

Isolation and sequence analysis of peptides from the skin secretion of the Middle East tree frog *Hyla savignyi*

Markus Langsdorf · Alireza Ghassempour ·
Andreas Römpf · Bernhard Spengler

Received: 8 June 2010 / Revised: 12 August 2010 / Accepted: 15 August 2010 / Published online: 11 September 2010
© Springer-Verlag 2010

Abstract Novel peptides were identified in the skin secretion of the tree frog *Hyla savignyi*. Skin secretions were collected by mild electrical stimulation. Peptides were separated by reversed-phase high-performance liquid chromatography. Mass spectra were acquired by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS), and fragment ion spectra were obtained after collision-induced dissociation and electron capture dissociation. Peptides were analyzed by manual de novo sequencing and composition-based sequencing (CBS). Sequence analyses of three so far undescribed, structurally unrelated peptides are presented in this paper, having the sequences DDSEEEVE-OH, P*EEVEEERJK-OH, and GJJDPJTGJVGJJ-NH₂. The glutamate-rich sequences are assumed to be acidic spacer peptides of the prepropeptide. One of these peptides contains the modified amino acid hydroxyproline, as identified and localized by high-accuracy FTICR-MS. Combination of CBS and of experience-based manual sequence analysis as complementary and database-independent sequencing strategies resulted in peptide identification with high reliability.

Keywords Frog peptides · Composition-based sequencing · Peptide fragmentation · Sequence rearrangements · Proline effect

Introduction

Skin secretions of the dorsal glands of amphibians provide a cocktail of bioactive compounds which protect the animals from microorganisms and predators. The granular glands produce a secretion containing various compounds of biological activity. The main secretion components are peptides [1, 2] and alkaloids [3].

Amphibian peptides have a large variety of biological functions, such as antimicrobial [4–7], antiviral [8, 9], and anticancer [10]. Biological activity is induced by peptide–lipid interactions with target membranes, leading to membrane penetration and cell death. The mode of action, effected by cationic amphiphilic peptides, is described by diverse suggested mechanisms [11–14]. A sequence database of antimicrobial peptides from various biological sources was founded in 2003 [15] and updated recently [16].

The suitability of bioactive peptides in mammalian systems, however, is very limited so far. Pharmacokinetic and pharmacodynamic barriers, instability against proteases, systemic and local toxicity, or reduced activity based on salt and pH sensitivity make a medical application difficult [17, 18]. An additional problem is the high cost of chemical synthesis in the production of antimicrobial peptides.

A review on structure determination and bioactivity of frog peptides, mainly isolated from Australian species, was published recently [19, 20]. The frog species of our studies, *Hyla savignyi*, belongs to the Hylidae family and was collected in the southern part of Meshkin Shahr in the Ardabil province of Iran. We recently analyzed peptides of this species and classified them as tryptophyllin-like

Published in the special issue *Mass Spectrometry (DGMS 2010)* with Guest Editors Andrea Sinz and Jürgen Schmidt.

Electronic supplementary material The online version of this article (doi:10.1007/s00216-010-4131-8) contains supplementary material, which is available to authorized users.

M. Langsdorf · A. Römpf · B. Spengler (✉)
Institute of Inorganic and Analytical Chemistry,
Justus Liebig University,
35392 Schubertstr. 60, Bldg. 16, Giessen, Germany
e-mail: Bernhard.Spengler@anorg.chemie.uni-giessen.de

A. Ghassempour
Medicinal Plants and Drugs Research Institute,
Shahid Beheshti University,
Evin, Tehran 1983963113, Iran

peptides [21]. Sequence analysis was presented and specific fragmentation features, such as formation of rearrangement products and the proline effect, were described. Peptides were found to show structural similarities to those found in the species *Litoria rubella* and *Litoria electrica*, called tryptophyllins L [22, 23]. Tryptophyllin-like peptides were first isolated from *Phyllomedusa rohdei* [24].

Peptide characterization without employing database searching programs, called de novo sequencing [25–28], has become an important process in peptide analysis of non-standard species. The procedure often works reliably and fast if tryptic peptides having basic residues at the C-terminus are investigated by tandem mass spectrometry. If peptides are not digestion products or have non-basic amino acids at the C-terminal end, however, this procedure often fails [29]. Basic amino acids in the central part of a peptide sequence may cause highly complex fragmentation spectra [30]. Incomplete fragment ion information and posttranslational modifications can be an obstacle in finding the complete and correct sequence. Rearrangement products in collision-induced dissociation (CID) measurements, formed after cyclization reactions and subsequent ring opening at various peptide bonds, can lead to misinterpretations [31, 32]. CID experiments of proline-containing peptides can generate very specific MS/MS spectra, depending on the position of proline along a peptide chain [33–37], and may complicate peptide characterization as well. Novel frog peptides of the animal's skin secretion are usually not listed in protein/peptide sequence databases. This requires an efficient database-independent sequencing strategy.

Peptides investigated in this study were characterized by manual de novo sequencing and by composition-based sequencing (CBS) [38]. The latter method is a database-independent sequencing strategy and includes two steps. In the first step, the amino acid composition (AAC) is determined by the evaluation of accurate MS and MS/MS mass values without employing probability functions. The second step determines the peptide sequence for the given AAC by scoring the agreement between the observed and expected fragment ion signals of permuted sequence propositions. The feasibility of the CBS strategy requires sufficient mass accuracy and (as every sequencing strategy) sufficient fragment ion information. In our previous studies of the tryptophyllins H [21], CBS turned out to be a valuable method since CID spectra were complex and most high-performance liquid chromatography (HPLC) fractions did not provide satisfying results when performing Edman degradation, probably due to sensitivity problems.

Peptide sequences of the present study were determined by CBS and manual spectra evaluation based on CID and electron capture dissociation (ECD) spectra. Since the amount of sample was limited, we were not able to extend the investigation by improving our Edman methodology. Deci-

phering I/L ambiguity was thus beyond the scope of the work presented here. Sequences of characterized peptides were confirmed by CID measurements of analog synthetic peptides.

Experimental

Collection of skin secretions

Adult specimens of *H. savignyi* were collected in the southern part of Meshkin Shahr in the Ardabil province of Iran. Frogs were submerged in a beaker containing deionized water and exposed to gentle electrical stimulation in ambient atmosphere using copper electrodes (3 V, 2 Hz, 4-ms pulse duration). Electrodes were rubbed on the skin surface for 40 s and a milky secretion was excreted and absorbed by the deionized water. Afterwards, animals were manually and gently submerged for an additional 60 s in the beaker to complete the extraction of skin secretion. The resulting aqueous solution was lyophilized with a freeze drying device and stored at -20°C . Animals were released in their natural habitat after treatment.

HPLC separation

One milligram of the lyophilized sample was dissolved in 500 μl water containing 5% acetonitrile. An insoluble residue was removed by centrifugation. For nano-HPLC-MS analysis, a binary nano-HPLC pump/autosampler system (Ultimate, Dionex/LCPackings, Idstein, Germany) was employed. Of a 1:10 dilution of the sample, 1 μl was pre-focused on a trap column (C18 PepMap, I.D. 300 μm , length 5 mm, Dionex) and separated on a fused-silica C18 PepMap100 capillary column (3 μm , 100 \AA ; I.D. 75 μm ; length 150 mm, Dionex). A multistep gradient from 10% to 50% solvent B [80% acetonitrile (v/v), 20% water (v/v), and 0.1% formic acid (v/v)] over a period of 25 min, followed by a 15-min isocratic step, and another gradient from 50% to 70% solvent B for a further 20 min was applied. Water containing 2% acetonitrile (v/v) and 0.1% formic acid (v/v) was used as solvent A. The flow rate was 0.2 $\mu\text{l}/\text{min}$. The eluent was monitored by ultraviolet absorbance at 214 nm.

A Summit microHPLC system (Dionex) was used for fractionation. Ten microliters was loaded onto a Microbore column (Macherey-Nagel, C18, 5 μm , 300 \AA , I.D. 1 mm, length 150 mm). The flow rate was 40 $\mu\text{l}/\text{min}$. Multistep gradient and solvents were used as in the nano-HPLC separation.

Mass spectrometry

Mass spectra were acquired on a "Finnigan LTQ FT Ultra" hybrid instrument (Thermo Scientific GmbH, Bremen,

Germany) consisting of a linear quadrupole ion trap and a Fourier transform ion cyclotron resonance (FTICR) mass spectrometer equipped with a 6-T magnet. A nanospray ionization source was used for analyses (capillary temperature 250 °C, capillary voltage 10 V, tube lens 100 V). Nanospray needles for off-line measurements were employed (spray voltage 1.4 kV, O.D. 1.2 mm, I.D. 5.0 µm, DNU-MS, Berlin, Germany). For online measurements, Silica Tips (Tip I.D. 10 µm, New Objective, Woburn, MA, USA) were used at a spray voltage of 1.9 kV. Mass resolving power was set to 100,000 at $m/z=400$, allowing us to record one spectrum per second and thus several spectra per chromatographic peak. Mass accuracy after external calibration was better than ± 2.0 ppm during both online and off-line measurements. Instrument-specific parameter settings for collision-induced dissociation in MS/MS experiments were as follows: activation energy 30 (normalized; manufacturer-specific units); activation duration 30 ms; activation Q 0.25 (manufacturer-specific units); wide scan range, three microscans per spectrum. The data acquisition software Xcalibur 2.0 SR2 (Thermo Fisher Scientific) was employed.

Sequence analysis

Calculations of accurate mass values of peptide ions and peptide fragment ions were performed using the computer program Peptide Composer 1.0 (copyright Bernhard Spengler). Peptide composition analysis, CBS, and manual de novo sequencing are supported by this program. Calculations were performed with a mass tolerance setting of ± 2.0 ppm for precursor and fragment ions. Based on possible compositions of fragment ions and precursor ion, AACs of the peptides were determined. The composition analysis algorithm of CBS performs a combinatorial analysis and does not contain experience-based evaluation steps. Because of the identical mass values of leucine and isoleucine, all calculations were performed with treating both amino acids as being equal ($L/I=J$). The number of all amino acids (including phosphorylation) was set to unknown. Starting the AAC calculations, peptides were treated as having one of three possible structures: free acid, acid amide, or cyclic structure. A decision on the structural property was established as soon as no possible AACs remained for two of the three C-terminal structures during the composition analysis. After precursor ion calculation and generation of an AAC list, composition analysis of fragment ions (within the candidate AACs of the precursor) was carried out starting with the CID spectra signal of the lowest mass value. The AAC list of the precursor ion was then checked against the AAC list of the fragment ion. This was executed in several iteration steps after each fragment composition analysis until all fragment ion mass values were evaluated.

Lists of sequence propositions were generated from determined AACs by exhaustive permutation. After in silico fragmentation of sequence propositions, theoretical and observed fragment ions were compared and matching ions were registered by the CBS procedure. Each hit was weighed with a fragment type-specific factor. In addition, we carried out a manual evaluation of the CID and ECD spectra.

Syntheses of peptides

Peptides were synthesized by GL Biochem (Shanghai) Ltd. (Shanghai, China) and were delivered with a purity of $>80\%$.

Results and discussion

Figure 1 shows the HPLC separation of the skin secretion. HPLC fractions were labeled A to L, and the corresponding identified peptides are listed in Table 1. Fractions C to K contained peptides that were characterized and classified as tryptophyllins H in a previous study [21]. They share sequence similarities with peptides found earlier in the skin secretion of the frogs *L. rubella* and *L. electrica*, called tryptophyllins L [22, 23]. Fractions between 24 and 30 min were investigated by MS/MS analysis. No peptides were found in these fractions. Sequence analysis of peptides found in fractions A, B, and L are discussed in this paper, and the results are listed in Table 1. Peptides were named savignins by us, according to the taxonomic name *H. savignyi*. CBS was carried out for the peptides of fractions A and L (savignin 1+3). The fraction B peptide (savignin 2) could be easily sequenced by manual evaluation of the ECD spectrum. All peptide sequences were confirmed by CID measurements of analog synthetic peptides. First bioassays, implemented on antibacterial activity against *Micrococcus luteus* and *Escherichia coli*, gave no significant results and are not described in this paper.

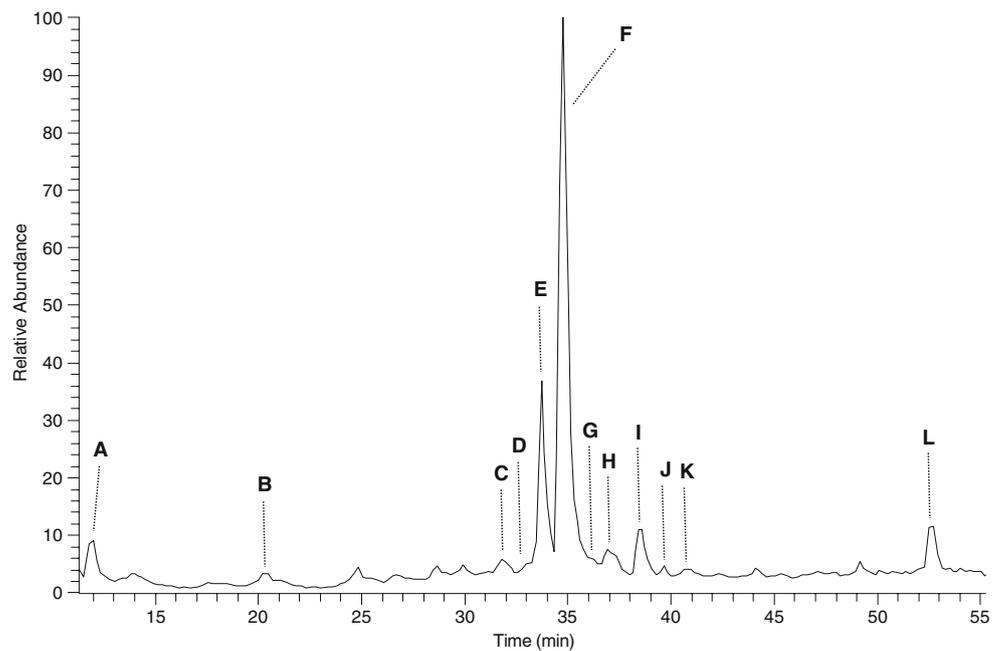
According to the accepted nomenclature [39, 40], b-type fragment ions include the N-terminus of the peptide. In our studies, we also observed two-step (internal) b-type fragments that are labeled as “ $b_{i(x-y)}$ ” with x and y being the start and end positions of the terminating amino acids of the fragment ion within the peptide chain [41]. The same nomenclature is applied to a- and z-type ions.

Savignin 1

Composition analysis of the fraction A peptide (savignin 1) resulted in 1,276 possible AAC candidates belonging to a C-terminal free acid, 997 AAC candidates of acid amide, and 1,591 AAC candidates of cyclic structure (Table 2).

After precursor ion calculation, the list of possible peptide amino acid compositions was narrowed down by

Fig. 1 Online nano-LC-MS separation (plot of total ion count) of the skin secretion of *H. savignyi*. Fractions C to K contain tryptophyllin-like peptides (tryptophyllins H) that were described recently [21]



fragment composition analyses (Table 3). Most fragment ion signals gave no hits for acid amide structure, so that this structure could be excluded in an early stage and is not considered in the table. Starting with the signal of the lowest mass value in the CID spectrum (Fig. 2a), calculation of fragment ion 247.12898 u gave a y-type ion of possible compositions (D_1J_1) and (E_1V_1) when checked against the precursor AAC list. Based on this result of fragment composition analysis, the precursor AAC list was updated and gave 172 possibilities of cyclic and 113 possibilities of free acid structures. The procedure was carried out in several iteration steps, and the precursor AAC list was updated after each fragment ion calculation (as shown in Table 3). After the composition analysis procedure, the peptide was found to have one out of two possible AACs of free acid ($D_3E_4J_1S_1$ or $D_2E_5S_1V_1$) and one cyclic structure ($D_2E_3J_1M_2S_1$). Fragment ions of high signal intensity at 834.26340 u and 933.33114 u gave no hit for a cyclic peptide structure. Furthermore, compositions of y-type fragment ions that were found within the peptide composition list of cyclic

structure could only be interpreted as (isobaric) $b+H_2O$ ions since cyclic peptides have no C-terminal end. The corresponding b ions, however, were not found in the MS/MS spectrum. Thus, the cyclic structure could be discarded. CBS analysis of the two remaining AAC candidates with C-terminal free acid gave DDSEEEVE as the overall highest sequence score. The precursor mass value (Table 2) therefore had a deviation of +0.94 ppm compared to the calculated mass value (540.69621 u). The CBS results (top 4 of each sequence analysis) are presented in Table 4.

CBS analysis of the ECD spectrum (Fig. 2b) confirmed the sequence found after an evaluation of CID mass data. The corresponding fragment ion mass list is shown in Table 5. Mass accuracy was improved by internal fine calibration as described earlier [38].

The ECD spectrum is unexpectedly dominated by a series of b and y ions (besides the expected c ions). With the exception of z_5' , ions of z' type occurred as two-step fragments. The signal at 687.25841 u had a considerable intensity, but was not matched in the CBS procedure. Based

Table 1 Savignin peptides from the skin secretion of *H. savignyi*

Peptide	Sequence	m/z $[M+H]^+$ (monoisotopic)		Deviation (ppm)	HPLC fraction
		Experimental	Theoretical		
Savignin 1	DDSEEEVE-OH	540.6967	540.6962	+0.94	A
Savignin 2	P*EEVEERJK-OH ^a	637.3166	637.3172	-1.83	B
Savignin 3	GJJDPJTGJVGJJ-NH ₂ ^a	668.9128	668.9136	-1.15	L

P* hydroxyproline

^aJ = L/I

Table 2 Number of possible AACs of savignin 1 (with a mass tolerance setting of ± 2.0 ppm) before fragment ion calculation

Precursor $[M+2H]^{2+}$ (monoisotopic)	Δ (ppm)	Number of possible AACs (outset)		
		Free acid	Acid amide	Cyclic
540.69672	2.00	1,276	997	1,591

on the high mass accuracy, this fragment was identified as internal z' type ion with a loss of water ($z_{i(3-8)}'-H_2O$), and no further fragment candidate was found. When the sequence ladder is pursued, it is obvious that (due to the identical amino acid composition) the signal at 600.22652 u may consist of both $z_5'-H_2O$ and $z_{i(4-8)}'-H_2O$. Accordingly,

it is assumed that the signal at 471.18378 u has a high ratio of $z_{i(5-8)}'-H_2O$ compared to $z_4'-H_2O$.

To understand the formation of internal z ions, we examined the C-terminal amino acid glutamic acid. Pyroglutamic acid is a common modification in peptides containing glutamic acid. This amino acid is known to

Table 3 Calculation of fragment ion mass values of savignin 1 (generated by CID) with a mass tolerance setting of ± 2.0 ppm

Fragment mass (u)		δ (ppm)	Fragment ion type	Possible fragment AACs ^a	Number of possible peptide AACs ^b	
Exp.	Calc.				Cyclic	Free acid
247.12898	247.12885	0.53	y	D ₁ J ₁ E ₁ V ₁	172	113
358.16073	358.16088	-0.42	b	D ₁ E ₁ J ₁ E ₂ V ₁	82	37
376.17135	376.17144	-0.24	y	D ₁ E ₁ J ₁ E ₂ V ₁	82	37
447.13498	447.13578	-1.79	a b	D ₃ E ₁ D ₂ E ₁ S ₁ D ₃ T ₁	11	11
505.21371	505.21403	-0.63	y	D ₁ E ₂ J ₁ E ₃ V ₁	2 - ^c	4
576.17724	576.17838	-1.98	a b	D ₃ E ₂ D ₂ E ₂ S ₁ D ₃ E ₁ T ₁	2	4
634.25594	634.25663	-1.09	y	D ₁ E ₃ J ₁ E ₄ V ₁	1 - ^c	4
705.22039	705.22097	-0.82	a b	D ₃ E ₃ D ₂ E ₃ S ₁ D ₃ E ₂ T ₁	- ^c 1 - ^c	4
763.29942	763.29922	0.26	b y	E ₃ J ₁ M ₂ D ₁ E ₄ J ₁ E ₅ V ₁	1 - ^c	- ^c 2
834.26340	834.26356	-0.19	a b	D ₃ E ₄ D ₂ E ₄ S ₁	- ^c	2
850.33130	850.33125	0.06	a b y	D ₁ E ₃ J ₁ M ₃ E ₃ J ₁ M ₂ S ₁ D ₁ E ₄ J ₁ S ₁ E ₅ S ₁ V ₁	1 - ^c	- ^c 2
933.33114	933.33198	-0.90	b	D ₃ E ₃ J ₁ S ₁ D ₂ E ₄ S ₁ V ₁	- ^c	2

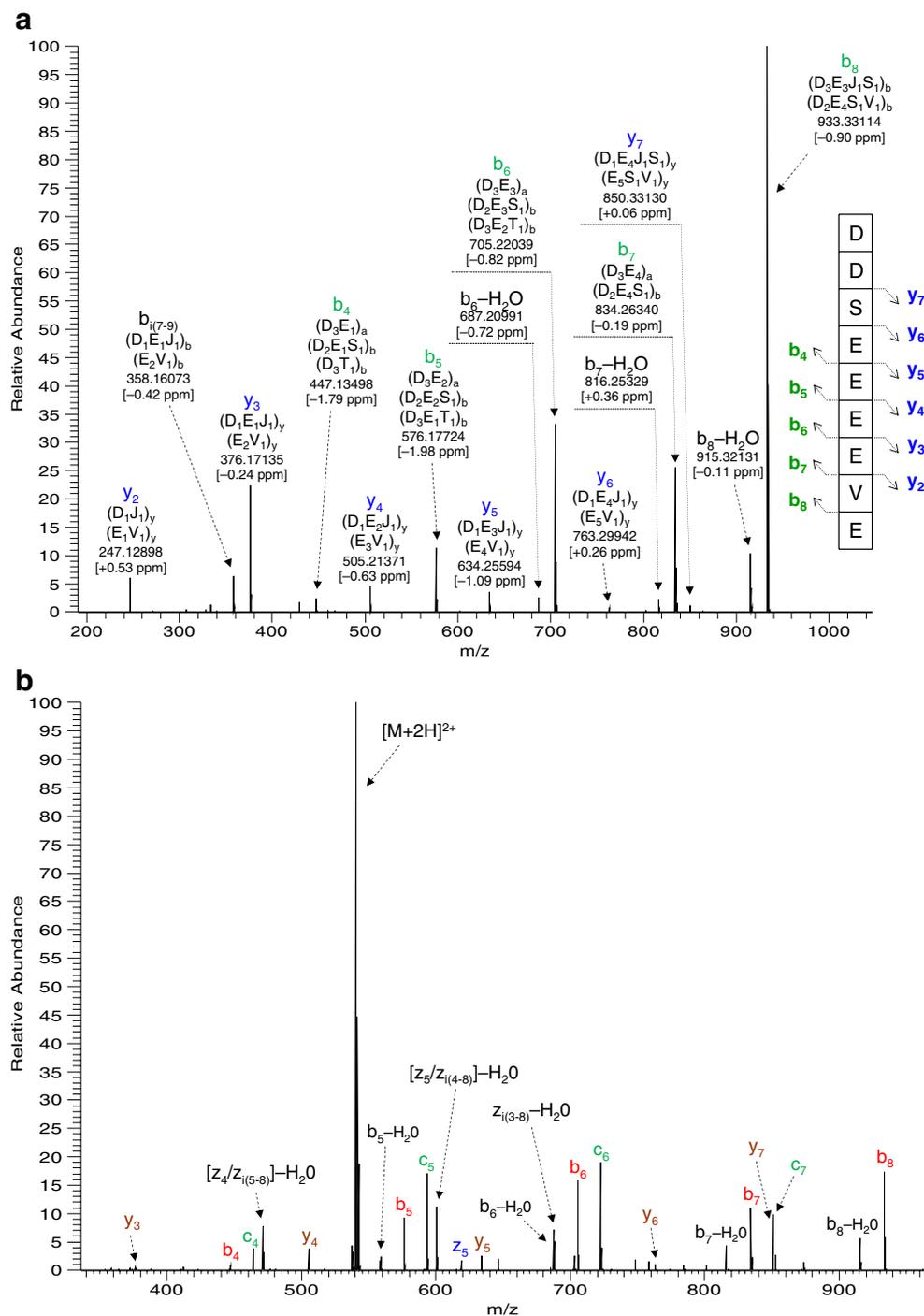
Each composition is given in alphabetical order with the corresponding ion type

^a Checked against precursor AAC list that was updated in each iteration step

^b Checked against fragment AAC list after each fragment calculation

^c The corresponding fragment ion compositions were not found within the peptide composition list of the respective structure

Fig. 2 CID (a) and ECD (b) spectrum of savignin 1. Composition analysis results (CID spectrum) are given in brackets with indicated fragment types. Fragment ion signals were assigned after sequence determination (colored labels)



form a lactam after cyclization and elimination of water. Since this amino acid is found at the C-terminus of savignin 1, we propose an analogous mechanism (Scheme 1) derived from the mentioned cyclization reaction. After the nucleophilic attack of the amide nitrogen to the side chain carboxyl group, pyroglutamic acid is split off as a neutral compound and an internal z ion remains.

Savignin 2

The CID spectrum of savignin 2 (Electronic supplementary material (ESM) Fig. S1) showed a series of internal fragment ions of types 'b' and 'y'. The ECD (Fig. 3) spectrum instead expressed a lower complexity and was used for sequence determination. The sequence could be

Table 4 CBS results of savignin 1 using CID fragment ions

Score	Sequence (free acid)	No. of observed ions matched
0.7665	DDSEEEVE	15
0.5877	DDESEEEVE	14
0.5442	SDDEEEVE	14
0.4098	DDSEEEEV	13
0.5880	DDSEEEEDJ	13
0.5880	DDSEEEEDJ	13
0.5445	DDSEEEEDJE	13
0.5445	DDSEEEJDE	13

easily read from the spectrum and no CBS analysis was carried out. The z' and c ions gave complementary and complete sequence information. Signals of c_9 and z_9' could be resolved by FTICR-MS as shown in the spectrum. The doubly charged precursor mass value (637.31658 u) had a deviation of -1.83 ppm compared to the theoretical value (637.31720 u). The mass value of hydroxyproline (113.04768 u) is rather close to that of leucine/isoleucine (113.08407 u). Accurate mass data allowed us to identify hydroxyproline at the N-terminal end and gave the sequence P*EEVEERJK-OH. Mass data of the ECD measurement are given in Table 6.

Savignin 3

Sequence analysis of the fraction L peptide (savignin 3) turned out to be rather difficult. CBS analysis was carried out as described above. The results of composition analysis (precursor ion evaluation before fragment ion evaluation) are shown in Table 7 considering free acid, acid amide, and cyclic structure.

After precursor ion evaluation, the list of amino acid composition candidates was reduced as much as possible by an evaluation of all CID fragments detected by FT-MS (Fig. 4). MS³ analysis (MS/MS of 399.22360 u) gave an

Table 5 Fragment ions after ECD of savignin 1: experimental mass values are those determined after internal fine calibration

Fragment mass (u)		$\delta_{\text{exp/calc}}$ (ppm)	Fragment ion
Exp.	Calc.		
376.17100	376.17144	-1.18	y_3
447.13568	447.13578	-0.21	b_4
464.16201	464.16232	-0.68	c_4
471.18436	471.18475	-0.82	$[z_4'/z_{4(5-8)}]-H_2O$
505.21403	505.21403	-0.01	y_4
558.16778	558.16782	-0.07	b_5-H_2O
576.17842	576.17838	0.08	b_5
593.20486	593.20492	-0.09	c_5
600.22726	600.22735	-0.14	$[z_5'/z_{5(4-8)}]-H_2O$
618.23833	618.23791	0.67	z_5'
634.25562	634.25663	-1.59	y_5
687.20982	687.21041	-0.86	b_6-H_2O
687.25841	687.25938	-1.41	$z_{i(3-8)}-H_2O$
705.22031	705.22097	-0.94	b_6
722.24672	722.24751	-1.09	c_6
763.29765	763.29922	-2.06	y_6
816.25210	816.25300	-1.10	b_7-H_2O
834.26303	834.26356	-0.64	b_7
850.33039	850.33125	-1.01	y_7
851.28965	851.29010	-0.53	c_7
915.32091	915.32142	-0.56	b_8-H_2O
933.33139	933.33198	-0.63	b_8

additional fragment ion at 171.11270 u (not shown). Composition analysis gave a list of 43 fragment ions considered in the evaluation procedure (ESM Table S1). Peptide composition analysis resulted in 86 AAC candidates of acid amide, eight AAC candidates of free acid, and ten AAC candidates of cyclic peptides after fragment ion evaluation. The final amino acid composition was determined by increment analysis as H(D₁G₄J₆P₁T₁V₁)NH₂.

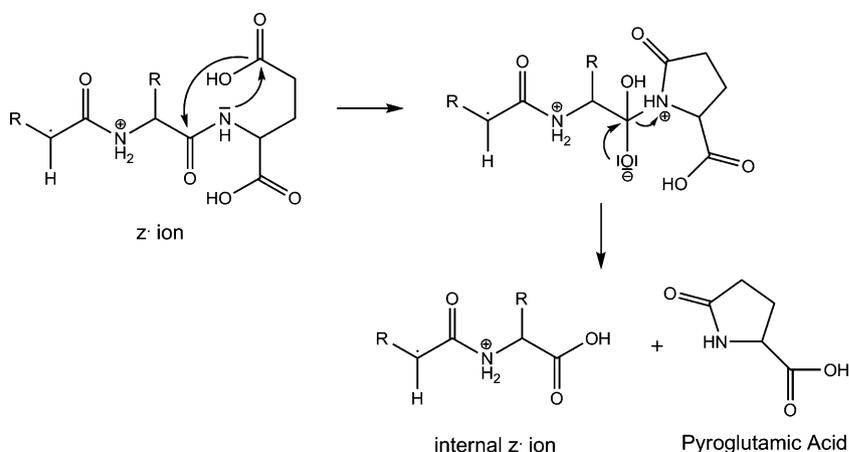
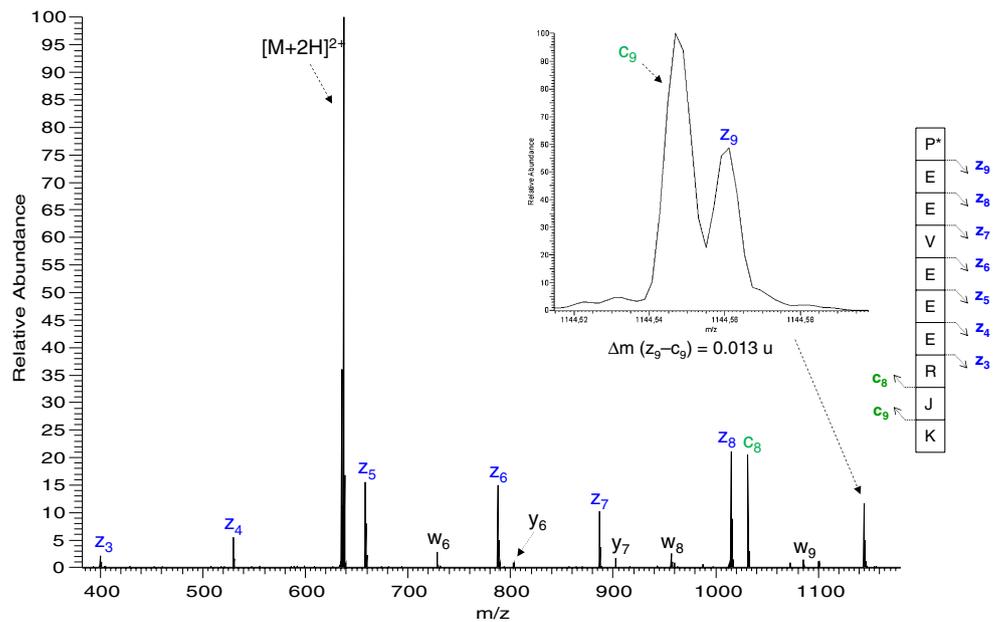
Scheme 1 Proposed mechanism of the internal z' ion formation after ECD of savignin 1

Fig. 3 ECD spectrum of savignin 2

Amino acids were identified by an evaluation of accurate increment mass values obtained in the MS/MS spectrum (see labeled sequence ladder in Fig. 5). The C-terminal amide structure was found based on a series of two-step fragment ions ($b_{i(x-14)}$) derived from y-type ions after loss of ammonia. The precursor mass value of Table 7 had a deviation of -1.15 ppm compared to the calculated value (668.91357 u). All fragment ions had a deviation within ± 2.0 ppm.

Knowing the amino acid composition of the sample peptide, the number of sequences to be tested is $P=k!$ (P is the number of permutations and k the peptide length). Therefore, the number of permutations (possible sequences) of $H(D_1G_4J_6P_1T_1V_1)NH_2$ is $P=14!=87,178,291,200$. Since the determined composition contains two multiplets of amino acids, the number of permutations reduces to $P = \frac{14!}{4! \cdot 6!} = 5045040$.

The number is reduced further by grouping of several amino acids. The intense b ion at 399.22360 u has the composition ($D_1G_1J_2$). The corresponding intense y ion at 938.60370 u includes the remaining amino acids of the peptide and is of composition ($G_3J_4P_1T_1V_1$). Within the latter AAC, the compositions ($G_2J_2V_1$) and (J_1P_1) were grouped additionally. The input string for permutational sequence analysis was then $\{[DGJJ][(GGJ)V]GJ(JP)T\}$ [38]. The number of sequences to be tested was thus $P = \frac{4!}{2!} [5! \cdot \frac{5!}{2! \cdot 2!} \cdot 2!] = 86400$.

CBS analysis gave sequence results with rather close score values (top 4 shown in Table 8). The highest and second highest scored sequence was tested in more detail by manual spectra analysis as discussed in the following. The number of matched fragment ions conforms to the number of assigned ion signals in the CID spectrum (+

fragment ion at 171.11270 u as described above) less the number of fragments that are rearrangement products and do not directly originate from the peptide sequence (fragment ion signals at 994, 1,109 and 1,220 u). Fragment ions b_5 and y_9 were not detected in the MS/MS spectrum. These ion signals would enable us to distinguish partial sequences “JP” and “PJ” at positions 5 and 6. The ECD spectrum (not shown) did not give additional sequence information. Responsible for different score values of the first two scored sequences was the b type fragment ion at

Table 6 Fragment ions after ECD of savignin 2: experimental mass values are those determined after internal fine calibration

Fragment mass (u)		$d_{\text{exp}/\text{calc}}$ (ppm)	Fragment ion
Exp.	Calc.		
400.27847	400.27925	-1.94	z_3
529.32139	529.32184	-0.85	z_4
658.36456	658.36443	0.19	z_5
728.39389	728.39426	-0.50	w_6
787.40726	787.40702	0.31	z_6
803.42571	803.42574	-0.04	y_6
886.47563	886.47543	0.22	z_7
902.49473	902.49415	0.65	y_7
956.50516	956.50526	-0.11	w_8
1015.51830	1015.51802	0.27	z_8
1031.46442	1031.46397	0.44	c_8
1085.54772	1085.54784	-0.11	w_9
1144.54860	1144.54804	0.49	c_9
1144.56156	1144.56060	0.84	z_9

Table 7 Number of possible AACs of savignin 3 (with a mass tolerance setting of ± 2.0 ppm) before fragment ion evaluation

Precursor [M+2H] ²⁺ (monoisotopic)	$\Delta m/z$ setting (ppm)	Number of possible AACs (outset)		
		Free acid	Acid amide	Cyclic
668.91280	2.00	787	1,057	1,076

808.49315 u having the composition (G₃J₃P₁T₁V₁). Regarding the top scored sequence, this fragment would be assigned to an internal b ion of sequence “PTGJVGGJJ” (b_{i(6-14)}) and to another internal b ion of sequence “JPTGJVGGJ” (b_{i(5-13)}). Both ions are mass identical due to identical AAC. Consequently, the fragment ions were (correctly) assigned twofold by CBS and the mass signal was matched twice. Manual analysis allows the assumption that the signal was composed of b_{i(5-13)} since b_{i(6-14)} is consistent with y₉-NH₃. The corresponding y₉ ion, however, was not observed. Within the second sequence candidate of score value 0.7924, only “PJTGJVGGJ” (b_{i(5-13)}) was a possible partial sequence of this ion signal and was matched only once. This explains the differently scored sequence candidates.

By a more detailed spectrum evaluation, we found that there is direct evidence in the CID spectrum for a partial sequence “PJ” (as in the second scored sequence candidate). The b type fragment ion at 994.62844 u of composition (G₄J₅T₁V₁) was found to be a rearrangement product (not matched by CBS) of the second scored sequence candidate and was manually identified as b₁₃-D-P based on accurate mass values. Fragment ion at 1109.65744 u of composition (D₁G₄J₅T₁V₁) was assigned as b₁₃-P accordingly. It is assumed that these ions were formed after cyclization of b₁₃ and ring opening C-terminal to proline. Elimination of P or DP, respectively, from the top scored sequence candidate is unlikely since D and P are not neighbored in this sequence. An additional support for the “PJ” instead of the “JP” subsequence is derived from an observed proline effect (described later).

Nevertheless, the top sequence candidate was synthesized as L⁶ and I⁶ peptide and their CID spectra were compared to that of the natural peptide (ESM Fig. S2). Both synthesized peptides provided identical CID spectra (I⁶ peptide not shown) and clearly disagreed with the MS/MS spectrum of the natural peptide, confirming that JGJDPTGJVGGJJ-NH₂ is not the correct sequence. Fragment ions b₅ and y₉, which were not observed in the spectrum of the natural peptide, predominate the spectrum of the synthetic peptide, again indicating that “JP” is not the correct subsequence. The CID spectrum of the natural peptide instead contains the y₁₀ ion signal with the highest intensity as a result of a proline effect [21]. Proline-containing peptides with an even number of amino acids N-terminal to proline show the feature that CID spectra are dominated by y-type ions with proline at the N-terminal

end [33] (called proline effect). This occurred in the MS/MS spectrum of the natural peptide (y_{n-4}, Fig. 4b), and the corresponding b₄ ion was detected with high abundance too. The corresponding y-type fragment ion of the synthetic peptide, on the other hand, carries a “J” at the N-terminal end and consequently is of low abundance. The observed behavior again confirms that the proline effect, which commonly suppresses other ion signals and sometimes impedes spectra evaluation, can on the other hand be very useful for sequence determination [37].

The N-terminal amino acid of the highly scored sequence candidates was leucine/isoleucine (sequence “J”), followed by glycine at the second position. This subsequence is untypical for reported frog peptides, while the subsequence “GJ” appears to be rather common [19]. To further investigate this inconsistency, we performed MS⁴ measurements of the b₂ ion (171.11270 u) that was generated by MS³ of b₄ (399.22360 u). We chose this fragmentation pathway to make sure that the fragmented ion at 171.11270 u was exclusively composed of b₂ ions and not of internal “GJ” fragments, as would have been possible in an MS² measurement. Fragmenting b₂ of the synthetic peptide (which had the sequence “JG”) resulted in the observation of the immonium ion of leucine/isoleucine (a₁) in high abundance (ESM, Fig. S3). Fragmentation of the b₂ ion of the natural peptide, on the other hand, gave no measurable signal of the corresponding immonium ion of glycine. It is thus obvious that the sequence at the N-terminal end of the sample peptide is “GJ” and not “JG”. The b-type ion at 286.13946 u of composition (G₁J₁D₁) (Fig. 4a) was matched as internal ion in CBS analysis, implying that the three amino acids were attached to each other. After subsequence “GJ” is now known, this fragment was identified as a non-direct sequence ion that misdirected the N-terminal structure analysis.

After the determination of the two questionable subsequences “GJ” and “PJ”, the final peptide sequence candidate was synthesized and CID spectra were compared (Fig. 5). The spectra of synthesized and natural peptide perfectly match and the sequence of savignin 3 was thus confirmed to be GJJDPTGJVGGJJ-NH₂.

Conclusions

Three novel, structurally unrelated peptides were reliably characterized by database-independent sequencing strate-

Fig. 4 CID spectrum of savignin 3 (**a** mass range 270–700 u; **b** mass range 700–1,000 u; **c** mass range 1,000–1,250 u). Fragments labeled in red are discussed in the text. Fragments at 994, 1,109, and 1,220 u are rearrangement products and are not matched by CBS

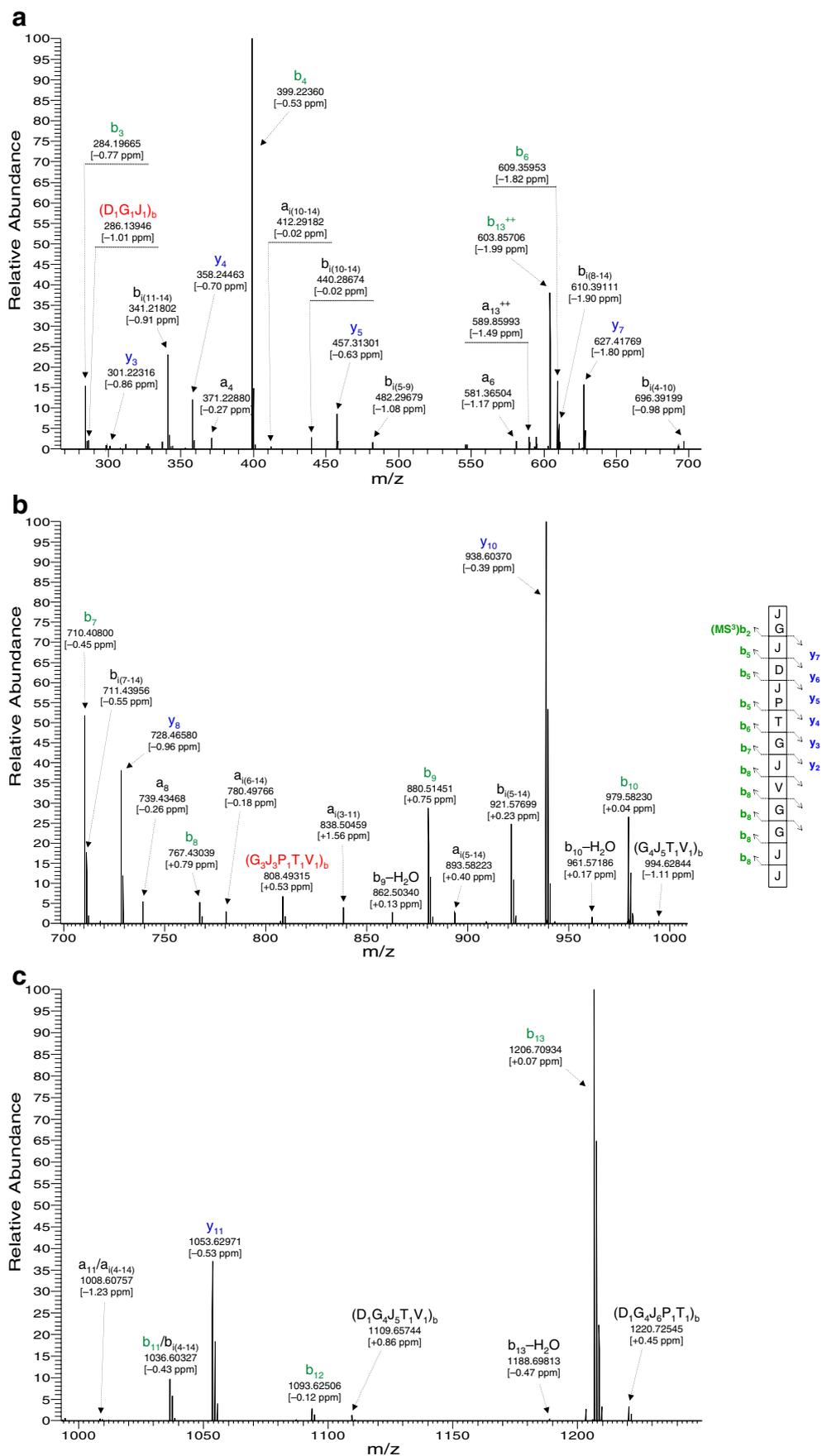
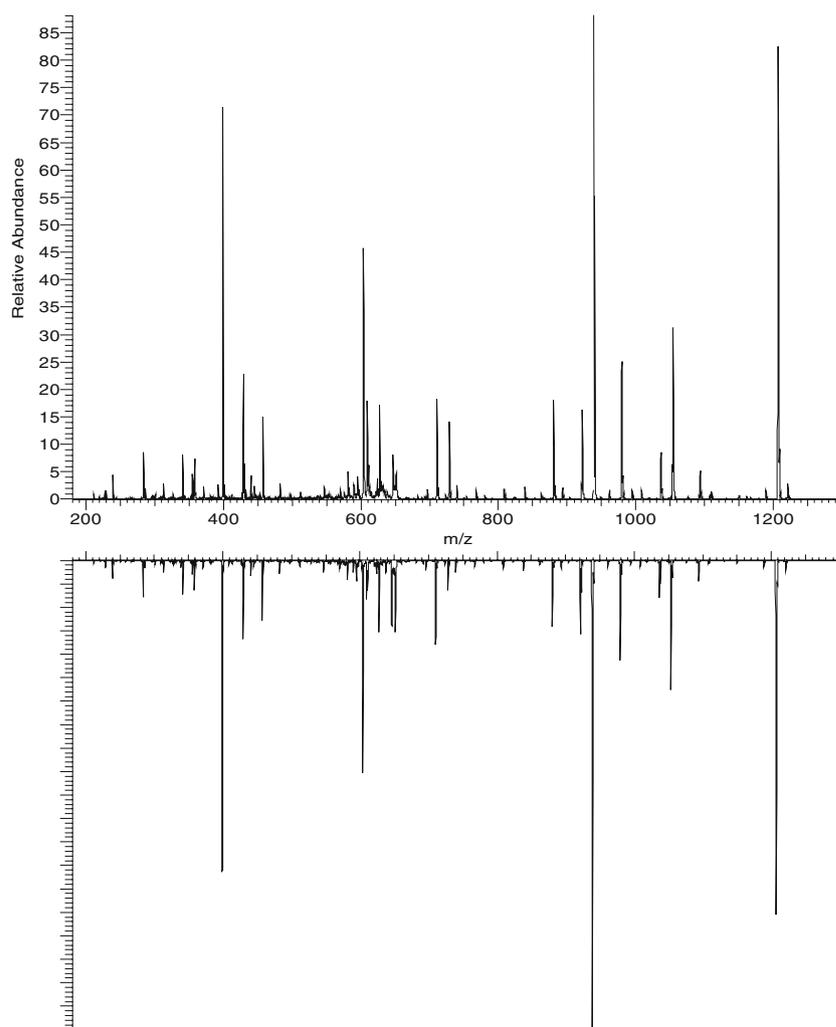


Fig. 5 CID spectra (ion trap) of natural (*above*) and of synthetic savignin 3 (*below*, with J = L). Confirmed sequence: GJJDPJTGJVGGJJ-NH₂



gies. Savignin 1 was easily sequence-analyzed by CBS using CID and ECD mass data as having the sequence DDSEEEVE-OH. The ECD spectrum unexpectedly showed a high number of b and y ions. Internal ions of z type were observed as well, and a mechanism of ion formation was proposed by us. Sequence determination of

savignin 2 was carried out based on ECD data, and the sequence could unequivocally be determined from the spectrum as P*EEVEERJK-OH. Hydroxyproline was identified and localized by accurate FT-MS analysis. Characterization of savignin 3 turned out to be rather difficult. In-depth sequence analysis by CBS and accurate mass increment evaluation led to the proposed sequence GJJDPJTGJVGGJJ-NH₂. This sequence could finally be verified by comparing the CID spectra of the natural peptide and the synthesized homologue.

Table 8 CBS results of savignin 3 (GJJDPJTGJVGGJJ-NH₂)

Score	Sequence (acid amide)	No. of observed ions matched
0.8009	JGJDPTGJVGGJJ	40
0.7924	JGJDPJTGJVGGJJ	40
0.7839	JGJDPTGJVJGJ	40
0.7839	JJGDPTGJVGGJJ	40

Sequences marked in red are discussed in the text

Sequence analysis of savignin 3 showed that the occurrence of rearrangement products and multiple assignable fragment ions of identical mass and composition are fundamental obstacles in de novo sequencing in general. A non-direct sequence ion led to a misinterpretation of the N-terminal sequence part of savignin 3. Another rearrange-

ment product, however, was helpful in identifying the sequential order of proline and leucine/isoleucine in the central part of the peptide chain.

The feasibility of the CBS method depends (as every sequencing strategy) on the availability of sufficient sequence information *and* on a limited complexity of MS/MS spectra. The software provides sequence results that are exclusively based on the mass calculation of fragment ions and precursor ion without employing database information. The process is experience-independent, only leaving the necessity to interpret the occurrence of rearrangement products and of the proline effect manually. However, the software performs a very comprehensive and exhaustive sequence analysis, including ion type characterization and composition analyses without “gaps” where all fragment ions are taken into account.

Savignin 1 and 2 are assumed to be non-active spacer or signal peptides [42, 43], being proteolytically cleaved off from a prepropeptide upon stimulation. Extensive series of bioactivity tests of savignin 3 on various pathogens will be implemented in the future.

Acknowledgments We would like to thank Faraham Ahmadzadeh and Hossein Hashempour from Shahid Beheshti University, Tehran, for collecting the skin secretions. Financial support by the Deutsche Forschungsgemeinschaft (DFG, Sp314/10-1) and by the Bundesministerium für Bildung und Forschung (BMBF, NGFN project 0313442) is gratefully acknowledged.

This publication represents a component of the doctoral (Dr. rer. nat.) thesis of Markus Langsdorf in the Faculty of Biology and Chemistry at the Justus Liebig University Giessen, Germany.

References

- Erspamer V, Erspamer GF, Mazzanti G, Endean R (1984) Active peptides in the skins of one hundred amphibian species from Australia and Papua New Guinea. *Comp Biochem Physiol* 77C:99–108
- Erspamer V, Erspamer GF, Cei JM (1986) Active peptides in the skins of two hundred and thirty American amphibian species. *Comp Biochem Physiol* 85C:125–137
- Habermehl G (1969) *Chemie und Biochemie von Amphibiengiften*. *Naturwissenschaften* 56:615–622
- Gomes A, Giri B, Saha A, Mishra R, Dasgupta SC, Debnath A, Gomes A (2007) Bioactive molecules from amphibian skin: their biological activities with reference to therapeutic potentials for possible drug development. *Indian J Exp Biol* 45:579–593
- Li J, Xu X, Xu C, Zhou W, Zhang K, Yu H, Zhang Y, Zheng Y, Rees HH, Lai R, Yang D, Wu J (2007) Anti-infection peptidomics of amphibian skin. *Mol Cell Proteomics* 6:882–894
- Nascimento ACC, Fontes W, Sebben A, Castro MS (2003) Antimicrobial peptides from anurans skin secretions. *Prot Pep Lett* 10:227–238
- Simmaco M, Mignogna G, Barra D (1998) Antimicrobial peptides from amphibian skin: what do they tell us? *Biopolymers* 47:435–450
- VanCompernelle SE, Taylor RJ, Oswald-Richter K, Jiang J, Youree BE, Bowie JH, Tyler MJ, Conlon JM, Wade D, Aiken C, Dermody TS, KewalRamani VN, Rollins-Smith LA, Unutmaz D (2005) Antimicrobial peptides from amphibian skin potently inhibit human immunodeficiency virus infection and transfer of virus from dendritic cells to T cells. *J Virol* 79:11598–11606
- Chinchar VG, Bryan L, Silphadaung U, Noga E, Wade D, Rollins-Smith L (2004) Inactivation of viruses infecting ectothermic animals by amphibian and piscine antimicrobial peptides. *Virology* 323:268–275
- Lu CX, Nan KJ, Lei Y (2008) Agents from amphibians with anticancer properties. *Anticancer Drugs* 19:931–939
- Shai Y (1999) Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by α -helical antimicrobial and cell non-selective membrane-lytic peptides. *Biochim Biophys Acta* 1462:55–70
- Matsuzaki K (1999) Why and how are peptide-lipid interactions utilized for self-defense? Magainins and tachyplesins as archetypes. *Biochim Biophys Acta* 1462:1–10
- Brogden KA (2005) Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? *Nat Rev Microbiol* 3:238–250
- Matyus E, Kandt C, Tieleman DP (2007) Computer simulation of antimicrobial peptides. *Curr Med Chem* 14:2789–2798
- Wang Z, Wang G (2004) APD: the antimicrobial peptide database. *Nucleic Acids Res* 32:D590–D592
- Wang G, Li X, Wang Z (2009) APD2: the updated antimicrobial peptide database and its application in peptide design. *Nucleic Acids Res* 37:D933–D937
- Koczulla AR, Bals R (2003) Antimicrobial peptides. *Drugs* 63:389–406
- Gordon YJ, Romanowski EG (2005) A review of antimicrobial peptides and their therapeutic potential as anti-infective drugs. *Curr Eye Res* 30:505–515
- Pukala TL, Bowie JH, Maselli VM, Musgrave IF, Tyler MJ (2006) Host-defence peptides from the glandular secretions of amphibians: structure and activity. *Nat Prod Rep* 23:368–393
- Apponyi MA, Pukala TL, Brinkworth CS, Maselli VM, Bowie JH, Tyler MJ, Booker GW, Wallace JC, Carver JA, Separovic F, Doyle J, Llewellyn LE (2004) Host-defence peptides of Australian anurans: structure, mechanism of action and evolutionary significance. *Peptides* 25:1035–1054
- Langsdorf M, Ghassempour A, Roempp A, Spengler B (2010) Characterization of a peptide family from the skin secretion of the middle east tree frog *Hyla savignyi* by composition-based de novo sequencing using positive and negative ion mass spectrometry. *Rapid Commun Mass Spectrom*. doi:10.1002/rem.4717
- Steinborner ST, Wabnitz PA, Waugh RJ, Bowie JH, Gao CW, Tyler MJ, Wallace JC (1996) The structures of new peptides from the Australian red tree frog *Litoria rubella*. The skin peptide profile as a probe for the study of evolutionary trends of amphibians. *Aust J Chem* 49:955–963
- Wabnitz PA, Bowie JH, Wallace JC, Tyler MJ (1999) Peptides from the skin glands of the Australian buzzing tree frog *Litoria electrica*. Comparison with the skin peptides of the red tree frog *Litoria rubella*. *Aust J Chem* 52:639–645
- Renda T, D’Este L, Buffa R, Usellini L, Capella C, Vaccaro R, Melchiorri P (1985) Tryptophyllin-like immunoreactivity in rat adenohypophysis. *Peptides* 6:197–202
- Papayannopoulos IA (1995) The interpretation of collision-induced dissociation tandem mass spectra of peptides. *Mass Spectrom Rev* 14:49–73
- Horn DM, Zubarev RA, McLafferty FW (2000) Automated de novo sequencing of proteins by tandem high-resolution mass spectrometry. *PNAS* 97:10313–10317
- Samgina TY, Artemenko AK, Gorshkov VA, Poljakov NB, Lebedev AT (2008) Oxidation versus carboxamidomethylation of S–S bond in Ranid frog peptides: pro and contra for de novo MALDI-MS sequencing. *J Am Soc Mass Spectrom* 19:479–487

28. Chen W, Lee PJ, Shion H, Ellor N, Gebler JC (2007) Improving de novo sequencing of peptides using a charged tag and C-terminal digestion. *Anal Chem* 79:1583–1590
29. Thieu VA, Kirsch D, Flad T, Mueller C, Spengler B (2006) Direct protein identification from nonspecific peptide pools by high-accuracy MS data filtering. *Angew Chem Int Ed* 45:3317–3319
30. Spengler B, Luetzenkirchen F, Metzger S, Chaurand P, Kaufmann R, Jeffery W, Bartlet-Jones M, Pappin DJC (1997) Peptide sequencing of charged derivatives by postsorce decay MALDI mass spectrometry. *Int J Mass Spectrom Ion Process* 169/170:127–140
31. Harrison AG, Young AB, Bleiholder C, Suhai S, Paizs B (2006) Scrambling of sequence information in collision-induced dissociation of peptides. *J Am Chem Soc* 128:10364–10365
32. Jia C, Qi W, He Z (2007) Cyclization reaction of peptide fragment ions during multistage collisionally activated decomposition: an inducement to lose internal amino-acid residues. *J Am Soc Mass Spectrom* 18:663–678
33. Unnithan AG, Myer MJ, Veale CJ, Danell AS (2007) MS/MS of protonated polyproline peptides: the influence of N-terminal protonation on dissociation. *J Am Soc Mass Spectrom* 18:2198–2203
34. Harrison AG, Young AB (2005) Fragmentation reactions of deprotonated peptides containing proline. The proline effect. *J Mass Spectrom* 40:1173–1186
35. Grewal RN, El Aribi H, Harrison AG, Siu KWM, Hopkinson AC (2004) Fragmentation of protonated tripeptides: the proline effect revisited. *J Phys Chem B* 108:4899–4908
36. Schwartz BL, Bursey MM (1992) Some proline substituent effects in the tandem mass spectrum of protonated pentaalanine. *Biol Mass Spectrom* 21:92–96
37. Kaufmann R, Chaurand P, Kirsch D, Spengler B (1996) Post-source decay and delayed extraction in matrix-assisted laser desorption/ionization-reflectron time-of-flight mass spectrometry. Are there trade-offs? *Rapid Commun Mass Spectrom* 10:1199–1208
38. Spengler B (2004) De novo sequencing, peptide composition analysis, and composition-based sequencing: a new strategy employing accurate mass determination by Fourier transform ion cyclotron resonance mass spectrometry. *J Am Soc Mass Spectrom* 15:703–714
39. Roepstorff P, Fohlmann J (1984) Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed Mass Spectrom* 11:601
40. Johnson RS, Martin SA, Biemann K (1988) Collision-induced fragmentation of $(M+H)^+$ ions of peptides. Side chain specific sequence ions. *Int J Mass Spectrom Ion Proc* 86:137–154
41. Schlosser A, Lehmann WD (2002) Patchwork peptide sequencing: extraction of sequence information from accurate mass data of peptide tandem mass spectra recorded at high resolution. *Proteomics* 2:524–533
42. Schatz G, Dobberstein B (1996) Common principles of protein translocation across membranes. *Science* 271:1519–1526
43. Pfanner N, Geissler A (2001) Versatility of the mitochondrial protein import machinery. *Nat Rev Mol Cell Biol* 2:339–349