DOI 10.1002/pmic.200800569

RESEARCH ARTICLE

Isolation, purification and *de novo* sequencing of TBD-1, the first beta-defensin from leukocytes of reptiles

Christin Stegemann[1], Alexander Kolobov[2]Jr.*, Yulia F. Leonova[3], Daniel Knappe[1], Olga Shamova[4], Tatiana V. Ovchinnikova[3], Vladimir N. Kokryakov[2, 4] and Ralf Hoffmann[1]*

[1] Institute of Bioanalytical Chemistry, Center for Biotechnology and Biomedicine, Faculty of Chemistry and Mineralogy, Leipzig University, Leipzig, Germany
[2] St. Petersburg State University, Department of Biochemistry, Faculty of Biology and Soil Science, St. Petersburg, Russia
[3] Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russia
[4] Institute for Experimental Medicine RAMS, Department of General Pathology and Pathophysiology, St. Petersburg, Russia

A novel peptide with antimicrobial activity was isolated from leukocytes of the European pond turtle *Emys orbicularis* and purified to homogeneity by preparative gel electrophoresis followed by reversed phase chromatography. It was highly active *in vitro* against *Escherichia coli*, *Listeria monocytogenes*, methicillin-resistant *Staphylococcus aureus*, and *Candida albicans*. The isolated peptide was sequenced *de novo* by tandem mass spectrometry using both collision-induced and electron-transfer dissociation in combination with different chemical derivatization techniques. The 40-residue peptide, called TBD-1 (turtle β-defensin 1), represents the first defensin isolated from reptilian leukocytes. It contains three disulfide bonds and shows high structural similarities to β-defensins isolated from birds and mammals.

Received: July 6, 2008
Revised: September 3, 2008
Accepted: October 5, 2008

Keywords:
Antibacterial assay / Antimicrobial peptide / Mass spectrometry / Turtle

Correspondence: Professor Dr. Ralf Hoffmann, Institute of Bioanalytical Chemistry, Center for Biotechnology and Biomedicine, Faculty of Chemistry and Mineralogy, Leipzig University, Deutscher Platz 5, 04103 Leipzig, Germany
E-mail: Hoffmann@chemie.uni-leipzig.de
Fax: +49- 341-973-1339

Abbreviations: **AMA**, antimicrobial activity; **AMP**, antimicrobial peptide; **AU**, acid urea; **ETD**, electron-transfer dissociation; **HFBA**, heptafluorobutyric acid; **MIC**, minimal inhibitory concentration; **MRSA**, methicillin-resistant *Staphylococcus aureus*; **Qq**, double quadrupole; **SACA**, 2-sulfobenzoic acid cyclic anhydride; **SPITC**, 4-sulfophenyl isothiocyanate sodium salt monohydrate; **TBD-1**, turtle β-defensin 1

1 Introduction

Antimicrobial peptides (AMP) are gene encoded peptide antibiotics isolated from plants, animals and microbes ranging in size from 12 to more than a hundred amino acid residues [1]. A single animal produces a set of peptide antibiotics, each peptide often exhibiting a completely different biological activity spectrum [2]. They are produced by all higher organisms as an important component of their host defense, the innate immune system. Most AMP are positively charged

* Both authors contributed equally to this work.

and amphipathic [3]. The bacteria are killed by different mechanisms, such as formation of pores in the membrane, or the targeting of specific intracellular or membrane bound targets [4]. Whereas some AMP possess a broad activity spectrum against a wide range of microorganisms, such as Gram-positive and Gram-negative bacteria, fungi and sometimes even enveloped viruses and protozoa, others are only active against a specific group of microorganisms [3, 5].

AMP can be divided into several classes based on their amino acid composition and structure, such as anionic peptides, linear cationic alpha-helical peptides, cationic peptides enriched for specific amino acids, peptide fragments of larger proteins, and peptides containing cysteines forming intramolecular disulfide bonds [5]. Defensins, which were named for their function in the host defense, belong to the last family [3]. They are a superfamily of evolutionary related AMP in invertebrates and vertebrates with a characteristic beta-sheet-rich secondary structure containing three intramolecular disulfide-bridges. Defensins have been isolated from different cell types of diverse animals (Table 1) and are classified into alpha (α) or beta (β) defensins based on the sequence length and the disulfide bridging as well as theta (θ) defensins, which are cyclic and contain three disulfide bridges. A fourth class of defensins was found in insects and other invertebrates.

So far, more than 300 defensins have been identified (http://defensins.bii.a-star.edu.sg/). Beta-defensins are considered the oldest of the three defensin subfamilies found in vertebrates [12]. The α-defensins are probably derived from a beta-defensin ancestor, despite their low sequence similarities and different cysteine connectivities [13]. They diverged during evolution into the α- and β-defensins, which are both found in mammals and birds [12, 13]. In 2005 and 2006, two defensin-like AMP, named TEWP and pelovaterin, were isolated from the eggs of the marine turtle *Caretta caretta* [14] and the eggshell of the Chinese soft-shelled turtle *Pelodiscus sinensis* [15]. TEWP shows some common structural features of defensins, including six cysteine residues, but its disulfide connectivity was different from all previously known defensins. Due to the structural differences, the authors concluded that TEWP does not belong to the defensin family, but might be the first representative of a new sub-class of defensins, related to avian defensins, that may have emerged from reptilian defensins. Pelovaterin is a glycine-rich acidic 42 amino acid residue peptide with no sequence homology to other proteins or peptides. The 3-D structure stabilized by three intramolecular disulfide bridges, however, is very similar to β-defensins, which may explain its antimicrobial activity [16].

Here, we describe the biological activity and *de novo* sequencing of a novel antimicrobial peptide isolated from leukocytes of the European pond turtle *Emys orbicularis*. The peptide was digested with different enzymes, chemically modified to allow easier interpretation of the fragment spectra, and the sequence finally obtained by combining partial sequence information retrieved from MS/MS spectra relying on PSD, CID, and electron-transfer dissociation (ETD). To the best of our knowledge, this sequence represents the first beta-defensin isolated from the leukocytes of reptiles.

2 Materials and methods

ACN (LC-MS grade) was obtained from Biosolve B.V. (Valkenswaard, NL) and CHCA from Bruker Daltonik (Bremen, Germany). Porcine trypsin ("Sequencing Grade Modified Trypsin") was purchased from Promega (Madison, WI), sequence grade Lys-C from Roche Diagnostics (Roche Applied Science, Mannheim, Germany) and Arg-C from Serva Electrophoresis (Heidelberg, Germany). Pipette tips for sample preparation (C_{18}-ZipTip™) were from Millipore (Billerica, MA) and off-line emitters for ESI-MS from Proxeon A/S (Odense, Denmark). All other reagents and solvents were obtained from Fluka Chemie (Buchs, Switzerland) in the highest-grade analytical purity and used without further purification.

Table 1. Classification of defensins[a)]

Name	Characteristics
α-Defensins [6, 7, 8]	Isolated from mammals, 29 to 35 amino acid residues
	Cys1 and Cys2 near the N-terminus, spaced by one residue
	Cys5 and Cys6 adjacent near the C-terminus
	Three intramolecular disulfide bonds between Cys1-Cys6, Cys2-Cys4, and Cys3-Cys5
β-Defensins [6, 7, 8]	Isolated from mammals and birds, 38 to 42 amino acid residues
	Cys5 and Cys6 adjacent near the C-terminus
	Three intramolecular disulfide bonds between Cys1-Cys5, Cys2-Cys4, and Cys3-Cys6
θ-Defensins [9]	Isolated from rhesus monkey
	Macrocyclic peptide
	Three intramolecular disulfide bonds
Insect defensins [10, 11]	Sapecins or royalisins, 29 to 34 amino acid residues
	Three intramolecular disulfide bonds between Cys1-Cys4, Cys2-Cys5, and Cys3-Cys6

a) Cys residues are numbered from the N to the C terminus from 1 to 6.

2.1 Isolation of cationic peptides

Blood from three male and three female European pond turtles from the Astrakhan region of Russia was collected, anticoagulated with heparin (20 U/mL blood), and immediately incubated at 37°C for 2–3 h. The layer of leukocytes (buffy coat) was collected with a micropipette. Leukocytes were homogenized in 10 mL of 10% v/v aqueous acetic acid at 4°C for 18 h, and centrifuged (25 000 × *g,* 45 min). The supernatant was collected, and the pellet was dissolved in 10 mL of 10% v/v aqueous acetic acid at 4°C. After 18 h, the suspension was centrifuged (25 000 × *g,* 45 min) and the supernatant was collected. The pellet was dissolved for a second time and the complete extraction procedure was repeated again. All three supernatants were combined and ultrafiltered (Amicon YM-10 filter, Millipore). The flow-through was concentrated and desalted on an Amicon YM-1 filter (Millipore) using 10% v/v aqueous acetic acid.

2.2 Electrophoresis

Peptides were dissolved in 3 mol/L aqueous urea containing 5% acetic acid v/v and analyzed by electrophoresis using a polyacrylamide gel (T=12.5%, C = 2.5%) at 20 V for 60 min (Hoeffer, San Francisco, CA) and 5% aqueous acetic acid (pH 2.2) as electrode buffer (AU-PAGE) [17]. The same conditions were also applied on a preparative scale to purify the peptides using a Model 491 Prep Cell chamber (Bio-Rad Laboratories, Philadelphia, PA). Five-microliter fractions were collected at a flow rate of 20 mL/h using an SC fraction collector (Beckman Coulter, Fullerton, CA). Peptide purities were determined by SDS-PAGE (T = 16%, C = 3%) according to Schägger *et al.* [18].

Peptides and proteins were stained in the gels with 0.095% w/v CBB (Sigma Aldrich, St. Louis, MO) solution containing 25.6% v/v methanol and 14.3% v/v formaldehyde for 30 min and destained in 5% v/v aqueous acetic acid for 3 h. Alternatively, a more sensitive silver nitrate stain was then applied [19]. Briefly, gels were washed five times with 50% v/v aqueous methanol for 15 min each, and incubated in staining solution for 15 min, which was prepared freshly by adding a 20% w/v aqueous silver nitrate solution to 0.08% w/v aqueous sodium hydroxide containing 2% NH_3 until the mixture turned brown. Stained gels were washed in water five times for 3 min each and incubated in 0.5% w/v aqueous citric acid containing 0.02% v/v formaldehyde until the appropriate staining appeared. The gel was washed with water for 30 s and incubated for 30 min in 50% v/v aqueous methanol containing 10% v/v acetic acid to stop the staining.

2.3 RP-HPLC

The peptide samples obtained by preparative gel electrophoresis were further separated on a Vydac C_{18} column (4.6-mm id, 25-cm length, particle size 5 μm) using an HPLC System Gold equipped with a 166A UV detector (Beckman Coulter). Peptides were eluted by a linear gradient from 100% eluent A (10 mM aqueous tetrabutylammonium hydrogen sulfate) to 60% eluent B (ACN) at a flow rate of 1 mL/min for 60 min. Fractions containing the targeted peptide were further purified on an Alltech C_{18} column (4.6-mm id, 25-cm length, particle size 5 μm) using 0.13% v/v) heptafluorobutyric acid (HFBA) as ion pair reagent in water (eluent C) and ACN (eluent D). Peptides were eluted by a linear gradient from 0 to 60% eluent D in 60 min at a flow rate of 1 mL/min.

2.4 Antimicrobial activity

The antimicrobial activity of peptides was determined after AU-PAGE in the gel by an overlay assay [19]. Thus, the sample to be tested was split in four equal parts and loaded on four separate gels. After completion of the electrophoretic separation one gel was stained with CBB G-250, while the other three gels were washed in sodium phosphate buffer (10 mM, pH 7.4) three times for 15 min and overlaid in petri dishes with agarose gels (1% v/v agarose, 0.01 M sodium phosphate buffer, pH 7.4) containing 4x 10^5 CFU/mL of *Escherichia coli* (strain ML-35p), *Listeria monocytogenes* (strain EGD) or *Candida albicans* (strain 820). This overlay was incubated at 37°C for 3 h before 1% v/v agarose containing 6% w/v tryptic soy broth (Sigma Aldrich) was poured into the petri dishes and incubated again at 37°C for 16 h.

Fractions obtained after RP-HPLC were tested by a radial diffusion assay for their antimicrobial activity against *E. coli* (strain ML35p), *L. monocytogenes* (strain EGD) and methicillin resistant *S. aureus* (strain ATCC 33591) [19]. Five microliter of purified peptides and standards dissolved in 0.01% v/v acetic acid were added as duplicates in a serial dilution directly to the pores of the underlay agars containing sodium phosphate buffer (10 mM, pH 7.4), agarose (1% w/v), and trypticase soy broth powder (0.3 mg/mL, Sigma Aldrich). The minimal inhibitory concentrations (MIC) were determined for low-salt (10 mM sodium phosphate buffer, pH 7.4) and high-salt conditions (10 mM sodium phosphate buffer, 100 mM NaCl) in the underlay agars. The antimicrobial activity (AMA) was calculated as:

$$\mathrm{AMA[units]} = 10x[d_{\mathrm{pore}}\,(\mathrm{mm}) - d_{\mathrm{lysis\ zone}}\,(\mathrm{mm})] \quad (1)$$

The lysis zone was defined as the zone of agarose gel where growth of microorganisms was inhibited by antimicrobial peptides. Minimal inhibitory concentrations were estimated by plotting the AMA against the peptide concentration on a semi logarithmic scale. The intersections of the linear regression lines and the x-axis were given as MIC-values.

2.5 Hemolytic activity test

Ten microliters of heparin (5 U) was added to 1 mL of human blood and resuspended in 10 mL of PBS (20 mM sodium phosphate buffer, 150 mM sodium chloride, pH 7.4)

containing 4 mM EDTA and centrifuged (10 000 × *g*, 4 °C) for 10 min. The supernatant was discarded and the pellet was washed three times with PBS. The erythrocyte suspension was diluted with PBS to a final concentration of 2.8% v/v. Three microliters of the peptide solution (0.01% v/v aqueous acetic acid) was added to 27 μL of the erythrocyte suspension. The positive control contained 3 μL of 10% v/v aqueous triton X-100 and the negative control 3 μL of 0.01% v/v aqueous acetic acid instead of the peptide solution. All samples were incubated at 37°C in parallel and the reaction was stopped after 30 min by addition of 75 mL cold PBS buffer and centrifuged (1500 × *g*, 10 min, 4°C). Supernatants were transferred to a 96-well plate and the optical density at 540 nm was measured on a Multiscan MS microtiter plate reader (Labsystems, Helsinki, Finland). The hemolytic activity was calculated by the following equation:

$$\text{Hemolysis}[\%] = 100\text{x}\frac{\text{OD}_{540,\text{sample}} - \text{OD}_{540,\text{negativecontrol}}}{\text{OD}_{540,\text{positivecontrol}} - \text{OD}_{540,\text{negativecontrol}}} \quad (2)$$

2.6 MALDI-MS

Sample solution (0.5 μL) was mixed with 0.5 μL of CHCA (4 or 7 mg/mL; 50% v/v aqueous ACN containing 0.1% v/v TFA) on a stainless steel target and air-dried at room temperature. Mass analysis was performed on a MALDI-TOF/TOF-MS (4700 Proteomics Analyzer, Applied Biosystems Applera Deutschland, Darmstadt, Germany) in reflector mode at an acceleration voltage of 20 kV, 70% grid voltage and a delay of 1.277 ns. Mass spectra were recorded in positive and negative ion mode accumulating 2000 laser shots. MS/MS spectra were recorded with higher laser energy in positive ion-mode by accumulation of 3000 laser shots.

2.7 ESI-MS

The peptides were dried in vacuum, dissolved in 3 to 8 μL of 60% v/v aqueous ACN containing 1% v/v formic acid, transferred to a nanospray needle, and analyzed on an ESI-QqTOF-MS (QSTAR Pulsar I, Applied Biosystems Applera Deutschland) equipped with a static nanospray source (Protana, Odense, Denmark). All mass spectra were acquired in positive ion-mode using the following instrumental setup: ion source gas 1 was 6, spray voltage was 1100 V and curtain gas was 25. Fragment spectra relied on collision-induced dissociation using a CE from 30 to 60 V. Alternatively, the nanospray needle was coupled to an Esquire HCTultra PTM Discovery System equipped with a static nanospray source (Bruker Daltonik). Fragment spectra on this instrument were recorded in electron-transfer dissociation (ETD) mode. Thus, fluoranthene reactant anions were accumulated (2.5 ms) by negative chemical ionization in negative ion mode. The spray voltage was 680 V, the solvent was evaporated at 150°C, and the nitrogen stream was 4 L/min.

2.8 Modification of cysteine residues

Peptides dissolved in 0.01% aqueous acetic acid (9.5 μL) were mixed with the same volume of 50 mmol/L ammonium bicarbonate buffer (pH 7.8) and reduced by addition of 1 μL of aqueous 1,4-DTT (1 mol/L) at 95°C. After 5 min the solution was divided into two equal parts and 2 μL of either 40% w/v aqueous acrylamide or 27.5% v/v aqueous iodoacetamide solution was added to alkylate the thiol groups at 24°C for 30 min. The reduced and alkylated peptides were desalted on a C_{18} ZipTip™ according to the manufacturer's protocol. Eluted peptides were dried in vacuum, dissolved in 60% v/v aqueous ACN containing 1% v/v formic acid and analyzed by ESI-MS. The number of cysteines and disulfide bonds was calculated from the mass difference between the original sample and the alkylated peptide. For all following analyses, the iodoacetamide-modified peptides were used.

2.9 Edman degradation

The N-terminal amino acid sequence was determined with a Procise cLC 491 Protein Sequencing System (PE Applied Biosystems, USA). Phenylthiohydantoin derivatives of the amino acids were identified by a 120A PTH Analyzer (PE Applied Biosystems).

2.10 Enzymatic digestions

The desalted peptide was dried in vacuum and digested by one of the following three protocols with trypsin, Lys-C or Arg-C in solution. The dried peptide was dissolved in 17 μL of ammonium bicarbonate buffer (3 mmol/L, pH 7.8). Two μL trypsin solution (20 ng/μL in 3 mmol/L ammonium bicarbonate buffer) were added and incubated at 37°C for 1 h. The reaction was stopped by addition of 1 μL formic acid. Alternatively, the dried peptide was dissolved in 10 μL Tris-HCl buffer (25 mmol/L, 1 mmol/L EDTA, pH 8) and digested with 5 μL Lys-C solution (0.01 μg/μL in Tris-HCL buffer) for 1 h at 37°C. The reaction was stopped with 2 μL formic acid. For the Arg-C digest, the peptide was dissolved in 50 μL incubation buffer (100 mmol/L Tris-HCl, 10 mmol/L $CaCl_2$, pH 7.6) before 1 μL of Arg-C solution (20 ng/μL in 50 mmol/L Tris-HCl, 10 mmol/L CaCl, 5 mmol/L EDTA, pH 8) and 5 μL activation buffer (50 mmol/L DTT, 5 mmol/L EDTA) were added. After 3 h at 37°C, the reaction was stopped by addition of 2 μL formic acid. Whereas the tryptic digest was dried in a Speed Vac, the Lys-C and Arg-C digests were desalted first on a C_{18} ZipTip™ according to the manufacturer's protocol. The dried digests were dissolved in 4 μL of 60% v/v aqueous ACN containing 1% v/v formic acid and analyzed by MALDI- and ESI-MS.

2.11 Peptide derivatizations

The ε-amino groups of lysine residues were guanidated according to Samyn [20], dried in vacuum and desalted on a

C_{18}-ZipTip™ according to the manufacturer's protocol. The lyophilized tryptic or Lys-C digests were either directly, or after guanidation, derivatized with 2-sulfobenzoic acid cyclic anhydride (SACA) in solution [20] or after adsorption on a C_{18}-ZipTip™ with 4-sulfophenyl isothiocyanate sodium salt monohydrate (SPITC) [21]. Mass spectra of the desalted samples were recorded in both positive and negative ion mode, whereas MS/MS spectra were recorded only in positive ion mode.

2.12 Disulfide connectivity

The lyophilized native peptide sample was dissolved in 5 μL aqueous ammonium bicarbonate buffer (3 mmol/L, pH 7.8), mixed with 4 μL trypsin solution (20 ng/μL in 3 mmol/L ammonium carbonate buffer, pH 7.8), and incubated at 37°C overnight. The reaction was stopped by addition of 1 μL formic acid. Tryptic peptides were analyzed directly in positive ion mode by MALDI-MS [22, 23].

3 Results

3.1 Peptide isolation

So far, no leukocyte-derived antimicrobial peptides have been described for reptiles in the literature, although they are common in many other species. Thus, proteins and peptides were extracted from leukocytes obtained from turtles with acetic acid and separated by AU-PAGE. The band patterns obtained by the CBB stain were identical for female and male turtles (Fig. 1, panel A). Several bands containing proteins or peptides showed a significant antimicrobial activity against *C. albicans* and *L. monocytogenes* (strain EGD) indicated by the clear zones of the agar overlay (Fig. 1, panels B and C). Again, the zone patterns for both microbes were independent of the turtles' gender. Two bands appeared especially interesting, as they were active against both tested microbes.

In this study, we focused on the isolation and sequence analysis of the faster migrating band (lower band in Fig. 1) containing, most likely, a relatively small and highly basic peptide. The acetic acid extracts were first ultrafiltrated against a 10-kDa cut-off membrane to remove all proteins. The filtrate was concentrated by ultrafiltration (1-kDa cut-off membrane) and the peptide mixture, with an assumed molecular mass range between 1 and 10 kDa, was separated by preparative gel electrophoresis (see Supporting Information Fig. S1). Fractions 37-43 showed the greatest antimicrobial activity against *E. coli* ML35p, *L. monocytogenes* EGD, and *C. albicans* in a radial diffusion assay (see Supporting Information Fig. S1). These seven fractions were combined, dried, and fractionated by RP-HPLC using a linear ACN gradient in the presence of tetrabutylammonium hydrogen sulfate as ion pair reagent (Fig. 1, panel D).

All fractions were tested for their antimicrobial activities in a radial diffusion assay and the most active sample eluting at 22 min (marked with an arrow in Fig. 1, panel D) was fur-

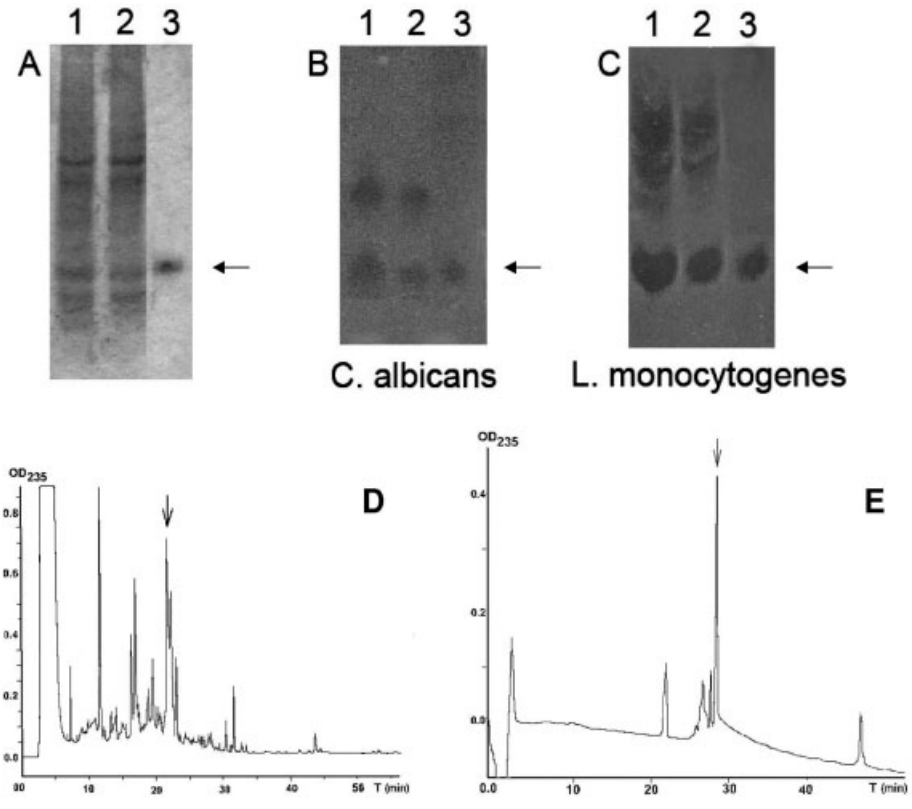

Figure 1. AU-PAGE of acid extracts (A) from leukocytes of male (lane 1) and female (lane 2) turtles and the corresponding overlay assays for *C. albicans* (B) and *L. monocytogenes* (C). Gels (T = 12.5%) were stained with CBB G-250 and rabbit α-defensin NP-4 was added as control (lane 3). Reversed-phase chromatograms of combined fractions 37-43 from a preparative gel electrophoresis using a Vydac C_{18}-column (D) and the 22-min fraction (marked with an arrow) of this chromatographic separation on an Alltech C_{18}-column using an alternative eluent system (E). Peptides were eluted by a linear gradient from 0 to 60% eluent B (D) or 0 to 60% eluent D (E) for 60 min at a flow rate of 1 mL/min. Fractions with highest antimicrobial activities used for further studies are marked with arrows.

ther purified on a different C_{18}-column using 0.13% v/v HFBA as ion pair reagent to retard the presumably basic peptide more strongly (Fig. 1, panel E). Finally, approximately 41 μg of a mostly homogeneous peptide with a monoisotopic mass of 4540.3 g/mol was obtained from 395 mg of total protein present in the acid leukocyte extract. This peptide was named turtle β-defensin 1 (TBD-1). Besides the dominant peptide, there were several minor impurities in the same mass range that could indicate further defensins present at lower quantities.

3.2 Antimicrobial and hemolytic activities

TBD-1 showed a high antimicrobial activity in a radial diffusion-agarose plate assay in the nmol/L-range against *E. coli* ML35p and *L. monocytogenes* EGD and in the low μmol/L-range against methicillin resistant *S. aureus* (MRSA) ATCC 33591 and *C. albicans* at low-salt concentrations (Table 2). Porcine protegrin PG-1 and human defensin HNP-1, which were tested as controls in parallel, showed similar activities as described [24]. The activity of both TBD-1 and HNP-1 decreased significantly at higher salt concentrations against all four tested organisms, whereas PG-1 was expectedly only slightly influenced. Thus, TBD-1 exhibited a similar activity pattern as reference peptide HNP-1, which is a typical representative of the defensin superfamily.

Whereas PG-1 showed a high hemolytic activity starting at 3 μmol/L further increasing to a hemolysis rate of about 80% at 25 μmol/L, neither TBD-1 nor HNP-1 were hemolytically active against human erythrocytes up to a concentration of 25 μmol/L (Supporting Information Fig. S2), which is about five times above the MIC value of TBD-1. Due to the low peptide amounts isolated from turtle leukocytes, it was not possible to extend the studies to higher peptide concentrations.

3.3 *De novo* sequencing by MS

Due to the peptide length and the low peptide amounts isolated, it appeared most promising to analyze the peptide by tandem mass spectrometry, which was challenging due to the expectedly low sequence homologies to known defensins from other species. Thus, the TBD-1 sequence had to be retrieved by *de novo* sequencing, combining different enzymes and MS/MS techniques. The relatively high charge states between five and eight in ESI-MS indicated a sequence rich in basic residues. After reduction with DTT and alkylation with iodoacetamide or acrylamide the charge state was retained, whereas the mass increased by 342.3 u for iodoacetamide and 426.7 u for acrylamide. Together with the mass shift of 6 u after reduction, both derivatization reagents indicated six cysteine residues forming three intramolecular disulfide bridges, which is typical for defensins and defensin-like peptides. Incubation with acrylamide yielded also minor byproducts, containing up to three additional acrylamide molecules, which resulted probably from alkylation of three amino groups, *e.g.* two lysine residues and a free N-terminus.

TBD-1 alkylated with iodoacetamide was digested with trypsin, endoprotease Arg-C or endoprotease Lys-C. Only partial sequences were obtained for all resulting peptides without sequence overlaps by PSD and CID in MALDI-MS. For most peptides, the sequence information was limited by incomplete fragment series, or the spectra were too complex to assign a certain sequence. The MS/MS spectra of the doubly and triply charged ions in ESI-MS displayed, at least for some peptides, dominant y-series, allowing for an easier interpretation. Thus, it was possible to deduce some sequences solely by ESI-MS/MS (Fig. 2), or in some cases, to close gaps in sequences only partially retrieved from MALDI-MS.

Most MS/MS spectra, however, did not allow retrieving much sequence information, and the obtained peptide masses did not sum up to the mass of the full-length peptide, *i.e.* TBD-1, indicating that some peptides were lost during sample preparation. So far, only two sequence stretches had been identified corresponding to a little more than half of the expected sequence, and it was unknown how to combine these sequences owing to overlapping peptide sequences. To overcome these limitations and to acquire a second independent dataset, the peptides of all three digests were chemically modified. Several published derivatization methods were evaluated with synthetic peptides at the low-pmol level and their influence on the fragmentation pattern studied. Finally, lysine residues were guanidated to enhance the gas phase basicity of C-terminal lysine residues obtained by trypsin or Lys-C [20] cleavage to pronounce the y-series. Additionally, the N-termini were derivatized with either

Table 2. MIC values of purified antimicrobial peptides TBD-1, HNP-1 and PG-1 in μmol/L

	E. coli ML35p		*Listeria monocytogenes* EGD		MRSA ATCC 33591		*Candida albicans* 820	
	LS[a)]	HS[b)]	LS	HS	LS	HS	LS	HS
TBD-1	0.65	>20	0.65	>20	5.6	>20	5.2	>20
HNP-1	0.90	>20	1.1	1.2	2.9	>20	3.6	>20
PG-1	0.95	1.1	0.95	0.45	0.85	0.75	2.2	2.0

a) LS (low salt) – 10 mmol/L sodium phosphate buffer (pH 7.4).
b) HS (high salt) – 10 mmol/L sodium-phosphate buffer (pH 7.4), 0.1 mol/L sodium chloride.

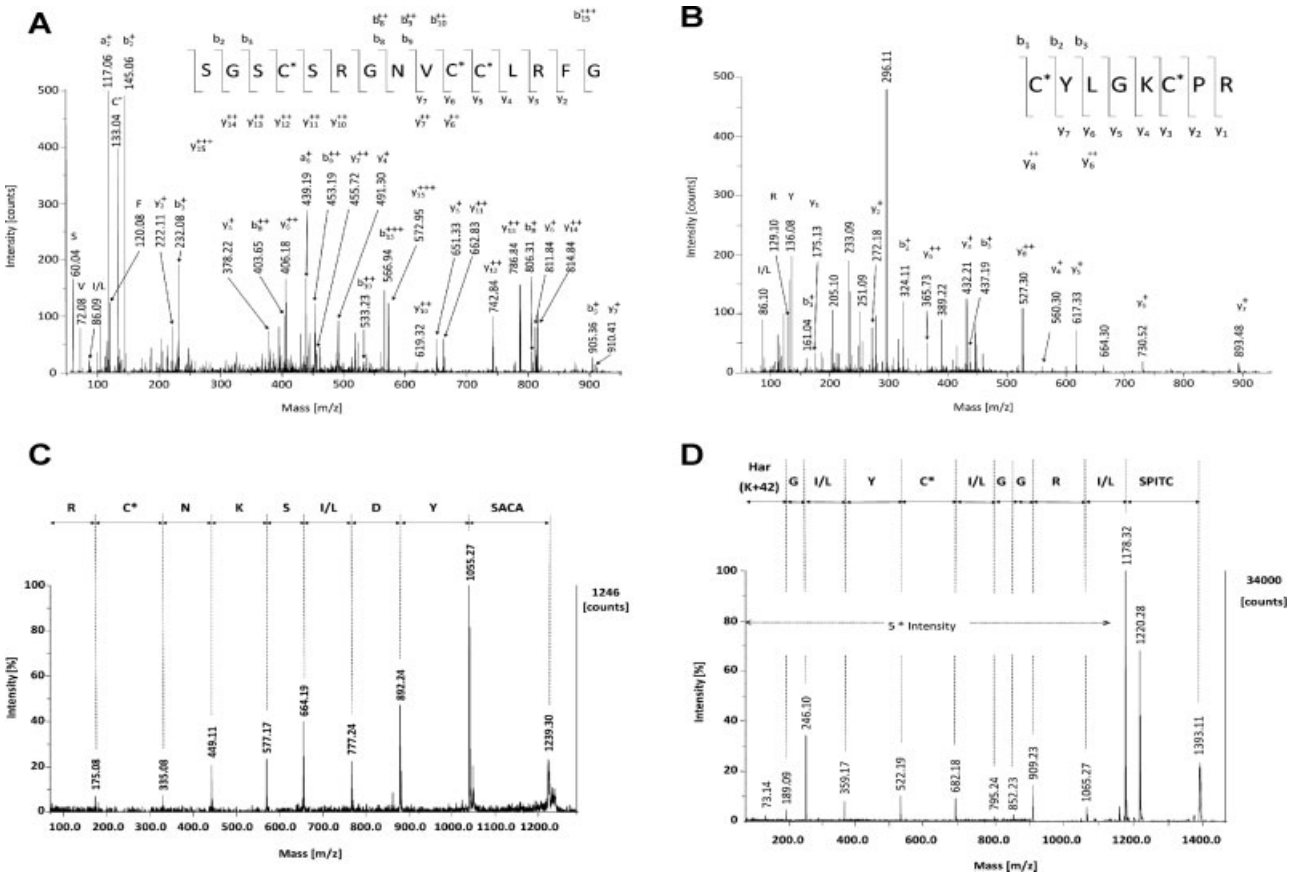

Figure 2. ESI-CID (panels A and B) and MALDI-PSD (panels C and D) MS/MS spectra of peptides obtained by an ArgC- (panels A to C) or a tryptic digest (panel D). Selected were the triply charged ion at *m/z* 572.95 (A) and the doubly charged ion at *m/z* 527.30 (B) of the digest or the singly charged ion at *m/z* 1239.3 after N-terminal derivatization with SACA (C) and the singly charged ion at *m/z* 1393.1 obtained by N-terminal derivatization with SPITC of a guanidated tryptic digest (D). TBD-1 was reduced with DTT and alkylated with iodoacetamide (C*) before enzymatic digestion. Deduced sequences and corresponding fragment ions are shown as inserts (L denotes either Leu or Ile)

SACA or SPITC [20, 21] to suppress the b-series by a negative charge located at the N-terminus of the peptide, allowing easier interpretation of the remaining y-ions in the MS/MS spectra. The obtained spectra indeed displayed enhanced y-ions that allowed for the complete sequencing of several peptides of the enzymatic digests, or to complete the partial sequences obtained before by direct analysis, as exemplified for parent ions at *m/z* 1239.3 and 1393.1 of an Arg-C and a trypsin digest, respectively (Fig. 2).

The combination of all data obtained so far resulted in a presumably N-terminal 21-residue ($YD^{I}/_{L}SKNCR^{I}/_{L}RGG^{I}/_{L}CY^{I}/_{L}GKCPR$) sequence, and a C-terminal 15-residue ($SGSCSRGNVCC^{I}/_{L}RFG\text{-}NH_2$) sequence. Between these two peptides remained a sequence gap corresponding to a molecular mass of 606.8 Da, which could match to peptide RFFR identified by MALDI-PSD in the tryptic digest after derivatization with SPITC. However, this middle sequence was not confirmed by any of the other mass spectra. The sequences deduced from the different mass spectra and their correct combination were finally verified by ESI-ETD using an IT. The fragment spectra of the five and six times protonated molecular ions of the alkylated full-length TBD-1 peptide confirmed the total sequence except the RFFR sequence (see Supporting Information Fig. S3). The z-series indicated GVFFR starting at position 22 of TBD-1, which is isobar to RFFR.

As none of the recorded mass spectra could solve this discrepancy, both peptides were synthesized on solid-phase, derivatized with SPITC, and their fragment spectra were recorded by MALDI-PSD. The spectra of both synthetic peptides displayed complete y-series, thus confirming the tetrameric RFFR sequence for the original tryptic peptide. At this point, only the question about leucine or isoleucine in positions 3, 9, 13, 16 and 37 remained to be solved. However, none of the techniques used so far gave any indication for either of the two residues. Even high-energy CID spectra recorded on the MALDI-TOF/TOF-MS did not display fragment ions indicative for either leucine or isoleucine. Therefore, TBD-1 was N-terminally sequenced by Edman degradation to determine the first four Leu/Ile positions. Thus, the complete sequence of TBD-1 was determined to be YDLSKNCRLRGGICYIGKCPRRFFRSG $SCSRGNVCC^{L}/_{I}RFG\text{-}NH_2$.

The initial reduction and alkylation experiments to determine the number of cysteine residues indicated already that all six cysteine residues form intramolecular disulfide bridges. Thus, the tryptic peptide mass fingerprint for each of the theoretically possible 15 different disulfide connectivities was calculated in silico [22, 23]. The MALDI-MS data obtained for the tryptic digest of the native non-reduced TBD-1 peptide corresponded to two possible isomers (Fig. 3, Supporting Information Table S1), *i.e.* α- or β-defensins (Table 1). All other 13 structures did not fit to the recorded mass spectra. However, it was not possible to distinguish between α- and β-defensins by MALDI-TOF/TOF-MS and it is not possible to cleave the bond between Cys35 and Cys36 enzymatically or chemically at such a small scale.

Having retrieved the complete sequence analysis including the likely disulfide-bridging we synthesized TBD-1 on solid phase by 9-fluorenylmethoxycarbonyl/*tert*-butyl (Fmoc/^{t}Bu)-chemistry using three orthogonal cysteine-protecting groups, *i.e.* trityl (Trt), *tert.*-butyl (^{t}Bu) and acetamidomethyl (Acm). The three Cys-protecting groups cleaved stepwise after the previous disulfide bond was obtained [26]. All intermediates as well as the final product were purified by RP-HPLC and analyzed by MS (see Supporting Information Fig. S4, panel A). The overall yield was 100 μg (about 2%), which was enough to show that the synthetic peptide was active against *E. coli* BL21AI, *M. luteus* 10 240 und *B. subtilis* 347 (see Supporting Information Fig. S4, panels B and C) in an Agar diffusion assay. Currently, the synthesis is further optimized and up-scaled to study the antimicrobial activity in more detail and to determine the MIC-values.

4 Discussion

A novel defensin called TBD-1 was isolated from leukocytes of the European pond turtle *Emys orbicularis* and purified in three consecutive steps by AU-PAGE and RP-HPLC using first a positively charged and then a negatively charged ion pair reagent on different C18-phases. The obtained peptide showed an interesting activity against both Gram-positive and Gram-negative bacteria including MRSA as well as fungi with MIC values in the low μM to even nM range. Although these antimicrobial activities, the peptide mass, and the number of cysteine residues indicated a defensin, or at least defensin-like peptide, the following sequence analysis was

Figure 3. Monoisotopic masses of the cysteine containing $[M+H]^+$ quasi-molecular ions theoretically obtained from a tryptic digest of the TBD-1 peptide with native intramolecular disulfide bridges.

challenging due to the limited peptide amounts accessible only to sensitive MS techniques. The unfavorable distribution of enzymatic cleavage sites in the highly basic sequence, as well as the poor fragmentation of the complete TBD-1 sequence, yielding only partial b- and y-series were also major challenges that had to be overcome. Furthermore, the identified sequences were not listed in databases and had to be retrieved *de novo*, which required overlapping sequences by different enzymatic digests. This was finally achieved with trypsin, endoprotease Lys-C and endoprotease Arg-C. The first indication to use this combination of endoproteases came from the high charge states in ESI-MS and the byproducts obtained after the acrylamide treatment to determine the number of cysteine residues. However, most peptides showed poor fragmentation patterns, with either not enough signals to give access to the full b- or y-series, or so many signals that it was impossible to retrieve a single reliable sequence. Mascot was also not able to identify any peptide sequence from the recorded tandem mass spectra with a reasonable score, which is now obvious from the low sequence homologies to all other defensin sequences listed in protein databases. Thus, different chemical derivatization techniques were tested to obtain longer and dominant y-series by guanidation of C-terminal lysine residues and derivatization of the N-terminal amino groups to suppress the b-series. Whereas guanidation worked well even for low peptide amounts, most tested N-terminal derivatization techniques described in the literature did not work in our hands at the low-pmol scale. Different reagents and methods were therefore evaluated first for low amounts of synthetic peptides. Based on the number of impurities, peptidic byproducts, peptide loss and especially interpretation of the obtained MS/MS spectra, SPITC and SACA were finally selected for the *de novo* approach. Indeed, both reagents worked for the TBD-1 digests and the derivatized peptides and showed superior fragmentation characteristics in MALDI-PSD, allowing an easier and more reliable sequence

Figure 4. Combination of the partial sequences obtained from different enzymatic digests and chemical modifications using ESI- and MALDI-MS/MS.

analysis for most peptides. The acquired datasets including the initial studies on synthetic peptides (data not shown) clearly indicated that both reagents might be of general use for miniscule peptide amounts down to one pmol for MALDI-MS analyses. Overall, each part of the sequence was obtained independently from at least two different fragmentation techniques (PSD, CID and ETD), or derivatization methods (SPITC, SACA, homoarginine and unmodified) (Fig. 4). Only for positions 22 to 25 (or 26) were two different isobaric sequences identified by MALDI-PSD (RFFR) and ETD (GVFFR). As it was not possible to solve this sequencing problem from any of the obtained spectra, we synthesized both short peptides, derivatized them, and confirmed that the sequence RFFR was indeed correct. Edman degradation identified also the ambiguous Leu/Ile-position and confirmed some parts of the N-terminal sequence, although several residues were below the LOD due to the low peptide amounts loaded. The calculated monoisotopic mass of the derived TBD-1 sequence of 4540.1 Da was in good agreement with the deconvoluted mass of 4540.3 Da obtained by ESI-MS.

From the 15 combinations theoretically possible for three disulfide bridges formed by six cysteine residues, we could exclude 13 based on the recorded MALDI mass spectra. Only two structures resembling α- and β-defensin remained. A further discrimination between these two classes was not possible, as Cys-5 and Cys-6 are neighbored. Although it is possible to determine the disulfide connectivity for such peptides, much larger peptide amounts, not available for TBD-1, would be required. However, the sequence length (40 residues) clearly indicates that TBD-1 belongs to the class of β-defensins (Table 1) and not to the smaller class of α-defensins (29- to 35- residue peptides). Additionally, the first two cysteine residues are not located near the N-terminus, and are not separated by one amino acid, which both would be indicative for α-defensins. This is further supported by structural features common to other β-defensins identified, especially the spacing among the cysteine-residues, in mammals and birds as well as partial sequence homologies (Fig. 5).

This sequence alignment showed that altogether 28 positions of the TBD-1 sequence are homologous to other avian or mammalian β-defensin sequences. The highest sequence homologies were found for Gal-1 and Gal-9 with 14 residues, as well as Gal-5, Gal-6 and Gal-12 with 13 residues. Moreover, these defensins have similar peptide lengths and several conservative mutations compared to the TBD-1 sequence. This is also illustrated by the dendrogram of TBD-1 (Fig. 6) based on a phylogenetic and molecular evolutionary analyses with MEGA (version 3.1) [24].

In contrast, the recently reported antibacterial peptide TEWP from reptile eggs [14] is similar to human β-defensin 3 (hBD-3) based on the NMR structure. Its disulfide connectivity is clearly different from α- and β-defensins isolated from mammals and birds (Table 1), which is also obvious from the sequence alignment (Fig. 5). In conclusion, we

```
TBD-1   YDLSKNCRLRGGICYIGKCPRRFFR-SGSCSRGNV-CCLRFG-NH2
Gal-1     GRKSDCFRKSGFCAFLKCPSLTLI-SGKCSRFYL-CCKRIWG
Gal-9     GLPQDCERRGGFCSHKSCPPGIGR-IGLCSKEDF-CCRSRWYS
Gal-5     RPIDTCRLRNGICFPGICRRPYYW-IGTCNNGIGSCCARGWRS
Gal-6     SPIHACRYQRGVCIPGPCRWPYYR-VGSCGSGLKSCCVRNRWA
Gal-1     NNEAQCEQAGGICSKDHCFHLHTRAPGHCQRG-VPCCRTVYD
Gal-14  ESDTV-TCRKMKGKCSFLLCPFFK-RSSGTCYNGLAKCCRPFW
THP-1     GKREKCLRRNGFCAFLKCPTLSVI-SGTCSRFQV-CC
THP-2        LFCK-R-GTCHFGRCPSHLIK-VGSCFGFRS-CCKWPWDA
Ost-1        LFCR-K-GTCHFGGCPAHLVK-VGSCFGFRA-CCKWPWDV
Ost-4   LPVNEAQCRQVGGYCGLRICNFPSRF-LGLCTRNHP-CCSRVWV
mBD-1     DQY-KCLQHGGFCLRSSCPSNTKL-QGTCKPDKPNCCKS
hBD-1      DHYNCVSSGGQCLYSACPIFTKI-QGTCYRGKAKCCK
BNBD-1     DFASCHTNGGICLPNRCPGHMIQ-IGICFRPRVKCCRSW
TEWP       pyroEKKCPGRCTLKCGLHERPTLPYNCGK-YI-CCVPVKVK
```

Figure 5. Sequence alignment of TBD-1 with β-defensins from chicken (Gallinacins, Gal), turkey (THP) and ostrich (Ostricacins, Ost) as well as mouse β-defensin 1 (mBD-1), human β−defensin 1 (hBD-1), bovine β-defensin 1 (BNBD-1) and the defensin-like peptide from marine turtle (TEWP). Bold positions indicate conserved cysteine residues and underlined positions indicate sequence homologies to TBD-1.

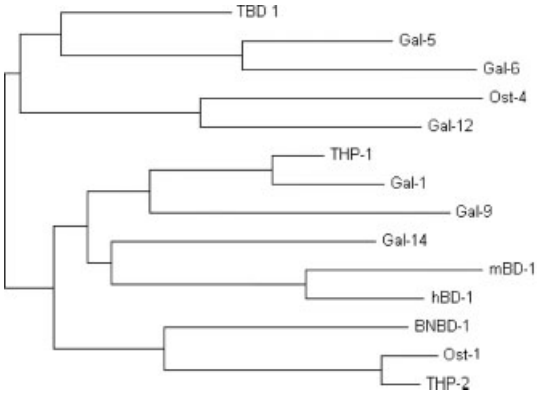

Figure 6. Dendrogram of TBD-1 with β-defensins from chicken (Gallinacins, Gal), ostrich (Ostricacins, Ost) and turkey (THP-2) as well as bovine β-defensin 1 (BNBD-1), mouse β-defensin 1 (mBD-1) and human beta defensin 1 (hBD-1) using MEGA (version 3.1) [25].

were able to isolate and sequence the first β-defensin from reptiles, which we suggest be named TBD-1 for turtle beta defensin 1. To the best of our knowledge, TBD-1 is the first defensin isolated from leukocytes of reptiles.

We would like to thank Dr. Andrea Schneider (Bruker Daltonik GmbH, Bremen) for recording the ETD mass spectra and for providing the BioTools software to analyze the spectra. We would also like to thank Dr. Thole Züchner for critical reading of this manuscript. This work was supported by the European Fond for Regional Structure Development (EFRE, European Union and Free State Saxony), the Leonhard-Euler-Program of the German Academic Exchange Service (DAAD), the Russian

Foundation of Basic Research (project numbers 09-04-01202, 06-04-49416 and 07-04-01759) and the Russian Federal Agency for Science and Innovations (state contract No. 02.512.11.2234 of 07.07.2008).

The authors have declared no conflict of interest.

5 References

[1] Boman, H. G., Peptide antibiotics and their role in innate immunity. *Annu. Rev. Immunol.* 1995, *13*, 61–92.

[2] Barra, D., Simmaco, M., Boman, H. G., Gene encoded peptide antibiotics and innate immunity. Do 'animalcules' have defense budgets? *FEBS Lett.* 1998, *430*, 130–134.

[3] Ganz, T., The role of antimicrobial peptides in innate immunity. *Integr. Comp. Biol.* 2003, *43*, 300–304.

[4] Bulet, P., Stocklin, R., Menin, L., Anti-microbial peptides: from invertebrates to vertebrates. *Immunol. Rev.* 2004, *198*, 169–184.

[5] Brogden, K. A., Antimicrobial peptides: Pore formers or metabolic inhibitors in bacteria? *Nat. Rev.* 2005, *3*, 238–250.

[6] Lehrer, R. I., Lichtenstein, A. K., Ganz, T., Defensins: antimicrobial and cytotoxic peptides of mammalian cells. *Annu. Rev. Immunol.* 1993, *11*, 105–128.

[7] White, S. H., Wimley, W. C., Selsted, M. E., Structure, function, and membrane integration of defensins. *Curr. Opin. Struct. Biol.* 1995, *5*, 521–527.

[8] Zhao, C., Nguyen, T., Liu, L., Sacco, R. E. *et al.*, Gallinacin-3, an inducible epithelial beta-Defensin in the chicken. *Infect. Immun.* 2001, *69*, 2684–2691.

[9] Tang, Y.Q., Yuan, J., Osapay, G., Osapay, K. *et al.*, A cyclic antimicrobial peptide produced in primate leucocytes by the ligation two truncated α-defensins. *Science* 1999, *286*, 498–502.

[10] Otvos, L. Jr., Antibacterial peptides isolated from insects. *J. Pept. Sci.* 2000, *6*, 497–511.

[11] Hoffmann, J., Hetru, C., Insect defensins: Inducible antibacterial peptides. *Immunol. Today* 1992, *13*, 309–316.

[12] Lehrer, R. I., Ganz, T., Defensins of vertebrate animals. *Cur. Opin. Immun.* 2002, *14*, 96–102.

[13] Liu, L., Zhao, C., Heng, H. H. Q., Ganz, T., The human beta-Defensin-1 and alpha-Defensins are encoded by adjacent genes: Two peptide families with differing disulfide topology share a common ancestry. *Genomics* 1997, *46*, 316–320.

[14] Chattopadhyay, S., Sinha, N. K., Banerjee, S., Roy, D. *et al.*, Small cationic protein from a marine turtle has beta-defensin-like fold and antibacterial and antiviral activity. *Proteins* 2006, *64*, 524–31.

[15] Lakshminarayanan, R., Chi-Jin, E. O., Loh, X. J., Kini, R. M.,Vivekanandan, S., Purification and characterization of a vaterite-induced peptide, pelovaterin, from the eggshells of *Pelodiscus sinensis*. (Chinese soft-shelled turtle). *Biomacromolecules* 2005, *6*, 1429–1437.

[16] Lakshminarayanan, R., Vivekanandan, S., Samy, R. P., Banerjee, Y. *et al.*, Structure, self-assembly, and dual role of a β-defensin-like peptide from the Chinese soft-shelled turtle eggshell matrix. *J. Am. Chem. Soc.* 2008, *130*, 4660–4668.

[17] Panyim, S., Chalkley, R., High resolution acrylamide gel electrophoresis of histones. *Arch. Biochem. Biophys.* 1969, *130*, 337–346.

[18] Schägger, H., von Jagow, G., Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* 1987, *166*, 368–379.

[19] Lehrer, R. I., Rosenman, M., Harwig, S. S., Jackson, R., Eisenhauer, P., Ultrasensitive assays for endogeneous antimicrobial polypeptides. *J. Immunol. Methods* 1991, *137*, 167–173.

[20] Samyn, B., Debyser, G., Sergeant, K., Devreese, B., Van Beeumen, J. A., Case study of *de novo* sequence analysis of N-sulfonated peptides by MALDI TOF/TOF mass spectrometry. *J. Am. Soc. Mass Spectrom.* 2004, *15*, 1838–1852.

[21] Chen, P., Nie, S., Mi, W., Wang, X.-C., Liang, S.-P., *De novo* sequencing of tryptic peptides sulfonated by 4-sulfophenyl isothiocyanate for unambiguous protein identification using post-source decay matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* 2004, *18*, 191–198.

[22] Patterson, S. D., Katta, V., Prompt fragmentation of disulfide-linked peptide during matrix-assisted laser desorption ionization mass spectrometry. *Anal. Chem.* 1994, *66*, 3727–3732.

[23] Bean, M. F., Carr, S. A., Characterization of disulfide bond position in proteins and sequence analysis of cysteine-bridged peptides by tandem mass spectrometry. *Anal. Biochem.* 1992, *201*, 216–225.

[24] Kokryakov V. N., Harwig S. S. L., Panyutich E. A., Shevchenko A. A. *et al.*, Protegrins: Leukocyte antimicrobial peptides combine features of corticostatic defensins and tachyplesins. *FEBS Lett.* 1993, *327*, 231–236.

[25] Kumar, S., Tamura, K., Nei, M., MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief. Bioinform.* 2004, 150–163.

[26] Fields, G. B., Noble, R. L., Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *Int. J. Pept. Protein Res.* 1990, *35*, 161–214.

