

# SCIENTIFIC REPORTS



OPEN

## Identification and Characterization of Roseltide, a Knottin-type Neutrophil Elastase Inhibitor Derived from *Hibiscus sabdariffa*

Received: 01 June 2016  
Accepted: 22 November 2016  
Published: 19 December 2016

Shining Loo\*, Antony Kam\*, Tianshu Xiao, Giang K.T. Nguyen, Chuan Fa Liu & James P. Tam

Plant knottins are of therapeutic interest due to their high metabolic stability and inhibitory activity against proteinases involved in human diseases. The only knottin-type proteinase inhibitor against porcine pancreatic elastase was first identified from the squash family in 1989. Here, we report the identification and characterization of a knottin-type human neutrophil elastase inhibitor from *Hibiscus sabdariffa* of the Malvaceae family. Combining proteomic and transcriptomic methods, we identified a panel of novel cysteine-rich peptides, roselptides (rT1–rT8), which range from 27 to 39 residues with six conserved cysteine residues. The 27-residue roselptide rT1 contains a cysteine spacing and amino acid sequence that is different from the squash knottin-type elastase inhibitor. NMR analysis demonstrated that roselptide rT1 adopts a cystine-knot fold. Transcriptome analyses suggested that roselptides are bioprocessed by asparagine endopeptidases from a three-domain precursor. The cystine-knot structure of roselptide rT1 confers its high resistance against degradation by endopeptidases, 0.2 N HCl, and human serum. Roseltide rT1 was shown to inhibit human neutrophil elastase using enzymatic and pull-down assays. Additionally, roselptide rT1 ameliorates neutrophil elastase-stimulated cAMP accumulation *in vitro*. Taken together, our findings demonstrate that roselptide rT1 is a novel knottin-type neutrophil elastase inhibitor with therapeutic potential for neutrophil elastase associated diseases.

*Hibiscus sabdariffa*, commonly known as roselle or red sorrel, is a shrub belonging to the Malvaceae family. The roselle plant is native to Central and West Africa and is cultivated extensively in subtropical and tropical regions, such as in India and Southeast Asia<sup>1</sup>. Its rose-color petals are used in drinks as food coloring and its fruit is used to make fruit preservatives, syrups, and jams. As a traditional medicine, roselle is primarily used as a remedy for coughs<sup>2</sup> but is also used as anti-hypertensive<sup>3,4</sup>, anti-hyperlipidemia<sup>5,6</sup>, anti-nociceptive, anti-pyretic, anti-inflammatory<sup>7</sup>, and diuretic agents<sup>8,9</sup>. Previous studies have shown that the roselle calyx extract contains proteinase inhibitors against angiotensin I converting enzyme, elastase, trypsin, and alpha-chymotrypsin<sup>10</sup>; however, the chemical nature of these proteinase inhibitors remains unexplored. To date, the known bioactive compounds of *Hibiscus sabdariffa* are limited to metabolites, including anthocyanins, flavonoids, saponins, tannins, phenols, glycosides, and alkaloids<sup>11–13</sup>.

Plant-derived peptides represent a promising group of natural products in drug discovery, as they fulfill the neglected chemical space between small-molecule metabolites (<1 kDa) and proteins (>8 kDa); however, these peptides have not received much attention as putative active compounds in medicinal plants and in drug development. Within the chemical space of 2–8 kDa, cysteine-rich peptides (CRPs), which possess multiple disulfide bridges to enhance both structural and physical stabilities, fulfill the criteria of putative bioactive peptides in medicinal plants<sup>14</sup>. Cysteine-rich peptides are classified into families primarily based on their cysteine motifs<sup>15–17</sup>. However, plant knottins are characterized structurally by their cystine-knot arrangement and their bioactivity as inhibitors. In particular, some knottins function as proteinase inhibitors against carboxypeptidase<sup>18</sup>, trypsin<sup>19,20</sup>, amylase<sup>21–23</sup>, and elastase<sup>24,25</sup>. Elastase inhibitors are of therapeutic interest since human neutrophil elastase is involved in several inflammatory diseases, including chronic obstructive pulmonary diseases (COPD), asthma, and cystic fibrosis<sup>26</sup>. Elastases are a class of serine proteinases that enzymatically degrade insoluble, highly

School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, 637551, Singapore. \*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to J.P.T. (email: JPTam@ntu.edu.sg)

cross-linked elastins. Serine proteinases have been reported to cleave and activate proteinase-activated receptors (PARs), a family of G protein-coupled receptors (GPCRs)<sup>27</sup>. Neutrophil elastase is recognized as a biased agonist of PAR2 and causes inflammation and pain<sup>28,29</sup>, as well as cough exaggeration<sup>30</sup>. Unlike trypsin inhibitors, elastase inhibitors appear to be a rarity. The only plant knottin-type elastase inhibitor was isolated from the squash family of *Mormodica charantia*<sup>24,25</sup>. The full length sequence MCEI-IV was identified in 1995, but three of its N-terminal truncated analogs MCEI-I to MCEI-III was first isolated in 1989. All four elastase inhibitors have been shown to inhibit porcine pancreatic elastases; however, an understanding of their inhibitory effect on human neutrophil elastases remains unknown.

In a mass-spectrometry-driven profiling program for CRPs in medicinal plants, we found a cluster of CRPs in the range of 2 to 4 kDa from the aqueous extracts of roselle plants. Herein, we report the identification and characterization of a knottin-type neutrophil elastase inhibitor (KNEI), roselptide rT1, from the medicinal plant *Hibiscus sabdariffa* of the Malvaceae family. Using a combination of proteomic and transcriptomic methods, we identified a panel of cysteine-rich peptides, collectively named roselptides (rT1-rT8). Transcriptomic analysis demonstrated that roselptides are bioprocessed from a three-domain precursor with Asn at the bioprocessing site to yield a mature roselptide. The prototypic and smallest member of roselptides, the 27-residue roselptide rT1, was shown to be a human neutrophil elastase inhibitor. Roselptide rT1 has a cystine-knot disulfide connectivity with a cysteine spacing that differs from the squash knottin-type elastase inhibitors. Taken together, our findings report the discovery and characterization of roselptide rT1, a novel plant-derived knottin-type neutrophil elastase inhibitor.

## Materials and Methods

**Materials.** All chemicals and solvents, unless otherwise mentioned, were purchased from Sigma-Aldrich, USA and ThermoFisher Scientific, USA. pGlosensor-20F vector was purchased from Promega, SG. pCMV6-XL5 encoding PAR2 receptor (NM 005242.3) was purchased from Origene, USA. Anti-PAR2 antibody (SAM11) Alexa Fluor 647 was purchased from Santa Cruz Biotechnology, USA.

**Plant materials.** *Hibiscus sabdariffa* were collected from the Nanyang Community Herb Garden, Nanyang Technological University, Singapore (courtesy of Mr. Ng Kim Chuan). The authenticity of samples was determined taxonomically by Lee, S. and Lam, H.J. of the Singapore Botanic Gardens and voucher specimens were deposited at the Singapore Herbarium in Singapore Botanic Gardens (code number: SING 2015-144). Dried calyces of *Hibiscus sabdariffa* were purchased from Hung Soon Medical Trading Pte Ltd, Singapore.

**Screening and profiling.** Fresh plant parts of *Hibiscus sabdariffa* were extracted with water for 15 min at room temperature in 1:10 ratio. The aqueous extract was vortexed vigorously and centrifuged at 16,000 × g for 5 min at 4°C and subjected to flash chromatography by C18 solid phase extraction (SPE) columns (Waters, USA). The fractions were eluted with 60% ethanol/0.01% trifluoroacetic acid (TFA) and analyzed by matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF MS) (AB SCIEX 4700 MALDI-TOF/TOF).

**Scale-up extraction and purification of Roselptide rT1.** Dried calyces (1 kg) of *Hibiscus sabdariffa* were extracted for 15 min with water and centrifuged at 9,000 rpm for 10 min at 4°C (Beckman Coulter, USA) and the supernatant was filtered through 1 μm pore size glass fiber filter paper (Sartorius, Singapore). The filtrate was then loaded onto a C18 flash column (Grace Davison, US) and eluted with 60% ethanol/0.01% TFA. The eluted fractions were then loaded onto an SP Sepharose resin column (GE Healthcare, UK), eluted with 1 M NaCl (pH 3.0), and followed by ultrafiltration (ViVaflow 200, 2000 MWCO hydrostat). Further purification was performed by reversed-phase high performance liquid chromatography (RP-HPLC) (Shimadzu, Japan). A linear gradient of mobile phase A (0.05% TFA/H<sub>2</sub>O) and mobile phase B (0.05% TFA/ACN) was used on the C18 column (250 × 22 mm, 5 μm, 300 Å) (Grace Davison, US). MALDI-TOF MS was used to identify the presence of roselptide rT1 in the eluted fractions. The eluted fractions were lyophilized for storage at room temperature.

**S-reduction and S-alkylation.** Purified roselptide rT1 (1 mg/mL) was S-reduced by 10 mM dithiothreitol (DTT) in ammonium bicarbonate buffer (25 mM) pH 8 at 37°C for 30 min, followed by S-alkylation with 60 mM of iodoacetamide (IAM) at 37°C for 45 min. MALDI-TOF MS was used to confirm the mass shift after S-reduction and S-alkylation.

**De novo peptide sequencing.** S-alkylated roselptide rT1 (1 mg/mL) were digested with trypsin or chymotrypsin in 5:1 (v/v) ratio in ammonium bicarbonate buffer (25 mM), pH 8 at 37°C for 10 min. The digested peptide fragments were then analyzed by MALDI-TOF MS followed by MS/MS (AB SCIEX 4700 MALDI-TOF/TOF). *De novo* peptide sequencing was performed using the *b*-ions and *y*-ions.

**Total RNA isolation and next generation transcriptome sequencing.** RNA isolation from fresh *Hibiscus sabdariffa* calyces was performed based on the protocol by Djami-Tchatchou and Straker (2011) using CTAB extraction buffer (2% cetyltrimethylammonium bromide, 2% polyvinylpyrrolidone, 100 mM Tris-HCl (pH 8.0), 2 mM EDTA, 2 M NaCl, 2% 2-mercapthoethanol)<sup>31</sup>. RNA library construction was performed using 1 μg of total RNA (RIN value >7.0) by Illumina TruSeq mRNA Sample Prep kit (Illumina, Inc., San Diego, CA, USA). Briefly, poly-A containing mRNA molecules were purified using poly-T-attached magnetic beads. Following purification, mRNA fragmentation was performed using divalent cations under elevated temperature. RNA fragments were reverse-transcribed into first strand cDNA using SuperScript II reverse transcriptase (Invitrogen) and random primers, followed by second strand cDNA synthesis using DNA Polymerase I and RNase H. These cDNA fragments were subjected to end repair process, the addition of a single 'A' base, and ligation of the

indexing adapters. The products were then purified and enriched using PCR to create the final cDNA library. The libraries were quantified using qPCR according to the qPCR Quantification Protocol Guide (KAPA Library Quantification kits for Illumina Sequencing platforms) and qualified using the TapeStation D1000 ScreenTape (Agilent Technologies, Waldbronn, Germany). Indexed libraries were sequenced using the HiSeq2500 platform (Illumina, San Diego, USA) by MacroGen Inc. (Korea).

**Peptide mapping using a peptidomic approach method.** Identification of putative rosetide sequences and confirmation of sequence of rosetide rT1 were performed using methods described by Serra et al (2016)<sup>32</sup>. One-pot reduction and alkylation were performed on the fractionated peptides. The peptide fractions were subjected to 30 mM dithiothreitol (DTT) and 60 mM bromoethylamine (BrEA) in 0.2 M Tris-HCl buffer (pH 8.6) at 55 °C for 60 min and the reaction was quenched using HCl. The reduced alkylated peptide samples were desalted using a C18 Sep-pack column (Waters, USA) and dried using SpeedVac (without heating). After re-dissolving the peptide solutions in H<sub>2</sub>O, they were analyzed using LC-MS/MS performed using an Orbitrap Elite mass spectrometer (Thermo Scientific Inc., Bremen, Germany) coupled with a Dionex UltiMate 3000 UHPLC system (Thermo Scientific Inc., Bremen, Germany). Samples were sprayed using a Michrom's Thermo CaptiveSpray nanoelectrospray ion source (Bruker-Michrom Inc, Auburn, USA) and separation was performed using a reverse phase Acclaim PepMap RSL column (75 μm ID × 15 cm, 2 μm particles; Thermo Scientific). The mobile phase was 0.1% formic acid (FA) as eluent A and 90% ACN 0.1% FA as eluent B, with a flow rate of 0.3 μL/min. A 60 min gradient was used for the elution as follows: 3% B for 1 min, 3–35% B over 47 min, 35–50% B over 4 min, 50–80% B over 6 s, 80% for 78 s; then, it was reverted to the initial state over 6 s and maintained for 6.5 min.

The Thermo Scientific Orbitrap Elite mass spectrometer was set to positive ion mode using LTQ Tune Plus software (Thermo Scientific Inc., Bremen, Germany) for data acquisition, alternating between a Full FT-MS (350–3000 m/z, resolution 60,000, with 1 μscan per spectrum) and a FT-MS/MS scan applying 65, 80, and 95 ms ETD activation times (150–2000 m/z, resolution 30,000, with 2 μscan averaged per MS/MS spectrum). The 3 most intense precursors with a charge >2+ were isolated with a 2 Da mass isolation window and fragmented. The automatic gain control (AGC) for Full MS and MS<sup>2</sup> was set to  $1 \times 10^6$  and the reagent AGC was  $5 \times 10^5$ . Data analysis was performed using PEAKS studio (version 7.0, Bioinformatics Solutions, Waterloo, Canada), where 10 ppm MS and 0.05 Da MS/MS tolerances were applied. A false discovery rate of 0.1% was applied to accept the sequences.

**Acid and proteolytic stability of rosetide rT1.** *Acid stability.* 0.1 M of purified rosetide rT1 was dissolved in 0.2 M HCl or phosphate buffered saline (PBS) (control) and incubated at 37 °C. Samples were collected at various time points (0, 15, 30, 45, 60 and 120 min).

*Enzymatic stability.* 0.1 M of purified rosetide rT1 was dissolved in 100 mM Tris buffer (pH 8) and incubated with trypsin (Catalog number: T1426, Roche Applied Science, US) in a 50:1 (w/v) ratio at 37 °C. 0.1 M of purified rosetide rT1 was dissolved in 0.2 M HCl and incubated with pepsin (Catalog number: 03117901001, Roche Applied Science, US) in a 50:1 (w/v) ratio at 37 °C. Trypsin substrate and DALK (KRPPGFSPL-NH<sub>2</sub>) were used as controls, respectively. Samples were collected at various time points (0, 15, 30, 45, 60 and 120 min).

*Human serum stability.* 0.1 M of purified rosetide rT1 was prepared in 25% human serum in DMEM medium without phenol red. The test samples were incubated at 37 °C. DALK was used as a positive control. Samples were collected at various time points (0, 24 and 48 h). The collected samples were subjected to protein precipitation with 100% ethanol and centrifugation at 180,000 g for 5 min at 4 °C. The supernatant was collected for analysis.

*Analysis for stability assays.* All collected samples from various stability assays were analyzed by RP-UHPLC with a linear gradient of mobile phase A (0.05% TFA/H<sub>2</sub>O) and mobile phase B (0.05% TFA/ACN) on C18 column (3.6 μm particle size; 2.1 mm internal diameter; 10 cm length) (Phenomenex, US). The resulting peaks were collected and identified by MALDI-TOF MS. The results were expressed as percentage of initial concentration using the peak area of the UHPLC profile.

**Human neutrophil elastase inhibition assay.** Human neutrophil elastase (HNE) activity was determined by measuring the release of p-nitroanilide at 405 nm at 37 °C using N-Methoxysuccinyl-Ala-Ala-Pro-Val p-nitroanilide as substrates. Purified rosetide rT1 was incubated with 1.75 U/mL HNE and 0.6 mM substrate in 100 mM Tris buffer (pH 8.0) at 37 °C; absorbance was measured continuously for 1 h. A synthetic elastase inhibitor, N-methoxysuccinyl-Ala-Ala-Pro-Val-chloromethyl ketone, was used as a positive control. The results were presented as normalized initial velocity.

**Peptide biotinylation and pull-down assay.** Purified rosetide rT1 was biotinylated with EZ-Link NHS-LC-biotin (Thermo Fisher Scientific, US) in 100 mM phosphate buffer at a pH of 7.8. Biotinylation was carried out at room temperature for 2 h and the biotinylated peptide was then identified and purified by MALDI-TOF MS and RP-HPLC.

Pull-down assay was performed using NeutrAvidin UltraLink Resin (Thermo Fisher Scientific, US). Briefly, the resin was washed with PBS three times and incubated with biotinylated rosetide rT1 or biotin (control) at room temperature with rotation for 1 h. The resin was washed again with excess PBS six times before incubation with HNE overnight at 4 °C with rotation. The resin was washed with excess PBS for six times and dissociation was performed by the addition of 6x loading dye with 2-mercaptoethanol and was heated at 85 °C for 10 min. The samples were then resolved in 12% SDS-PAGE for 1.5 h. Silver staining was performed for protein visualization.

**Cell culture.** Huh7 (human liver carcinoma cells), A549 (human lung adenocarcinoma epithelial cells) cells, and Chinese Hamster Ovary-K1 cells (CHO-K1) were cultured in Dulbecco's Modified Eagle's Medium

(DMEM)/Ham's F12 containing 15 mM HEPES and L-glutamine and supplemented with 10% fetal bovine serum, 100 U/mL of penicillin, and streptomycin. The cells were grown in a 5% CO<sub>2</sub> humidified incubator at 37 °C. CHO-K1 cells were transfected with pGlosensor-20F plasmid by electroporation and selected using 500 µg/mL hygromycin. CHO-K1 cells stably expressing the pGlosensor-20F construct (cAMP-CHO) were then transfected with pCMV6-XL5 encoding PAR2 receptor (NM 005242.3) and selected using 500 µg/mL G418. Stable cell lines co-expressing both pGlosensor-20F and PAR2 receptor constructs (PAR2-cAMP-CHO) were maintained with 500 µg/mL hygromycin and 500 µg/mL G418.

**Flow cytometry analysis.** PAR2-cAMP-CHO and cAMP-CHO cells were harvested and collected by centrifugation at 500 g for 5 min at 4 °C. The cell pellet was stained with the anti-PAR2 antibody (SAM11) Alexa Fluor 647 (1:100, Santa Cruz Biotechnology, USA) in serum-containing medium for 30 min on ice. The pellet was then washed three times and samples were analyzed by flow cytometry. 10,000 cells were analyzed using the BD LSRII Fortessa X-20 flow cytometer.

**Immunofluorescence analysis.** PAR2-cAMP-CHO and cAMP-CHO cells were seeded on an 8-well chamber slide (ibidi, Germany). The slides were washed gently with PBS, fixed in 4% paraformaldehyde for 10 min, and permeabilized in PBS containing 0.25% triton X-100 for 15 min. The slides were blocked in PBS containing 3% BSA for 1 h and then incubated with anti-PAR2 antibody (SAM11) Alexa Fluor 647 (1:50, Santa Cruz Biotechnology, USA) for 1 h. After PBS washing, the slides were mounted with Fluoroshield containing DAPI (Sigma, USA) and observed under a Zeiss LSM 710 confocal microscope (ZEISS, Germany).

**Glosensor cAMP assay.** Glosensor cAMP assays were performed according to the manufacturer's instructions. Briefly, PAR2-cAMP-CHO cells were grown until they reached confluency in a white-walled, clear-bottom 96 well plate. The culture medium was replaced with 6% (v/v) Glosensor cAMP substrate in CO<sub>2</sub>-independent medium in the dark at room temperature for 1 h. Luminescent intensity was measured using a microplate reader (Tecan Infinite 200 Pro, Switzerland); the plate was pre-read for 30 s to establish basal luminescent level. Following treatment, luminescent levels were continuously monitored for 20 min. The results are presented as fold change relative to basal luminescent levels and were quantified using the area under the curve.

**Cell viability assay.** Cell viability was measured using 3-(4,5-dimethylthiazolyl)-2,5-diphenyltetrazolium bromide (MTT) dye reduction assay. Briefly, cells were treated with rosetide rT1 or 0.1% triton X-100 (positive control) for 24 h. MTT (final concentration 0.5 mg/mL) was added and incubated for 3 h at 37 °C. Dimethyl sulfoxide was then added to dissolve the insoluble formazan crystal. The absorbance was measured at 550 nm using a microplate reader (Tecan Infinite 200 Pro, Switzerland).

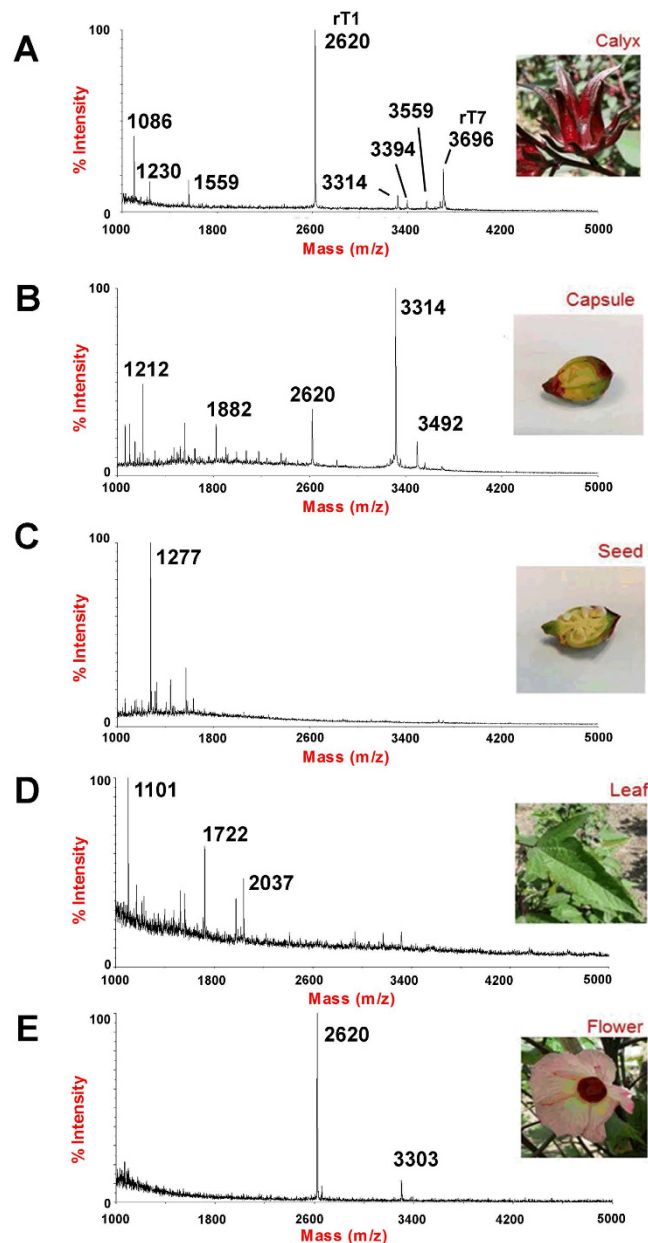
**NMR spectroscopy and structure determination of rosetide rT1.** A sample of rosetide rT1 for NMR spectroscopy was prepared by dissolving the lyophilized peptide in water containing 5% D<sub>2</sub>O at a final peptide concentration of 1.5 mM. For H/D exchange NMR experiment, the sample was dissolved in solution with 100% D<sub>2</sub>O immediately before the experiment. All NMR spectra were collected at a sample temperature of 298 K on a Bruker AVANCE II 600 MHz NMR spectrometer equipped with four RF channels and a 5 mm z-gradient TCI cryoprobe. Phase-sensitive two-dimensional <sup>1</sup>H, <sup>1</sup>H-TOCSY and NOESY spectra were recorded with a spectral width of 12 ppm. For water suppression, excitation sculpting with gradients was applied to all NMR experiments. TOCSY and NOESY spectra were obtained with mixing times of 80 ms and 200 ms respectively. The proton chemical shifts were referenced to external sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS). All measurements were recorded with 2048 complex data points and zero-filled to 2048 × 512 data matrices. Time domain data in both dimensions were multiplied by a 90°-shifted squared sine bell window function prior to Fourier transformation. Baseline correction was applied with a fifth order polynomial. NMR data were acquired and processed by TopSpin (Bruker BioSpin). The NMR spectra were processed with NMRpipe<sup>33</sup>. Sequence-specific assignments were achieved with 2D TOCSY and NOESY and NOEs were performed using SPARKY<sup>34</sup>. Distance restraints were derived based on the intensities of NOE cross peaks, which were divided to three classes: strong, 1 < d ≤ 1.8; medium, 1.8 < d ≤ 3.4; weak, 3.4 < d ≤ 5. Hydrogen bond restraints were determined based on H/D exchange 1D NMR experiment, in which amide protons exchanged with solvent deuterium for 24 hours in 298 K. The hydrogen bond restraints were defined as: N-O, 0.8~3.3 Å; HN-O, 0.6~2.2 Å. Dihedral angle restraints were derived from the <sup>3</sup>J<sub>HN-Hα</sub> coupling constant measured in 1D <sup>1</sup>H NMR spectrum. The backbone Φ angle was considered between -100° to -160° if the coupling constant was larger than 8 Hz. Three-dimensional structures were reconstructed using CNSsolve 1.3<sup>35</sup>. The 6 cysteines were assumed to form disulfide bonds in structure calculation. Structures were displayed with Chimera<sup>36</sup> and Pymol<sup>37</sup> and validated with the online server PDBsum<sup>38</sup>.

**In silico modeling.** The *in silico* docking was performed using automatic protein-protein docking server ClusPro Version 2.0<sup>39,40</sup>. Both the NMR structure of rosetide rT1 and the crystal structure of HNE (PDB entry: 1HNE)<sup>41</sup> were uploaded to the server. It uses rigid body docking protocol and the model was generated based on electrostatic potentials.

**Statistical analyses.** Statistical comparisons were performed using GraphPad Version 6.0d (USA). Data were analyzed using one-way analysis of variance (ANOVA) followed by Newman-Keuls *post hoc* tests. Data were expressed as mean ± S.E.M and P < 0.05 was considered to be statistically significant.

## Results

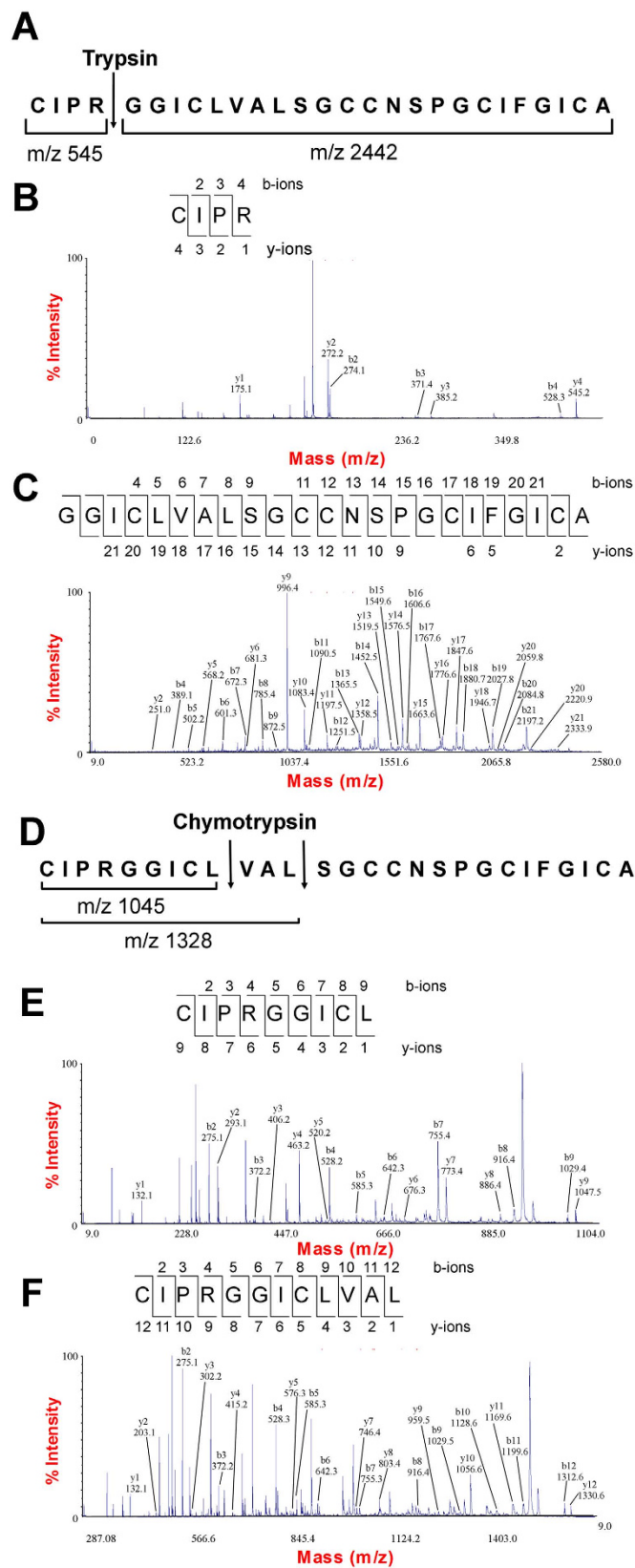
**Isolation of rosetide rT1 from *Hibiscus sabdariffa*.** Mass spectrometry profiling of the aqueous extracts of *Hibiscus sabdariffa* revealed the presence of a cluster of strong signals in the mass range of 2–4 kDa



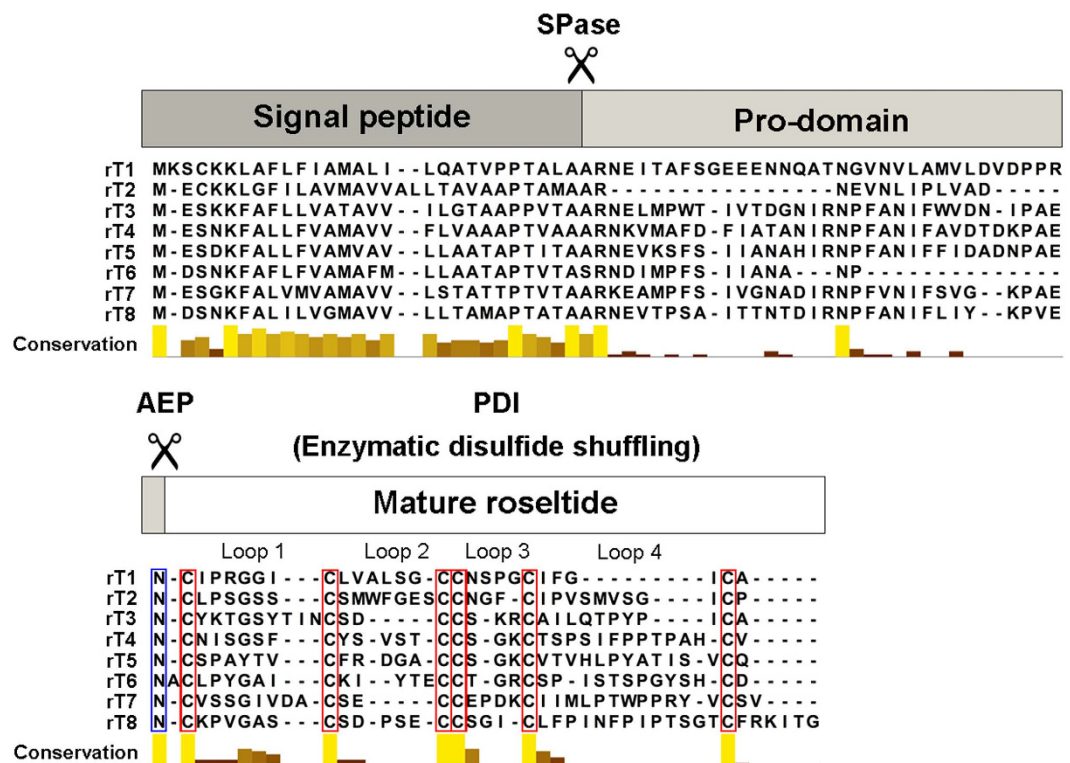
**Figure 1.** MS profiles of different plant parts of *Hibiscus sabdariffa*. (A) calyces, (B) capsule, (C) seed, (D) leaves, and (E) flower of *Hibiscus sabdariffa* were collected and profiled using MALDI-TOF MS.

in the calyces, capsules and flowers (Fig. 1). We focused on one of the strongest signals of this cluster, the 2620 Da-peak which was found in the calyces, capsules and flower extracts. The 2620 Da-peak was shown to be a CRP with six cysteine residues based on a mass increase of 348 Da after S-reduction by dithiothreitol and S-alkylation by iodoacetamide (Supplementary S1). To characterize this CRP, a scale-up aqueous extraction was carried out using the calyces of *Hibiscus sabdariffa*. The crude aqueous extract was fractionated by C18 reversed-phase and strong cation-exchange flash chromatography, followed by ultrafiltration using a membrane with molecular weight cut-off of 2000 Da. The CRP-enriched fraction was further purified by RP-HPLC (Supplementary data S2), and the 2620-Da CRP was designated roseltide rT1. 6–8 mg of purified roseltide rT1 was obtained per Kg of dried calyces.

To determine the amino acid sequence of roseltide rT1, the purified roseltide rT1 was fully S-reduced and S-alkylated followed by digestion with either trypsin or chymotrypsin. The digested peptide fragments were analyzed by MALDI-TOF MS, followed by MS/MS sequencing. Analysis using the *b*-ions and *y*-ions revealed that roseltide rT1 is a 27-residue peptide with six cysteine residues (Fig. 2). The amino acid sequence of rT1 was confirmed by transcriptomic analysis. Transcriptome analyses showed that roseltide rT1 was biosynthesized as a 90-residue precursor with three domains: a 28-residue N-terminal signal peptide, a 35-residue pro-domain and a 27-residue C-terminal mature peptide (Fig. 3). Using the asparaginyl endopeptidase cleavage site and the cysteine spacing pattern of roseltide rT1 in its transcriptomic database search, additional seven putative roseltide



**Figure 2.** *De novo* sequencing of roseltide rT1. (A) *S*-alkylated roseltide rT1 was digested with trypsin, resulting in two tryptic fragments with the *m/z* of 545 and 2442; (B) MS/MS spectra of 545 Da fragment; (C) MS/MS spectra of 2442 Da fragment; (D) *S*-alkylated roseltide rT1 was digested with chymotrypsin at two sites, resulting in fragments with the *m/z* of 1045 and 1328. The third peptide fragment of *m/z* 1640 could not be detected.; (E) MS/MS spectra of 1045 Da fragment; (F) MS/MS spectra of 1328 Da fragment.



**Figure 3.** Roseltide transcripts from *Hibiscus sabdariffa*. The histogram depicts conservation among the putative amino acid sequences of roselptides by amino acid property grouping as determined by Jalview software. AEP: asparagine endopeptidase; PDI: protein disulfide isomerase; SPase: signal peptidase.

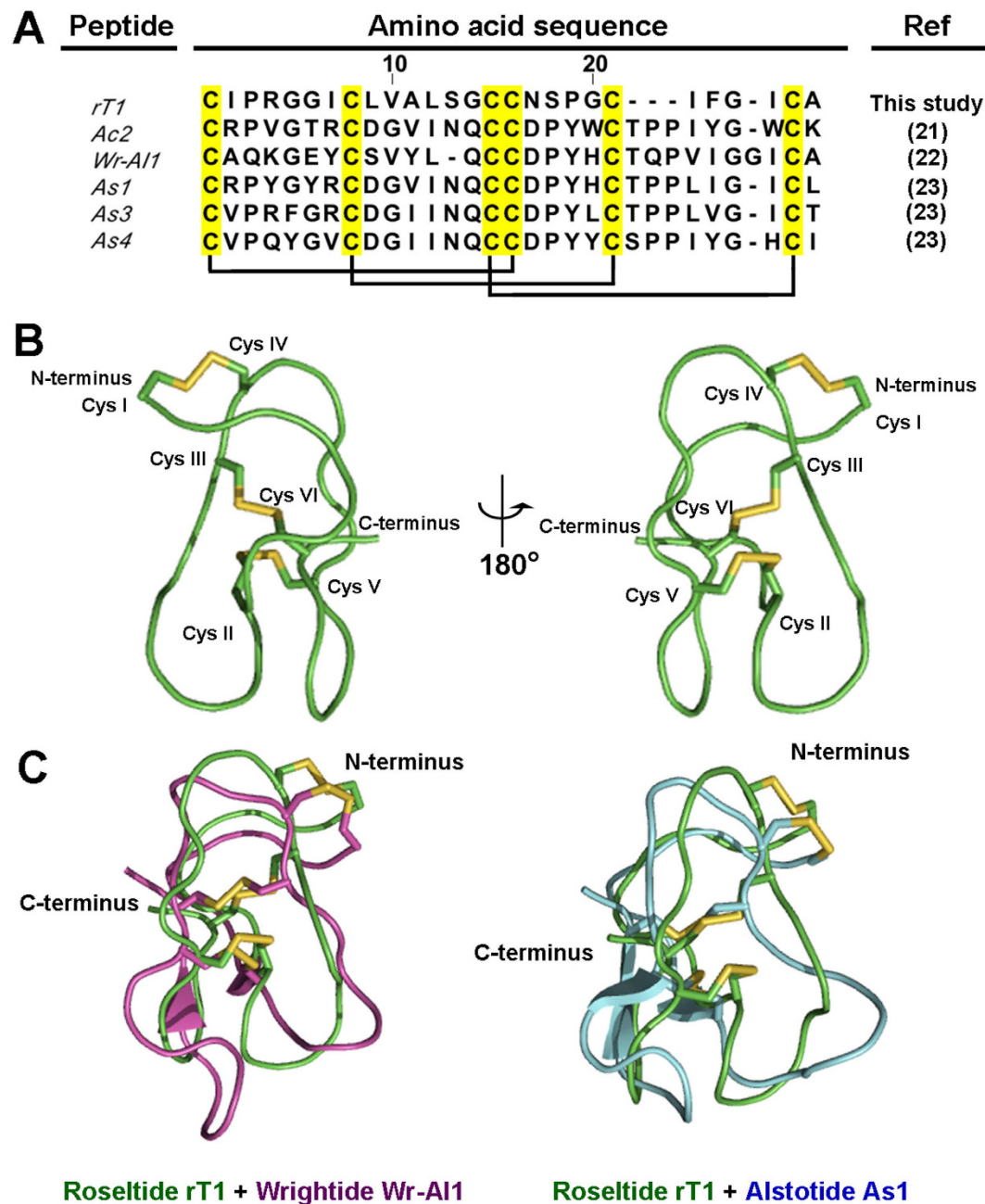
Roseltide	Amino acid sequence	Calculated MW	Expression level observed
rT1	-CIPRGGI---CLVALSG--CCNSPGCIFG-----ICA-----	2620	mRNA, protein
rT2	-CLPSGSS---CSMWFGESCCNGF-CIPVSMVSG---ICP-----	3281	mRNA
rT3	-CYKTGSYTI NCSD-----CCS-KRCAILQTPYP---ICA-----	3298	mRNA
rT4	-CNISGSF---CYS-VST-CCS-GKCTSPSIFPPTPAH-CV-----	3376	mRNA
rT5	-CSPAYTV---CFR-DGA-CCS-GKCVTVHLPYATIS-VCQ-----	3475	mRNA
rT6	ACLPGYAI---CKI--YTECC-TGRGSP-ISTSPGYSH-CD-----	3492	mRNA
rT7	-CVSSGIVDA-CSE-----CCEPDKCIIMLPWPPRY-VCSV-----	3696	mRNA, protein
rT8	-CKPVGAS---CSD-PSE-CCSGI-CLFPINFP IPTSGT CFRKITG	4028	mRNA

**Table 1.** Putative amino acid sequences of roselptides.

sequences (rT2-rT8) were identified (Fig. 3 & Table 1). Using a high-throughput peptidomic method for peptide sequencing developed by our laboratory<sup>32</sup>, we identified the presence of rT7 in the aqueous extract of *Hibiscus sabdariffa* calyces (Supplementary data S3).

**NMR structure of roselptide rT1.** The cysteine spacing of all eight roselptides contains a CC motif which provides a clue to their putative disulfide connectivity as a cystine-knot. Using peptide mapping, NMR, and X-ray crystallography, our laboratory has previously demonstrated in cystine-knot  $\alpha$ -amylase inhibitors (CKAI) (Fig. 4A) that the six-Cys-containing CRPs having a CC motif of CIII and CIV such as roselptides are often arranged in a cystine-knot motif (Cys I-IV, Cys II-V and Cys III-VI)<sup>21-23</sup>.

To characterize the structural fold of roselptide rT1, the solution structure of rT1 was determined using 2D <sup>1</sup>H, <sup>1</sup>H-TOCSY and NOESY NMR spectra. The sequential assignment was done based on the NOE cross peaks between HN<sub>i</sub> and H $\alpha_{i-1}$  as well as the other side chain protons of residue i-1 (Supplementary data S4 and S5). When performing the assignment, we compared the NOESY spectrum and the TOCSY spectrum to differentiate the intra-residue and inter-residue NOEs. The amide proton of residue i should have NOE cross peaks with the side chain protons of the residue i-1. The pattern of the peaks in TOCSY of each amide proton stripe should correspond to its specific residue. Based on these strategies, the sequential assignment was completed. More than 95% of the peaks in the NOESY spectrum were assigned. H/D exchange NMR experiment indicated that the residue C26 and G5, I25 and I22 as well as I2 and C16 should be involved in hydrogen bonds, which was consistent



**Figure 4.** (A) Sequence alignment of CKAI with roselptide rT1; allotide Ac2 from *Allamanda cathartica*; wrightide Wr-A11 from *Wrightia religiosa*; alstotides As1, As3 and As4 from *Alstonia scholaris* (B) Solution structure of roselptide rT1; (C) Superimposition of roselptide rT1 (green) on wrightide Wr-A11 (PDB entry 2MAU) (purple) and alstotides As1 (PDB entry 2MM6) (blue).

with the NOE cross peaks between the amide protons of them accordingly. This observation further confirms the sequential assignment. Twenty structures with the lowest energy among 100 structures were generated by CNSsolve 1.3<sup>35</sup> (Supplementary data S6). The 20 structures are highly converged, of which the backbone RMSD and the heavy atom RMSD are  $0.71 \pm 0.14 \text{ \AA}$  and  $1.19 \pm 0.26 \text{ \AA}$  respectively (Table 2). The ensemble of the 20 best structures has been deposited to Protein Data Bank (PDB) with the accession number 5GSF. The chemical shifts have been submitted to Biological Magnetic Resonance Bank (BMRB), of which the accession number is 26874 (Supplementary data S7).

The proton chemical shifts were uploaded to the online server CSI 3.0 (<http://csi3.wishartlab.com/cgi-bin/index.php>) to predict the secondary structure, indicating that only residue F23-C26 might be edge  $\beta$  strand while the other residues were coil<sup>42</sup>. The structure of rT1 generated by simulated annealing contains no  $\alpha$  helix or  $\beta$  strand. The two prolines both adopt trans form, which are supported by the NOE cross peak between  $H\delta_i$  and  $HN_{i+1}$  of the proline and the previous residue respectively. The disulfide bonds Cys8-Cys21 and Cys15-Cys26 cross in the center of rT1. The disulfide bond Cys1-Cys16 makes the N terminus bend (Fig. 4B).



Experimental Restraints and Structural Statistics of 20 Lowest-Energy Structures of GB4242 among the 100 Structures Generated by CNSsolve 1.3		
NMR Distance Restraints		
Intra-Residue NOE	431	
Sequential NOE ( $ i-j =1$ )	97	
Medium-Range NOE ( $1 <  i-j  \leq 5$ )	127	
Long-Range NOE ( $ i-j  > 5$ )	74	
Hydrogen Bonds	133	
Dihedral Angle Restraints	6	
Dihedral Angle Restraints		
Structural Statistics (27 residues, C <sup>1</sup> -A <sup>27</sup> )		
Violations per Structure		
NOE Violation (Å)	0.025 ± 0.002	
Torsion Angles (°)	0.064 ± 0.04	
Ramachandran Plot Region (27 residues)		
Residues in Most Favored Regions	5	27.8%
Residues in Additional Allowed Regions	10	55.6%
Residues in Generously Allowed Regions	3	16.7%
Residues in Disallowed Regions	0	0%
Number of End-Residues (excl. Gly and Pro)		
Number of Glycine Residues	2	
Number of Proline Residues	5	
Mean RMSD from the Average Coordinates (27 residues, C <sup>1</sup> -A <sup>27</sup> )		
Backbone Atoms(Å)	0.71 ± 0.14	
Heavy Atoms(Å)	1.19 ± 0.26	

**Table 2. Parameters and restrains of structure calculation of roseltide rT1.**

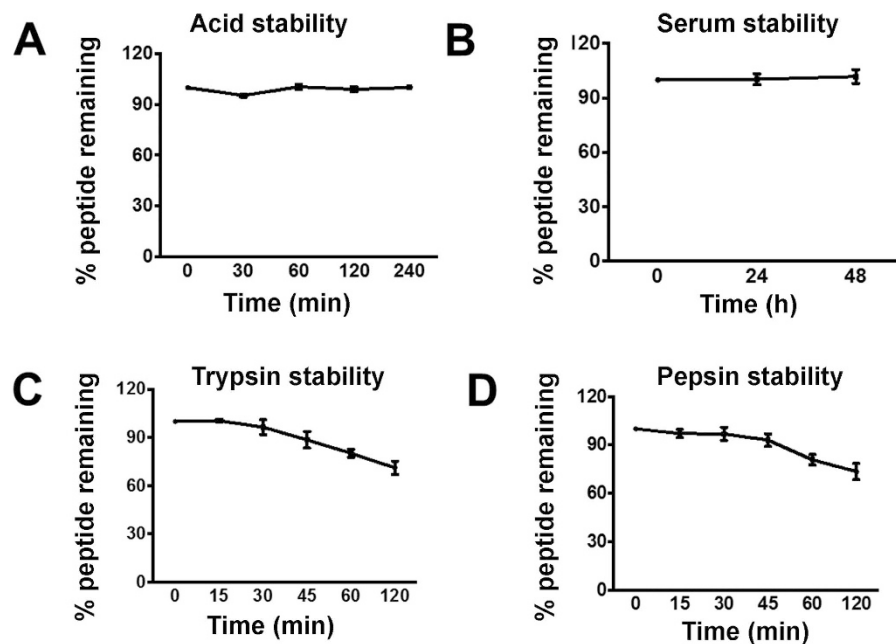
Further evidence supporting the Cys1-Cys16, Cys8-Cys21 and Cys15-Cys26 disulfide linkages were observed from the H $\beta$ -H $\beta$  NOE cross peaks (Supplementary data S8). To confirm the disulfide connectivity, all the six cysteines were assumed reduced in the simulated annealing by CNSsolve 1.3<sup>35</sup>. The 100 structures were highly converged. The backbone RMSD and the heavy atom RMSD of the 20 best structures are  $0.73 \pm 0.27$  Å and  $1.25 \pm 0.26$  Å respectively. The averaged energy of the 20 best structures is similar to that of the 20 best structures with the 3 disulfide bonds imposed (Cys1-Cys16, Cys8-Cys21 and Cys15-Cys26). The structures generated with and without the disulfide bond imposed are very similar, except for the sidechains of the six cysteines (Supplementary data S6). Moreover, another 14 disulfide patterns were assumed in structure calculations respectively. The average energy of the 20 best structures for each combination was compared with the one with the pattern: Cys1-Cys16, Cys8-Cys21 and Cys15-Cys26. The disulfide pattern of: Cys1-Cys16, Cys8-Cys21 and Cys15-26 has the lowest average energy among the 15 combinations (Supplementary data S9). These results strongly suggest that the disulfide connectivity of roseltide rT1 exists in a cystine-knot motif.

Protein tertiary structure comparison was conducted using the SuperPose software Version 1.0<sup>43</sup> for the wrightide Wr-AI1 (PDB entry: 2MAU)<sup>22</sup> and alstotide As1 (PDB entry: 2MM6)<sup>23</sup> which displays similar cystine-knot fold (Fig. 4C). The RMSD values between the superimposed structure of wrightide Wr-AI1 and roseltide rT1 were 0.515 Å and 1.036 Å for all C $\alpha$  and heavy atoms, respectively. The RMSD values between the superimposed structures of alstotide As1 and roseltide rT1 were 0.534 Å and 1.003 Å for all C $\alpha$  and heavy atoms, respectively. Based on the electrostatic potential surface of roseltide rT1, a negatively-charged region was observed. This is created by the side-chain of Arg4 residue which is positioned outwards.

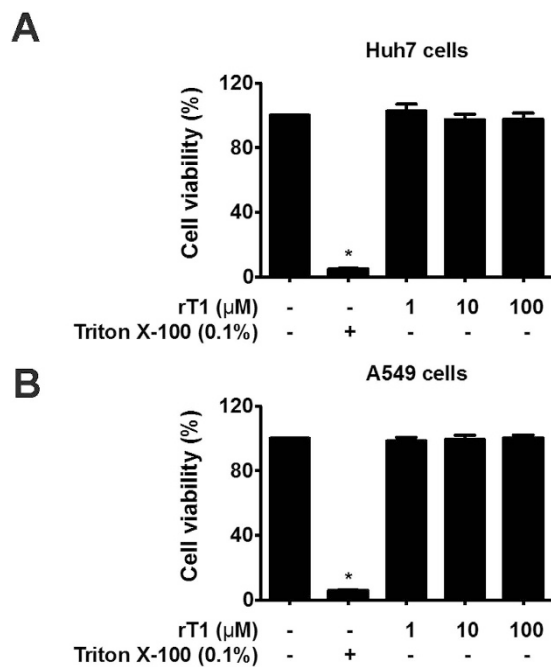
**Acid and proteolytic stability of roseltide rT1.** Intact CRPs are highly cross-linked by multiple disulfide bridges which confers their high stability. To determine the acid and proteolytic stability of roseltide rT1, roseltide rT1 was incubated in 0.2 N HCl, or with proteinases (trypsin or pepsin), or in 25% human serum. The results demonstrated that roseltide rT1 was resistant against acid, proteinase and human serum-mediated degradation (Fig. 5).

**Roseltide rT1 is not cytotoxic.** To determine the cytotoxicity of roseltide rT1, cell viability was measured by MTT assay. Treatment with roseltide rT1 of concentrations up to 100  $\mu$ M for 24 h did not affect the viability of Huh7 (human liver carcinoma cells) and A549 (human lung adenocarcinoma epithelial cells) cells (Fig. 6).

**Roseltide rT1 inhibited human neutrophil elastase.** Previous studies have shown that the extract of *Hibiscus sabdariffa* calyces contains proteinase inhibitors against elastase<sup>10</sup>, however, the active compounds are yet to be reported. To characterize the biological activities of roseltide rT1, its effects on the enzymatic activities of human neutrophil elastase was examined. As shown in Fig. 7A, roseltide rT1 inhibited the enzymatic activities of human neutrophil elastase in a dose-dependent manner, with an IC<sub>50</sub> of 0.47  $\mu$ M. This was comparable to the activity of a synthetic elastase inhibitor (MeOSu-AAPV-CMK). Roseltide rT1 was also screened against trypsin and porcine pancreatic elastase without observable inhibitory effects up to 10  $\mu$ M (data not shown).

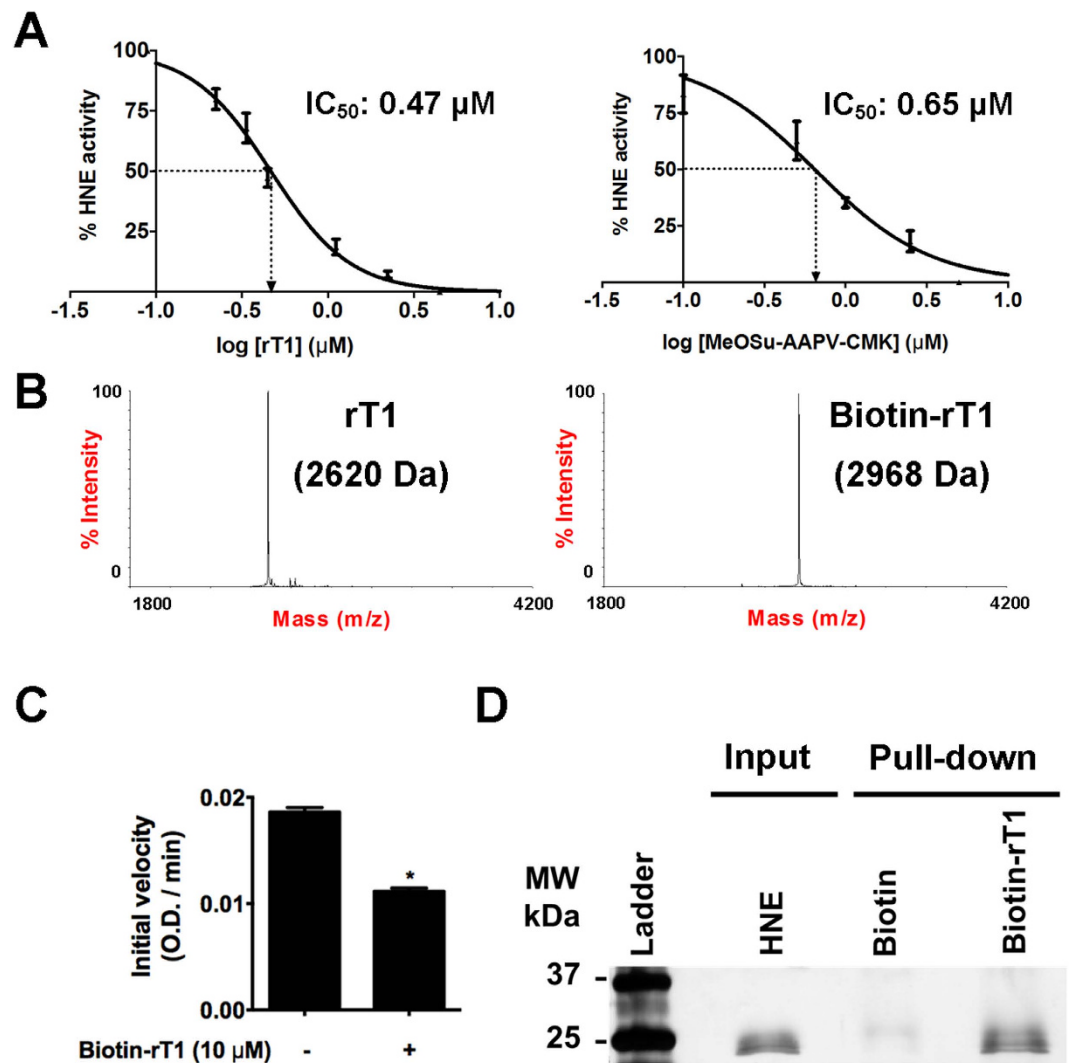


**Figure 5.** (A) Acid, (B) human serum, (C) trypsin, and (D) pepsin stability of rosetide rT1. All results were expressed as mean  $\pm$  S.E.M. (n = 3).



**Figure 6.** Rosetide rT1 does not show cytotoxic activities. Effects of rosetide rT1 on (A) Huh7 and (B) A549 cells. All results were expressed as mean  $\pm$  S.E.M. (n = 3). \* $P$  < 0.05 compared to control group.

**Rosetide rT1 showed protein interactions to human neutrophil elastase.** Pull-down assays were performed to determine the binding of rosetide rT1 to human neutrophil elastase. Purified biotin-rT1 was characterized by MALDI-TOF MS and RP-HPLC (Fig. 7B). Following biotinylation, the MS profile showed an increase in mass from 2620 Da to 2960 Da. Biotin-rT1 inhibits human neutrophil elastase (Fig. 7C) and was able to pull-down human neutrophil elastase with a band of approximately 25 kDa (Fig. 7D). Pull-down assay was also performed using porcine pancreatic elastase however, biotin-rT1 does not show protein interactions with porcine pancreatic elastase (Supplementary data S10).

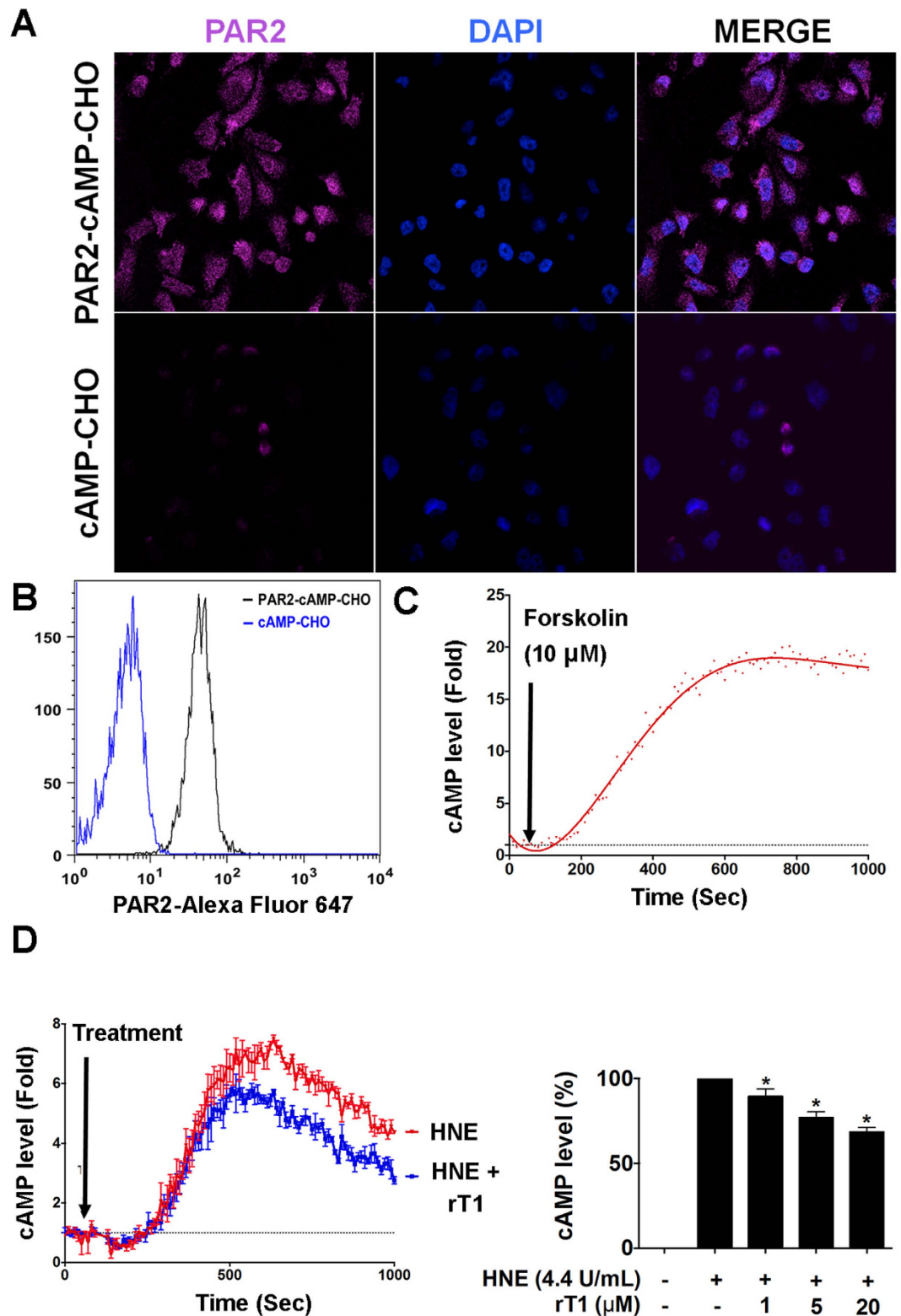


**Figure 7. The effects of rosetide rT1 on human neutrophil elastase.** (A) The effects of different concentrations of rosetide rT1 or synthetic elastase inhibitor, N-methoxysuccinyl-Ala-Ala-Pro-Val-chloromethyl ketone, on human neutrophil elastase (HNE) activity was measured at 405 nm at 37 °C using N-methoxysuccinyl-Ala-Ala-Pro-Val p-nitroanilide as a substrate. All results were expressed as mean  $\pm$  S.E.M. ( $n = 3$ ); (B) MS spectra showed biotinylation of rosetide rT1; (C) The effects of biotin-rT1 on HNE activity. All results were expressed as mean  $\pm$  S.E.M. ( $n = 3$ ). \* $P < 0.05$  compared to control; (D) Silver-stained SDS-PAGE of the pull-down complex between HNE and biotin-rT1. The left-most lane shows a protein marker (Bio-rad, US); HNE-only lane: purified HNE only; biotin lane: purified HNE incubated with biotin and NeutrAvidin resin (control), biotin-rT1 lane: purified HNE incubated with biotin-rT1 and NeutrAvidin resin. The full-length gel image is provided in Supplementary data S11.

**Rosetide rT1 inhibited neutrophil elastase-induced cAMP accumulation.** PAR2 is a GPCR responsible for the cellular effects of neutrophil elastase. To demonstrate the cellular effects of neutrophil elastase inhibition *in vitro*, the effects of rosetide rT1 on neutrophil elastase-stimulated cAMP accumulation were evaluated. In this study, CHO-K1 cells stably transfected with Glosensor cAMP biosensor (cAMP-CHO cells) was used to provide direct and real-time measurement of cAMP accumulation in live cells. The cAMP-CHO cells were further overexpressed with the gene of PAR2 receptor (PAR2-cAMP-CHO) cells. The expression levels of PAR2 were confirmed by flow cytometry and confocal microscopy while the function of cAMP biosensor was assessed using a cAMP activator, forskolin ( $10 \mu\text{M}$ ) (Fig. 8A–C). Similar to previous findings, human neutrophil elastase stimulated intracellular cAMP accumulation in PAR2-cAMP-CHO cells. Co-incubation of human neutrophil elastase with rosetide rT1 significantly suppressed cAMP accumulation (Fig. 8D).

## Discussion

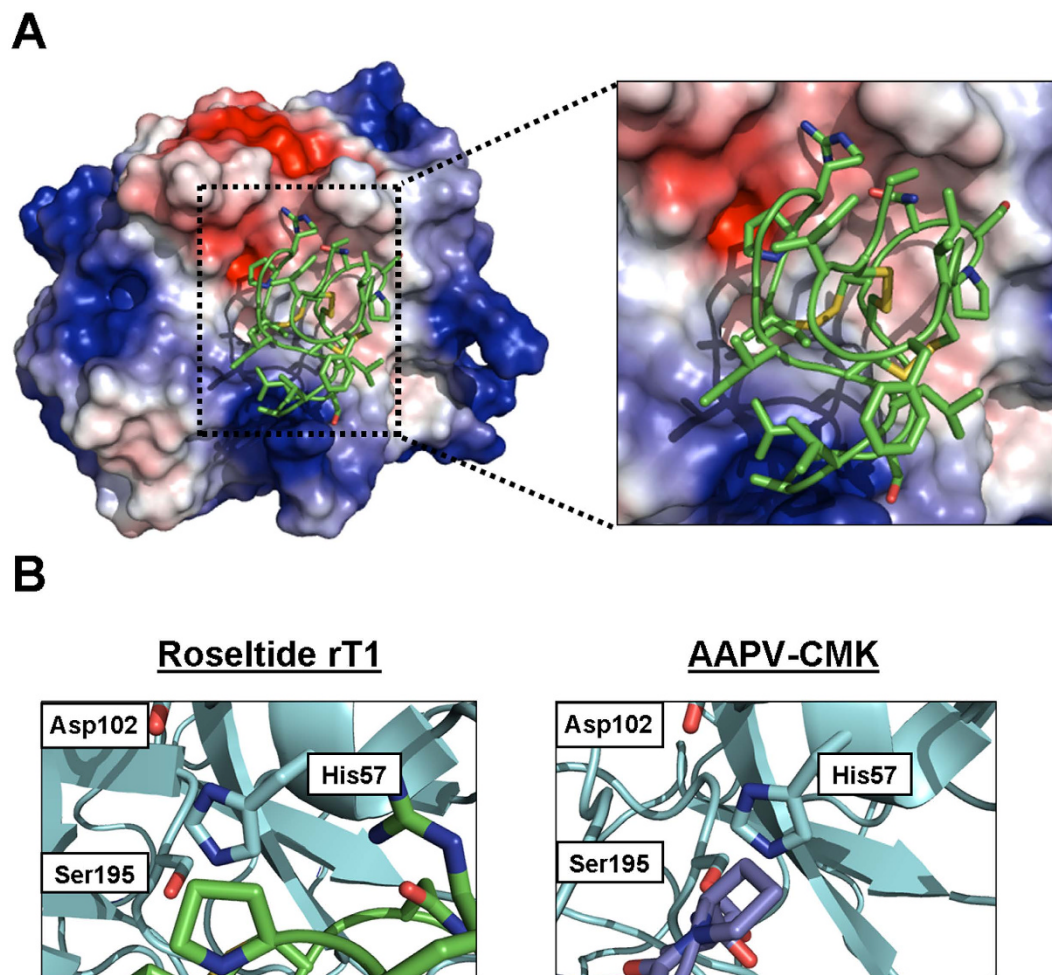
Plants are rich in proteinase inhibitors that serve as defense mechanisms against pests and pathogens<sup>44</sup>. This study employed a peptidomic approach to identify a panel of CRPs, rosetides (rT1–rT8), from the medicinal plant *Hibiscus sabdariffa*. Rosetide rT1, the smallest member of rosetide family, was identified as a knottin-type



**Figure 8.** Roseltide rT1 ameliorates human neutrophil elastase (HNE)-stimulated cAMP accumulation in CHO-K1 cells co-expressed with PAR2 receptor and Glosensor cAMP biosensor constructs. Comparison of PAR2 receptor expressions in cAMP-CHO cells using (A) confocal microscopy and (B) flow cytometry; (C) Effects of cAMP activator Forskolin (10 μM) on cAMP accumulation in PAR2-cAMP-CHO cells; (D) Effects of HNE with or without rosetide rT1 on cAMP accumulation in PAR2-cAMP-CHO cells. All results were expressed as mean ± S.E.M. (n = 3). \* $P < 0.05$  compared to HNE-treated group.

Roseltide	Amino acid sequence	Reference
rT1	----CIPRGGICLVALSGC---CNSPGCIFGICA	This study
MCEI-I	---RICPLIWMECKRDS-CLAQCI---CVDGHCG	24
MCEI-II	--ERICPLIWMECKRDS-CLAQCI---CVDGHCG	24
MCEI-III	-EERICPLIWMECKRDS-CLAQCI---CVDGHCG	24
MCEI-IV	EEERICPLIWMECKRDS-CLAQCI---CVDGHCG	25

**Table 3.** Sequence comparison between Roseltide rT1 and MCEI-I-IV.



**Figure 9.** (A) Modeling the interaction between roseltide rT1 and the human neutrophil elastase (PDB entry: 1HNE) using the ClusPro Version 2.0 server. Blue: Negatively-charged, Red: Positively-charged and White: Neutral.; (B) Interactions between roseltide rT1 and peptide chloromethyl ketone inhibitor (AAPV-CMK) (PDB entry: 1HNE) with the catalytic triad of human neutrophil elastase formed by His57, Asp102, and Ser195.

inhibitor against human neutrophil elastase. There has been only one reported knottin-type porcine elastase inhibitor (MCEI-IV) with three N-terminal truncated analogs (MCEI-I to MCEI-III) within the chemical space of 2–6 kDa from *Momodica charantia* of the squash family<sup>24,25</sup>. However, plant knottin-type elastase inhibitors are likely more prevalent than previously believed.

Transcriptomic and proteomic analyses showed that roseltides (rT1–rT8) contain six Cys residues with a cysteine spacing pattern of  $CX_{6-8}CX_{2-7}CCX_{3-4}CX_{4-13}C$  and four inter-cysteine loops (Fig. 3). Roseltides range in length from 27 to 39 amino acid residues with highly variable amino acid sequences in loops 2 and 4; for example, loop 4 is particularly variable in length, ranging from 4 residues in rT1 to 13 residues in rT8. It is important to note that the highly variable loop lengths of roseltides could be used as a grafting strategy for the engineering new bioactive peptides using roseltides as a scaffold<sup>45,46</sup>.

Transcriptome analyses demonstrated that all roseltide precursors possess a three-domain precursor of ER signal peptide, pro-domain, and mature domain. Sequence analyses showed that these precursors share similar signal peptide sequences and bioprocessing sites. This bioprocessing pathway was also observed in the

precursor sequence of cystine-knot  $\alpha$ -amylase inhibitors<sup>21–23</sup>. The N-terminal cleavage site for the mature peptide of rosetide rT1 occurs between Asn and Cys, suggesting that removal of the pro-domain is mediated by an asparaginyl endopeptidase, a vacuolar processing enzyme<sup>47</sup>.

The smallest and most abundant member of rosetlides, rosetide rT1, is a single, positively charged peptide with six cysteine residues and a cysteine spacing pattern of CX<sub>6</sub>CX<sub>6</sub>CCX<sub>4</sub>CX<sub>4</sub>CX, a Cys-spacing motif which is prevalent in CKAI<sup>21–23</sup>. Of the 27 amino acids, seven are Ile/Val/Leu and a total of 85% of its residues are hydrophobic. The disulfide bridges of rosetide rT1, arranged in a cystine-knot motif, confer its high resistance against acid and proteolytic degradation. Many plant knottins have been identified as proteinase inhibitors, including the potato carboxypeptidase inhibitor (PCI), *Momordica cochinchinensis* trypsin inhibitor II (MCoTI-II), and *Ecballium elaterium* trypsin inhibitor II (EETI-II)<sup>48–50</sup>. Elastases are a class of serine proteinases that cleave C-terminal of small hydrophobic amino acids including Gly, Val, and Ala. Neutrophil elastase is one of the most potent elastases and has gained increased attention for its involvement in airway inflammatory diseases, particularly cystic fibrosis, asthma, COPD, and pulmonary emphysema<sup>26</sup>. Its underlying mechanisms have been attributed to the tissue-damaging, hypersecretory, and pro-inflammatory effects of neutrophil elastases<sup>51</sup>. Our results suggest that rosetide rT1 binds to human neutrophil elastase and inhibits its proteolytic activities. In comparison to the knottin-type elastase inhibitors from the squash inhibitors family (MCEI-I to MCEI-IV), there is no sequence homology to rosetide rT1 other than classification as a cystine-knot peptide<sup>24,25</sup> (Table 3).

Previous studies have demonstrated that the S1 pocket of elastases is narrowed by bulky side-chains, resulting in the recognition of small or medium-sized aliphatic P1 residues<sup>52</sup>. Compared to porcine pancreatic elastase (PPE), the S1 pocket of HNE is more flexible and is negatively charged<sup>52,53</sup>, which may explain the recognition of Lys or Arg at the P1 position of the inhibitor, leading to the selectivity of biotin-rT1 towards HNE (supplementary data S10). The S2 pocket of HNE is hydrophobic and can accept peptidyl inhibitors with Pro residue at its P2 position<sup>54–56</sup>. The S3 pocket of HNE is also hydrophobic and can tolerate small lipophilic residues, including Val and Ile, at the P3 position<sup>54</sup>; thus, it is likely that Ile-Pro-Arg in the inter-cysteine loop 1 of rosetide rT1 plays an important role in the interaction between the S1, S2, and S3 pockets of HNE. *In silico* modeling using the ClusPro server suggests that the Pro residue of rosetide rT1 interacts with the catalytic triad of HNE formed by His57, Asp102, and Ser195, which is similar to the known inhibitor AAPA-CMK (PDB entry: 1HNE) (Fig. 9)<sup>39–41,57</sup>. Additional studies are necessary to investigate key amino acid residues and the molecular mechanisms of the inhibitory activities of rosetide rT1 against neutrophil elastase using point mutation and X-ray crystallography.

Neutrophil elastase is a biased agonist of PAR2<sup>28</sup>. It cleaves the extracellular domain of PAR2 at Ser<sup>67</sup> and Val<sup>68</sup> to trigger its receptor activation and induces cAMP accumulation<sup>58</sup>. A previous study reported that PAR2 activation leads to TRPV1-triggered cough exaggeration in guinea pigs<sup>30</sup>. PAR2 activation by neutrophil elastase has also been shown to trigger airway hyperresponsiveness<sup>59</sup>, as well as TRPV4-mediated pain and inflammation<sup>58</sup>. Therefore, PAR2 activation plays a significant role in the deleterious effects of neutrophil elastase on inflammation. To investigate the cellular effects of neutrophil elastase inhibition by rosetide rT1, the intracellular level of cAMP was monitored in CHO-K1 cells co-expressed with PAR2 and cAMP biosensor constructs. The results showed that rosetide rT1 ameliorates neutrophil elastase-stimulated cAMP accumulation in PAR2-overexpressed cells. Hence, these findings suggest that metabolically-stable rosetide rT1 has therapeutic potential for neutrophil elastase-driven diseases, particularly in airway inflammatory diseases.

In conclusion, this study reports the identification of rosetlides from *Hibiscus sabdariffa* and characterizes rosetide rT1 as a novel KNEI, 27 years after the report of the first squash knottin-type elastase inhibitor. Our findings highlight the therapeutic potential of rosetide rT1, a novel KNEI of the Malvaceae family, for neutrophil elastase-driven diseases, including COPD, asthma, and cystic fibrosis.

## References

1. Ali, B. H., Al Wabel, N. & Blunden, G. Phytochemical, pharmacological and toxicological aspects of *Hibiscus sabdariffa* L.: a review. *Phytother. Res.* **19**, 369–375 (2005).
2. Wilson, F. D. & Menzel, M. Y. Kenaf (*Hibiscus cannabinus*), roselle (*Hibiscus sabdariffa*). *Econ. Bot.* **18**, 80–91 (1964).
3. Herrera-Arellano, A. *et al.* Clinical effects produced by a standardized herbal medicinal product of *Hibiscus sabdariffa* on patients with hypertension. A randomized, double-blind, lisinopril-controlled clinical trial. *Planta Med.* **73**, 6–12 (2007).
4. Ajay, M., Chai, H. J., Mustafa, A. M., Gilani, A. H. & Mustafa, M. R. Mechanisms of the anti-hypertensive effect of *Hibiscus sabdariffa* L. calyces. *J. Ethnopharmacol.* **109**, 388–393 (2007).
5. Chen, C. C. *et al.* *Hibiscus sabdariffa* extract inhibits the development of atherosclerosis in cholesterol-fed rabbits. *J. Agric. Food Chem.* **51**, 5472–5477 (2003).
6. Ajiboye, T. O. *et al.* *Hibiscus sabdariffa* calyx palliates insulin resistance, hyperglycemia, dyslipidemia and oxidative rout in fructose-induced metabolic syndrome rats. *J. Sci. Food Agric.* 1522–1531 (2015).
7. Ali, M. K., Ashraf, A., Biswas, N. N., Karmakar, U. K. & Afroz, S. Antinociceptive, anti-inflammatory and antidiarrheal activities of ethanolic calyx extract of *Hibiscus sabdariffa* Linn. (Malvaceae) in mice. *J. Chin. integr. med.* **9**, 626–631 (2011).
8. Jimenez-Ferrer, E. *et al.* Diuretic effect of compounds from *Hibiscus sabdariffa* by modulation of the aldosterone activity. *Planta Med.* **78**, 1893–1898 (2012).
9. Alarcon-Alonso, J. *et al.* Pharmacological characterization of the diuretic effect of *Hibiscus sabdariffa* Linn (Malvaceae) extract. *J. Ethnopharmacol.* **139**, 751–756 (2012).
10. Jonadet, M. *et al.* *In vitro* enzyme inhibitory and *in vivo* cardioprotective activities of hibiscus (*Hibiscus sabdariffa* L.). *J. Pharm. Belg.* **45**, 120–124 (1990).
11. Carvajal-Zarrabal, O. *et al.* *Hibiscus sabdariffa* L., roselle calyx, from ethnobotany to pharmacology. *J. Exp. Pharmacol.* **4**, 25–39 (2012).
12. Hirunpanich, V. *et al.* Hypocholesterolemic and antioxidant effects of aqueous extracts from the dried calyx of *Hibiscus sabdariffa* L. in hypercholesterolemic rats. *J. Ethnopharmacol.* **103**, 252–260 (2006).
13. Edema, M. O. & Alaga, T. O. Comparative evaluation of bioactive compounds in *Hibiscus sabdariffa* and *Syzygium samarangense* juice extracts. *Afr. Crop Sci. J.* **20**, 179–187 (2012).
14. Nguyen, G. K. T. *et al.* Discovery of linear cyclotides in monocot plant *Panicum laxum* of Poaceae family provides new insights into evolution and distribution of cyclotides in plants. *J. Biol. Chem.* **288**, 3370–3380 (2013).
15. Broekaert, W. F. *et al.* Antimicrobial peptides from plants. *Crit. Rev. Plant Sci.* **16**, 297–323 (1997).

16. Broekaert, W. F., Terras, F. R., Cammue, B. P. & Osborn, R. W. Plant defensins: novel antimicrobial peptides as components of the host defense system. *Plant Physiol.* **108**, 1353–1358 (1995).
17. García-Olmedo, F., Molina, A., Alamillo, J. M. & Rodríguez-Palenzuela, P. Plant defense peptides. *Biopolymers* **47**, 479–491 (1998).
18. van den Hooven, H. W. *et al.* Disulfide bond structure of the AVR9 elicitor of the fungal tomato pathogen *Cladosporium fulvum*: evidence for a cystine knot. *Biochemistry* **40**, 3458–3466 (2001).
19. Felizmenio-Quimio, M. E., Daly, N. L. & Craik, D. J. Circular proteins in plants solution structure of a novel macrocyclic trypsin inhibitor from *Momordica cochinchinensis*. *J. Biol. Chem.* **276**, 22875–22882 (2001).
20. Isaacs, N. W. Cystine knots. *Curr. Opin. Struct. Biol.* **5**, 391–395 (1995).
21. Nguyen, P. Q. *et al.* Allotides: Proline-rich cystine knot alpha-amylase inhibitors from *Allamanda cathartica*. *J. Nat. Prod.* **78**, 695–704 (2015).
22. Nguyen, P. Q. *et al.* Discovery and characterization of pseudocyclic cystine-knot alpha-amylase inhibitors with high resistance to heat and proteolytic degradation. *FEBS J.* **281**, 4351–4366 (2014).
23. Nguyen, P. Q. T. *et al.* Antiviral cystine knot  $\alpha$ -amylase inhibitors from *Alstonia scholaris*. *J. Biol. Chem.*, 31138–31150 (2015).
24. Hamato, N. *et al.* Trypsin and elastase inhibitors from bitter melon (*Momordica charantia* LINN.) seeds: Purification, amino acid sequences, and inhibitory activities of four new inhibitors. *J. Biochem.* **117**, 432–437 (1995).
25. Hara, S., Makino, J. & Ikenaka, T. Amino acid sequences and disulfide bridges of serine proteinase inhibitors from bitter melon (*Momordica charantia* LINN.) seeds. *J. Biochem.* **105**, 88–92 (1989).
26. Tsai, Y. F. & Hwang, T. L. Neutrophil elastase inhibitors: a patent review and potential applications for inflammatory lung diseases (2010–2014). *Expert Opin. Ther. Pat.* **25**, 1–14 (2015).
27. Ossovskaya, V. S. & Bunnett, N. W. Protease-activated receptors: contribution to physiology and disease. *Physiol. Rev.* **84**, 579–621 (2004).
28. Hollenberg, M. D. *et al.* Biased signalling and proteinase-activated receptors (PARs): targeting inflammatory disease. *Br. J. Pharmacol.* **171**, 1180–1194 (2014).
29. Schmidlin, F. & Bunnett, N. W. Protease-activated receptors: how proteases signal to cells. *Curr. Opin. Pharmacol.* **1**, 575–582 (2001).
30. Gatti, R. *et al.* Protease-activated receptor-2 activation exaggerates TRPV1-mediated cough in guinea pigs. *J. Appl. Physiol.* **101**, 506–511 (2006).
31. Djami-Tchatchou, A. T. & Straker, C. J. The isolation of high quality RNA from the fruit of avocado. *S. Afr. J. Bot.* **78**, 44–46 (2011).
32. Aida Serra, X. H., Nguyen, G. K., Nguyen, N. T., Sze, S. K. & Tam, J. P. A high-throughput peptidomic strategy to decipher the molecular diversity of cyclic cysteine-rich peptides. *Sci. Rep.* **6** (2016).
33. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
34. Goddard, T. & Kneller, D. SPARKY 3 University of California, San Francisco 15 (2004).
35. Brünger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54**, 905–921 (1998).
36. Huang, C. C., Couch, G. S., Pettersen, E. F. & Ferrin, T. E. Chimera: an extensible molecular modeling application constructed using standard components. *In Pacific symposium on biocomputing.* 1519–1523 (1996).
37. DeLano, W. L. The PyMOL molecular graphics system (2002).
38. Laskowski, R. A., Chistyakov, V. V. & Thornton, J. M. PDBsum more: new summaries and analyses of the known 3D structures of proteins and nucleic acids. *Nucleic Acids Res.* **33**, D266–D268 (2005).
39. Comeau, S. R., Gatchell, D. W., Vajda, S. & Camacho, C. J. ClusPro: a fully automated algorithm for protein–protein docking. *Nucleic Acids Res.* **32**, W96–W99 (2004).
40. Comeau, S. R., Gatchell, D. W., Vajda, S. & Camacho, C. J. ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics* **20**, 45–50 (2004).
41. Navia, M. A. *et al.* Structure of human neutrophil elastase in complex with a peptide chloromethyl ketone inhibitor at 1.84-Å resolution. *Proc. Natl. Acad. Sci. USA* **86**, 7–11 (1989).
42. Hafsa, N. E., Arndt, D. & Wishart, D. S. CSI 3.0: a web server for identifying secondary and super-secondary structure in proteins using NMR chemical shifts. *Nucleic Acids Res.*, gkv494 (2015).
43. Maiti, R., Van Domselaar, G. H., Zhang, H. & Wishart, D. S. SuperPose: a simple server for sophisticated structural superposition. *Nucleic Acids Res.* **32**, W590–W594 (2004).
44. Ryan, C. A. Protease inhibitors in plants: genes for improving defenses against insects and pathogens. *Annu. rev. phytopathol.* **28**, 425–449 (1990).
45. Gracy, J. *et al.* KNOTTIN: the knottin or inhibitor cystine knot scaffold in 2007. *Nucleic Acids Res.* **36**, D314–319 (2008).
46. Wong, C. T. *et al.* Orally active peptidic bradykinin B1 receptor antagonists engineered from a cyclotide scaffold for inflammatory pain treatment. *Angew. Chem. Int. Ed.* **51**, 5620–5624 (2012).
47. Abe, Y. *et al.* Asparaginyl endopeptidase of jack bean seeds. Purification, characterization, and high utility in protein sequence analysis. *J. Biol. Chem.* **268**, 3525–3529 (1993).
48. Heitz, A. *et al.* Solution structure of the squash trypsin inhibitor MC $\alpha$ TI-II. A new family for cyclic knottins. *Biochemistry* **40**, 7973–7983 (2001).
49. Le Nguyen, D. *et al.* Molecular recognition between serine proteases and new bioactive microproteins with a knotted structure. *Biochimie* **72**, 431–435 (1990).
50. Rees, D. C. & Lipscomb, W. N. Refined crystal structure of the potato inhibitor complex of carboxypeptidase A at 2.5 Å resolution. *J. Mol. Biol.* **160**, 475–498 (1982).
51. Korkmaz, B., Horwitz, M. S., Jenne, D. E. & Gauthier, F. Neutrophil elastase, proteinase 3, and cathepsin G as therapeutic targets in human diseases. *Pharmacol. Rev.* **62**, 726–759 (2010).
52. Garcia-Fernandez, R. *et al.* Three-dimensional structure of a Kunitz-type inhibitor in complex with an elastase-like enzyme. *J. Biol. Chem.* (2015).
53. Hernandez Gonzalez, J. E., Garcia-Fernandez, R. & Valiente, P. A. Polar desolvation and position 226 of pancreatic and neutrophil elastases are crucial to their affinity for the kunitz-type inhibitors ShPI-1 and ShPI-1/K13L. *PLoS one* **10**, e0137787 (2015).
54. Cregge, R. J. *et al.* Inhibition of human neutrophil elastase. 4. design, synthesis, x-ray crystallographic analysis, and structure–activity relationships for a series of P2-modified, orally active peptidyl pentafluoroethyl ketones. *J. Med. Chem.* **41**, 2461–2480 (1998).
55. Nakajima, K., Powers, J. C., Ashe, B. M. & Zimmerman, M. Mapping the extended substrate binding site of cathepsin G and human leukocyte elastase. Studies with peptide substrates related to the alpha 1-protease inhibitor reactive site. *J. Biol. Chem.* **254**, 4027–4032 (1979).
56. Kasperkiewicz, P. *et al.* Design of ultrasensitive probes for human neutrophil elastase through hybrid combinatorial substrate library profiling. *Proc. Natl. Acad. Sci. USA* **111**, 2518–2523 (2014).
57. von Nussbaum, F. *et al.* Freezing the bioactive conformation to boost potency: The identification of BAY 85-8501, a Selective and Potent Inhibitor of Human Neutrophil Elastase for Pulmonary Diseases. *ChemMedChem* **10**, 1163–1173 (2015).
58. Zhao, P. *et al.* Neutrophil elastase activates PAR2 and TRPV4 to cause inflammation and pain. *J. Biol. Chem.* (2015).
59. Koga, H. *et al.* Inhibition of neutrophil elastase attenuates airway hyperresponsiveness and inflammation in a mouse model of secondary allergen challenge: neutrophil elastase inhibition attenuates allergic airway responses. *Respir. Res.* **24**, 1 (2013).

## Acknowledgements

This research was supported in part by the Competitive Research Grant by National Research Foundation in Singapore (NRF-CRP8-2011-05). We thank Dr. Aida Serra Maqueda for performing the peptide mapping and Professor Sze Siu Kwan for providing the proteomic technology platform.

## Author Contributions

S.L. and A.K. performed the experiments and wrote the manuscript. T.X. collected, analyzed and determined the NMR structure of roseltide rT1. G.K.T.N. performed the *de novo* sequencing of roseltide rT1. L.C.F. oversaw the analysis of NMR data. J.P.T. conceived the idea and edited the manuscript. All authors have read and approved the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Loo, S. *et al.* Identification and Characterization of Roseltide, a Knottin-type Neutrophil Elastase Inhibitor Derived from *Hibiscus sabdariffa*. *Sci. Rep.* **6**, 39401; doi: 10.1038/srep39401 (2016).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016