RAPID COMMUNICATIONS IN MASS SPECTROMETRY
*Rapid Commun. Mass Spectrom.* 2010; **24**: 2885–2899
Published online in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/rcm.4717

RCM

# Characterization of a peptide family from the skin secretion of the Middle East Tree Frog *Hyla savignyi* by composition-based *de novo* sequencing

**Markus Langsdorf[1], Alireza Ghassempour[2], Andreas Römpp[1] and Bernhard Spengler[1]***

[1]Institute of Inorganic and Analytical Chemistry, Justus Liebig University, Giessen, Germany
[2]Medicinal Plants and Drugs Research Institute, Shahid Beheshti University, Tehran, Iran

Received 18 May 2010; Revised 16 July 2010; Accepted 20 July 2010

**A new tryptophyllin-like peptide family was found in the skin secretion of the tree frog *Hyla savignyi*. Peptides were characterized by database-independent sequencing strategies and specific ion fragmentation features were investigated. Skin secretions from specimens of *Hyla savignyi* were collected by mild electrical stimulation. Peptides were separated by reversed-phase nano-high-performance liquid chromatography (nanoHPLC) and mass spectra were acquired online by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS). Peptides were characterized by manual *de novo* sequencing and by composition-based sequencing (CBS), appearing mostly as C-terminal free acids and as their acid amide analogs. Amide peptides yielded lower intensities of y-type ions after collision-induced dissociation (CID) than their acid analogs. A mechanism of internal b-ion formation (positive ion mode) and of $CO_2$ elimination (negative ion mode) is proposed. We also exemplified phenomena such as the proline effect and formation of non-direct sequence ions after sequence rearrangements. The occurrence of rearrangement products, of internal ions and of the proline effect made the CID spectra highly complex. CBS analysis nevertheless resulted in successful and highly reliable sequence analysis. Copyright © 2010 John Wiley & Sons, Ltd.**

In recent years, amphibians have emerged as valuable model systems in medical applications, since the demand for bioactive agents in pharmaceutics and antibiotic therapy is still increasing. Skin secretions of the dorsal glands of amphibians provide a rich chemical pool of host-defence compounds, which are part of the animal's own defence system against microorganisms and predators. There are two types of glands in the amphibian skin, mucous and granular (venom) glands.[1] The mucous secretion is responsible for thermoregulation and for protection against predators. The granular glands produce a secretion containing various compounds of biological activity. The main secretion components are peptides,[2,3] alkaloids[4]and amines.[5]

Amphibian peptides exhibit a wide range of biological activities, such as antimicrobial,[6–9] antiviral,[10,11] anticancer[12]and spermicidal.[13] 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 diverging mechanistic models.[14–17] The application of such peptides in mammalian systems has to overcome barriers resulting from pharmacokinetic and pharmacodynamic aspects, instability against proteases, systemic and local toxicity or reduced activity based on salt and pH sensitivity.[18,19]

A number of amphibian peptides have been characterized in the past, predominantly from Australian species. The structures and bioactivity of these compounds were reviewed recently.[20,21] A sequence database of antimicrobial peptides from various biological sources was set up in 2003[22] and was updated recently.[23]

Tryptophyllin-like peptides were first isolated from *Phyllomedusa rohdei*.[24] Based on structural similarities, Chen *et al*. classified tryptophyllin-like peptides into three different structural groups, T-1 (8 amino acids), T-2 (4–7 amino acids) and T-3 (13 amino acids).[25] T-2 peptides that are structurally related to the peptide family of our studies were found in the Australian red tree frog *Litoria rubella* and in the Australian buzzing tree frog *Litoria electrica* and are called tryptohyllins L.[26,27] The tryptohyllins L were reported to have no significant antimicrobial activity. Recent studies, however, showed that two major peptides of this group have opioid activity.[28]

*Hyla savignyi* (Fig. 1) is a frog species belonging to the Hylidae family. Specimens of *Hyla savignyi* are mainly found on the Arabian Peninsula and in adjacent regions. Frogs used in this study were collected in the southern part of Meshkin Shahr in the Ardabil province of Iran.

*Correspondence to: B. Spengler, Institute of Inorganic and Analytical Chemistry, Justus Liebig University, 35392 Giessen, Germany.
E-mail: Bernhard.Spengler@anorg.chemie.uni-giessen.de

**Figure 1.** *Hyla savignyi* (young animal).

*De novo* sequencing[29–32] has become a highly efficient method in the field of amphibian peptidomics, since novel peptides from glandular secretions of amphibians are in general not listed in protein/peptide sequence databases. Peptides therefore have to be analyzed by database-independent sequencing strategies. Non-database-assisted peptide sequencing often works reliably and fast, if tryptic peptides having basic residues at the C-terminus are investigated by tandem mass spectrometry (MS/MS). If, on the other hand, peptides are not digestion products, have non-basic C-termini or if the cleavage enzyme is unknown, this procedure often fails.[33] Basic amino acids in the central part of a peptide sequence may cause highly complex fragmentation spectra.[34] Incomplete sequence ladders and post-translational modifications can also be an obstacle to finding the complete and correct sequence. Rearrangement processes in collision-induced dissociation (CID) experiments, comprising peptide cyclization and subsequent ring opening at various peptide bonds, are known to form non-direct sequence ions, leading to ambiguous interpretations.[35,36] The fragmentation of proline-containing peptides can generate characteristic MS/MS spectra, depending on the position of proline along a peptide chain,[37–41] and may also complicate peptide characterization.

The peptides investigated in this study were characterized by manual *de novo* sequencing and by composition-based sequencing (CBS).[42] The latter method is a database-independent sequencing strategy and includes two steps (Fig. 2). In the first step, the amino acid composition (AAC) is determined by 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 observed and expected product ion signals of permuted sequence propositions. Successful CBS analysis requires sufficient mass accuracy and (as every sequencing strategy) sufficient product ion information.

The aim of this paper is to elucidate the complexity of peptide fragmentation patterns of frog skin peptides and to employ accurate mass values in sequence determination, resulting in simplified and reliable spectrum interpretation and database-independent peptide sequencing.

## EXPERIMENTAL

### Collection of skin secretions

Adult specimens of *Hyla 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, the animals were manually and gently submerged for an additional 60 s in the beaker to complete the extraction of the skin secretion. The resulting aqueous solution was lyophilized with a freeze-drying device and stored at −20°C. Animals were released back into their natural habitat after treatment.

### HPLC separation

Lyophilized samples (1 mg) were dissolved in 500 μL water containing 5% acetonitrile. An insoluble residue was removed by centrifugation. For nanoHPLC-MS analysis, a binary nano-HPLC pump/autosampler system (Ultimate, Dionex/LCPackings, Idstein, Germany) was employed. A volume of 1 μL of a 1:10 dilution of the sample 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 Å; i.d. 75 μm; length 150 mm, Dionex). A multi-step 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 μL/min. The eluent was monitored by ultraviolet absorbance at 214 nm.

A Summit microHPLC system (Dionex) was used for fractionation and 10 μL were loaded onto a Microbore column (Macherey-Nagel, Düren, Germany, C18, 5 μm, 300 Å, i.d. 1 mm, length 150 mm). The flow rate was 40 μL/min. The multi-step gradient and solvents were used as in the nanoHPLC separation.

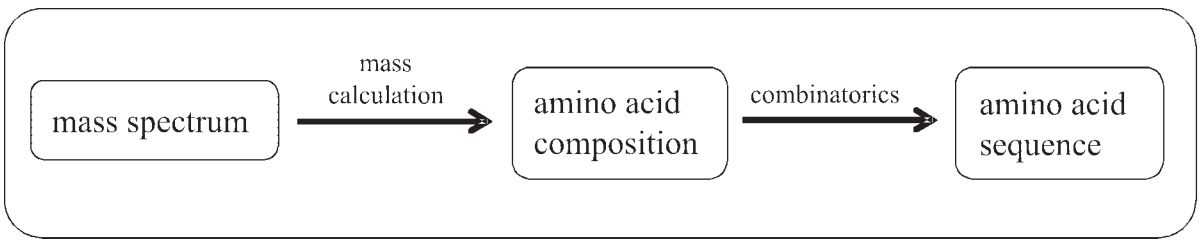



**Figure 2.** Composition-based sequencing (CBS) scheme.





DOI: 10.1002/rcm

### Mass spectrometry

Mass spectra were acquired on a Finnigan LTQ FT Ultra hybrid instrument (Thermo Scientific, Bremen, Germany) consisting of a linear quadrupole ion trap and a Fourier transform ion cyclotron resonance mass spectrometer with a 6 Tesla magnet. A nanospray ionization source was used for analyses (capillary temperature 250°C, capillary voltage 10 V, tube lens voltage 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 on-line measurements, Silica Tips (Tip i.d. 10 μm, New Objective, Woburn, MA, USA) were used at a spray voltage of 1.9 kV. The 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. The mass accuracy after external calibration was better than ±2.0 ppm during both on-line and off-line measurements. The instrument-specific parameter settings for CID in the 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 (3 microscans per spectrum). Xcalibur 2.0 SR2 data acquisition software (Thermo Scientific) was employed.

### Sequence analysis

All calculations of accurate mass values of peptide ions and peptide product ions were performed using the computer program Peptide Composer 1.0 (copyright Bernhard Spengler 1996–2010). Peptide composition analysis, CBS and manual *de novo* sequencing are supported by this program. Calculations were performed with a mass tolerance value set to ±1.5 ppm for precursor ions and ±2.0 ppm for product ions. Based on possible compositions of product ions and the precursor ion, the 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 the calculations treated both amino acids as being equal (L/I = J). The number of all amino acids (including phosphorylation) was set to 'unknown'. At the start of the AAC calculations, peptides were treated as having one of three possible structures: free acid, acid amide or cyclic structure. A decision on this structural property was established as soon as 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, compositional analysis of the product ions (within the candidate AACs of the precursor) was carried out starting with the CID spectra signal of lowest mass value. The AAC list of the precursor ion was then checked against the AAC list of that product ion. This was executed in several iteration steps after each fragment mass calculation until the AAC of the peptide was determined unambiguously.

Lists of sequence propositions were generated from determined AACs by exhaustive permutation. After *in silico* fragmentation of sequence propositions, theoretical and observed product ions were compared and matching ions were registered by the CBS procedure. Each hit was weighed with a fragment-type specific factor, so that for example internal ions were considered to a lower extent than N- or C-terminal ions.

In addition, automated Edman sequencing was performed on HPLC fractions. An Applied Biosystems 491 Procise Sequencer (Foster City, CA, USA) with pulsed liquid-phase delivery was used.

### Syntheses of peptides

After sequence determination, peptides were synthesized by GL Biochem Ltd. (Shanghai, China) with a purity of >90%.

## RESULTS AND DISCUSSION

### Sequence analysis

The HPLC fractions shown in Fig. 3 are labeled 'A' to 'H'. Table 1 shows the results of the AAC analyses of peptides found in the HPLC fractions. Each precursor ion resulted in a number of possible AACs belonging to either a free acid, acid amide or cyclic peptide structure. Mass analysis of each product ion of type a, b, or y[43,44] gave a number of possible AACs, mostly only one possible AAC. All the precursor and product ions evaluated were singly charged. Compositional analysis of the product ions finally resulted in only one possible C-terminal structure of the respective peptide and only one possible AAC.

Seven peptides were separated and characterized by CBS. One additional precursor ion signal was found to contain two chromatographically and mass spectrometrically unseparated peptides, which could be separated and interpreted only by MS/MS sequence analysis (Table 2).

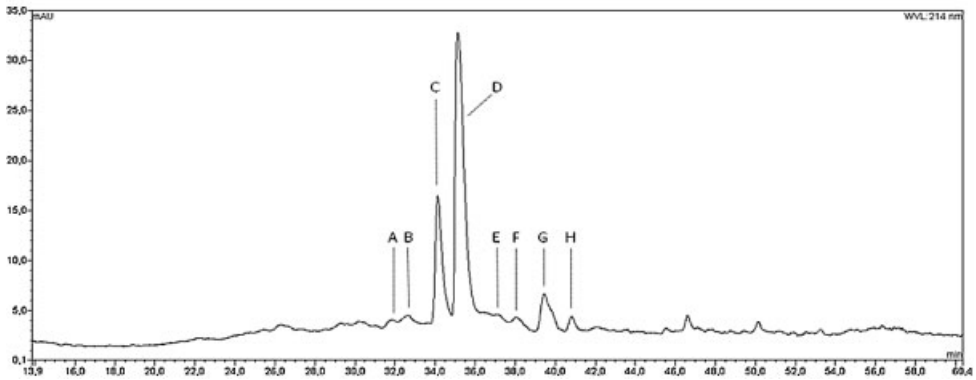



**Figure 3.** NanoHPLC separation of the skin secretion of *Hyla savignyi*.





DOI: 10.1002/rcm

**Table 1.** AAC result list of peptides found in HPLC fractions A–H. Compositions were determined with a CBS mass tolerance setting of 1.5 ppm and are listed in alphabetical order

| | *m/z* $[M+H]^+$ (monoisotopic) | | | | Structure | | |
|---|---|---|---|---|---|---|---|
| HPLC fraction | experimental | theoretical | Deviation [ppm] | AAC | free acid | acid amide | cyclic |
| A | 489.3075 | 489.3071 | 0.82 | $F_1J_2P_1$ | X | | |
| B | 488.3227 | 488.3231 | −0.82 | $F_1J_2P_1$ | | X | |
| C | 527.3339 | 527.3340 | −0.19 | $J_2P_1W_1$ | | X | |
| D | 561.3183 | 561.3184 | −0.18 | $F_1J_1P_1W_1$ | | X | |
| E | 528.3184 | 528.3180 | 0.76 | $J_2P_1W_1$ | X | | |
| F | 562.3027 | 562.3024 | 0.53 | $F_1J_1P_1W_1$ | X | | |
| G | 674.4023 | 674.4024 | −0.15 | $F_1J_2P_1W_1$ | | X | |
| H | 675.3858 | 675.3865 | −1.04 | $F_1J_2P_1W_1$ | X | | |

The nine peptides show sequence similarities with peptides found earlier in the skin secretion of the frogs *Litoria rubella* and *Litoria electrica*, which were named tryptophyllins L.[26,27] In this paper, we adapt this nomenclature to peptides of *Hyla savignyi*, naming them tryptophyllins H. Three sequences were obtained both as a free acid and a C-terminal acid amide. According to the classification of Chen *et al.*,[25] our peptide structures belong to the T-2 group of tryptophyllin-like peptides.

Edman degradation was mostly unsuccessful in independent sequence analysis, so that I/L was set to 'J' in general. We assume suppression effects as a possible reason, since the chromatogram showed a considerable background presumably caused by other compound classes that were not identified in this study. Electron capture dissociation (ECD), a promising method to distinguish Ile and Leu,[45,46] could not be applied in our studies because the peptide precursor ions unfortunately were detected as singly charged species only.

All peptide sequences (excepting tryptophyllin H 5) were determined by CBS analysis and were subsequently confirmed by CID measurements of analog synthetic peptides (setting J to I). As mentioned in the Experimental section, the CBS algorithm weighs each hit with a fragment-type specific factor. According to the accepted nomenclature,[43,44] product ions of type 'b' include the N-terminus of the peptide. Internal product ions with structures equivalent to b-type ions are formed by a two-step fragmentation process. They do not contain the N-terminus of the peptide and are labeled '$b_{i(x\text{-}y)}$' with x and y being the start and end positions of the terminating amino acids of the product ion within the peptide chain.[47] Product ions formed by rearrangement processes are labeled '$b_{r(\ldots)}$' if homologous to b-type ions. The same nomenclature is applied to a-type ions.

A nomenclature for product ions generated in the negative ion mode was proposed by Bowie *et al.*, who classified y-type ions as α ions containing the C-terminus and b-type ions as β ions containing the N-terminus.[48] In this paper, however, we maintain the Roepstorff and Fohlmann nomenclature[43] for product ions in the negative ion mode to keep the classification uniform.

Sequence determination by CBS was unequivocal in all cases and resulted in matching all observed product ion signals except those identified as rearrangement products. The number of matched product ions conforms to the number of assigned ion signals in the CID spectra (discussed later).

The CBS procedure worked efficiently, even if all product ions of various structural types and of the complete intensity range had to be interpreted. Manual *de novo* sequencing was employed only to back up the CBS results, but gave no new sequence information.

**Table 2.** Tryptophyllins H from the skin secretion of *Hyla savignyi*

| | | *m/z* $[M+H]^+$ (monoisotopic) | | | | Determined by | |
|---|---|---|---|---|---|---|---|
| Tryptophyllin | Sequence | experimental | theoretical | Deviation [ppm] | HPLC fraction | CBS (+) | (−) |
| H 1a | FPJJ−$NH_2$ | 488.3227 | 488.3231 | −0.82 | B (+traces in A) | yes | yes |
| H 1b | JJPF−OH | 489.3075 | 489.3071 | 0.82 | A (+traces in B) | yes | yes |
| H 2a | JPWJ−$NH_2$ | 527.3339 | 527.3340 | −0.19 | C | yes | no |
| H 2b | JPWJ−OH | 528.3184 | 528.3180 | 0.76 | E | yes | no |
| H 3a | JPFW−$NH_2$ | 561.3183 | 561.3184 | −0.18 | D (+traces in E,F) | yes | yes |
| H 3b | JPFW−OH | 562.3027 | 562.3024 | 0.53 | F | yes | yes |
| H 4a* | JJPFW−$NH_2$ | 674.4023 | 674.4024 | −0.15 | G (+traces in H) | yes | no |
| H 4b | JJPFW−OH | 675.3858 | 675.3865 | −1.04 | H | yes | no |
| H 5* | FJPWJ−$NH_2$ | 674.4023 | 674.4024 | −0.15 | G (+traces in H) | no | no |

* Peptides H 4a and H 5 could not be separated by LC or MS.
(+) = positive ion MS/MS (CID). (−) = negative ion MS/MS (CID). J = I or L.





DOI: 10.1002/rcm

*Tryptophyllins H 1a and H 1b*

Compositional analysis of peptide mass values from HPLC fractions A and B gave two AAC candidates in each case, one with a linear and one with a cyclic structure. Positive ion MS/MS spectra of both peptides are shown in Fig. 4. The AAC of the fraction B peptide (H 1a) was found to be $H(F_1J_2P_1)NH_2$ and this was confirmed after compositional analysis of product ion 211.1442 u. This ion was identified as a b-type ion of composition $(J_1P_1)$, so that the alternative, cyclic peptide composition $(F_1J_1K_1V_1)_{cyclic}$ could be discarded.

The b-type ion 227.1757 u of the peptide of fraction A (H 1b) has an AAC $(J_2)$ that fits both precursor composition candidates, $H(F_1J_2P_1)OH$ and $(J_2V_1Y_1)_{cyclic}$. Compositional analysis of product ion 245.1288 u gave a b-type ion of composition $(F_1P_1)$ pointing to peptide composition $H(F_1J_2P_1)OH$.

These two peptides were called H 1a and 1b since they were found to have the same AAC. The peptides, however, do not have identical amino acid sequences. The sequence of 1a (FPJJ) is in a reverse order to that of 1b (JJPF). This was

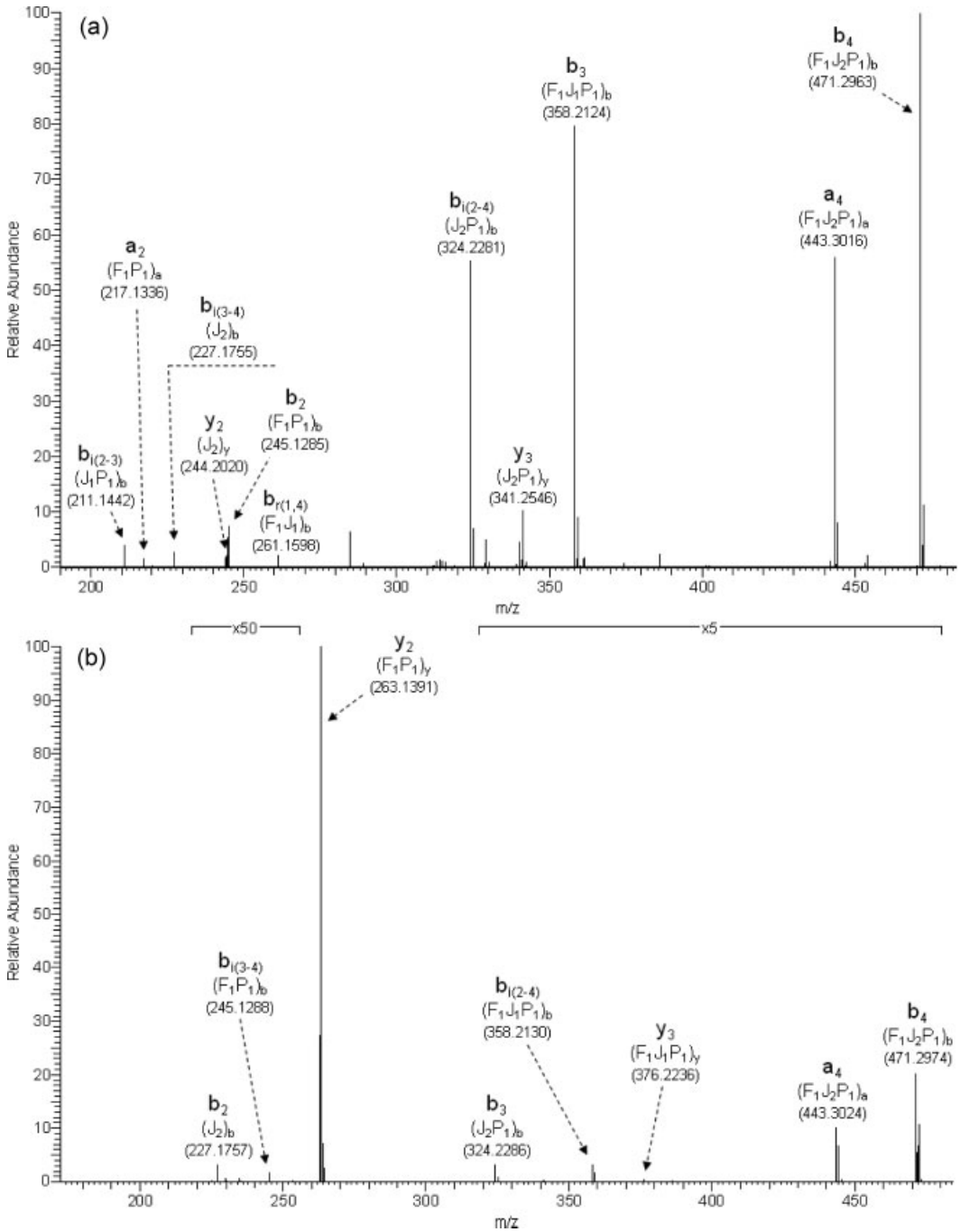



**Figure 4.** Positive ion MS/MS spectra of protonated (a) tryptophyllin H 1a ($FPJJ-NH_2$) and (b) tryptophyllin H 1b ($JJPF-OH$). Indicated masses are experimental data. Results from compositional analysis of product ions are given in brackets with fragment types as indices. Product ion identification after sequence analysis is given on top of labels.





DOI: 10.1002/rcm

RCM

first astonishing to us, especially since most other identified peptides were found to have the same sequence as a free acid and acid amide. The MS/MS spectrum of protonated H 1a does not show a predominant $y_2$ ion at *m/z* 262 resulting from a proline effect[49] (discussed later) as in the case of H 1b (*m/z* 263.1391). This indicates that proline is not located at the same position within the two peptides. The sequence of H 1a was finally verified by the synthetic analog providing the identical CID spectrum including the low-abundance rearrangement product ion at 261.1598 u. The sequence of H 1a was also confirmed by negative ion MS/MS through product ions $y_2$, $y_3$ and $b_3$, and the sequence of H 1b by the y ion series and by product ions $b_2$ and $b_3$ (Fig. 5).

### *Tryptophyllins H 2a and H 2b*

Compositional analysis of the peptide found in HPLC fraction C resulted in two AAC candidates, and of fraction E in three AAC candidates (Table 1). Positive ion CID spectra are shown in Fig. 6. Candidate composition $(J_1K_1V_1W_1)_{cyclic}$ of fraction C peptide H 2a conflicts with composition $(J_1P_1)$ of the b-type ion 211.1440 u and was thus discarded, so that AAC $H(J_2P_1W_1)NH_2$ remained. Fraction E peptide H 2b was found to have the same free acid composition as H 2a, based on determined composition $(P_1W_1)$ of a-type ion 256.1447 u, which excluded the cyclic candidate compositions $(A_1F_1J_1P_1V_1)_{cyclic}$ and $(F_1G_1J_2P_1)_{cyclic}$.

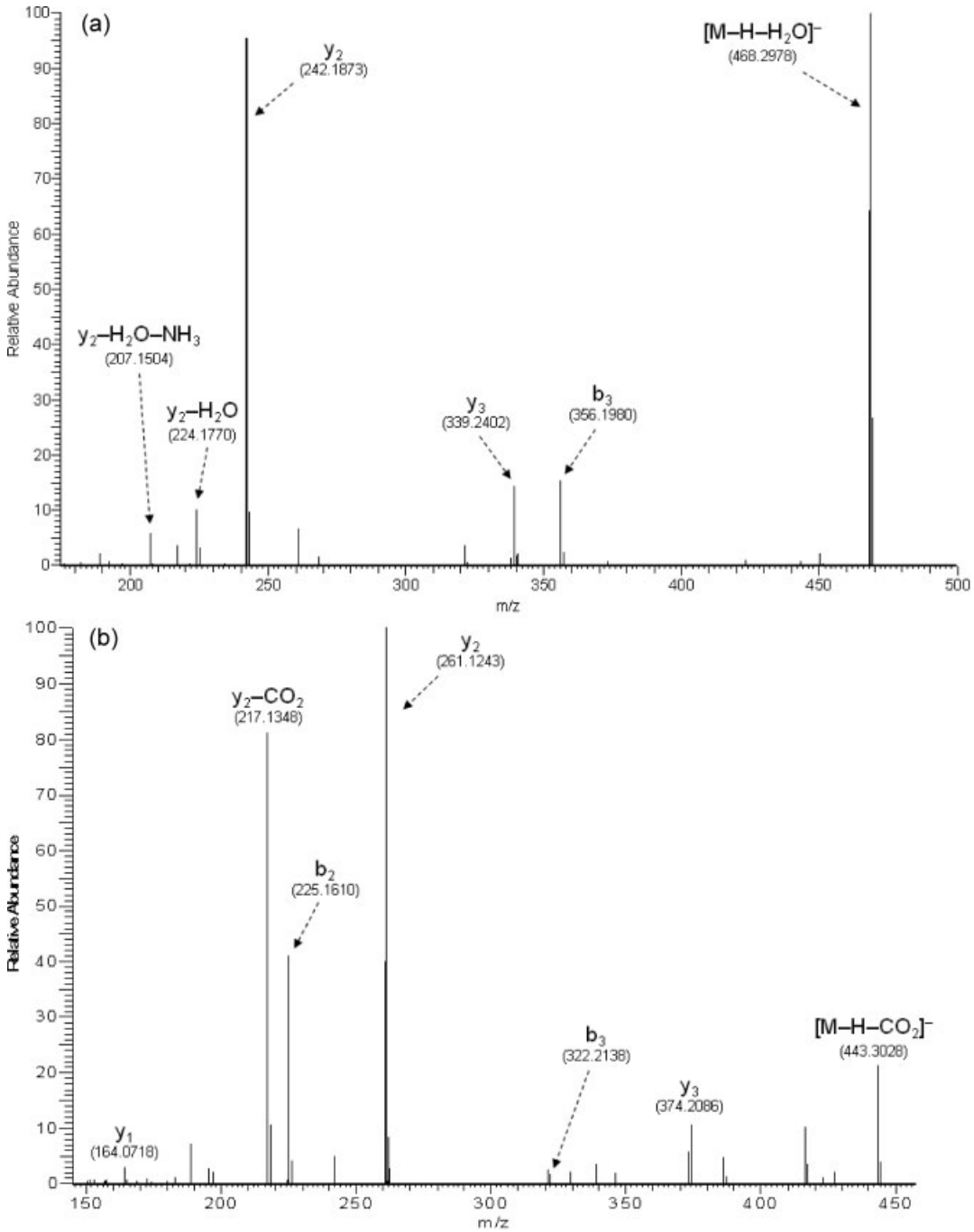



**Figure 5.** Negative ion MS/MS spectra of deprotonated (a) tryptophyllin H 1a ($FPJJ-NH_2$) and (b) tryptophyllin H 1b (JJPF−OH). Indicated masses are experimental data.





DOI: 10.1002/rcm

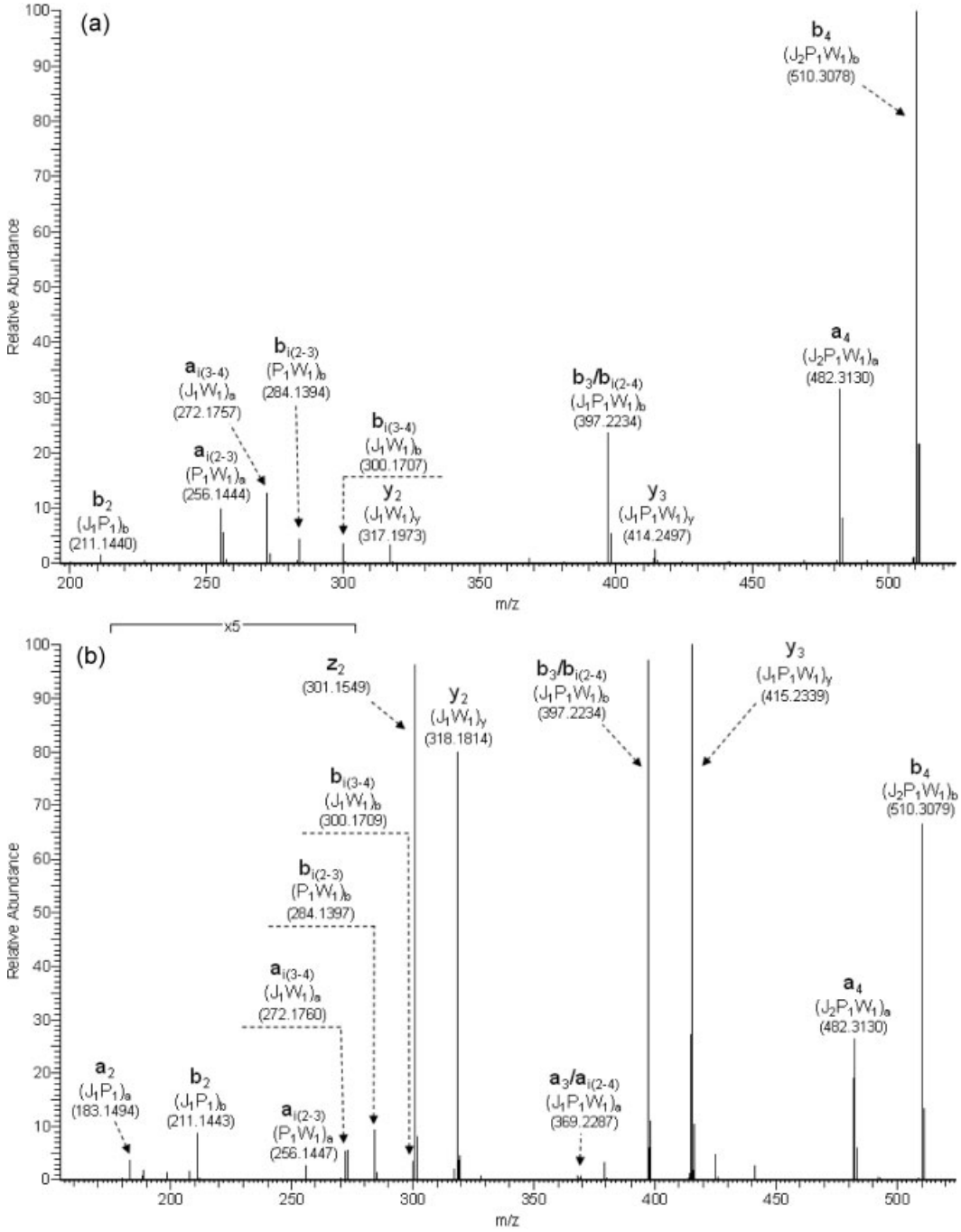



**Figure 6.** Positive ion MS/MS spectra of protonated (a) tryptophyllin H 2a (JPWJ$-NH_2$) and (b) tryptophyllin H 2b (JPWJ$-$OH). Indicated masses are experimental data. Results from compositional analysis of product ions are given in brackets with fragment types as indices. Product ion identification after sequence analysis is given on top of labels.

Manual spectra evaluation of tryptophyllins H 2a and 2b was carried out for confirmation. The sequence of H 2a conforms to that of tryptophyllin L 2.1 (IPWL$-NH_2$) of the frog *Litoria rubella*.[26,27] The negative ion MS/MS spectra of H 2a and 2b did not contain complete sequence information and are not shown here.

### *Tryptophyllins H 3a and H 3b*

Compositional analysis of the precursor mass of fraction D peptide H 3a resulted in four AAC candidates of acid peptides, 13 AAC candidates of amide peptides and one AAC candidate with a cyclic structure. Compositional analysis of ion 183.1494 u (Fig. 7) gave ($J_1P_1$) with an a-type structure, which conforms to peptide structure H($F_1J_1P_1W_1$)$NH_2$ only. Compositional analysis of fraction F peptide H 3b gave one AAC candidate as a free acid and three AAC candidates with cyclic structures. Two of the latter AACs could be excluded with the first fragment mass value of 183.1494 u (Fig. 7), conforming to an a-type ion of composition ($J_1P_1$). The remaining composition candidates



DOI: 10.1002/rcm

RCM

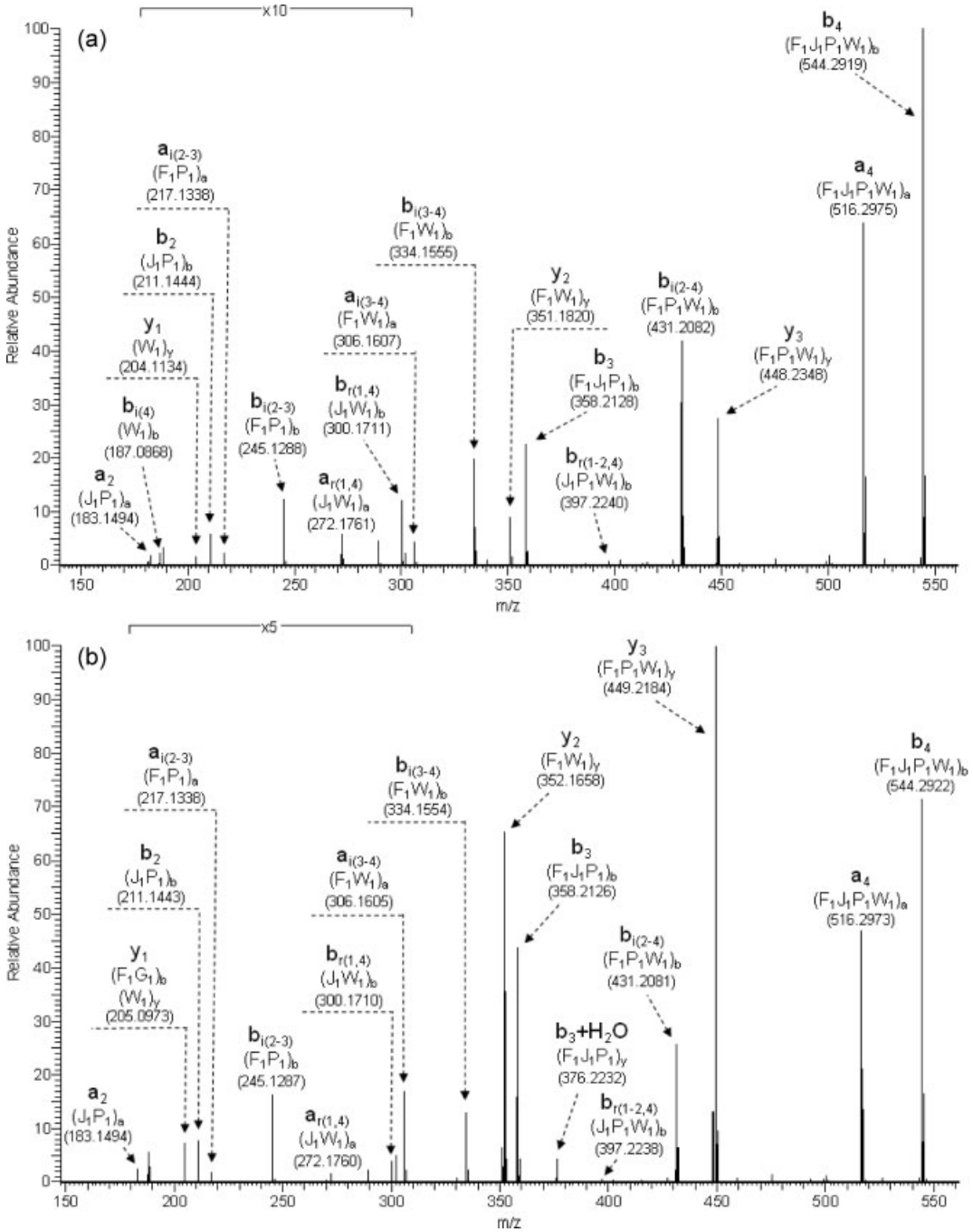



**Figure 7.** Positive ion MS/MS spectra of protonated (a) tryptophyllin H 3a (JPFW−$NH_2$) and (b) tryptophyllin H 3b (JPFW−OH). Indicated masses are experimental data. Results from compositional analysis of product ions are given in brackets with fragment types as indices. Product ion identification after sequence analysis is given on top of labels.

$H(F_1J_1P_1W_1)OH$ and $(F_2G_1J_1P_1)_{cyclic}$ competed until the sixth iteration, where fragment 272.1760 u was determined as a-type ion of composition ($J_1W_1$), thus confirming that the peptide has the same AAC as tryptophyllin H 3a as a C-terminal free acid. Ions formed by rearrangement processes (such as $a_{r(1,4)}$ and $b_{r(1,4)}$) impeded manual *de novo* spectra evaluation, but not compositional analysis. Rearrangement processes, leading to ions of type $a_r$ and $b_r$, were found to be preferred in amide peptides over acid peptides (discussed in the next section).

The results for tryptophyllin H 3a and H 3b were also checked against published sequence assignments of tryptophyllin L 1.2 (FPWL−$NH_2$) of *Litoria rubella* [26,27] as a possible candidate. Referring to the published structure, labeled $b_r$ ions of H 3a at 300.1711 u ($J_1W_1$) and 397.2240 u ($J_1P_1W_1$) would correspond to more common internal ions and the intense ion signal at 431.2082 u of composition ($F_1P_1W_1$) would relate to an N-terminal $b_3$ ion. We therefore synthesized this peptide and subjected it to CID measurements (Fig. 8). In addition, we synthesized all possible





DOI: 10.1002/rcm

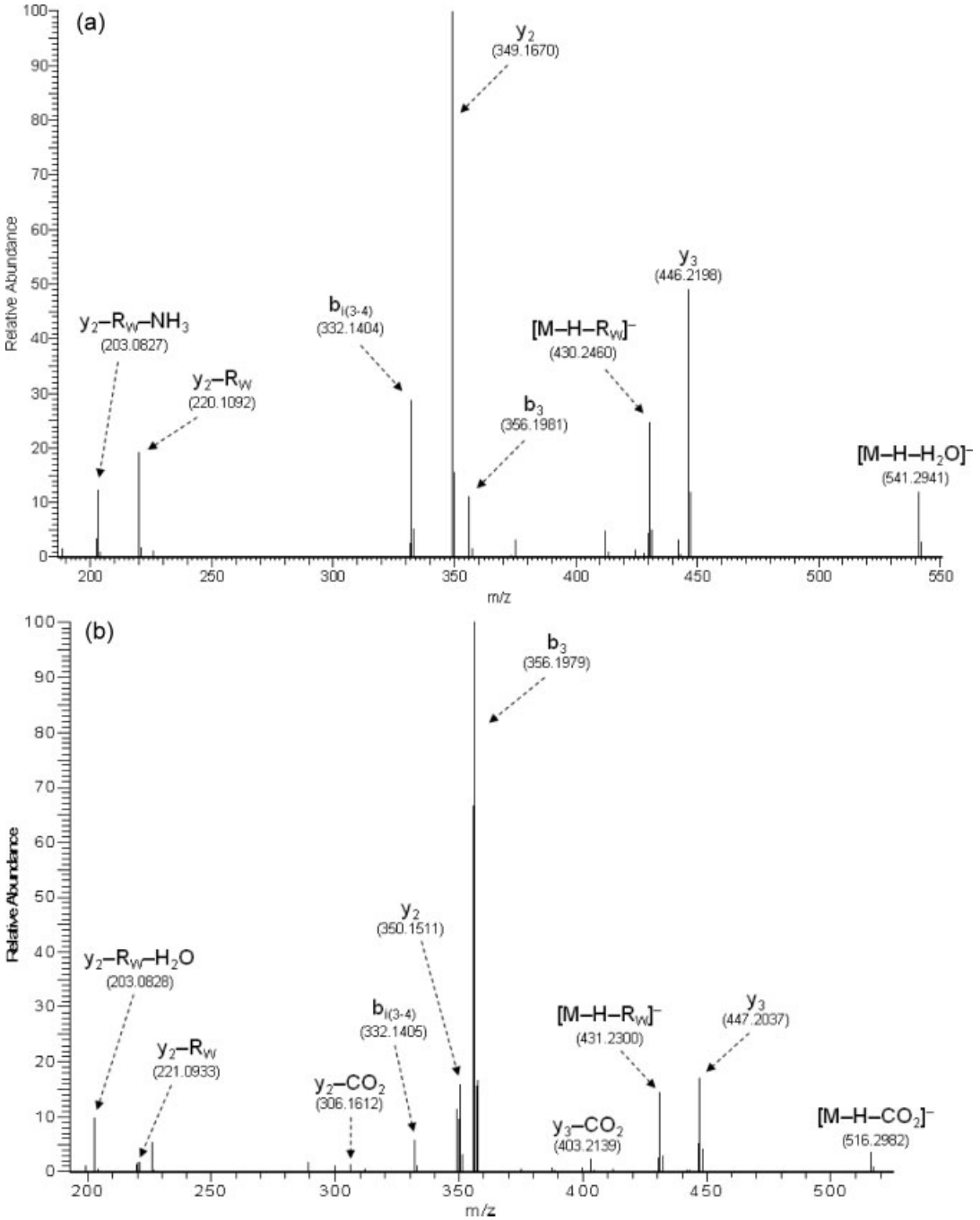



**Figure 8.** Negative ion MS/MS spectra of deprotonated (a) tryptophyllin H 3a (JPFW$-NH_2$) and (b) tryptophyllin H 3b (JPFW$-$OH). Indicated masses are experimental data. $R_W$ = side-chain residue of tryptophan ($C_9H_7N$, $\Delta m = 129$).

isomers of our determined sequence JPFW (setting J to I, see Scheme 2) and compared the CID spectra (see Supporting Information). Only the fragmentation pattern of sequence JPFW was in agreement with that of the natural peptide and the same rearrangement products were observed. Comparison with the pure synthetic peptide furthermore confirmed that the analyzed peptide of fraction D with an $[M+H]^+$ ion at *m/z* 561.3183 was a pure compound.

The peptides were also characterized by negative ion mass spectrometry. Tryptophyllin H 3a and 3b (Fig. 8) both showed the $y_2$, $y_3$ and $b_3$ product ions confirming the positive ion CBS results. Negative ion specific fragmentations included a loss of the side chain from tryptophan ($C_9H_7N$, $\Delta m = 129$, labeled as $y_n–R_W$) and a loss of $CO_2$ from the $[M-H]^-$ ion of peptide H 3b.

### *Tryptophyllins H 4a, H 4b and H 5*

Compositional analysis of the fraction G peptide (H 4a) resulted in seven AAC candidates belonging to a C-terminal free acid, 19 AAC candidates of amide peptides and four AAC candidates with cyclic structures. The total number of candidates could be reduced to one by the first iteration, when checking against product ion 245.1283 u with the AAC



DOI: 10.1002/rcm

RCM

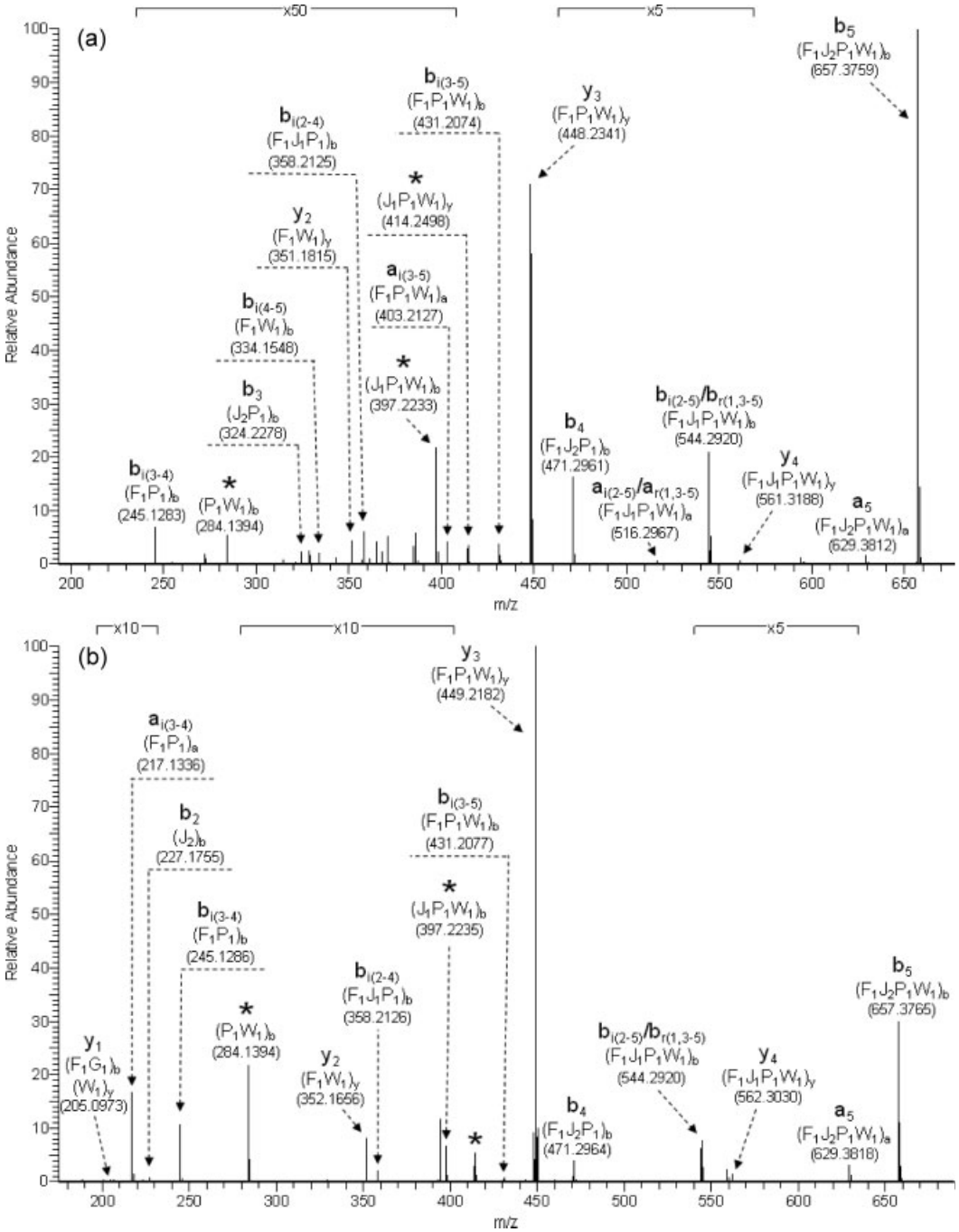



**Figure 9.** Positive ion MS/MS spectra of protonated (a) tryptophyllin H 4a (JJPFW−$NH_2$) and (b) tryptophyllin H 4b (JJPFW−OH). Indicated masses are experimental data. Results from compositional analysis of product ions are given in brackets with fragment types as indices. Product ion identification after sequence analysis is given on top of labels. Asterisks indicate product ions belonging to a second component tryptophyllin H 5.

$(F_1P_1)$ and b-type structure (Fig. 9). The peptide composition was found to be $H(F_1J_2P_1W_1)NH_2$.

Compositional analysis of fraction H peptide H 4b resulted in 19 AAC candidates for acid peptides, 33 for amide peptides and six for cyclic peptides. Considering product ion 205.0973 u, which had already been identified in tryptophyllin H 3b, reduced the AAC candidate list to $H(F_1J_2P_1W_1)OH$, $(F_2G_1J_2P_1)_{cyclic}$ and $(J_2V_1W_1Y_1)_{cyclic}$ for the theoretical mass of *m/z* 675.3865. The latter composition was eliminated when checking against product ion 217.1336 u, determined as an a-type ion of composition $(F_1P_1)$. The other cyclic peptide composition candidate was deleted after evaluation of the ion signal at *m/z* 284.1394. The b-type product ion of composition $(P_1W_1)$ determined the peptide composition finally as $H(F_1J_2P_1W_1)OH$.

The peptide sequences were determined by CBS scoring and were found to be those of tryptophyllins H 3a and 3b with an additional leucine/isoleucine at the N-terminus. The





DOI: 10.1002/rcm

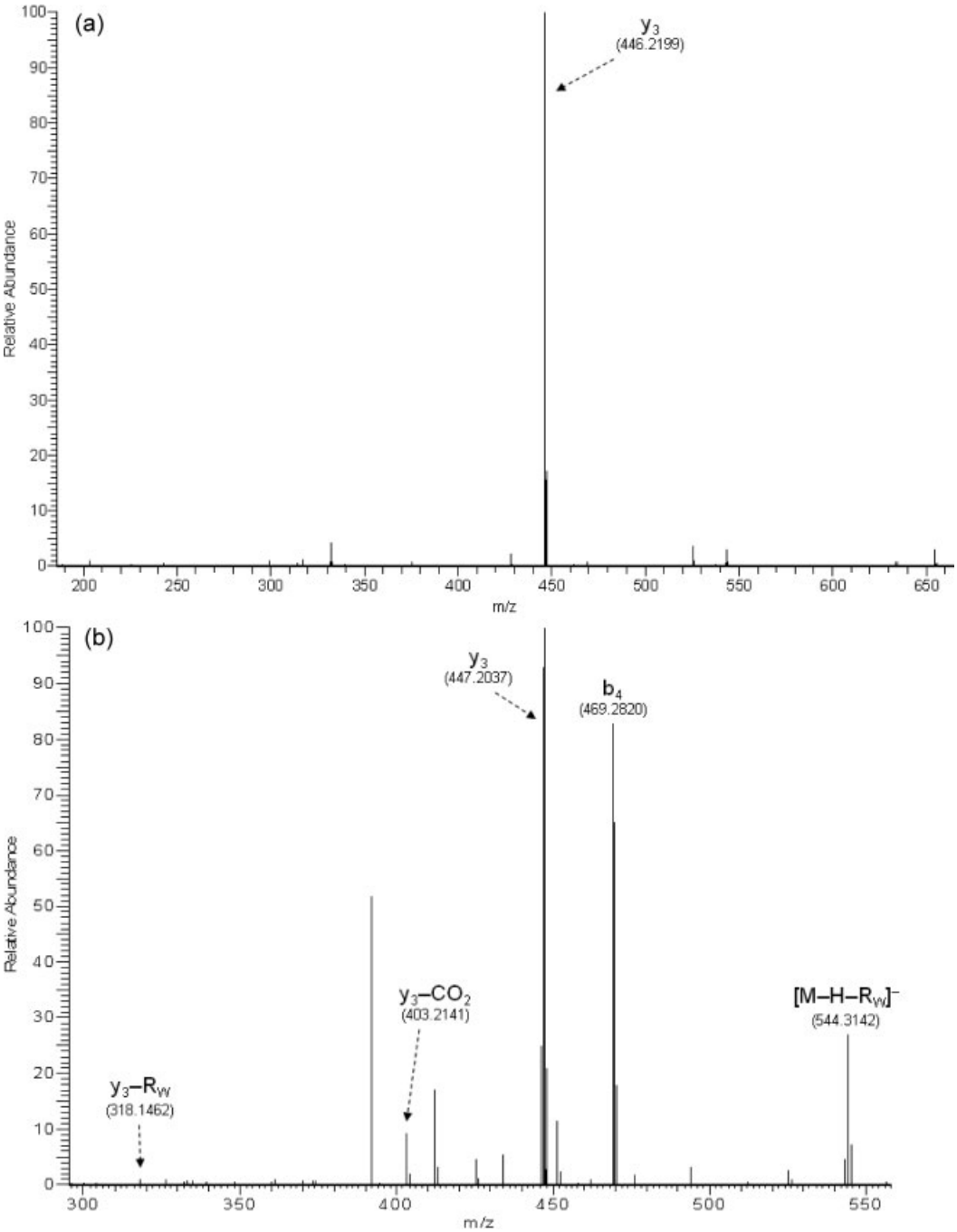



**Figure 10.** Negative ion MS/MS spectra of deprotonated (a) tryptophyllin H 4a (JJPFW$-NH_2$) and (b) tryptophyllin H 4b (JJPFW$-$OH). Indicated masses are experimental data. $R_W$ = side-chain residue of tryptophan ($C_9H_7N$, $\Delta m = 129$).

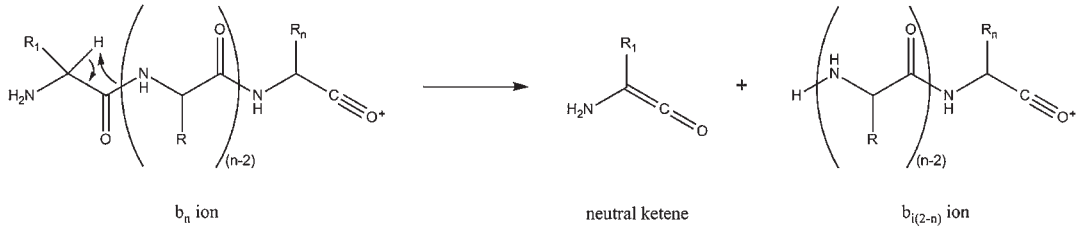
positive and negative ion CID spectra (Figs. 9 and 10) were found to differ strongly from those of H 3a and H 3b. The negative ion CID spectrum of H 4a impressively showed proline-related fragmentation behavior[48] resulting in only one intense product ion ($y_3$). Both spectra, of H 4a and H 4b, were useless for sequence interpretation.

Positive ion CID spectra of synthetic tryptophyllins H 4a and 4b (not shown) did not provide product ions at *m/z* 284, 397 and 414 as found for the precursor ion of the natural peptide. This indicated that the HPLC fraction contained a second peptide of the same mass and the same composition, but a different sequence, that was not separated from tryptophyllins H 4a and 4b. We could confirm this hypothesis after more detailed analysis of the nanoLC fraction G. The corresponding peak of the nanoLC separation (Fig. 3) had a shoulder at the right side. This shoulder peak showed the expected product ions of tryptophyllin H 4a and also the above-mentioned product ion signals not found in synthetic H 4a. The main peak, however, only provided product ions assigned to H 4a. The second peptide (tryptophyllin H 5) was thus found in this shoulder. For the CID spectrum of the fractionated microLC sample (Fig. 9), we



DOI: 10.1002/rcm

RCM




**Scheme 1.** Proposed mechanism for the N-terminal elimination of amino acid units in b-type ions (no oxazolone ring formation).

assigned a y-type ion at 414.2498 u as being of composition ($J_1P_1W_1$). 397.2233 u ($J_1P_1W_1$) is the corresponding y$-NH_3$ product ion while the b-type ion at 284.1394 u ($P_1W_1$) is another product ion belonging to tryptophyllin H 5 rather than H 4. After analysis of these ions, it can be suggested that the second peptide (of lower concentration) has the sequence (FJ)PWJ$-NH_2$. Decisive for this approach is the relatively high intensity of the signal at 414 u (not enlarged in Fig. 9) that can also be explained by the proline effect. That means that the proline occurs in the same position as in the determined sequence JJPFW$-NH_2$. With regard to the sequence of tryptophyllin H 2a (JPWJ), the structural relation allows the assumption that the peptide has the sequence FIPWI rather than IFPWI. Edman degradation provided traces of phenylalanine when analyzing the first (N-terminal) amino acid and, thus, supports this suggestion.

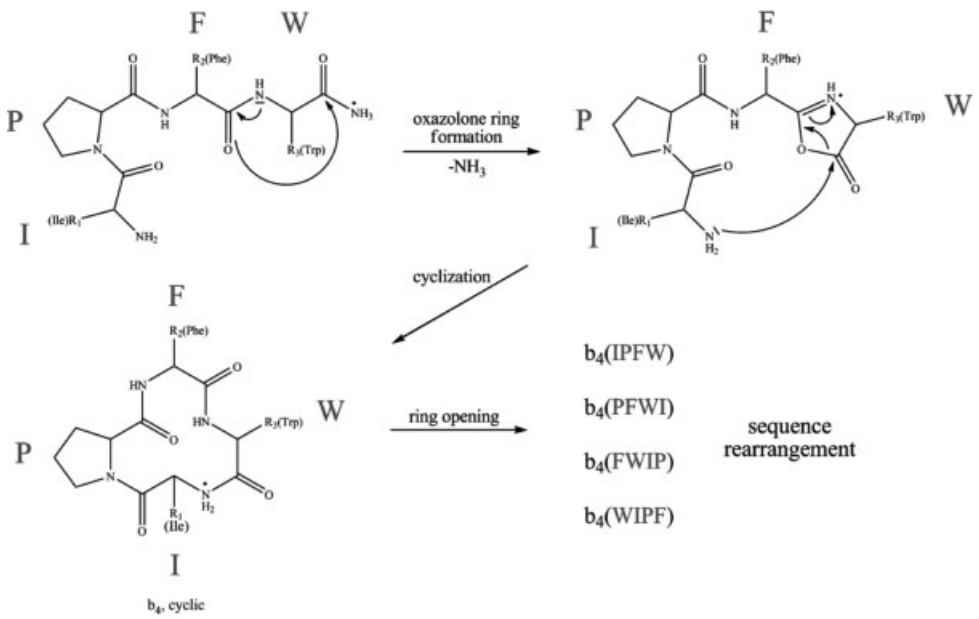
After CID of the corresponding synthetic peptide (Supporting Information), the proposed sequence could be confirmed. A cross-check of an HPLC run of synthetic tryptophyllins H 4a and H 5 (1:1 mixture) gave no separation under the same conditions as chosen for the natural sample.

## Fragmentation features

An overview of the fragmentation pathways of protonated peptides can be found in Paizs and Suhai.[50] Pathways of b-type ion formation, however, are only described for N-terminal species in this review. MS/MS spectra of tryptophyllin peptides show high abundances of internal b-type ions (see Figs. 4–9). Schlosser and Lehmann[47] explained the formation of internal b ions via the accepted oxazolone mechanism,[51] which includes a multistep fragmentation of y or b ions.

In our experiments, we selected y- and b-type product ions as precursors and subjected them to $MS^n$ measurements. Based on these experiments it was found that internal b ions, which retain the C-terminal amino acid residues, can be formed either via y-type ions releasing water or ammonia, or via b-type ions detaching its N-terminal amino acid as neutral ketene, as proposed in Scheme 1. $MS^n$ spectra of y and b ions showed the corresponding internal b ions in high abundance. The $MS^3$ spectrum (Supporting Information) of the $b_4$ ion of tryptophyllin H 1b (see Fig. 4) for example is predominated by the $b_{i(2-4)}$ and $b_{i(3-4)}$ ions (in addition to $a_4$), whereas the $b_3$ ion signal has a rather low intensity.

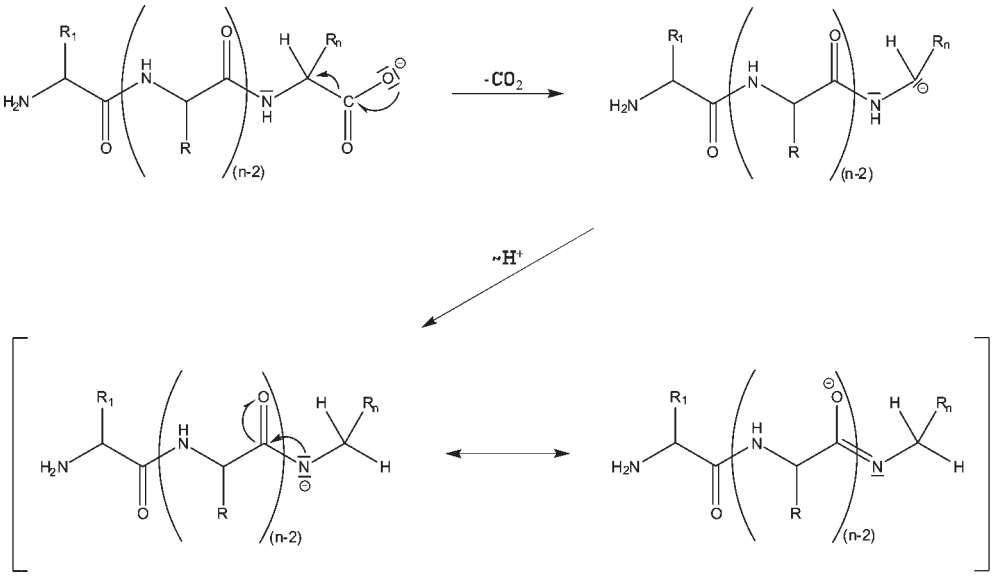
It has been observed that the acid amide peptides tryptophyllin H 2a and 3a yielded lower relative intensities of y-type ions by CID than their free acid analogs (Figs. 6 and 7). The corresponding internal b ion signals ($b_i$) of the amide peptides, at the same time, occurred with higher intensities. Amide peptides thus yielded significantly




**Scheme 2.** Formation of non-direct sequence ions resulting from sequence rearrangements in protonated peptides.[35] Illustration at tryptophyllin H 3a (JPFW-$NH_2$, setting J to I).



DOI: 10.1002/rcm




**Scheme 3.** Proposed mechanism for the $CO_2$ elimination of deprotonated peptides.

higher $b_i/y$ ratios than their acid analogs. Regarding the fragmentation pathway via y-type ions as precursors of internal b ions, this means that acid amide peptides eliminate ammonia more easily than free acid peptides eliminate water. This can be explained by the higher basicity of the acid amide than of the free acid. Considering the proton mobility model,[52] the mobile proton thus is attached to the amide group (releasing ammonia) with higher probability than to the acid group (releasing water). CID spectra of tryptophyllins H 4a and 4b (Fig. 9), which are dominated by the proline effect, still show this phenomenon to a lesser extent.

Another observation is the appearance of non-direct sequence ions as shown in Fig. 4 (tryptophyllin H 1a, $b_{r(1,4)}$) and Fig. 7 (tryptophyllin H 3a, $a_{r(1,4)}$, $b_{r(1,4)}$, $b_{r(1\text{-}2,4)}$), and this was confirmed by CID measurements of synthetic peptides. These ions are believed to be formed via peptide cyclization, followed by ring opening at any peptide bond, causing sequence rearrangements with respect to the original peptide sequence. These peptide fragmentation pathways have been investigated by experimental and modeling techniques.[35,36] The formation of sequence rearrangements is elucidated in Scheme 2 for the sequence of tryptophyllin H 3a. After the oxazolone ring formation, the peptide cyclization reaction is initiated by nucleophilic attack of the N-terminal amino group on the oxazolone carbonyl group. This leads to a cyclic isomer of the $b_4$ ion that can open at various peptide bonds. C-terminal or N-terminal fragmentations of these rearranged peptides produce ions that do not directly correspond to the original sequence. The free acid analog showed lower abundance of non-direct sequence ions after CID. Considering the oxazolone ring formation in Scheme 2, where $NH_3$ is split off from the amide peptide to induce cyclization, this affirms that the elimination of water in free acid peptides is more impeded. With regard to the pathway of Scheme 2, it is obvious that internal b ions can also be formed through cyclization.

In our studies, rearrangement products were only observed as low-abundance species compared with internal or regular b ions. The scheme of possible isomers is important to decide whether observed ions in MS/MS spectra derive from the determined peptide sequence or if they are rearrangement products formed after CID. A new application-related aspect of this scheme is that observed product ions, which do not fit the scheme, can be identified as belonging to a different (coeluted, isobaric) peptide. The latter case was found in the CID spectra of tryptophyllin H 4a and 4b (Fig. 9). Product ion signals labeled with their composition only are those that cannot be deduced from the determined peptide sequence. Thus, we can show that the sample contained another peptide (tryptophyllin H 5) having the same mass and same composition, but a different sequence, as mentioned earlier.

Deprotonated peptides that are not amidated at the C-terminus showed loss of $CO_2$ in product ion spectra (Figs. 8 and 10). $CO_2$ elimination can be depicted by the proposed mechanism in Scheme 3. After decarboxylation, the carbanion is transformed into an amide anion by a proton shift. $CO_2$ elimination has been described earlier for deprotonated peptides containing disulfide links.[53]

## Proline effect

MS/MS spectra of proline-containing peptides with at least two amino acids N-terminal to proline are dominated by the y-type ion with proline at the N-terminal end. This was observed both in the positive and in the negative ion mode. (Figs. 4 and 5 (tryptophyllin H 1b) and Figs. 9 and 10 (tryptophyllin H 4a and 4b)). This phenomenon is called the 'proline effect'.[41,54] MS/MS spectra of peptides with only one amino acid N-terminal to proline do not show this effect (Fig. 4 (H 1a), Fig. 6 (H 2a and 2b) and Fig. 7 (H 3a and 3b)).

Due to the proline effect, most product ion signals of tryptophyllin H 4a had to be enlarged up to 50-fold to make them visible (Fig. 9). The fragmentation ability in the negative ion mode (Fig. 10) was marginal and the MS/MS spectra of both H 4a and H 4b were unusable for assisting peptide characterization.





DOI: 10.1002/rcm

The mechanism of the proline effect is not yet understood. Schwartz *et al.*[49] discussed a high proton affinity of proline due to its greater basicity than that of other amino acids. They assumed that this would lead to a preferred attachment of the proton at the N-terminal proline site leading to corresponding y ions. Vaisar *et al.*[54] found that an increased proton affinity of proline is not the only reason for this effect. Previous studies[37] showed that the charge-carrying proton in the positive ion mode plays an important role and that sodiated peptides do not exhibit the proline effect. Grewal *et al.* [40] explored the fragmentation of proline-containing tripeptides by density functional theory (DFT) calculations. In addition to the higher basicity of the prolyl nitrogen, they took steric reasons into account. Since we also observe the proline effect in the negative ion mode, the argument based on high proton affinity would indicate that there is also proton mobility within a deprotonated peptide. It can be assumed that proton mobility also occurs in deprotonated peptides causing an associated mobility of the resulting negative charge. Molecular modeling of the reaction coordinates of deprotonated peptides supports this assumption.[48] Harrison *et al.* [39] explored the fragmentation behavior of deprotonated peptides containing proline at different positions. In accordance with their results, we observed the formation of y-type ions with favored cleavage at the C-terminal side of proline, if there is only one amino acid N-terminal to proline. This is shown in the MS/MS spectra of the deprotonated acid amides H 1a and 3a (Figs. 5 and 8), where the $y_2$ ion occurs in high abundance. The CID spectrum of the free acid analogue H 3b, however, shows the $b_3$ ion as the most intense peak.

## CONCLUSIONS

Combining CBS and manual *de novo* peptide sequence analysis allowed us to characterize amphibian peptides efficiently. Edman sequencing instead did not provide satisfying results for most HPLC fractions, so that I/L was set to 'J' in general. Sequence analysis of peptides in this study showed that CBS worked successfully and reliably using CID spectra of high complexity. Sequences were confirmed by comparing MS/MS spectra of natural and synthesized peptides. In our studies, negative ion mass spectrometry was employed as an assisting, additional sequencing method, providing information that is complementary to positive ion MS/MS data. Negative ion MS/MS can be particularly useful if complex fragmentation patterns are obtained in the positive ion mode. Investigation of fragmentation mechanisms helped to unequivocally identify unknown peptides by composition-based sequencing.

The peptides characterized in this study show sequence similarities to tryptophyllin-like peptides from the skin secretion of the Australian red tree frog *Litoria rubella* and of the Australian buzzing tree frog *Litoria electrica*, called tryptophyllins L.[26,27] The tryptophyllins L were reported to have no significant antimicrobial activity. Bioactivity testing of peptides from *Hyla savignyi* against *Escherichia coli* (Gram negative) and *Micrococcus luteus* (Gram positive) gave results in accordance with their findings. The natural function of tryptophyllin-like peptides is thus still unknown.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article.


## Acknowledgements

We would like to thank Faraham Ahmadzadeh and Hossein Hashempour from Shahid Beheshti University, Tehran, for collecting the skin secretions. We also like to thank Hans-Günter Welker and Dr Sigrid Schmitt (Biochemistry Institute, Medical Faculty of Justus Liebig University, Giessen) for Edman degradation. 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


1. Toledo RC, Jared C. *Comp. Biochem. Physiol.* 1995; **111A**: 1.
2. Erspamer V, Erspamer GF, Mazzanti G, Endean R. *Comp. Biochem. Physiol.* 1984; **77C**: 99.
3. Erspamer V, Erspamer GF, Cei JM. *Comp. Biochem. Physiol.* 1986; **85C**: 125.
4. Habermehl G. *Naturwissenschaften* 1969; **56**: 615.
5. Roseghini M, Erspamer V, Endean R. *Comp. Biochem. Physiol.* 1976; **54C**: 31.
6. Gomes A, Giri B, Saha A, Mishra R, Dasgupta SC, Debnath A, Gomes A. *Indian J. Exp. Biol.* 2007; **45**: 579.
7. Li J, Xu X, Xu C, Zhou W, Zhang K, Yu H, Zhang Y, Zheng Y, Rees HH, Lai R, Yang D, Wu J. *Mol. Cell. Proteomics* 2007; **6**: 882.
8. Nascimento ACC, Fontes W, Sebben A, Castro MS. *Prot. Pep. Lett.* 2003; **10**: 227.
9. Simmaco M, Mignogna G, Barra D. *Biopolymers* 1998; **47**: 435.
10. VanCompernolle 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. *J. Virol.* 2005; **79**: 11598.
11. Chinchar VG, Bryan L, Silphadaung U, Noga E, Wade D, Rollins-Smith L. *Virology* 2004; **323**: 268.
12. Lu CX, Nan KJ, Lei Y. *Anti-cancer drugs* 2008; **19**: 931.
13. Zairi A, Serres C, Tangy F, Jouannet P, Hani K. *Bioorg. Med. Chem.* 2008; **16**: 266.
14. Shai Y. *Biochim. Biophys. Acta* 1999; **1462**: 55.
15. Matsuzaki K. *Biochim. Biophys. Acta* 1999; **1462**: 1.
16. Brogden KA. *Nat. Rev. Microbiol.* 2005; **3**: 238.
17. Matyus E, Kandt C, Tieleman DP. *Curr. Med. Chem.* 2007; **14**: 2789.
18. Koczulla AR, Bals R. *Drugs* 2003; **63**: 389.
19. Gordon YJ, Romanowski EG. *Curr. Eye Res.* 2005; **30**: 505.
20. Pukala TL, Bowie JH, Maselli VM, Musgrave IF, Tyler MJ. *Nat. Prod. Rep.* 2006; **23**: 368.
21. Apponyi MA, Pukala TL, Brinkworth CS, Maselli VM, Bowie JH, Tyler MJ, Booker GW, Wallace JC, Carver JA, Separovic F, Doyle J, Llewellyn LE. *Peptides* 2004; **25**: 1035.
22. Wang Z, Wang G. *Nucleic Acids Res.* 2004; **32**: D590.
23. Wang G, Li X, Wang Z. *Nucleic Acids Res.* 2009; **37**: D933.
24. Renda T, D'Este L, Buffa R, Usellini L, Capella C, Vaccaro R, Melchiorri P. *Peptides* 1985; **6**: 197.
25. Chen T, Orr DF, O'Rourke M, McLynn C, Bjourson AJ, McClean S, Hirst D, Rao P, Shaw C. *Regulatory Peptides* 2004; **117**: 25.
26. Steinborner ST, Wabnitz PA, Waugh RJ, Bowie JH, Gao CW, Tyler MJ, Wallace JC. *Aust. J. Chem.* 1996; **49**: 955.
27. Wabnitz PA, Bowie JH, Wallace JC, Tyler MJ. *Aust. J. Chem.* 1999; **52**: 639.
28. Jackway RJ, Maselli VM, Musgrave IF, Maclean MJ, Tyler MJ, Bowie JH. *Rapid Commun. Mass Spectrom.* 2009; **23**: 1189.
29. Chen W, Lee PJ, Shion H, Ellor N, Gebler JC. *Anal. Chem.* 2007; **79**: 1583.






DOI: 10.1002/rcm

30. Racine AH, Payne AH, Remes PM, Glish GL. *Anal. Chem.* 2006; **78**: 4609.
31. Beardsley RL, Sharon LA, Reilly JP. *Anal. Chem.* 2005; **77**: 6300.
32. Papayannopoulos IA. *Mass Spectrom. Rev.* 1995; **14**: 49.
33. Thieu VA, Kirsch D, Flad T, Mueller C, Spengler B. *Angew. Chem. Int. Ed.* 2006; **45**: 3317.
34. Spengler B, Luetzenkirchen F, Metzger S, Chaurand P, Kaufmann R, Jeffery W, Bartlet-Jones M, Pappin DJC. *Int. J. Mass Spectrom. Ion Proc.* 1997; **169/170**: 127.
35. Harrison AG, Young AB, Bleiholder C, Suhai S, Paizs B. *J. Am. Chem. Soc.* 2006; **128**: 10364.
36. Jia C, Qi W, He Z. *J. Am. Soc. Mass Spectrom.* 2007; **18**: 663.
37. Unnithan AG, Myer MJ, Veale CJ, Danell AS. *J. Am. Soc. Mass Spectrom.* 2007; **18**: 2198.
38. Smith LL, Herrmann KA, Wysocki VH. *J. Am. Soc. Mass Spectrom.* 2006; **17**: 20.
39. Harrison AG, Young AB. *J. Mass Spectrom.* 2005; **40**: 1173.
40. Grewal RN, El Aribi H, Harrison AG, Siu KWM, Hopkinson AC. *J. Phys. Chem. B* 2004; **108**: 4899.
41. Kaufmann R, Chaurand P, Kirsch D, Spengler B. *Rapid Commun. Mass Spectrom.* 1996; **10**: 1199.
42. Spengler B. *J. Am. Soc. Mass. Spectrom.* 2004; **15**: 703.
43. Roepstorff P, Fohlmann J. *Biomed. Mass Spectrom.* 1984; **11**: 601.
44. Biemann K. *Methods Enzymol.* 1990; **193**: 886.
45. Savitski MM, Nielsen ML, Zubarev RA. *Anal. Chem.* 2007; **79**: 2296.
46. Williams JP, Creese AJ, Roper DR, Green BN, Cooper HJ. *J. Am. Soc. Mass. Spectrom.* 2009; **20**: 1707.
47. Schlosser A, Lehmann WD. *Proteomics* 2002; **2**: 524.
48. Bowie JH, Brinkworth CS, Dua S. *Mass Spectrom. Rev.* 2002; **21**: 87.
49. Schwartz BL, Bursey MM. *Biol. Mass Spectrom.* 1992; **21**: 92.
50. Paizs B, Suhai S. *Mass Spectrom. Rev.* 2005; **24**: 508.
51. Yalcin T, Khouw C, Csizmadia IG, Peterson MR, Harrison AG. *J. Am. Soc. Mass Spectrom.* 1995; **6**: 1165.
52. Dongré AR, Jones JL, Somogyi Á, Wysocki VH. *J. Am. Chem. Soc.* 1996; **118**: 8365.
53. Bilusich D, Maselli VM, Brinkworth CS, Samguina T, Lebedev AT, Bowie JH. *Rapid Commun. Mass Spectrom.* 2005; **19**: 3063.
54. Vaisar T, Urban J. *J. Mass Spectrom.* 1996; **31**: 1185.





DOI: 10.1002/rcm