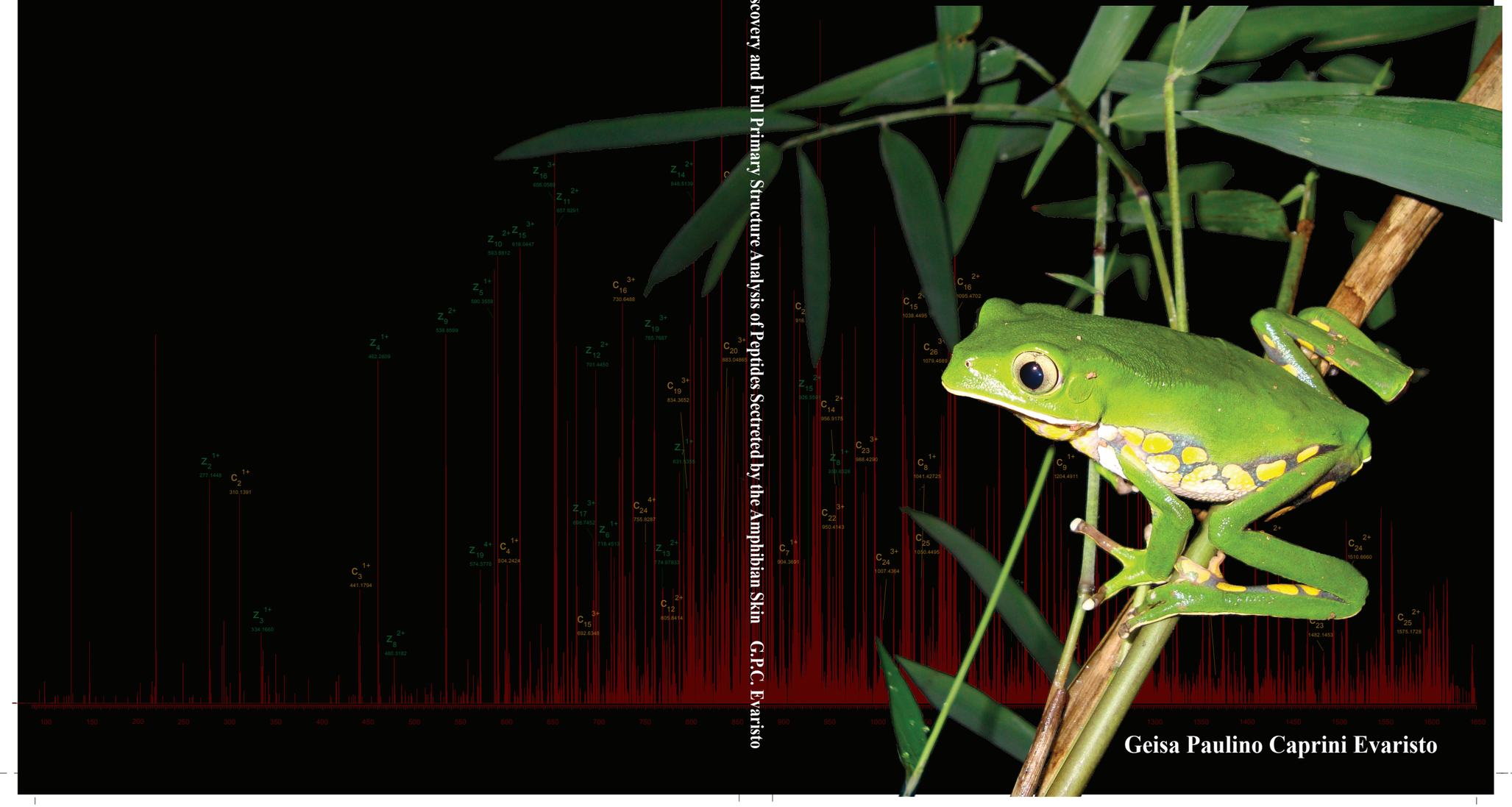
MS Approaches for the Discovery and Full Primary Structure Analysis of Peptides Secreted by the Amphibian Skin G.P.C. Evaristo

Mass Spectrometry Approaches for the Discovery and Full Primary Structure Analysis of Peptides Secreted by the Amphibian Skin



Geisa Paulino Caprini Evaristo

Invitation

To the public defence of my doctoral thesis entitled

Mass Spectrometry Approaches for the Discovery and Full Primary Structure Analysis of Peptides Secreted by the Amphibian Skin

Monday October 29, 2012 at 12:30h
in the Frans van Hasseltzaal of the Aula
of the Delft University of Technology,
Mekelweg 5, Delft.

A short presentation of this work
will precede the defence at 12:00h

Paranymphs:

Michaela Greiler
michaela.greiler@gmail.com

Eda Bener
edabeneraksam@gmail.com

Geisa Evaristo
geisapc@gmail.com

Mass Spectrometry Approaches for the Discovery and Full Primary Structure Analysis of Peptides Secreted by the Amphibian Skin

Geisa Paulino Caprini Evaristo

Cover photo: Joseph Albert Medeiros Evaristo

Cover layout: Martijn W. H. Pinkse and Geisa P. C. Evaristo

Printed by: Uitgeverij BOXPress, 's-Hertogenbosch

Mass Spectrometry Approaches for the Discovery and Full Primary Structure Analysis of Peptides Secreted by the Amphibian Skin

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan te Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.C.A.M.Luyben,
voorzitter van het College voor Promoties
in het openbaar te verdedigen op maandag 29 oktober 2012 om 12:00 uur
door

Geisa Paulino Caprini EVARISTO

Master of Science in Biosciences and Biotechnology, State University of Norte Fluminense "Darcy Ribeiro", Brazil
geboren te Cachoeiro de Itapemirim, Espírito Santo, Brazil

Dit proefschrift is goedgekeurd door de promotor:
Prof. Dr. P. D. E. M. Verhaert

Copromotor:
Dr. M. W. H. Pinkse

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. Dr. P. D. E. M. Verhaert,	Technische Universiteit Delft, promotor
Dr. M. W. H. Pinkse,	Technische Universiteit Delft, copromotor
Prof. Dr. J. H. de Winde,	Technische Universiteit Delft
Prof. Dr. W. R. Hagen,	Technische Universiteit Delft
Prof. Dr. A. J. R. Heck,	Universiteit Utrecht
Prof. Dr. C. Shaw,	Queen's University of Belfast, North Ireland
Prof. Dr. J. Schripsema,	State University of Norte Fluminense, Brazil
Prof. Dr. Ir. J. J. Heijnen,	Technische Universiteit Delft, reservelid

This study was financially supported by National Council for Scientific and Technological Development (CNPq-Brazil), grant number 200847/2007-4, and the Netherlands Genomics Initiative.

ISBN: 9789088915093

To my husband, Joseph

Table of Contents

Chapter 1

General introduction

1.1 The amphibian skin and its venom

Frog and toads are particularly vulnerable to many of their predators; they have a soft skin, lack claws to defend themselves and often they are brightly colored. It is the skin of amphibians which plays a pivotal role in their survival. It does not only helps in camouflage, it also regulates several physiological processes, including respiration, temperature control and defense. Most of these functions are supported by mucous and granular (or serous) glands distributed in the dermis. The mucous glands serve to keep the skin surface moist and slippery, protecting it during prolonged contact with water, and retarding evaporative water loss when on land. The granular glands of anurans are located mainly in the dorsal region. In response to stress they secrete compounds, which give the frog a bad taste and/or are poisonous or even lethal to their predators. This secretion is a complex cocktail of bioactive molecules, of which an important portion is actually encoded in the amphibian genome as peptides and small proteins. As frogs do not have stings, fangs or any other mechanically invasive tool, an elaborate chemistry based defensive strategy against big predators/carnivores has developed through evolution. Amphibian skin secretions contain large arsenals of cytolytic peptides that constitute the first frontline barrier against infections by microorganism[1, 2]. A few classes of opportunistic bacteria (gram-negative) constitute the natural flora of frogs and apparently the constant production of mucus and cytolytic peptides helps to 'wash off' growing microbial colonies from the amphibian skin[3]. They also serve to penetrate (in)vertebrate predator skin/epithelia to provide access into the circulatory systems. The co-secretion of myoinhibitory peptides with vasodilatory (smooth muscle dilation) effects[4-6] further facilitate/stimulate the faster uptake of the venom through endothelia (blood vessel walls). Other classes of peptides have specific activities directed to deter predators which are both vertebrate (fish, other amphibians, reptiles, birds, mammals) or invertebrate (such as carnivorous arthropods (insects, spiders)). Once they are in the blood circulatory system they are transported throughout the whole body, targeting for example proteins/receptors, leading to the deregulating of important physiological processes of the predator. During millions of years, each frog species has evolved its own specialized set of chemical warfare agents against their local environmental challengers and/or predators.

The complex frog skin secretory cocktails have been the subject of intense research for many years, both by academic as well as pharmaceutical groups. Constituents of amphibian 'venoms' have been recognized to possess potential applicability as new drugs by

biopharmaceutical research, yielding several lead compounds. As mentioned, the majority of peptides found in amphibian secretions have cytolytic activity. Due to their amphipatic characteristic they interact with phospholipid bilayers leading to the formation of pores in cell membranes, consequently lysing the cells[7-14]. Several of these cytolytic peptides are polycationic, conferring affinity with the more acidic bacterial cell-membranes, rather than the more neutral mammalian cell membrane. For this feature they are typically described as "amphibian antimicrobial peptides"[15-17]. During the past decade, the emergence of multiple-drug resistant strains of many pathogenic microorganisms has stimulated the search for new classes of antimicrobial compounds with potential clinical applications[18]. From studying skin secretions of various frog species, several families of cytolytic frog defense peptides have been characterized. Members of these families tend to act upon many different types of membranes and therefore they are considered to be an interesting lead for the development of antimicrobials[19-22].

Since the class of cytolytic peptides forms the main constituent of the frog secretion, they are easily picked up by (antimicrobial) bioassays, which is reflected by the numerous publications on this subject found in the literature. However, as mentioned previously, a wide variety of non-cytolytic peptides is also present in frog skin secretions. Their abundance is considerably lower than that of the cytolytic ones. The most interesting characteristic of these amphibian skin peptides is that they often are analogous to mammalian regulatory peptides. For example, caeruleins show sequence and activity similarity with cholecystokinin (CCK-8) and gastrin from mammals[23-26], and peptide tyrosine-tyrosine (PYY), a peptide hormone originally identified in pigs, is also present in Phyllomedusinae and Ranidae skins[27]. This prompted Vittorio Erspamer, one of the pioneers in frog defensive peptide research, to formulate his theory that every peptide present in the frog skin has its equivalent in the mammalian brain and/or gut[23, 28]. In special cases, concentrations of frog skin peptides are sometimes more than 3 orders of magnitude higher than in mammalian tissues[23]. Therefore, the study of this class of peptides may provide, next to useful information on their pharmacology and biochemistry in the frogs themselves[19, 29], also a basis for a new research in higher vertebrates, including man[19]. From this point of view, the frog skin is a remarkable 'treasure chest' containing highly interesting peptides and it is therefore often valued as a rich source of potential drug leads. Particularly this feature has increased the awareness that amphibians deserve a much broader interest than they enjoy today. It has also triggered growing emphasis on the conservation of specific and unique amphibian species and their habitats, which are currently rapidly decreasing[30].

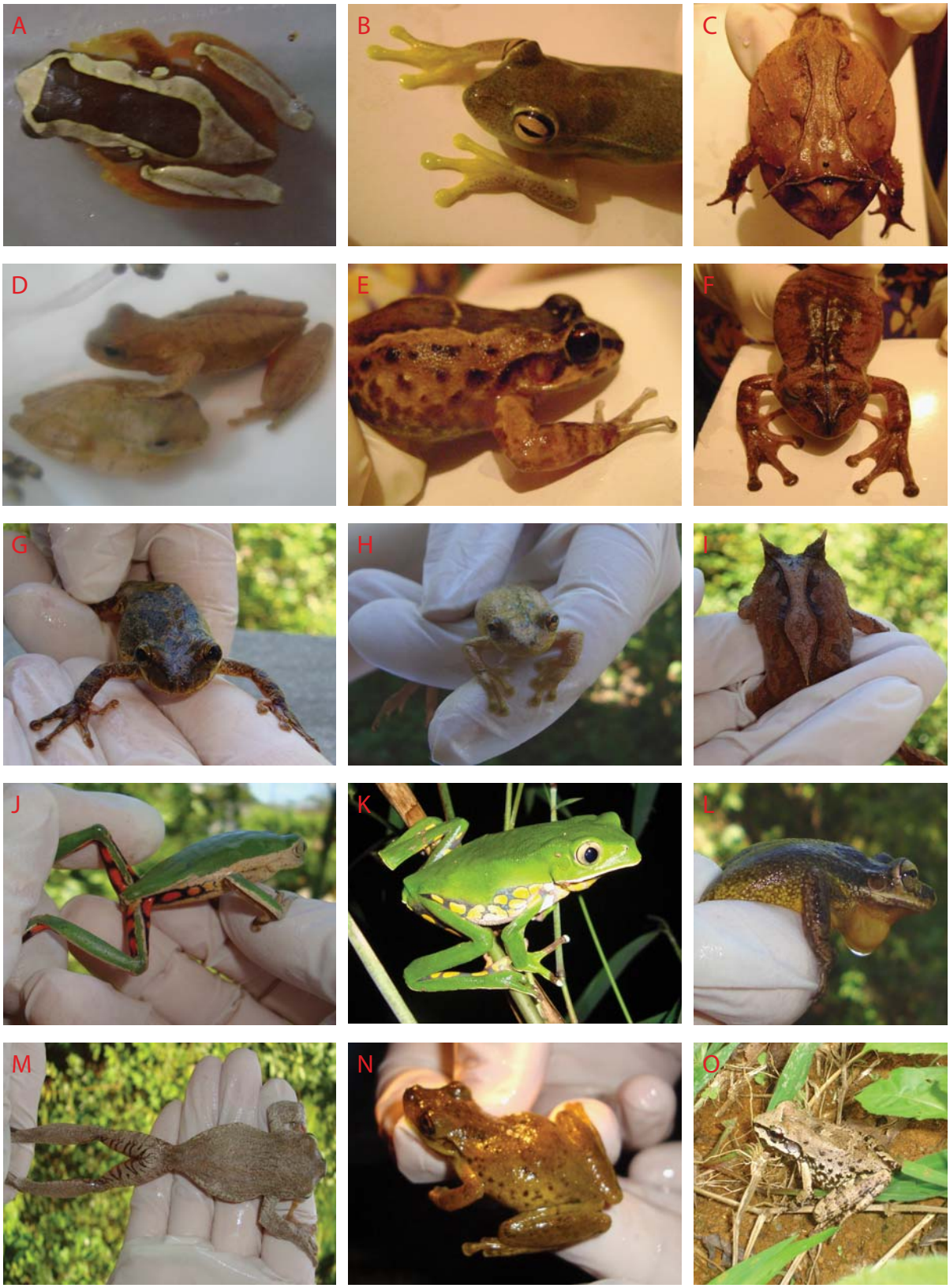


Figure 1: Brazilian frogs found during several frog hunting expeditions in Brazil. (A) Dendrosophus elegans, *(B)* Hypsiboas albomarginatus, *(C)* Proceratophrys sp., *(D)* Dendrosophus minutus, *(E)* Scinax havi, *(F)* Thoropa miliaris, *(G)* Hypsiboas faber, *(H)* Scinax sp., *(I)* Proceratophrys boiei, *(J)* Phyllomedusa rohdei, *(K)* Phyllomedusa burmeisteri *and (L-O) four unidentified species*

Amphibian biodiversity is the greatest in South America, and the most toxic and interesting frog secretions are found in these areas. The start for this PhD was to study several frog skin secretions of Brazilian frogs (Figure 1). This country has so far the largest known diversity in frog species (913 species, http://www.sbherpetologia.org.br). Urbanization and agricultural activities, together with the accompanying loss and isolation of natural habitats threatens the survival of many frog species. It is clear that with every species which gets extinct a unique library of bioactive compounds vanishes from the planet with no possibility whatsoever to ever recuperate the biologically interesting information contained in it. Mankind should not let this happen and, in this context we see this PhD thesis as a small contribution towards maintaining the rich biodiversity of frogs and their biotopes.

1.2 Biosynthesis and processing

Granular gland amphibian skin peptides are synthesized as pre-propeptides, typically consisting of three parts. The open reading frame (ORF) encodes a hydrophobic putative "signal peptide" of around 22 residues, followed by a 16-25 residues "acid spacer", terminating in a conserved dibasic processing signal (Lys-Arg), followed by the actual bioactive peptide. Pre-propeptides are processed by first the removal of the signal peptide to the proform, which is stored in the large granules of the glands. The bioactive peptide gets finally released by C-terminal cleavage after the dibasic site (see examples in Fig. 2).

	Putative signal // acid spacer // bioactive peptide
Pre-propeptide	MAFVKKSLLLVLFLGLVSFSIC // EEEKRETEEDENEDEIEEESEEKKR // ENREVPAGFTALIKTLRKCKII
Propeptide	EEEKRETEEDENEDEIEEESEEKKR // ENREVPAGFTALIKTLRKCKII
Mature peptide	ENREVPAGFTALIKTLRKCKII

Figure 2: Translated sequence from ORF encoding a frog skin peptide (distinctin cloned from Phyllomedusa burmeisteri, *see Chapter 7), showing pre-proforms and mature (bioactive) peptide.*

Many pre-proregions of amphibian skin peptide precursors share sequence similarities, suggesting a common ancestral gene origin. During evolution, amphibian skin peptides have evolved independently in each species, as a result of continuous interaction with the environment and its respective selection-pressure[31, 32]. Consequently, the diversity of frog skin peptides, both intra and inter-species, is extensive, with a plethora of species specific peptides differing in primary structure, size, charge and hydrophobicity. In addition, many peptides are post-translationally modified, which often enhances their bioactivity.

1.3 Post-translational modifications

Post-translational modifications (PTMs) are the result of chemical reactions in a protein after the ribosomal translation process, with functional groups added (and later also cleaved off) to and from specific residues. There are a few hundred types of PTMs and new ones are still discovered[33]. PTMs change the amino acid primary structure, thereby also altering the higher order peptide structure and, consequently, its function. The role of each PTM is attributed to its type. In peptides secreted by the amphibian skin, 6 major types of PTMs have been described in the literature:

1- C-terminal amidation ($–NH_2$): Exchange of the hydroxyl group of the last residue to amine. The amide is donated by an obligatory carboxyterminal Gly residue encoded in the pre-propeptide. This is one of the most common PTM found in bioactive peptides (of amphibian and other origin) and results in a mass decrease of 0.98 Da.

2- Disulfide bridges (S-S): Oxidation of the sulfhydryl groups of two Cys residues, yielding a cystine. This PTM implies the loss of 2 hydrogens, yielding a 2.02 Da decrease from the calculated mass of the peptide. Disulfide bonds stabilize the secondary structure of a peptide and, therefore, are often responsible for an altered bioactivity.

3- Hydroxyproline (Hyp): Addition of hydroxyl group to a Pro residue, yielding a 15.99 Da mass increase.

4- Sulfation of tyrosine (SO_3): Addition of a sulfate group to Tyr, with a consequent 79.96 Da mass increase.

5- Pyroglutamic acid (pGln): the free amino function of an aminoterminal Gln cyclizes, with the loss of an NH_3 group (17.03 Da mass decrease).

6- D-amino acid residues: Enzymatic (isomerase) conversion of the natural L-isomer into its D-variant. As D- and L-amino acids have the same molecular mass, this PTM is difficult to spot directly by mass spectrometry.

1.4 Overview of amphibian peptides

In view of the species under investigation in this PhD work, we here present a brief overview of the peptide families reported in Raninae and Phyllomedusinae frogs, which are classified on the basis of general sequence features and specific PTMs.

- Rana box peptides: Ranid antimicrobial peptides containing a conserved C-terminal disulfide bridged heptapeptide, the so-called "*Rana* box" motif. These include gaegurins (24–37 residues), brevinins-1 (17–24 residues) and -2 (30–34 residues), ranalexins (20 residues), ranatuerins-1 (25 residues) and -2 (33 residues), esculentins-1 (46 residues) and -2 (37 residues), palustrins (31 residues), japonicins-1 (14 residues) and -2 (21 residues), nigrocins-2 (21 residues) and rugosins (33–37 residues)[31, 32, 34, 35].
- Ranid antimicrobial peptides without *Rana* box motif: temporins (10–14 residues) [32]. They are slightly cationic and C-terminally amidated[36].
- Antimicrobial phyllomedusinid peptides: dermaseptins and -related peptides (24-34 residues), phylloxins (19 residues), dermatoxins (32 residues), phylloseptins (19-21 residues), plasticinins (23 residues) and skin polypeptides YY (36 residues) [32, 37-40]. Several of these peptides are C-terminally amidated.
- Myotropic phyllomedusinid peptides (with effects on smooth muscle contractions) as listed in the NCBI protein sequence database: phyllokinins (8-11 residues, some with sulfoTyr and/or hydroxyPro), bradykinins (8-14 residues), tryptophyllins (5-13 residues), litorins (9 residues), phyllolitorins (9 residues), phyllomedusins (10 residues), phyllocaeruleins (9 residues), sauvagines (40-42 residues) and sauvatides (14 residues) [38, 41]. The N-terminal Gln of bradykinin-potentiating peptides is often post transitionally modified to pyroGlu (pyroglutamic acid or pyrrolidone carboxylic acid).
- Phyllomedusinid peptides with (opioid like) central nervous system effects: deltorphins (7-17 residues) and dermorphins (5-13 residues)[42-45]. Besides sharing structural similarities with other opioids, both have a D-amino acid as second residue from the aminoterminus.
- Phyllomedusinid peptides with yet unkown activity: hyposins (11-15 residues)[46].
- *Phyllomedusa* polypeptides with protease inhibition activity: kazal peptides (6 kDa mass ranges). They have 3 intramolecular disulfide bridges[47].

1.5 Peptide primary structure analysis

In view of the analysis of frog skin peptides, the most common ways to elucidate the primary structure of a peptide (or protein) is to make use of [1] Edman degradation, [2] molecular cloning or [3] tandem mass spectrometry (MS/MS) *de novo* sequencing. All 3 methods can be used independently or in combination and they typically yield complementary data. Advantages and disadvantages/limitations of each approach are summarized in Table 1.

Table 1: Pros and cons of available techniques to sequence peptides from frog skin secretion.

	Pros	Cons
Edman degradation	• Accurate • Cheap	• low sensitivity • Purified peptide required (>95%) • Max. 30 residues per peptide • No PTM recognition • Relatively slow
Molecular cloning	• Accurate • Full precursor	• Intact mRNAs required • Expensive extraction kit • Analysis takes weeks • No PTM recognition
Mass spectrometry	• Fast spectra acquisition • PTM identification • No purification required • High sensitivity	• Expensive instruments • Ile/Leu cannot be distinguished • Discrimination of Lys Gln requires high resolution

Edman degradation is an automated technique where the amino-terminal residue is derivatized and cleaved from the peptide without disrupting peptide bonds between the other amino acid residues. It is a well-established technique (dating from 1967[48]) and still considered a standard for sequencing peptides. It is an accurate way to elucidate straightforward peptide sequences discriminating between the isobaric amino acids Leu and Ile. Disadvantages of Edman sequencing are poor sensitivity, requirement for highly purified sample (typically about 1 pmole of >95% pure peptide/protein is required), the inability to identify PTMs, relatively low analysis speed, and the restriction that only peptides with a free amino-terminus can be successfully analyzed. Peptides with 'blocked' N-termini, e.g. pyroglutamic acid cannot be directly sequenced by Edman degradation.

Molecular cloning is currently the preferred technique to obtain the primary structure of frog skin peptides. Pioneering work by Chen, Shaw and coworkers has shown that skin secretions still contain the mRNAs encoding for the amphibian defense peptides[49]. Frogs

secrete peptides through a holocrine process which involves rupturing of the mature granular cells in the center of the gland, which releases all of their cytosolic components in the secretion, including intact polyadenylated mRNAs[49]. The intact mRNA molecules present in the venom can be isolated by virtue of their polyA tail, using oligo dT magnetic beads. Subsequently, a cDNA library is made by converting all mRNAs into complementary DNA (cDNA) by reverse transcriptase. From this library, cDNAs encoding for specific peptides sequences can be multiplied by polymerase chain reaction (PCR) amplification using either a highly specific primer (when a partial sequence of the peptide of interest is known), or a degenerated primer, designed from a conserved sequence, such as from a signal peptide. The amplified DNA is inserted into a vector, which is introduced in a bacterial system. Colonies containing the recombinant cDNA sequences are selected for DNA purification, PCR and finally DNA sequencing. The major advantage of molecular cloning is that it does not only identify (predict) the active peptide sequence, also the sequence of signal peptide and acid spacer (precursor peptide) are elucidated. Also, Leu and Ile isobaric residues are easily differentiated because of their different codons. Although a particular amino acid can be specified by more than one codon, each specific codon represents a single amino acid. Therefore, the cDNA sequencing provides sufficient data to identify a particular primary peptide structure. One of the disadvantages of this molecular cloning approach is that in order to be successful the venom must contain intact polyadenylated mRNAs of all peptides and, unfortunately not all peptide encoding mRNAs are equally abundant. PTMs present in peptides, except for C-terminal amidation (see above), cannot be readily predicted from the cDNA sequencing results. Additionally, molecular cloning is often considered quite costly (expensive RNA extraction kits) and time consuming.

In both Edman degradation and the molecular cloning approach, extensive peptide purification and or specific primer designs are needed to target the sequence of only one certain peptide in the skin secretion, where the majority of the other peptides remain unnoticed. From this point of view mass spectrometry has a huge advantage, since it can make an inventory in terms of size, complexity and abundances of the peptides present in the skin secretion. Furthermore, with tandem mass spectrometry it is possible to obtain in a short period of time fragmentation spectra of many peptides, from which partial or full primary amino acid structure could be obtained. Due to the sensitivity of modern mass spectrometers, only a few milligrams of crude secretion are already enough to do multiple analysis. This issue is fundamental in this type of research as sample material is limited. The amount of frog secretion obtained during hunting expeditions largely depends on the number of frogs

captured, and the amount of material collected per specimen also varies for different frog species. Frog secretions can be analyzed in more detail using (nano)liquid chromatography in combination with mass spectrometry, either online via ESI or off-line via MALDI. In addition an extra LC separation can be included for highly complex mixtures (so-called multidimensional LC separation). Another unique characteristic of MS-based peptide analysis is the capability to specifically detect and identify PTMs on peptides. In Edman degradation the residues containing the PTM are typically not determined, or sometimes can lead to an erroneous amino acid call. From molecular cloning results, some PTMs, like amidation or intramolecular disulfide bonds, can be predicted on basis of conserved motifs and or similarity with other know modified peptides. However, these predictions always need confirmation by complementary analyses, and for this MS is highly suitable. Compared to the other techniques, the analysis time in MS is relatively fast. In general, a few hours are enough to achieve sample preparation and analysis. Acquisition of (tandem) MS spectra is very fast nowadays, which allows the acquisition of hundreds of spectra per LC-MS(/MS) run. Consequently, the amount of spectral data produced is enormous. Bioinformatic tools can help identifying peptides from tandem MS spectra, provided that the peptide's sequence has been deposited in a database, like NCBI or UniProt. Spectra of unknown/novel peptides should be "*de novo* sequenced", which is often done manually. In many cases this yields only partial sequence information, but this could be still be valuable information for homology/similarity searches with other peptides present in sequence databases. Major drawback of mass spectrometry is the high instrumentation cost, especially for high sensitivity and high resolution apparatus, such as those equipped with orbitrap analyzers. Specific limitations of MS with respect to peptide sequencing are [1] the difficulty of MS instruments with limited resolving power to differentiate amino acid residues with the same nominal mass, e.g. Gln (128.0586) and Lys (128.0950) residues; and [2] the impossibility to distinguish isobaric amino acid residues such as Leu and Ile in conventional instruments (although ECD fragmentation spectra containing the side chain loss information of these amino acids[50] can be used to discriminate between both).

Many different instrument setup and chemical tools are available for more detailed structural characterization. Due to this wide flexibility MS-based methods are growing in popularity and are more and more used to identify and *de novo* sequence frog skin peptides. However, not all MS/MS spectra contain sufficient or unambiguous information to complete the full amino acid sequence. The absence of genomic information forms another drawback when elucidating frog peptide sequences by this approach. But then again, as mentioned

above, for frog skin peptides genomic/transcriptomic information can in some cases be obtained by molecular cloning. From this point of view, MS/MS based sequencing could provide additional sequence information for the design of novel primers. In this thesis different (tandem) MS techniques are explored and employed for *de novo* sequencing of peptides secreted by the skin granular glands of amphibians.

1.6 Selection of amphibian species studied

As soon as we started our project we initiated the procedure to get authorization by IBAMA (Brazilian Institute of Environment and Natural Resources, license number 010453/2010-5) to study selected Brazilian frog species. While waiting for the permit to come through, we trained ourselves on the analysis of frog peptides within the framework of an ongoing collaboration between our Delft Laboratory and the Laboratory of Natural Drug Discovery (Prof. C. Shaw) of the School of Pharmacy of Queen's University of Belfast (U.K.). We obtained crude venom of the Chinese odorous frog, *Odorrana schmackeri* (Fig. 3), and the aim was to employ the latest peptidomic techniques available in Delft to this material. Venom of *O. schmackeri* had already been extensively studied and characterized using more conventional methodology by the Shaw group. Yet, using our high end peptidomic approach, we were able to confirm but also further complement the existing data. To test the applicability of our specific peptide analysis protocols, Prof. Shaw also provided us with skin secretion from two other anura species, the African frog *Kassina senegalensis* (Fig. 4) and the European fire bellied toad *Bombina variegata* (Fig. 5).

Fig. 3: The Chinese odorous frog Odorrana schmackeri.

Fig. 4: The African frog Kassina senegalensis.

Fig. 5: The European fire bellied toad Bombina variegata.

Frogs from Brazil we focused on are two *Phyllomedusa* species (*burmeisteri* and *rohdei,* Figures 6 and 7). This had three main reasons. First of all, Vittorio Erspamer (1985), the pioneer of frog venom peptide research called *Phyllomedusa* a true "treasure store" for bioactive peptides: "*No other amphibian skin can compete with that of the Phyllomedusae*". Second, the Shaw lab provided us access to primers designed to known *Phyllomedusa* (pre-pro)peptides, and last but not least, we had the opportunity to sample these frogs in the field ourselves in Espírito Santo State.

*Fig. 6: The Brazilian walking leaf frog*s Phyllomedusa burmeisteri*. The three specimens on right picture are doing thanatosis (playing dead), a defensive mechanism common in Hylidae frogs.*

Fig. 7: Brazilian tree-frogs Phyllomedusa rohdei.

During our work, good contacts have been established between our group at Delft and two relevant scientists in Brazil, including Prof. C. Haddad (University of São Paulo at Rio Claro) and Dr. D. Pimenta Carvalho (Butantan, São Paulo). Both agreed to provide us deeper access to the extensive Latin American amphibian fauna, so that in the future many more species which so far have not yet been investigated with respect to their venom 'peptidomes', can be studied using the methodology outlined in this PhD thesis.

1.7 Collection of *Phyllomedusa* venoms

At the beginning of this project I had the opportunity to participate in two expeditions headed by Prof. Célio Haddad, who taught me the basics of 'frog hunting' (Fig. 8). Prof. Haddad's expertise and motivation to find the target frog species hiding in the forests, ponds, on beaches, and any other probable frog habitat was an amazing experience to me, giving me the confidence to setup a subsequent 'hunting trip' on my own.

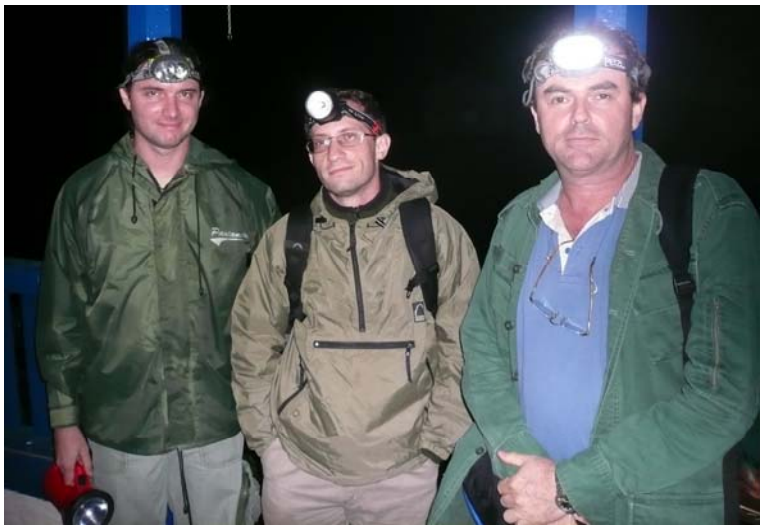


Fig. 8: Prof. Célio Haddad (right in picture) and his team during a frog hunting expedition to Serra da Bocaina National Park in November 2007.

The expeditions to collect *Phyllomedusa* venoms took place at Brunoro's Farm (latitude 20,2626 s; longitude 41,1049 w) and at National Forest of Pacotuba (latitude 20,4406 s; longitude 41,1805 w), both in southern Espírito Santo State, Brazil. They were guided by Mr. Helimar Rabello, a local biologist specialized in the fauna and flora of Espírito Santo State, and we gratefully acknowledge here the collaboration of some of his students (Fig. 9).

The hunting-trips were in the Brazilian summer season, with typical high temperatures and torrential rains. During that period tree frogs descend from the trees to complete their reproduction cycle in the abundant puddles and ponds filled with rain water. We typically started in the afternoon (in daylight) by identifying the possibly reproducing spots, near to ponds and 'dry boxes' (human-made cavities in the mountain that collect/retain rain water and prevent erosion, Fig. 9). Once daylight completely disappeared, the male frogs start to vocalize, by which the desired specimens can be identified and localized. This still remains difficult due to the excellent camouflage of certain species and their hiding skills in the

vegetation. Next to that, the presence of other nocturnal animals, like poisonous snakes (Fig. 10), and noisy mammals somewhat discourage a thoroughly extensive exploration of the area. On these trips, we were successful in collecting a good number of *Phyllomedusa burmeisteri* and *Phyllomedusa rohdei* specimens (Fig. 11).

Fig. 9: Helimar Rabello (top left) and some of his biology students who helped with Phyllomedusa *hunting. A so-called "dry box" (bottom right) was found to be a frog-rich spot.*

Captured frogs were cleaned from vegetation and mud/dust. The secretion from the dermal glands was then harvested by a non-invasive method, unharmful to the 'donor specimens' (mild electrical stimulation of 5 V, 4 ms pulses at 25 Hz[51], followed by a gentle manual (wearing gloves!) massage of the frog's upper back). Secretions were rinsed from the animal surface with a jet wash of Milli-Q water, after which samples were filtered through a 0.2 µm cellulose acetate filter, frozen, lyophilized and stored at -20˚C prior to analysis (Fig. 12).

Fig. 10: A young Bothrops jararaca *found in hunting area.*

Fig. 11: Tree-frogs Phyllomedusa burmeisteri *captured to collect the venom. All of them were found around a dry box (Fig. 9) at Brunoro's Farm. After collection of the skin secretion frogs were released back in their natural environment.*

Fig. 12: Collection of skin secretion by gentle electrical stimulation followed by manual massage, resulting in visible foam-like secretion. Secretion is rinsed from the frog with MilliQ water and filtered into clean 50 ml tubes.

Aim and outline of the thesis

The aim of this PhD thesis was to develop mass spectrometry based methods and strategies to study the peptide content of frog venoms. Making innovative use of modern mass spectrometers, in combination with peptide chemistry, new peptidomic approaches are established that aid the *de novo* sequencing and characterization of PTMs on frog skin peptides. The approaches outlined in this thesis can be adopted to study other complex and PTM containing peptide-rich samples.

In Chapter 2 an overview is given about MS approaches that could deal with all PTM types reported in the frog secretion literature: amidation, disulfide bonds, hydroxyproline, sulfotyrosine, pyroglutamic acid and D-amino acids are exemplified with the respective MS procedures and tools, from the peptides secreted by our models of study.

Besides MS techniques, transcriptomics was another approach employed to obtain the sequence of certain peptides. In Chapter 3 the study of peptidome and transcriptome of the Chinese odorous frog *Odorrana schmackeri* skin secretion is described focusing on the cloning and MS characterization of new nigrocin peptides. The known peptides from this frog were easily spotted in the crude venom by their molecular masses and their structure confirmed by MS/MS *de novo* sequencing. In this chapter, we introduce a LC/MS 2-dimensional display that gives a complete overview of the peptide content and complexity.

In Chapter 4 on yet an extra type of PTM identified in *O. schmackeri* venom is reported: O-linked glycosylation. Identification of the monosaccharide units, localization of the anchoring residue and *de novo* sequencing were achieved using ETD and CID fragmentation.

More in depth MS-based analysis on *O. schmackeri* venom revealed, unexpectedly, a different class of peptide called calcitonin-like peptides from *Odorrana schmackeri* (OsCLP-1 and 2) as reported in Chapter 5. They were manually spotted out of hundreds of MS/MS on basis of an unusual N-terminal disulfide bond. Their full primary amino acid sequence were obtained from ETD and CID spectra of intact, tryptic and labeled peptides.

Chapter 6 focuses on the method to specifically spot peptides containing disulfide bonds by a PTM-driven differential peptide display of LC/MS runs of crude and reduced frog secretion samples. Proof-of-principle of this method was obtained on the skin secretions of five different frog species.

Chapter 7 describes the combined use of molecular cloning and PTM-driven differential peptide display, for the characterization of dimeric distinctin peptides from the skin secretion of *Phyllomedusa burmeisteri*.

Chapter 8 deals with *de novo* sequencing of large intact kazal-like inhibitor peptides from *P. burmeisteri* venom, using ETD and CID fragmentation and chemical treatments.

Finally, concluding remarks and future perspectives are outlined in Chapter 9.

References

1. Simmaco, M.; Mignogna, G.; Barra, D., Antimicrobial peptides from amphibian skin: what do they tell us? *Biopolymers* **1998,** 47, (6), 435-50.
2. Bevins, C. L., Antimicrobial peptides as agents of mucosal immunity. *Ciba Found Symp* **1994,** 186, 250-60; discussion 261-9.
3. Barra, D.; Simmaco, M.; Boman, H. G., Gene-encoded peptide antibiotics and innate immunity. Do 'animalcules' have defence budgets? *FEBS Lett* **1998,** 430, (1-2), 130-4.
4. Wang, L.; Chen, Y.; Yang, M.; Zhou, M.; Chen, T.; Sui, D. Y.; Shaw, C., Peptide DV-28 amide: An inhibitor of bradykinin-induced arterial smooth muscle relaxation encoded by Bombina orientalis skin kininogen-2. *Peptides* **2010,** 31, (5), 979-82.
5. Chen, X.; Wang, L.; Wang, H.; Chen, H.; Zhou, M.; Chen, T.; Shaw, C., A fish bradykinin (Arg0, Trp5, Leu8-bradykinin) from the defensive skin secretion of the European edible frog, Pelophylax kl. esculentus: structural characterization; molecular cloning of skin kininogen cDNA and pharmacological effects on mammalian smooth muscle. *Peptides* **2011,** 32, (1), 26-30.
6. Zhou, X.; Wang, L.; Zhou, M.; Chen, T.; Ding, A.; Rao, P.; Walker, B.; Shaw, C., Amolopkinins W1 and W2--novel bradykinin-related peptides (BRPs) from the skin of the Chinese torrent frog, Amolops wuyiensis: antagonists of bradykinin-induced smooth muscle contraction of the rat ileum. *Peptides* **2009,** 30, (5), 893-900.
7. Mor, A.; Amiche, M.; Nicolas, P., Structure, synthesis, and activity of dermaseptin b, a novel vertebrate defensive peptide from frog skin: relationship with adenoregulin. *Biochemistry* **1994,** 33, (21), 6642-50.
8. Mor, A.; Hani, K.; Nicolas, P., The vertebrate peptide antibiotics dermaseptins have overlapping structural features but target specific microorganisms. *J Biol Chem* **1994,** 269, (50), 31635-41.
9. Maisetta, G.; Mangoni, M. L.; Esin, S.; Pichierri, G.; Capria, A. L.; Brancatisano, F. L.; Di Luca, M.; Barnini, S.; Barra, D.; Campa, M.; Batoni, G., In vitro bactericidal activity of the N-terminal fragment of the frog peptide esculentin-1b (Esc 1-18) in combination with conventional antibiotics against Stenotrophomonas maltophilia. *Peptides* **2009,** 30, (9), 1622-6.
10. Al-Ghaferi, N.; Kolodziejek, J.; Nowotny, N.; Coquet, L.; Jouenne, T.; Leprince, J.; Vaudry, H.; King, J. D.; Conlon, J. M., Antimicrobial peptides from the skin secretions of the South-East Asian frog Hylarana erythraea (Ranidae). *Peptides* **2010,** 31, (4), 548-54.
11. Ashcroft, J. W.; Zalinger, Z. B.; Bevier, C. R.; Fekete, F. A., Antimicrobial properties of two purified skin peptides from the mink frog (Rana septentrionalis) against bacteria isolated from the natural habitat. *Comp Biochem Physiol C Toxicol Pharmacol* **2007,** 146, (3), 325-30.
12. Bevier, C. R.; Sonnevend, A.; Kolodziejek, J.; Nowotny, N.; Nielsen, P. F.; Conlon, J. M., Purification and characterization of antimicrobial peptides from the skin secretions of the mink frog (Rana septentrionalis). *Comp Biochem Physiol C Toxicol Pharmacol* **2004,** 139, (1-3), 31-8.
13. Kim, J. B.; Iwamuro, S.; Knoop, F. C.; Conlon, J. M., Antimicrobial peptides from the skin of the Japanese mountain brown frog, Rana ornativentris. *J Pept Res* **2001,** 58, (5), 349-56.
14. Thomas, P.; Vineeth Kumar, T. V.; Reshmy, V.; Kumar, K. S.; George, S., A mini review on the antimicrobial peptides isolated from the genus Hylarana (Amphibia: Anura) with a proposed nomenclature for amphibian skin peptides. *Mol Biol Rep* **2012,** 39, (6), 6943-7.

15. Conlon, J. M.; Sonnevend, A., Antimicrobial peptides in frog skin secretions. *Methods Mol Biol* **2010,** 618, 3-14.
16. Batista, C. V.; da Silva, L. R.; Sebben, A.; Scaloni, A.; Ferrara, L.; Paiva, G. R.; Olamendi-Portugal, T.; Possani, L. D.; Bloch, C., Jr., Antimicrobial peptides from the Brazilian frog Phyllomedusa distincta. *Peptides* **1999,** 20, (6), 679-86.
17. Spencer, J. H., Antimicrobial peptides of frog skin. *Adv Enzyme Regul* **1992,** 32, 117-29.
18. Neu, H. C., The crisis in antibiotic resistance. *Science* **1992,** 257, (5073), 1064-73.
19. Bevins, C. L.; Zasloff, M., Peptides from frog skin. *Annu Rev Biochem* **1990,** 59, 395-414.
20. Rinaldi, A. C., Antimicrobial peptides from amphibian skin: an expanding scenario. *Curr Opin Chem Biol* **2002,** 6, (6), 799-804.
21. Clarke, B. T., The natural history of amphibian skin secretions, their normal functioning and potential medical applications. *Biol Rev Camb Philos Soc* **1997,** 72, (3), 365-79.
22. Kreil, G., Antimicrobial peptides from amphibian skin: an overview. *Ciba Found Symp* **1994,** 186, 77-85; discussion 85-90.
23. Erspamer, V., Amphibian skin peptides in mammals — looking ahead. *Trends in Neurosciences* **1983,** 6, (6), 200-201.
24. Dimaline, R., Is caerulein amphibian CCK? *Peptides* **1983,** 4, (4), 457-62.
25. Johnsen, A. H., Phylogeny of the cholecystokinin/gastrin family. *Front Neuroendocrinol* **1998,** 19, (2), 73-99.
26. Roelants, K.; Fry, B. G.; Norman, J. A.; Clynen, E.; Schoofs, L.; Bossuyt, F., Identical skin toxins by convergent molecular adaptation in frogs. *Curr Biol* **2009,** 20, (2), 125-30.
27. Mor, A.; Chartrel, N.; Vaudry, H.; Nicolas, P., Skin peptide tyrosine-tyrosine, a member of the pancreatic polypeptide family: isolation, structure, synthesis, and endocrine activity. *Proc Natl Acad Sci U S A* **1994,** 91, (22), 10295-9.
28. Erspamer, V.; Melchiorri, P.; Broccardo, M.; Erspamer, G. F.; Falaschi, P.; Improota, G.; Negri, L.; Renda, T., The brain-gut-skin triangle: new peptides. *Peptides* **1981,** 2 Suppl 2, 7-16.
29. Erspamer, V. a. Melchiorri, P. , Active polypeptides: from amphibian skin to gastrointestinal tract and brain of mammals. *Trends in pharmacological sciences* **1980,** 1, (14), 391-395.
30. Pounds, J. A.; Bustamante, M. R.; Coloma, L. A.; Consuegra, J. A.; Fogden, M. P.; Foster, P. N.; La Marca, E.; Masters, K. L.; Merino-Viteri, A.; Puschendorf, R.; Ron, S. R.; Sanchez-Azofeifa, G. A.; Still, C. J.; Young, B. E., Widespread amphibian extinctions from epidemic disease driven by global warming. *Nature* **2006,** 439, (7073), 161-7.
31. Conlon, J. M.; Kolodziejek, J.; Nowotny, N., Antimicrobial peptides from ranid frogs: taxonomic and phylogenetic markers and a potential source of new therapeutic agents. *Biochim Biophys Acta* **2004,** 1696, (1), 1-14.
32. Duda, T. F., Jr.; Vanhoye, D.; Nicolas, P., Roles of diversifying selection and coordinated evolution in the evolution of amphibian antimicrobial peptides. *Mol Biol Evol* **2002,** 19, (6), 858-64.
33. Seo, J.; Lee, K. J., Post-translational modifications and their biological functions: proteomic analysis and systematic approaches. *J Biochem Mol Biol* **2004,** 37, (1), 35-44.
34. Li, J.; Xu, X.; Xu, C.; Zhou, W.; Zhang, K.; Yu, H.; Zhang, Y.; Zheng, Y.; Rees, H. H.; Lai, R.; Yang, D.; Wu, J., Anti-infection peptidomics of amphibian skin. *Mol Cell Proteomics* **2007,** 6, (5), 882-94.

35. Matutte, B.; Storey, K. B.; Knoop, F. C.; Conlon, J. M., Induction of synthesis of an antimicrobial peptide in the skin of the freeze-tolerant frog, Rana sylvatica, in response to environmental stimuli. *FEBS Lett* **2000,** 483, (2-3), 135-8.
36. Mangoni, M. L.; Rinaldi, A. C.; Di Giulio, A.; Mignogna, G.; Bozzi, A.; Barra, D.; Simmaco, M., Structure-function relationships of temporins, small antimicrobial peptides from amphibian skin. *Eur J Biochem* **2000,** 267, (5), 1447-54.
37. Wechselberger, C., Cloning of cDNAs encoding new peptides of the dermaseptin-family. *Biochim Biophys Acta* **1998,** 1388, (1), 279-83.
38. Azevedo Calderon, L.; Silva Ade, A.; Ciancaglini, P.; Stabeli, R. G., Antimicrobial peptides from Phyllomedusa frogs: from biomolecular diversity to potential nanotechnologic medical applications. *Amino Acids* **2011,** 40, (1), 29-49.
39. Fleury, Y.; Vouille, V.; Beven, L.; Amiche, M.; Wroblewski, H.; Delfour, A.; Nicolas, P., Synthesis, antimicrobial activity and gene structure of a novel member of the dermaseptin B family. *Biochim Biophys Acta* **1998,** 1396, (2), 228-36.
40. Leite, J. R.; Silva, L. P.; Rodrigues, M. I.; Prates, M. V.; Brand, G. D.; Lacava, B. M.; Azevedo, R. B.; Bocca, A. L.; Albuquerque, S.; Bloch, C., Jr., Phylloseptins: a novel class of anti-bacterial and anti-protozoan peptides from the Phyllomedusa genus. *Peptides* **2005,** 26, (4), 565-73.
41. Erspamer, V.; Melchiorri, P.; Falconieri Erspamer, G.; Montecucchi, P. C.; de Castiglione, R., Phyllomedusa skin: a huge factory and store-house of a variety of active peptides. *Peptides* **1985,** 6 Suppl 3, 7-12.
42. Erspamer, V.; Melchiorri, P.; Falconieri-Erspamer, G.; Negri, L.; Corsi, R.; Severini, C.; Barra, D.; Simmaco, M.; Kreil, G., Deltorphins: a family of naturally occurring peptides with high affinity and selectivity for delta opioid binding sites. *Proc Natl Acad Sci U S A* **1989,** 86, (13), 5188-92.
43. Barra, D.; Mignogna, G.; Simmaco, M.; Pucci, P.; Severini, C.; Falconieri-Erspamer, G.; Negri, L.; Erspamer, V., [D-Leu2]deltorphin, a 17 amino acid opioid peptide from the skin of the Brazilian hylid frog, Phyllomedusa burmeisteri. *Peptides* **1994,** 15, (2), 199-202.
44. Montecucchi, P. C.; de Castiglione, R.; Erspamer, V., Identification of dermorphin and Hyp6-dermorphin in skin extracts of the Brazilian frog Phyllomedusa rhodei. *Int J Pept Protein Res* **1981,** 17, (3), 316-21.
45. Lazarus, L. H.; Bryant, S. D.; Attila, M.; Salvadori, S., Frog skin opioid peptides: a case for environmental mimicry. *Environ Health Perspect* **1994,** 102, (8), 648-54.
46. Thompson, A. H.; Bjourson, A. J.; Orr, D. F.; Shaw, C.; McClean, S., Amphibian skin secretomics: application of parallel quadrupole time-of-flight mass spectrometry and peptide precursor cDNA cloning to rapidly characterize the skin secretory peptidome of Phyllomedusa hypochondrialis azurea: discovery of a novel peptide family, the hyposins. *J Proteome Res* **2007,** 6, (9), 3604-13.
47. Gebhard, L. G.; Carrizo, F. U.; Stern, A. L.; Burgardt, N. I.; Faivovich, J.; Lavilla, E.; Ermacora, M. R., A Kazal prolyl endopeptidase inhibitor isolated from the skin of Phyllomedusa sauvagii. *Eur J Biochem* **2004,** 271, (11), 2117-26.
48. Edman, P.; Begg, G., A protein sequenator. *Eur J Biochem* **1967,** 1, (1), 80-91.
49. Chen, T.; Farragher, S.; Bjourson, A. J.; Orr, D. F.; Rao, P.; Shaw, C., Granular gland transcriptomes in stimulated amphibian skin secretions. *Biochem J* **2003,** 371, (Pt 1), 125-30.
50. Savitski, M. M.; Nielsen, M. L.; Zubarev, R. A., Side-chain losses in electron capture dissociation to improve peptide identification. *Anal Chem* **2007,** 79, (6), 2296-302.

51. Tyler, M. J.; Stone, D. J.; Bowie, J. H., A novel method for the release and collection of dermal, glandular secretions from the skin of frogs. *J Pharmacol Toxicol Methods* **1992,** 28, (4), 199-200.

Chapter 2

Mass spectrometry approaches to identify and *de novo* sequence amphibian defense peptides with post-translational modifications

Skin secretions of anura are a rich source of biologically active peptides with high pharmaceutical potential. Over the last decades several frog secretions have been subject of intense research and this has led to the discovery and characterization of several peptides with interesting biological activity. In the majority of cases, these bioactive peptides are first identified by specific bioassay screening and after this their structural sequence is further elucidated using Edman degradation and mass spectrometry. Many of the peptides identified and characterized by this approach so far are post-translationally modified. These post-translational modifications (PTMs) contribute to the bioactivity of these peptides, either by enhancing receptor binding affinity or by extending their life time due to increase resistance against proteolytic degradation. In this traditional discovery approach, a relatively big amount of venom is required for individual peptide purification and bioassays screening. However, the amphibian venoms are complex cocktails of peptides and many potentially interesting peptides risk to be missed in the bioassays and thus to escape from the investigator's attention. In this thesis, we suggest alternative, mass spectrometry based approaches to study these complex mixtures in order to discover interesting peptides. The objective of this chapter is to present an overview of the different MS-based methods we have used or developed during our research of frog skin secretions.

The overall aim is to first sequence and characterize as many peptides as possible using modern (tandem) mass spectrometry (MS), prior to performing bioassays to elucidate their activity, as suggested by Vittorio Erspamer[1]. This, however, is an enormous, labor intensive task, especially in view of the large number of different peptides which are found in several frog families. Therefore, we elected to initially focus on those peptides that carry a certain PTM, as this may be an indicator for potential bioactivity. For this we use specialized MS techniques to deal with the amphibian PTMs by [1] locating specifically these 'potentially active peptides' in a crude venom peptide profile, prior to *de novo* sequencing; [2] using chemical treatments; [3] setting the MS program to search for specific neutral losses; [4] taking advantage of the high resolution of instruments to distinguish between isobaric regular residues and modified residues; and [5] finding (novel) PTMs by *de novo* MS/MS sequencing abundant peptides that yielded information rich fragmentation patterns.

In a first step, we analyze the crude sample (washed-off skin gland secretion) by LC-MS to obtain a full profile of the sample and to evaluate its complexity. Fig. 1 shows the LC-MS based chromatogram of the crude venom of *Phyllomedusa burmeisteri* as typical example. The chromatogram contains many peaks from multiple peptides in varying charge states (Fig. 2A). Some of these peptides yield good peptide fragmentation spectra from which

partial or full sequence information can be readily obtained (as demonstrated in Fig. 2B). By performing a database search using search engines for proteomics data such as Mascot (www.matrixscience.com) or Sequest (http://fields.scripps.edu/sequest) usually several peptide identifications can be made, however, these are obviously peptides that were identified before and had been submitted to public databases such as Uniprot and NCBI. Manual interpretation of the MS/MS spectra that are not initially matched by these database searches is tremendously labor intensive and time consuming, however, at the moment this is the only way to solve an unknown (non-tryptic!) peptide sequence. To enhance the success-rate, we developed approaches to specifically target peptides that contain 'predefined' PTMs.

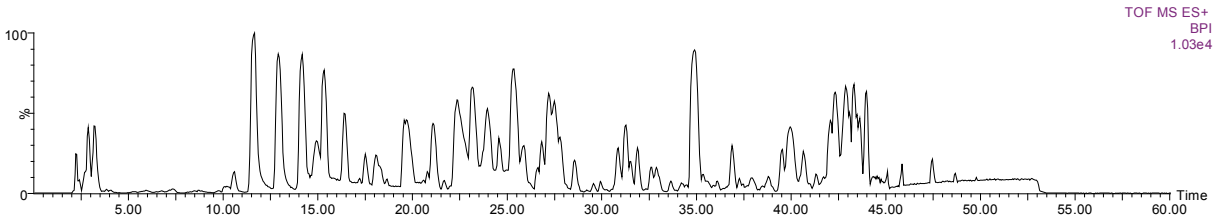


Figure 1: LC-MS Chromatogram of Phyllomedusa burmeisteri *crude venom.*

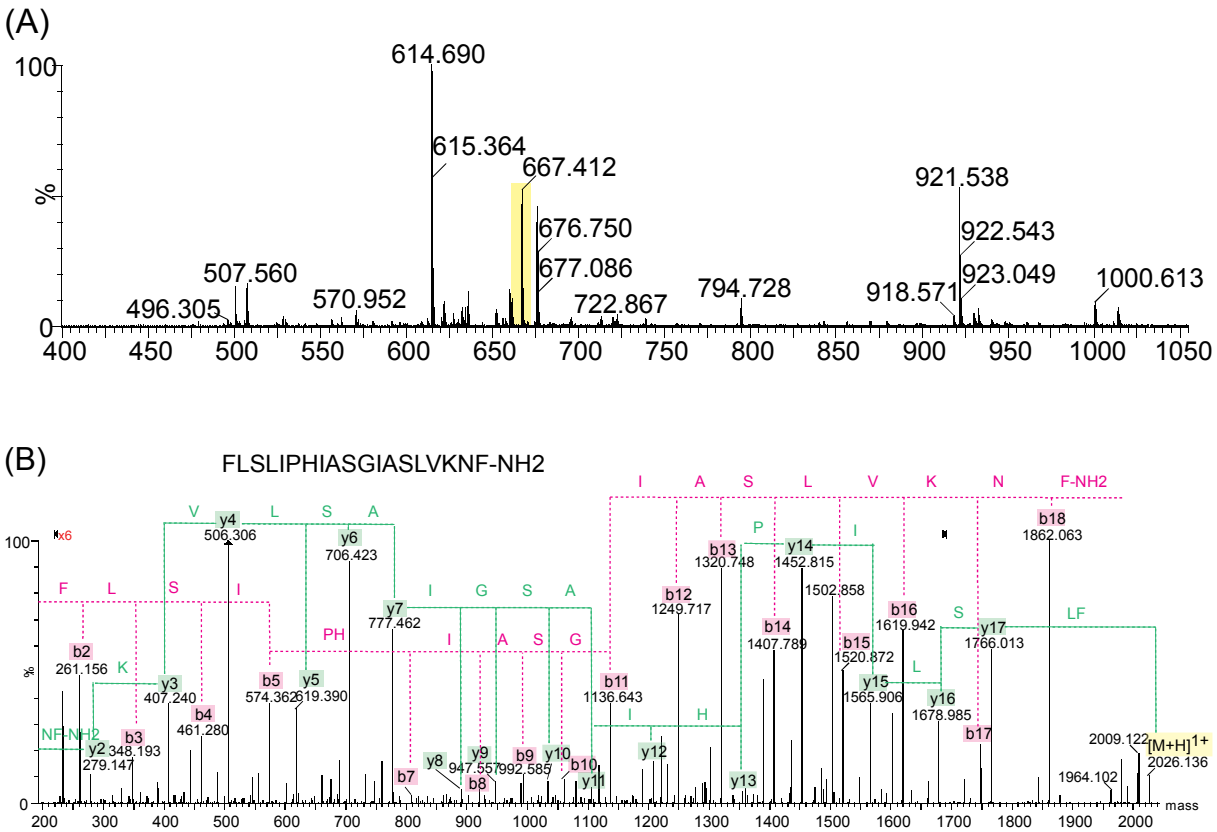


Figure 2: MS-based characterization of a peptide from P. burmeisteri. *(A) Example of MS spectrum acquired at a certain time during HPLC separation, showing several peptide ions. (B) MS/MS spectrum of phylloseptin peptide* de novo *sequenced.*

To get a more comprehensive overview of the complexity of the sample we convert the data from the LC-MS analysis into a 2-dimensional (2-D) display, in which both retention time (RT) and *mass-to-charge* ratio *(m/z)* are plotted (Fig. 3). This is done using the publicly available software tool MSight (SIB, Switzerland), which allows one to obtain a broad and detailed overview of the peptide content in the (crude) sample.

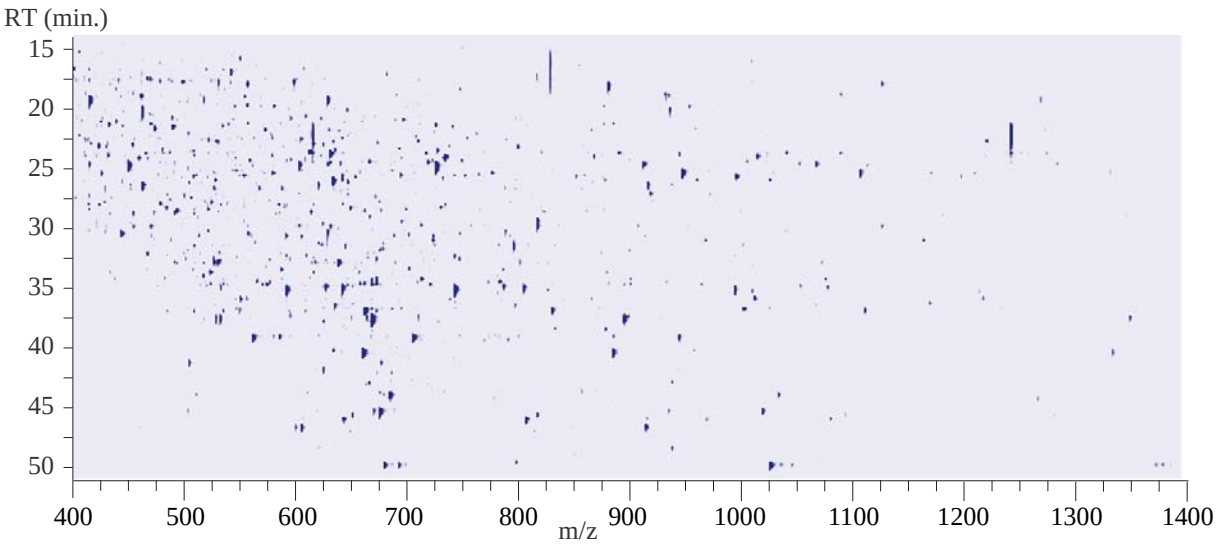


Figure 3: 2-dimensional LC-MS peptide profile of P. burmeisteri *skin secretion (picture generated using MSight), illustrating sample complexity.*

In addition, this display also enables the experienced eye to 'spot' peptides with a particular PTM. For example, a chemical or enzymatic treatment affecting certain PTMs prior to a second LC-MS analysis results in specific shifts in RT and/or *m/z*. As a consequence, peptides carrying these PTMs are easily detected by their altered location in the 2-D profile, when compared to the non-treated sample. This 2-D (differential) profiling can thus be employed to get an inventory of those peptides with selected PTMs that can subsequently be characterized by tandem mass spectrometry in more detail. To demonstrate to power of this approach we picked some illustrative examples.

[1] C-terminal amidation of peptides is one of the PTMs that can be quickly seen using the 2-D display. In many cases a small percentage of the peptide remains without amidation. The amidated and non-amidated peptide do not differ too much in chromatographic elution behavior and in most cases they elute near each other. Next to that, the mass difference between the free carboxyl and the amide group in the carboxyterminal residue is 1Da (more exactly -0.984 Da). Due to this, C-terminally amidated peptides are easily recognized in the 2-D display as two ions (isotope clusters) with a -0.984 Da shift in *m/z* and small shift in RT as illustrated in (Fig. 4A). Once this feature is detected, the fragmentation spectra of both forms

of this peptide can be analyzed and compared to acquire the full sequence (Fig. 4B). For *de novo* sequencing purposes this comparative analysis appears ideal since this PTM occurs at the final C-terminal residue and, therefore, both amidated and non-amidated peptide tandem mass spectra yielded identical *b-ions* series. On the other hand, all *y-ions* shift 0.984 Da (change of –OH to $–NH_2$ from free acid to amide). This feature greatly facilitates the sequencing as illustrated in Fig. 4B.

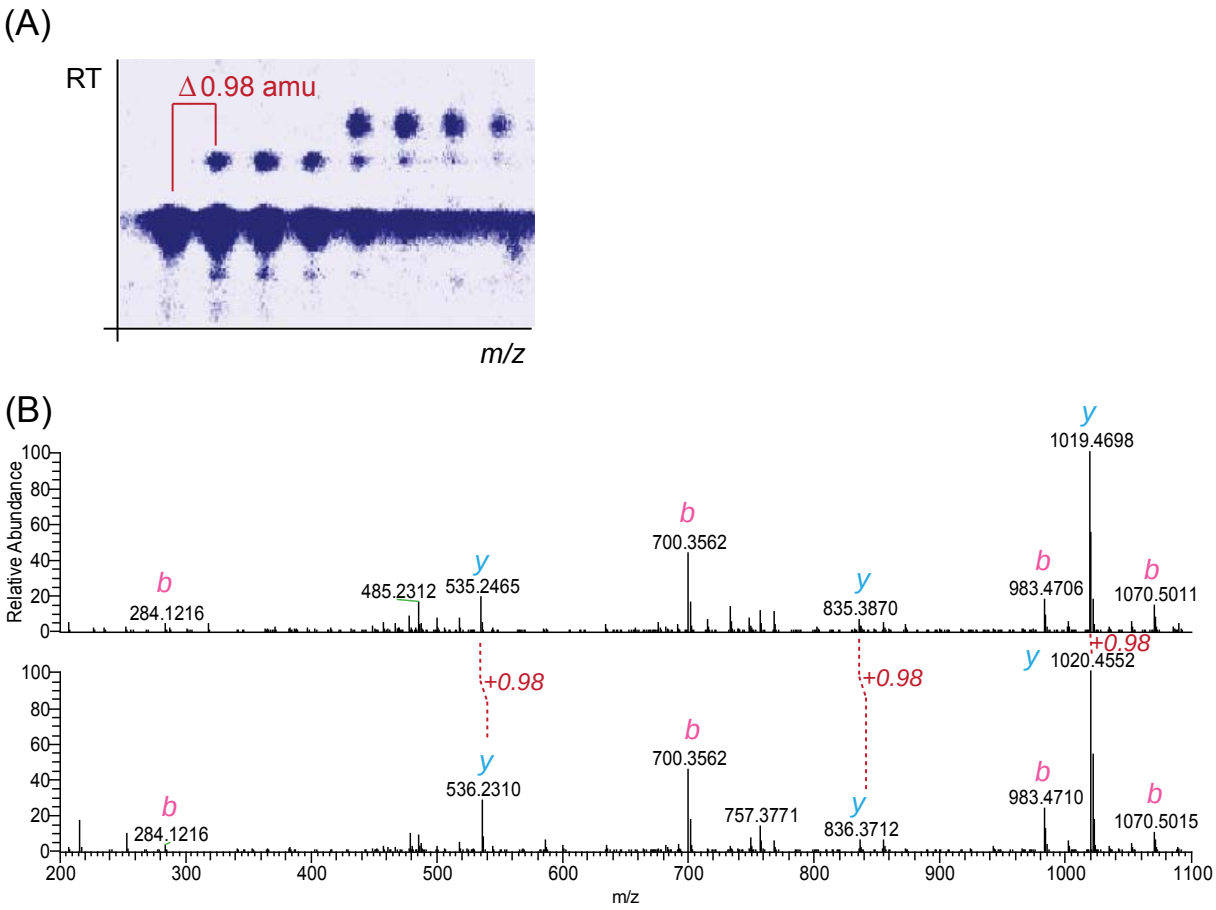


Figure 4: Peptides C-terminally amidated and non-amidated. (A) Zoom in area of 2-D display depicting peptide ion isotopes with and without amidation. Characteristic is close retention time and 0.98 amu mass difference. (B) MS/MS spectra showing similar fragmentation pattern. Masses of b-ions are identical, y-ion masses of non-amidated peptide are shifted by 0.98 amu.

[2] Using the same 2-D display, also other PTMs encountered on frog peptides can be observed. Several frog peptides have been characterized to contain a single D-amino acid at the second N-terminal position, for example deltorphin[2, 3]. Although all peptides are initially fully synthesized in their L-amino acid form, peptidyl-aminoacyl-L/D isomerases, which have been identified in frog skin before[4] may specifically alter the stereochemistry of –typically-

one amino acid within a peptide. This conversion is, however, seldom complete and therefore, both versions of the peptide (the original full L- isoform and the one with a D-residue) can be spotted. Both obviously have an identical mass, however their chromatographic behavior is slightly different. For example, we could detect deltorphin at two different RTs in *P. burmeisteri* venom with both differently eluting peptides displaying identical peptide fragmentation. Also in *Bombina variegata* skin secretion we detected several peptides that show this behavior, suggesting the presence of a D-amino acid. As example, Fig. 5A highlights five L- and D- peptide pairs in the 2-D display of *B. variegata* crude venom. The mass spectrometer cannot distinguish between the L- or D- residue as in both cases the MS/MS spectra are identical (Fig. 5B). Additional amino acid analysis is needed to solve this, however, the 2-D display allows one to quickly detect those peptides that potentially contain a D-amino acid.

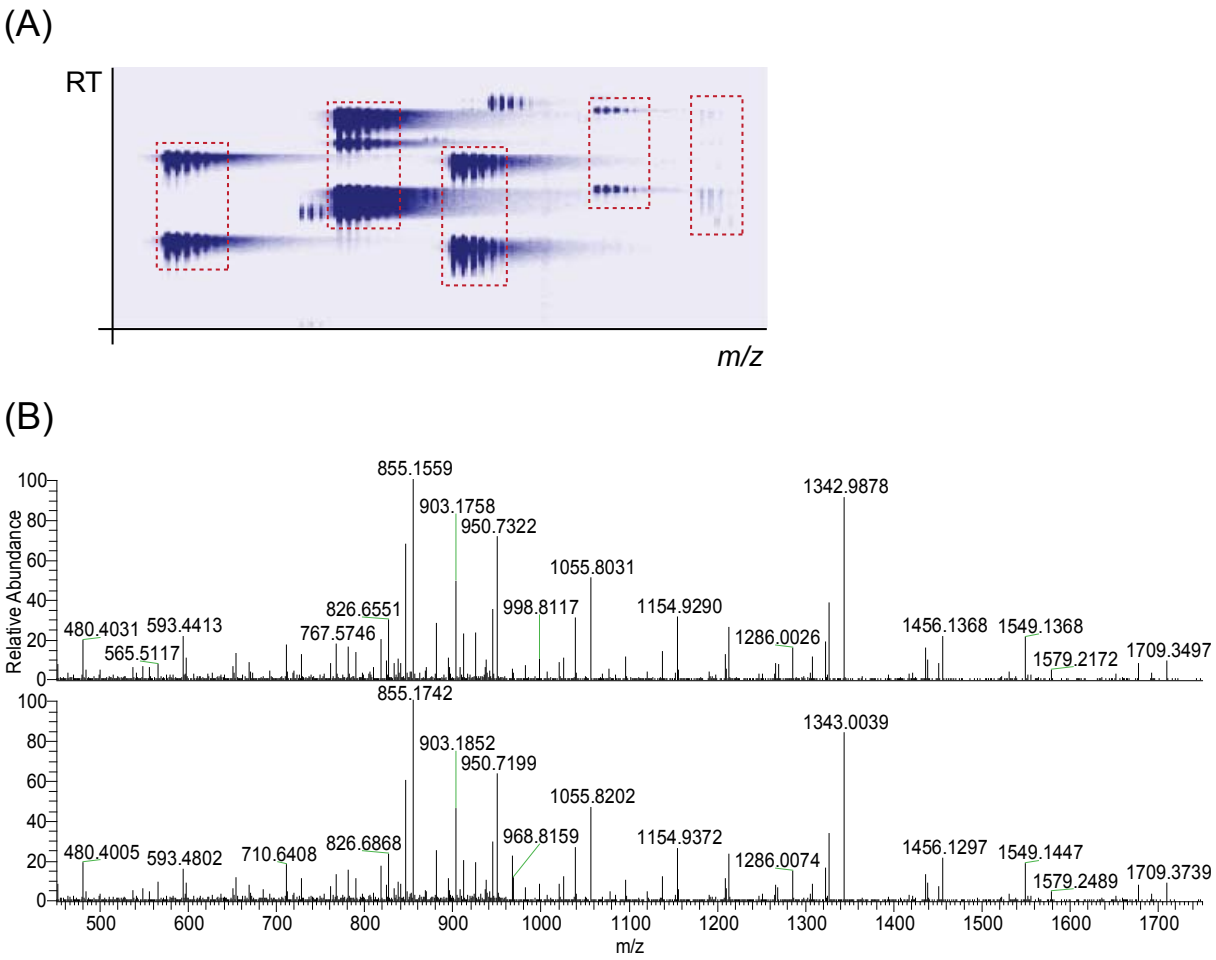


Figure 5: Five peptides pair found in B. variegata *venom, with identical mass, but with different retention times. This feature is indicative for the presence of a D-amino acid in one of two variants. (A) 2-D display focusing on peptides with identical* m/z *value and different RT. (B) Example of identical MS/MS spectra of two such peptides.*

[3] In the course of our PTMs searching and verification of MS/MS spectra, we observed that many peptides could not be efficiently fragmented, with MS/MS spectra being composed of only few ion peaks, which did not reveal the full primary structure (Fig. 6A). This is often caused by peptides with Cys residues forming disulfide bonds – a cystine. The cyclic nature of the peptide hampers the backbone cleavage and the MS/MS spectra are relatively poor in fragment ions. By using chemical reagents, such as dithiothreitol (DTT) or tris 2-carboxyethyl phosphine (TCEP), the thiol groups from Cys can be reduced, and subsequently the peptides are much better fragmented, as illustrated in Fig. 6B. Additionally, a method based on the 2-D display to spot all the peptides containing cystine was developed and will be described in more detail in Chapter 6.

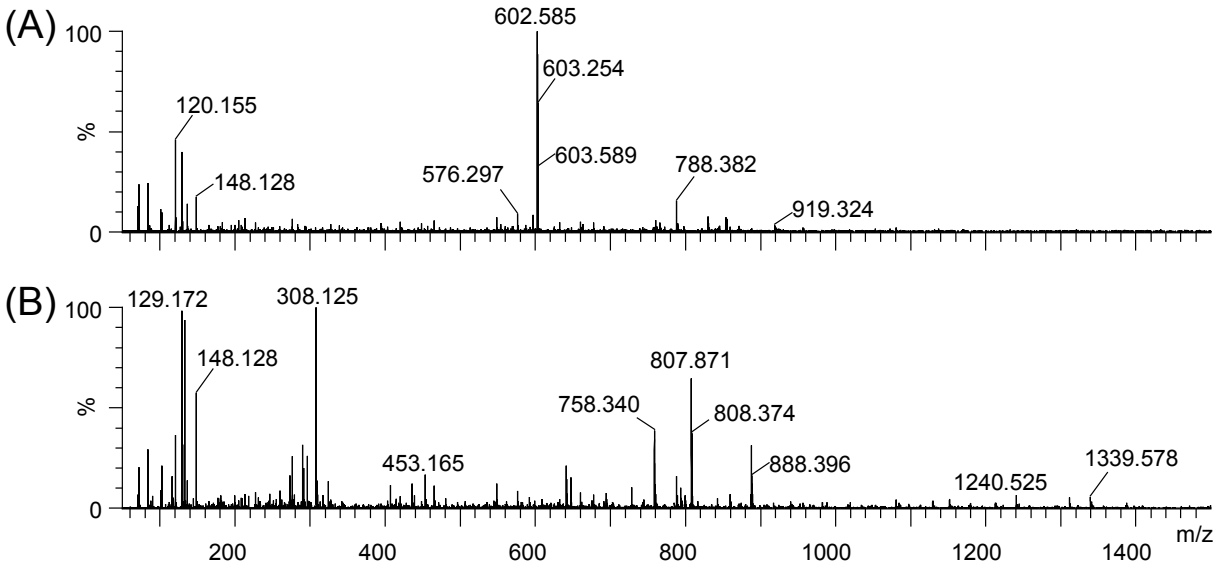


Figure 6: Peptide containing disulfide bridges. (A) MS/MS spectrum of peptide with S-S bonds, showing deficient fragmentation. (B) MS/MS spectrum of same peptide after reduction and alkylation of the free thiol groups, showing good peptide backbone fragmentation.

[4] Other PTMs which we occasionally encountered on frog peptides are hydroxyproline (Hyp or HyPro), Tyr-sulfation and pyroglutamate (pGlu). For example, two phyllokinins were sequenced from *P. burmeisteri* with identical structure to the original peptide identified previously in *P. hypochondrialis*[5] – RPPGFTPFRIY and RPPGFSPFRIY. We also found both peptides in their modified version, with the third proline residue modified to Hyp and the C-terminal residue modified to a sulfated tyrosine. The Hyp residue in phyllokinins could be distinguished from Leu/Ile using the high resolution (7.500) orbitrap analysis. As indicated in the spectrum of Fig. 7, the difference of *y8* and *y9-ions* is 113.0481 Da, which does not fit to the Ile/Leu mass (113.0841 Da). A second example is the peptide

originally described from *P. hypochondrialis* as "phypo Xa"[6], pEFRPSYQIPP. This peptide was found to contain an aminoterminal pyroglutamic acid residue.

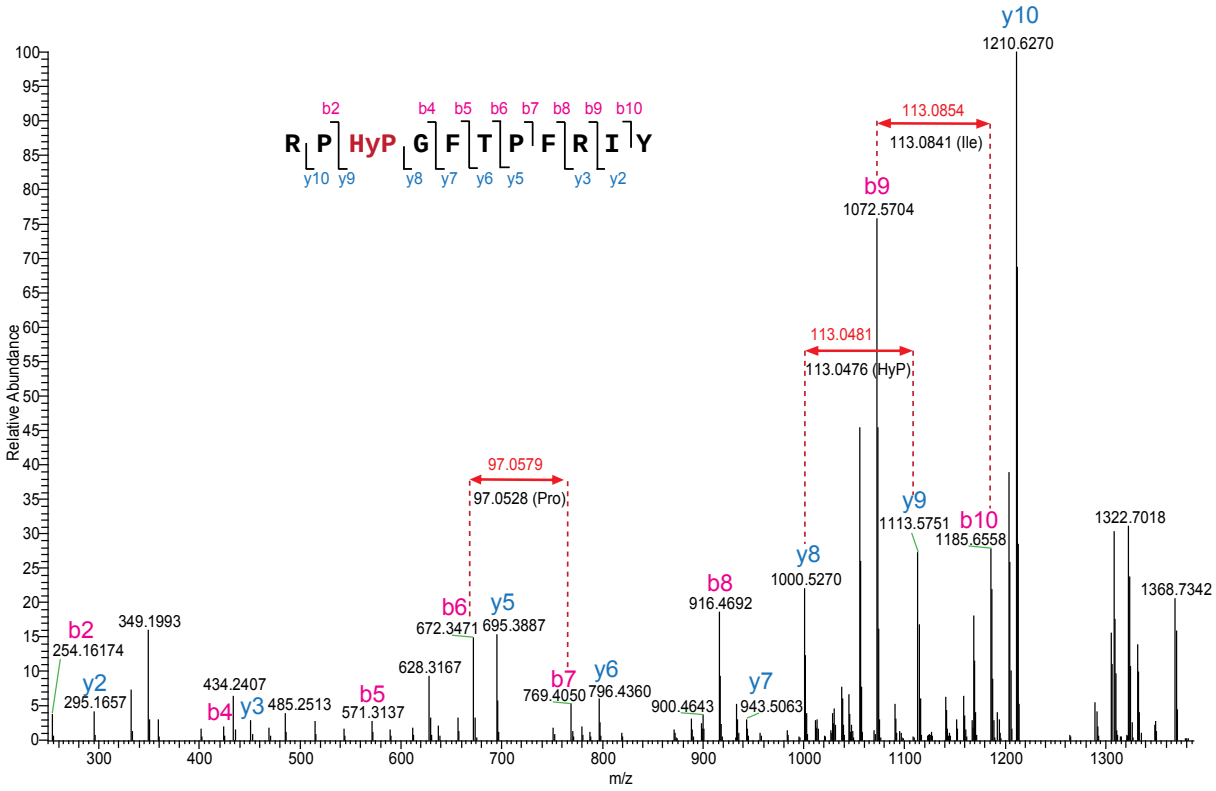


Figure 7: Fragmentation spectrum of peptide containing hydroxyproline (phyllokinin). MS resolution allowed differentiating HyP from Ile/Leu.

None of these PTMs (HyPro, Tyr sulfations and pGlu) can be directly "spotted" on a 2-D peptide display. However, a 2-D differential display method (also see Chapter 6) could in principle be developed to spot the peptides carrying these PTMs. For this, the crude venom is analyzed twice, once without any treatment and once after it is either chemically, or enzymatically treated to alter the PTM of interest. From the comparison of the 2-D displays, the peptides with altered RT and or *m/z* ratio are easily detected as new "spots". In the case of pGlu-containing peptides, treatment of the crude venom with the pyroglutamate aminopeptidase from *Pyrococcus furiosus* could be done. This enzyme removes the pyroglutamic acid from the N-terminus and is frequently used to deblock pGlu peptides prior to Edman degradation. As another example, sulfotyrosine containing peptides could be detected in a similar manner after enzymatic removal of the sulfate group[7]. In both these examples, the treatment will lead to a mass decrease of either 111 or 80 Da for pGlu acid and sulfoTyr, respectively. Additionally, such approaches can also be developed to find other PTMs and to screen the venoms of other organisms.

[5] Finally, low energy CID spectra of sulfated peptides are often dominated by neutral loss of the SO_3 group (-79.9595 amu), leading to a relatively poor sequence information for identification. However, the specific neutral loss scan can be used to trigger an MS^3 experiment (Fig. 8). In a data dependent neutral loss driven analysis of *P. burmeisteri* crude venom, all doubly charged peptides that showed a loss of 39.9798 in terms of *m/z* in their MS^2 spectrum, were used as a trigger for a second stage of fragmentation analysis on the daughter ion (MS^3 spectra, Fig. 9). To improve the efficiency of this method by increasing the chance of the selection of a sulfated peptides, the same sample was analyzed 4 times, in each analysis a specific mass range was set: from *m/z* 400 to 600, 600 to 800, 800 to 1000 and 1000 to 1200. In this case more MS^2/MS^3 triggers were observed when compared to analysis performed with 1 broad mass band (*m/z* 400 to 1500). An example of a sulfated peptide discovered using this MS-based method is a phyllokinin, which shows no backbone fragmentation in MS^2 but a neutral loss of sulfated. From the MS^3 it was possible to *de novo* sequence this tyrosine sulfated peptide (Fig. 10). Due to the high resolving power and accuracy of the orbitrap analyzer only peptides losing the sulfate group lead to trigger for MS^3 making it a highly specific screening strategy for this PTM.

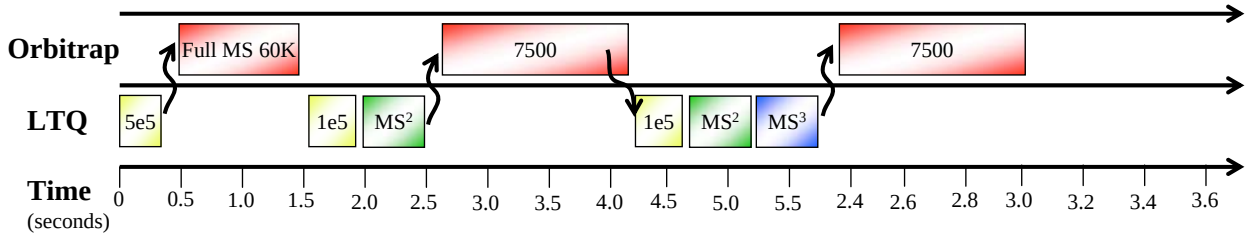


Figure 8: Schematic overview of neutral loss driven data-dependent MS^2/MS^3 scans-settings used for detection of tyrosine-sulfated peptides.

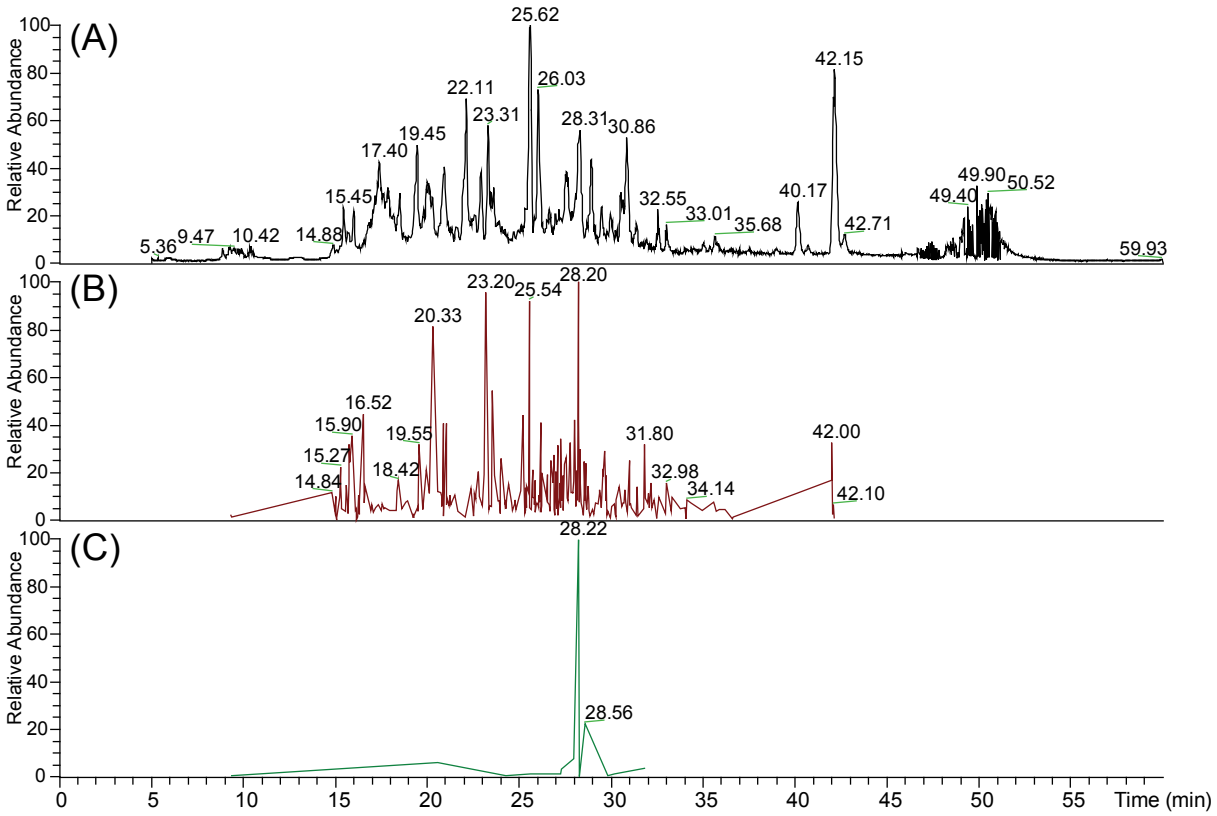


Figure 9: Chromatograms of data dependent analysis using sulfate neutral loss driven MS^3 analysis of sulfated peptides from skin secretion of P. burmeisteri. *Mass range of this measurement was between 600 to 800* m/z. *Shown are the filtered trace of (A) MS, (B) MS^2 and (C) MS^3 spectra.*

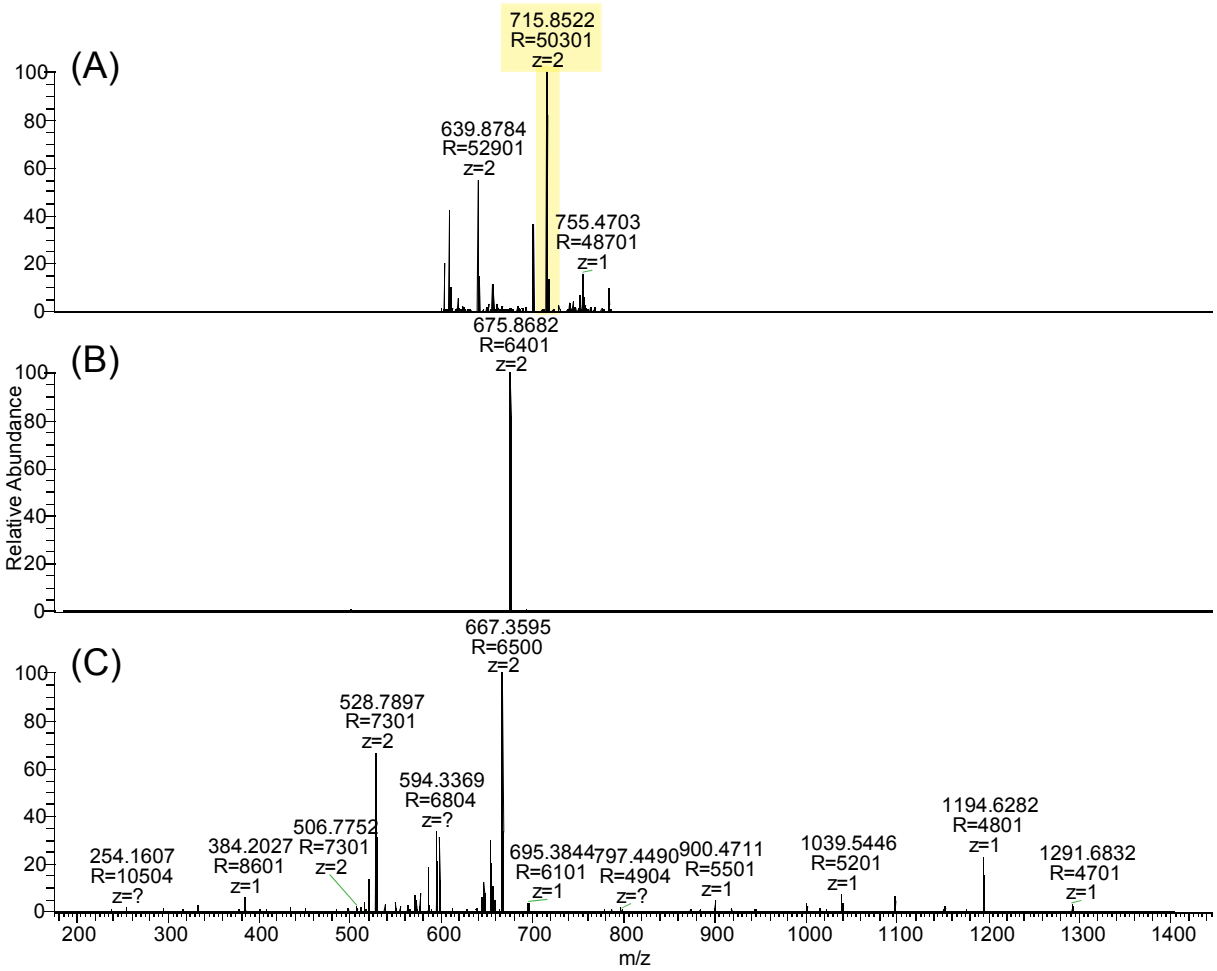


Figure 10: (A) Full MS spectrum of phyllokinin [715.8522+2H]$^{2+}$, (B) MS^2 spectrum showing predominant sulfate neutral loss [675.8682+2H]$^{2+}$and (C) MS^3 spectrum with rich fragmentation, from which the primary amino acid sequence could be derived.

References

1. Erspamer, V.; Melchiorri, P.; Falconieri Erspamer, G.; Montecucchi, P. C.; de Castiglione, R., Phyllomedusa skin: a huge factory and store-house of a variety of active peptides. *Peptides* **1985,** 6 Suppl 3, 7-12.
2. Barra, D.; Mignogna, G.; Simmaco, M.; Pucci, P.; Severini, C.; Falconieri-Erspamer, G.; Negri, L.; Erspamer, V., [D-Leu2]deltorphin, a 17 amino acid opioid peptide from the skin of the Brazilian hylid frog, Phyllomedusa burmeisteri. *Peptides* **1994,** 15, (2), 199-202.
3. Erspamer, V., The opioid peptides of the amphibian skin. *Int J Dev Neurosci* **1992,** 10, (1), 3-30.
4. Jilek, A.; Mollay, C.; Lohner, K.; Kreil, G., Substrate specificity of a peptidyl-aminoacyl-L: /D: -isomerase from frog skin. *Amino Acids* **2012,** 42, (5), 1757-64.
5. Brand, G. D.; Krause, F. C.; Silva, L. P.; Leite, J. R.; Melo, J. A.; Prates, M. V.; Pesquero, J. B.; Santos, E. L.; Nakaie, C. R.; Costa-Neto, C. M.; Bloch, C., Jr., Bradykinin-related peptides from Phyllomedusa hypochondrialis. *Peptides* **2006,** 27, (9), 2137-46.
6. Conceicao, K.; Konno, K.; de Melo, R. L.; Antoniazzi, M. M.; Jared, C.; Sciani, J. M.; Conceicao, I. M.; Prezoto, B. C.; de Camargo, A. C.; Pimenta, D. C., Isolation and characterization of a novel bradykinin potentiating peptide (BPP) from the skin secretion of Phyllomedusa hypochondrialis. *Peptides* **2007,** 28, (3), 515-23.
7. Vries, M. P. d.; Rullens-Ligtvoet, N. C. C. J. M.; Hoeberichts, W. J.; Gool, A. J. v.; Zeeland, M. J. M. v.; Kouwijzer, M.; Verheijden, G. F.; Boots, A. M. H.; Verhaert, P. D. E. M.; Bos, E. S., Identification of Tyrosine-O-Sulfation as Post-Translational Modification of Melanoma Inhibitory Activity (Mia) Protein. *Biomacromolecular Mass Spectrometry* **2005,** 1, (1), 57-73.

Chapter 3

Nigrocin-2 peptides from Chinese *Odorrana* frogs – integration of UPLC/MS/MS with molecular cloning in amphibian skin peptidome analysis

Abstract

Peptidomics is a powerful set of tools for the identification, structural elucidation and discovery of novel regulatory peptides and for monitoring the degradation pathways of structurally and catalytically important proteins. Amphibian skin secretions, arising from specialized granular glands, often contain complex peptidomes containing many components of entirely novel structure and unique site-substituted analogues of known peptide families. Following the discovery that the granular gland transcriptome is present in such secretions in a PCR-amenable form, we designed a strategy for peptide structural characterization involving the integration of 'shotgun' cloning of cDNAs encoding peptide precursors, deduction of putative bioactive peptide structures, and confirmation of these structures using tandem MS/MS sequencing. Here, we illustrate this strategy by means of elucidation of the primary structures of nigrocin-2 homologues from the defensive skin secretions of four species of Chinese *Odorrana* frogs, *O. schmackeri, O. livida, O. hejiangensis* and *O. versabilis*. Synthetic replicates of the peptides were found to possess antimicrobial activity. Nigrocin-2 peptides occur widely in the skin secretions of Asian ranid frogs and in those of the *Odorrana* group, and are particularly well-represented and of diverse structure in some species. Integration of the molecular analytical technologies described provides a means for rapid structural characterization of novel peptides from complex natural libraries in the absence of systematic online database information.

This Chapter is published as:

Wang, L.; Evaristo, G.; Zhou, M.; Pinkse, M.; Wang, M.; Xu, Y.; Jiang, X.; Chen, T.; Rao, P.; Verhaert, P.; Shaw, C., Nigrocin-2 peptides from Chinese Odorrana frogs-integration of UPLC/MS/MS with molecular cloning in amphibian skin peptidome analysis. *Febs J* **2010,** 277, (6), 1519-31.

3.1 Introduction

Proteomics has become a field of biochemical research in its own right, whose techniques are widely employed by biological/biomedical scientists from many disciplines[1, 2]. Studies on 'model organisms', including humans, are facilitated by the fact that, in most cases, the entire genome has been sequenced, and, consequentially, translated protein datasets are readily available through public online databases[1, 2]. High-resolution protein separation technologies and mass spectrometry have been fundamental to development of this discipline, and a typical scheme involves 2D gel electrophoresis, tryptic digestion of resolved protein 'spots', MS/MS fragmentation sequencing of several tryptic fragments, and identification of the parent protein from generated primary structural data through database interrogation[3-5]. Peptidomics, a daughter discipline, focuses on characterization of endogenous low-molecular-mass peptides (no trypsin digestion is necessary prior to MS analysis), and differs from proteomics in the initial separation technology, whereby 2D HPLC is used rather than 2D gel electrophoresis[6, 7]. However, both techniques are less successful and their outputs massively reduced if the source of the proteins and peptides is a more exotic organism whose online sequence database coverage is poor or non-existent[3-7]. The peptidomes of amphibian defensive skin secretions have been the focus of our research for over a decade, and many species fall into this poorly represented category.

Many anuran amphibians secrete a complex peptide-based skin secretion from specialized granular glands in response to stress, which is most often related to attack by a predator but may be elicited by a stimulus as simple as handling[8, 9]. Among the plethora of biologically active peptides, those that display broad-spectrum antimicrobial activity usually predominate. For this reason, and because novel therapeutics are urgently required to address the global emergence of multiple drug resistance in human pathogens, this class has been the most studied to date[10-14]. Typical of the multiple structural classes of antimicrobial peptides elucidated to date are the bombinins from Eurasian bombinid toads, the magainins from African pipid toads, the dermaseptins from South / Central American phyllomedusine frogs, the caerins from Australasian
litorid frogs, and the esculentins, brevinins and temporins from the ranid frogs of Asia, North America and Europe[8, 9, 11]. Many of these peptides have broad-spectrum activity against Gram-positive and Gram-negative bacteria, some have limited effects on fungi and some have additional undesirable cytolytic effects on red blood cells[11, 12]. These attributes are mostly a

consequence of their amphipathic helical and cationic nature – a feature exhibited by some protein fragments that share the membrane-interacting and perturbing effects on bacteria[15-17]. Recently, an antimicrobial amphibian skin peptide of an anionic nature has been described[18], but this class of peptide would not appear to be common in amphibian defensive skin secretions.

Nigrocins 1 and 2 are cationic peptides that were originally discovered in the skin secretions of the common and widely distributed Oriental frog, *Rana nigromaculata* (now *Pelophylax nigromaculatus*) as a consequence of their antimicrobial activity[19]. While nigrocin-1 showed a high degree of primary structural similarity to the established brevinin-2 family, raising doubts about its novel name[20], nigrocin-2 was of sufficient primary structural novelty to warrant its name and status as the prototype of a novel family of amphibian skin antimicrobial peptides. The nomenclature of these skin peptides has become a significant topic of debate in recent times[19-21], as the numbers of primary structures available that pay no heed to established peptide family names create a significant degree of confusion for specialists and non-specialists alike. As an example, nigrocin-2 related peptides were isolated from the skin secretion of the odorous frog, *Odorrana grahami*, and the very same peptides were named as nigrocin-2GRa-c[22], grahamins[23] and nigrocin OG17 and 20[24].

Here, we describe a multidisciplinary approach that serves to rapidly and effectively circumvent the bottle-neck that occurs in novel peptide identification and structural characterization. This involves the integration of [1] 'shotgun' cloning of skin secretion-derived cDNAs to predict encoded peptide primary structures with [2] HPLC or UPLC analysis of the very same skin secretion sample, combined with mass spectrometry (MS) technology to confirm the actual presence of the translationally mature predicted peptides. Whereas in our conventional analysis scheme, candidate peptides of interest were located after chromatography by MALDI-TOF MS with their primary structures separately confirmed by Edman degradation, we demonstrate here that an alternative scheme based upon online UPLC-MS/MS technology can achieve unequivocal sequence confirmation of actual amphibian skin secretion from approximately three orders of magnitude less material and without prior peptide purification. We illustrate this here with 3 novel nigrocin-2 analogues whose structures were predicted from cloned skin cDNAs of the Oriental frog *Odorrana schmackeri* and unambiguously identified in the skin secretions by UPLC MS/MS. Finally, these peptides could be attributed a biological function. Using synthetic replicates of each peptide, antimicrobial activity was demonstrated of a potency and spectrum consistent with previous reports for other members of the family. In addition to revealing widespread

distribution of multiple nigrocin-2 analogues in the skin secretions of Oriental ranid frogs, these data illustrate the applicability of our multi-disciplinary approach to rapidly identify and characterize novel peptides in complex peptidomes.

3.2 Materials and Methods

3.2.1 Specimen biodata and secretion harvesting.

All frog specimens investigated in this study were captured in their natural habitats in the People's Republic of China. *Odorrana schmackeri,* (n=4, snout-to-vent length 5-7 cm) were captured during expeditions in Wuyishan National Park, Fujian Province. Specimens of *Odorrana livida* (n=4) and *Odorrana hejiangensis* (n=6) of similar size, were collected in various locations in Fujian and Shaanxi Provinces as were four specimens of *Pelophylax nigromaculatus*. Specimens of *Odorrana versabilis* (n=3, snout-to-vent length 7-10 cm) were collected in the Five Fingers Peak Nature Reserve in Hainan. All frogs were adults of undetermined sex, and secretion harvesting was performed by gentle transdermal electrical stimulation as described previously[25]. Stimulated secretions were washed from the frogs with deionized water, snap-frozen in liquid nitrogen, lyophilized, and stored at -20 ^{o}C prior to analyses.

3.2.2 Molecular cloning of nigrocin-2 peptide biosynthetic precursor cDNAs from skin secretion-derived libraries

Five milligram samples of lyophilized skin secretion from each species of *Odorrana schmackeri* frog were separately dissolved in 1 ml of cell lysis/mRNA stabilisation solution (Dynal, Biotech, Bromborough, UK). Polyadenylated mRNA was isolated using magnetic oligo(dT) beads as described by the manufacturer (Dynal Biotech, UK). The isolated mRNA was subjected to 5' and 3' RACE procedures to obtain full-length nigrocin-2 precursor nucleic acid sequence data using a SMART-RACE kit (Clontech, Basingstoke, UK), essentially as described by the manufacturer. Briefly, the 3'-RACE reactions employed a nested universal primer (supplied with the kit) and a degenerate sense primer (N2-S1; 5'-GGIYTIYTIWSIAARGT -3') (I = deoxyinosine, Y = C or T, W = A or T, S = C or G, R = A or G) complementary to the N-terminal sequence of nigrocin-2, GLLSKV, from *Pelophylax nigromaculatus*[19]. The products of 3' RACE reactions were gel-purified and cloned using a

pGEM-T vector system (Promega Corporation, Madison, WI, USA) and sequenced using an ABI 3100 automated sequencer. The sequence data obtained from these 3' RACE products were used to design a gene-specific antisense primer, N2-AS1 (5'-CCACATMAGATKATTTCYGATTYAA-3') (M = A or C, K = T or G), to a common region of the 3' non-translated regions. 5' RACE was performed using this specific primer in conjunction with the nested universal primer, and the generated products were gel-purified, cloned, and sequenced as described above. Following acquisition of these data, a second gene-specific sense primer (N2S2, 5'–GTTCACCWYGAAGAAATCCMTKYTACT–3') was designed to a region in the putative signal peptide domain, and was employed in 3' RACE reactions. Products were gel-purified, cloned and sequenced as described previously.

3.2.3 Identification and structural analysis of nigrocins

3.2.3.1 Workflow 1: conventional HPLC, MALDI-TOF MS and automated Edman sequencing

Five milligram samples of lyophilized skin secretion from each species of frog were separately dissolved in 0.5 ml of 0.05/99.5 (v/v) trifluoroacetic acid (TFA)/water and clarified of microparticles by centrifugation at 1500 g for 10 min at 4 °C. The clear supernatant was separately fractionated by injecting directly onto a reverse phase HPLC column (Phenomenex C-18, 25 cm in length × 0.45 cm in width; Phenomenex, Torrance, CA, USA), and peptides were eluted using a gradient formed from 0.05/99.5 (v/v) TFA/water to 0.05/19.95/80.0 (v/v/v) TFA/water/acetonitrile over 240 min at a flow rate of 1 ml/min. A Cecil CE4200 Adept gradient reverse-phase HPLC system (Cambridge, UK) was used, and fractions were collected automatically at 1 min intervals. One microlitre of each chromatographic fraction was prepared for mass analysis using MALDI-TOF MS on a linear time-of-flight Voyager DE mass spectrometer (Perseptive Biosystems, Framinham, MA, USA) in positive detection mode using α-cyano-4-hydroxycinnamic acid as the matrix. Internal mass calibration of the instrument was achieved using standard peptides of established molecular mass providing a determined accuracy of ±0.1%. Peptides in the molecular mass range of nigrocin-2 peptides (1.9 – 2.2 kDa) were subjected to primary structural analysis by automated Edman degradation using an Applied Biosystems 491 Procise sequencer in pulsed-liquid mode. Nigrocin-2-related peptides identified by this approach were synthesized by standard solid-phase Fmoc chemistry using a Protein Technologies (Tucson, AZ, USA) PS3^{TM} automated

peptide synthesizer. Following cleavage from the synthesis resin, impurities were removed from each synthetic replicate by reverse phase HPLC and molecular masses of purified products were confirmed by MALDI-TOF mass spectrometry as identical to the corresponding natural peptides. The structural identities between respective natural and synthetic peptides were also confirmed subsequently by MS/MS fragmentation.

3.2.3.2 Workflow 2: on-line nanoUPLC-MS and MS/MS

One milligram of the lyophilized skin secretion from *O. schmackeri* was reduced (30 min in 0.2 mM DTT, 25 mM NH_4HCO_3) and alkylated (30 min in 0.4 mM iodacetamide, 25 mM NH_4HCO_3) to break disulphide bridges and to stabilise the resulting cysteine side-chain sulfhydryl groups, respectively. Two micrograms aliquots of this sample were analysed on a nano-UPLC (nanoAcquity, Waters, Manchester, UK) directly coupled to a Q-TOF hybrid tandem mass spectrometer (QTOF Premier, Waters).

For nanoLC-MS, a 2 µg equivalent of reduced and alkylated peptide sample was delivered to a trap column packed in-house with Phenomenex W-POREX C4 (5 µm particle size, 200 Å, Phenomenex, Torrance, CA, USA) at a flow rate of 5 µl/min. After 10 min. trapping, the trap column was switched in-line with the C4 analytical column (50 µm ID x 200 mm length, packed in-house). The flow was reduced to 150 nl/min. and a linear gradient from 0 to 40 % Solvent B (0.1 M acetic acid in 8:2, v/v, acetonitrile/water) at 1% min^{-1} analytically separated the content of the trap column. The column effluent was directly electrosprayed into the ESI-source of the mass spectrometer using a nano-ESI emitter (New Objective, Woburn, MA, USA). A first LC MS "survey" run served to establish the specific retention times of the three potential nigrocin-2 related peptides according to their predicted *m/z* values (Figure 7). Using MSight software (Swiss Institute of Bioinformatics), two-dimensional images were also generated by plotting peptide retention time versus mass spectrum (*m/z* values). These representations of a LC MS run, so-called "peptide displays", can be used to illustrate the retention time of different peptides as well as the total complexity of a sample in terms of detectable MS signals (Figure 11). In a second (replicate 2 µg equivalent injection) LC MS/MS run, the nigrocin peaks were manually selected at their retention times for collision-induced fragmentation. Multiply-charged precursor fragmentation spectra were deconvoluted using MaxEnt3 (Waters) software to assist in sequence interpretation.

3.2.4 Antimicrobial activity of nigrocin-2 peptides

Each nigrocin-2 related peptide was subjected to a range of antimicrobial assays to determine minimal inhibitory concentrations (MICs) using non-pathogenic standard laboratory strains of a Gram-positive bacterium (*Staphylococcus aureus*) and a Gram-negative bacterium (*Escherichia coli*). Samples of each novel peptide were initially reconstituted in 200 µl of NaCl/Pi, pH 7.2, to achieve a concentration of 100 µM, and doubling dilutions were prepared from this stock solution in nutrient broth. MICs for each peptide against both test organisms were assessed by incubating peptides in nutrient broth following inoculation with 50 µl of overnight standard cultures (containing 10^4 CFU) into 96-well microtiter cell-culture plates. Plates were incubated for 18 h at 37 ^{o}C in a humidified atmosphere. The growth of bacteria was determined by measuring attenuance at 550 nm using a microtiter plate reader (MA Bioproducts, Walkersville, MD, USA, model MA308). MICs were taken as the lowest concentration of peptides for which no visible growth was observed.

3.3 Results

3.3.1 Molecular cloning of nigrocin-2 precursor cDNAs from skin secretion-derived cDNA libraries

Nigrocin-2 related peptide precursor cDNAs were consistently cloned from each skin secretion-derived library. A single transcript was amplified from the *P. nigromaculatus* library, and this encoded the biosynthetic precursor of the archetypal nigrocin-2[17]. Single transcripts encoding nigrocin-2 related peptide precursors were also cloned from skin secretion cDNA libraries of *O. hejiangensis* and *O. versabilis*. In contrast, two transcripts encoding different nigrocin-2 related peptide precursors were consistently cloned from the *O. livida* skin secretion cDNA library, and three from the library generated from *O. schmackeri* skin secretions. The nucleotide sequences of each cloned nigrocin-2 biosynthetic precursor cDNA and the amino acid sequences of their open-reading frame are shown in Figures 1 and 2. In common with many other amphibian skin peptide precursor transcripts, especially those encoding structurally related analogues, high degrees of nucleotide sequence identity were found within the open reading frames of nigrocin-2 precursors (Fig. 3). The domain architecture of each open reading frame was likewise highly conserved, and was consistent with the organization found in many other amphibian skin peptide precursors (Fig. 4). The putative signal peptide of 22 amino acid residues preceded an acidic amino acid residue-rich spacer peptide that contained a putative dibasic amino acid (-RR-) propeptide convertase cleavage site in all nigrocin-2 transcripts. Further into the sequence, a typical convertase cleavage site (-KR-) was found to be present, and cleavage of this generated the C-terminally located mature nigrocin peptide in each case. Of major interest was the finding that the mature peptides nigrocin-2N, nigrocin-2LVa and nigrocin-2HJ are of identical primary structure. The nigrocin-2-related peptides described here are aligned with those isolated previously from *Odorrana* frogs in Fig. 5. It is noteworthy that the same peptides from one species are sometimes described under disparate names (see peptides from *Odorrana grahami*, Fig. 5C, D). This study identified nigrocin-2 peptides of identical primary structure from different species. This is a most unusual finding, as the primary structures of most amphibian skin antimicrobial peptides are typically poorly conserved between closely related species[26].

```
(A)     Nigrocin-2N
         M  F  T   T  K  K  S   I  L  L   L  F  F   L  G  M  V
  1   ATGTTCACCA CAAAGAAATC CATTCTACTC CTTTTCTTTC TTGGGATGGT
         S  L  A   L  C  E   Q  E  R  D   A  N  E   E  E  R  R
 51   CTCCTTAGCT CTCTGTGAGC AAGAGAGAGA TGCCAATGAG GAAGAAAGAA
           D  E  L   D  E  R   D  V  E   A  I  K  R   G  L  L
101   GAGATGAGCT TGATGAAAGG GATGTTGAAG CGATAAAGCG AGGTCTATTA
         S  K  V   L  G  V  G   K  K  V   L  C  G   V  S  G  L
151   AGCAAAGTCC TCGGCGTCGG GAAGAAAGTA CTATGTGGAG TTAGTGGGCT
         C  *
201   TTGTTAA

(B)     Nigrocin-2SCa
         M  F  T   L  K  K  S   I  L  L   L  F  F   L  G  T  I
  1   ATGTTCACCT TGAAGAAATC CATTCTACTC CTTTTCTTCC TTGGGACCAT
         N  L  S   L  C  Q   D  E  T  N   A  E  E   E  R  R  D
 51   CAACTTATCT CTCTGTCAGG ATGAGACAAA TGCCGAAGAA GAAAGAAGAG
           E  E  V   A  K  M   E  E  I   K  R  G  I   L  S  G
101   ATGAAGAAGT TGCTAAAATG GAAGAGATAA AACGCGGTAT ATTAAGTGGC
         I  L  G   A  G  K  S   L  V  C   G  L  S  G   L  C  *
151   ATCCTCGGTG CGGGGAAGAG CTTAGTATGT GGACTTAGCG GGCTGTGCTA
201   A

(C)     Nigrocin-2SCb
         M  F  T   L  K  K  S   L  L  L   L  F  F   L  G  T  I
  1   ATGTTCACCT TGAAGAAATC CCTGTTACTC CTTTTCTTCC TTGGGACCAT
         N  L  S   L  C  Q   D  E  T  N   A  E  E   E  R  R  D
 51   CAACTTATCT CTCTGCCAGG ATGAGACAAA TGCCGAAGAA GAAAGAAGAG
           E  E  V   A  K  M   E  E  I   K  R  G  I   L  S  G
101   ATGAAGAAGT TGCTAAAATG GAAGAGATAA AACGCGGTAT ATTAAGTGGC
         V  L  G   M  G  K  K   I  V  C   G  L  S   G  L  C  *
151   GTCCTCGGTA TGGGGAAGAA AATAGTATGT GGACTTAGCG GGCTGTGCTA
201   A

(D)     Nigrocin-2SCc
         M  F  T  T   K  K  S   L  L  L   L  F  F  L   G  T  I
  1   ATGTTCACCA CGAAGAAATC CCTGTTACTC CTTTTCTTCC TTGGGACCAT
         N  L  S   L  C  Q   D  E  T  N   A  E  E   E  R  R  D
 51   CAACTTATCT CTCTGCCAGG ATGAGACAAA TGCCGAAGAA GAAAGAAGAG
           E  E  V   A  K  M   E  E  I   K  R  G  I   L  S  N
101   ATGAAGAAGT TGCTAAAATG GAAGAGATAA AACGCGGTAT TTTAAGTAAC
         V  L  G   M  G  K  K   I  V  C   G  L  S   G  L  C  *
151   GTCCTCGGTA TGGGGAAGAA AATAGTATGT GGACTTAGCG GGCTGTGCTA
201   A
```

*Figure 1: Nucleotide sequences and superimposed translated open reading frames of cloned cDNAs encoding the biosynthetic precursors of (A) nigrocin-2N (*P. nigromaculatus*) and (B_D) nigrocin-2SCa-c (*O. schmackeri*). Mature peptides are underlined with a single line, putative signal peptides are underlined with a double line, and stop codons are indicated by asterisks.*

```
(A)    Nigrocin-2HJ
        M  F  T   L  K  K  S   L  L  L   L  F  F   L  G  M  V
  1   ATGTTCACCT TAAAGAAATC CCTGCTACTC CTTTTCTTTC TTGGGATGGT
        S  L  A   L  C  E   Q  E  R  D   A  N  E   E  E  R  R
 51   CTCCTTAGCT CTCTGTGAGC AAGAGAGAGA TGCCAATGAG GAAGAAAGAA
          D  E  L   D  E  R   D  V  E   A  I  K  R   G  L  L
101   GAGATGAGCT TGATGAAAGG GATGTTGAAG CGATAAAGCG AGGTCTATTA
        S  K  V   L  G  V  G   K  K  V   L  C  G  V   S  G  L
151   AGCAAAGTCC TCGGCGTCGG GAAGAAAGTA CTATGTGGAG TTAGTGGGCT
        C  *
201   TTGTTAA

(B)    Nigrocin-2LVa
        M  F  T   L  K  K  S   L  L  L   L  F  F   L  G  M  V
  1   ATGTTCACCT TGAAGAAATC CCTGCTACTC CTTTTCTTTC TTGGGATGGT
        S  L  A   L  C  E   Q  E  R  D   A  N  E   E  E  R  R
 51   CTCCTTAGCT CTCTGTGAGC AAGAGAGAGA TGCCAATGAG GAAGAAAGAA
          D  E  L   D  E  R   D  V  E   A  I  K  R   G  L  L
101   GAGATGAGCT TGATGAAAGG GATGTTGAAG CGATAAAGCG AGGTCTATTA
        S  K  V   L  G  V  G   K  K  V   L  C  G   V  S  G  L
151   AGCAAAGTCC TCGGCGTCGG GAAGAAAGTA CTATGTGGAG TTAGTGGGCT
        C  *
201   TTGTTAA

(C)    Nigrocin-2LVb
        M  F  T   L  K  K  S   L  L  L   L  F  F   L  G  T  I
  1   ATGTTCACCT TGAAGAAATC CCTGTTACTC CTTTTCTTTC TTGGGACCAT
        N  L  S   L  C  Q   E  E  R  N   A  E  E   E  R  R  D
 51   CAACTTATCT CTCTGTCAGG AAGAGAGAAA TGCCGAAGAA GAAAGAAGAG
          E  E  V   A  K  V   E  E  I   K  R  G  I   L  S  G
101   ATGAAGAAGT TGCTAAAGTG GAAGAGATAA AACGCGGTAT TTTAAGTGGC
        I  L  G   M  G  K  K   L  V  C   G  L  S   G  L  C  *
151   ATCCTCGGTA TGGGGAAGAA ACTAGTATGT GGACTTAGCG GGCTGTGCTA
201   A

(D)    Nigrocin-2VS
        M  F  T   L  K  K  S   M  L  L   L  F  F   L  G  T  I
  1   ATGTTCACCT TGAAGAAATC CATGTTACTC CTTTTCTTTC TTGGAACCAT
        S  L  S   L  C  E   Q  E  R  N   A  D  E   E  E  R  R
 51   CTCCTTATCT CTCTGTGAGC AAGAGAGAAA TGCTGATGAG GAAGAAAGAA
          D  E  E   V  A  K   V  E  E   I  K  R  S   I  L  S
101   GAGATGAAGA AGTTGCTAAA GTGGAAGAGA TAAAACGCAG TATTTTAAGT
        G  N  F   G  V  G  K   K  I  V   C  G  L   S  G  L  C
151   GGCAATTTCG GTGTGGGGAA GAAAATAGTA TGTGGACTTA GCGGGCTGTG
        *
201   CTAA
```

*Figure 2: Nucleotide sequences and superimposed translated open reading frames of cloned cDNAs encoding the biosynthetic precursors of (A) nigrocin-2HJ (*O. hejiangensis*), (B,C) nigrocin-2LVa and LVb (*O. livida*), and (D) nigrocin-2 VS (*O. versabilis*). Mature peptides are underlined with a single line, putative signal peptides are underlined with a double line, and stop codons are indicated by asterisks.*

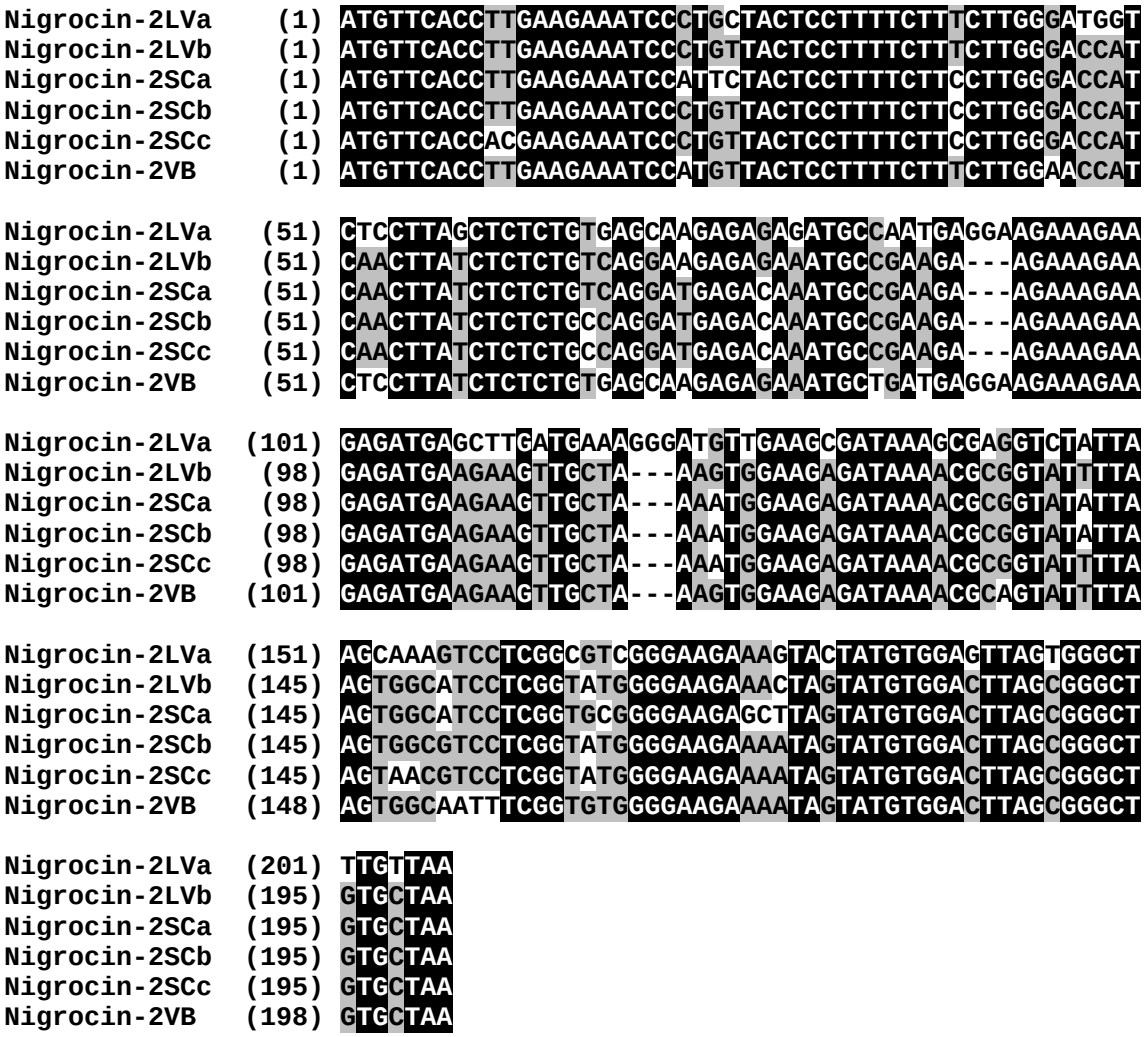


Fig. 3: Aligned open reading frame nucleotide sequences of clones encoding the biosynthetic precursors of nigrocin-2-related peptides from the skins of selected Odorrana *frogs used in this study (LV,* O. livida*; SC,* O. schmackeri*; VB,* O. versabilis*). Note the particularly high degree of nucleotide sequence conservation at both 5' and 3' ends. Conserved nucleotides are shaded in black, and consensus nucleotides are shaded in grey.*

```
                ←----------1----------→ ←----------2----------→  3 ←----------4----------→
Nigrocin-2N     MFTTKKSILLLFFLGMVSLALC  EQERDANEEE RR DELDERDVEAI KR GLLSKVLGVGKKVLCGVSGLC
Nigrocin-2SCa   MFTLKKSILLLFFLGTINLSLC  QDETNAEEE- RR DEEVAKMEEI- KR GILSGILGAGKSLVCGLSGLC
Nigrocin-2SCb   MFTLKKSLLLLFFLGTINLSLC  QDETNAEEE- RR DEEVAKMEEI- KR GILSGVLGMGKKIVCGLSGLC
Nigrocin-2SCc   MFTTKKSLLLLFFLGTINLSLC  QDETNAEEE- RR DEEVAKMEEI- KR GILSNVLGMGKKIVCGLSGLC
Nigrocin-2LVa   MFTLKKSLLLLFFLGMVSLALC  EQERDANEEE RR DELDERDVEAI KR GLLSKVLGVGKKVLCGVSGLC
Nigrocin-2LVb   MFTLKKSLLLLFFLGTINLSLC  QEERNAEEE- RR DEEVAKVEEI- KR GILSGILGMGKKLVCGLSGLC
Nigrocin-2HJ    MFTLKKSLLLLFFLGMVSLALC  EQERDANEEE RR DELDERDVEAI KR GLLSKVLGVGKKVLCGVSGLC
Nigrocin-2VS    MFTLKKSMLILFFLGTISLSLC  EQERNADEEE RR DEEVAKVEEI- KR SILSGNFGVGKKIVCGLSGLC
                *** *** * *****   * **    *  * **  ** **      *   **   **   * **   ** ****
```

Fig. 4: Domain architecture of nigrocin-2 biosynthetic precursors reported in this study. 1, putative signal peptide; 2, proximal acidic residue rich spacer peptide; 3, putative dibasic residue propeptide convertase processing site; 4, mature active antimicrobial peptide encoding domain (underlined). Disulfide-bridged Rana *boxes at the C-termini of mature nigrocins are italicized. Conserved amino acid residue sites are indicated by asterisks.*

```
(A)
Nigrocin-2SCa      GILSGILGAGKSLVCGLSGLC
Nigrocin-2SCb      GILSGVLGMGKKIVCGLSGLC
Nigrocin-2SCc      GILSNVLGMGKKIVCGLSGLC
                   ****  ** **  *******

(B)
Nigrocin-2GRa      GLLSGILGAGKHIVCGLSGLC
Grahamin-1         GLLSGILGAGKHIVCGLSGLC
Nigrocin-OG1       GLLSGILGAGKHVVCGLSGLC
Nigrocin-OG2       GLLGKILGVEKKVLCGLSGMC
Nigrocin-OG3       GLLSGILGVGKHIVCGLSGLC
Nigrocin-OG4       GLLSGILGAGKHIICGLSGLC
Nigrocin-OG6       GLLSGILGTGKHIVCGLSGLC
Nigrocin-OG8       GLLSGILGAGKHIVCGLSRLC
Nigrocin-OG13      GLLSGILSAGKHIVCGLSGLC
Nigrocin-OG15      GLLRGILGAGKHIVCGLSGLC
Nigrocin-OG18      GLLSGILGAGKNIVCGLSGPC
Nigrocin-OG19      GLLSGILGAGKHTVCGLSGLC
Nigrocin-OG20(1)   GLLSGILGAGKHIVCGLSGLC
Nigrocin-OG21      GLLSGVLGVGKKVLCGLSGLC
Nigrocin-OG22      GLLSKILGVGKKVLCGLSGMC
                   ***   **  *   ****  *

(C)
Nigrocin-2GRc      GLLSGILGAGKNIVCGLSGLC  (Conlon et al., 2006)
Grahamin-2         GLLSGILGAGKNIVCGLSGLC  (Xu et al., 2006)
Nigrosin-OG17      GLLSGILGAGKNIVCGLSGLC  (Li et al., 2007)
Nigrosin-OG20(2)   GLLSGILGAGKNIVCGLSGLC  (Li et al., 2007)
                   *********************

(D)
Nigrocin-2GRa      GLLSGILGAGKHIVCGLSGLC  (Conlon et al., 2006)
Grahamin-1         GLLSGILGAGKHIVCGLSGLC  (Xu et al., 2006)
Nigrocin-OG20(1)   GLLSGILGAGKHIVCGLSGLC  (Li et al., 2007)
                   *********************
```

Fig. 5: Alignment of amino acid residues in (A) nigrocin-2-related peptides from O. schmackeri *skin, (B) nigrocin-2-related peptides from* O. grahami *skin, (C and D) identical nigrocin-2-related peptides from* O. grahami *skin with different names. Identical amino acid residues indicated by asterisks.*

3.3.2 Isolation and structural characterization of nigrocins from reverse-phase HPLC fractions of skin secretions

We previously reported our standard approach to identify and structurally characterize frog venom peptides[27, 28]. This involves HPLC fractionation, MALDI TOF MS monitoring of the separation of the peptides, and automated Edman sequencing of the purified material (i.e. Workflow 1- see Materials and Methods). This approach was followed for all novel nigrocin-2-related peptides described in this study. Mature nigrocin-2-related peptide primary structures were predicted from open reading frames of the cloned skin secretion-derived cDNA from each species. The calculated molecular masses of the predicted peptides were then used to interrogate MALDI-TOF MS data for reverse phase HPLC fractions of the same species. Eight peptides in total, within the molecular mass range of 1.9 – 2.2 kDa and with associated antimicrobial activities, were resolved within the reverse phase HPLC fractions of the species studied. One peptide with a molecular mass of 2027.17 Da, very similar to that of the prototype nigrocin-2[19], was located in fractions of *P. nigromaculatus* skin secretion and was named nigrocin-2N to reflect assignment to *nigromaculatus*. Two peptides (2027.15 and 2016.95 Da) were found in *O. livida* skin secretion fractions, three peptides (1915.25, 2002.12 and 2060.15 Da) were found in *O. schmackeri* fractions and single peptides of 2050.10 Da and 2027.20 Da in *O. versabilis* and *O. hejiangensis* fractions, respectively. Co-incident molecular masses of peptides, implying identity, and discrepant molecular masses, implying lack of identity, were observations confirmed following automated amino acid sequencing (Table 1). A sample reverse phase HPLC chromatogram from fractionation of *O. schmackeri* skin secretion, generated using Workflow 1, is shown in Fig. 6. The elution positions (retention times) of respective nigrocin-2 peptides in this species are indicated, with comparable abundance of two peptides and the significantly lower abundance of a third.

Table 1: Primary structures and molecular masses of nigrocin-2 related peptides identified in this study from Odorrana *frog skin secretions. The disulfide bridged domain between Cys[15] and Cys[21] is underlined.*

Peptide	Original fraction	Mass obs. (Da)	Mass calc. (Da)	Amino acid sequence
Nigrocin 2LVa	**178**	**2027.15**	**2027.54**	**GLLSKVLGVGKKVLCGVSGLC**
Nigrocin 2LVb	**186**	**2016.95**	**2016.53**	**GILSGILGMGKKLVCGLSGLC**
Nigrocin 2SCa	**174**	**1915.25**	**1915.32**	**GILSGILGAGKSLVCGLSGLC**
Nigrocin 2SCb	**176**	**2002.12**	**2002.50**	**GILSGVLGMGKKIVCGLSGLC**
Nigrocin 2SCc	**187**	**2060.15**	**2059.55**	**GILSNVLGMGKKIVCGLSGLC**
Nigrocin 2VB	**180**	**2050.10**	**2049.46**	**SILSGNFGVGKKIVCGLSGLC**

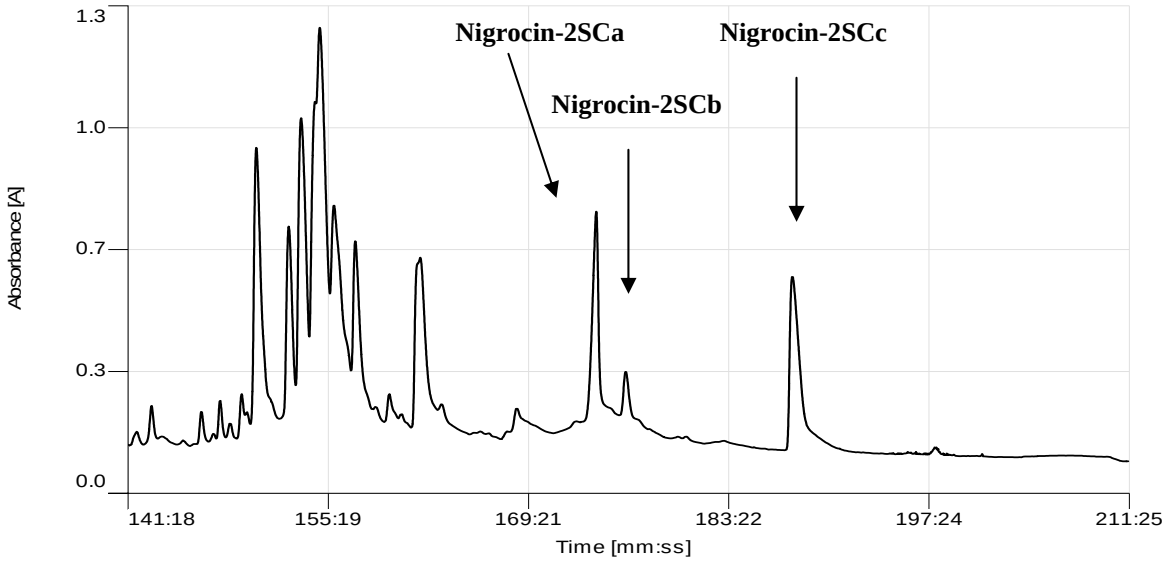


Figure 6: Expanded region of reverse-phase HPLC chromatogram by workflow 1 of skin secretion from O. schmackeri *indicating absorbance peaks corresponding in molecular mass to nigrocin-2-related peptide masses deduced from the respective cloned biosynthetic precursors.*

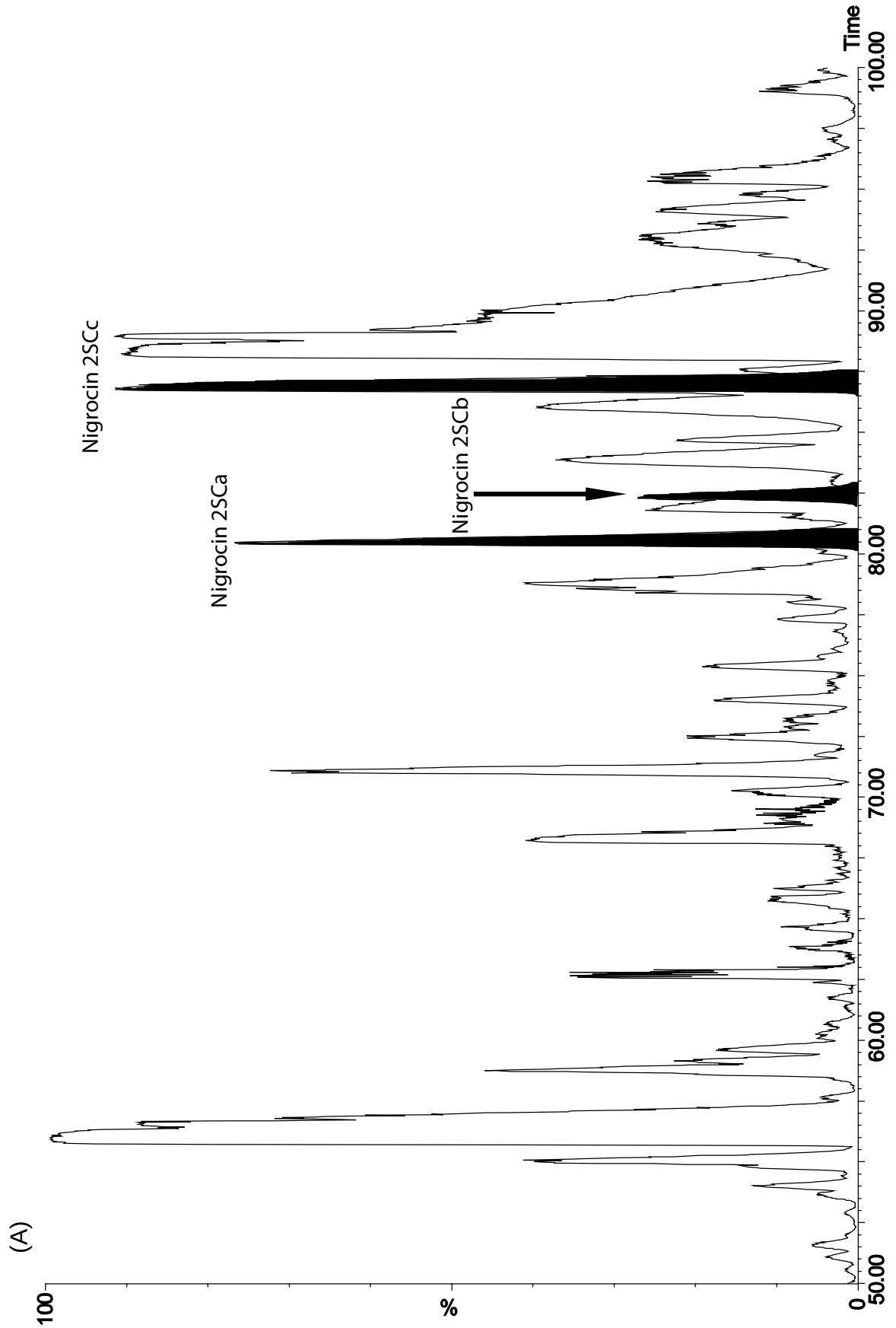
Interrogation of contemporary protein/peptide databases within the National Center for Biotechnology Information using FASTA and BLAST algorithms established that the primary structures of one of the peptides from *O. livida* (nigrocin-2LVa) and of nigrocin-2HJ from *O. hejiangensis*, were identical to that of the original nigrocin-2N peptide from the skin of the Oriental ranid frog, *Pelophylax nigromaculatus,* and that the remaining five peptides were novel nigrocin-2 variants. These were named in accordance with previously established nomenclature as nigrocin-2LVb (LV, *livida*), nigrocin-2SCa, b and c (SC, *schmackeri*) and nigrocin-2VB (V, *versabilis*). The estimated quantities of nigrocins recovered from each species following transdermal electrical stimulation ranged between 10-14 nmol/mg of lyophilized skin secretions as determined by amino acid analysis.

As today's proteomics/peptidomics technology allows peptide sequence information to be obtained by tandem mass spectrometry (so-called MS/MS) directly from peptide mixtures, without prior purification, we decided to evaluate a second fully LC-MS/MS-based approach (*workflow 2*, see Experimental procedures), i.e. without Edman degradation. To demonstrate the efficacy of this LC-MS/MS approach, we selected the *O. schmackeri* skin secretion sample.

The molecular cloning data indicated that it contained three different nigrocin-2-related peptides.

Using *workflow 2*, all three predicted *O. schmackeri* nigrocin-2-related peptides were located in reverse phase nanoUPLC fractions after calculation of the predicted molecular masses from their cloned precursor- encoding cDNAs (Fig. 7A). The primary structure of each peptide was unambiguously confirmed by MS/MS fragmentation sequencing (Figs 8–10). Interestingly, almost complete b- and y-ion series were obtained, which facilitated confirmation of primary structures obtained through Edman microsequencing and predicted from cloned precursor cDNAs. It is noteworthy that, in addition to the qualitative agreement between the data from workflows 1 and 2, the relative abundances of the three nigrocins observed using the *workflow 2* approach are similar to those observed using *workflow 1* (Fig. 6). This finding is of particular interest as the quantity of starting material used in *workflow 2* is three orders of magnitude lower than that required for *workflow 1* – a factor of some significance when working with rare and precious natural materials from threatened species.

Generating 2D images from the LC run (Fig. 7B) nicely illustrates the complexity of the samples. In addition to the peptide ions representing the three nigrocins identified in this study, other peptides previously identified in *O. schmackeri* were also observed and their sequences confirmed by MS/MS. These include peptides previously described as brevinin 1HS, esculentin 1S, esculentin 2S, odorranin C7HSa and odorranin PB.


(A)
100
%
0
50.00
60.00
70.00
80.00
90.00
100.00
Time
Nigrocin 2SCa
Nigrocin 2SCb
Nigrocin 2SCc

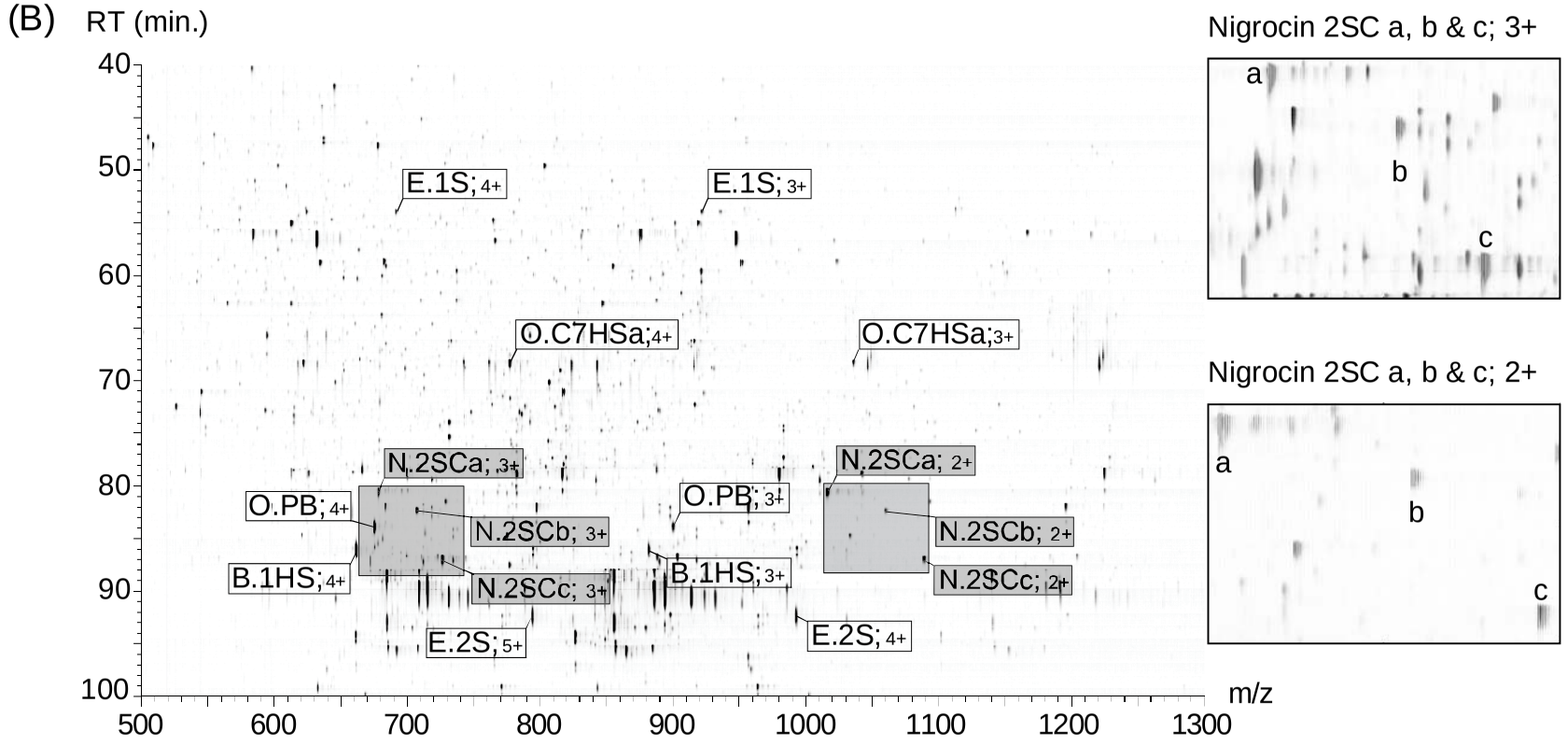


Figure 7: (A) Base peak intensity chromatogram of reverse-phase (C4) nanoUPLC-QTOF MS analysis of 2 µg O. schmackeri *skin secretion obtained by workflow 2, after reduction and alkylation. Peaks coincident in molecular mass with nigrocin-2SCa–c are indicated by black shading. (B) Two-dimensional map (so-called 'peptide display') of this nanoUPLC-QTOF MS analysis with retention time (y axis) plotted against mass-over-charge ratio (x axis). The image was produced using MSight version 1.0.1 (Swiss Institute of Bioinformatics, Switzerland). The positions of known* O. schmackeri *skin peptides and novel nigrocin-2-related peptide ions identified in this study are indicated. Peptide display areas containing nigrocin-2SCa–c (N.2SCa, N.2SCb and N.2SCc, respectively) are highlighted in grey. The zoomed inserts show doubly and triply protonated peptides ($[M + 2H+]^{2+}$ and $[M + 3H+]^{3+}$, respectively). Abbreviations: B.1HS, brevinin 1HS; E.1S, esculentin 1S; E.2S, esculentin 2S; O.C7HSa, odorranin C7HSa; O.PB, odorranin PB. Note that many peptide signatures remain unidentified. 2+, 3+, 4+ and 5+ indicate peptides that are protonated twice, three, four and five times ($[M+2H^+]^{2+}$, $[M+3H^+]^{3+}$, $[M+4H^+]^{4+}$ and $[M+5H^+]^{5+}$, respectively).*

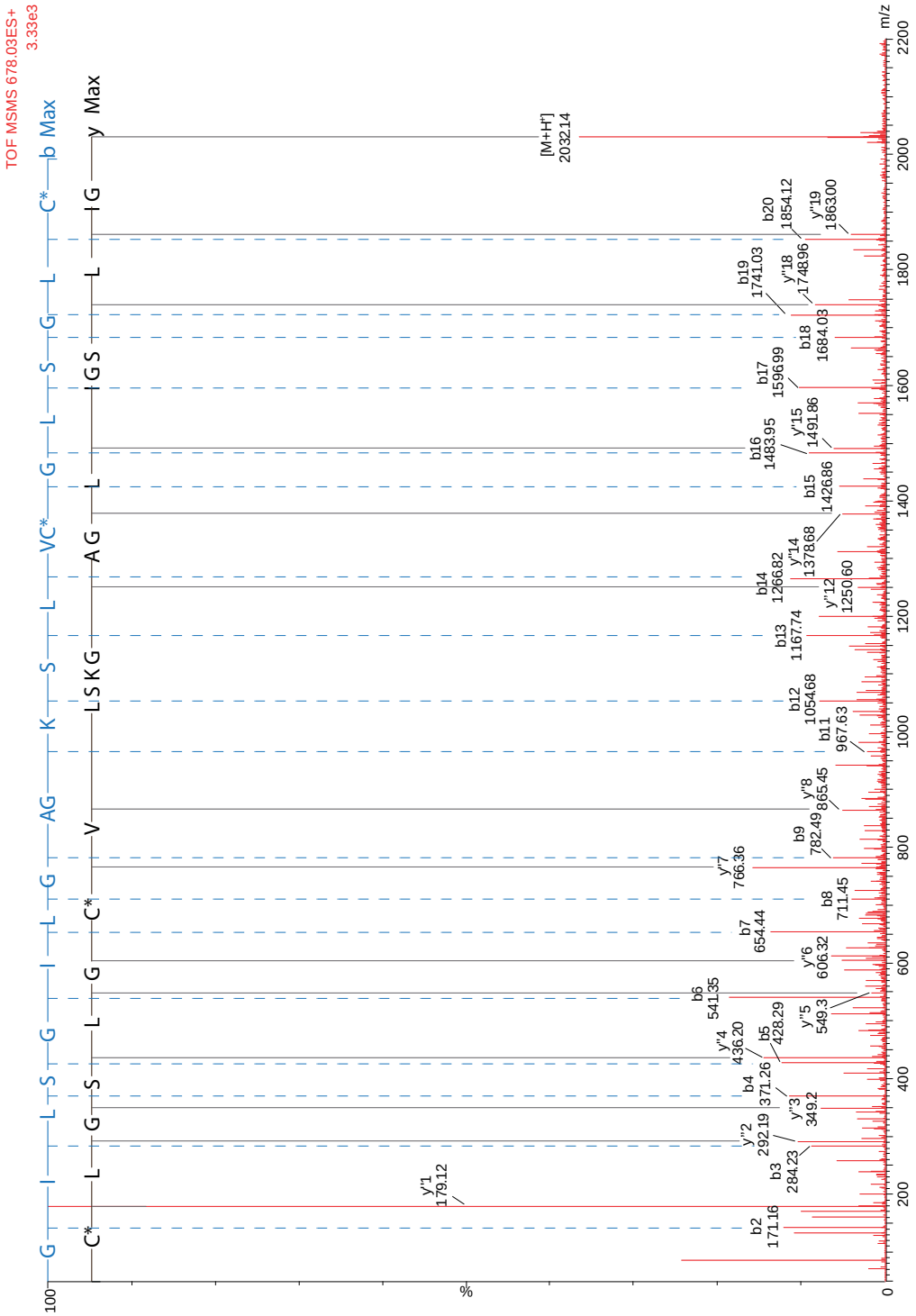


Figure 8: Deconvoluted QTOF MS/MS low-energy CID spectra of reduced and alkylated Nigrocin 2SCa triply charged peptide ($[M+3H]^{3+}$ at m/z 678.03). The b-ion and y-ion series are labeled. C represents carbamidomethyl cysteine residue. Note that isobaric I/L residue assignments are not possible from MSMS data, and these assignments were made on the basis of Edman sequencing and cloned biosynthetic precursor deduced primary structures.*

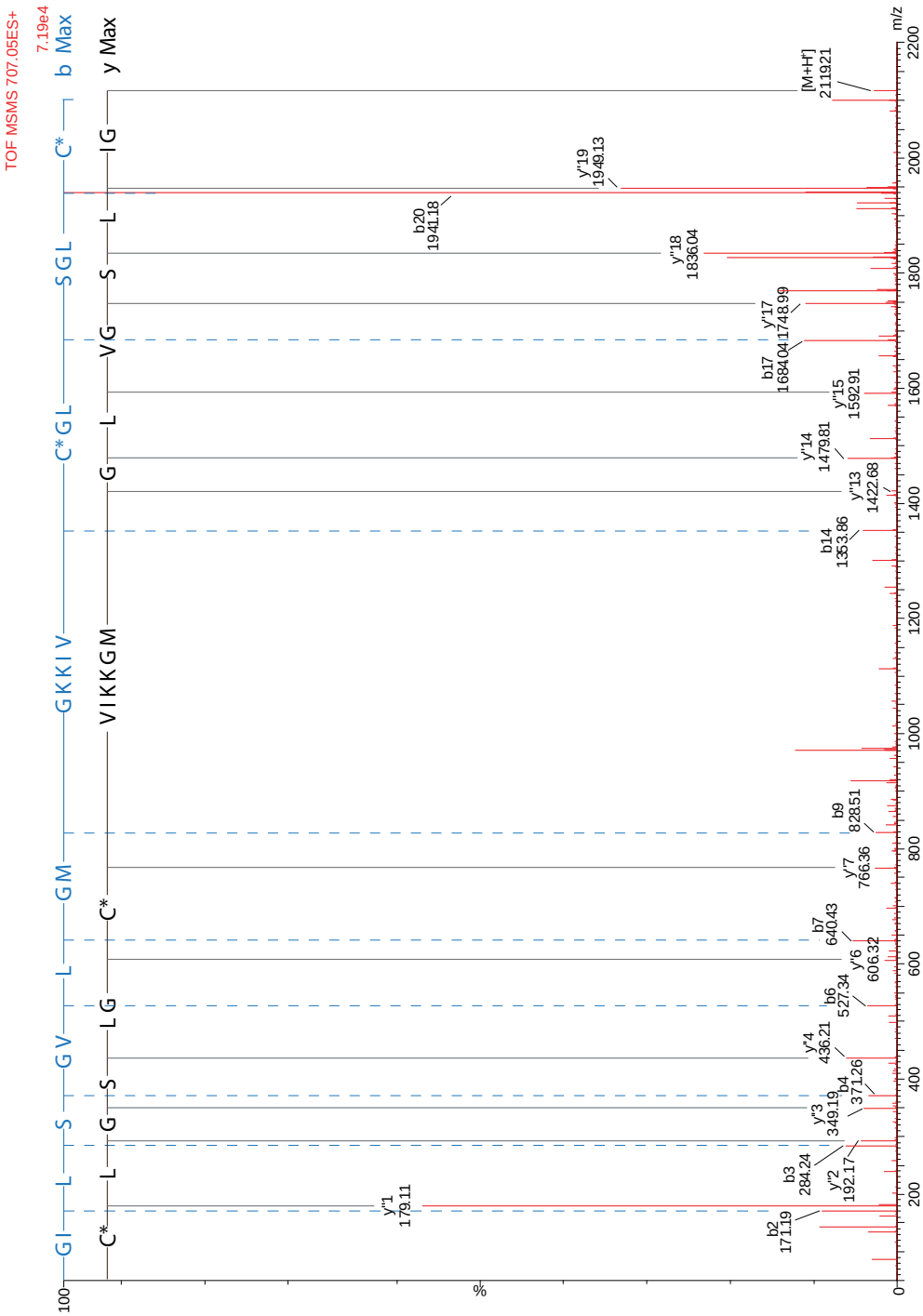


Figure 9: Deconvoluted QTOF MS/MS low-energy CID spectra of reduced and alkylated Nigrocin 2SCb triply charged peptide ($[M+3H]^{3+}$ at m/z 707.05). The b-ion and y-ion series are labeled. C represents carbamidomethyl cysteine residue. Note that isobaric I/L residue assignments are not possible from MS/MS data, and these assignments were made on the basis of Edman sequencing and cloned biosynthetic precursor deduced primary structures.*

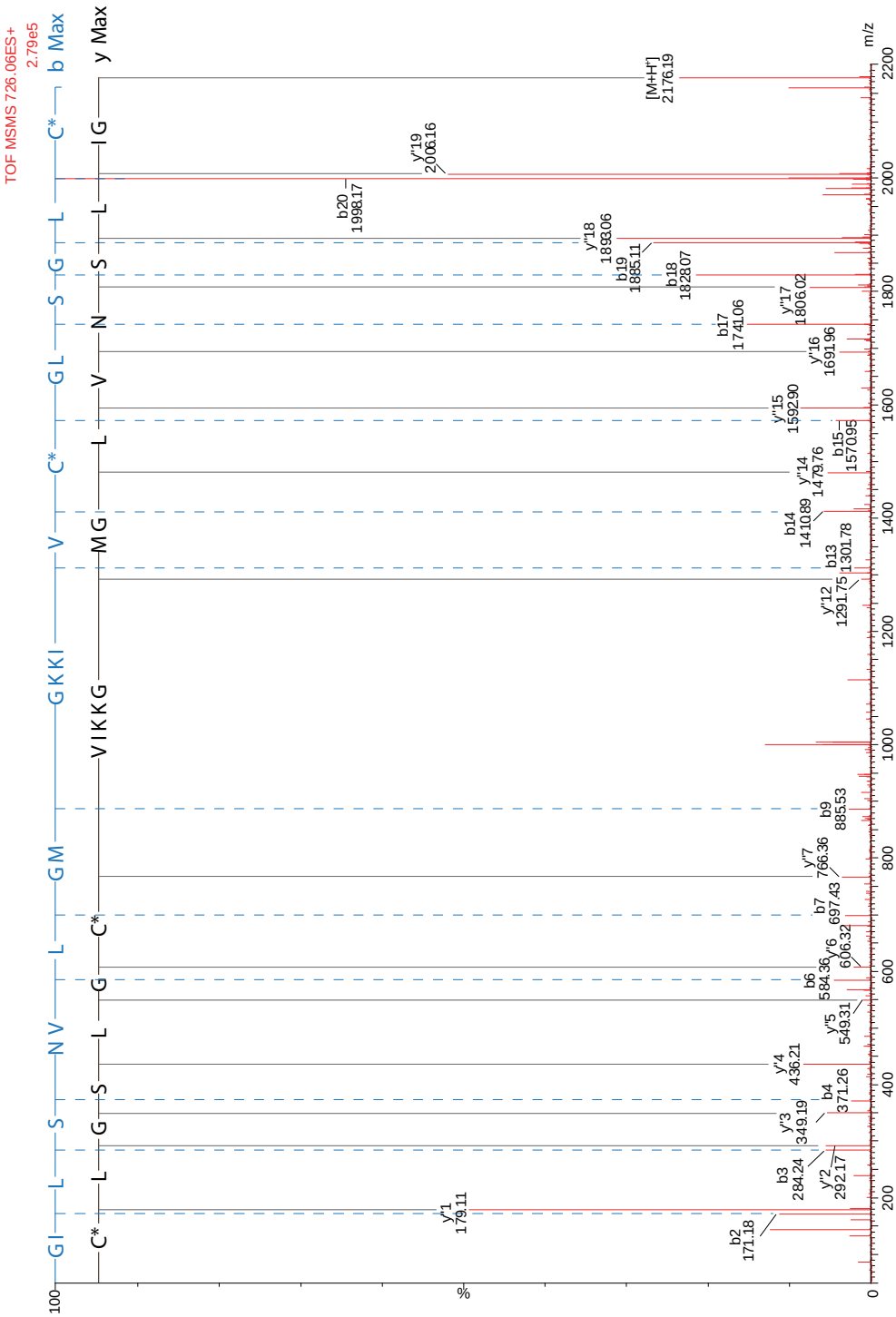


Figure 10: Deconvoluted QTOF MS/MS low-energy CID spectra of reduced and alkylated Nigrocin 2SCc triply charged peptide ($[M+3H]^{3+}$ at m/z 706.26). The b-ion and y-ion series are labeled. C represents carbamidomethyl cysteine residue. Note that isobaric I/L residue assignments are not possible from MS/MS data, and these assignments were made on the basis of Edman sequencing and cloned biosynthetic precursor deduced primary structures.*

3.3.3 Antimicrobial activity of nigrocin-related peptides

The minimum inhibitory concentrations obtained for each peptide against *Staphylococcus aureus* and *Escherichia coli* are summarized in Table 2. Like several previously characterized nigrocin-2 peptides[19, 22-24], including the archetypal peptide from *P. nigromaculatus*[19], synthetic replicates of the nigrocin-2-related peptides reported in the present study were not as potent as members of other classes of ranid frog skin antimicrobial peptide, such as brevinins or temporins[26]. However, unusually, each appeared to be more potent against the model Gram-negative bacterium, E. *coli*, than against the model Gram-positive bacterium, *S. aureus*. There also appeared to be a relationship between net positive charge and efficacy in this respect.

*Table 2: Minimal inhibitory concentrations (μM) of novel nigrocin-2-related peptides against a Gram-positive bacterium (*S. aureus*) and a Gram-negative bacterium* (E. coli*). Nigrocin 2LVa and 2LVb are from* Odorrana livida*; nigrocin 2SCa, 2SCb and 2SCc are from* Odorrana schmackeri*; nigrocin 2VB is from* Odorrana versabilis*). Note from text that nigrocin-2LVa is of identical primary structure to nigrocin-2N and nigrocin-2HJ.*

Peptide	(charge)	*S. aureus*	*E. coli*
Nigrocin-2LVa	**(+3)**	**15.0**	**5.0**
Nigrocin-2LVb	**(+2)**	**50.0**	**15.0**
Nigrocin-2SCa	**(+1)**	**>100**	**50.0**
Nigrocin-2SCb	**(+2)**	**50.0**	**25.0**
Nigrocin-2SCc	**(+2)**	**>100**	**70.0**
Nigrocin-2VB	**(+2)**	**>100**	**25.0**

3.4 Discussion

While antimicrobial peptides are of widespread occurrence within the defensive skin secretions of anuran amphibians, the taxon that has the greatest diversity of structurally defined classes of this type of peptide is undoubtedly frogs of the family Ranidae[11, 26]. This amphibian family has many members and a wide global distribution[8, 9, 11, 26]. The hotspot for ranid frog diversity is undeniably within Asia, and several research groups have found that the primary structures of skin antimicrobial peptides can be a useful taxonomic adjunct when used with an appropriate measure of caution[26]. One of the fundamental prerequisites for such use is that the peptides themselves have a standardized and rational nomenclature scheme that

is widely- if not universally adopted by researchers in the field. Until now such a scheme has not been available. However, one has recently been proposed that is both logical and systematic for ranid frog skin antimicrobial peptides[26]. The peptides identified in the present study are unequivocally members of the nigrocin-2 peptide family whose prototype was originally described from the skin of the Oriental black-spotted pond frog, *Rana nigromaculata*[19] (now *Pelophylax nigromaculatus*)[29]. They have been named in accordance with the new scheme as nigrocin-2LVa and b (*O. livida*), nigrocin-2SCa, b and c (*O. schmackeri*), nigrocin-2VB (*O. versabilis)* and nigrocin-2HJ (*O. hejiangensis*). In accordance, the archetypal nigrocin-2 from *Pelophylax nigromaculatus* was renamed nigrocin-2N.

In common with the nigrocin-2-related peptides isolated from the skin of *Odorrana grahami*[24], these *Odorrana* nigrocin-2 homologues exhibited a relatively low activity against Gram-positive and Gram-negative bacteria. Secondary structural predictions indicated a lack of helicity in these peptides when modeled in aqueous environments (data not shown) – a finding that is in accordance with CD studies performed on the prototype[19]. However, in membrane mimetic environments, the peptides become helical to a high degree, as indicated by CD studies on the prototype[19]. The mode of action with regard to inhibition of bacterial growth by these peptides may thus differ markedly from that of other established classes of skin antimicrobial peptides, and, in fact, the actual biological target(s) of nigrocin-2-related peptides may not be prokaryotic at all. In other words, the antibacterial activity displayed may be a consequence of structure rather than biological design. While nigrocin-2 peptides have a *Rana* box at the C-terminus[26], this is most unusual in its lack of basic amino acid residues and highly hydrophobic character. The cationicity of the *Rana* box motif is thought to play a fundamental role in the initial interaction of these peptides with the anionic glycocalyx of bacterial cells, and this motif alone has a potent effect on mast cell degranulation[26]. This unusual structural feature of nigrocin-2 peptides is reminiscent of the glycine / leucine-rich dermaseptin orthologues from the skins of neotropical phyllomedusine leaf frogs, named plasticins[30]. Nigrocin-2 peptides could be regarded as glycine/leucine-rich brevinin orthologues, as they share this structural feature with the plasticins (plasticin PBN2KF, 52.1% Gly/Le; nigrocin-2N, 47.6% Gly/Leu). The skin secretions of amphibian taxa that contain antimicrobial activity thus contain a broad range of peptides varying in the numbers of amino acid residues, net charge and hydrophobic characteristics, as well as a range of isomeric forms from each class. This can effect complex molecular interactions both between discrete peptides themselves and with molecular targets, maximizing the overall antimicrobial efficacy of the secretion – a factor that is often overlooked in the biochemical reductionist approach of

studying single molecular entities. The study of complex amphibian skin secretion peptidomes has therefore an enormous potential to address and potentially solve many complex technological problems arising from a holistic integration of modern analytical tools in high-precision molecular characterization. Moreover, amphibians are in a global decline[31], and such studies can obtain structural and functional data from unique natural peptide libraries whose donors may be approaching the verge of extinction, and thus may provide new therapeutic leads.

The large number of biologically active peptides, often of unique primary structure, within the peptidomes of amphibian defensive skin secretions (usually available in limited supply like their donors), has presented the biological chemist with the problem of enhancing the rate of discovery of novel chemical entities through primary structural characterization in the absence of substantive and relevant online structural databases. This was the compelling reason why we developed and evaluated a novel analysis scheme, based upon online UPLC MS/MS, that had the potential to rapidly and effectively identify and structurally characterize new peptidic components of complex and relatively unstudied natural peptidomes. We have used the nigrocin-2-related peptides, a subset of antimicrobial peptides from the skin secretions of the Oriental frog, *O. schmackeri*, to illustrate this. The resultant data show that, in combination with molecular cloning technology, UPLC MS/MS allows unambiguous detection and sequence confirmation and/or characterization of bioactive peptides from several thousand-fold lower quantities of source material than required for HPLC/Edman degradation analytics. In addition, a peptide display of the LC-MS data (Fig. 7B) clearly shows that the majority of the skin peptides in this species await structural/functional elucidation, and serves to direct and focus the attention of the analyst to individual molecules within this category.

Acknowledgements

Lei Wang is in receipt of an Overseas Studentship at Queen's University, Belfast. We gratefully acknowledge the excellent technical assistance of graduate students Joel Fulton, Rory O'Donnell and Ryan J. Graham. This study was partly funded by the Netherlands Genomics Initiative (NGI) and the Brazilian National Counsel of Technological and Scientific Development (CNPq – grant GDE-200847/2007-04).

References

1. Qoronfleh, M. W., Role and challenges of proteomics in pharma and biotech: technical, scientific and commercial perspective. *Expert Rev Proteomics* **2006,** 3, (2), 179-95.
2. Ashburner, M.; Goodman, N., Informatics--genome and genetic databases. *Curr Opin Genet Dev* **1997,** 7, (6), 750-6.
3. Nagele, E.; Vollmer, M.; Horth, P.; Vad, C., 2D-LC/MS techniques for the identification of proteins in highly complex mixtures. *Expert Rev Proteomics* **2004,** 1, (1), 37-46.
4. Tangrea, M. A.; Wallis, B. S.; Gillespie, J. W.; Gannot, G.; Emmert-Buck, M. R.; Chuaqui, R. F., Novel proteomic approaches for tissue analysis. *Expert Rev Proteomics* **2004,** 1, (2), 185-92.
5. Whitfield, E. J.; Pruess, M.; Apweiler, R., Bioinformatics database infrastructure for biotechnology research. *J Biotechnol* **2006,** 124, (4), 629-39.
6. Vollmer, M.; Nagele, E.; Horth, P., Differential proteome analysis: two-dimensional nano-LC/MS of E. coli proteome grown on different carbon sources. *J Biomol Tech* **2003,** 14, (2), 128-35.
7. Ivanov, V. T.; Yatskin, O. N., Peptidomics: a logical sequel to proteomics. *Expert Rev Proteomics* **2005,** 2, (4), 463-73.
8. Bevins, C. L.; Zasloff, M., Peptides from frog skin. *Annu Rev Biochem* **1990,** 59, 395-414.
9. Erspamer, V., Bioactive secretions of the integument. In *Amphibian Biology Volume 1: the Integument*, integument., B. s. o. t., Ed. Surrey Beatty & Sons (Heatwole H & Barthalmus GT eds): Chipping Norton, Australia., 1994; Vol. 1, pp 179-350.
10. Clarke, B. T., The natural history of amphibian skin secretions, their normal functioning and potential medical applications. *Biol Rev Camb Philos Soc* **1997,** 72, (3), 365-79.
11. Conlon, J. M., The therapeutic potential of antimicrobial peptides from frog skin. *Rev Med Microbiol* **2004**, (15), 1-9.
12. Nicolas, P.; Mor, A., Peptides as weapons against microorganisms in the chemical defense system of vertebrates. *Annu Rev Microbiol* **1995,** 49, 277-304.
13. Rybak, M. J., Resistance to antimicrobial agents: an update. *Pharmacotherapy* **2004,** 24, (12 Pt 2), 203S-15S.
14. Levy, S. B.; Marshall, B., Antibacterial resistance worldwide: causes, challenges and responses. *Nat Med* **2004**, (10), 122–129.
15. Boman, H. G., Peptide antibiotics and their role in innate immunity. *Annu Rev Immunol* **1995,** 13, 61-92.
16. Barra, D.; Simmaco, M., Amphibian skin: a promising resource for antimicrobial peptides. *Trends Biotechnol* **1995,** 13, (6), 205-9.
17. Simmaco, M.; Mignogna, G.; Barra, D., Antimicrobial peptides from amphibian skin: what do they tell us? *Biopolymers* **1998**, (47), 435–450.
18. Lai, R.; Liu, H.; Hui Lee, W.; Zhang, Y., An anionic antimicrobial peptide from toad Bombina maxima. *Biochem Biophys Res Commun* **2002,** 295, (4), 796-9.
19. Park, S.; Park, S. H.; Ahn, H. C.; Kim, S.; Kim, S. S.; Lee, B. J.; Lee, B. J., Structural study of novel antimicrobial peptides, nigrocins, isolated from Rana nigromaculata. *FEBS Lett* **2001,** 507, (1), 95-100.
20. Conlon, J. M., Reflections on a systematic nomenclature for antimicrobial peptides from the skins of frogs of the family Ranidae. *Peptides* **2008,** 29, (10), 1815-9.
21. Dubois, A., Naming taxa from cladograms: a cautionary tale. *Mol Phylogenet Evol* **2007,** 42, (2), 317-30.
22. Conlon, J. M.; Al-Ghaferi, N.; Abraham, B.; Jiansheng, H.; Cosette, P.; Leprince, J.; Jouenne, T.; Vaudry, H., Antimicrobial peptides from diverse families isolated from the skin of the Asian frog, Rana grahami. *Peptides* **2006,** 27, (9), 2111-7.

23. Xu, X.; Li, J.; Han, Y.; Yang, H.; Liang, J.; Lu, Q.; Lai, R., Two antimicrobial peptides from skin secretions of Rana grahami. *Toxicon* **2006,** 47, (4), 459-64.
24. Li, J.; Xu, X.; Xu, C.; Zhou, W.; Zhang, K.; Yu, H.; Zhang, Y.; Zheng, Y.; Rees, H. H.; Lai, R.; Yang, D.; Wu, J., Anti-infection peptidomics of amphibian skin. *Mol Cell Proteomics* **2007,** 6, (5), 882-94.
25. Tyler, M. J.; Stone, D. J.; Bowie, J. H., A novel method for the release and collection of dermal, glandular secretions from the skin of frogs. *J Pharmacol Toxicol Methods* **1992,** 28, (4), 199-200.
26. Conlon, J. M.; Kolodziejek, J.; Nowotny, N., Antimicrobial peptides from ranid frogs: taxonomic and phylogenetic markers and a potential source of new therapeutic agents. *Biochim Biophys Acta* **2004,** 1696, (1), 1-14.
27. Wang, L.; Zhou, M.; Zhou, Z.; Chen, T.; Walker, B.; Shaw, C., Sauvatide--a novel amidated myotropic decapeptide from the skin secretion of the waxy monkey frog, Phyllomedusa sauvagei. *Biochem Biophys Res Commun* **2009,** 383, (2), 240-4.
28. Wang, L.; Zhou, M.; Chen, T.; Walker, B.; Shaw, C., PdT-2: a novel myotropic type-2 tryptophyllin from the skin secretion of the Mexican giant leaf frog, Pachymedusa dacnicolor. *Peptides* **2009,** 30, (8), 1557-61.
29. Fei, L.; Ye, C. Y.; Jiang, J. P.; Xie F.; Huang, Y. Z., *An Illustrated Key to Chinese Amphibians.* Sichuan Publishing Group/Sichuan Publishing House of Science and Technology.: Chengdu, China., 2005.
30. Conlon, J. M.; Abdel-Wahab, Y. H.; Flatt, P. R.; Leprince, J.; Vaudry, H.; Jouenne, T.; Condamine, E., A glycine-leucine-rich peptide structurally related to the plasticins from skin secretions of the frog Leptodactylus laticeps (Leptodactylidae). *Peptides* **2009,** 30, (5), 888-92.
31. Houlahan, J. E.; Findlay, C. S.; Schmidt, B. R.; Meyer, A. H.; Kuzmin, S. L., Quantitative evidence for global amphibian population declines. *Nature* **2000,** 404, (6779), 752-5.

Chapter 4

Discovery of O-glycosylated peptides secreted by the skin of the Chinese frog *Odorrana schmackeri*

Abstract

The amphibian skin secretion is a rich source of peptides with a wide variety in bioactivity, and many of them are post-translationally modified. Because of this, MS-based analysis of these peptides requires prior treatment of the venom for full primary structure elucidation. *Odorrana schmackeri* belongs to the Ranidae family which is known to secrete many peptides containing intramolecular disulfide bonds. This is reflected in the observation that LC-MS(/MS) analyses of crude venom peptides yield a majority of CID spectra which cannot be solved due to poor peptide backbone fragmentation. Fortuitously, during this analysis our attention was drawn to 3 tandem MS spectra that did show rich fragment ions information. These spectra were found to originate from 3 peptides (with no S-S linked) with similar chromatographic retention and molecular masses. During *de novo* sequencing it was observed that these peptides exhibited a-typical peptide fragmentation and we discovered that these amphibian skin peptides are in fact glycosylated peptides, with either 2, 3 and 4 monosaccharide units each. Further CID/ETD experiments enabled us to obtain the full primary structure, including the identification of location of the (O-linked) glycosylation site at a threonine residue. The primary sequence of this peptide was determined to be AVP(I/L)(I/L)YNRPG(I/L)YVTKRPK-amide. This shows high similarity with other Ranid peptides, including odorranainin-O-RA and lividin-8, both predicted to be antimicrobial peptides and discovered from molecular cloning experiments. Our findings in *O. schmackeri* suggest that these peptides are in fact post-translationally modified by O-glycosylation. This is the first example of an O-glycosylated peptide to be present in a frog skin secretion.

Manuscript in preparation

Evaristo, G. P. C.; Pinkse, M. W. H.; Shaw, C.; Verhaert, P. D. E. M., Discovery of O-Glycosylated Peptides Secreted by the Skin of the Chinese Frog *Odorrana schmackeri.*

4.1 Introduction

The granular glands of amphibian release a broad spectrum of biologically active peptides. So far, hundreds of peptides have been structurally characterized and these have been classified into several groups based upon the primary sequence features. Among these several different post-translational modifications (PTMs) have been characterized. These include C-terminal amidation, disulfide bridges, hydroxyproline, N-terminal pyroglutamic acid, sulfotyrosine and D-amino acids. It is assumed that these modifications contribute in some way or the other to the bioactivity of the peptides, e.g. by increasing structural stability, enhancing receptor binding or prolonging peptide half-life times (improved resistance to proteolytic degradation).

Glycosylation is a PTM that is frequently encountered on secreted proteins and also several secreted toxic peptides/proteins in other species are either N-linked or O-linked glycosylated. For example, several antibacterial glycopeptides have been discovered throughout the years. Many of them inhibit the cell-wall synthesis in gram-positive bacteria. Probably the best known of these is vancomycin, which is a branched tricyclic glycosylated nonribosomal peptide, produced by *Amycolatopsis orientates*. Also in higher organisms glycosylation of toxic peptides has been observed. Phospholipases found in the venom of arthropods and snakes are examples. The Phospholipase A2 from honeybees, the brown spider and snakes are all N-glycosylated proteins. Also O-glycosylated toxic peptides are found in venoms of the green mamba snake *Dendroaspis angusticeps*[1], the lizard *Heloderma horridum* (helospectin)[2] and marine cone (*Conus sp.*) snails[3].

Here we report on the discovery of glycosylated peptide present in the venom of the Chinese odorous frog *Odorrana schmackeri*. This is the first time that this type of post translational modification is described in frog skin secreted peptides.

By combining the information obtained from CID and ETD fragmentation spectra of the glycopeptides we were able to characterize their primary amino acid sequence and unambiguously assign the site of glycosylation to a threonine residue.

4.2 Materials and Methods

4.2.1 Venom collection

Odorrana schmackeri specimens (n = 4, snout-to-vent length 5–7 cm, adults and undetermined sex) were captured during expeditions in Wuyishan National Park, Fujian Province, in China. Collection of the skin secretion was performed by gentle transdermal electrical stimulation as described previously[4]. Secretions were washed from the frogs using de-ionized water, snap-frozen in liquid nitrogen, lyophilized, and stored at -20 °C prior to analyses.

4.2.2 Fractionation and peptide characterization

One milligram of lyophilized crude venom was solubilized in water:acetonitrile (90:10) and particulate matter was removed by centrifugation. HPLC fractionation was performed using a 2695 Alliance HPLC module (Waters, Manchester, UK) using a C8 column (Zorbax Eclipse XDB C8, 5 μm particles, 4.6 mm x 150 mm, Agilent, USA). The column was eluted with a flow rate of 1 mL/min. (solvent A, 0.05 % TFA in milliQ water; and B, 0.05 % TFA in methanol) by a linear gradient from 10 to 80% B at 1.75 % min^{-1}, followed by 10 min. in 100 % B. The peptide elution was monitored by UV at 214 nm and 10% of effluent was directly electrosprayed in the ion source a Q-TOF mass spectrometer (QTof Premier, Waters, Manchester, UK) and 90% of the effluent was directed to a fraction collector (1 min. fractions). Low energy CID MS/MS spectra were acquired by data dependent analysis and a collision energy ramp was applied from 5 to 35 eV. Multiply charged precursor fragmentation spectra were deconvoluted using Max-Ent3 software (Waters) to assist in sequence interpretation.

4.2.3 *De novo* sequencing

Five nanogram of the crude venom (reduced and alkylated) were electrosprayed (1.7 kV) into a LTQ Orbitrap XL (ThermoFisher). MS data acquisition was performed automatically, in the positive ion mode and CID and ETD fragmentation were performed on the each ion precursor selected. To facilitate the *de novo* sequencing of spectra from multiply charged precursors, they were deconvoluted (Xtract software, ThermoFisher).

4.3 Results and Discussion

4.3.1 Glycosylation feature

Initial analyses of the skin secretion of *O. schmackeri* by LC-MS/MS showed that low-energy CID spectra of the majority of peptides displayed little of no backbone fragmentation. This was rationalized by the observation that most peptides contain one if not multiple intramolecular disulfide bonds, a common PTM for skin secretions peptides of the genus *Rana*[5]. The presence of a disulfide bond within a peptide seriously hinders its fragmentation by low energy CID[5, 6, 7]. Among these low-energy CID spectra, a few exceptions triggered our attention. In total three candidate peptides that eluted quite close to each other (between 18 to 20 min., see Fig. 1), with calculated monoisotopic masses of 2448.39, 2594.45 and 2756.50 Da (Fig. 2), yielded considerably good fragmentation, allowing sequence elucidation. The low energy collision induced dissociation fragmentation spectra of these putative peptides are given in Fig. 3. Manual interrogation of these 3 spectra showed they display a similar fragmentation pattern with a-typical peptide fragmentation. In fact, dominant neutral losses corresponding to the loss of 203, 162 and 146 Da were observed. These neutral losses are commonly observed in CID spectra of glycosylated peptides and are indicative for the carbohydrate loss of N-acetyl hexosamine (HexNAc), hexose (Hex) and deoxyhexose (DeoxyHex), respectively. Next to these abundant losses of carbohydrate moieties, typical peptide fragmentation was observed. An abundant a2/b2 pair was observed at *m/z* ratios of 143 and 171 respectively. From a corresponding y”-ions series it was deduced that all three peptides share the same N-terminal amino acid sequence of AVP(I/L)(I/L)Y. However the remaining primary structure could not be unambiguously assigned from this CID spectrum and also the site and type of glycosylation was at this point unknown.

4.3.2 *De novo* sequencing and glycosylation localization

In a second stage of analysis these three glycopeptides were further characterized by employing a different type of peptide fragmentation, namely Electron Transfer Dissociation (ETD). In this technique positively charged peptide ions are mixed with radical anions. Electrons are transferred from the radical anions onto the peptide cations inducing random cleavages along the peptide backbone. It has been observed that the majority of the side chains of amino acids and more importantly post-translational modification such as phosphorylation stay intact during this fragmentation, making it very suitable to locate PTM-

sites. The ETD spectra of the three peptides are given in Fig. 4. From these spectra it is clear that the glycan moieties are indeed not cleaved from the peptide backbone during fragmentation. Instead using the partial amino acid sequence obtained from the CID spectra, the corresponding c and z ion series could easily be found and eventually be extended to a full length primary structure of AVP(I/L) (I/L)YNRPG(I/L)YVTKRPK-amide. As illustrated in Fig. 4 an atypically large mass difference between c4-c5 and z12-z13 was observed. This mass difference (774.29 Da) corresponds to the combined masses of 4 monosaccharides added to the mass of a Thr residue, suggesting O-linked glycosylation. Manual *de novo* sequencing of the ETD spectra, showed that ions from N-terminal Pro cleavage were absent, as expected[8], whereas in CID spectra these yielded high y"-ion peaks, confirming the positions of the Pro residues.

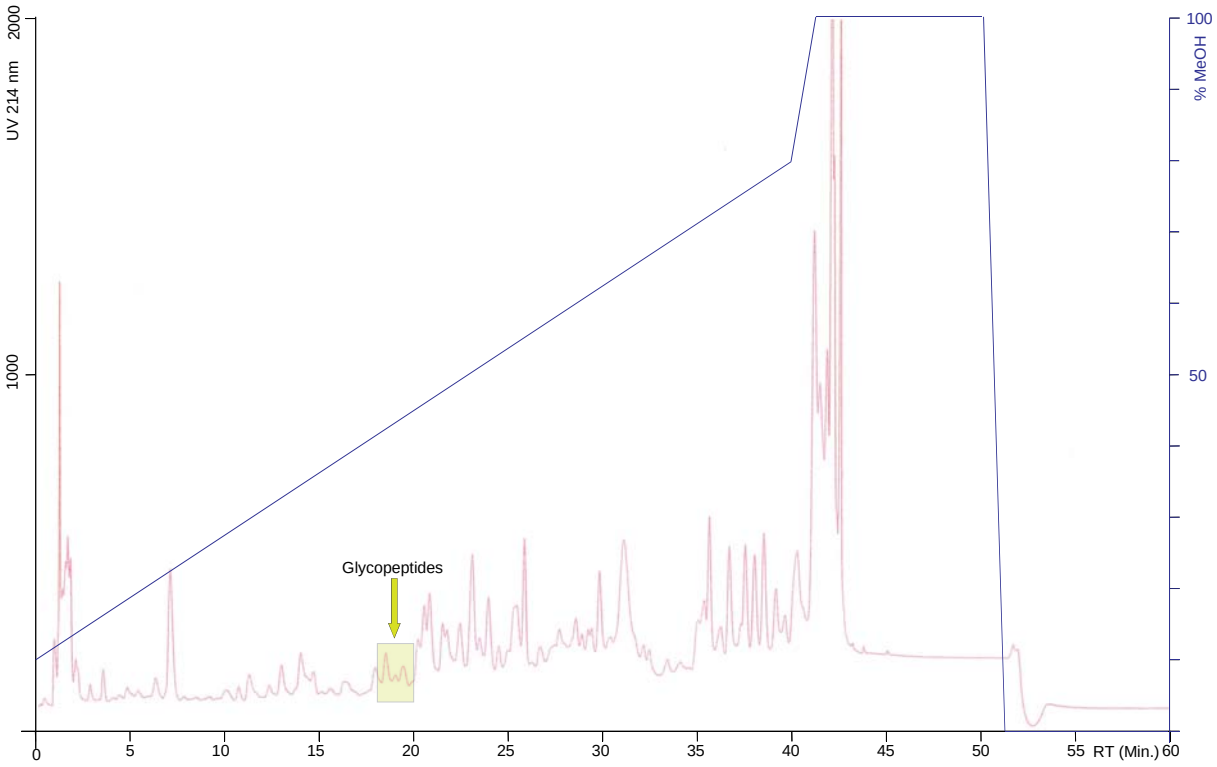


Figure 1: UV chromatogram of Odorrana schmackeri *crude venom monitored at 214 nm with elution time of glycosylated peptides indicated.*

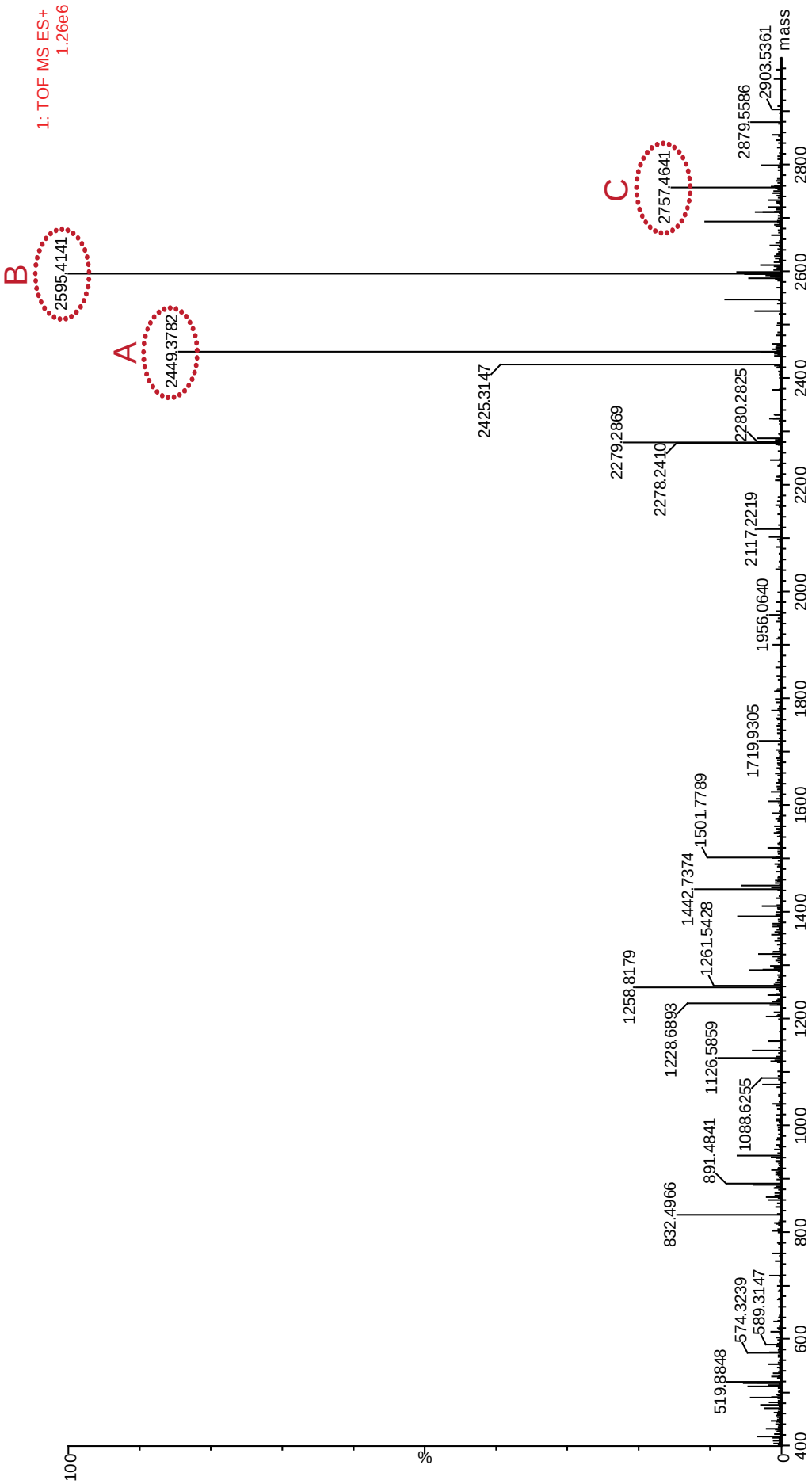


Figure 2: Deconvoluted (MaxEnt3) QTOF MS spectrum displaying 3 different glycopeptides. MS spectra of elution time highlighted in Fig. 1 were combined prior to deconvolution.

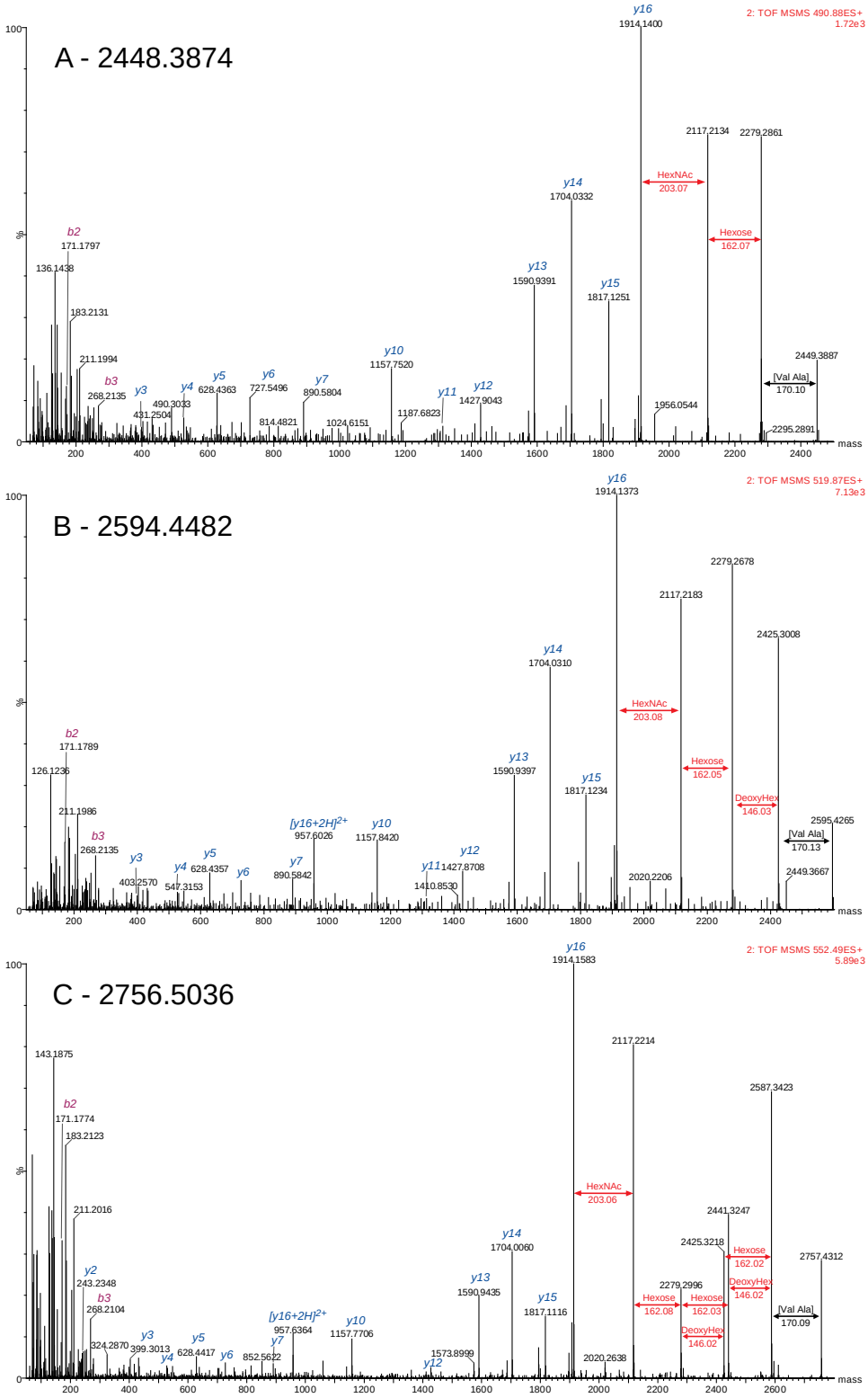


Figure 3: Deconvoluted (MaxEnt3) low energy CID spectra of peptides A, B and C showing abundant neutral losses of carbohydrates. All 3 peptides share peptide backbone fragmentation pattern, differing only at glycan content: glycan of peptide A is N-acetyl hexosamine (HexNAc) linked to hexose; glycan of peptide B has an extra deoxyhexose (DeoxyHex); and glycan of peptide C has yet another hexose. These spectra do not allow determination of glycosylated residue, as sugar moieties are cleaved from peptide backbone during CID fragmentation.

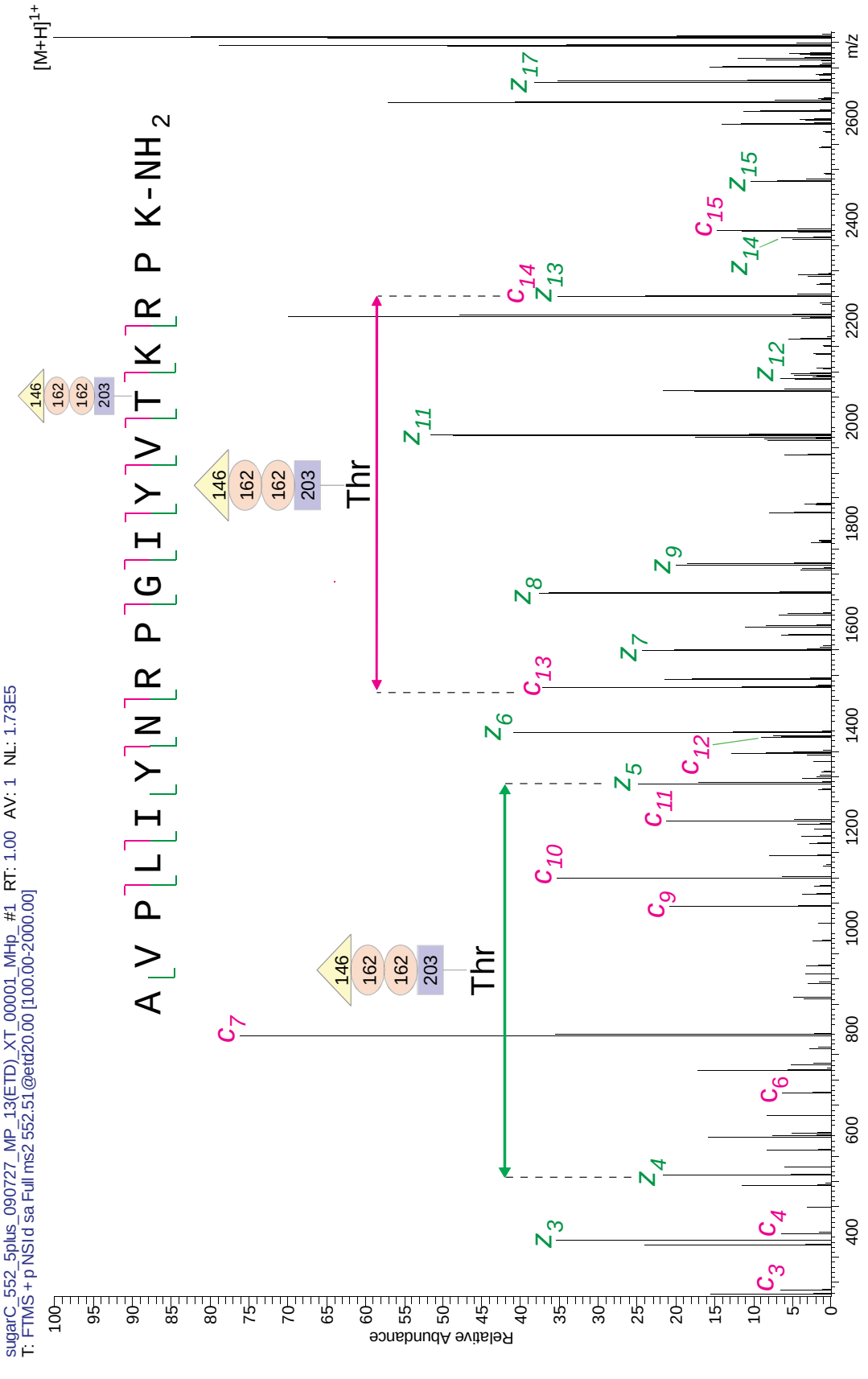


Figure 4: Deconvoluted (Xtract) ETD spectrum of five times charged glycopeptide C with sequenced residues annotated indicating glycosylation at threonine residue.

With the full primary sequence solved, the potential presence of the non-glycosylated version of this peptide could be investigated. Extracted ion chromatograms were generated for different charge states of the unmodified peptide (both amidated and non-amidated). This showed that the unmodified peptide does not occur at detectable levels within the skin secretion. However, in addition to the majority of the glycosylated peptides being C-terminally amidated, also low levels of the non-amidated isoforms of all three glycosylated peptides were detectable in the venom (data not shown). Further interrogation of the primary amino acid sequence by means of a BLAST against the NCBI non-redundant protein database showed that the *O. schmackeri* glycopeptide sequence shows near 100% sequence similarity with sequences from *R. andersonii*, *R. livida*, *R grahami*, *A. mantzoum* and *R. tiannanensis.* All of these primary structures are in fact predicted from prepropeptide precursor sequences obtained by molecular cloning[9] (e.g. mature peptide cleavage at KR, etc.). For all these precursors processing of the C-terminus has not been determined at the peptide/protein level. Our peptide analysis in *O. schmackeri* suggests the GK in the precursors listed in Table 1 is in fact removed, yielding a Lys-amide C-terminus in the mature peptide. More importantly, the molecular cloning did not predict other PTMs, and our data suggest that, with the exception of the *R. grahami* peptide (in which Thr is substituted by Pro), the mature form of these peptides are O-glycosylated on the conserved Thr.

Table 1. Alignment of primary structure of O. schmackeri *glycopeptide with translated cDNA sequences of similar Rana antimicrobial peptides.*

1 *O. schmackeri* **[*]**	**AVPLIYNRPGIYVTKRPK**-amide
2 *Rana andersonii*	**AVPLIYNRPGIYVTKRPK**GK
3 *Rana livida*	**AVPLIYNRPGIYVTKRPK**GK
4 *Rana grahami*	**AVPLIYNRPGIY**AP**KRPK**GK
5 *Amolops mantzoum*	**AVPLIYNRP**S**IYVTKRPK**GK
6 *Rana tiannanensis*	**AVPLIYNRP**S**IYVTKRPK**GK

[*] Leu and Ile residues are predicted on the basis of alignment with the other *Rana* peptides.

4.4 Conclusion

Three novel O-glycosylated peptides were discovered and characterized from the skin secretion of the Chinese odorous frog *Odorrana schmackeri.* CID spectra showed they contain of 2, 3 or 4 different monosaccharides (hexose, deoxyhexose and N-acetyl hexosamine) in each peptide. The full primary sequence was obtained from ETD spectra, which also allowed the assignment of the site of glycosylation to the threonine residue. All peptides were also C-terminally amidated. Sequence analyses showed that these peptides are in fact highly conserved between various members of the *Rana* genus. Our data are the first to report/suggest that in their mature form these frog venom peptides occur predominantly as O-linked glycopeptides. Both the function as well as the exact structure of the glycan moieties remain at the moment unknown and this will be the subject of further studies.

Acknowledgments

We acknowledge CNPq (Brazilian National Counsel of Technological and Scientific Development) for a PhD grant to G. Evaristo. (Bolsista CNPq- 200847/2007-4).

References

1. Quinton, L.; Gilles, N.; Smargiasso, N.; Kiehne, A.; De Pauw, E., An unusual family of glycosylated peptides isolated from Dendroaspis angusticeps venom and characterized by combination of collision induced and electron transfer dissociation. *J Am Soc Mass Spectrom* **2011,** 22, (11), 1891-7.
2. Vandermeers-Piret, M. C.; Vandermeers, A.; Gourlet, P.; Ali, M. H.; Waelbroeck, M.; Robberecht, P., Evidence that the lizard helospectin peptides are O-glycosylated. *Eur J Biochem* **2000,** 267, (14), 4556-60.
3. Craig, A. G.; Zafaralla, G.; Cruz, L. J.; Santos, A. D.; Hillyard, D. R.; Dykert, J.; Rivier, J. E.; Gray, W. R.; Imperial, J.; DelaCruz, R. G.; Sporning, A.; Terlau, H.; West, P. J.; Yoshikami, D.; Olivera, B. M., An O-glycosylated neuroexcitatory conus peptide. *Biochemistry* **1998,** 37, (46), 16019-25.
4. Tyler, M. J.; Stone, D. J.; Bowie, J. H., A novel method for the release and collection of dermal, glandular secretions from the skin of frogs. *J Pharmacol Toxicol Methods* **1992,** 28, (4), 199-200.
5. Wang, L.; Evaristo, G.; Zhou, M.; Pinkse, M.; Wang, M.; Xu, Y.; Jiang, X.; Chen, T.; Rao, P.; Verhaert, P.; Shaw, C., Nigrocin-2 peptides from Chinese Odorrana frogs--integration of UPLC/MS/MS with molecular cloning in amphibian skin peptidome analysis. *Febs J* **2010,** 277, (6), 1519-31.
6. Samgina, T. Y.; Artemenko, K. A.; Gorshkov, V. A.; Poljakov, N. B.; Lebedev, A. T., 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* **2008,** 19, (4), 479-87.
7. Samgina, T. Y.; Gorshkov, V. A.; Vorontsov, E. A.; Bagrov, V. V.; Nifant'ev, I. E.; Lebedev, A. T., New cysteine-modifying reagents: Efficiency of derivatization and influence on the signals of the protonated molecules of disulfide-containing peptides in matrix-assisted laser desorption/ionization mass spectrometry. *Jounal of Analytical Chemistry* **2010,** 65, (13), 1320-1327
8. Mikesh, L. M.; Ueberheide, B.; Chi, A.; Coon, J. J.; Syka, J. E.; Shabanowitz, J.; Hunt, D. F., The utility of ETD mass spectrometry in proteomic analysis. *Biochim Biophys Acta* **2006,** 1764, (12), 1811-22.
9. Li, J.; Xu, X.; Xu, C.; Zhou, W.; Zhang, K.; Yu, H.; Zhang, Y.; Zheng, Y.; Rees, H. H.; Lai, R.; Yang, D.; Wu, J., Anti-infection peptidomics of amphibian skin. *Mol Cell Proteomics* **2007,** 6, (5), 882-94.

Chapter 5

MS-based discovery of two novel peptides with homology to calcitonin-like peptides, from the skin secretion of the Chinese frog, *Odorrana schmackeri*.

Abstract

The primary structures of two new 34 residue peptides were elucidated from the skin secretion of the Chinese frog *Odorrana schmackeri.* This chapter describes our mass spectrometry based discovery and analytical strategy to solve the complete sequence of both peptides, including the elucidation of their post-translational modifications. *De novo* sequencing was achieved by the analysis of complementary high resolution CID and ETD fragmentation spectra of full length peptides as well as of selected tryptic fragments. Heavy and light isotope dimethyl labeling assisted with the assignment of tandem MS sequence ions. The identified peptide sequences are: GCDLSTCATHNLVNELNKFDKSKPSSGGVGPESP-NH_2 and SCNLSTCATHNLVNELNKFDKSKPSSGGVGPESF-NH_2. The exact sequence (including I/L discriminations) was finally confirmed by Edman degradation. Both peptides have a disulfide bond in the N-terminal region, which is quite uncommon for frog defense peptides, and both sequences are not encountered in frog venom before and share sequence homology with calcitonin, calcitonin gene related peptide (CGRP) and adrenomedullin from other vertebrates. Detailed sequence analysis as well as the 34 residue length of both *O. schmackeri* peptides, suggests they may represent members of a novel peptide family within the calcitonin gene related peptide superfamily, as they do not fully qualify as either calcitonins (32 residues), CGRPs (37 amino acids) or adrenomedulins (>50 residues).

Manuscript in preparation

Evaristo, G. P. C.; Pinkse, M. W. H.; Verhaert, P. D. E. M., MS-based discovery of two novel peptides with homology to calcitonin-like peptides, from the skin secretion of the Chinese Frog, *Odorrana schmackeri*.

5.1 Introduction

In our study of the peptidome secreted by *O. schmackeri* defense glands[1, 2] (Chapters 3 and 6, this thesis), our attention was triggered towards one peptide with a single disulfide bond located within its N-terminal domain. Disulfide bonds are quite common post-translational modifications encountered on frog defense peptides, however defense peptides from ranid frogs often have a C-terminal S-S loop, the so called 'rana-box'. The occurrence of the disulfide bond at the N-terminus of a frog defense peptide is quite unique. Due to this and since this post-translational modification in general renders a peptide its specific tertiary structure and is, therefore, often associated with bioactivity, we decided to perform a detailed investigation into this peptide in order to obtain its full primary amino acid sequence utilizing mass spectrometry, HPLC purification and Edman degradation. During this investigation a second peptide was discovered that, although much lower in abundance, showed similar retention behavior in HPLC and fragmentation in tandem MS analysis. Both peptides were the subject of intensive and thorough MS-based analysis to unravel their full primary amino acid sequence as described below. Both peptides may be members of a novel class of peptides within the family of frog skin defensive peptides and it will be intriguing to investigate their bioactivity.

5.2 Materials and Methods

5.2.1 LC MS/MS analysis: ETD and CID fragmentation

Three equivalent batches (0.5 mg) of crude venom of *Odorrana schmackeri* were prepared for nano HPLC MS/MS analysis on a LTQ Orbitrap XL (ThermoFisher Scientific) equipped with ETD fragmentation. All three samples were reduced with 5 mM dithiothreitol (DTT), one batch was alkylated with 10 mM iodoacetamide (IAM), the second with 10 mM bromoethylamine (BrEA), and the third was not alkylated. During LC MS/MS analysis, the two most intense ions from the survey scan were automatically selected for fragmentation by both CID and ETD. FT MS and FT MS/MS spectra were recorded respectively at 60.000 and 30.000 resolution at *m/z* 400, in positive ionization mode and as profile data. For fragmentation analysis, ions were accumulated in the linear ion trap with an isolation width of 3 amu and a target value of $3e^5$. For CID fragmentation a normalized collision energy of 35% and an activation time of 30 ms were used. For ETD fragmentation the anion target was set to

$2e^5$ and activation time was set 50 ms. Supplemental activation was enabled and the energy was set to 25%.

5.2.2 Peptide purification

A 4 mg aliquot of *Odorrana schmackeri* skin secretion was fractionated by reverse phase HPLC (Waters 2695 Alliance, Manchester, UK) on a C8 column (Zorbax- XD, 5 μm particles, 4.6 x 150 mm, Agilent, Amstelveen, The Netherlands) employing a 40 min. linear gradient from 10 to 80% of methanol with 0.05% trifluoroacetic acid (TFA). The flow rate was 1 mL/min. and fractions were automatically collected every 2 minutes. The column effluent was split in two; 90% going to the fraction collector and 10% going to the ion-source of a Q-TOF MS system (Waters QTof Premier, Manchester, UK) yielding a *m/z*-based chromatogram of the eluate.

Fractions containing the peptide of interest were vacuum-dried and reconstituted in 50 μL of 25 mM ammonium bicarbonate buffer, pH 8.0. Disulfide bonds were reduced with DTT (final concentration 2 mM, for 30 min. at room temperature) and free thiol groups were blocked with IAM (iodoacEtamide, final concentration 4 mM, for 45 min. at room temperature in the dark). This sample was fractionated again by HPLC (Waters 2695 Alliance, Manchester, UK) on a C4 column (Reprosil 5 μm particles, 2.0 x 150 mm, Dr. Maisch, Ammerbuch-Entringen, Germany). The flow rate was set to 0.25 mL/min. and peptides were eluted using a linear gradient, starting 5 min. after injection, of 1% B/min. in 50 min. Solvent A was 0.005% TFA in milliQ water and B was 80% acetronitrile, 20% milliQ, 0.004% TFA. Elution was monitored at 215 nm, the fractions were collected every minute and the peptide presence was confirmed by high resolution MS (Orbitrap Velos) measurements. One half of the fraction containing the peptide of interest was vacuum dried, and subsequently trypsinized with 1 μg enzyme in 100 mM tri-ethyl ammonium bicarbonate buffer (TEAB)for 1.5 h at room temperature. Subsequently it was submitted to another fractionation using the same C4 column, gradient and solvents as described above. After this purification, fractions of interest were subjected to further MS analysis and were submitted for automated Edman degradation on a pulsed liquid-phase sequencer (Procise 492 cLC, Life Technologies, Grand Island, N.Y., USA).

5.2.3 Dimethyl labeling

The purified tryptic fragments of the peptides of interest were diluted in TEAB, and split in two aliquots and subjected to dimethyl labeling[3]. Either 4 μL normal or 4 μL

deuterated formaldehyde (4% solution) were added to the peptide solution together with 4 μL of 0.6 M cyanoborohydrate solution and this mixture was left to react for 1 h at room temperature, after which the labeling was stopped by addition of 16 μL 1% ammonia solution. Heavy and light labeled peptides were mixed and analyzed by nano HPLC coupled to an Orbitrap Velos. Peptides were fragmented by CID, using a normalized collision energy of 35%, fragments were analyzed in the orbitrap analyzer (parameters: resolution (at *m/z* 400) 30.000, CID activation time 30 ms, isolation width 2.5 amu).

5.2.4 Data analysis

All the obtained mass spectra were manually analyzed within Thermo Xcalibur Qual Browser (version 2.1). Spectra were deconvoluted using the Xtract option within Xcalibur software. The obtained primary amino acid sequences were aligned against the NCBI non-redundant protein database using the BLASTp algorithm or against the reference genomic sequence (refseq_genomic) using the tBLASTn algorithm (http://blast.ncbi.nlm.nih.gov/).

5.3 Results and Discussion

5.3.1 Stepwise sequence elucidation

Peptides from the skin extract of *Odorrana schmackeri* were analyzed by nano HPLC coupled online to mass spectrometry. In order to recognize peptides within the frog skin extract that carry a disulfide bond, three different samples were prepared; peptides solely reduced with dithiotreitol, peptides reduced with dithiotreitol and subsequently free-thiol conversion to ethylamine-cysteine, or peptides reduced with dithiotreitol and subsequently free-thiol conversion to carbamidomethyl cysteines. All three samples were separated and analyzed by nano HPLC coupled online to a high resolution mass spectrometer, capable of both collision induced dissociation (CID) and electron transfer dissociation (ETD) fragmentation. From these analyses hundreds of high quality fragmentation spectra were obtained. Several of these spectra could be directly linked to peptides previously sequenced from frog skin secretion by automated database searching, however many others remained unresolved and were further manually interrogated. One peptide attracted our interest in particular. This peptide had a mass of 3563.692 Da in the natural form, and showed after reduction a mass increase of 2.027 mass units, indicative for the presence of 1 disulfide bond

(see Fig. 1A and B). After reaction of the free thiols with both alkylating reagents, a mass increase was observed that showed it has a total of 2 cysteines per peptide (Fig. 1C and D). By comparing the ETD fragmentation spectra of this peptide with ethylamine-modified cysteines (EtA-Cys) or carbamidomethyl cysteines (CaM-Cys) an interesting observation was made. In the low molecular weight region of the spectra a small difference was observed. The EtA-Cys peptide showed three fragment ions at *m/z* 251.117, 365.160 and 478.243, while the CaM-Cys peptide showed three fragment ions at *m/z* 265.096, 379.138 and 492.222 (see Fig. 2 A and B). This fragment ion-series is shifted with a mass difference of exactly 13.978 amu, which fits perfect with the mass difference between a single CaM-Cys (160.031) and EtA-Cys (146.051). The interesting part here is that the smallest ion observed with this characteristic mass difference is the 251.117/265.096 ion pair. Since this is an ETD spectrum, this ion is most likely a c2 or a z2 ion of the first two N-terminal or the last two C-terminal amino acids of this peptide, respectively. If we assume it is the c2-ion, with one cysteine, the remaining mass can be assigned to a serine. In the case it is a z-ion, no other amino acid could be fitted in combination with the cysteine, even when C-terminal amidation is included as possible additional modification. In summary this is how this peptide attracted our attention.

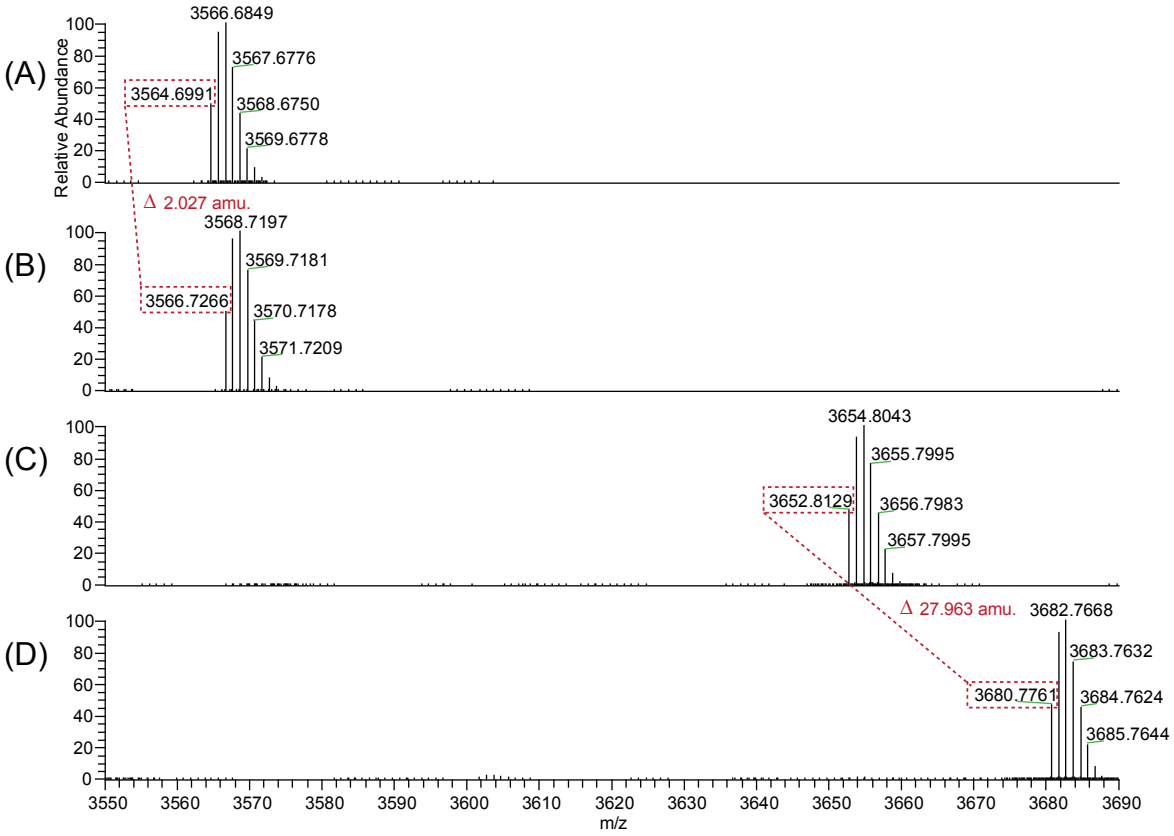


Figure 1: Deconvoluted MS spectra of OsCLP-2, (A) non-reduced, (B) reduced, (C) EtA-derivatized and (D) CAM-derivatized.

Therefore, the peptide has a disulfide bond and at this stage we could locate it at the N-terminus for at least one of the cysteines. This by itself is quite uncommon for defense peptides from ranid frogs, of which the majority have the disulfide bond at the C-terminal part of the peptide. Due to this atypical location of the cysteine/disulfide bond we decided to put more effort into this particular peptide and try to sequence as much as possible. Using the CID and ETD fragmentation spectra of the peptide with either EtA- or CaM-cysteines, the first 19 N-terminal amino acids, namely SCN(I/L)STCATHN(I/L)VNELNKF, of this peptide could be sequenced with high confidence. For illustration, Fig. 3 displays the CID and the ETD spectrum of the $[M+5H]^{5+}$ of the EtA-Cys-modified peptide with its partially elucidated sequence. It should be noted that at this point no distinction could be made between leucine and isoleucines, also not on basis of side-chain fragmentation in the ETD spectrum[4, 5]. Using this partial sequence, a BLAST search was done against the NCBI non-redundant protein database (Update date:2012/05/09, Number of sequences:17919084). The results indicated the presence of a putative conserved domain of the calcitonin, CGRP, IAPP family (pfam00214), however, no significant match to any of the sequences in the database could be made.

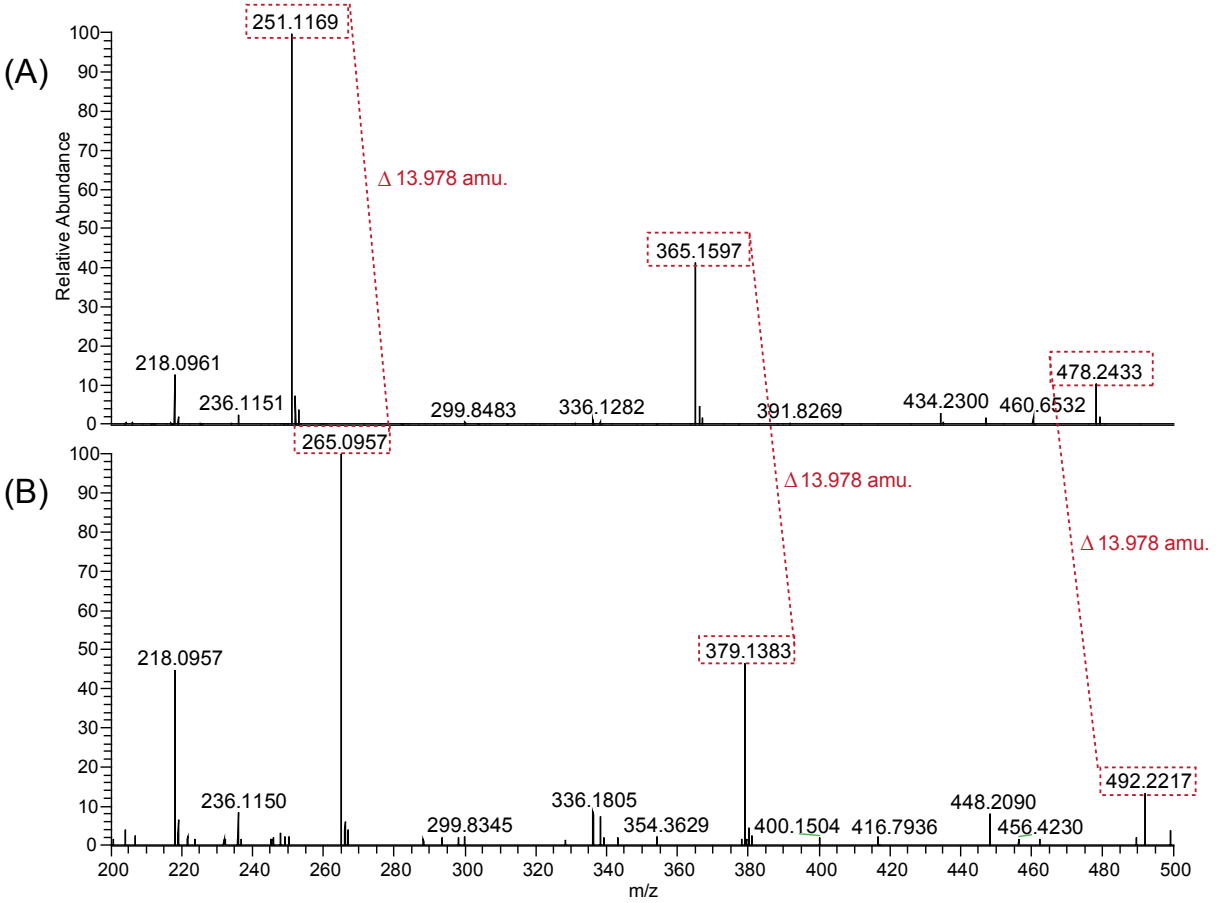


Figure 2: Zoom in of ETD-MS2 spectra of the $[M+5H]^{5+}$ of OsCLP-2 after (A) EtA and (B) IAM derivatization showing ions shift of 13.978 amu.

Since the sequence obtained so far was only the N-terminal part of the peptide representing about 60% of the entire sequence, it was decided to further purify the peptide from the venom in order to perform more detailed analysis. Four mg of crude venom was fractionated by HPLC using a C8 column. The eluent of the column was split to the fraction collector and Q-TOF mass spectrometer. In this way the exact location of the peptide could be traced back to the fraction number. After the HPLC-MS analysis an extracted ion chromatogram was made for the $[M+3H]^{3+}$ of the peptide of interest. A high abundant signal for the peptide was detected at a retention time of around 32 minutes. At this stage of the analysis a closely eluting peptide was noticed that displayed a similar charge state distribution and a similar but not identical molecular mass as the peptide of interest. We re-analyzed the initial nano HPLC-MS data acquired on the Orbitrap Velos which has higher mass accuracy and from this it was found that the second peptide is 79.046 amu smaller in molecular mass, also has an N-terminal disulfide bond and its CID/ETD fragmentation is largely comparable to the other peptide (Fig. 4 displays the ETD spectra of these two peptides).

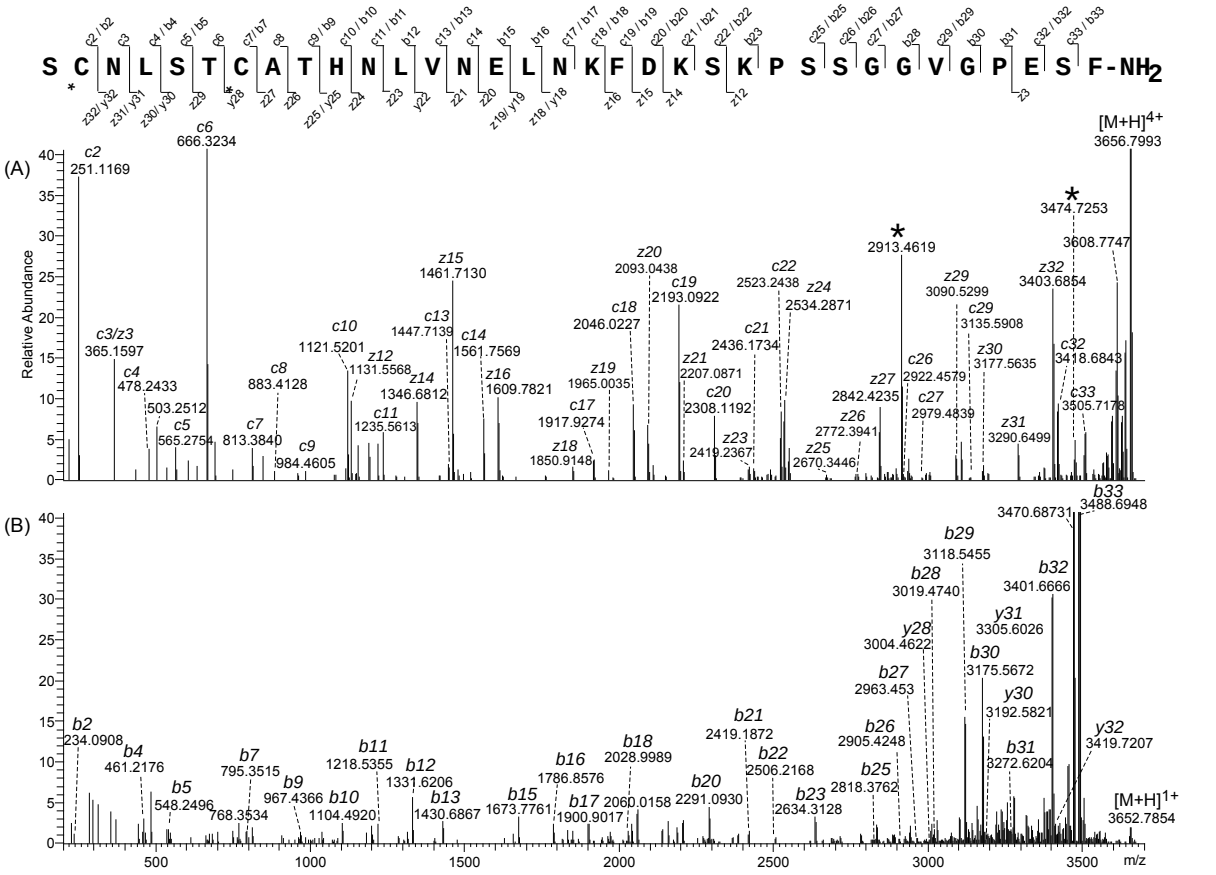


Figure 3: Deconvoluted ETD (A, cysteines modified with BrEA) and CID (B, cysteines modified with IAM) spectra of OsCLP-2. Ions de novo *sequenced are assigned in spectra and labeled on peptide sequence. Ions detected below 5 % are not annotated in the spectra, but are shown on sequence. [*]Ions yielding specific side chain loss of modified cysteine (-76.022 amu).*

From the ETD spectrum of this second peptide only a small part of the N-terminal sequence could be elucidated as GCDLSTCATH. Also this sequence shows high similarity to the Calcitonin, CGRP, IAPP family (pfam00214). From now on we designated these two peptides OsCLP-1 and OsCLP-2 (*Odorrana schmackeri* calcitonin-like peptide 1 and 2).

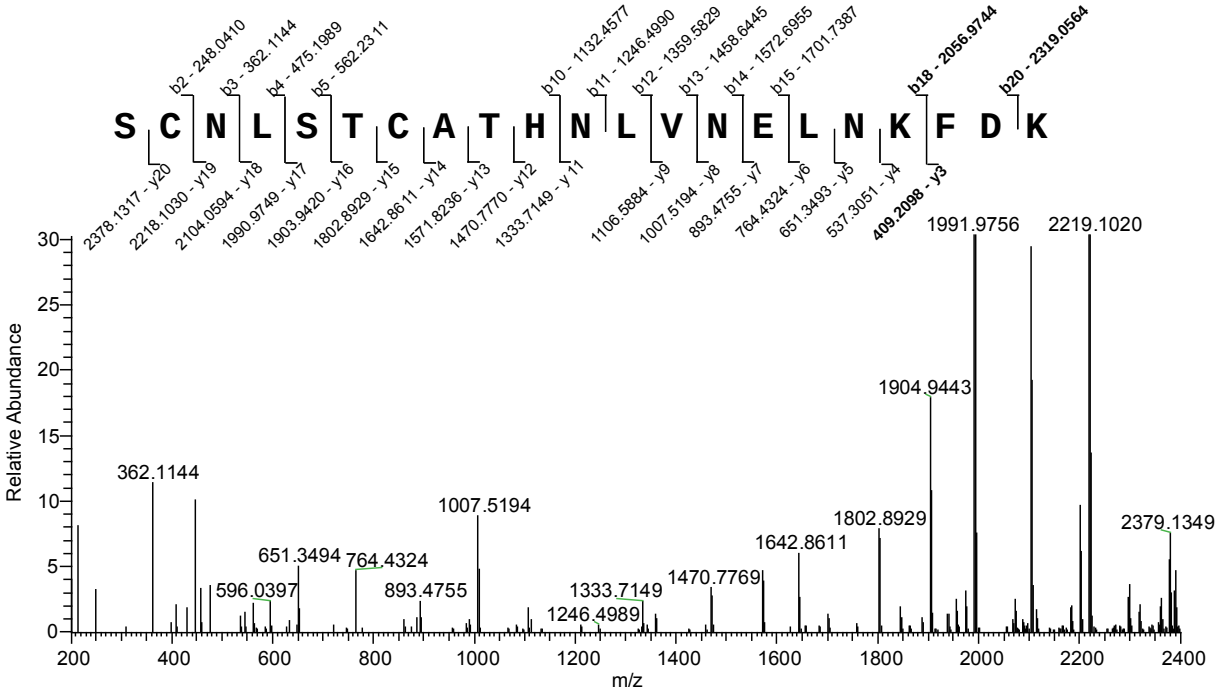


Figure 4: Deconvoluted CID spectrum of the $[M+4H]^{4+}$ at m/z *617.0465 of OsCLP-2 N-terminus tryptic fragment containing carbamidecysteines. The* b18, b20 *and* y3 *ions confirm the sequence [FDK].*

The two fractions containing both peptides (#15 and #16) still contained several other peptides with different masses. Both fractions were vacuum dried and after reconstitution subjected to reduction by DTT and alkylation with IAM. It was anticipated that by reduction and alkylation of the cysteines the retention times of all peptides might differ in a next round of HPLC fractionation, yielding more pure fractions containing the peptide of interest. After this next round of separation, OsCLP-2 was indeed obtained in higher purity, whereas OsCLP-1 was unfortunately not detected anymore. In a last purification step, the fraction containing the most of the carbamidomethyl modified OsCLP-2 was digested with trypsin. Trypsin was chosen since, at least, one tryptic cleavage site was present in the sequence obtained sofar. The resulting digest was again purified by HPLC and this yielded three major UV absorbing peaks (210 nm) with molecular masses 1233.604, 2073.960 and 2464.156 Da. The occurrence of 3 peaks at this point could indicate that there is a second trypsin cleavage site. The two bigger peptides have a mass difference of 390.196 amu. Assuming that the difference is of 3 amino acids and by calculating the masses of all 3 amino acid combination

using the 20 amino acids, several possible combinations were found; [FSR] or [AYR] (390.2016), [PHR] (390.2128), [VYK] (390.2267) or [FDK] (390.1903). There are two other options [YVQ] a [(I/L)YN] which have and identical mass of 390.1903, however these do not have a lysine of arginine residues, and hence no second tryptic cleavage site. Within this part of the peptide, there is a phenylalanine and by assuming there is indeed another trypsin cleavage site (i.e. a lysine or an arginine) this leads to two possibilities (FDK, calculated mass 390.1903 or FSR mass 390.2016). All three peptides were subjected to nano HPLC-Orbitrap MS/MS analysis. From this it was confirmed on basis of the accurate mass and the fragmentation spectrum that the N-terminal part of the peptide is SCN(I/L)STCATHN(I/L)VNE(I/L)NKFDK (Fig. 4). For the smaller C-terminal peptide of molecular mass 1233.604 Da the fragmentation spectrum yielded a partial sequence of SSGGVGPESF-amide (Fig. 5). The first 3 amino acids of this tryptic fragment remained unclear due to absence of the b1 and its complementary y"-ion. On basis of the accurate mass of the b3 ion, and calculation of the mass of all b3-ions possible with the 20 natural amino acids, only one combination was found to fit within the error of the measurement, namely [SPK]. However, only on basis of this calculation the order of these amino acids within the peptide could obviously not be made. It was decided to do a labeling experiment in combination with tandem MS analysis on this tryptic peptide via reductive methylation with formaldehyde. In the presence of $NaCNBH_3$, all the α- and ε-amino groups of a peptide are di-methylated, yielding a mass increase of 28 amu per reaction. After labeling, the C-terminal tryptic peptide was increased by 56.050 amu, indicating there are two sites available for labeling, the α-amino group of the N-terminus and most likely the ε-amino group of an additional lysine.

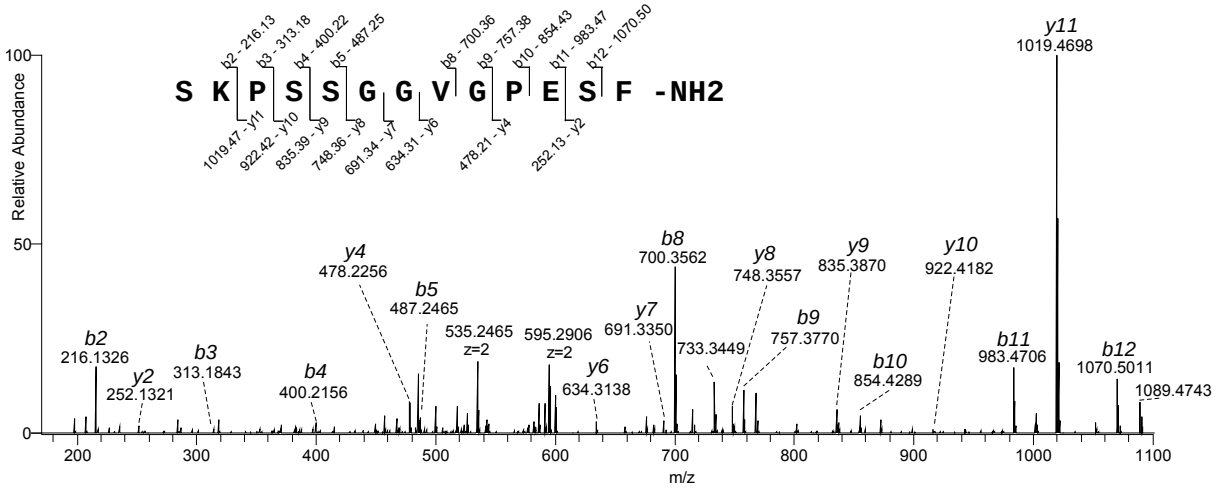


Fig. 5: CID spectrum of the $[M+2H]^{2+}$ at m/z *617.8037 of the C-terminal tryptic fragment of OsCLP-2.*

For sequencing purposes, the peptide was also labeled with deuterated formaldehyde (d2-formaldehyde). Figure 6 shows the MS/MS spectra of this C-terminal tryptic fragment after labeling with normal formaldehyde and deuterated formaldehyde. With this labeling experiment using high resolution orbitrap measurements, the assignment of the C-terminal sequence could be completed. The first three residues of the C-terminal tryptic peptides are SKP. The dimethyl labeling revealed the presence of two sites for labeling, next to the N-terminus a lysine residues is labeled. This lysine residue was not cleaved during the tryptic digestion due to the adjacent proline. The complete amino acid sequence of OsCLP-2 is SCNLSTCATHNLVNELNKFDKSKPSSGGVGPESF-amide. Additionally, the purified tryptic fragments of OsCLP-2 were of sufficient purity and amount for Edman degradation analysis. The results from this analysis confirmed our proposed primary amino acid sequence and unambiguously assigned the presence of either leucine or isoleucine within it. Finally the high sequence homology between OsCLP1 and 2 prompts us to speculate that the final sequence of OsCLP-1 also has all Leu rather than Ile residues and that its full length sequence is: GCDLSTCATHNLVNELNKFDKSKPSSGGVGPESP-NH_2.

5.3.2 Sequence homology analysis: members of a novel calcitonin-like peptide family?

Sequence homology analysis (BLAST) of the OsCLP primary structures showed they have some common features with adrenomedullin (AM), calcitonin (CT) and calcitonin gene-related peptide (CGRP) from different animal species. Other members of this family are calcitonin receptor-stimulating peptide (CRSP), and amylin (AMY). These peptides are grouped by their sequence homology, and originate from the same or related genes. CT and CGRP are both products of the same calcitonin (CALC) gene, which is by alternative exon splicing transcribed into two different mRNAs. In the mammalian thyroid gland the CALC gene is primarily expressed as CT, whereas in the central and peripheral nervous system neurons, the CALC gene is primarily expressed as CGRP [6, 7].

The alignment of the OsCLPs with several animal CTs, CGRPs and AMs is illustrated in Table 1 and 2. All CTs have 32 residues with a conserved disulfide bridge between Cys_1-Cys_7, and have a carboxyterminal phenylamine-amide. All CGRPs have 37 residues, the disulfide bond is in between Cys_2-Cys_7 and they have a proline-amide at the C-terminus. AM are peptides of 52-57 amino acids with a disulfide bond located in the middle of the peptides and a tyrosine-amide C-terminus. All three AM, CT and CGRP have strongly conserved amino acid residues in both N and C terminal part between various organisms, as highlighted in Table 1 and 2 [8, 9].

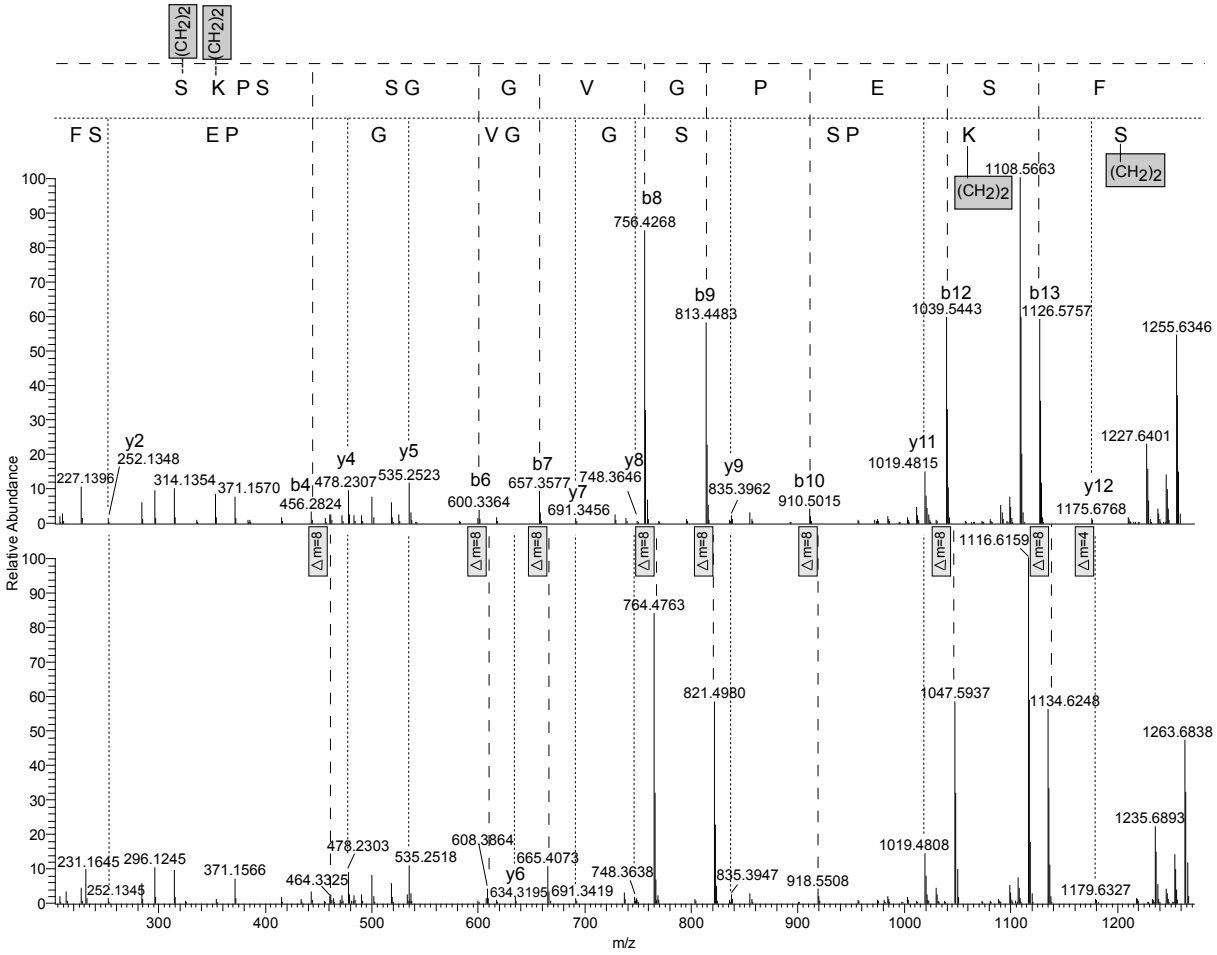


Figure 6: Deconvoluted CID spectra of the $[M+2H]^{2+}$ at m/z *645.8420 (top) and $[M+2D]^{2+}$ at* m/z *649.8661 (bottom) of the C-terminal tryptic fragment of OsCLP-2 labeled light with and heavy dimethyl. b4-ion had the mass increased in 56.067 Da on top due to label of Ser^1 and Lys^2 with light dimethyl, and 64.117 Da on bottom due to labeling with deuterated dimethyl. Therefore, a difference in mass of 8.050 Da was observed in all b-ion peaks. y-ions of heavy and light labeled peptide showed the same values until y12, increasing mass to 156.195 Da (light dimethylation) and 160.152 Da (heavy dimethylation) due to the label of Lys. Non-annotated y3 (381.1776) and y10 (922.4301 and 922.4270) ions were present in the native spectra (non-deconvoluted) and helped to confirm the peptide sequence.*

Table 1: Alignment of OsCLP 1 and 2 with (A) mature calcitonin (CT), (B) calcitonin gene related peptide (GCRP) and (C3) adrenomedullin (AM). CT, CGRP and AM sequences from different species were derived from translation of the CALC or ADM gene (in NCBI refseq_genomic) of some vertebrates or derived from the UNIPROT protein database.

A

Calcitonins

```
Zebrafish*                  CSSLSTCVLGKLSQELHKLQTYPRTNVGAGTP
Xenopus**                   CSNLSTCVLGKLSQELHKLQSYPRTDVGAGTP
Rana catesbeiana (P0C234)   CSGLSTCALMKLSQDLHRFNSYPRTNVGAGTP
Lizard***                   CANLSTCVLGKLSQELHKMQTYPRTDVGAGTP
Chicken (P07660)            CASLSTCVLGKLSQELHKLQTYPRTDVGAGTP
Ox (P01260)                 CSNLSIYVLGIYTQDLNKIYIFYLTKTGDRMP
Rat (P01257)                CGNLSTCMLGTYTQDLNKFHTFPQTSIGVGAP
Human (P01258)              CGNLSTCMLGTYTQDFNKFHTFPQTAIGVGAP
Pig (P25117)                CSNLSTCVLSAYWRNLNNFHRFSGMGFGPETP
Horse (Q9N0V5)              CSNLSTCVLGTYTQDLNKFHTFPQTAIGVGAP
Dog (P41547)                CSNLSTCVLGTYSKDLNNFHTFSGIGFGAETP
O. schmackeri OsCLP 1       GCDLSTCATHNLVNELNKFDKSKPSSGGVGPESP
O. schmackeri OsCLP 2       SCNLSTCATHNLVNELNKFDKSKPSSGGVGPESF
```

B

CGRP

```
Zebrafish*                  ACNTATCVTHRLADFLSRSGGIGSSKFVPTNVGSQAF
Xenopus**                   ACNTATCVTHRLADFLSRSGGMGKNNFVPTNVGSKAF
Rana ridibunda[9]           ACNTATCVTHRLADFLSRSGGMAKNNFVPTNVGSKAF
Phyllomedusa bicolor (P81564) SCDTSTCATQRLADFLSRSGGIGSPDFVPTDVSANSF
Lizard***                   GCNTATCVTHRLADFLSRSGGVGKNNFVPTNVGSKGF
Chicken (P10286)            ACNTATCVTHRLADFLSRSGGVGKNNFVPTNVGSKAF
Rat (P01256)[4]             SCNTATCVTHRLAGLLSRSGGVVKDNFVPTNVGSEAF
Human A (P06881)            ACNTATCVTHRLAGLLSRSGGMVKSNFVPTNVGSKAF
Human B (P10092)            ACDTATCVTHRLAGLLSRSGGVVKNNFVPTNVGSKAF
Pig (P30880)                SCNTATCVTHRLAGLLSRSGGMVKSNFVPTDVGSEAF
O. schmackeri OsCLP 1       GCDLSTCATHNLVNELNKFDKSKPSSGGVGPESP
O. schmackeri OsCLP 2       SCNLSTCATHNLVNELNKFDKSKPSSGGVGPESF
```

C

Adrenomedullins

```
Zebrafish*                  KRSKNSINQSR---RSGCSLGTCTVHVLAHRLHDLNNKLKIGNAPVDKINPYGY
Xenopus**                   YRHTFHHLQSV---RVGCRFGTCTVQNLAHQIFQYTDKDKDSTAPVNKISSQGY
Lizard***                           LKSLGCSLGTCSLQKLRAMQIQVYQIQSKTKLHFNSR
Chicken 4*                  YRQSVNSFPHLPTFRMGCRFGTCTVQKLAHQLYQLTDKVKDGAAPVNKISPQGY
Ox (O62827)                 YRQSLNNFQGL--RSFGCRFGTCTVQKLAHQIYHFTDKDKDGSAPRSKISPQGY
Rat (P43145)                YRQSMN--QGS--RSTGCRFGTCTMQKLAHQIYQFTDKDKDGMAPRNKISPQGY
Pig (A5LHG2)                YRQSMNNFQGL--RSFGCRFGTCTVQKLAHQIYQFTDKDKDGVAPRSKISPQGY
Human (E9PL83)              YRQSMNNFQGL--RSFGCRFGTCTVQKLAHQIYQFTDKDKDNVAPRSKISPQGY
Horse (Q9N0F0)              YRQSMNNFQGL--RSFGCRFGTCTVQKLAHQIYQFTDKDKDGVAPRSKISPQGY
Dog (O77559                 YRQSMNNFQGP--RSFGCRFGTCTVQKLAHQIYQFTDKDKDGVAPRSKISPQGY
Rabbit 5*                   YRQSMKNFQGS--RSFGCRFGTCTVQNLAHQIYQFTDKDKDDTAPRNKISPQGY
O. schmackeri OsCLP 1                   GCDLSTCATHNLVNELNKFDKSKPSSGGVGPESP
O. schmackeri OsCLP 2                   SCNLSTCATHNLVNELNKFDKSKPSSGGVGPESF
```

Protein sequences are obtained from the UNIPROT database. The other gene fragments were obtained from NCBI 'tblastn' using a human CT, CGRP or AM peptide sequence against 'reference genomic sequences' and selecting the organisms listed above as reference. The respective NCBI accession references are *Zebrafish: CT NW_003334572.1: 5336105 until 5336010, CGRP from 5335378 until 5335268 and AM NW_003039701.2: 180936 until 180784; **Xenopus: CT NW_003164150.1: 317312 until 317407, CGRP from 319334 until 319444 and AM NW_003163561.1:487390 until 487542; ***Lizard: CT NW_003338554.1: 106604 until 106699, CGRP from 109059 until 109169 and AM NW_003338578.1: 3644722 until 3644832; 4*Chicken: NW_003763785.1: 2625206 until 2625045; and 5*Rabbit:NW_003159229.1: 69231932 until 69232087.

Table 2: Alignment of Adrenomedullin (AM) fragment, translated from cDNA molecular cloned, from salivary glands of Ornithodoros parkeri *(NCBI accession gb|EF633931.1) and CTLP 1 and 2 from* O. schmackeri.

Ornithodoros parkeri	GCSLSTCVLQKLSDKLNHFTDDSKNKSGGTGPDSY
O. schmackeri CTLP 1	GCDLSTCATHNLVNELNKFDKSKPSSGGVGPESP
O. schmackeri CTLP 2	SCNLSTCATHNLVNELNKFDKSKPSSGGVGPESF

For frogs, apart from the *Xenopus* sequence, two CTs and a CGRPs have been described at the protein level; CGRP from *Rana ridibunda* isolated from brain and intestine [8], another CGRP isolated from the skin exudate of *Phyllomedusa bicolor* [10], and CT isolated from *Rana catesbiana* ultimobranchial glands[11]. All these three peptides show small sequence variations, but they all fit within the CT/CGRP family (Table 1 A and B). Alignment of the OsCLP 1 and 2 found in these studies all show some similarity, but clearly not enough overlap to classify them as either the *O. schmackeri* CT, CGRP or AM. For example, the total number of residues (34) in both OsCLTPs differs from that of the typical calcitonin gene family members: CT (32 residues), CGRP (37) or AM (~52 residues). The OsCLPs have a disulfide bridge between Cys-2 and Cys-7, which they share with CGRPs, whereas OsCLP-1 has a C-terminal Proline-amide, which fits more with CTs.

Since these new peptides do not fit with other animal CT, CGRP or AM, together with their skin origin, we assume they are not the typical *O. schmackeri* brain-gut CALC gene products. In other words, we find it more likely that these peptides represent specific frog skin translated defense peptides. About the function of these novel peptides in the skin of *O. schmackeri*, we can only speculate. As vasodilatory peptides comprise part of the defense strategy of frogs against predators (e.g. bradykinins, bradykinin potentiating peptides[12-14]), it is likely that the OsCLPs would be 'useful' to *O. schmackeri* in this way. Of the many bioactivities ascribed to peptides belonging to the calcitonin gene-related peptide superfamily, particularly the dramatic effects of CGRP on vasodilation (the most potent so far described) and heart beat regulation[15-17] as well as the vasodilatory effect of AM[18-20] would point in that direction.

5.4 Conclusion

Here we report on the discovery of two novel 34 residue peptides from the skin secretion of the Chinese odorous frog *Odorrana schmackeri*. They were fully *de novo* sequenced using a combination of different chemical and analytical tools: the usage of two alkylation reagents combined with CID and ETD tandem MS fragmentation, trypsin digestion; dimethyl labeling, HPLC-based peptide purification and Edman degradation. CID and ETD spectra are known to be provide complementary information for the *de novo* sequencing of peptides [5], as illustrated in this chapter (Figs. 3-4). Cysteine alkylation with bromoethylamine promoted a slightly higher charging of the peptide ion (data not shown), which turned out to be beneficial for improved ETD fragmentation [21-24]. Higher intensity of b-/y-ions (CID) aminoterminally from Pro-residues, and the absence of this cleavage in ETD were additional mass spectral features that assisted with sequence assignment in case of proline residues (see Figs. 3). The identification of one Lys-residues in the middle of the initially obtained long sequence-tag, prompted us to attempt to trypsinize the peptide and perform additional CID fragmentations on the smaller proteolytic fragments to get full sequence coverage. High resolution MS (Orbitrap) data helped to accurately calculate and predict the possible solutions to fill 'gaps' of 2-3 residues in incomplete b/y-ion series. Labeling with heavy and light dimethyl also yielded small but valuable clues in this case. By comparing heavy and light labeled spectra, y and b-ions were easily distinguished and consequently assigned (Fig. 6). Finally, the entire amino acid sequences was confirmed for OSCLP-2 by Edman degradation

Follow-up research has been initiated to elucidate the OsCLP mRNA sequences from an *O. schmackeri* skin secretion cDNA library constructed from reverse transcribed poly-adenylated mRNAs, using a primer derived from the sequences established.

In addition both OsCLP peptides are synthesized so that they can be tested for bioactivity. These could include smooth muscle myotropic assays as well as bone calcification assays.

Acknowledgements

This study was partly funded by the Netherlands Genomics Initiative (NGI) and the Brazilian National Council of Technological and Scientific Development (CNPq – grant GDE-200847/2007-04).

References

1. Wang, L.; Evaristo, G.; Zhou, M.; Pinkse, M.; Wang, M.; Xu, Y.; Jiang, X.; Chen, T.; Rao, P.; Verhaert, P.; Shaw, C., Nigrocin-2 peptides from Chinese Odorrana frogs--integration of UPLC/MS/MS with molecular cloning in amphibian skin peptidome analysis. *Febs J* **2010,** 277, (6), 1519-31.
2. Evaristo, G., Pinkse, M., Pimenta, D. C., Chen, T., Shaw, C., & Verhaert, P., PTM-driven Differential Peptide Display: Elucidation of Inter/Intra-molecular Disulfide Bridges in Frog Peptides. *Ook Press Ltd* **2010**, 579-581.
3. Boersema, P. J.; Aye, T. T.; van Veen, T. A.; Heck, A. J.; Mohammed, S., Triplex protein quantification based on stable isotope labeling by peptide dimethylation applied to cell and tissue lysates. *Proteomics* **2008,** 8, (22), 4624-32.
4. Savitski, M. M.; Nielsen, M. L.; Zubarev, R. A., Side-chain losses in electron capture dissociation to improve peptide identification. *Anal Chem* **2007,** 79, (6), 2296-302.
5. Zubarev, R. A.; Zubarev, A. R.; Savitski, M. M., Electron capture/transfer versus collisionally activated/induced dissociations: solo or duet? *J Am Soc Mass Spectrom* **2008,** 19, (6), 753-61.
6. Amara, S. G.; Jonas, V.; Rosenfeld, M. G.; Ong, E. S.; Evans, R. M., Alternative RNA processing in calcitonin gene expression generates mRNAs encoding different polypeptide products. *Nature* **1982,** 298, (5871), 240-4.
7. Rosenfeld, M. G.; Mermod, J. J.; Amara, S. G.; Swanson, L. W.; Sawchenko, P. E.; Rivier, J.; Vale, W. W.; Evans, R. M., Production of a novel neuropeptide encoded by the calcitonin gene via tissue-specific RNA processing. *Nature* **1983,** 304, (5922), 129-35.
8. Conlon, J. M.; Tonon, M. C.; Vaudry, H., Isolation and structural characterization of calcitonin gene-related peptide from the brain and intestine of the frog, Rana ridibunda. *Peptides* **1993,** 14, (3), 581-6.
9. Mol, J. A.; Kwant, M. M.; Arnold, I. C.; Hazewinkel, H. A., Elucidation of the sequence of canine (pro)-calcitonin. A molecular biological and protein chemical approach. *Regul Pept* **1991,** 35, (3), 189-95.
10. Seon, A. A.; Pierre, T. N.; Redeker, V.; Lacombe, C.; Delfour, A.; Nicolas, P.; Amiche, M., Isolation, structure, synthesis, and activity of a new member of the calcitonin gene-related peptide family from frog skin and molecular cloning of its precursor. *J Biol Chem* **2000,** 275, (8), 5934-40.
11. Yoshida, A.; Kaiya, H.; Takei, Y.; Watanabe, T. X.; Nakajima, K.; Suzuki, N.; Sasayama, Y., Primary structure and bioactivity of bullfrog calcitonin. *Gen Comp Endocrinol* **1997,** 107, (2), 147-52.
12. Chen, X.; Wang, L.; Wang, H.; Chen, H.; Zhou, M.; Chen, T.; Shaw, C., A fish bradykinin (Arg0, Trp5, Leu8-bradykinin) from the defensive skin secretion of the European edible frog, Pelophylax kl. esculentus: structural characterization; molecular cloning of skin kininogen cDNA and pharmacological effects on mammalian smooth muscle. *Peptides* **2011,** 32, (1), 26-30.
13. Wang, L.; Chen, Y.; Yang, M.; Zhou, M.; Chen, T.; Sui, D. Y.; Shaw, C., Peptide DV-28 amide: An inhibitor of bradykinin-induced arterial smooth muscle relaxation encoded by Bombina orientalis skin kininogen-2. *Peptides* **2010,** 31, (5), 979-82.
14. Zhou, X.; Wang, L.; Zhou, M.; Chen, T.; Ding, A.; Rao, P.; Walker, B.; Shaw, C., Amolopkinins W1 and W2--novel bradykinin-related peptides (BRPs) from the skin of the Chinese torrent frog, Amolops wuyiensis: antagonists of bradykinin-induced smooth muscle contraction of the rat ileum. *Peptides* **2009,** 30, (5), 893-900.
15. Ghatta, S. a. N., D., Calcitonin gene-related peptide: Understanding its role. *Indian J Pharmacol* **2004,** 36, (5), 277-283.

16. Recober, A.; Russo, A. F., Calcitonin gene-related peptide: an update on the biology. *Curr Opin Neurol* **2009,** 22, (3), 241-6.
17. Smillie, S. J.; Brain, S. D., Calcitonin gene-related peptide (CGRP) and its role in hypertension. *Neuropeptides* **2011,** 45, (2), 93-104.
18. Hinson, J. P.; Kapas, S.; Smith, D. M., Adrenomedullin, a multifunctional regulatory peptide. *Endocr Rev* **2000,** 21, (2), 138-67.
19. Cai, D. Y.; Yu, F.; Jiang, W.; Jiang, H. F.; Pan, C. S.; Qi, Y. F.; Chang, L.; Zhao, J.; Yang, J. H.; Zhu, M. J.; Jia, Y. X.; Geng, B.; Ma, T. M.; Pang, Y. Z.; Tang, C. S., Adrenomedullin(27-52) inhibits vascular calcification in rats. *Regul Pept* **2005,** 129, (1-3), 125-32.
20. Takao, M.; Tomita, M.; Tanahashi, N.; Kobari, M.; Fukuuchi, Y., Transient vasodilatory effects of adrenomedullin on cerebral parenchymal microvessels in cats. *Neurosci Lett* **1999,** 268, (3), 147-50.
21. Xia, Y.; Gunawardena, H. P.; Erickson, D. E.; McLuckey, S. A., Effects of cation charge-site identity and position on electron-transfer dissociation of polypeptide cations. *J Am Chem Soc* **2007,** 129, (40), 12232-43.
22. Campbell, J. L.; Hager, J. W.; Le Blanc, J. C., On performing simultaneous electron transfer dissociation and collision-induced dissociation on multiply protonated peptides in a linear ion trap. *J Am Soc Mass Spectrom* **2009,** 20, (9), 1672-83.
23. Frese, C. K.; Altelaar, A. F.; Hennrich, M. L.; Nolting, D.; Zeller, M.; Griep-Raming, J.; Heck, A. J.; Mohammed, S., Improved peptide identification by targeted fragmentation using CID, HCD and ETD on an LTQ-Orbitrap Velos. *J Proteome Res* **2011,** 10, (5), 2377-88.
24. Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W., Electron capture dissociation for structural characterization of multiply charged protein cations. *Anal Chem* **2000,** 72, (3), 563-73.

Chapter 6

PTM-driven differential peptide display: survey of peptides containing inter/intra-molecular disulfide bridges in frog venoms

Abstract

Amphibian defensive skin secretions are complex species-specific mixtures of biologically active molecules, including many uncharacterized peptides. Many of these peptides are post-translationally modified and amongst the modifications discovered so far on amphibian defense peptides, disulfide bonds are quite frequently encountered. The presence of this PTM often complicates the MS-based sequencing and reduction of the S-S bonds is thus required. Here we demonstrate a method to target peptides containing specific PTMs like cystines (disulfide bridged cysteines) applying a PTM-driven differential display. Upon reduction of the disulfide bond both the molecular mass and the HPLC retention time of a peptide are altered. Assembling the LC-MS data by orthogonal plotting of the *m/z* data against the retention time generates a peptide display and overlaying peptide displays of untreated and DTT-reduced material yields a differential display. From such a differential display, peptides originally carrying a disulfide bond are easily distinguished from the others due to the shift in both retention time and *m/z* values, whereas non cystine containing peptides remain unaltered in the differential display. The success of this approach is demonstrated by the visualization of the cystines-containing peptides in the skin secretion of *Odorrana schmackeri, Phyllomedusa burmeisteri, Phyllomedusa rohdei, Kassina senegalensis,* and *Bombina variegata*. The venoms from these different species yield complicated displays, showing interesting peptide features, allowing one to target them specifically for more detailed structural characterization.

This chapter is published in the Journal of Proteomics as:

Evaristo, G. P. C.; Verhaert, P. D. E. M.; Pinkse, M. W. H., PTM-driven differential peptide display: Survey of peptides containing inter/intra-molecular disulfide bridges in frog venoms, *J Prot* **2012**. (http://dx.doi.org/10.1016/j.jprot.2012.09.001)

6.1 Introduction

Amphibian skin secretions are a rich source of biologically active peptides and, so far, hundreds of peptides have been structurally characterized. Several methods and approaches are at hand to identify and isolate these bioactive compounds from the amphibian venom. Often they are first discovered on the basis of biological activity (physiological tests or bioassays). Subsequently, molecular characterization frequently involves structure determination via amino acid analysis, Edman degradation or via tandem mass spectrometry[1, 2]. In the last two decades molecular cloning of the complementary DNA encoding the precursor of amphibian peptides has often been the preferred approach to unravel the primary structure of peptides in frog skin secretions[3]. Also, mass spectrometry is becoming more and more a popular tool for the characterization of this class of peptides [4-6]. It should be noted however that the peptidomic analysis of the frog skin secretions is in general a complex endeavour because the frog venom contains tens to hundreds of different peptide sequences, varying both in size, concentration and post-translational modifications (PTMs). Due to the many PTMs and the general absence of genomic information, the MS-based sequencing of these peptides is often very challenging. In some instances it is possible to fully *de novo* sequence the primary structure of a peptide, but in most cases only a partial sequence tag is obtained. The true strength of the MS-based approach is that it allows one to visualize the entire amphibian skin peptidome after one single LC-MS analysis. In addition MS also makes it possible to first detect those peptides carrying a specific PTM, before attempts are made to elucidate their primary structure by tandem MS. In proteomics analysis this is frequently done using specific PTM isolation or detection strategies[7]. Amongst the PTMs found in frog peptides the disulfide bond is one of the most common. Intra-molecular disulfide bonds induce a specific (relatively rigid) conformation to a peptide, often essential for its biological activity[8]. Several antimicrobial peptides from the genus *Rana*, are known to have a conserved motif with a disulfide bond at the carboxyl terminus, spaced by 5-8 other amino acids, called "Rana box"[9, 10]. In other frog species different intra-molecular cystine motifs are found in the peptides secreted by skin, such as in serine protease inhibitors, Kazal, Kunitz, Bowman-Birk-like inhibitors[11-17], and calcitonin gene related peptide[18]. Moreover, an inter-molecular disulfide bond forming the distinctin heterodimer was found in the skin secretion of *Phyllomedusa distincta*[19].

In this Chapter we describe a straightforward method that allows for immediate visualization of peptides carrying disulfide bridges from a two-dimensional peptide display. Skin secretions of 5 different anurans were subjected to this method: *Odorrana schmackeri*, a Chinese Ranid frog with many Rana-box containing peptides; the African *Kassina senegalensis;* the European toad *Bombina variegata;* and two South-American tree frogs *Phyllomedusa burmeisteri* and *Phyllomedusa rohdei.* These last two belong to a genus known for its highly bioactive venom and yet there is a lack of knowledge about cystine containing peptides from these species in the literature[19-22].

6.2 Materials and Methods

6.2.1 Frog secretions collection

Venoms of the Asian odorous frog, *Odorrana schmackeri*; two South-American tree-frogs, *Phyllomedusa burmeisteri* and *Phyllomedusa rohdei;* the African frog, *Kassina senegalensis;* and the European toad, *Bombina variegata*, were used in this study. All specimens were sampled in their natural habitats and immediately released after venom collection. To access the genetic heritage information of the Brazilian anurans, an official license was obtained from the IBAMA - Brazilian Institute of Environment and Renewable Natural Resources - number 010453/2010-5.

All frogs were adults of undetermined sex. The secretion harvesting was performed by gentle transdermal electrical stimulation as described previously[23]. Secretions were jet-washed from the frogs using de-ionized water, filtered through a membrane (0.45 μm), snap-frozen in liquid nitrogen, lyophilized, and stored at -20°C prior to analysis.

6.2.2 Samples preparation

One milligram of lyophilized skin secretion from each species of frogs were separately dissolved in 100 μL of 0.05/99.95 v/v trifluoroacetic acid/water and clarified of microparticles by centrifugation at 13000 g for 15 min. at 4°C. Each clear supernatant was diluted in ammonium bicarbonate buffer (25 mM, pH 8.0) in a final concentration of 2 mg/mL. This sample was splitted in two and one part remained without any further treatment (from here on

referred to as the *untreated* venom); and the other part was treated with 2 mM dithiothreitol (DTT) in order to reduce the S-S bonds (from here on referred to as the *DTT-reduced* venom).

6.2.3 PTM-driven differential peptide display

In order to visualize those amphibian defense peptides that contain one or more disulfide bridge(s), the following three main steps were followed to generate a 2-dimensional differential peptide display. First, 50 micrograms of untreated and DTT-reduced samples of the frog venom were individually separated by HPLC, using a Waters 2695 Alliance HPLC module equipped with a microbore C4 column (Dr. Maisch, Ammerbuch-Entringen, Germany; 5 μm particles, 2.0 mm ID x 150 mm length). The column was kept at 40°C and the flow rate was set to 0.2 mL/min., solvent A was 0.1 M acetic acid in MilliQ, solvent B was 0.1 M acetic acid in 8/2 v/v MeCN/water. Peptides were eluted with a linear gradient until 70% of solvent B in 70 minutes (1% B/min.), followed by 5 min. of 4% B/min. and 5 min. cleaning with 90% of B. The effluent of the column was directly sprayed into the ion source of a Q-TOF mass spectrometer (Waters QTof Premier). MS spectra were acquired in positive ion mode over the *m/z* range of 400-1500 Da in profile format (scan time 3 s, inter-scan time 0.1 s). The raw data of both DTT-reduced and untreated samples were converted into a text file using the DataBridge tool within the MassLynx 4.1 software. The files were directly imported into MSight (http://www.expasy.org/MSight). The contrast of the images was adjusted manually (with color inversion). Finally, the PTM-driven differential peptide displays were generated using the overlay option.

6.2.4 Peptide identifications

After interrogation of the images, peptides of interest were further analysis by targeted MS-MS analysis. Prior to these measurements, the crude venom was reduced and alkylated with iodoacetamide (IAM). Samples were first separated by nano-HPLC coupled either to a Waters QTOF Premier, a Thermo LTQ-Orbitrap XL equipped with ETD or a Thermo Orbitrap Velos equipped with HCD capabilities. For each peptide different fragmentation spectra were acquired using CID, HCD and ETD collision cells and selecting different charge states. Peptides were *de novo* sequenced by manual interpretations of the spectra. The obtained amino acid sequences were blasted against NCBI non redundant protein database selecting anura (tax ID # 8342) as restriction.

6.3 Results

Figure 1 displays the two ion chromatograms of the LC-MS analysis of the skin secretion of *O. schmackeri* that was untreated (Fig. 1A) and reduced with DTT (Fig. 1B). Direct comparison of these chromatograms shows the elution pattern is different after DTT reduction. In order to extract more information from the LC-MS data, two-dimensional displays were constructed. In such a display, the acquired *m/z* spectra are orthogonally plotted against the retention time and the ion intensity is displayed in a colour scale. Figure 2 shows the peptide displays of the two analyses. From these images it is evident that the LC-MS data of the skin secretion of *O. schmackeri* contains many ion signals. By constructing an overlay of the two peptide displays a differential ion map is generated that allows one to distinguish between changed and unchanged ion features. Figure 3 shows the differential display of the untreated and DTT-reduced skin secretion of *O. schmackeri.* The untreated sample is shown in blue and the DTT-reduced in red. If a peptide is unaffected by DTT, it has the same retention time and *m/z* value in both chromatograms. As a result, the blue and the red ion signals overlap, leading to a white ion feature in the overlay (Fig. 4A). If a peptide is reduced by DTT, its *m/z* value and its retention time are different between the two analyses [24]. In the overlay this is reflected by red and blue ion features that are separated with a small difference in retention time and *m/z* ratio. In figure 4B-D typical examples are shown of peptides with single or multiple S-S bonds that are detected in this way. The number of disulfide bonds can be deduced from total increase of molecular weight considering that the reduction of one single disulfide bond increases the molecular weight of a peptide with 2 amu.

Closer inspection of the differential display of *O. schmackeri* (Fig. 3) clearly shows that in this venom only a few peptides are unchanged (white spots, no S-S bonds) and that the majority of the observed peptides have single disulfide bonds. Several peptides from this venom have been previously characterized, including nigrocins, esculentins, brevinins and odorranains, all with a C-terminal Rana-box motif[9]. Individual members from these peptide families often have similar sequences varying in 1 or 2 amino acid residues. In the LC/MS analysis they are observed in the same *m/z* range and retention time. As mentioned, upon reduction of the S-S bond, retention time is shifted. For peptides of the same family, the direction of the shift is highly similar. The correlation between the number of S-S bonds, overall molecular weight and retention time shift aids in recognizing that certain ion features originates from peptides that belong to the same family. This is illustrated in Figure 5 for

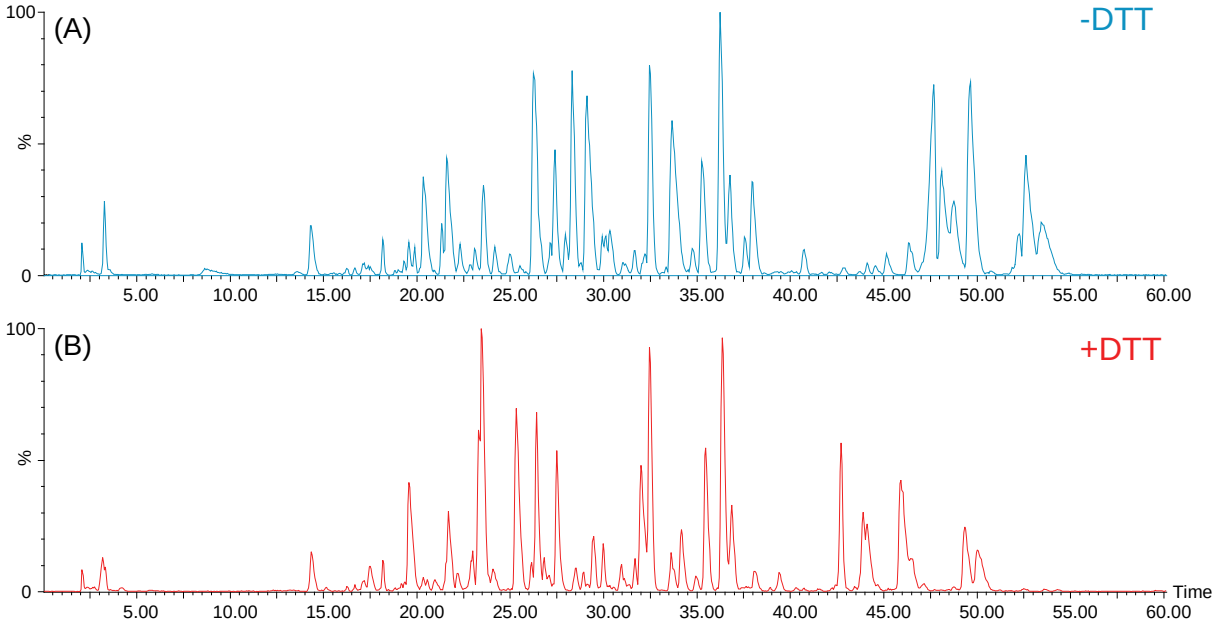


Figure 1: LC-MS chromatograms of untreated (in blue, A) and DTT-reduced (in red, B) O. schmackeri *venom.*

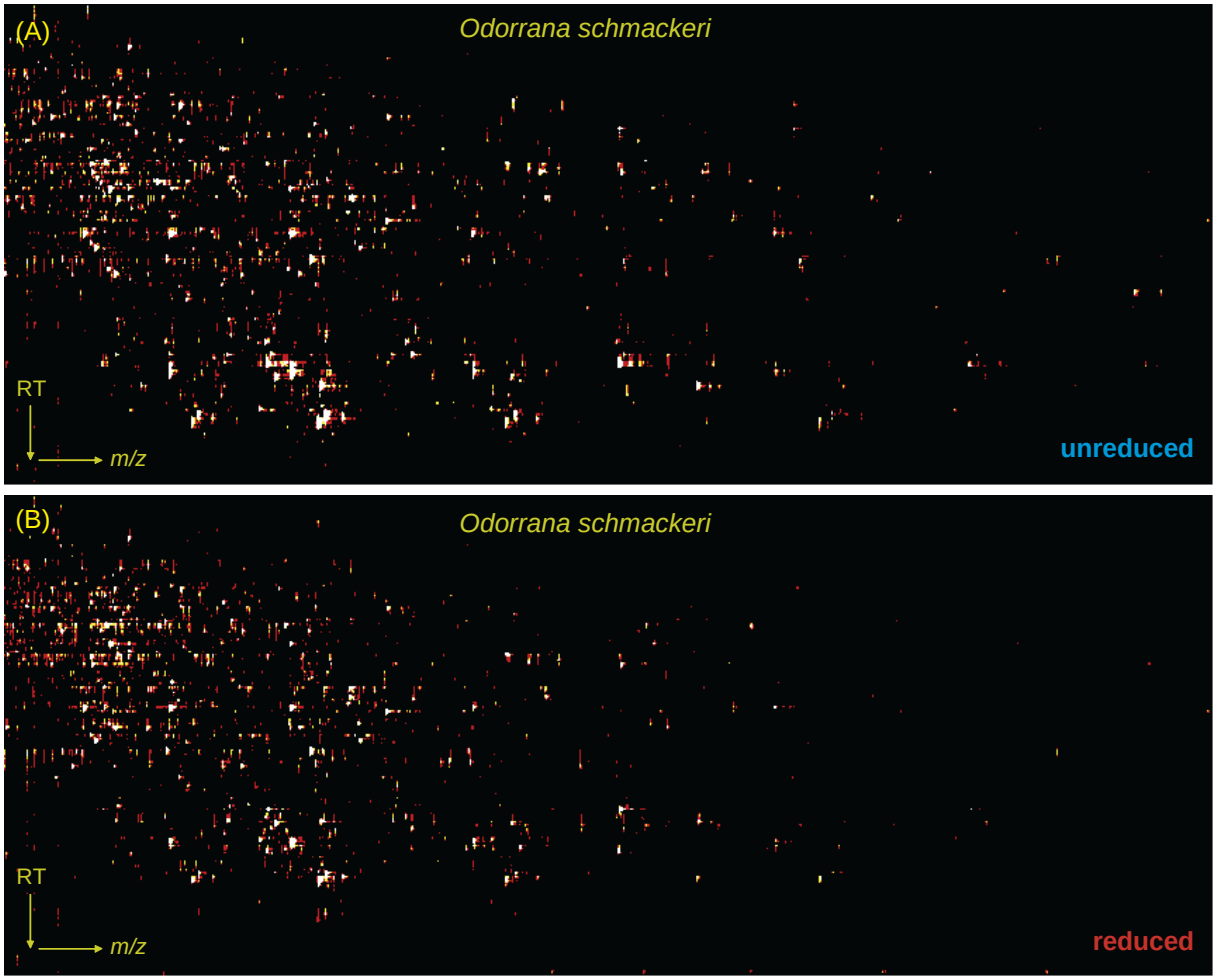


*Fig. 2: 2-D display of the LC/M chromatograms from Fig. 1 by plotting retention time (RT 10-60 min.) versus mass-to-charge ratio (*m/z *400-1500).*

members of the nigrocin family. These peptides have a molecular mass of ~2 kDa and are observed as doubly charged ions in between *m/z* 920 and 1080. Both N.2SCa, N.2SCb and N.2SCc [9] are highlighted in the display and the *m/z* and retention shift are indicated with an arrow. All have an identical PTM-driven shift in *m/z* and retention time in the differential display (Fig. 5).

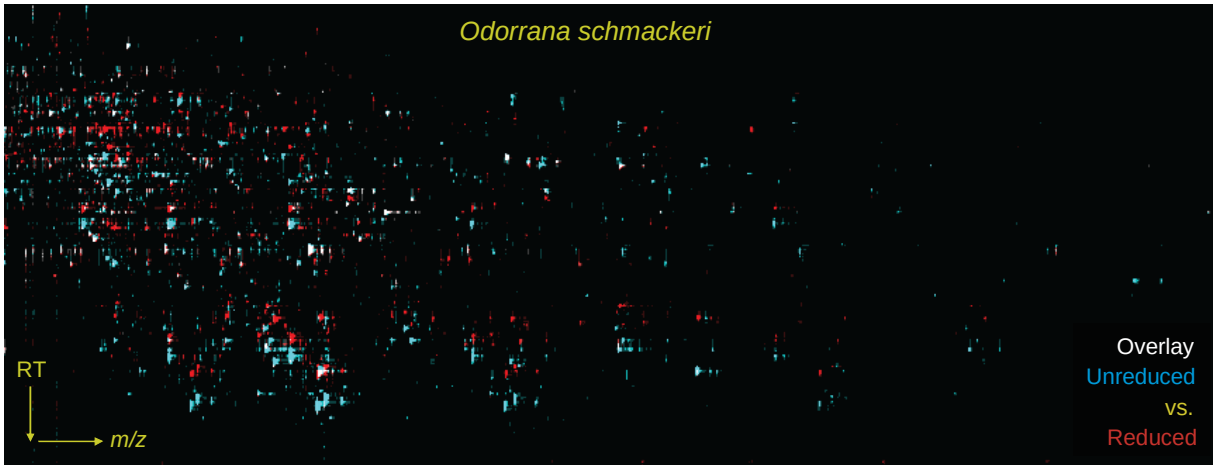


Fig. 3: PTM-driven differential peptide display generated by overlaying the 2-D displays from Fig. 2 (RT 10 - 60 min. and m/z *400 - 1500). In this image, white spots represent molecular ions that do not show any change in mass and retention time after DTT treatment, indicating they do not contain cystines. Blue and red spots represent peptide ions that do contain one or more disulfide bonds.*

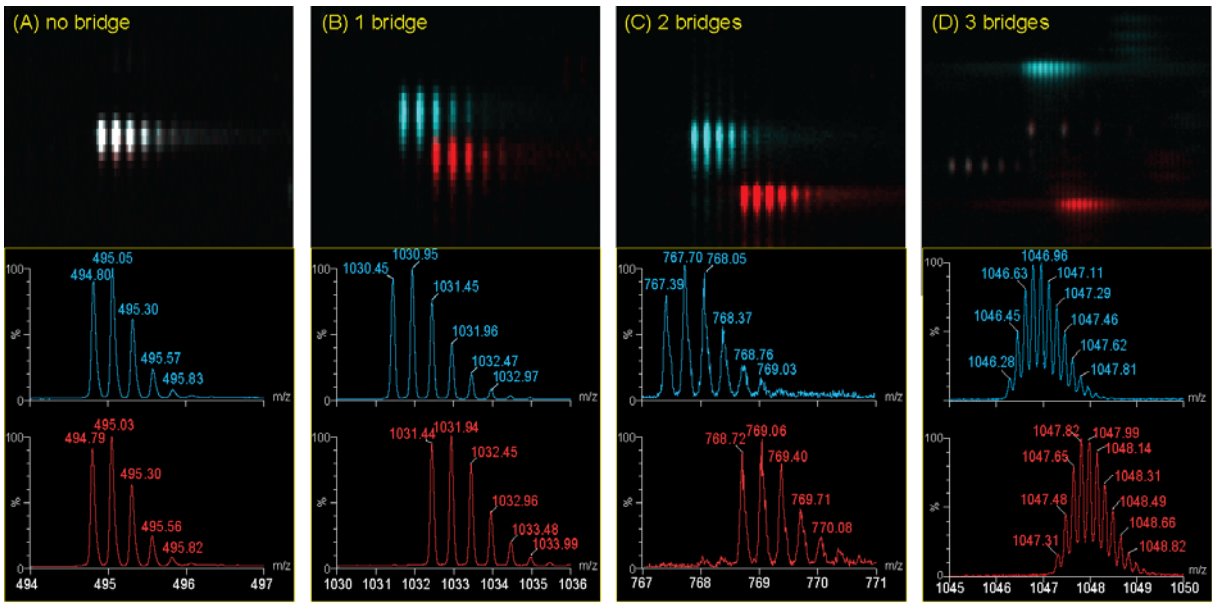


Fig. 4: Zoom-ins of the PTM-driven differential peptide displays depicting peptides without cystines (A) and containing 1 (B), 2 (C) and 3 (D) disulfide bridges, represented by the respective m/z *shift of 2, 4 and 6.*

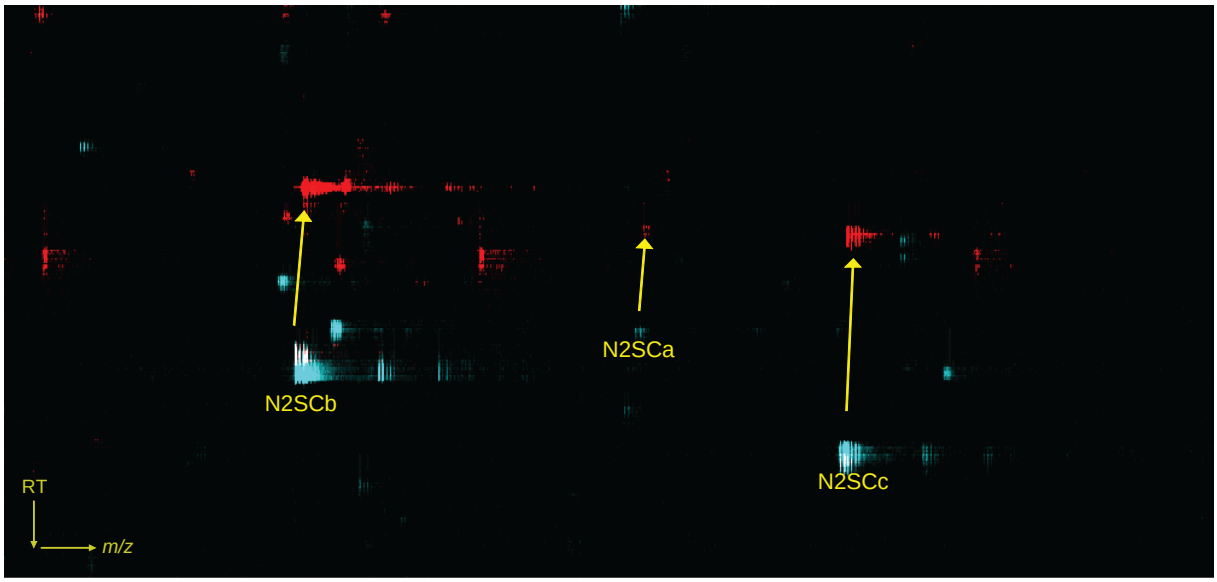
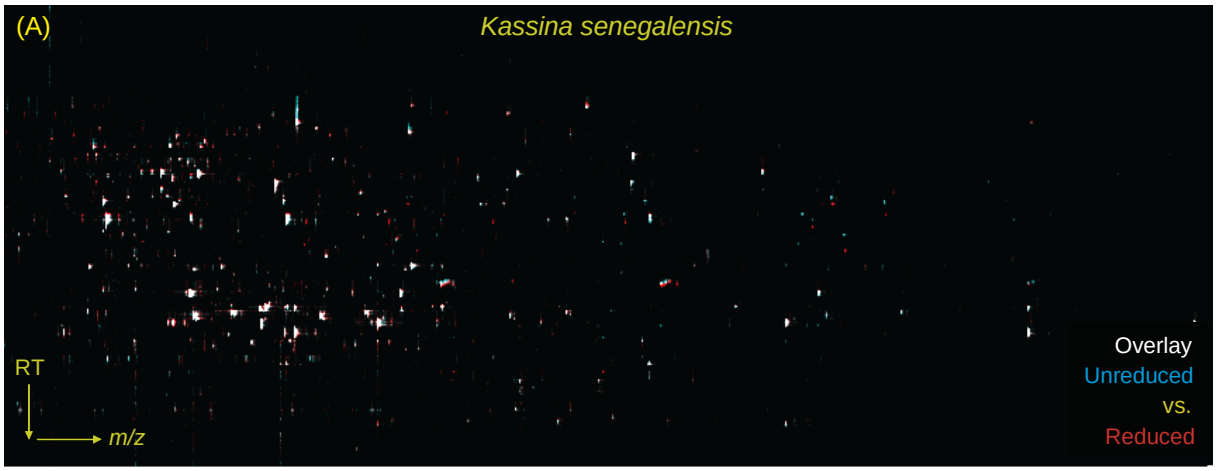


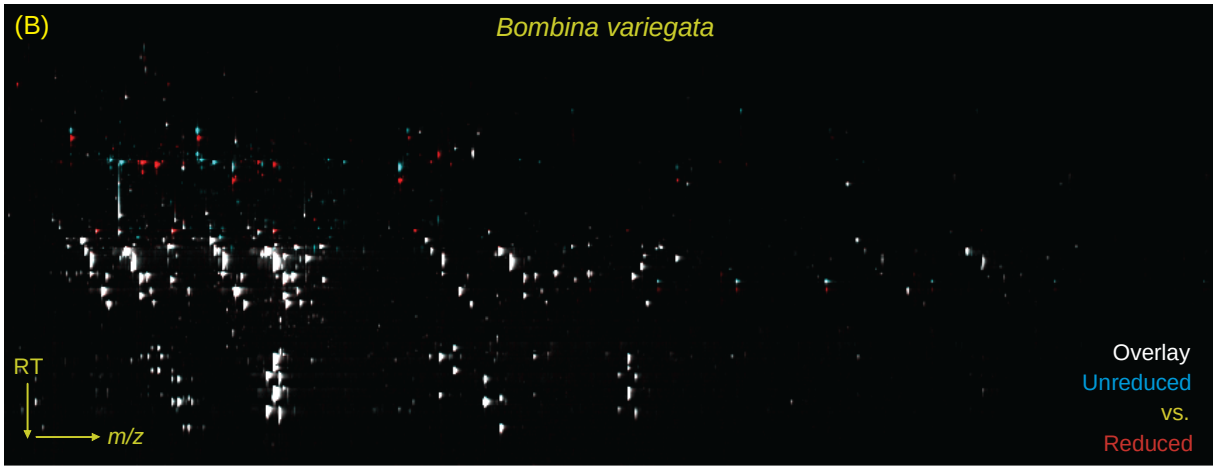
*Fig. 5: Enlarged part (*m/z *from 920 to 1080 and retention time from 38 to 53 min.) of the* O. schmackeri *differential display (from Fig. 3). This zoom in shows the doubly charged ions of nigrocins 2SCa, 2SCb and 2SCc untreated (in blue) and DTT-reduced (in red). All nigrocin peptides have a similar tendency to elute earlier upon reduction.*

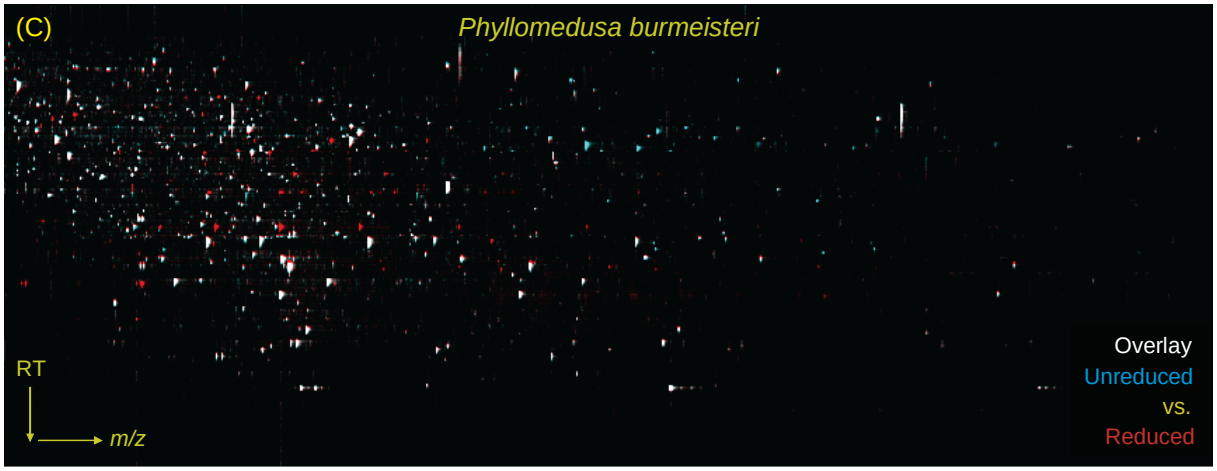
Figure 6 shows the PTM-driven differential peptide displays of the other amphibians analyzed in this study: *Phyllomedusa burmeisteri*, *Phyllomedusa rohdei*, *Kassina senegalensis* and *Bombina variegata*. From these images it is clear that each frog species has a unique set of peptides. The PTM-driven differential displays reveal the level of cystine containing peptide for each of the frog species. For example, from the display of *O. schmackeri* (Fig. 3) mainly contains blue/red ion-pairs and, as mentioned above, the majority originate from peptides with a single S-S bond. Differently, the displays of the other 4 frogs (Fig. 6 A-D) exhibit, on average, more white spots, indicating the majority of skin peptides of these species do not have a S-S bond. In the differential display of *K. senegalensis* (Fig 6A) several low abundant ion feature were observed that originate from peptides with either 2 or 3 disulfide bonds. Recently a peptide with a mass of 6.7 kDa was characterized as a Kunitz inhibitor (KSCI) [25]. In the display, this peptide was easily spotted due to its PTM-driven shift from the reduction of 3 disulfide bond, which is a conserved motif in Kunitz type inhibitors. In the differential display of *B. variegata* (Fig. 6B) blue and red ion features are observed in the upper part of the display. Some of these are due to multiply charge states of two peptides with a molecular mass of approximately 3.4 kDa. In *B. variegata*, the majority of cystine containing peptides have a single S-S bonds, but also a peptide with 2 disulfide bonds was observed. Fig 6C and 6D show the differential display of *P. burmeisteri* and *P. rohdei*. Species from the *Phyllomedusa* genus have long been known for their highly toxic and

complex secretions[26]. Various families of Phyllomedusidae peptides have been identified: phyllocaerulein, phyllokinin, phyllolitorin, phyllomedusin, phylloseptin, dermaseptin, dermorphin, tryptophillin and hyposin[5, 26, 27]. Interestingly, none of these contain disulfide bonds and this is reflected in the PTM-driven differential peptide display by the occurrence of many white ion features. The differential display of *P. burmeisteri* and *P. rohdei* slightly differ in the overall number of peptides ion signals observed, whereas for *P. rohdei* a smaller amount of ion signals are observed compared to *P. burmeisteri*. However, in the PTM-driven differential display of both *P. burmeisteri* and *P. rohdei* a series of red/blue ion pairs are observed. For all these peptides, the calculated mass differences correspond to the presence of 3 S-S bonds. By analyzing the images in more detail it turned out that approximately 20 peptides within the skin secretion of *P. burmeisteri*, and also in *P. rohdei* venom, have this characteristic pattern (i.e. molecular mass around 6 kDa and 3 internal S-S bonds (Fig. 7). As previously demonstrated for the nigrocin peptides in *O. schmackeri*, the identical behaviour in the PTM driven display, suggest that these peptide in *P. rohdei* and *P. burmeisteri* belong to the same family. Additionally, in both *P. burmeisteri* and *P. rohdei* a few peptides are observed that have single disulfide bonds.

Although inter-molecular disulfide bridges are quite rare for amphibian skin peptides, with the above described approach it is also possible to detect these from the differential displays. The recognition of inter-molecular disulfide bonds is less obvious than intra-molecular disulfide bonds, mainly due to the unpredictable difference in term of *m/z* shift. Using the PTM-driven display, we discovered dimeric peptides in the venom of *P. burmeisteri* that are linked via an intra-molecular S-S bridge (Chapter 7)[28]. An intact heterodimeric peptide with a total mass of 5449.10 Da, eluted in the untreated sample at 52.5 min. After reduction no ion signals of similar mass were observed close around this retention time. Instead, in the reduced sample ion signal for the monomers, with masses of 2499.42 and 2951.68 Da were observed at the retention times of 49.5 and 48.5 min. (Fig. 8). The masses of both monomeric peptides, minus 2 amu for a disulfide bond, are in agreement with the total mass of the large peptide detected in the untreated sample. Targeted MS/MS analysis of the all peptide ions confirmed the presence of the heterodimer and that this peptide showed high sequence similarity with distinctin, a heterodimeric peptide previously discovered in the skin secretion of *P. distincta*[19].


(A)
Kassina senegalensis
RT
m/z
Overlay
Unreduced
vs.
Reduced


(B)
Bombina variegata
RT
m/z
Overlay
Unreduced
vs.
Reduced


(C)
Phyllomedusa burmeisteri
RT
m/z
Overlay
Unreduced
vs.
Reduced

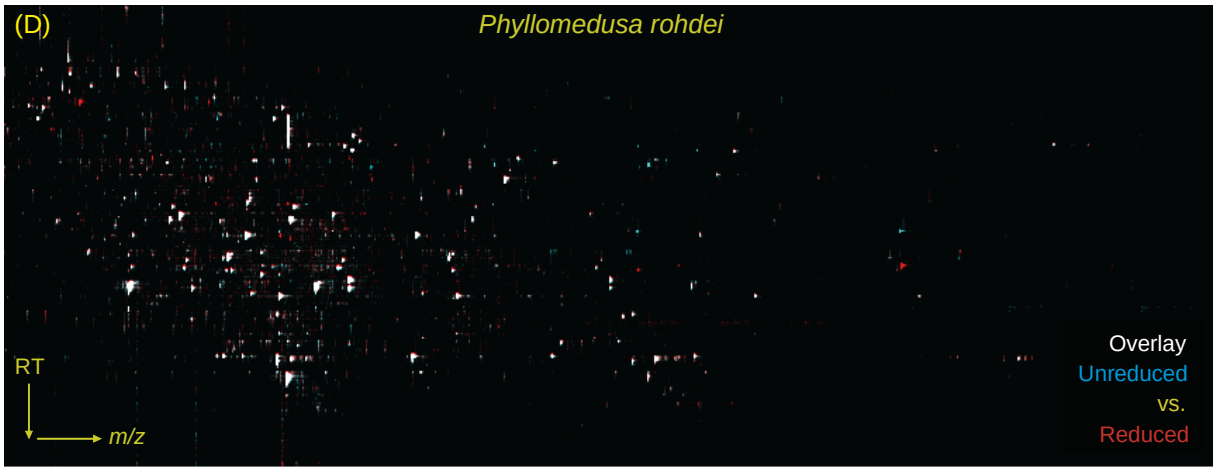


Fig. 6: PTM-driven differential peptide displays of (A) Kassina senegalensis, *(B)* Bombina variegata, *(C)* Phyllomedusa burmeisteri *and (D)* Phyllomedusa rohdei. *Retention time is from 10 to 60 min. and* m/z *range from 400 to 1500.*

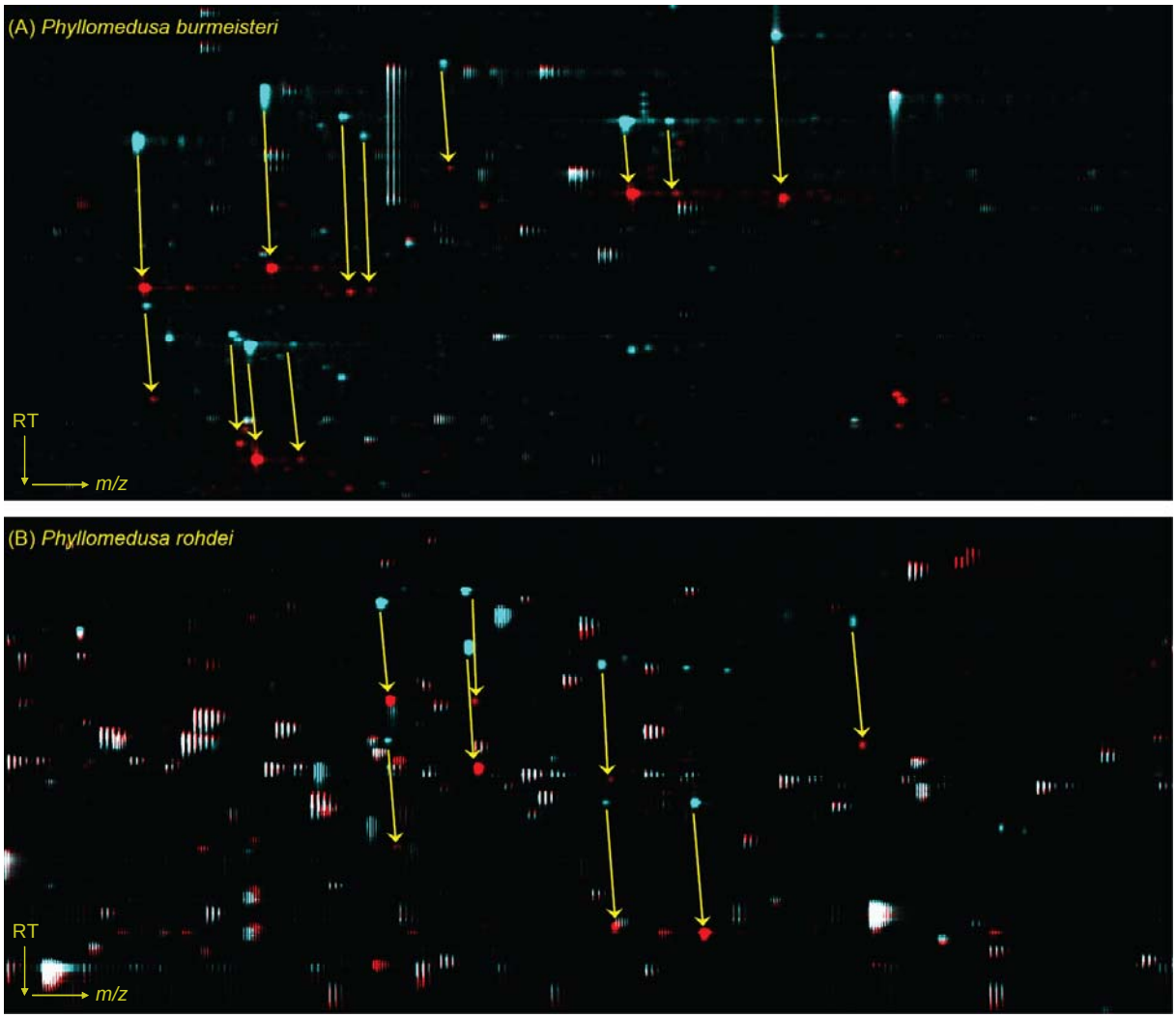


Fig. 7: Enlarged part of the PTM-driven differential display of (A) P. burmeisteri *and (B)* P. rohdei *highlighting several large peptides that are affected by DTT. All peptides have a molecular mass of ~6kDa, a similar shift in retention time and 3 S-S bonds, indicating they belong to one peptide family.*

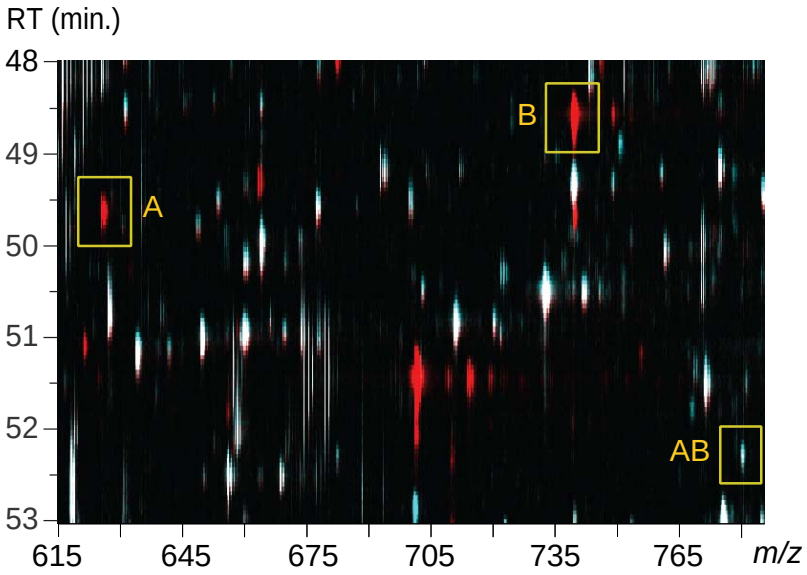


Figure 8: Enlarged part of the differential display of P. burmeisteri *(Fig. 5A) illustrating a peptide with an inter-molecular disulfide bond. A and B are, respectively, the* $[M+4H]^{4+}$ *of distinctin chains A and B, DTT-reduced (in red), and AB is* $[M+7H]^{7+}$ *distinctin heterodimeric peptide untreated (in blue) in crude sample. Mass sum of A and B minus 2H is exact the mass of AB.*

6.4 Discussion

This Chapter describes a method to directly detect peptides containing a disulfide bridge in complex samples such as frog venoms. This approach is rapid and efficient and applicable to any kind of venom/mixture of biological peptides. As the retention of a peptide on a reversed phase column depends both on its hydrophobicity and its secondary structure, reduction of a S-S bond does not only increase the *m/z* ratio, in most cases the peptides will also have a different elution time. By acquiring 2 LC-MS runs of an untreated and treated (DTT-reduced) sample of the venom, a two dimensional peptide display can be constructed. Dithiotreitol was chosen as reducing agent due to its strong action, making sure that all the peptides containing cystines are reduced, which is a requirement in this approach. Previous experiments using TCEP showed that it was less efficient in reducing all cystines (data not shown), hence DTT is preferred. From the differential peptide display both the complexity of the venom in term of total number of peptides as well as the number of peptides containing cystines can be deduced. This is illustrated for the venom of *O. schmackeri*, in which many single S-S bond containing peptides were observed. *P. burmeisteri* and *P. rohdei* showed

several large peptides containing 3 S-S bonds. *K. senegalensis* showed less peptide content and some peptides with 1 S-S bonds. In *B. variegata* venom several peptides with single or double S-S bonds were detected. Although less obvious to detect by this method, inter-molecular S-S bonds can also be found by this approach, as illustrated in the venom of *P. burmeisteri*. For the peptidome analysis of amphibian skin secretion in general, this strategy provides very useful information and it assists in the selection of further suitable treatments and MS/MS analysis.

6.5 Conclusion

An efficient and quick way to visualize the entire content of a frog skin gland secretion is to construct a peptide display of the acquired LC-MS data. In a next step, a differential 2-dimensional display of DTT-reduced and untreated samples allows one to detect peptides with disulfide bridges. Using this approach we showed that several peptides with inter/intra-molecular disulfide bridges could be recognized in highly complex samples prior to sequencing their primary amino acid structures. This is a powerful method considering that this principle could also be applied to monitor other PTMs and/or in different kind of samples.

Acknowledgements

We thank Prof. Dr. Chris Shaw and Dr. Tianbao Chen for providing *Odorrana schmackeri, Bombina variegata* and *Kassina senegalensis* venom. This project is funded by CNPq-Brazil (Geisa P. C. Evaristo PhD grant 200847/2007-4) and the Netherlands Genomics Initiative (NGI).

References

1. Montecucchi, P. C.; de Castiglione, R.; Piani, S.; Gozzini, L.; Erspamer, V., Amino acid composition and sequence of dermorphin, a novel opiate-like peptide from the skin of Phyllomedusa sauvagei. *Int J Pept Protein Res* **1981,** 17, (3), 275-83.
2. Montecucchi, P. C., Isolation and primary structure determination of amphibian skin tryptophyllins. *Peptides* **1985,** 6 Suppl 3, 187-95.
3. Zhou, M.; Wang, L.; Owens, D. E.; Chen, T.; Walker, B.; Shaw, C., Rapid identification of precursor cDNAs encoding five structural classes of antimicrobial peptides from pickerel frog (Rana palustris) skin secretion by single step "shotgun" cloning. *Peptides* **2007,** 28, (8), 1605-10.
4. Bai, B.; Zhang, Y.; Wang, H.; Zhou, M.; Yu, Y.; Ding, S.; Chen, T.; Wang, L.; Shaw, C., Parallel Peptidome and Transcriptome Analyses of Amphibian Skin Secretions Using Archived Frozen Acid-Solvated Samples. *Mol Biotechnol* **2012**.
5. Thompson, A. H.; Bjourson, A. J.; Orr, D. F.; Shaw, C.; McClean, S., Amphibian skin secretomics: application of parallel quadrupole time-of-flight mass spectrometry and peptide precursor cDNA cloning to rapidly characterize the skin secretory peptidome of Phyllomedusa hypochondrialis azurea: discovery of a novel peptide family, the hyposins. *J Proteome Res* **2007,** 6, (9), 3604-13.
6. Samgina, T. Y.; Gorshkov, V. A.; Vorontsov, Y. A.; Artemenko, K. A.; Zubarev, R. A.; Lebedev, A. T., Mass spectrometric study of bradykinin-related peptides (BRPs) from the skin secretion of Russian ranid frogs. *Rapid Commun Mass Spectrom* **2011,** 25, (7), 933-40.
7. Zhao, Y.; Jensen, O. N., Modification-specific proteomics: strategies for characterization of post-translational modifications using enrichment techniques. *Proteomics* **2009,** 9, (20), 4632-41.
8. Zasloff, M., Antimicrobial peptides of multicellular organisms. *Nature* **2002,** 415, (6870), 389-95.
9. Wang, L.; Evaristo, G.; Zhou, M.; Pinkse, M.; Wang, M.; Xu, Y.; Jiang, X.; Chen, T.; Rao, P.; Verhaert, P.; Shaw, C., Nigrocin-2 peptides from Chinese Odorrana frogs--integration of UPLC/MS/MS with molecular cloning in amphibian skin peptidome analysis. *Febs J* **2010,** 277, (6), 1519-31.
10. Simmaco, M.; Mignogna, G.; Barra, D., Antimicrobial peptides from amphibian skin: what do they tell us? *Biopolymers* **1998,** 47, (6), 435-50.
11. Conlon, J. M.; Kim, J. B., A protease inhibitor of the Kunitz family from skin secretions of the tomato frog, Dyscophus guineti (Microhylidae). *Biochem Biophys Res Commun* **2000,** 279, (3), 961-4.
12. Wu, J.; Liu, H.; Yang, H.; Yu, H.; You, D.; Ma, Y.; Ye, H.; Lai, R., Proteomic analysis of skin defensive factors of tree frog Hyla simplex. *J Proteome Res* **2011,** 10, (9), 4230-40.
13. Mignogna, G.; Pascarella, S.; Wechselberger, C.; Hinterleitner, C.; Mollay, C.; Amiconi, G.; Barra, D.; Kreil, G., BSTI, a trypsin inhibitor from skin secretions of Bombina bombina related to protease inhibitors of nematodes. *Protein Sci* **1996,** 5, (2), 357-62.
14. Ali, M. F.; Lips, K. R.; Knoop, F. C.; Fritzsch, B.; Miller, C.; Conlon, J. M., Antimicrobial peptides and protease inhibitors in the skin secretions of the crawfish frog, Rana areolata. *Biochim Biophys Acta* **2002,** 1601, (1), 55-63.
15. Zhao, Y.; Jin, Y.; Wei, S. S.; Lee, W. H.; Zhang, Y., Purification and characterization of an irreversible serine protease inhibitor from skin secretions of Bufo andrewsi. *Toxicon* **2005,** 46, (6), 635-40.

16. Zhang, Y.; Wang, M.; Wei, S., Isolation and characterization of a trypsin inhibitor from the skin secretions of Kaloula pulchra hainana. *Toxicon* **2010,** 56, (4), 502-7.
17. Gebhard, L. G.; Carrizo, F. U.; Stern, A. L.; Burgardt, N. I.; Faivovich, J.; Lavilla, E.; Ermacora, M. R., A Kazal prolyl endopeptidase inhibitor isolated from the skin of Phyllomedusa sauvagii. *Eur J Biochem* **2004,** 271, (11), 2117-26.
18. Seon, A. A.; Pierre, T. N.; Redeker, V.; Lacombe, C.; Delfour, A.; Nicolas, P.; Amiche, M., Isolation, structure, synthesis, and activity of a new member of the calcitonin gene-related peptide family from frog skin and molecular cloning of its precursor. *J Biol Chem* **2000,** 275, (8), 5934-40.
19. Batista, C. V.; Scaloni, A.; Rigden, D. J.; Silva, L. R.; Rodrigues Romero, A.; Dukor, R.; Sebben, A.; Talamo, F.; Bloch, C., A novel heterodimeric antimicrobial peptide from the tree-frog Phyllomedusa distincta. *FEBS Lett* **2001,** 494, (1-2), 85-9.
20. Conceição, K.; Bruni, F. M.; Sciani, J. M.; Konno, K.; Melo, R. L.; Antoniazzi, M. M.; Jared, C.; Lopes-Ferreira, M.; Pimenta, D. C., Identification of bradykinin-related peptides from Phyllomedusa nordestina skin secretion using electrospray ionization tandem mass spectrometry after a single-step liquid chromatography. *Journal of Venomous Animals and Toxins including Tropical Diseases* **2009,** 15, (4), 633-52.
21. Kuckelhaus, S. A.; Leite, J. R.; Muniz-Junqueira, M. I.; Sampaio, R. N.; Bloch, C., Jr.; Tosta, C. E., Antiplasmodial and antileishmanial activities of phylloseptin-1, an antimicrobial peptide from the skin secretion of Phyllomedusa azurea (Amphibia). *Exp Parasitol* **2009,** 123, (1), 11-6.
22. Conceicao, K.; Konno, K.; Richardson, M.; Antoniazzi, M. M.; Jared, C.; Daffre, S.; Camargo, A. C.; Pimenta, D. C., Isolation and biochemical characterization of peptides presenting antimicrobial activity from the skin of Phyllomedusa hypochondrialis. *Peptides* **2006,** 27, (12), 3092-9.
23. Tyler, M. J.; Stone, D. J.; Bowie, J. H., A novel method for the release and collection of dermal, glandular secretions from the skin of frogs. *J Pharmacol Toxicol Methods* **1992,** 28, (4), 199-200.
24. Jakubowski, J. A.; Sweedler, J. V., Sequencing and mass profiling highly modified conotoxins using global reduction/alkylation followed by mass spectrometry. *Anal Chem* **2004,** 76, (22), 6541-7.
25. Wang, H.; Wang, L.; Zhou, M.; Yang, M.; Ma, C.; Chen, T.; Zhang, Y.; Zeller, M.; Hornshaw, M.; Shaw, C., Functional peptidomics of amphibian skin secretion: A novel Kunitz-type chymotrypsin inhibitor from the African hyperoliid frog, Kassina senegalensis. *Biochimie* **2012,** 94, (3), 891-9.
26. Erspamer, V.; Melchiorri, P.; Falconieri Erspamer, G.; Montecucchi, P. C.; de Castiglione, R., Phyllomedusa skin: a huge factory and store-house of a variety of active peptides. *Peptides* **1985,** 6 Suppl 3, 7-12.
27. Erspamer, V.; Falconieri Erspamer, G.; Cei, J. M., Active peptides in the skins of two hundred and thirty American amphibian species. *Comp Biochem Physiol C* **1986,** 85, (1), 125-37.
28. Evaristo, G.; Pinkse, M.; Wang, L.; Zhou, M.; Wu, Y.; Wang, H.; Chen, T.; Shaw, C.; Verhaert, P., The chains of the heterodimeric amphibian skin antimicrobial peptide, distinctin, are encoded by separate messenger RNAs. *J Proteomics*. **2012,** (in press).

Chapter 7

The chains of the heterodimeric amphibian skin antimicrobial peptide, distinctin, are encoded by separate messenger RNAs

Abstract

Certain bioactive peptides found in the venoms of snakes, insects, scorpions and frogs occur as dimers (separate peptide chains linked by disulfide bonds). This dimerization may produce homo- or heterodimers and often such molecular organizations can enhance the toxins biological activity. Using a primer to a conserved 5' untranslated nucleotide sequence of previously-cloned *Phyllomedusa* species skin peptides, two distinct cDNAs were "shotgun" cloned from a skin secretion-derived cDNA library of the frog, *Phyllomedusa burmeisteri*, that encoded separate chains (A and B) of an analog of the previously-reported heterodimeric peptide, distinctin. LC MS/MS peptide analysis of native versus reduced and alkylated skin secretion samples, both fully confirmed the predicted amino acid sequences as well as the cystine link between the monomers. Targeted nanoLC MS/MS showed that distinctin predominantly exists in frog secretion as a heterodimer (A-B). Neither of the constituent peptides were detected as monomer, whereas of the two possible homodimers (A-A or B-B), only B-B was detected in comparatively low quantity. High resolution nanoLC MS/MS analysis of *in vitro* dimerization of synthetic replicates of the peptide monomers, demonstrated that both homodimers (A-A and B-B) were formed. Distinctin is thus the first example of an amphibian skin dimeric peptide that is formed by covalent linkage of two chains that are the products of different mRNAs. How this phenomenon occurs *in vivo*, to exclude significant homodimer formation, is unclear at present but a "favored steric state" type of interaction between both chains is most likely.

This chapter is published in the Journal of Proteomics as:

Evaristo, G. P.C.; Pinkse, M. W. H.; Wang, L.; Zhou, M.; Wu ,Y.; Wang, H.; Chen, T.; Shaw, C.; Verhaert, P. D. E. M., The chains of the heterodimeric amphibian skin antimicrobial peptide, distinctin, are encoded by separate messenger RNAs. *J Prot* **2012**. (http://dx.doi.org/10.1016/j.jprot.2012.09.016)

7.1 Introduction

Amphibians are known to release bioactive compounds from their skin granular glands that include alkaloids, steroids, biogenic amines, proteins and peptides[1, 2]. Many hundreds of different peptides, differing in size, charge, hydrophobicity, conformation, primary structure, as well as in post-translational modifications, have been demonstrated in frog defensive skin secretions. As already reported by Vittorio Erspamer, one of the original frog skin peptide researchers, frogs from the genus *Phyllomedusa* have skins that are "a treasure trove of biologically-active peptides"[3], due to the large quantities and structural diversity of the peptides that he found in members of this taxon. Several *Phyllomedusa* skin peptides bear striking identities to mammalian (neuro)peptides[3, 4], and exhibit bioactivities in mammalian systems that are consistent with interaction with and activation of endogenous receptors. Other peptide classes include highly-potent antivirals and antimicrobials that are effective against bacteria as well as protozoa (recently reviewed[5, 6]). These rich sources of bioactive peptides are of particular interest to pharmacologists who are continuously searching for potential novel drug leads from unusual sources. Peptides thus far isolated from phyllomedusine skin, belong to diverse families, among which the dermaseptins and phylloseptins represent the largest both in terms of the numbers of peptide sequences identified and in the numbers of species in which they have been found[6]. The emergence of high throughput techniques such as tandem mass spectrometry based "peptidomics" and cDNA cloning have substantially assisted in enlarging the collection of sequences of bioactive (frog skin) peptides. Performance of high resolution mass spectrometry linked to efficient and robust (nano)HPLC or UPLC, is a very powerful combination strategy to collect (partial) amino acid sequence information straight from complex peptide mixtures. The data thus generated can then be validated and/or completed as soon as the nucleic acid sequences of the corresponding mRNAs become available. The latter has been facilitated by use of an efficient molecular cloning approach, developed by Chen and coworkers[7], based on the use of magnetic oligo(dT) beads to extract polyadenylated mRNA sequences directly from only a few milligrams of the actual frog skin secretion. This technique does not require dissected tissue(s) for mRNA extraction, and hence does not necessitate sacrificing secreting donor specimens. Indeed, harmless 'gland milking' appears sufficient to collect frog skin secretions rich in both peptides and (their) corresponding messenger RNAs.

One successful strategy to "shotgun" clone novel skin peptide precursor-encoding cDNAs from frog skin cDNA libraries, has proven to be possible by the design of primers from highly-conserved nucleic acid sequences derived from skin peptide precursor-encoding cDNAs from related frog species[8, 9]. We here report the cloning and full structural characterization, directly from lyophilized *P. burmeisteri* skin secretion, of two peptide precursor-encoding cDNAs and their encoded peptides which are homologous to chains A and B from the heterodimeric peptide, distinctin, originally isolated from the skin of the related specie, *P. distincta*[10]. Differential analysis of native and reductively-alkylated *P. burmeisteri* distinctin demonstrated that the intermolecular disulfide bridge between both chains is an endogenous and specific post-translational modification (PTM). Among the PTMs that are characteristic for amphibian skin peptides, disulfide bonds are not uncommon features. However, most of the cystines in frog peptides are intramolecular. The so-called *Rana* box, for example, is a conservative C-terminal motif of 5-8 residues between a cystine[11]. This is typical for certain classes of antimicrobial peptides that are active against Gram-positive/-negative bacteria and the yeast, *Candida albicans*, and is found in peptide groups including the brevinins, gaegurins, nigrocins, odoranains, esculentins, ranalexins, dybowskins, japonicins, palustrins and ranatuerins[12-16]. Other conserved internal disulfide-bridged domains in frog skin peptides are known, such as in the skin calcitonin-gene related peptide and the Kazal protease inhibitor peptides, found in *P. bicolor* and *P. sauvagei,* respectively[17, 18]. Intermolecular disulfide bonds, however, are quite rare in amphibian skin peptides. The only example so far is distinctin, a 5.4 kDa heterodimer composed of two different peptide chains containing 22 and 25 residues, respectively, that was originally identified in the skin secretion of the phyllomedusine frog, *P. distincta*. It has antimicrobial activity against Gram-negative and Gram-positive bacteria[10, 19, 20], and the dimerization was found to enhance the peptides' bioactivity with respect to that of each monomer. In this context, heterodimeric peptides may represent a new class of amphibian skin peptide with potent biological/pharmacological activity that relies upon the formation of intermolecular complexes. This has inspired other researchers to actually engineer disulfide bridges into natural monomeric and linear amphibian skin peptides to form such complexes. Bioassays indeed subsequently confirmed that synthetic heterodimers, such as between magainin 2 and PGLa, and between magainin and its analogue pexiganan (i.e., MSI-78), showed a greatly enhanced antimicrobial activity when compared to the original native (monomeric) peptides[21].

7.2 Materials and Methods

7.2.1 Skin secretion collection

About 10 specimens of *Phyllomedusa burmeisteri*, one of the Brazilian walking leaf frogs, were captured during expeditions (January 2009) in the Pacotuba Forest (Cachoeiro de Itapemirim City), and at a local farm (Brunoro's, at Venda Nova do Imigrante City), both in the State of Espírito Santo (Brazil). Skin secretion samples were collected in the field by gentle transdermal electrical stimulation, essentially as described by Tyler and coworkers [22]. The secretions were jet-washed from the sampled frog with deionized water, after which the donors were immediately released back into their natural environment. Samples were filtered over cellulose acetate (0.2 μm pore size), frozen, lyophilized, and stored at -20 C prior to analysis. The required permit to access the genetic heritage information was obtained from the Brazilian Institute of Environment and Renewable Natural Resources (IBAMA license number 010453/2010-5).

7.2.2 Peptide separation and mass spectrometry

One mg of lyophilized crude skin secretion was diluted in 25 mM NH_4HCO_3. This sample was divided in two. One part was reduced (2 mM dithiothreitol (DTT)) and alkylated (4 mM iodoacetamide (IAM)). Equivalents of 200 micrograms of both samples – native and reduced – were separately analyzed by HPLC (Waters 2695 Alliance, Manchester, UK), on a C4 column (Reprosil C4, 5 μm particles, 2 mm x 150 mm, Dr. Maisch, Germany) coupled on-line to a tandem mass spectrometer. The column was eluted (solvent A, 0.05% TFA, 0.1 M HAc; and B, 0.05% TFA, 0.1 M HAc in acetonitrile) by a linear gradient from 0 to 60% B at 0.75 % min^{-1}, and the eluate was directly nanosprayed into a Q-TOF (QTof Premier, Waters, Manchester, UK). Of the crude secretion, 0.1 μg was also analyzed by nanoLC using a C8 capillary column (Reprosil-Pur 120 C8, 5 μm particles, 150 mm, Dr. Maisch, Germany) coupled to a linear trap-Orbitrap (Orbitrap Velos, ThermoFisher Scientific, Bremen, Germany). To facilitate the location of peptides containing cystine, PTM-driven differential peptide displays were generated by MSight (SIB, Geneva, Switzerland) as described earlier [23]. This proved an elegant way to assess the presence of the various distinctin dimers and monomers in both untreated and reduced secretions. For extensive targeted collision-induced fragmentation analysis of all distinctin ions, the respective peaks were manually selected at their retention times in a separate replicate LC MS/MS run.

7.2.3 Molecular cloning of distinctin peptides chains A and B

Five mg of lyophilized skin secretion were dissolved in 1 ml of cell lysis/mRNA stabilization solution (Dynal, UK). Polyadenylated mRNA was isolated using magnetic oligo-dT beads as described by the manufacturer (Dynal Biotech, UK). The isolated mRNA was subjected to 5' and 3'- RACE procedures to obtain full-length peptide precursor nucleic acid sequence data using a SMART-RACE kit (Clontech UK). Briefly, the 3'-RACE reactions employed a nested universal primer (NUP - supplied with the kit) and a degenerate sense primer (N2-S1; 5'- ACTTTCYGAWTTRYAAGMCCAAABATG-3'), that was complementary to a conserved sequence 5' to the putative signal peptide (including the start codon ATG). This primer had been successfully used to clone different skin secretion peptides from other *Phyllomedusa* species [8, 9]. 3'-RACE reaction products were gel-purified, cloned using a pGEM-T vector system (Promega Corporation) and sequenced using an automated DNA sequencer (ABI 3100).

7.2.4 Peptide synthesis and oxidation reactions

A and B chains were synthesized by standard solid-phase Fmoc chemistry on an automated peptide synthesizer (Protein Technologies PS3, Tucson, AZ, USA). Molar equivalents of synthetic chains A and B were 'incubated' (one week at RT) with themselves or in combination in two different conditions: i) in H_2O and ii) in 50 mM NH_4HCO_3 pH 8.0, containing 5% DMSO. In the H_2O condition the synthetic chains were merely diluted in water individually and A plus B mixed, and analyzed by LC MS/MS. In the other condition the three possible combinations of chains (only A, only B and A plus B) were induced to oxidize in bicarbonate buffer in 5% DMSO pH 8.0, and measured by MS.

7.3 Results

7.3.1 "Shotgun" cloning of distinctin chain precursor-encoding cDNAs

To amplify the distinctin precursor-encoding mRNA sequences present in a cDNA library constructed from lyophilized *P. burmeisteri* skin secretion, a degenerate primer pool designed to a conserved nucleotide sequence upstream of the putative signal sequence of previously cloned *Phyllomedusa* skin peptides, was employed for its interrogation[8, 9]. Two full size preprodistinctin cDNAs were thus "shotgun" cloned yielding two different open reading frames (ORFs) with separate start and stop codons. The prepropeptide sequences can be predicted on the basis of putative prepropeptide convertase motifs and comparisons with similar peptides reported in the literature[7]. The open-reading frame (ORF) of the cDNA encoding the precursor of chain A, contains 69 amino acid residues, the first 22 of which, comprise a putative signal peptide, followed by a 25-residue acidic 'spacer' peptide and finally the 22-residue mature distinctin A chain peptide. The ORF of the cDNA encoding the precursor of chain B contains 67 residues, that includes a 20-residue putative signal peptide, a 22- residue acidic spacer and a 25-residue mature distinctin B chain peptide (Fig. 1). The mature distinctin A- and B-chain peptide sequences are structurally very similar to those described for the original *P. distincta* peptide [10]. A significant difference in the primary structure is at residue 7 of the A-chain, which is an alanine (A) in the *P. burmeisteri* peptide, whereas the *P. distincta* homolog has a proline (P) in this position.

7.3.2 Identification of homo- and heterodimers by MS

MS analysis of crude *P. burmeisteri* skin secretion confirmed the existence of distinctin predominantly as a heterodimer (Fig. 2). The amino acid sequences of the chains were determined by collision-induced dissociation (CID) of both monomers that had been generated by reduction and alkylation of the venom components. MS/MS analysis was carried out using Q-TOF MS as well as by linear ion trap-Orbitrap tandem mass spectrometry. *De novo* sequencing of both chains, fully confirmed the sequences predicted from the "shotgun" cloned precursor cDNAs: a 22-residue mature A-chain: ENREVPAGFTALIKTLRKCKII, and a 25-residue mature B-chain, NLVSGLIEARKYLEQLHRKLKNCKV. In the DTT-treated skin secretion, as expected, no dimers remained (neither homo- nor heterodimers), which can easily be seen in the PTM-driven differential peptide display of reduced versus untreated (native) sample (Fig. 3).

```
(A)      M  A  F  V   K  K  S   L  L  L   V  L  F  L   G  L  V ·
    1    ATGGCTTTCG TTAAAAAATC TCTTCTCCTT GTACTTTTCC TTGGATTGGT
         TACCGAAAGC AATTTTTTAG AGAAGAGGAA CATGAAAAGG AACCTAACCA
         · S  F  S   I  C  E  E   E  K  R   E  T  E   E  D  E  N ·
   51    CTCCTTTTCC ATCTGTGAAG AAGAGAAAAG AGAGACTGAA GAGGACGAGA
         GAGGAAAAGG TAGACACTTC TTCTCTTTTC TCTCTGACTT CTCCTGCTCT
         · E  D  E   I  E  E   E  S  E  E   K  K  R   E  N  R
  101    ATGAGGATGA AATAGAGGAA GAAAGTGAAG AGAAGAAAAG AGAGAATCGA
         TACTCCTACT TTATCTCCTT CTTTCACTTC TCTTCTTTTC TCTCTTAGCT
          E  V  P  A   G  F  T   A  L  I   K  T  L  R   K  C  K ·
  151    GAAGTACCTG CAGGATTCAC TGCATTGATT AAAACATTAA GAAAGTGTAA
         CTTCATGGAC GTCCTAAGTG ACGTAACTAA TTTTGTAATT CTTTCACATT
         · I  I  *
  201    GATTATATAA TCTAAGTAGT ACAGTTATCA ATGATTATGC CAAAACCATA
         CTAATATATT AGATTCATCA TGTCAATAGT TACTAATACG GTTTTGGTAT
  251    TTAAAGCATA TTTAATGTAA AAAAAAAAAA AAAAAAAAAA AAAAAAAA
         AATTTCGTAT AAATTACATT TTTTTTTTTT TTTTTTTTTT TTTTTTTT

(B)      M  A  F  L   K  K  S   L  F  L   V  L  F  L   V  F  L ·
    1    ATGGCTTTCC TTAAAAAATC TCTTTTCCTT GTACTATTCC TTGTATTCCT
         TACCGAAAGG AATTTTTTAG AGAAAAGGAA CATGATAAGG AACATAAGGA
         · S  L  C   E  E  E  K   R  E  E   E  N  E   E  K  Q  E ·
   51    TTCTCTCTGT GAAGAAGAGA AAAGAGAAGA GGAAAATGAG GAAAAACAAG
         AAGAGAGACA CTTCTTCTCT TTTCTCTTCT CCTTTTACTC CTTTTTGTTC
         · D  D  Q   S  E  E   K  R  N  L   V  S  G   L  I  E
  101    AAGACGATCA AAGTGAAGAG AAGAGAAATC TGGTGTCAGG TCTAATAGAA
         TTCTGCTAGT TTCACTTCTC TTCTCTTTAG ACCACAGTCC AGATTATCTT
          A  R  K  Y   L  E  Q   L  H  R   K  L  K  N   C  K  V ·
  151    GCAAGAAAAT ACCTTGAACA GCTGCATCGT AAACTAAAAA ATTGTAAAGT
         CGTTCTTTTA TGGAACTTGT CGACGTAGCA TTTGATTTTT TAACATTTCA
         . *
  201    TTAAGAAAAT GTAAAATCTA AGACCTCTAA GAGTGTTTTC ACACGGTGTG
         AATTCTTTTA CATTTTAGAT TCTGGAGATT CTCACAAAAG TGTGCCACAC
  251    TTTTTGGTGT GTTTTTTGAT GCATTTTTTG TGCAGAGAGG CACAGCAATT
         AAAAACCACA CAAAAAACTA CGTAAAAAAC ACGTCTCTCC GTGTCGTTAA
  301    AATGCTTGCA TTTCTGCACA GGGAAGGCGT CGGAAAACGC ACCAAAAGCA
         TTACGAACGT AAAGACGTGT CCCTTCCGCA GCCTTTTGCG TGGTTTTCGT
  351    CACCTTATGA AACCACCCTA AGGAGCACAA TTACCAATCA TTGTGCCAAA
         GTGGAATACT TTGGTGGGAT TCCTCGTGTT AATGGTTAGT AACACGGTTT
  401    ATAAAATCCA GCATATTTAA ACAAAAAAAA AAAAAAAAAA AAAAAAAAAA
         TATTTTAGGT CGTATAAATT TGTTTTTTTT TTTTTTTTTT TTTTTTTTTT
```

Figure 1: Nucleotide sequences of precursor cDNAs encoding Phyllomedusa burmeisteri *distinctin chain A (A) and distinctin chain B (B) cloned from a skin secretion-derived library. Putative signal peptides (double-underlined), mature processed peptides (single-underlined) and stop codons (asterisks) are indicated.*

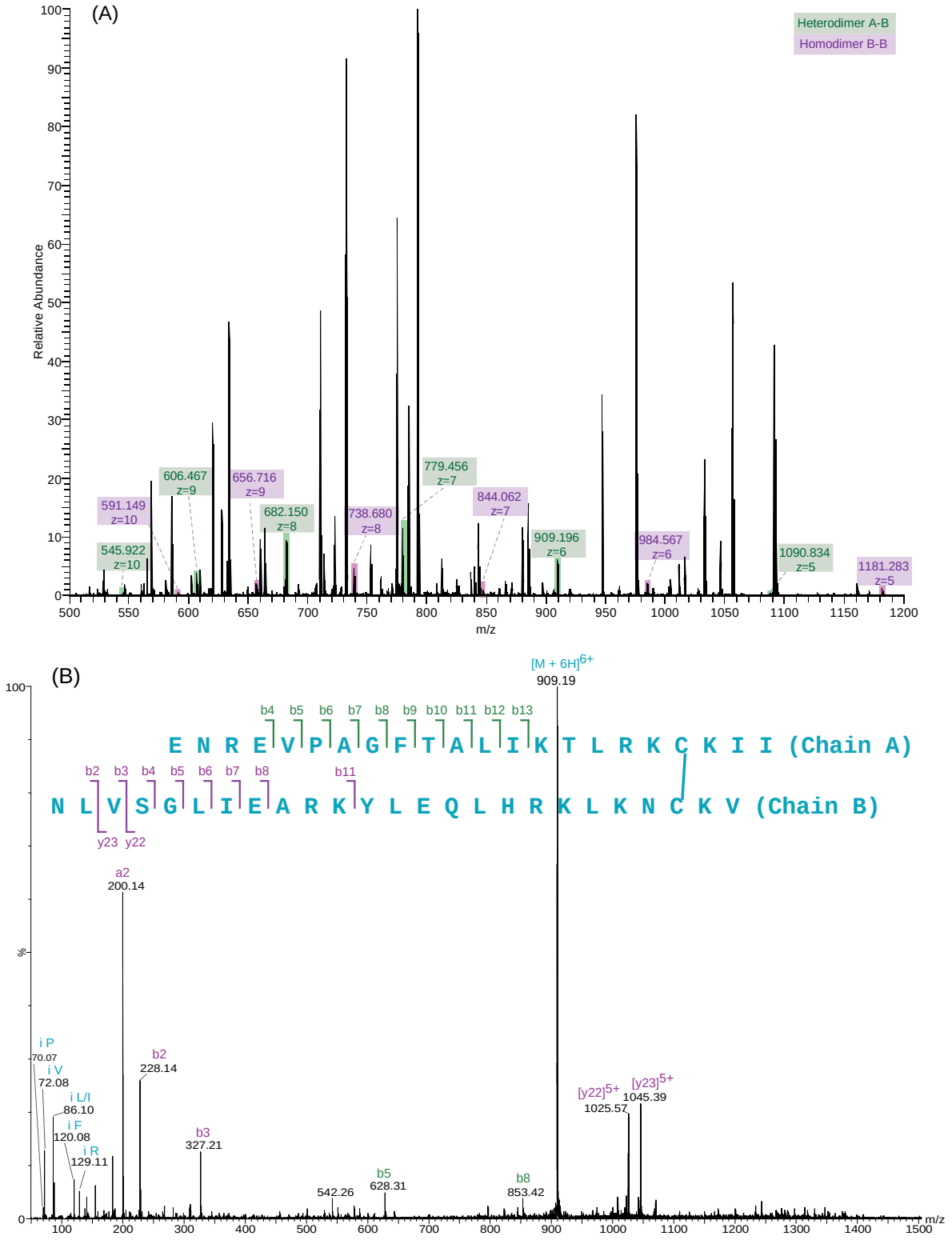


Fig. 2: Distinctin dimers in native P. burmeisteri *skin secretion. (A) Combined MS spectrum of elution times in nanoLC chromatogram of both distinctin dimers. Spectrum shows presence of heterodimer (green peaks) and homodimer B-B (magenta peaks), among several other (higher and lower abundant) non-distinctin peptide peaks. Both distinctin dimers appear as multiply charged ions (z=5 until z=10) with heterodimer being most abundant. (B) Tandem MS of heterodimer (precursor ion at m/z 909.190 selected ($[M+6H]^{6+}$) showing b-, y-, and i-(immonium) ions from both chains A and B. Insert shows primary structure of chain A and B with all detected b- and y- ions indicated.*

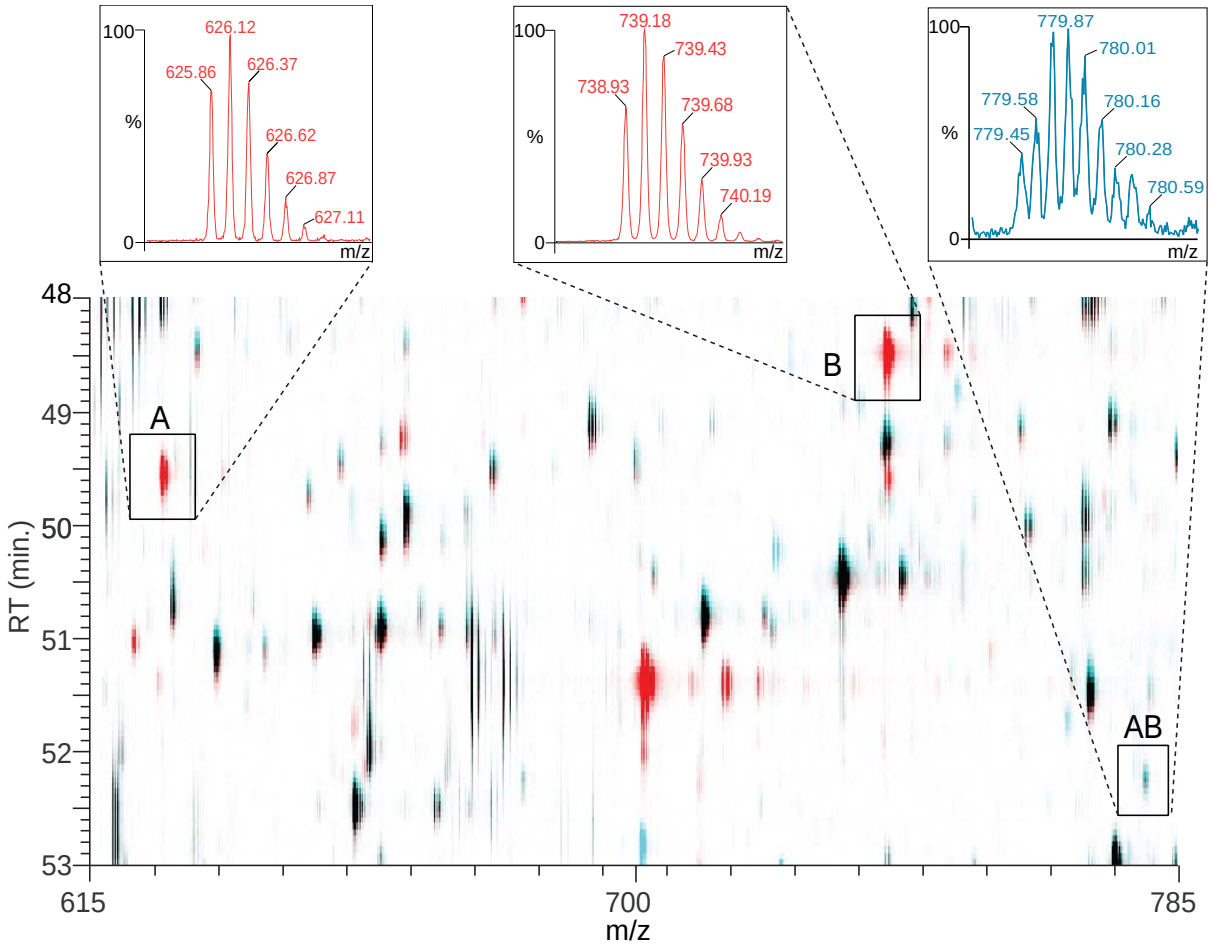


Figure 3: Zoomed-in area of PTM-driven differential peptide display of native crude P. burmeisteri *venom (in blue) and DTT-reduced venom (in red): inserts show* $[M+4H]^{4+}$ *ions of distinctin monomeric chains A (A) and B (B) and* $[M+7H]^{7+}$ *ions of distinctin heterodimer (AB). Note that sum of mass (A) and mass (B) minus mass (2H) exactly equals mass (AB). Ions common in both samples have blue and red colors superimposed, yielding blackish spots. Red color of monomers A and B indicate that these ions are not observed in native venom, whereas, conversely, blue heterodimer ions reveal that no heterodimer was detected in DTT-reduced venom. (Full PTM-driven differential peptide display is provided as Supplementary Fig. 1).*

The distinctin heterodimer, (A-B, *m/z* 779.45 (7+ charge state)), in the zoomed-in area of Fig. 3) was present exclusively in the native sample (complete absence in the DTT-treated venom). On the other hand, the monomeric A-chain (*m/z* 625.96) and B-chain (*m/z* 738.92, four times charged), were entirely absent from crude native skin secretion indicating that in this, they are cross-linked (by cystine bonds) and that the native heterodimer was efficiently reduced by the DTT treatment. Likewise, targeted analysis identified the homodimer B-B in the native skin secretion (high resolution FTMS analyses; Fig. 2A). The heterodimer is clearly present in a higher concentration than the homodimer B-B. No monomeric chain A or B, or

any of the homodimer A-A, were detected in native skin secretion. The heterodimer was observed at *m/z* 909.19 $[M + 6H]^{6+}$, or parent mass 5,449.12 Da. This exactly equals the sum of the masses of chains A and B, minus the mass of 2 protons (disulfide formation from two cysteines). CID of this heterodimer ion precursor indeed yielded both b- and y-ions (as well as immonium ions) representing the N-terminal sequences of chains A and B (peaks highlighted in Fig. 2B). Sequence ions representing residues close to the cystine connection could not be identified, due to the covalent bond between the chains. Similarly the homodimer B-B, with an observed molecular mass of 5,901.34 Da was selected and CID fragmented, showing only chain B amino terminal sequence ions. Together with the precursor mass this proved that it is indeed the covalent dimer.

7.3.3 *In vitro* oxidation of synthetic replicates of distinctin A- and B - chains

In an attempt to answer the intriguing questions as to how/why two peptide chains originating from two different ORFs preferentially form heterodimers, synthetic replicates of the peptide monomers were made and their *in vitro* dimerizations were studied by high resolution nano LC MS/MS. ynthetic A- and B-chains were either left to react in H_2O or in bicarbonate buffer containing 5% DMSO. In H_2O, no dimer was formed (Fig. 4A-C). However, air oxidation of the chains in bicarbonate buffer containing 5% DMSO, induced partial dimerization. Chain A appeared nearly 100 % oxidized when left to react with itself (Fig. 4D), whereas the oxidation to form homodimer B-B, was incomplete (Fig. 4E). The oxidation reaction containing both chains, resulted in the formation of all possible dimers: heterodimer A-B, homodimer B-B and a little homodimer A-A. Only traces of chain A and B monomer remained (Fig. 4F). These oxidation reactions showed no clear preferential dimer formation. They also demonstrated that, in the absence of the other chain, both homodimers can be formed. To reiterate, the native frog skin secretion contained significantly more heterodimer A-B than homodimer B-B, whereas homodimer A-A was not detected.

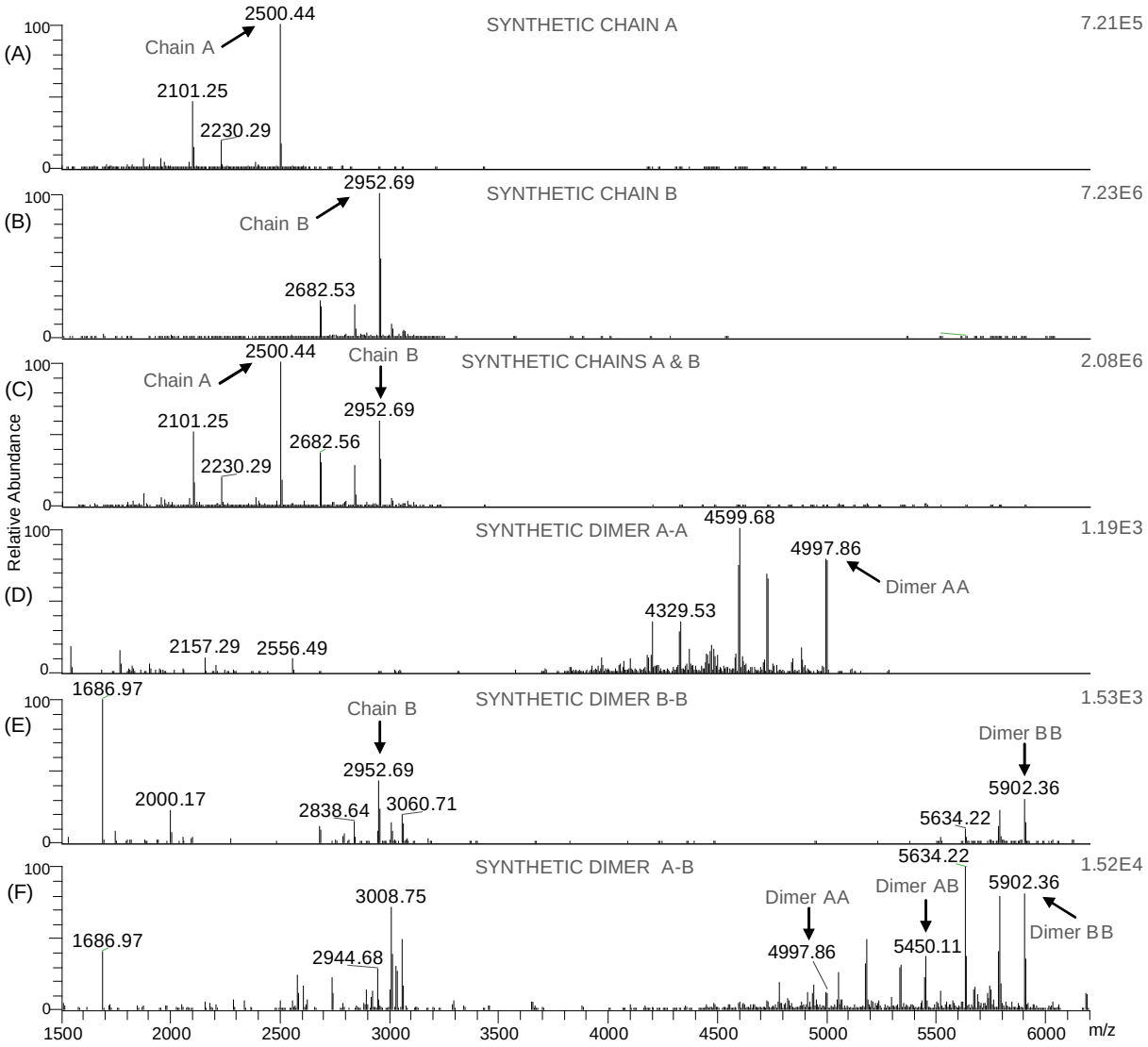


Figure 4: Deconvoluted MS spectra of in vitro dimerization of synthetic distinctin chains A and B. (A-C) "SYNTHETIC CHAIN(S) A/B/A&B" represent the mixture after one week of incubation of the respective chains in water. (D-F) "SYNTHETIC DIMERS A-A/B-B/A-B" show the oxidation products after one week in bicarbonate buffer pH 8.0. Masses of respective chains (monomers) and dimers are indicated by arrows.

7.4 Discussion

Here, we report the molecular cloning of cDNAs encoding the precursor sequences of two chains, A and B, that constitute the heterodimeric peptide, distinctin, from the skin secretion of *Phyllomedusa burmeisteri*. The corresponding mature peptide gene products were detected in chromatographic fractions of the skin secretion as a distinctin heterodimer using MS. The only heterodimeric peptide described to date in amphibian skin secretions is the peptide, distinctin, from *Phyllomedusa distincta*, with a fully-established primary structure and antimicrobial activity [10]. The homodimer B-B, was also identified in the skin secretion of *P. burmeisteri,* and this is the first report of a naturally-occurring distinctin homodimer (Fig. 2A). These data indicate that *in vivo,* distinctin has a strong tendency to exist as a disulfide-bridged dimer. Heterodimerization seems to be preferred, but the detection of the homodimer suggests that rather than to exist as a monomer, excess peptide chains will homodimerize. Interestingly, the cloning and sequencing data indicate that the chains which *in vivo* are detected as heterodimers, are encoded by two different mRNA-encoded precursors, each with their own start and stop codon, that probably represent two distinct gene products. Scrutinizing the protein sequence database (UniProt) found that *P. burmeisteri* chain B exhibits a 71% sequence homology with a distinctin-like peptide chain sequence previously cloned from *Phyllomedusa azurea*[24]. This sequence does not contain any cysteinyl residues and a chain A equivalent in this species was not reported. Thus, this species of phyllomedusine frog provides further, albeit indirect evidence, that chain B appears not to be encoded by the same ORF as chain A.

The mechanism behind the specific dimerization of the two *P. burmeisteri* chains from a pool of peptides containing single free cysteine residues, remains intriguing. Other natural hetero- and/or homodimeric peptides from venoms have been cloned and their nucleic acid sequences reported. These include the paralytic heterodimer, pimplin, from the wasp, *Pimpla hypochondriaca*[25]; the histamine-releasing homodimer, pilosulin 5, from the ant, *Myrmecia pilosula*[26]; hetero- and homo-dimeric αD-conotoxins (VxXXA, VxXXB and VxXXC) found in the venoms of several marine cone snails of the genus *Conus*[27]; salmorin, a thrombin-induced fibrinogen-clotting inhibitor from the snake, *Agkistrodon halys*[28]; numerous dimeric disintegrins, cloned from *Vipera* and *Echis* snakes[29-31]; irditoxin, a neurotoxic dimer from the brown treesnake, *Boiga irregularis*[32]; and various secretory enzymes of the phospholipase A2 family, such as the heterodimers, imperatoxin I and phospholipidin from *Pandinus*

imperator[33, 34] –, and HDP-1P and HDP-2P subunits of the neurotoxic and anti-coagulant heterodimer from *Vipera nikolskii*[35]. As in the case of the well-known example insulin[36], most of these heterodimers arise from a single mRNA precursor, which is translated in one continuous peptide sequence. This is thought to be a crucial step in the heterodimer formation, which occurs during or immediately after translation. A connecting peptide (of varying length in insulin and the various venom heterodimers) between both mature chains is spliced out by specific endoproteases, and this occurs after the disulfide bridge(s) formation.

The cDNAs of disintegrins, salmorin and irditoxin chains represent an exception to this rule, with both monomers encoded by different open reading frames. However, in all these cases, the pre-peptide sequences of A and B chains are very similar or almost identical (typically >70% sequence identity), which suggests that a gene duplication lies at the origin of both ORFs. In disintegrins, for example, it is suggested that the A chain precursor gene was derived from a B chain gene duplication. Subsequent deletion of a continuous and large base pair sequence of the ORF, encoding a C-terminal metalloprotease domain in the B chain pre-peptide gene and of a N-terminal portion of the disintegrin domain, are thought to produce the final structure of the A chain[30]. Therefore, the disulfide connection between the subunits would proceed similarly as in homodimers (which are also detected in these venoms), where the dimerization takes place by virtue of the simultaneous presence of the chains in the endoplasmic reticulum (ER) during the translation process. The case of the heterodimeric *P. burmeisteri* distinctin is markedly different. The peptides representing the signal peptide, acidic spacer and mature peptide of each chain all exhibit different lengths as well as many sequence dissimilarities, yielding an overall sequence identity of only 47%. This suggests that a different mechanism may be behind the dimerization of these dissimilar subunit chains, perhaps comparable to what happens when antibody molecules (immunoglobulins) are maturing, with heavy and light chains translated from different chromosomes "finding" each other and oxidizing in the ER[37]. This is a hypothesis, and the actual mechanism of *in vivo* heterodimer formation is still unclear.

In an attempt to obtain information on potential physicochemical constraints in the formation of the possible dimers from the two subunits identified, *in vitro* dimerization experiments were performed with synthetic replicates of the monomer peptides. However, this did not provide clues as to whether (sequence specific) physiochemical phenomena would preferably direct the formation of one dimer over the other. Our results indicate that both the heterodimer and the two homodimers can be formed *in vitro*. This is similar to what has been described for magainin and conotoxin synthetic dimers[38, 39]. Although, in the native *P.*

burmeisteri skin secretion the main compound detected was the heterodimer, followed by the homodimer B-B, no homodimer A-A, or A or B monomers, were detected. Why the heterodimer prevails in the *P. burmeisteri* skin secretion remains an unanswered question, and we can only speculate at this time. Structure modeling and NMR studies on the similar distinctin chains of *P. distincta* has indicated that matching hydrophobic areas on the outer surfaces of chain A and B and alpha-helices, may play a role in bringing both chains together[19, 40-42]. These studies, however, did not look at the structure of the homodimers.

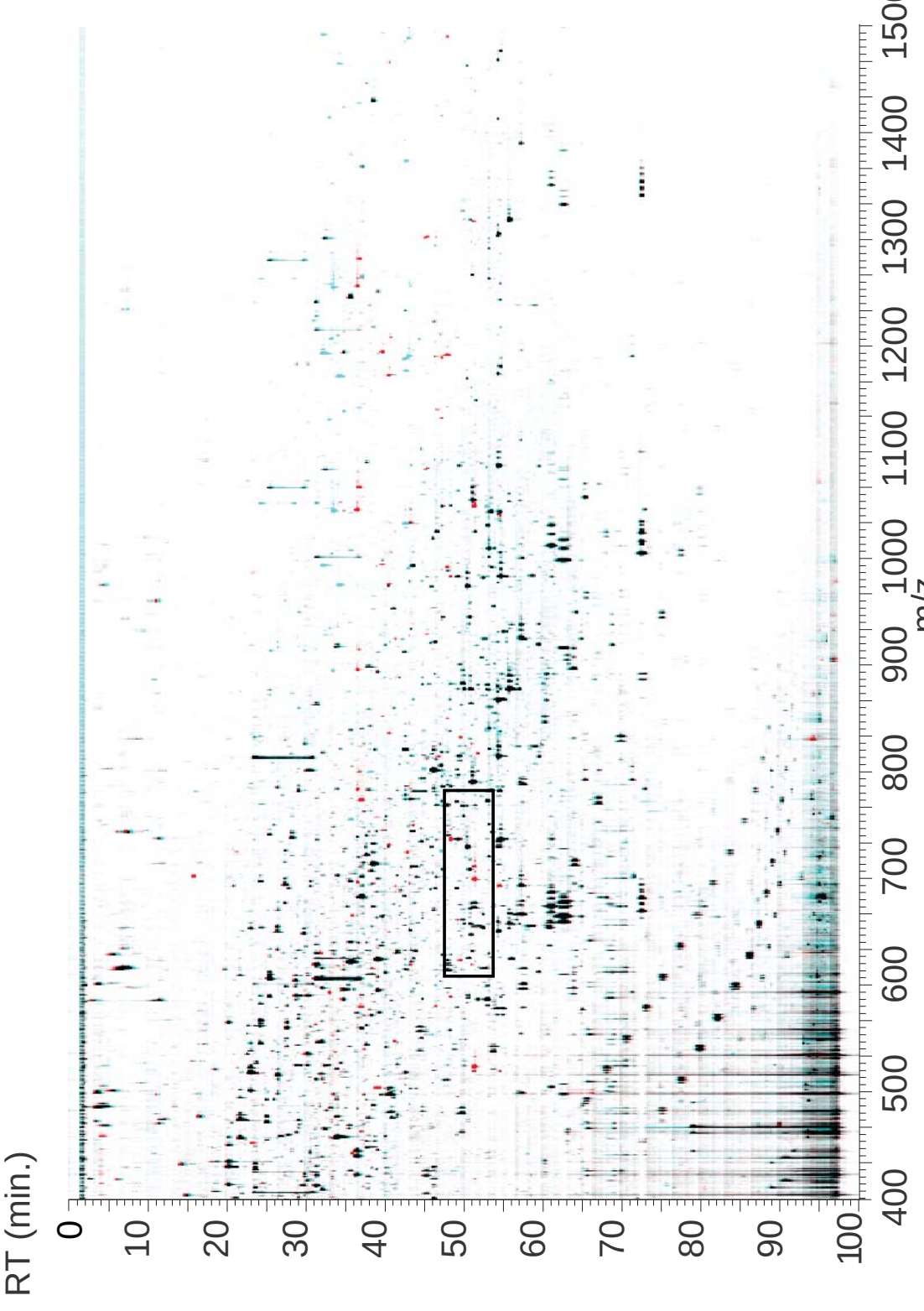
Comparative proteolytic degradation assays of distinctin and synthetic peptide analogues reported that distinctin is more resistant to serine protease breakdown as a heterodimer than as monomers or homodimers[19, 20]. The situation in the *P. burmeisteri* skin secretion, with most of the distinctin occurring as a heterodimer, may reflect a similar situation. Previous investigations have assessed the antimicrobial activity of all possible forms of the closely-related *P. distincta* distinctins[19, 20]. They found that the heterodimer has broad-spectrum antimicrobial activity, and that the dimers (hetero- or homo-) are more active than the separate monomers. With the *P. burmeisteri* distinctins differing in only a single amino acid residue, the bioactivity is likely to be very similar. This would mean that the relative amounts of distinctin chains in the *P. burmeisteri* skin secretion is consistent with maximization of its bioactivity. To assess this, we performed pilot bioactivity studies (*E. coli* growth inhibition test) using dilution series of equivalent amounts of monomeric and *in vitro* dimerized synthetic peptides (see Fig. 4). Our preliminary data indicate that indeed the *P. burmeisteri* dimers have higher specific bioactivity than the monomers, and that the sample containing the heterodimer exhibits the highest antibiotic effect, confirming what was reported for *P. distincta* distinctin. With the present emergence of multiple-drug resistant strains of many pathogenic micro-organisms and diseases requiring a pharmaceutical solution, the development of novel and potently-active pharmaceutical agents with potential clinical and therapeutic applications could be exemplified by natural molecules like the *P. burmeisteri* distinctin heterodimer reported here.

7.5 Conclusion

Distinctin, a 5.4 kDa heterodimeric peptide with antimicrobial activity was identified within the skin secretion of *Phyllomedusa burmeisteri*. Molecular cloning of the cDNA encoding the two chains that constitute this heterodimeric peptide reveals they originate from two different mRNA-encoded precursors, each with their separate start and stop codon. MS analysis showed that in the crude venom the two chains are primarily present as the heterodimer A-B, a subfraction as homodimer B-B, whereas neither homodimer AA nor the individual A of B chains were observed. *In vitro* oxidation experiments with synthetic chains showed no preferential formation of hetero- of homodimer, suggesting there is an alternative mechanism for dimer formation *in-vivo.* To our knowledge this is the first example of an amphibian skin dimeric peptide that is formed by covalent linkage of two chains that are the products of different mRNAs.

Acknowledgements

We are grateful to Prof. Dr. Celio F. B. Haddad for identification of the frogs, and to Mr. Helimar Rabello and his students for guiding and helping with frog capturing. This study was partly funded by the Netherlands Genomics Initiative (NGI) and the Brazilian National Council of Technological and Scientific Development (CNPq – grant GDE-200847/2007-04).



Supplementary Figure 1: Full PTM-driven differential peptide display of P. burmeisteri *crude venom. Selected area is zoomed-in in Fig. 3.*

References

1. Daly, J. W., The chemistry of poisons in amphibian skin. *Proc Natl Acad Sci U S A* **1995,** 92, (1), 9-13.
2. Daly, J. W.; Caceres, J.; Moni, R. W.; Gusovsky, F.; Moos, M., Jr.; Seamon, K. B.; Milton, K.; Myers, C. W., Frog secretions and hunting magic in the upper Amazon: identification of a peptide that interacts with an adenosine receptor. *Proc Natl Acad Sci U S A* **1992,** 89, (22), 10960-3.
3. Erspamer, V.; Melchiorri, P.; Falconieri Erspamer, G.; Montecucchi, P. C.; de Castiglione, R., Phyllomedusa skin: a huge factory and store-house of a variety of active peptides. *Peptides* **1985,** 6 Suppl 3, 7-12.
4. Erspamer, V., Amphibian skin peptides in mammals — looking ahead. *Trends in Neurosciences* **1983,** 6, (6), 200-201.
5. Leite, J. R.; Silva, L. P.; Rodrigues, M. I.; Prates, M. V.; Brand, G. D.; Lacava, B. M.; Azevedo, R. B.; Bocca, A. L.; Albuquerque, S.; Bloch, C., Jr., Phylloseptins: a novel class of anti-bacterial and anti-protozoan peptides from the Phyllomedusa genus. *Peptides* **2005,** 26, (4), 565-73.
6. de Azevedo Calderon, L.; Silva , A. d. A. E.; Ciancaglini, P.; Stabeli, R. G., Antimicrobial peptides from Phyllomedusa frogs: from biomolecular diversity to potential nanotechnologic medical applications. *Amino Acids* **2010,** 40, (1), 29-49.
7. Chen, T.; Zhou, M.; Chen, W.; Lorimer, J.; Rao, P.; Walker, B.; Shaw, C., Cloning from tissue surrogates: antimicrobial peptide (esculentin) cDNAs from the defensive skin secretions of Chinese ranid frogs. *Genomics* **2006,** 87, (5), 638-44.
8. Chen, T.; Tang, L.; Shaw, C., Identification of three novel Phyllomedusa sauvagei dermaseptins (sVI-sVIII) by cloning from a skin secretion-derived cDNA library. *Regul Pept* **2003,** 116, (1-3), 139-46.
9. Chen, T.; Zhou, M.; Gagliardo, R.; Walker, B.; Shaw, C., Elements of the granular gland peptidome and transcriptome persist in air-dried skin of the South American orange-legged leaf frog, Phyllomedusa hypocondrialis. *Peptides* **2006,** 27, (9), 2129-36.
10. Batista, C. V.; Scaloni, A.; Rigden, D. J.; Silva, L. R.; Rodrigues Romero, A.; Dukor, R.; Sebben, A.; Talamo, F.; Bloch, C., A novel heterodimeric antimicrobial peptide from the tree-frog Phyllomedusa distincta. *FEBS Lett* **2001,** 494, (1-2), 85-9.
11. Conlon, J. M.; Kolodziejek, J.; Nowotny, N., Antimicrobial peptides from ranid frogs: taxonomic and phylogenetic markers and a potential source of new therapeutic agents. *Biochim Biophys Acta* **2004,** 1696, (1), 1-14.
12. Chen, T.; Zhou, M.; Rao, P.; Walker, B.; Shaw, C., The Chinese bamboo leaf odorous frog (Rana (Odorrana) versabilis) and North American Rana frogs share the same families of skin antimicrobial peptides. *Peptides* **2006,** 27, (7), 1738-44.
13. Park, S. H.; Kim, H. E.; Kim, C. M.; Yun, H. J.; Choi, E. C.; Lee, B. J., Role of proline, cysteine and a disulphide bridge in the structure and activity of the anti-microbial peptide gaegurin 5. *Biochem J* **2002,** 368, (Pt 1), 171-82.
14. Wang, L.; Evaristo, G.; Zhou, M.; Pinkse, M.; Wang, M.; Xu, Y.; Jiang, X.; Chen, T.; Rao, P.; Verhaert, P.; Shaw, C., Nigrocin-2 peptides from Chinese Odorrana frogs--integration of UPLC/MS/MS with molecular cloning in amphibian skin peptidome analysis. *Febs J* **2010,** 277, (6), 1519-31.
15. Kim, S. S.; Shim, M. S.; Chung, J.; Lim, D. Y.; Lee, B. J., Purification and characterization of antimicrobial peptides from the skin secretion of Rana dybowskii. *Peptides* **2007,** 28, (8), 1532-9.
16. Conlon, J. M.; Leprince, J.; Vaudry, H.; Jiansheng, H.; Nielsen, P. F., A family of antimicrobial peptides related to japonicin-2 isolated from the skin of the chaochiao

brown frog Rana chaochiaoensis. *Comp Biochem Physiol C Toxicol Pharmacol* **2006,** 144, (1), 101-5.
17. Seon, A. A.; Pierre, T. N.; Redeker, V.; Lacombe, C.; Delfour, A.; Nicolas, P.; Amiche, M., Isolation, structure, synthesis, and activity of a new member of the calcitonin gene-related peptide family from frog skin and molecular cloning of its precursor. *J Biol Chem* **2000,** 275, (8), 5934-40.
18. Gebhard, L. G.; Carrizo, F. U.; Stern, A. L.; Burgardt, N. I.; Faivovich, J.; Lavilla, E.; Ermacora, M. R., A Kazal prolyl endopeptidase inhibitor isolated from the skin of Phyllomedusa sauvagii. *Eur J Biochem* **2004,** 271, (11), 2117-26.
19. Raimondo, D.; Andreotti, G.; Saint, N.; Amodeo, P.; Renzone, G.; Sanseverino, M.; Zocchi, I.; Molle, G.; Motta, A.; Scaloni, A., A folding-dependent mechanism of antimicrobial peptide resistance to degradation unveiled by solution structure of distinctin. *Proc Natl Acad Sci U S A* **2005,** 102, (18), 6309-14.
20. Dalla Serra, M.; Cirioni, O.; Vitale, R. M.; Renzone, G.; Coraiola, M.; Giacometti, A.; Potrich, C.; Baroni, E.; Guella, G.; Sanseverino, M.; De Luca, S.; Scalise, G.; Amodeo, P.; Scaloni, A., Structural features of distinctin affecting peptide biological and biochemical properties. *Biochemistry* **2008,** 47, (30), 7888-99.
21. Gottler, L. M.; Ramamoorthy, A., Structure, membrane orientation, mechanism, and function of pexiganan--a highly potent antimicrobial peptide designed from magainin. *Biochim Biophys Acta* **2009,** 1788, (8), 1680-6.
22. Tyler, M. J.; Stone, D. J.; Bowie, J. H., A novel method for the release and collection of dermal, glandular secretions from the skin of frogs. *J Pharmacol Toxicol Methods* **1992,** 28, (4), 199-200.
23. Evaristo, G.; Pinkse, M.; Pimenta, D. C.; Chen, T.; Shaw, C.; Verhaert, P., PTM-driven Differential Peptide Display: Elucidation of Inter / Intra-molecular Disulfide Bridges in Frog Peptides. *4th EuPA Scientific Meeting. A Proteomics Odyssey Towards Next Decades.* **2010**, 579-581.
24. Thompson, A. H., A genomic/proteomic approach to isolating and identifying bioactive peptides from the skin secretions of Phyllomedusa hypochondrialis azurea. *Uniprot acccession number Q17UZ0* **2006,** Thesis, (Thesis), University of Ulster Coleraine, United Kingdom.
25. Parkinson, N.; Smith, I.; Audsley, N.; Edwards, J. P., Purification of pimplin, a paralytic heterodimeric polypeptide from venom of the parasitoid wasp Pimpla hypochondriaca, and cloning of the cDNA encoding one of the subunits. *Insect Biochem Mol Biol* **2002,** 32, (12), 1769-73.
26. Inagaki, H.; Akagi, M.; Imai, H. T.; Taylor, R. W.; Wiese, M. D.; Davies, N. W.; Kubo, T., Pilosulin 5, a novel histamine-releasing peptide of the Australian ant, Myrmecia pilosula (Jack Jumper Ant). *Arch Biochem Biophys* **2008,** 477, (2), 411-6.
27. Loughnan, M. L.; Nicke, A.; Lawrence, N.; Lewis, R. J., Novel alpha D-conopeptides and their precursors identified by cDNA cloning define the D-conotoxin superfamily. *Biochemistry* **2009,** 48, (17), 3717-29.
28. Koh, Y.; Chung, K.; Kim, D., Purification and cDNA cloning of salmorin that inhibits fibrinogen clotting. *Thromb Res* **2000,** 99, (4), 389-98.
29. Vija, H.; Samel, M.; Siigur, E.; Aaspollu, A.; Tonismagi, K.; Trummal, K.; Subbi, J.; Siigur, J., VGD and MLD-motifs containing heterodimeric disintegrin viplebedin-2 from Vipera lebetina snake venom. Purification and cDNA cloning. *Comp Biochem Physiol B Biochem Mol Biol* **2009,** 153, (3), 253-60.
30. Calvete, J. J.; Marcinkiewicz, C.; Monleon, D.; Esteve, V.; Celda, B.; Juarez, P.; Sanz, L., Snake venom disintegrins: evolution of structure and function. *Toxicon* **2005,** 45, (8), 1063-74.

31. Calvete, J. J.; Moreno-Murciano, M. P.; Theakston, R. D.; Kisiel, D. G.; Marcinkiewicz, C., Snake venom disintegrins: novel dimeric disintegrins and structural diversification by disulphide bond engineering. *Biochem J* **2003,** 372, (Pt 3), 725-34.
32. Pawlak, J.; Mackessy, S. P.; Sixberry, N. M.; Stura, E. A.; Le Du, M. H.; Menez, R.; Foo, C. S.; Menez, A.; Nirthanan, S.; Kini, R. M., Irditoxin, a novel covalently linked heterodimeric three-finger toxin with high taxon-specific neurotoxicity. *Faseb J* **2009,** 23, (2), 534-45.
33. Zamudio, F. Z.; Conde, R.; Arevalo, C.; Becerril, B.; Martin, B. M.; Valdivia, H. H.; Possani, L. D., The mechanism of inhibition of ryanodine receptor channels by imperatoxin I, a heterodimeric protein from the scorpion Pandinus imperator. *J Biol Chem* **1997,** 272, (18), 11886-94.
34. Conde, R.; Zamudio, F. Z.; Becerril, B.; Possani, L. D., Phospholipin, a novel heterodimeric phospholipase A2 from Pandinus imperator scp6pion venom. *FEBS Lett* **1999,** 460, (3), 447-50.
35. Ramazanova, A. S.; Zavada, L. L.; Starkov, V. G.; Kovyazina, I. V.; Subbotina, T. F.; Kostyukhina, E. E.; Dementieva, I. N.; Ovchinnikova, T. V.; Utkin, Y. N., Heterodimeric neurotoxic phospholipases A2--the first proteins from venom of recently established species Vipera nikolskii: implication of venom composition in viper systematics. *Toxicon* **2008,** 51, (4), 524-37.
36. Bell, G. I.; Swain, W. F.; Pictet, R.; Cordell, B.; Goodman, H. M.; Rutter, W. J., Nucleotide sequence of a cDNA clone encoding human preproinsulin. *Nature* **1979,** 282, (5738), 525-7.
37. Uhr, J. W., Intracellular events underlying synthesis and secretion of immunoglobulin. *Cell Immunol* **1970,** 1, (2), 228-44.
38. Hara, T.; Mitani, Y.; Tanaka, K.; Uematsu, N.; Takakura, A.; Tachi, T.; Kodama, H.; Kondo, M.; Mori, H.; Otaka, A.; Nobutaka, F.; Matsuzaki, K., Heterodimer formation between the antimicrobial peptides magainin 2 and PGLa in lipid bilayers: a cross-linking study. *Biochemistry* **2001,** 40, (41), 12395-9.
39. Wu, X. C.; Zhou, M.; Peng, C.; Shao, X. X.; Guo, Z. Y.; Chi, C. W., Novel conopeptides in a form of disulfide-crosslinked dimer. *Peptides* 31, (6), 1001-6.
40. Resende, J. M.; Moraes, C. M.; Munhoz, V. H.; Aisenbrey, C.; Verly, R. M.; Bertani, P.; Cesar, A.; Pilo-Veloso, D.; Bechinger, B., Membrane structure and conformational changes of the antibiotic heterodimeric peptide distinctin by solid-state NMR spectroscopy. *Proc Natl Acad Sci U S A* **2009,** 106, (39), 16639-44.
41. Verardi, R.; Traaseth, N. J.; Shi, L.; Porcelli, F.; Monfregola, L.; De Luca, S.; Amodeo, P.; Veglia, G.; Scaloni, A., Probing membrane topology of the antimicrobial peptide distinctin by solid-state NMR spectroscopy in zwitterionic and charged lipid bilayers. *Biochim Biophys Acta* **2011,** 1808, (1), 34-40.
42. Munhoz, V. H. d. O.; Alcântara, A. F. d. C.; Piló-Veloso, D., Análise Conformacional por Cálculos Teóricos da Distinctina, Peptídeo Antimicrobiano Isolado de Anuros da Espécie *Phyllomedusa distincta*. *Química Nova* **2008,** 31, (4), 822-827.

Chapter 8

Kazal protease inhibitor-like peptides from *Phyllomedusa burmeisteri* elucidated by ETD/CID *de novo* sequencing

Abstract

Natural protease inhibitors can be found in all living organisms (even in viruses), being produced in specific tissues and under certain physiological conditions. In anuran they have been found in the skin secretion, as part of the frogs' defense strategy. Here we report a screen for anti-trypsin activity by testing HPLC fractions from the skin secretion of the tree frog *Phyllomedusa burmeisteri.* Five protease inhibitors with observed masses 5855.8, 5964.7, 6272.8, 6503.2 and 6407.2 Da., out of a pull of more than 12 inhibitors, were (partially) *de novo* sequenced. To *de novo* sequence these large peptides, containing three conserved intramolecular disulfide bonds, they were ethylamine-alkylated and fragmented by electron transfer dissociation (ETD) and collision induced dissociation (CID) and measured by high resolution mass spectrometry. The ETD and CID spectra yielded complementary data from which a large part of the primary sequence of these large peptides could be recovered with only little amounts of venom material 'consumed'. The primary structures show high similarity with Kazal type protease inhibitor found in *P. sauvagei*, which notably has no inhibitory effect on trypsin or chymotrypsin. These new Kazal-like peptides were named PbKP-1, PbKP-2, PbKP-3, PbKP-4 and PbKP-5.

Manuscript in preparation

Evaristo, G. P. C.; Pinkse, M. W. H.; Chen, T.; Mohammed, S.; Heck, J. R. A.; Shaw, C.; Verhaert, P. D. E. M., Kazal protease inhibitor-like peptides from *Phyllomedusa burmeisteri* elucidated by ETD/CID *de novo* sequencing.

8.1 Introduction

Many, if not all, amphibians rely on the bioactivity of substances secreted by their skin glands to cope with threatening environmental challenges. In frogs like *Phyllomedusa* many of these defensive chemicals have been identified as peptides and grouped in several families based on their conservative motifs. Phyllocaerulein, phyllokinin, phyllolitorin, phyllomedusin, phylloseptin, dermaseptin, dermorphin, tryptophillin and hyposin are examples of families of Phyllomedusinae peptides[1,2]. They exhibit diverse biological functions, such as cytolitic, antimicrobial, smooth muscle contraction induction, wound repair, act on nervous system and many other specific functions related to the recent peptide discoveries.

An interesting characteristic of the general anuran peptide families is that they often are analogues to mammalian peptides. Caerulein (or phyllocaerulein), for example, shows sequence and activity similarity with cholecystokinin (CCK-8) and gastrin from mammals[3-6]. And the peptide tyrosine-tyrosine (PYY) hormone, originally found in pigs, was also detected in Phyllomedusinae and Ranidae skins[7]. According to Erspamer, every peptide present in the frog skin is also present in mammalian brain and gut tissues[3, 8]. Moreover, concentrations of the frog skin peptides can reach thousands times higher than in mammalian tissues[3]. Therefore, frog venom research may contribute to discover new and/or low concentration mammalian peptides as it may aid in a better understanding of their pharmacological and biochemical functioning[9, 10]. These statements are supported by the discovery of unusual peptides in the frog skin secretion that do not belong to any specific frog peptide family described above, but are related to mammals metabolism, such as calcitonins, pancreatic polypeptides[7, 11, 12] and protease inhibitors.

Indeed, few protease inhibitors have been described in amphibian venoms, which are defined by a conserved motif in their primary sequence, representing different classes. Kunitz inhibitors have been found in the skin secretion of the tomato frog *Dyscophus guineti* [13, 14] and in *Rana chensinensis* (RCSPI2, PubMed accession GQ303267.1); Bowman–Birk like inhibitors in *Huia versabilis* (HV-BBI)[15]; *B. bombina* (BSTI), *B. maxima* (BMTI), *B. orientalis* (BOTI), *B. variegata* (BVTI), *B. microdeladigitora* (BMSI 1 and BMSI 2), two serine protease from *Hyla simplex* (hylaserpin S1 and S2)[16] and a trypsin inhibitor from *Rana aerolata* constitute a novel inhibitor of trypsin/thrombin that share conserved motif with an inhibitor from the parasitic nematode, *Ascaris suum,* classified in the ascaris PI3 family[15, 17-21]; another trypsin inhibitor (OGTI), from *Odorrana grahami,* a 17 amino acid peptide

containing 2 cysteines, that does not share similarity with the inhibitors family described[22]; irreversible serine protease inhibitors from *Bufo andrewsi* (BATI and baserpin)[23, 24]; a 23 kDa trypsin inhibitor from *Kaloula pulcra hainana* (partially sequenced)[18]; and Kazal inhibitors in *Phyllomedusa sauvagei* (PSKP-1 and PSKP-2)[25].

Protease inhibitors are found in diverse protein structures (and few microorganisms secrete non-proteinaceous compounds which block protease activity[26]) in numerous compartments and tissues of all living organisms and virus. They are responsible to regulate the proteolytic activity of proteases, acting as modulators, and therefore, are very important to keep normal the physiological functions. On the other hand if their synthesis or function are uncontrolled this could lead to serious diseases. In this case, therapeutic treatments using protease inhibitors are needed. At least fifty proteases are currently considered to be potential therapeutic targets, and more than thirty drugs specific for some of them have been approved for clinical use[27]. However, many diseases caused by protease dysfunction remain without cure/solution and, hence, new inhibitors need to be discovered.

The skin secretions of amphibians are a rich source of bioactive compounds and it is not a surprise that frogs together with the bioactive defense-peptides also secrete multiple peptidase/protease inhibitors, which helps to ensure an increased half-life of the defensive peptides and inhibit proteases of microrganisms[28]. Here we report the results from a search for protease inhibitors in the skin secretion of the walking leaf tree frog *Phyllomedusa burmeisteri* using a high throughput trypsin activity test. New polypeptides were discovered and (partial sequence) analysis showed sequence homology to the Kazal class of serine protease inhibitors. We named them as PbKP-1, PbKP-2, PbKP-3, PbKP-4 and PbKP-5 (*Phyllomedusa burmeisteri* kazal-like peptides).

8.2 Materials and Methods

8.2.1 Venom collection

About 10 specimens of the tree-frog *Phyllomedusa burmeisteri* (Pb) were captured in Venda Nova do Imigrante City, Espírito Santo State, Brazil and the skin venoms were locally harvested by a mild transdermal electrical stimulation[29] followed by a gentle massage. The foamy secretions were washed from the dorsal skin using deionized water, and the frogs were immediately released back into their natural environment. The skin secretions were filtered

over cellulose acetate (0.2 μm pore size), frozen, lyophilized and stored at −20 ◦C prior to analyses. The required permit to access the genetic heritage information was obtained from the Brazilian Institute of Environment and Renewable Natural Resources (IBAMA license number 010453/2010-5).

8.2.2 HPLC fractionation and trypsin inhibition assay

11 mg of lyophilized crude Pb secretion was dissolved in 0.5 ml of 0.05/99.95 (v/v) trifluoroacetic acid (TFA)/water and clarified of microparticles by centrifugation. The supernatant was then subjected to LC on an analytical column (Jupiter C-5, 5 μm particle, 300Å pore, 250 mm × 10 mm, Phenomenex, UK) and separated using a linear gradient from 0.05/99.95 (v/v) TFA/water to 0.05/19.95/80.0 (v/v/v) TFA/water/acetonitrile in 240 min. at a flow rate of 1 ml/min. The effluent absorbance was monitored at 214 nm. Fractions were collected every min. and 100 μL aliquots from each fraction were dried for the trypsin inhibition assay. 10 μl of a 43 μM trypsin in 1 mM HCl was added to the wells of a 96-well micro-titer plate containing 180 μL of a 50 μM Gly-Gly-Arg-Amc substrate and either 50% reconstituted chromatographic fraction in 10 mM phosphate buffer, pH 7.4, containing 2.7 mM KCl and 137 mM NaCl (final volume 200 μl). Each determination was carried out in duplicate. The rate of hydrolysis of substrate was monitored continuously, at 37 °C, by measuring the rate of increase of fluorescence, due to production of 7-amino-4-methylcoumarin (Amc) at 460 nm (excitation 360 nm) in a CYTOFLUOR® multi-well plate reader Series 4000 spectrofluorimeter.

8.2.3 MS analysis

20 μgram of the crude venom was dissolved in 50 mM ammonium bicarbonate and reduced with 2.5 mM dithiotreitol (DTT) for 30 min. One third of the reduced sample was further treated with 10 mM iodoacetamide (IAM) and another third of the sample was treated with 10 mM bromoethylamine (BrEA) and reacted for 45 min at room temperature in the dark. The alkylated and non-alkylated samples were individually analyzed by nanoLC-MS/MS. Eluted peptides were electrosprayed with 1.7 kV into the ion source of a LTQ Orbitrap XL with CID and ETD capability (Thermo Fisher Scientific, Bremen, Germany). Data acquisition was performed in the positive ion mode; AGC target values were $1e^6$ and $3e^4$ for respectively MS and MS^2 spectra. Spectra were acquired with a resolution at *m/z* 400 of 60.000 for MS and 30.000 for MS^2 spectra. The data dependent analysis was set to perform

MS^2 on the two most intense multiply charged ions. For each precursor both a CID and ETD spectrum was acquired. For CID a normalized collision energy of 35% and an activation time of 0.25 ms was used. While for ETD, the target value for the anion was set to $2e^5$, the reaction time was set to 50 ms and supplemental activation was enabled. An isolation width of 3.0 amu was applied for all MS^2 fragmentation. ETD and CID spectra from deconvoluted using the XTract option within the Xcalibur software and all spectra were interpreted manually.

The active fractions against trypsin were analyzed threefold by nanoHPLC coupled to a LTQ Orbitrap Velos with CID and HCD capabilities (ThermoFisher Scientific, Bremen, GE): [1] untreated, [2] DTT-reduced (2.5 mM) and [3] IAM-alkylated (5 mM). Full FT MS spectra were acquired with 30K resolution. FT-MS/MS spectra were acquired at a resolution of 7500 at *m/z* 400. For CID a normalized collision energies of 35% was used and for HCD a normalized collision energy of 40% was used. Other parameter values in the data depending settings were: the 2 most intense multiply charged ions with a minimum signal threshold of $1e^5$ were isolated with a width of 2 amu, for CID the activation time was 10 ms and for HCD the activation time was 30 ms.

8.3 Results

8.3.1 Protease inhibitor screen

In this study, a protease inhibitor screen was performed on fractionated skin secretion of the walking leaf frog, *Phyllomedusa burmeisteri.* The fluorimetric anti-trypsin assay of all 240 HPLC fractions (Figure 1) showed anti-trypsin activity in nine consecutive fractions, from #79 (78 min.) until #87 (86 min.). Subsequent MS analysis of these fractions revealed they contained 12 peptides in the mass range of 5900-6500 Da, with varying abundance. Upon reduction of these peptides with dithiotreitol their mass increased with 6.01 amu, indicating they contain at least 6 cysteines that form 3 intramolecular disulfide bonds (Fig. 2). For five peptides (from now on named P1 to P5) that showed a high abundance, high resolution CID, HCD and ETD spectra were acquired in order to obtain as much of their primary sequence as possible.

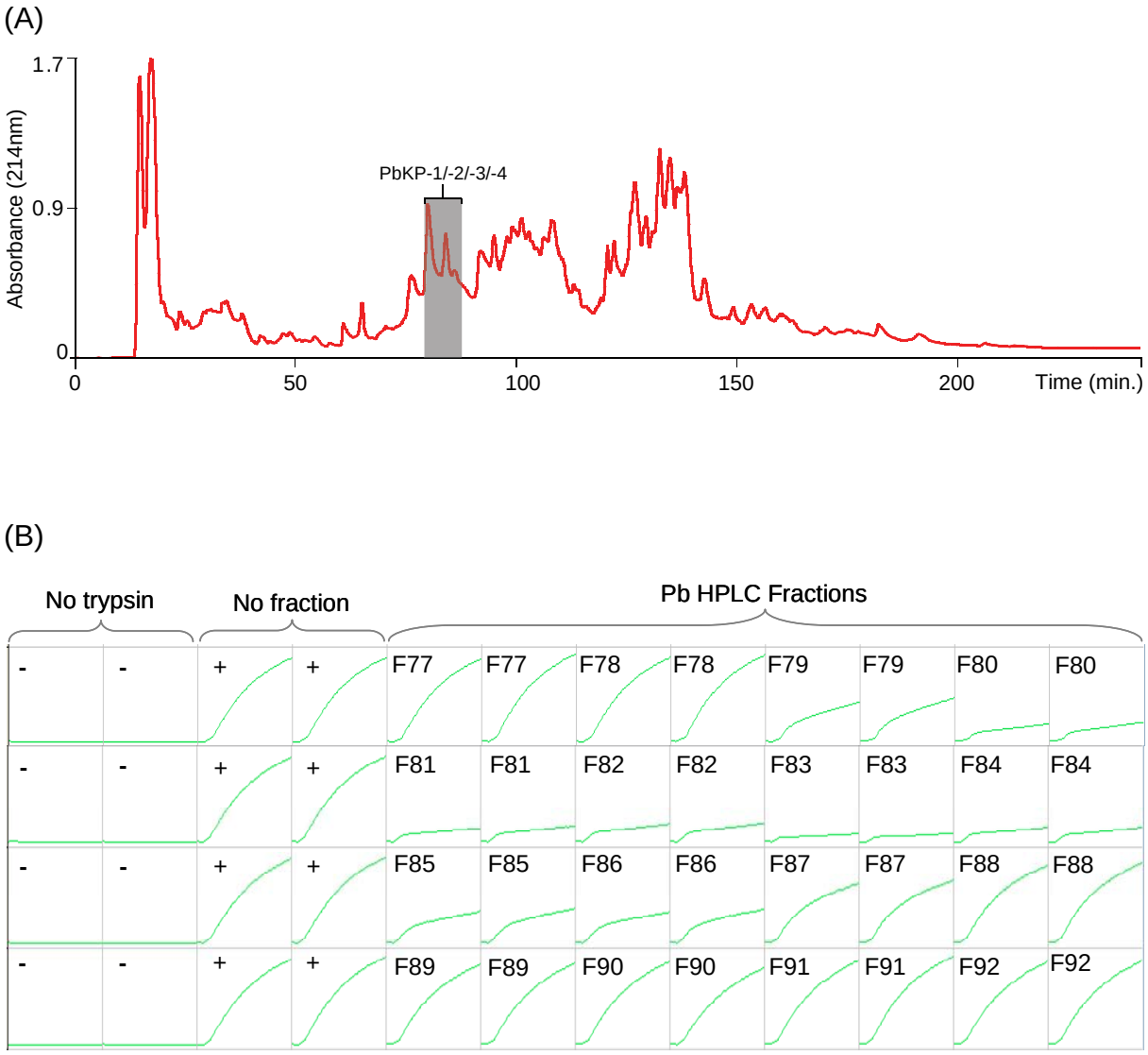


Figure 1: (A) Reversed phase HPLC UV-chromatogram of Phyllomedusa burmeisteri *skin secretion in which the retention time window of PbKP-1, PbKP-2, PbKP-3 and PbKP-4 kazal like peptide inhibitor between 78 and 86 min. are highlighted (grey color). Selected results of the trypsin inhibition assay showing activity in these PbKP containing fractions are shown below. (For explanation see text).*

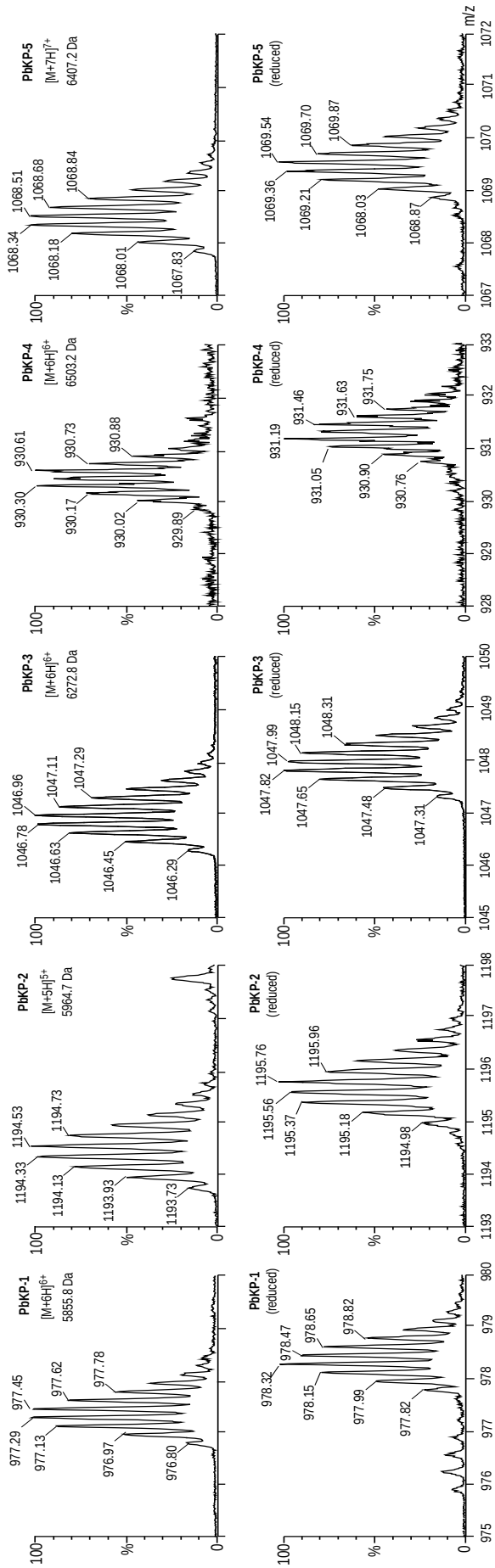


Figure 2: Mass spectra of untreated (top) and DTT-reduced (bottom) of the $[M+6H]^{6+}$ of PbKP-1, PbKP-3, PbKP-4 and the $[M+5H]^{5+}$ of PbKP-2 present in fraction #83, and $[M+7H]^{7+}$ of PbKP-5 found in crude venom. Differences in total mass of untreated and reduced peptide are 1 m/z, for peptides with 6 charge states, 1.25 m/z for PbKP-2, 5 times charged, and 0.87 m/z for PbKP-5 7 times charged, which correspond to a 6 Da difference for each peptide, indicative of 3 disulfide bonds.

8.3.2 *De novo* sequencing

Prior to LC/MS/MS analysis, peptides were reduced with DTT and the cysteines were either blocked with iodoacetamide (IAM) or with bromoethylamine (BrEA). By modifying the cysteines with BrEA, the number of observed charges in the ESI-MS spectra increased, as illustrated for one of the peptides (Fig. 3). This increase of the number of charges might be beneficial for the quality of the ETD fragmentation as previously reported[30]. The ETD spectra of four different charge states (from 7+ to 10+) of one and the same peptide are shown in Fig. 4. The ETD spectra of the 9+ and 10+ precursors yielded more abundant fragment ions as compared to those of the 7+ and 8+ precursors, in which primarily signals of multiple electron captures (but no dissociation) are dominant. For each individual peptide both ETD and CID spectra of the IAM-alkylated and BrEA-alkylated peptide were manually interpreted.

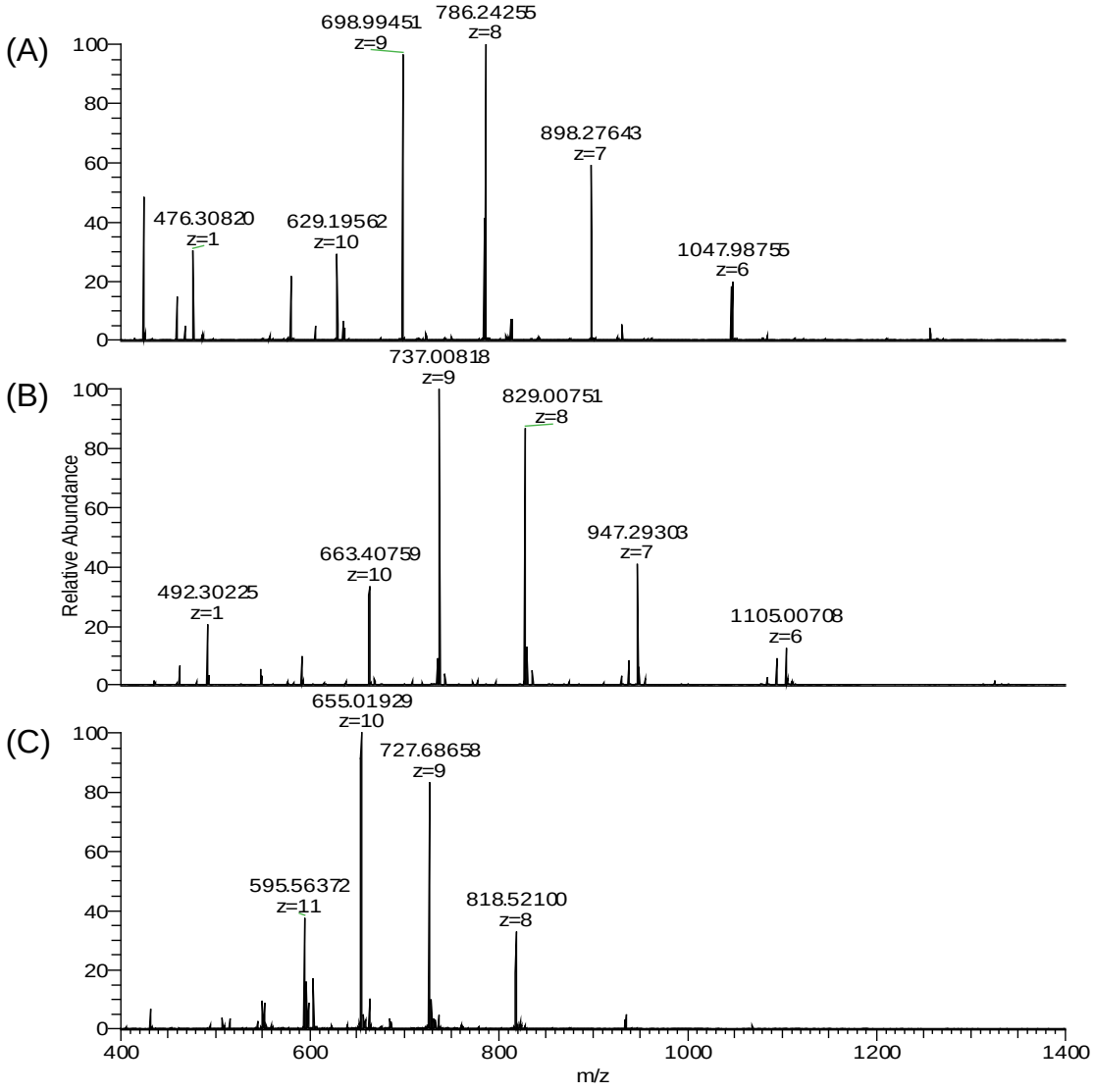


Fig. 3: Electrospray ionization mass spectra of PbKP-3 reduced with (A) DTT, (B) reduced with DTT and alkylated with IAM and (C) reduced with DTT and alkylated with BrEA.

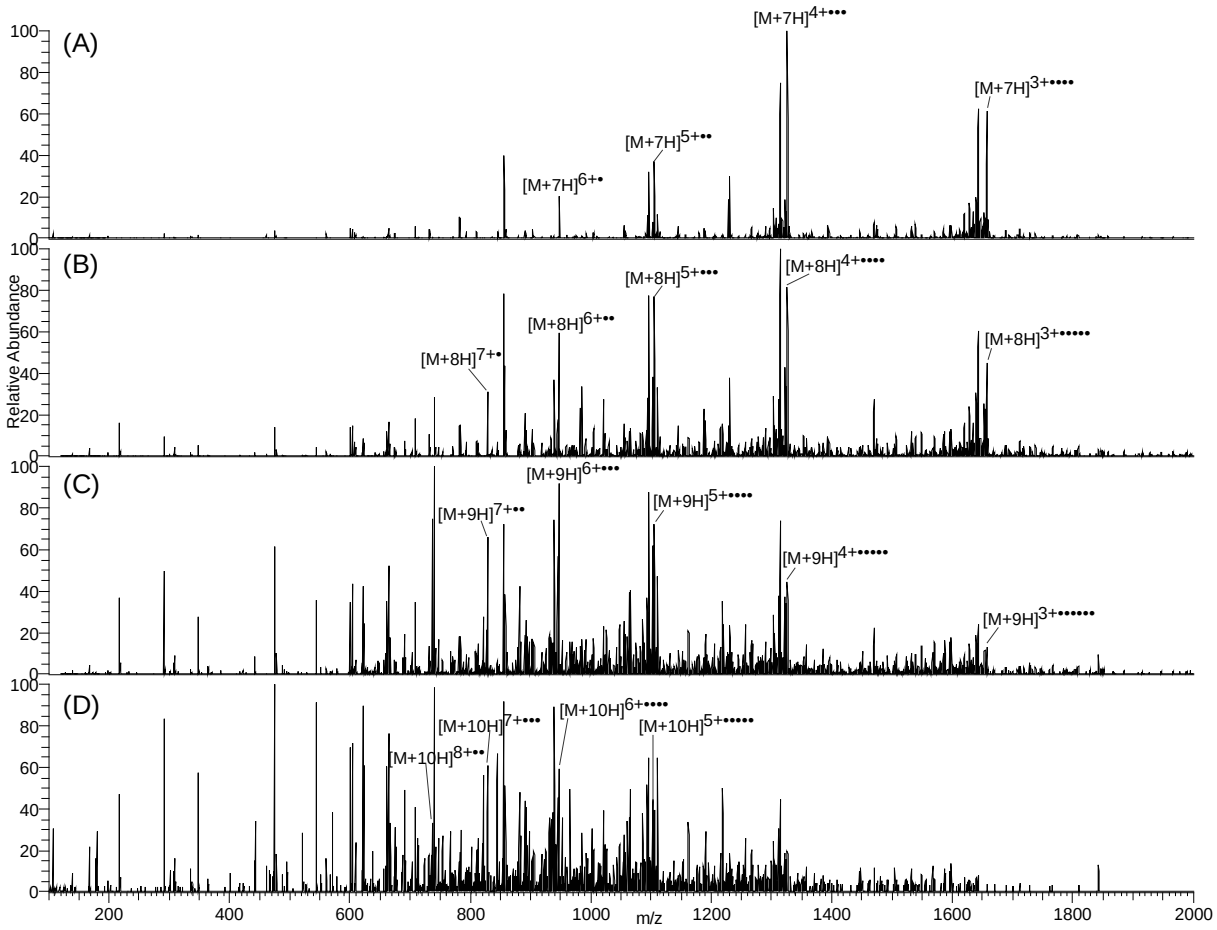


Figure 4: ETD spectra of the PbKP-3 with different precursor charge states: (A) 7+, (B) 8+, (C) 9+, (D) 10+ showing better fragmentation of higher charge state precursors.

As an example, the CID spectrum of PbKP-3 contained only a few fragment ions from which only a short part of the N-terminus of the peptide could be sequenced (Fig. 5A). In contrast to the CID spectrum, the ETD spectrum of PbKP-3 showed many *c*- and *z-ions* allowing us to deduce a great part of the sequence (Fig. 5B). By combining the information of both CID and ETD spectra of the 9+ precursor, a full sequence could be proposed for PbKP-3. In Figure 6 the sequence coverage for both fragmentation techniques is represented in a bar-graph. This graph shows superior sequence coverage from the ETD spectrum, however, information from the CID spectrum completed the total sequence assignment. Interestingly, comparing the intensity of the *c*- and *z-ions* of the BrEA-modified peptide and the IAM modified peptide revealed that the BrEA-modified peptides displayed more abundant *c*- and *z-ions*. Although the IAM spectra also showed these fragment ions, their relative intensity is much lower (Fig. 7). During the manual interpretation of the spectra it was also noticed that in the ETD spectra most of the *z-ions* from fragmentation at the N-terminal side of cysteines

were either very low or completely absent in the ETD spectra, whereas the complementary *c-ions* were present. For all these *z-ions* intense ion signals were observed at -76 amu and –90 amu for respectively ethyl-amine (EtA-Cys) and carbamidomethyl-cysteines (CAM-Cys). These signals correspond to the *z-ions* from N-terminal cleavage at a cysteine of which the cysteine side chain is also cleaved. Side-chain cleavage of modified cysteines in ETD has been previously observed[31, 32]. Peaks missing the side chain of EtA-Cys (76.02 amu) were observed and considered to sequence coverage (Fig. 6). Discrimination of leucine and isoleucine residues from all peptides on basis of the side chain loss in ETD spectra could not be made.

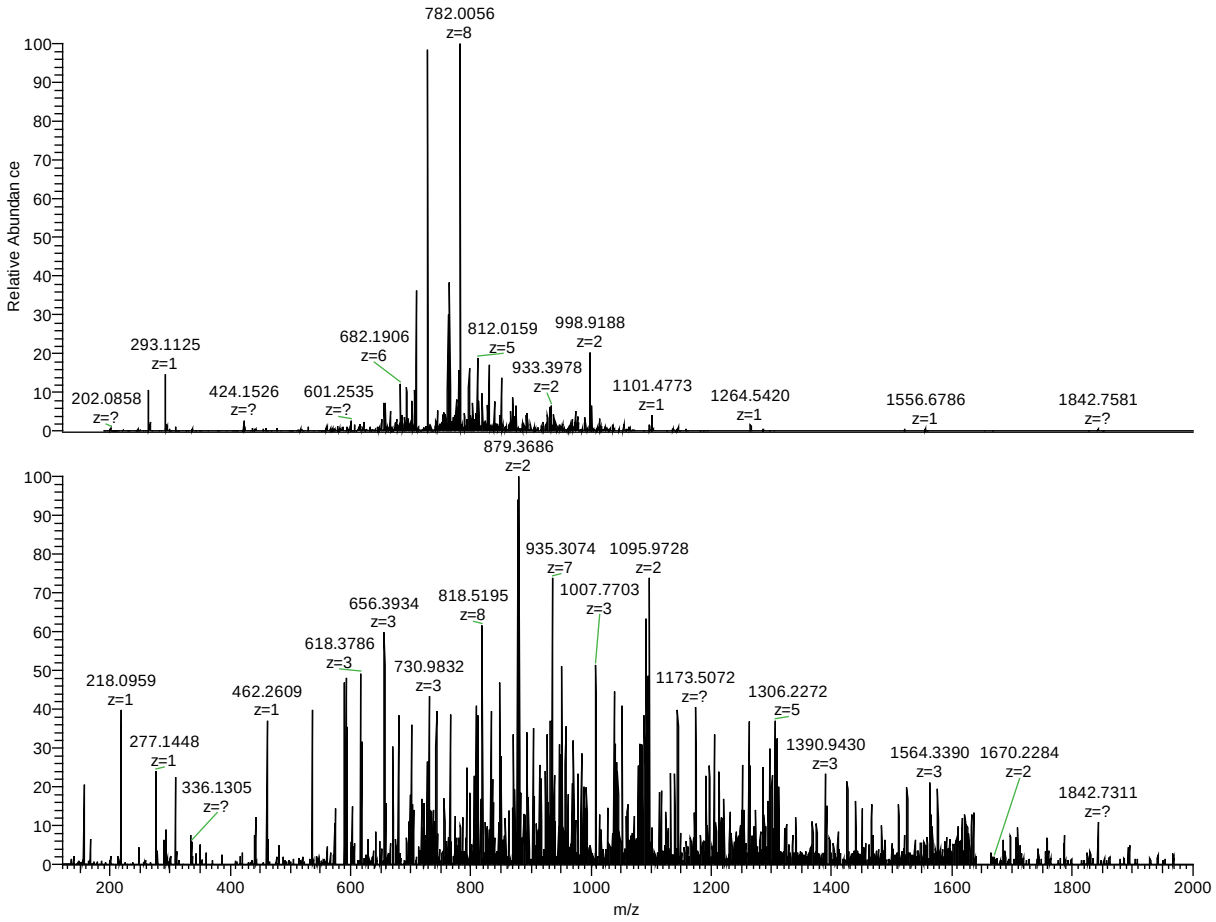


Figure 5: CID (top) and ETD (bottom) fragmentation spectra of $[M+9H]^{9+}$ PbKP-3 alkylated with bromoethylamine.

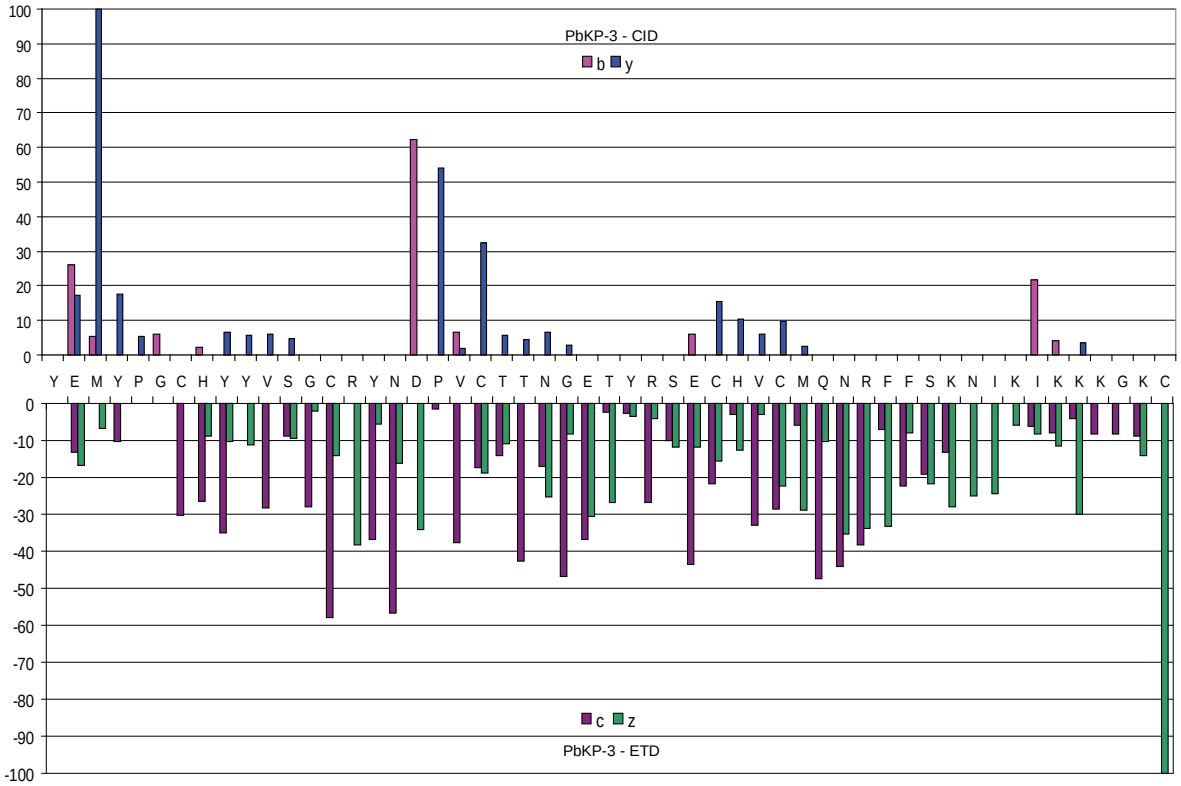


Fig. 6: Sequence coverage and relative ion abundances in CID and ETD spectra of PbKP-3: Shown are (top) the b *and* y' *ions from CID spectra and (bottom)* c *and* z *ions from the ETD spectrum of the* $[M+9H]^{9+}$ *of PbKP-3 alkylated with BrEA.*

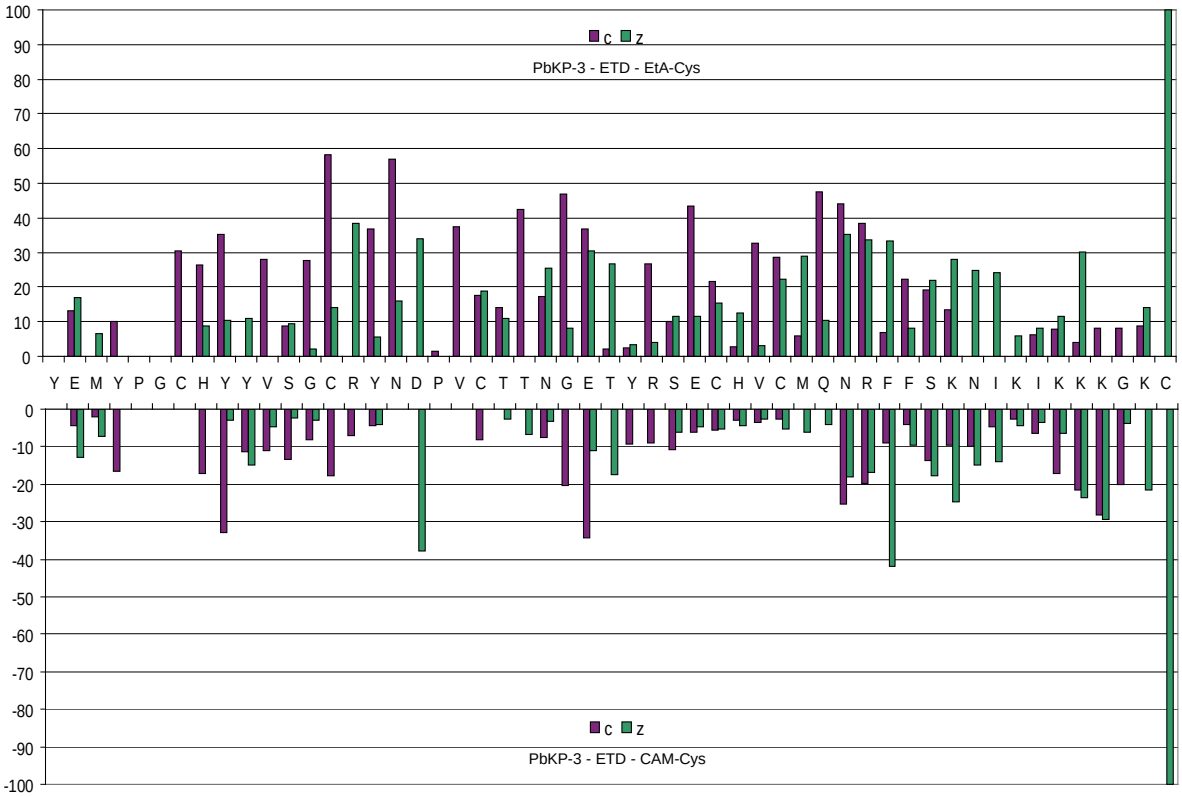


Fig. 7: Comparison of sequencing coverage and relative ion intensity of ETD spectra of $[M+9H]^{9+}$ *precursor of PbKP-3 (top) alkylated with BrEA and (bottom) alkylated with IAM.*

8.3.3 BLAST

After manual interpretation of all ETD, CID and HCD spectra of five peptides, a full sequence could be proposed for four of them and for one approximately 90 % of the sequence could be obtained from the ETD/CID spectra (Table 1). A basic local alignment search (NCBI-blastp) showed that all 5 sequences have conserved Kazal type serine protease inhibitors and follistatin-like domains (KAZAL_FS[cd00104]). PbKP-4 in particular has 84% identity similarity with PSKP-1 and 81% identity similarity with PSKP-2 from *Phyllomedusa sauvagei*[25]. PbKP-3 has 50% sequence similarity with the serine protease inhibitor Kazal type found in the pancreatic juice of the mouse *Rattus norvegicus*[33] (Table 1).

Table 1: Comparison of the five PbKPs from P. burmeisteri *with the primary sequences of Kazal peptides from* P. saugavii *PSKP-1/-2 (NCBI accession n. P83578 and P83579) and* Rattus norvegicus *(*R.nor*, NCBI accession n. 1409221A). Kazal domain pattern for residues between cysteines and highly conserved tyrosine residue position are highlighted on top.*

Kazal	Sequence
Kazal pattern -	C-$x_{(5-12)}$-C-$x_{(6-9)}$-C-$x_{(5-6)}$-Y$x_{(3)}$-C$x_{(2-6)}$C-----$x_{(7-21)}$-----C
PSKP-1	VIEPKCYKYEGKKCPPDINPVCGTDKRTYYNECALCVFIRQSTKKADKAIKIKKWGKC
PSKP-2	VIEPDCKKYEGKKCPPDIALVCGTNGREYYNECALCVFIRDSTLKADKAIKIKKWGKC
R.nor	GNPPAEVNGKTPNCPKQIMGCPRIYDPVCGTNGITYPSECSLCFENRKFGTSIHIGRRGTC
PbKP-1	VIEPKCSRYPGLCTFGPNALCGTNGKKYFNECTLCAFNKKNGKEVKIEKYGTC
PbKP-2	543+CPWFSSGCPSEQNCVFGTDGKSYMDEC-x-WNKNHGEKVKLEKKGRC
PbKP-3	YEMYPGCHYYVSGCRYNDPVCTTNGETYRSECHVCMQNRFFSKNIKIKKKGKC
PbKP-4	VIEPNCKKYEGKKCPPDIALLCGTDGREYYNECALCLFISQSTKKADKAVKIKKWGNC
PbKP-5	KKPRCEKYPPNEKCPLAKNVVCGTDGREYYNECMACVFIRQNKKQEVIKKDGEKC

8.4 Discussion

Here we propose primary amino acid sequences of five novel polypeptides secreted by skin of the Brazilian tree frog *Phyllomedusa burmeisteri*. Comparison of these novel peptides sequences with entries within the NCBI non redundant protein database indicates homology with kazal inhibitors from *P. sauvagei* and *Rattus norvegicus.* Structurally, Kazal is a polypeptide composed of 40 to 70 amino acid residues with a highly conserved pattern of 3 disulfide bonds, with a certain number of amino acid spaced between the cysteines[34]. An amino acid pattern between the cysteines of Kazal-type inhibitors is described in literature on the basis of many invertebrate Kazal sequences as $C–X_{(1-6)}–C–X_{(6-9)}–C–X_{(10)}–C–X_{(3-6)}–C–X_{(7-16)}–C$ [34]. Vertebrate Kazal domains are slightly larger than invertebrates. Here, therefore, we characterize general Kazal motif as $C–X_{(5-12)}–C–X_{(6-9)}–C–X_{(5-6)}–Y\text{-}X_{(3)}\text{-}C–X_{(2-6)}–C–X_{(7-21)}–C$ (Table 1), where '$X_{(n)}$' represents the number of residues between the cysteines and 'Y' is a highly conservative tyrosine residue. The 3 intramolecular disulfide bonds are formed by cysteines 1-5 / 2-4 / 3-6 and each Kazal inhibitor has a reactive site, called P1, which is located 2 residues up stream of the second cysteine (position P3). Another conserved motif is ProValCysGly, which contains the third cysteine of the sequence, and a proline residue as part of the reactive site - P4'. The alignments (Table 1) reveal, besides the conserved Kazal motifs and residues, the predicted reactive site (P1) of the PbKP-polypeptides, which are targeted against S1 pocket of the protease. In *P. sauvagei* PSKP-1 and PSKP-2 the P1 position is occupied by a proline and these peptides effectively inhibit a prolyl endoprotease[16]. In our analysis we performed first an anti-trypsin screen and this prompted us to analyze these large Kazal-like polypeptides in more detail. So far we have not identified this putative trypsin inhibitor. For example, in all the five *P. burmeisteri* peptides sequenced until now the P1 is occupied by Phe (PbKP-1), Ser (PbKP-2), Tyr (PbKP-3), Pro (PbKP-4) and Ile or Leu (PbKP-5), and on basis of this non of these are expected to inhibit trypsin[35]. It should be noted that in total we observed up to 12 ion signals in fractions 81 to 84, which all 'look' like Kazal inhibitors on basis of molecular mass and number of S-S bridges. Unfortunately most of them were too low of intensity to be selected in the data dependent analysis and as a consequence peptide fragmentation spectra for these are lacking.

8.5 Conclusion

Conventional procedures to determine primary structures of peptide include Edman degradation, tandem mass spectrometry, and/or molecular cloning. However, these strategies often require a relatively large amount of material and/or additional purification steps are needed. For molecular cloning, the presence of mRNA precursors in the sample and partial sequence information for the cloning is required. Often the concentration of a certain interesting peptide secreted by a frog is too low for the above described procedure or not enough crude venom is available. This prompted us to explore to capabilities of modern mass spectrometric techniques to fully *de novo* sequence large peptides using little amount of crude venom. A relatively new peptide fragmentation method, called electron transfer dissociation (ETD), has been successfully used in the sequencing (of multiply charged) peptides[36, 37]. The fragmentation is induced by converting the positively charged peptides into radicals during an electron transfer reaction using radical anions[38]. During peptide fragmentation, *c*- and *z*-ions are generated; instead of the classical *b*- and *y*-ions formed during collision induced dissociation (CID). ETD reactions are performed in ion trap mass spectrometers, which also have the CID fragmentation option. Here we were able to fully *de novo* sequence several 6kDa peptides by manual interpretation of both ETD/CID spectra. We observed that if cysteines are alkylated with bromoethylamine, rather than the commonly used iodoacetamide, the ETD fragmentation spectra displays more dominant fragment ions. Furthermore, ETD and CID fragmentation spectra yielded in many cases complementary information as illustrated for the $[M+9]^{9+}$ PbKP-3 and $[M+7]^{7+}$ PbKP-5 (Fig. 5 and 6). ETD gave more complete series of *c-/z*-ions, and CID showed important *b-/y*-ions to distinguish the right order of residues if these were not present in the ETD spectrum.

In summary, we describe here an approach to identify serine protease inhibitors in a complex sample as such as the *Phyllomedusa burmeisteri* skin secretion and *de novo* sequence these large peptides from their ETD and CID fragmentation spectra.

Acknowledgements

We thank Dr. Prof. Celio F. B. Haddad for identification of the frogs and to Mr. Helimar Rabello to help collect the frogs. This work was supported by a PhD grant 200847/2007-4 from the Brazilian National Counsel of Technological and Scientific Development (CNPq-Brazil) and funded by the Netherlands Genomics Initiative (NGI).

References

1. Erspamer, V.; Melchiorri, P.; Falconieri Erspamer, G.; Montecucchi, P. C.; de Castiglione, R., Phyllomedusa skin: a huge factory and store-house of a variety of active peptides. *Peptides* **1985,** 6 Suppl 3, 7-12.
2. Thompson, A. H.; Bjourson, A. J.; Orr, D. F.; Shaw, C.; McClean, S., Amphibian skin secretomics: application of parallel quadrupole time-of-flight mass spectrometry and peptide precursor cDNA cloning to rapidly characterize the skin secretory peptidome of Phyllomedusa hypochondrialis azurea: discovery of a novel peptide family, the hyposins. *J Proteome Res* **2007,** 6, (9), 3604-13.
3. Erspamer, V., Amphibian skin peptides in mammals — looking ahead. *Trends in Neurosciences* **1983,** 6, (6), 200-201.
4. Dimaline, R., Is caerulein amphibian CCK? *Peptides* **1983,** 4, (4), 457-62.
5. Johnsen, A. H., Phylogeny of the cholecystokinin/gastrin family. *Front Neuroendocrinol* **1998,** 19, (2), 73-99.
6. Roelants, K.; Fry, B. G.; Norman, J. A.; Clynen, E.; Schoofs, L.; Bossuyt, F., Identical skin toxins by convergent molecular adaptation in frogs. *Curr Biol* **2010,** 20, (2), 125-30.
7. Mor, A.; Chartrel, N.; Vaudry, H.; Nicolas, P., Skin peptide tyrosine-tyrosine, a member of the pancreatic polypeptide family: isolation, structure, synthesis, and endocrine activity. *Proc Natl Acad Sci U S A* **1994,** 91, (22), 10295-9.
8. Erspamer, V.; Melchiorri, P.; Broccardo, M.; Erspamer, G. F.; Falaschi, P.; Improota, G.; Negri, L.; Renda, T., The brain-gut-skin triangle: new peptides. *Peptides* **1981,** 2 Suppl 2, 7-16.
9. Erspamer, V. a. Melchiorri, P., Active polypeptides: from amphibian skin to gastrointestinal tract and brain of mammals. *Trends in pharmacological sciences* **1980,** 1, (14), 391-395.
10. Bevins, C. L.; Zasloff, M., Peptides from frog skin. *Annu Rev Biochem* **1990,** 59, 395-414.
11. Conlon, J. M., Peptide tyrosine-tyrosine (PYY) - An evolutionary perspective. *American Zoologist* **1995,** 35, (6), 466-473.
12. Nicolas, P.; Mor, A., Peptides as weapons against microorganisms in the chemical defense system of vertebrates. *Annu Rev Microbiol* **1995,** 49, 277-304.
13. Qi, R. F.; Song, Z. W.; Chi, C. W., Structural features and molecular evolution of Bowman-Birk protease inhibitors and their potential application. *Acta Biochim Biophys Sin (Shanghai)* **2005,** 37, (5), 283-92.
14. Conlon, J. M.; Kim, J. B., A protease inhibitor of the Kunitz family from skin secretions of the tomato frog, Dyscophus guineti (Microhylidae). *Biochem Biophys Res Commun* **2000,** 279, (3), 961-4.
15. Song, G.; Zhou, M.; Chen, W.; Chen, T.; Walker, B.; Shaw, C., HV-BBI--a novel amphibian skin Bowman-Birk-like trypsin inhibitor. *Biochem Biophys Res Commun* **2008,** 372, (1), 191-6.
16. Wu, J.; Liu, H.; Yang, H.; Yu, H.; You, D.; Ma, Y.; Ye, H.; Lai, R., Proteomic Analysis of Skin Defensive Factors of Tree Frog Hyla simplex. *J Proteome Res* **2011,** 10(9): 4230-40.
17. Mignogna, G.; Pascarella, S.; Wechselberger, C.; Hinterleitner, C.; Mollay, C.; Amiconi, G.; Barra, D.; Kreil, G., BSTI, a trypsin inhibitor from skin secretions of Bombina bombina related to protease inhibitors of nematodes. *Protein Sci* **1996,** 5, (2), 357-62.
18. Zhang, Y.; Wang, M.; Wei, S., Isolation and characterization of a trypsin inhibitor from the skin secretions of Kaloula pulchra hainana. *Toxicon* **2010,** 56, (4), 502-7.

19. Chen, T.; Shaw, C., Identification and molecular cloning of novel trypsin inhibitor analogs from the dermal venom of the Oriental fire-bellied toad (Bombina orientalis) and the European yellow-bellied toad (Bombina variegata). *Peptides* **2003,** 24, (6), 873-80.
20. Lu, X.; Ma, Y.; Wu, J.; Lai, R., Two serine protease inhibitors from the skin secretions of the toad, Bombina microdeladigitora. *Comp Biochem Physiol B Biochem Mol Biol* **2008,** 149, (4), 608-12.
21. Ali, M. F.; Lips, K. R.; Knoop, F. C.; Fritzsch, B.; Miller, C.; Conlon, J. M., Antimicrobial peptides and protease inhibitors in the skin secretions of the crawfish frog, Rana areolata. *Biochim Biophys Acta* **2002,** 1601, (1), 55-63.
22. Li, J.; Wu, J.; Wang, Y.; Xu, X.; Liu, T.; Lai, R.; Zhu, H., A small trypsin inhibitor from the frog of Odorrana grahami. *Biochimie* **2008,** 90, (9), 1356-61.
23. Zhao, Y.; Jin, Y.; Lee, W. H.; Zhang, Y., Isolation and preliminary characterization of a 22-kDa protein with trypsin inhibitory activity from toad Bufo andrewsi skin. *Toxicon* **2005,** 46, (3), 277-81.
24. Zhao, Y.; Jin, Y.; Wei, S. S.; Lee, W. H.; Zhang, Y., Purification and characterization of an irreversible serine protease inhibitor from skin secretions of Bufo andrewsi. *Toxicon* **2005,** 46, (6), 635-40.
25. Gebhard, L. G.; Carrizo, F. U.; Stern, A. L.; Burgardt, N. I.; Faivovich, J.; Lavilla, E.; Ermacora, M. R., A Kazal prolyl endopeptidase inhibitor isolated from the skin of Phyllomedusa sauvagii. *Eur J Biochem* **2004,** 271, (11), 2117-26.
26. Krowarsch, D.; Cierpicki, T.; Jelen, F.; Otlewski, J., Canonical protein inhibitors of serine proteases. *Cell Mol Life Sci* **2003,** 60, (11), 2427-44.
27. Turk, B., Targeting proteases: successes, failures and future prospects. *Nat Rev Drug Discov* **2006,** 5, (9), 785-99.
28. Moss, A. S.; Carty, N.; San Francisco, M. J., Identification and partial characterization of an elastolytic protease in the amphibian pathogen Batrachochytrium dendrobatidis. *Dis Aquat Organ* **2010,** 92, (2-3), 149-58.
29. Tyler, M. J.; Stone, D. J.; Bowie, J. H., A novel method for the release and collection of dermal, glandular secretions from the skin of frogs. *J Pharmacol Toxicol Methods* **1992,** 28, (4), 199-200.
30. Vasicek, L.; Brodbelt, J. S., Enhanced electron transfer dissociation through fixed charge derivatization of cysteines. *Anal Chem* **2009,** 81, (19), 7876-84.
31. Chalkley, R. J.; Brinkworth, C. S.; Burlingame, A. L., Side-chain fragmentation of alkylated cysteine residues in electron capture dissociation mass spectrometry. *J Am Soc Mass Spectrom* **2006,** 17, (9), 1271-4.
32. Zhang, Z., Prediction of electron-transfer/capture dissociation spectra of peptides. *Anal Chem* **2010,** 82, (5), 1990-2005.
33. Uda, K.; Ogawa, M.; Shibata, T.; Murata, A.; Mori, T.; Kikuchi, N.; Yoshida, N.; Tsunasawa, S.; Sakiyama, F., Purification, characterization and amino-acid sequencing of two pancreatic secretory trypsin inhibitors in rat pancreatic juice. *Biol Chem Hoppe Seyler* **1988,** 369 Suppl, 55-61.
34. Rimphanitchayakit, V.; Tassanakajon, A., Structure and function of invertebrate Kazal-type serine proteinase inhibitors. *Dev Comp Immunol* **2010,** 34, (4), 377-86.
35. Laskowski, M., Jr.; Kato, I., Protein inhibitors of proteinases. *Annu Rev Biochem* **1980,** 49, 593-626.
36. Good, D. M.; Wirtala, M.; McAlister, G. C.; Coon, J. J., Performance characteristics of electron transfer dissociation mass spectrometry. *Mol Cell Proteomics* **2007,** 6, (11), 1942-51.

37. Molina, H.; Matthiesen, R.; Kandasamy, K.; Pandey, A., Comprehensive comparison of collision induced dissociation and electron transfer dissociation. *Anal Chem* **2008,** 80, (13), 4825-35.
38. Syka, J. E.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F., Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A* **2004,** 101, (26), 9528-33.

Chapter 9

Concluding remarks and future perspectives

The amphibian skin is a fascinating source of biomolecules, which help the frog to defend itself against predators and invading microorganisms. The skin secretions contain a multitude of bioactive peptides that have diverse functions and this has triggered ample scientific research because of their pharmaceutical potential. Molecular cloning, Edman degradation and mass spectrometry are the most common, and quite often complementary techniques used in frog skin peptide research. The analysis of frog skin secretions is a great challenge due to the remarkable variety of peptides with different primary structures and post-translational modifications (PTMs) and varying abundances.

In this thesis we focus on the characterization of the peptide composition of the skin secretions of various anura using mass spectrometry based approaches. As this technology continuously improves, with new techniques being developed and more sensitive and accurate machines becoming available, it allows us to mine deeper into the peptidome of frog skin secretions than before. Making innovative use of these modern mass spectrometers, in combination with simple peptide chemistry, new peptidomic approaches are established during this PhD that aid in discovery and subsequent *de novo* sequencing of novel frog skin peptides. These approaches proved in particular very useful for the detection and characterization of post-translationally modified peptides. In this final chapter we summaries the mass spectrometry (MS) based methods and approaches we employed during this PhD and discuss the findings and discoveries we made using them.

During our study, we primarily focused on the characterization of the peptides from the skin secretion of two frogs, *Odorrana schmackeri* and *Phyllomedusa burmeisteri.* In addition venoms from other anurans (*Phyllomedusa rohdei*, *Kassina senegalensis* and *Bombina variegata*) were used in some instances to demonstrate the power of MS-based methods. A detailed outline of the MS capabilities and mass spectrometric ‘tricks’ to study amphibian peptides is presented in **Chapter 2**. Overall, the major advantages of using a mass spectrometer to study frog secretions can be summarized as follows:

[1] Without the need of peptide purification, MS provides an overview of the complexity and variety of peptides on basis of their unique molecular masses.

[2] With complementary fragmentation techniques such as ETD and CID, full primary amino acids sequences can be obtained via tandem mass spectrometry more easily then before.

[3] MS allows the characterization of peptides at unprecedentedly low concentrations, i.e. consuming more than 3 orders of magnitude less material than required for Edman degradation analysis after HPLC purification. Single or multi-dimensional miniaturized liquid

chromatography setups combined with mass spectrometric detection allows one to dig deeper and deeper in the peptidome.

[4] On the basis of specific mass-based characteristics, post-translationally modified peptides can be recognized out of a complex background and subsequently be studied in more detail by tandem mass spectrometric investigation.

In **Chapter 3**, a combined peptidomic/transcriptomic approach was undertaken that led to the discovery of novel members of the nigrocin peptide family from *Odorrana schmackeri*. First, molecular cloning of cDNA sequences reverse transcribed from mRNAs isolated from the skin secretion of *O. schmackeri* was performed, following a procedure described by Chen and coworkers from the School of Pharmacy at the Queen's University of Belfast. This yielded three novel cDNA sequences encoding peptide precursors. The mature peptide structures could be unambiguously confirmed by UPLC Q-TOF analysis of the crude secretion. Comparative MS analysis of the native, reduced and alkylated peptides revealed the presence of an intramolecular disulfide bond, the *Rana* box motif. These peptides have extensive sequence similarity with the nigrocin-2 family, so we named them nigrocin-2 SC (*O. schmackeri*) a, b and c. This study was the first in which we introduced a LC-MS 2-dimensional map ('peptide display'), in which all previously reported peptides from *O. schmackeri* venom could be readily spotted by their *m/z* ratio and tandem MS signature. One of the observations from this study was that a majority of peptides from *O. schmackeri* venom still await structural elucidation.

During the analysis of the MS-results obtained from the skin secretion of *Odorrana schmackeri*, several CID spectra, displaying a-typical peptide fragmentation, caught our attention **(Chapter 4)**. Closer inspection showed that CID spectra from several *O. schmackeri* peptides yielded evident monosaccharide fragmentation patterns (N-acetyl hexosamine, hexose and deoxyhexose). At this stage only a small part of the amino acid sequence could be *de novo* sequenced. Additional ETD fragmentation analysis of these peptides completed the full amino acid composition of these glycopeptides as AVP(I/L)(I/L)YNRPG(I/L)YVTKRPK-amide. Interestingly, from these spectra we discovered that the glycosylations were O-linked on the threonine residue. A BLAST search with this peptide sequence showed 100% identity with odorranain-9 (AVPLIYNRPGIYVTKRPKGK), a previously reported antimicrobial peptide from *Odorrana grahami*[1]. This peptide was predicted from in-silico processing of ORFs obtained by molecular cloning of cDNAs from five ranid frogs and additional validation on basis of molecular mass. Our data show that this prediction is not entirely correct, as we found that the

mature form is actually ending with -RPK-amide and not with the -PRKGK as previously proposed. More interesting is our observation that these peptides are, besides the C-terminal amidated lysine, O-glycosylated, carrying 3 to 5 monosaccharides. The supplementary material of the study by *Li et al.* contained MALDI-TOF-MS measurements of HPLC separated fractions of the skin secretion of *O. grahami*. In these MS spectra, we found the same series of molecular masses as we observed for these glycopeptides in *O. schmackeri*, which were not annotated in the study of *Li et al.* (Supplementary figure S25 and S26). In our opinion this is a very nice example of the added value of thorough mass spectrometric characterization, which helps in determining the mature form of peptides (including PTMs content) of those predicted from molecular cloning results. The results described in this chapter further strengthen our view that peptide detection and sequencing by MS is a highly important part in frog secretome peptide analysis. The finding of O-linked glycosylation is also intriguing, as this PTM had not been encountered on frog skin peptides before. O-linked glycosylation have been observed on peptides found in venoms of a lizard and of a snake[2, 3]. However, as in our case, neither the effect of the PTM nor the biological role of the glycopeptides is has been observed on, are at this moment unclear. In the case of the glycopeptides of *O. schmackeri*, we do not know the exact nature of the glyco-moieties. Substantional amounts of the natural material need to be purified to either do additional structural characterization (e.g. NMR) or perform bioassays (e.g. test antimicrobial activity) to learn more about these peptides.

Chapter 5 describes the structural characterization of yet two new peptides found in *O. schmackeri* skin secretion. These peptides were noticed by an a-typical disulfide bond at the N-terminal part of the peptide. Many of the amphibian skin peptides known to have a disulfide bond have this feature in their C-terminal domain. Thorough LC-MS/MS analysis using ETD and CID fragmentation on intact and tryptic peptides in combination with reduction, alkylation and even isotopic labeling (heavy and light dimethyl), were undertaken to get the full primary amino acid sequence. The two peptides differ in only 3 out of 34 residues, and both occur predominantly in their C-terminally amidated variants. Edman degradation of one the peptides fully confirmed the obtained sequence, and allowed the final assignment of Leu and Ile residues, which are likely to be conserved in the second peptide. Similarity searches in the NCBI non redundant proteins database showed that the *O. schmackeri* peptides look similar to two peptides found in the saliva of a thick[4, 5]. Moreover, these peptides share similarities with calcitonin, calcitonin gene-related peptide (CGRP) and adrenomedullin from other vertebrates. However, they do not fit within either of these peptide

families. We named them *O. schmackeri* calcitonin like peptides (OsCLP) 1 and 2. It is the first report of such novel vertebrate peptide structures (in length between calcitonin and CGRP) and about their bioactivity we can only speculate. In view of the activities of the other members of the calcitonin gene peptide family, we predict these peptides may have vasodilator properties, but this needs to be investigated in the future. For this purpose the necessary amounts of pure synthetic peptides have been produced.

In **Chapter 6** we introduce a novel method to specifically target peptides in frog skin secretions in which inter- or intra-molecular disulfide bonds are present. This method consists of three main steps: [1] individual LC-MS analysis of the crude venom untreated and dithiotreitol reduced material, [2] independent conversion of these runs into a 2-D LC-MS profile displays (RT plotted over *m/z* ratio) and [3] overlaying these two 2-D displays generating a PTM-driven differential peptide display. Peptides carrying S-S bond are easily spotted due to the shift in both RT and *m/z* values, whereas peptides without cystines remain unaltered in the 2-D differential display. This method was applied in skin secretions of 5 different amphibians yielding species specific characteristics. Venom of *O. schmackeri* appears predominantly composed of peptides with single intra-molecular disulfide bonds. *P. burmeisteri* and *P. rohdei* venoms show relatively few peptides containing a disulfide bond. Nevertheless they both have polypeptides with 3 intra-molecular S-S bridges and, in *P. burmeisteri*, 2 peptides making inter-molecular links were discovered (**Chapter 7**). In the PTM-differential peptide display of *Kassina senegalensis* a peptide with 2 disulfide bridges was observed, but the general profile showed less peptides with S-S bonds. *Bombina variegata* venom, on the other hand, was composed of very few peptides with 1 cystine. Interestingly, inspection of the PTM-driven differential peptide display of Bombina variegata an a-typical decrease of 319 Da after reduction was observed for two peptides of ~2.9 kDa. The amino–acid sequence of the reduced peptides could be partially solved as *L*QS*L*HK*L*RWPGKP*LLL*CENENGK*L* and *L*QS*L*HK*L*RWPGKP*LLL*CENENEK*L* (*L* = either isoleucine or leucine). This peptide has a single cysteine, and the loss of 319 Da could be due to an intra-molecular S-S bond. Unfortunately, attempts to solve the sequence of the other (smaller) part of this putative dimer have thus far not succeeded. Similarity searches of this sequence showed that the C-terminal part (P*LLL*CENE) has some similarity with acidic spacer regions in several ORFs of skin secreted frog peptides. Whether this is indeed a mature bioactive peptide or an acidic spacer remains unknown. Also it remains unknown whether it is a heterodimeric disulfide linked peptide or whether the cysteine is modified in another way. However, the PTM-driven differential peptide display immediately allowed us to focus on

this peptide and from the results obtained sofar it is intriguing to study it in more detail. Overall, we conclude that a peptide-display yields a very nice overview of the complexity and diversity of molecular masses within a frog skin secretion. The PTM-driven differential peptide display is a very elegant method to visualize instantaneously those peptides in which disulfide bonds are present.

The usefulness of this PTM-driven differential display methods is demonstrated in **Chapter 7** and **8.** In **Chapter 7**, we report the 'shotgun' cloning of two *P. burmeisteri* distinctin-like peptide chains (A and B) from a cDNA library constructed from mRNA isolated directly out of the skin secretion. This was achieved by using primers designed from a conserved 5' untranslated nucleotide sequence of previously-cloned *Phyllomedusa* species peptides. Both A and B distinctin chains have one Cys residue each. In the 2-D peptide display of the reduced sample the two (dissociated) chains (A and B) are readily spotted separately, whereas in the untreated sample, the predominant appearance was as heterodimer AB. Lower abundant ion signal of the homodimer BB were also observed, while individual (monomeric) chains were not detected. This is clear evidence for intermolecular disulfide bridges. An analogue of the heterodimeric distinctin was previously reported in *P. distincta*, however the homodimeric BB distinctin had not yet been reported. *In vitro* oxidation of the synthetic *P. burmeisteri* distinctin chains showed random dimer formation, with no preference for one dimer over the other. Why *in vivo* the heterodimer seems to be preferred remains to be clarified.

Other peptides containing S-S bonds that were easily spotted by the 2-D differential display were the Kazal inhibitor-like polypeptides, from *P. burmeisteri* (*Phyllomedusa burmeisteri* Kazal-like peptides (PbKP), **Chapter 8**). Their structural characterization was triggered by anti-trypsin activity detected in several *P. burmeisteri* skin secretion HPLC fractions. MS analysis established that the active fractions contained 6 kDa polypeptides. In the PTM-driven differential peptide display these large peptides shifted 6 Da after reduction, revealing 3 intra-molecular cystines in each of them. Alkylation with 2 different reagents, bromoethylamine (BrEA) and iodoacetamide (IAM), prior to analysis by MS/MS aided in sequence ion annotations. From the manual interpretations of ETD and CID spectra, 5 of more than 12 such polypeptides, were fully *de novo* sequenced. An interesting characteristic is that the polypeptides alkylated with BrEA yielded higher charge state in ESI and also display a better ETD fragmentation, compared to the IAM alkylated polypeptides. Moreover, we noticed that modified Cys residues yield a 'signature' side chain loss in ETD

fragmentation. Overall, this chapter combines bioassays and MS-based approaches to discover and characterize peptides from frog skin secretions.

In summary, in this thesis we demonstrate the capabilities of state-of-art mass spectrometers and associated tools and techniques to study complex biological samples such a frog skin secretions. We specifically focused on MS-methods and applications to detect specifically PTM-containing peptides. Once modified peptides are noticed, extended MS/MS analysis and interpretation of these allows for the sequence elucidation For this purpose it is highly recommended to use high resolution mass spectrometry, and it is beneficial to make use of multiple peptide fragmentation techniques such as ETD, CID and HCD, as they often yield complementary peptide backbone fragmentation, which is useful for *de novo* sequencing. Chemical and enzymatic treatments are also convenient and in some cases necessary, as illustrated for peptides with disulfide bonds in **Chapters 5** and **8**. Multiple stages of MS (MS^n) can in special cases be quite useful, such as with peptides yielding prominent neutral losses in tandem MS (e.g. sulfated peptides).

Although mass spectrometry is very capable of acquiring high resolution tandem mass spectra, we are very well aware of the fact that in the current stage not all peptides can be *de novo* sequenced by this approach. This is due to the fact that not all fragmentation spectra contain enough information to achieve this, as well as the fact that interpretation is mainly done manually, which is labor intensive and time consuming. The additional usage of information of molecular cloning is very useful, although this is not easily performed at large scale due to costs and time. Similarity searches of the obtained sequences against existing protein and DNA databases have potential as shown in **Chapters 4, 5**, **7** and **8**, however, it generally fails if the sequence tag is too small or if the database is too large (e.g. contains complete genomes). Searching against a sub-collection of specific toxic venom proteins/peptide sequences in a database format is highly desirable. The animal toxin annotation program of Uniprot is in this respect a vey welcome initiative.

Current improvements in transcriptomics, and in particular high-throughput sequencing techniques, (next generation sequencing), may allow in the near future the analysis of all mRNA sequences that can be retrieved for a frog skin secretion. In combination with mass spectrometry analysis and more standard proteomics database search platforms, such as MASCOT and Sequest, a higher throughput in terms of peptide identification, including PTM characterization can be expected. To complete the total picture of a novel peptide being identified and fully structurally characterized, the bioactivity/pharmacology of the peptides needs to be investigated. Structure similarities with other active peptides

(keeping in mind the brain-gut-skin triangle theory proposed by Vittorio Erspamer) can provide leads as to which bioassay(s) to use, as exemplified in **Chapter 5**. The analysis of the bioactivity and pharmacology of the novel peptides sequenced in this PhD thesis is part of the future research plans. For several of the peptides the necessary synthetic replicas of the peptides have already been produced.

Summing up, our work has yielded innovative MS approaches to study bioactive peptides from highly complex mixtures, such as frog skin secretions, using very little amount of sample. It shall be clear that the protocols described are also applicable to other biological peptidomic samples.

Geisa Paulino Caprini Evaristo

References

1. Li, J.; Xu, X.; Xu, C.; Zhou, W.; Zhang, K.; Yu, H.; Zhang, Y.; Zheng, Y.; Rees, H. H.; Lai, R.; Yang, D.; Wu, J., Anti-infection peptidomics of amphibian skin. *Mol Cell Proteomics* **2007,** 6, (5), 882-94.
2. Quinton, L.; Gilles, N.; Smargiasso, N.; Kiehne, A.; De Pauw, E., An unusual family of glycosylated peptides isolated from Dendroaspis angusticeps venom and characterized by combination of collision induced and electron transfer dissociation. *J Am Soc Mass Spectrom* **2011,** 22, (11), 1891-7.
3. Vandermeers-Piret, M. C.; Vandermeers, A.; Gourlet, P.; Ali, M. H.; Waelbroeck, M.; Robberecht, P., Evidence that the lizard helospectin peptides are O-glycosylated. *Eur J Biochem* **2000,** 267, (14), 4556-60.
4. Francischetti, I. M.; Mans, B. J.; Meng, Z.; Gudderra, N.; Veenstra, T. D.; Pham, V. M.; Ribeiro, J. M., An insight into the sialome of the soft tick, Ornithodorus parkeri. *Insect Biochem Mol Biol* **2008,** 38, (1), 1-21.
5. Francischetti, I. M.; Meng, Z.; Mans, B. J.; Gudderra, N.; Hall, M.; Veenstra, T. D.; Pham, V. M.; Kotsyfakis, M.; Ribeiro, J. M., An insight into the salivary transcriptome and proteome of the soft tick and vector of epizootic bovine abortion, Ornithodoros coriaceus. *J Proteomics* **2008,** 71, (5), 493-512.

Samenvatting

Vrijwel alle soorten kikkers scheiden een mengsel van min of meer giftige stoffen uit de huidklieren ter verdediging. Deze huidafscheiding is een rijke bron van biologisch actieve peptiden die een mogelijke toepassing hebben in de farmaceutische wereld. Onderzoek naar dit kikkergif wordt gedaan op drie verschillende niveaus, namelijk peptidoom (primaire structuur)-, transcriptoom- en biologische activiteit analyses **(Hoofdstuk 1)**. Hoewel het complete genoom van een kikkersoort inmiddels in kaart is gebracht, is het analyseren van kikkergifpeptiden van andere soorten toch nog steeds lastig. Dit komt doordat elke species gedurende de evolutie zijn eigen set van bioaktieve peptiden in zijn genen heeft vergaard. In dit proefschrift richten we ons op massaspectrometrische (MS) methoden die bijdragen aan het bestuderen van deze interessante klasse van peptiden. In **Hoofdstuk 2** wordt aan de hand van een aantal voorbeelden uitgelegd hoe massaspectrometrie kan bijdragen aan het karakteriseren van het kikkergif. Veel van de peptiden in het kikkergif zijn na hun synthese post-translationeel gemodificeerd en massaspectrometrie is een goede methode om deze in kaart te brengen. In **Hoofdstuk 2** is dit geïllustreerd voor een aantal veelvoorkomende post-translationele modificaties (PTMs) die eerder zijn gevonden op kikkergifpeptiden.

De huidafscheiding van de Chinese kikker, *Odorrana schmackeri*, was eerder op het transcriptoom niveau uitvoerig bestudeerd. Dit heeft een aantal open reading frames opgeleverd, waarmee de aminozuursequenties van de uiteindelijk biologische peptiden kunnen worden voorspeld. In een volgende stap moet worden aangetoond dat deze voorspelde peptiden ook daadwerkelijk in het gif aanwezig zijn. Nano-vloeistofchromatografie en massaspectrometrie is een ideale combinatie om het bestaan van deze voorspelde peptiden aan te tonen in het gif, aangezien er hiervoor zeer weinig materiaal nodig is . In **Hoofdstuk 3** is deze combinatie van transcriptoom- en peptidoom-analyse gebruikt voor het ontdekken en karakteriseren van verschillende nigrocine-peptiden in het gif van *Odorrana schmackeri*. De resultaten van de MS-metingen toonden aan dat alle nigrocine-peptiden een zwavelbrug bevatten en van de peptide-fragmentatiespectra kon de volledige aminozuursequentie van de drie nigrocine-peptiden worden bevestigd. Tijdens het doen van de metingen aan de nigrocine-peptiden viel een aantal andere peptiden in het oog doordat ze een afwijkende peptide-fragmentatie vertoonde. Vanuit hun CID- en ETD-fragmentatiespectra kon de volledige aminozuurvolgorde worden achterhaald en werd het tevens duidelijk dat het hier om

O-linked geglycosyleerde peptiden gaat **(Hoofdstuk 4)**. Alle 3 de peptiden hebben een identieke aminozuursequentie, waarbij de PTM op een threonine zit. Ze verschillen in het aantal monosaccharides (2, 3, en 4 groepen) dat ze bevatten. Het ongemodificeerde peptide (zonder suiker-groepen) kon niet worden terug gevonden in het kikkergif. Daarnaast is eerder dezelfde primaire structuur opgelost met behulp van transcriptoomanalyse aan het gif van andere Ranidae soorten, echter deze post-translationele modificatie kon niet worden voorspeld vanuit het DNA. Voor zover bij ons bekend is dit het eerste voorbeeld van een kikkergifpeptide welke deze modificatie draagt. De biologische activiteit van dit peptide evenals de rol van de suikergroepen moet nog verder worden onderzocht.

Tijdens dezelfde MS analyses van het gif van *Odorrana schmackeri* vielen nog twee andere peptiden op. De meerderheid van gifpeptiden van kikkers uit de familie Ranidae bevatten een zogenaamde Rana-box motief. Dit is een zwavelbrug in het C-terminale gedeelte van het peptide. Echter, deze twee peptiden van 34 aminozuren groot, hebben een zwavelbrug in het N-terminale gedeelte, iets wat vrij uitzonderlijk is. De zwavelbrug van beide peptiden kon worden verbroken met dithiotreitol en de vrijgekomen thiol groepen werden vervolgens geblokkeerd met zowel bromoethylamine als iodoacetamide. De aminozuursequentie van beide peptiden kon volledig worden achterhaald aan de hand van hoge resolutie CID- en ETD-peptide-fragmentatiespectra van zowel het intacte peptide evenals van fragmenten verkregen na behandeling met het protease trypsine en dimethyl-labeling. De volledige aminozuursequentie, evenals het onderscheid tussen Leu en Ile, kon worden bevestigd door Edman-degradatie. Beide peptiden bevatten delen die lijken op de geconserveerde motieven in calcitonine, adrenomeduline en het calcitonine-gen gerelateerde peptide (“CGRP”). Het is voor het eerst dat een dergelijk peptide is gevonden in Ranidae-kikkergif, en mogelijk vormen deze twee peptiden (OsCLP1 en 2) een nieuwe groep binnen de familie calcitonine-gerelateerde peptiden **(Hoofdstuk 5)**

Een ander kikkergif wat tijdens dit onderzoek is bestudeerd is het gif van *Phyllomedusa* kikkers, welke bekend staan voor hun uiterst giftige huidafscheiding, echter van sommige soorten is nog weinig bekend over de peptiden die in het gif zitten. Tijdens een expeditie in Brazilië werden meerdere exemplaren van zowel *Phyllomedusa burmeisteri* als *Phyllomedusa rohdei* gevangen. Ter plekke werd de huidafscheiding afgenomen waarna de kikkers weer werden vrijgelaten in hun natuurlijke omgeving. Het gif van beide kikkers is gebruikt voor het optimaliseren van de op massaspectrometrie gebaseerde methoden en analyses en op deze manier zijn een aantal verschillende peptiden gevonden. In **Hoofdstuk 6** is een generieke methode omschreven waarmee peptiden met een inter- of intra-moleculaire

zwavelbrug kunnen worden gevonden, alvorens hun primaire structuur kan worden opgelost. Deze methode is gebaseerd op de vergelijking van twee LC-MS metingen van het kikkergif; namelijk van het onbehandelde kikkergif en van dit kikkergif waarin alle zwavelbruggen werden verbroken met behulp van dithiotreitol. De LC-MS data wordt weergegeven in een zogenaamde tweedimensionale "peptide display", waarin de retentietijd wordt uitgezet tegen de massa-over-lading-verhouding. Door de peptide display van beide metingen te vergelijking kunnen peptiden met zwavelbruggen makkelijk worden herkend, omdat hun massa per verbroken zwavelbrug toeneemt met 2 Dalton en de retentietijd meestal ook iets varieert. Behalve het kikkergif van de *Phyllomedusa* kikker, zijn nog 3 andere kikkergiffen met deze methode geanalyseerd en dit toonde aan dat elk gif een andere samenstelling heeft. *O. schmackeri,* bijvoorbeeld, bevat heel veel peptiden met een enkele intra-moleculaire zwavelbrug. Het gif van zowel *P. burmeisteri* als *P. rohdei* bevat een groep van relatief grote peptiden met in totaal 3 intra-moleculaire zwavelbruggen. Het gif van *Kassina senegalensis* is minder complex in samenstelling en bevat een relatief laag gehalte aan peptiden met een enkele zwavelbrug. Tot slot, het gif van *Bombina variegata* bevat een aantal peptiden met een enkele en met een dubbele zwavelbrug. Tevens werd met deze methode een peptide gevonden met een inter-moleculaire zwavelbrug, iets wat vrij uitzonderlijk is voor kikkergif peptiden. In **Hoofdstuk 7** beschrijven we transcriptoom- en peptidoom-analyse van beide A en B keten van dit dimerische peptide genaamd distinctine. Een interessante waarneming is het feit dat beide ketens van distinctine door verschillende open reading frames zijn gecodeerd op het genoom. *In vitro* analyse met synthetische peptiden toont aan dat er geen voorkeur is voor dimerisatie terwijl distinctine voornamelijk als heterodimeer voorkomt.

Met behulp van de methode beschreven in **Hoofdstuk 6**, werden een aantal relatief grote peptiden van ~6 kDa waargenomen in zowel het gif van *P. burmeisteri* en *P. rohdei*. In **Hoofdstuk 8** worden deze peptiden verder gekarakteriseerd met behulp van verschillende peptide-fragmentatie-technieken. Het is uiteindelijk mogelijk geweest om de volledige sequentie van een aantal peptiden te achterhalen vanuit een combinatie van hun CID- en hun ETD-spectra. Het blijkt hier om peptiden te gaan die een sterke overeenkomst hebben met Kazal peptiden, een klasse van protease-remmers. Een interessante waarneming tijdens deze analyses wat het feit dat peptiden gemodificeerd met bromoethylamine een veel beter ETD-fragmentatiepatroon vertoonden dan peptiden gemodificeerd met iodoacetamide, het veel gebruikte cysteïne-blokkeer-reagens in proteomics. Een ander belangrijk aspect hierbij is dat, net zoals voor alle andere MS methoden beschreven in dit proefschrift, hiervoor slechts weinig materiaal nodig was, in tegenstelling tot andere methoden waarmee volledige

structuren kunnen worden opgelost. Dit is zeker een belangrijk punt, aangezien de hoeveelheid materiaal dat beschikbaar is vaak beperkt is of de toegang hiertoe lastig. Tot slot worden in **Hoofdstuk 9** alle belangrijke bevindingen uit dit proefschrift bediscussieerd.

Summary

Most, if not all, frog species are known to secrete a mixture of more or less toxic compounds from skin glands as defense mechanism. This skin secretion of amphibian constitutes a rich source of bioactive peptides with potential pharmaceutical applications. Research in this area includes peptidome (primary structure), transcriptome and bioassays analyses **(Chapter 1)**. Although the complete genome of a frog species has been mapped, the analysis of frog peptides remains difficult. This is mainly because during evolution each frog species has gathered its own set of bioactive peptides encoded in its genes. In this thesis we focus on MS-based methods that help to characterize the peptidome of frog venoms. In **Chapter 2** an overview is given of mass spectrometry methods and approaches to study frog skin secretions. Many of the frog skin peptides are post-translationally modified after their synthesis and mass spectrometry is an excellent tool to characterize them. This is illustrated in **Chapter 2** for several frequently occurring post-translational modifications previously found on frog peptides.

The skin secretion of the Chinese odorous frog *Odorrana schmackeri* had been studied before at the transcriptome level, leading to the discovery of full length open reading frames. The amino acid sequence of the mature peptide can be predicted from these open reading frames obtained by molecular cloning. After obtaining transcriptome results, mass spectrometry is the next logical step to proof the existence of the predicted peptides. Miniaturized nano liquid chromatography in combination with mass spectrometry is the preferred tool to do this as it requires minimal amounts of sample. In **Chapter 3** a combination of transcriptome and peptidome analyses lead to the discovery and characterization of several members of the nigrocin peptide family found in the skin secretion of *Odorrana schmackeri*. The results from the MS measurements showed that all nigrocin peptides have a disulfide-bridge and from their peptide fragmentation spectra the complete amino acid sequence for each of the three nigrocins peptides could be confirmed. During these measurements of the nigrocin peptides, three other peptides caught our attention because of their a-typical peptide fragmentation behavior. Using both CID and ETD fragmentation analyses we could fully *de novo* sequence them and establish that they are O-linked glycosylated peptides **(Chapter 4)**. All three peptides had the same primary structure and O-glycosylation at a Threonine residue. They differed in the monosaccharides

composition (2, 3 and 4 units). Interestingly, the unmodified peptide (i.e. with no glycosylation) was not detected in the venom. Interestingly the very same primary structure has been found by molecular cloning in other Ranid frogs, however, obviously the glycosylation could not be (and was not) predicted from the cDNA/mRNA analysis. To our knowledge, this is the first example of a glycosylation as a PTM in amphibian skin secretion. The bioactivity of the peptide and the role of the glycosylations have yet to be determined.

From the same analysis of the skin secretion of *Odorrana schmackeri*, yet two other peptides caught our interest. The majority of the peptides from Ranidae frogs contain a so called *Rana* box motif, which is a disulfide bridge in the C-terminal domain of the peptide. However, we discovered two relatively large peptides with a disulfide bridge at the N-terminal side, which is very uncommon for frog skin peptides. These peptides, with 34 residues, could be *de novo* sequenced using chemical treatments with DTT, iodoacetamide and bromoethylamine; high resolution CID and ETD fragmentation on full length peptides as well as on selected tryptic fragments; and labeling purified tryptic fragments with heavy and light dimethyl. The entire primary sequence, including Leu and Ile discriminations, was confirmed by Edman degradation. These two peptides share conservative motifs with calcitonin, adrenomedullin and calcitonin gene related peptides, but yet cannot be categorized in either of these existing peptide families. These unusual calcitonin-like peptide sequences (OsCLP1 and 2) present in the skin secretion of an anuran species may represent a novel peptide family within the calcitonin gene related peptide superfamily.

Another type of venom which was studied during this research was that of the genus *Phyllomedusa*, which are known for their highly poisonous skin secretions. Specimens of the walking leaf frog, *Phyllomedusa burmeisteri* and *Phyllomedusa rohdei* were captured during hunting trips in Brazil. Not much was known yet about the peptides present in their toxic secretions. Their skin secretions were collected by a non-invasive method and afterwards all specimen were immediately released in their natural surroundings. The skin secretions of both frogs were used to optimize our MS-based methods and approaches and this has led to the discovery of several novel peptides. In **Chapter 6** a generic method is presented to detect peptides with intra- or an inter-molecular disulfide bonds prior to MS based sequencing of their primairy structure. The method is based upon the comparative analysis by LC-MS of the untreated crude venom and the crude venom in which all S-S bonds are broken with a reducing agent, such as dithiotreitol. The LC-MS data are converted into a so called peptide display, in which the retention time is plotted against the mass-over-charge ratio. By comparing the peptide display of both analyses, peptides with a disulfide bond(s) are easily

recognized, because their mass increases by 2 Da per broken disulfide bond, and their retention times usually also slightly shift. Beside the two *Phyllomedusa* species, 3 other types of frog venoms were subjected to this method and it was observed that each venom shows different profiles. *O. schmackeri*, for example, is highly rich in intramolecular disulfide bonds. The skin secretion of *P. burmeisteri* and *P. rohdei* contain several large peptides containing 3 S-S bonds, *Kassina senegalensis* contains overall a less complex peptide content and only a few peptides with a S-S bridge, whereas *Bombina variegata* venom yielded several peptides with single or double S-S bonds. Inter-molecular disulfide bonds were also observed by this approach in the venom of *P. burmeisteri*, i.e., the distinctin homo- and heterodimers. In **Chapter 7** we describe studies involving the shotgun cloning of the distinctin chains from the skin secretion. cDNA cloning surprisingly revealed that the chains A and B are encoded by separate mRNAs. Oxidation experiments with the synthetic chains did not show any preferential formation over the dimerization, suggesting either that there is an alternative mechanism for dimer formation *in vivo* or that synthetic chains have different behavior than native peptides.

The *P. burmeisteri* and *P. rohdei* peptides described in **Chapter 6** that contain 3 disulfide bonds as determined by the 2-D differential peptide display method are further characterized in **Chapter 8**. Several MS-based approach were used to *de novo* sequence these large 6 kDa peptides, including complementary CID and ETD peptide fragmentation and high resolution orbitrap detection. These peptides showed several conservative motifs of the Kazal peptides, a group of protease inhibitors. An interesting observation made during these analyses is that peptides modified with bromoethylamine show an much better ETD fragmentation pattern compared to the same peptides modified with iodoacetamide, which is the most commonly used reagent to block free cysteines in proteomics. Another important issue to mention is that for this, and all the other MS based methods described in this thesis, only little amount of material was used, which is crucial for frog skin researches on limited samples which are difficult to acquire. Moreover, the methodologies showed in this thesis can be applied to other peptidomic samples. Finally in **Chapter 9** all important findings done during this research are discussed, several concluding remarks are given and a future outlook is presented.

Curriculum Vitae

Geisa Paulino Caprini Evaristo was born on 26 May 1981, in Cachoeiro de Itapemirim, Espírito Santo State, Brazil, and resided with her parents in Iconha. In 1998 she completed her high school studies from Colégio Jesus Cristo Rei, in Cachoeiro de Itapemirim. In March 2000 she moved to Campos dos Goytacazes, Rio de Janeiro State to study Biology at State University of Norte Fluminense "Darcy Ribeiro", and in March 2004 she obtained the Bachelor's degree. After this she immediately started her Master's studies in Biosciences and Biotechnology at the same University and obtained the MSc. Degree in May 2007. From August 2004 until December 2006 she also worked as teacher in Biology and Nutrition courses at the São Camilo University Centre, in Cachoeiro de Itapemirim. From July 2005 until December 2007 she was Biology teacher at Cecierj Consortion, in São Francisco de Itabapoana, Rio de Janeiro State. In August 2007 she got a scholarship from the Brazilian National Council for Scientific and Technological Development (CNPq) to do her PhD at the Biotechnology Department of the Delft University of Technology under supervision of Prof. Peter Verhaert, in the Analytical Biotechnology group. In January 2008 she moved to the Netherlands to start her PhD and during her research she collaborated with the group of Prof. C. Shaw of the Laboratory of Natural Drug Discovery at the School of Pharmacy of Queen's University of Belfast (U.K.). In 2011 she won the Lake Louise Student Travel Award to attend the 24th Annual Tandem Mass Spectrometry Workshop, in Lake Louise, Canada. After her doctorate study she will return to Brazil to start a post-doc, as following her CNPq scholarship.

List of Publications

Schripsema, J.; **Caprini, G. P.**; Dagnino, D., Revision of the structures of citrifolinin A, citrifolinoside, yopaaoside A, yopaaoside B, and morindacin, iridoids from Morinda citrifolia L. and Morinda coreia Ham. *Org Lett* **2006,** 8, (23), 5337-40.

Schripsema, J.; **Caprini, G. P.**; van der Heijden, R.; Bino, R.; de Vos, R.; Dagnino, D., Iridoids from Pentas lanceolata. *J Nat Prod* **2007,** 70, (9), 1495-8.

Wang, L.; **Evaristo, G.**; Zhou, M.; Pinkse, M.; Wang, M.; Xu, Y.; Jiang, X.; Chen, T.; Rao, P.; Verhaert, P.; Shaw, C., Nigrocin-2 peptides from Chinese Odorrana frogs--integration of UPLC/MS/MS with molecular cloning in amphibian skin peptidome analysis. *Febs J* **2010,** 277, (6), 1519-31.

Rezende, M. E. ; Jasmim, J. M. ; **Caprini, G. P.** ; Sousa, E. F. ; Schripsema, Jan ; Thiebaut, J. T. L., Teor e composição química do óleo essencial de alpínia em razão da adubação e da disponibilidade de água no solo. *Revista Ceres* **2011,** v. 58, p. 208-215.

Evaristo, G. P. C.; Verhaert, P. D. E. M.; Pinkse, M. W. H., PTM-driven differential peptide display: Survey of peptides containing inter/intra-molecular disulfide bridges in frog venoms. *J Proteomics* **2012,** in press. (http://dx.doi.org/10.1016/j.jprot.2012.09.001)

Evaristo, G. P. C.; Pinkse, M. W. H.; Wang, L.; Zhou, M.; Wu, Y.; Wang, H.; Chen, T.; Shaw, C.; Verhaert, P. D. E. M., The chains of the heterodimeric amphibian skin antimicrobial peptide, distinctin, are encoded by separate messenger RNAs. *J Proteomics* **2012,** in press. (http://dx.doi.org/10.1016/j.jprot.2012.09.016)

Evaristo, G. P. C.; Pinkse, M. W. H.; Shaw, C.; Verhaert, P. D. E. M., Discovery of O-Glycosylated Peptides Secreted by the Skin of the Chinese Frog *Odorrana schmackeri. In preparation.*

Evaristo, G. P. C.; Pinkse, M. W. H.; Chen, T.; Mohammed, S.; Heck, J. R. A.; Shaw, C.; Verhaert, P. D. E. M., Kazal protease inhibitor-like peptides from *Phyllomedusa burmeisteri* elucidated by ETD/CID *de novo* sequencing. *In preparation.*

Evaristo, G. P. C.; Pinkse, M. W. H.; Verhaert, P. D. E. M., MS-based discovery of two novel peptides with homology to calcitonin-like peptides, from the skin secretion of the Chinese Frog, *Odorrana schmackeri*. *In preparation.*

Acknowledgements

Living in The Netherlands was a fantastic experience to me. I could not even imagine that such experience of living abroad would affect me so much. When I remember all the situations that I had I feel so happy. Most of them are related to my friends. Even the ones related to my bad moments remind me that I could overcome them and therefore I feel powerful, confident, bright, ..., happy. When I think of the special aspects that made me fall in love with this country, the first thing that comes to my mind is the feeling of freedom to go anywhere by bike at anytime. In my memories I see a movie of myself crossing beautiful landscapes, full of flowers and animals, either with my husband, friends or even alone. It was also a pleasure to me just to see the Dutch enjoying simple things, like swimming in the canals or enjoying an icecream in front of the canals, and also feeding the ducks and swans. The simple way of living also gets my admiration for this culture. Other aspect that I appreciate is the Dutch style of the buildings, streets and cities. I like the unique and romantic combination of old buildings with plenty of canals on the side of the streets. And I can not forget about the food, of course. There is a popular saying that the Dutch do not have a traditional food (or at least for warm meals). The fact is that I increased 8 kg during my first 6 months here. Stroopwafels, oliebollen, appeltaart and many types of breads, cheeses and chocolate (Belgian or not here they are), and more stroopwafels were my heaven (or my lapse). For all these good remembrances I want to thank the Dutch people, and one in special who encouraged me to come here to do my PhD, Prof. Jan Schripsema. You were totally right when you told me that I would like to live here.

I also would like to thank Prof. Jan Schripsema for helping me with the application on CNPq to this PhD project in Peter's group, and for following my development during all these years, with scientific support and advices.

I am also very grateful to Prof. Peter Verhaert for giving me the opportunity to do my doctorate in his group. Thank you for all the efforts and enthusiasm in my project, for triggering scientific discussions, and being friendly and comprehensive with my situations.

And I would like to express my deepest gratitude to Dr. Martijn Pinkse, for introducing me to the "world" of mass spectrometry of peptides, for guiding me, teaching me to be critic when I write articles, for helping me with the instruments and whenever I needed, specially for working hard with me until late in the evening and during his holidays. It was a

pleasure to work with such clever person. Thank you for the support, sharing your life experiences and the advices, I really appreciated.

I am very grateful with the CNPq, which was my sponsor for the four years of my PhD research and also the staff, in particular Ms. Sandra Jurema Santos, Mr. Belmiro Salles, Ms. Elza Baduino, Mr. José Gomes and Mr. José Airton, who helped me with all the burocratic issues related to my scholarship during my PhD.

In the Analytical Biotechnology group I would like to thank Herman and Inez for helping me to improve my English and for their friendship; Mervin for the good lunches full of laughs; Johan, Claire, Merve, Qiao and Yu for the nice chats and company; Emanuel for the nice time (and parties) we had in The Netherland and in Austria; and Eda for all the support and care. I am also thankful to my students, Yingying, Ricardo, Marleen, Jessica, Weiyi, Rick and Jahir, for their contribution and for their enthusiasm in the lab to achieve the best results.

I want to thank my collaborator Prof. Chris Shaw for his scientific support and giving me the opportunity to work in his group. It was a nice experience. I also want to thank his team, Dr. Tianbao Chen, Mei Zhou (Michele) and Lei Wang (Fiona), who taught and guided me on the Transcriptomic and bioassay analyses, and for giving me frog skin secretions to perform my first experiments in Delft. Many thanks to Cherry, Larry, Monica and all the other Chinese students who helped me to perform the experiments in Belfast and for the kind hospitality. It was a pleasure work with you.

I am grateful to our collaborator Prof. Celio Haddad, who indicated the endemic areas of interesting (highly poisonous) frogs in Brazil, who taught me how to hunt for amphibians in different environments and who helped to identify the frogs that we “milked”. It was an amazing experience to hunt with such an expert in anuran species. I still remember him vocalizing as a male *Hypsiboas faber*, when all of a sudden an aggressive frog jumped from branch to branch into his directions.

Another collaborator that was fundamental to the development of this project was Mr. Helimar Rabello, who promptly accepted to help this project by indicating the best places in the South of Espírito Santo State to find the frogs of interest. He also helped to arrange a group of people for the expeditions with some of his bachelor students to hunt for the frogs with me. Therefore, I also want to thank his students (Débora, Dinho and many others) for all their efforts, leaving your homes during your holiday time or weekend to go frog hunting during the night amongst the mosquitos, snakes and other wild dangerous animals.

Many thanks to Prof. Antonia Elenir Oliveira, from UENF, who allowed me to use her equipment to process my samples in Brazil. Special thanks to Simone who helped me to process them.

Special thanks for Prof. Patrick van Dijck and Helene Tournu, from the Catholic University of Leuven, who helped me with bioassays.

I would like to thank Linda and Laura, from Enzymology group, for the technical support, and Rosario for helping me to process transcriptome data.

During my staying in The Netherland I met special friends and I would like to thank them for all their care and love. Lygia, thanks for the nice trips, holidays and the scientific discussions. Eda, thanks for your sincerity and support. And Michaela, thanks for your support, sharing your bright ideas, motivation, enthusiasm, friendly advices and the nice time together enjoying a drink in Delft (and of course for shopping on Friday evenings). Thanks also to Eda and Michaela for accepting being my paranimfs.

Agora eu gostaria de agradecer as pessoas mais importantes da minha vida, a minha família, por todo apoio, carinho e compreensão. Aos meus irmãos, Jocemar, Jean e Germana, pela confiança. Aos meu sobrinhos, Gabriele e Ian, pela paciência da tia não poder acompanhar de perto o seu crescimento. Às minhas cunhadas, Juliana, Isabela e Rosana, cunhados, Thiago e Cássio, e enteada, Hellory, pelo carinho e entusiamo. À minha sogra querida, Jussara, pela confiança, apoio, suporte e amor. E aos meus pais, Jucélia e José Antônio, por toda a confiança, dedicação e amor. Amo todos vocês.

And finally I would like to thank my dear husband, Joseph, for encouraging me to follow my dreams, for all the patience waiting 2 years to be together, for leaving everything behind and joining me in The Netherlands, and for all your love. I love you.

Geisa