meiyalaghan S, Latimer JM, Gatehouse ML, Monaghan KS, Vanga BR, Pitman AR, Jones EE, Conner AJ, Jacobs JME (2014) GSL2 over-expression confers resistance to *Pectobacterium atrosepticum* in potato. Theor Appl Genet 127:677–689. https://doi.org/10.1007/s00122-013-2250-2

Nahirñak V, Almasia NI, Fernandez PV, Hopp HE, Estevez JM, Carrari F, Vazquez-Rovere C (2012a) Potato snakin-1 gene silencing affects cell division, primary metabolism, and cell wall composition. Plant Physiol 158:252–263. https://doi.org/10.1104/pp.111.186544

Nahirñak V, Almasia NI, Hopp HE, Vazquez-Rovere C (2012b) Involvement in hormone crosstalk and redox homeostasis Snakin/GASA proteins. Plant Signal Behav 7:1004–1008. https://doi.org/10.4161/psb.20813

Nahirñak V, Rivarola M, Gonzalez de Urreta M, Paniego N, Hopp HE, Almasia NI, Vazquez-Rovere C (2016) Genome-wide analysis of the snakin/gasa gene family in *Solanum tuberosum* cv. Am J Potato Res, Kennebec. https://doi.org/10.1007/s12230-016-9494-8

Oliveira-Lima M, Benko-Iseppon AM, Ferreira Costa RJ, Rodríguez-Decuadro S, Kido EA, Crovella S, Pandolfi V (2017) Snakin: structure, roles and applications of a plant antimicrobial peptide. Curr Protein Pept Sci 18:1–7. https://doi.org/10.2174/1389203717666161606191

Padovan L, Crovella S, Tossi A, Segat L (2010) Techniques for plant defensin production. Curr Protein Pept Sci 11:231–235. https://doi.org/10.2174/138920310791112101

Pérez A, Luque FJ, Orozco M (2007) Dynamics of B-DNA on the microsecond time scale. J Am Chem Soc 129:14739–14745. https://doi.org/10.1021/ja0753546

Pestana-Calsa MC, Ribeiro ILAC Jr (2010) Bioinformatics-coupled molecular approaches for unravelling potential antimicrobial peptides coding genes in Brazilian native and crop plant species. Curr Protein Pept Sci 11:199–209. https://doi.org/10.2174/138920310791112138

Porto WF, Franco OL (2013) Theoretical structural insights into the snakin/GASA family. Peptides 44:163–167. https://doi.org/10.1016/j.peptides.2013.03.014

Richero M, Barraco-Vega M, Cerdeiras MP, Cecchetto G (2013) Development of SCAR molecular markers for early and late differentiation of *Eucalyptus globulus* ssp globulus from *E. globulus* ssp *maidenii*. Trees 27:249–257. https://doi.org/10.1007/s00468-012-0792-6

Roe DR, Cheatham TE III (2013) PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. J Chem Theory Comput 9:3084–3095. https://doi.org/10.1021/ct400341p

Rong W, Qi L, Wang J, Du L, Xu H, Wang A, Zhang Z (2013) Expression of a potato antimicrobial peptide SN1 increases resistance to take-all pathogen *Gaeumannomyces graminis* var. *tritici* in transgenic wheat. Funct Integr Genom 13:403–409. https://doi.org/10.1007/s10142-013-0332-5

Ryckaert J-P, Ciccotti G, Berendsen HJ (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. J Comput Phys 23:327–341. https://doi.org/10.1016/0021-9991(77)90098-5

Salomon-Ferrer R, Götz AW, Poole D, Le Grand S, Walker RC (2013) Routine microsecond molecular dynamics simulations with amber on gpus. 2. Explicit solvent particle mesh Ewald. J Chem Theory Comput 9:3878–3888. https://doi.org/10.1021/ct400314y

Savojardo C, Fariselli P, Alhamdoosh M, Martelli PL, Pierleoni A, Casadio R (2011) Improving the prediction of disulfide bonds in Eukaryotes with machine learning methods and protein subcellular localization. Bioinformatics 27:2224–2230. https://doi.org/10.1093/bioinformatics/btr387

Schägger G, von Jagow G (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal Biochem 379:368–379. https://doi.org/10.1016/0003-2697(87)90587-2

Segura A, Moreno M, Madueño F, Molina A, García-Olmedo F (1999) Snakin-1, a peptide from potato that is active against plant pathogens. Mol Plant Microbe Interact 12:16–23. https://doi.org/10.1094/MPMI.1999.12.1.16

Shi Y, Mowery RA, Ashley J, Hentz M, Ramirez AJ, Bilgicer B, Slunt-brown H, Borchelt DR, Shaw BF (2012) Abnormal SDS-PAGE migration of cytosolic proteins can identify domains and mechanisms that control surfactant binding. Protein Sci 21:1197–1209. https://doi.org/10.1002/pro.2107

Silva ON, Mulder KCL, Barbosa AEAD, Otero-Gonzalez AJ, Lopez-Abarrategui C, Rezende TMB, Dias SC, Franco OL (2011) Exploring the pharmacological potential of promiscuous host-defense peptides: from natural screenings to biotechnological applications. Front Microbiol 2:1–14. https://doi.org/10.3389/fmicb.2011.00232

Smith DE, Dang LX (1994) Computer simulations of NaCl association in polarizable water. J Chem Phys 100:3757–3762. https://doi.org/10.1063/1.466363

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods research resource. Mol Biol Evol 28:2731–2739. https://doi.org/10.1093/molbev/msr121

Terpe K (2003) Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. Appl Microbiol Biotechnol 60:523–533. https://doi.org/10.1007/s00253-002-1158-6

Thevissen K, Kristensen H, Thomma BPHJ, Cammue BPA, Francois IEJA (2007) Therapeutic potential of antifungal plant and insect defensins. Drug Discov Today 12:966–971. https://doi.org/10.1016/j.drudis.2007.07.016

Tossi A, Sandri L (2002) Molecular diversity in gene-encoded, cationic antimicrobial polypeptides. Curr Pharm Des 8:743–761. https://doi.org/10.2174/1381612023395475

Trapalis M, Li SF, Parish RW (2017) The Arabidopsis GASA10 gene encodes a cell wall protein strongly expressed in developing anthers and seeds. Plant Sci 260:71–79. https://doi.org/10.1016/j.plantsci.2017.04.003

Tuppo L, Alessandri C, Pomponi D, Picone D, Tamburrini M, Ferrara R, Petriccione M, Mangone I, Palazzo P, Liso M, Giangrieco I, Crescenzo R, Bernardi M, Zennaro D, Helmer-Cittrerich M, Mari A, Ciardiello M (2013) Peamaclein—a new peach allergenic protein: similarities, differences and misleading features compared to Pru p 3. Clin Exp Allergy 43:128–140. https://doi.org/10.1111/cea.12028

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 3:Research 0034. https://doi.org/10.1186/gb-2002-3-7-research0034

Vila-Perelló M, Sanchez-Vallet A, Garcıa-Olmedo F, Molina A, Andreu D (2005) Structural dissection of a highly knotted peptide reveals minimal motif with antimicrobial activity. J Biol Chem 280:1661–1668. https://doi.org/10.1074/jbc.M410577200

Webb B, Sali A (2014) Comparative protein structure modeling using modeller. Curr Protoc Bioinform. https://doi.org/10.1002/0471250953.bi0506s15

Wintjens R, Rooman M (1996) Structural classification of hth dna-binding domains and protein–dna interaction modes. J Mol Biol 262:294–313. https://doi.org/10.1006/jmbi.1996.0514

Xu D, Zhang Y (2012) *Ab initio* protein structure assembly using continuous structure fragments and optimized knowledge-based force field. Proteins 80:1715–1735. https://doi.org/10.1002/prot.24065

Yang J, He BJ, Jang R, Zhang Y, Shen HB (2015) Accurate disulfide-bonding network predictions improve ab initio structure prediction of cysteine-rich proteins. Bioinformatics 31:3773–3781. https://doi.org/10.1093/bioinformatics/btv459

Yeaman MR, Yount NY (2003) Mechanisms of antimicrobial peptide action and resistance. Pharmacol Rev 55:27–55. https://doi.org/10.1124/pr.55.1.2

Yeung H, Squire CJ, Yosaatmadja Y, Panjikar S, López G, Molina A, Baker EN, Harris PWR, Brimble MA (2016) Radiation damage and racemic protein crystallography reveal the unique structure of the gasa/snakin protein superfamily. Angew Chem Int Ed 55:7930–7933. https://doi.org/10.1002/ange.201602719

Zasloff M (2002) Antimicrobial peptides of multicellular organisms. Nature 415:389–395. https://doi.org/10.1038/415389a

Zhu J, Zhang L, Li W, Han S, Yang W, Qi L (2013) Reference gene selection for quantitative real-time pcr normalization in *Caragana intermedia* under different abiotic stress conditions. PLoS One 8:e53196. https://doi.org/10.1371/journal.pone.0053196