PART I

STUDIES OF NATURALLY OCCURRING PEPTIDES

1

ANALYSIS OF THE PEPTIDOMES OF AMPHIBIAN SKIN GRANULAR GLAND SECRETIONS—AN INTEGRATED FUNCTIONAL GENOMIC STRATEGY

CHRIS SHAW AND TIANBAO CHEN

1.1 INTRODUCTION

Amphibians possess an extraordinary anti-predator defence mechanism based upon the stress-induced secretion of noxious and often highly-toxic chemical cocktails from specialised structures, the granular glands, usually located in discrete regions of the tegument. Although the chemical composition of such secretions can vary widely between different taxa, peptides represent the major class of biochemicals in the majority of species studied thus far with the exception of the dendrobatids, the so-called arrow poison frogs of the Neotropics. With several thousand species secreting peptidomes that may contain several hundred different peptides, the potential natural library of novel actives is sizeable requiring the application of modern bioanalytical technologies to the tasks of data acquisition, warehousing and mining. Superimposed upon these problems are the perhaps more difficult issues that gravitate around eco- and etho-friendly sample acquisition from species that are often extremely difficult to obtain or even encounter and that are under considerable survival threats on a global scale. For several years our research

Peptidomics Methods and Applications, Edited by Mikhail Soloviev, Chris Shaw and Per Andrén

has been directed to solving these problems, both technically and ethically, in a manner that integrates modern technologies with non-invasive, non-destructive sample acquisition and that does not compromise, in any way, the robustness of the scientific data produced. The integrated functional genomic strategy that has resulted from this research stands as a model for unravelling, in a holistic manner, the pharmacological/physiological roles of bioactive peptides that represent the "smart" weaponry encoded by fit-for-purpose evolved genes within the Amphibia.

Amphibian skin has been known to possess potent pharmacological properties for millennia having been utilized in many ways, from traditional ethnic medicines through tipping of arrows for hunting purposes to the more sinister applications in poisons and potions [1,4,5,26,28,29,33]. Even today, the use of amphibian skin *per se* or secretions derived from the specialized dermal granular glands are used widely across the world especially in Asia and South America [1,4,5,28,29,33]. Amphibians, as a group of animals, are generally rather small, soft-skinned, and slow in movement that renders them a prime source of food for many predatory species among the fish, reptiles, birds, and mammals. In the absence of surface armor, spines, claws, stings, or fangs, the amphibians have developed a highly sophisticated chemical defense against predation in their skin secretions that ranges from the noxious and bitter tasting at one end of the scale to the self-permeabilizing and highly toxic at the other [3,23,24,26,29]. The latter strategy is exemplified by the arrow-poison frogs of the family Dendrobatidae, from the New World [23–25]. Dendrobatid frog skin secretions are rich in alkaloid toxins, that, current thought believes, are derived and sequestered from forest invertebrates such as ants, which constitute, for the most part, the diet of these tiny and often brightly colored frogs [25,31,32]. The skin secretions of typical toads are likewise rich in small alkaloid molecules such as bufodienolides, bufotoxins, and biogenic amines, but this case, such molecules are *de novo* synthesized within the dermal venom glands that are often concentrated behind the eyes and on the back legs of the animals themselves [23,24]. The vast majority of other frog species and some groups of toads produce, as a result of endogenous gene expression, a highly complex cocktail of bioactive peptides and their biosynthetic precursor-related fragments when subjected to stress such as predator attack [3,26,29]. Typically, there may be as many as 100 different peptides with some species producing as many as 600. This obviously presents the biochemist with a formidable problem in terms of structural analysis and identification even with state-of-the-art analytical hardware and software, and if this is considered as a major barrier to overcome, it is nothing when compared to the establishment of individual biological functions. These barriers to understanding thus necessitate novel approaches that integrate existing elements of analytical strategy with new modifications and applications, and it is the purpose of this chapter to delineate how we have achieved this in part to characterize the peptidomes of amphibian defensive skin secretions.

1.2 HISTORICAL PERSPECTIVE

Amphibian skin peptide research owes its origins to a superlative Italian scientist, Vittorio Erspamer, who was a classical pharmacologist by training [26]. He was intrigued by the obvious biological effects of both invertebrate and vertebrate venoms and was particularly impressed by the effect that the dermal venom of the common toad (*Bufo bufo*) had upon his dogs when they made a wrong decision to attack such a toad in his garden. This observation by a curious pharmacologist opened a treasure chest of molecular discovery that spanned the best part of half a century of productive ground breaking research. Many early works were focused on the characterization of the biogenic amines and alkaloids present in toad and salamander venoms, but a second chance observation on frog skin extracts led to the discovery of the defensive secretion peptides. Unlike extracts of toad skin, extracts of frog skin that were subjected to protease digestion lost much of their activity, indicating that the endogenous actives were of a proteinaceous character. In the early days of the evolution of chromatographic and electrophoretic techniques, this assertion was confirmed by the use of thin-layer and gel permeation chromatography [27]. Erspamer's method was for the most part bioassay driven, in that using a series of defined bioassays that included the use of a variety of smooth muscle preparations, actives were identified and purified to homogeneity using the standards of the time [26]. The starting material for these experiments was usually a methanolic extract of air-dried skins that were obtained in the field after sacrifice of often many hundreds if not thousands of specimens of a particular species (phyllokinin—(bradykinyl-IY sulfate) originally purified from 3555 skins of *Phyllomedusa rhodei*) [2]. While this approach was necessary to produce sufficient material for primary structural analysis and pharmacological workups, remembering that in the early 1960s manual sequencing was the order of the day and peptide synthesis was in its infancy [2], it was highly destructive of animal life and produced an extract that included many elements not derived from the granular gland secretion but rather from additional skin tissues, blood, nerves, muscle, and so on. So while this approach worked well in the beginning with an identifiable bioactivity, it was most inefficient, time-consuming, and would be considered unethical by contemporary standards. For the purposes of contemporary peptidome research, we have the benefits of much more highly advanced analytical technologies, and we are in possession of a greater body of knowledge and an ethical code in the use of live animals for research that can be integrated in a way that results in more rapid, complete, and robust data generation in a manner that is noninvasive and nondestructive to the donor animals [35]. A "classically" prepared skin extract would contain many thousands if not tens of thousands of peptides/proteins that are not components of the granular glands, and hence most of the investigators' valuable time and resources would be spent in structurally characterizing irrelevant molecules. This has to be overcome to enable a focus on the granular gland secretions themselves.

1.3 CONTEMPORARY METHODS OF SAMPLE ACQUISITION

Since the peptidome of the granular gland is the focused subject of study and analysis, methods of acquisition should ideally provide as pure a sample of the secretion of these structures as possible and ideally without significant damage or mortality to the donor. As the physiological control system for secretion in most species of amphibian relies on application of a stressor [26], the most user-friendly means of applying this has been investigated. Specimens in the field and those recently brought into captivity have a low threshold for such species but some adapt very rapidly to human handling and become readily tame such that more extreme stressors have to be applied after time to acquire defensive secretion. Some species of frog and toad are ideal candidates for study in that they secrete with the mildest of stresses that can be as mundane as lifting from their vivaria while others appear to be resistant to all but one might consider a severe stress. The granular glands are constructed in such a way that they are lined by myoepithelial cells that contract in the presence of stress-related catecholamines, such as norepinephrine, that are located in intrinsic nerve terminals [26,29,34]. This process can thus be mimicked by injection of the catecholamine directly into dorsal lymph sacs after which copious secretion is released from the granular glands [26,29,34]. This technique is reproducible and reliable but involves hypodermic injection of a substance whose effects will be systemic with possible side effects following prolonged usage. Another method involves placing the amphibian into a closed space with ether fumes acting as the stressor [26]. This requires the use of an inflammable and potentially explosive organic solvent that may cause harm to the specimens following prolonged usage. Perhaps the most efficient, reproducible, and harmless method is transdermal electrical stimulation as first described by Tyler et al. [35]. This technique applies a mild electrical stimulus (equivalent to two personal stereo batteries) via platinum electrodes to the moistened skin of the amphibian. Secretion is almost instant, and the whole procedure lasts not more than 1 min in total. It is sufficient to contract the myoepithelial cells and cause granular gland discharge with minimal stress to the whole animal, and it is very well tolerated with frogs even spawning shortly afterward in some cases. The procedure is highly reproducible and can be carried out at monthly intervals on the same individuals. As the stimulator is battery-powered, the technique is amenable to use in the field. Secretion is washed from the skin surface using a stream of deionized water and collected into a chilled glass beaker. The washings are snap-frozen in liquid nitrogen and lyophilized. Lyophilized material is stable indefinitely if stored under vacuum in a freezer at $-20^{\circ}C$ [6–22,38–41]. If secretions are collected in the field, then addition of trifluoroacetic acid to an approximate final concentration of 0.1% (v/v) is sufficient to stabilize the peptidome for subsequent analysis although lyophilization at the earliest opportunity is recommended [6–22,38–41]. Using this technique, the investigator obtains as complete a granular gland secretion as possible (minus volatiles) permitting objective

peptidomic analysis with minimal contamination of other skin tissue peptides/proteins.

1.4 THE INTEGRATED FUNCTIONAL GENOMIC STRATEGY

This may seem to be somewhat aberrant in a text that focuses on peptidomics, but biochemical analysis should always aim to procure the maximum amount of molecular information that is possible from a given biological sample. This is not a fanciful objective but rather can aid the researcher in producing a holistic understanding of the entire biological system from gene through protein/peptide through action—a necessary prerequisite in assessing academic, therapeutic, or biotechnological potential of new bioactive peptides. Our research on bioactive peptides from amphibian skin basically employs a three-prong approach with entry into the system being possible from each point. *The bioassay approach* duplicates the original method but tests individual fractions from reverse-phase HPLC fractionated skin secretion rather than crude material as a first pass. Positive fractions are repeated at least in triplicate to confirm effects following which further fractionation is performed if required to a homogeneous peptide, leading ultimately to structural characterization. *The peptidomic approach* treats the secretion sample like any other complex biological material and separates the components by tandem reversed-phase HPLC/electrospray MS or by off-line reversed-phase HPLC followed by MALDI-TOF analysis of individual fractions. In essence, to provide a more complete and readily accessible data set, both techniques are usually performed. Primary structures of peptides can be determined using MS/MS fragmentation with either online trawling of public databases to establish identity or *de novo* sequencing where this is not possible. Herein lies the major challenge that perhaps illustrates the major difference in peptidome analysis from exotic species rather than a peptidomic or indeed proteomic study from human, rat, mouse, chicken, *Xenopus*, zebrafish, and so on, whose genomes and hence translated potential proteomes have been fully or nearly fully sequenced. For the vast majority of species that we study, there is no relevant archive of either gene or protein sequence information publically available to provide a high throughput format that other proteome/peptidome researchers essentially take for granted for their "model" organisms. For this reason, we have to be particularly creative in problem solving to this end, and the finer points of this approach will be explained in greater detail in a following section. *The genomic approach* is one that is not new in this discrete field, but rather we employ this in a different manner to predecessors. In previous studies, the messenger RNAs encoding the biosynthetic precursors of novel peptides identified in amphibian skin secretions have been cloned from cDNA libraries manufactured from skin taken immediately post mortem from sacrificed specimens [37]. So no matter what good intentions had been met by obtaining the defensive secretion itself

for peptidome analysis, the amphibians were sacrificed in the quest for additional information on genetic coding, precursor organization, propeptide convertase sites, coencoded peptides, and so on. This invariably followed the acquisition of full primary structure of the peptide usually by automated Edman degradation. These data facilitated the generation of degenerate primer sets for the purpose of specifically cloning the precursor of the peptide following PCR amplification of its transcript. Following a brainstorming session over a Monday morning coffee, we had the idea that based on the syncytial nature of the granular gland and our *in cerebro* visualization of its stress-induced rupture and expulsion of entire contents, cytosolic peptide precursor-encoding polyadenylated mRNA might just be present in the skin secretion itself in a quantity and format that might facilitate PCR. The idea was correct. We first cloned from freshly extruded secretion, then lyophilized secretion, then lyophilized skin, and finally from air-dried skin [6–22,38–41]. The mechanism of polyadenylated mRNA protection within the amphibian skin has been proven to be very robust and essentially afforded by electrostatic interaction with the abundant cationic amphipathic peptides present in the secretions [22]. More recently, we have shown that there is sufficient DNA within the lyophilized secretion that after amplification can provide a template for specific gene structural mapping [12]. These discoveries on nucleic acid presence within the secretions meant that a holistic functional genomic study could readily be performed on a single, noninvasively obtained skin secretion sample: Bioactivity from reversed-phase HPLC fractions and establishment of primary structures of both the general peptidome and the specific active, cloning of precursor-encoding cDNA, and mapping of gene organization—all from as little as 10 mg of lyophilized secretion.

1.5 HOW THEN CAN ALL THREE APPROACHES BE INTEGRATED?

The bioassay approach gravitates around a whole organism, tissue, or cell that provides a defined target whose interaction with an active molecule in the skin secretion produces a defined and objectively measurable response. Indeed, the assay may be completely reductionist in that the target may simply be another biomolecule, such as a protease if one seeks inhibitors or naked DNA if one seeks nucleic acid interacting peptides. This approach requires no knowledge of the chemical structure of the active in the first instance. Even large pharma, with their considerable financial and human resources, are limited in the number of targets that molecular libraries can be screened against in a high throughput fashion and without doubt many have large numbers of potential actives/drug leads that remain undisclosed because active and appropriate targets have never been personally introduced. A relatively small research team thus has a discovery problem of several orders of magnitude more problematic than the one that necessitates a critical focus in target selection

with only a low-to-moderate throughput. Costs can be kept low if cell lines displaying the selected target are employed or lower still if target molecules themselves can be arrayed. The advantage of the bioassay approach is that the biological end point act as the selection system for active peptides within the fractionated cocktail with no *a priori* weighting on primary structure—a point that will become more apparent when sequencing of the peptidome produces peptides of novel structure with no obvious homologs in online databases. The bioassay data, within the previously stated limitations, will provide the investigator with a bioactivity that can then be used as a monitor for purification to homogeneity and ultimately to establishment of full primary structure. Once this has been achieved for a novel peptide, then by the employment of considered degenerate primer design, the cDNA encoding the biosynthetic precursor of the peptide can be cloned and its full primary structure deduced from the specific nucleic acid template. Thus, the sequence of events using the bioassay as the initiation point of discovery of a novel peptide is activity, purification, establishment of primary structure, degenerate primer design, cloning of biosynthetic precursor- encoding cDNA, and deduction of full primary structure. appropriate target has been selected and presented in a suitable format; then one can turn attention to procurement of sufficient fractionated samples for the screening experiments.

The peptidomic approach requires no prior knowledge of bioactivity but rather involves the systematic sequencing of peptides following reversed-phase HPLC fractionation of crude skin secretion. If automated Edman degradation of individual peptides is the chosen approach, then HPLC gradients are chosen that are long and gentle such that the majority of peptides are resolved into individual fractions. Alternatively, peptides that elute very closely following the initial separation can be subsequently resolved using different column chemistries. Using this approach, chromatography is usually performed off-line, and each fraction is subjected to mass analysis using MALDI-TOF mass spectrometry. This provides data both on molecular mass with single and doubly charged ions predominating and on purity/homogeneity of the peptide in the individual fraction. The observed molecular mass should correspond to the molecular mass deduced from the Edman degradation data set. One of the advantages of this approach is in the positional identification of stable posttranslationally modified residues whose elution characteristics have been determined for most modern sequencers. In addition, Leu/Ile assignations can be made unequivocally as can Lys/Gln and Leu/Ile/Hyp that may be difficult using dated mass spectrometric hardware. However, the contemporary approach is to perform reversed-phase HPLC/electrospray MS for complete data-dependent acquisition that provides MS/MS fragmentation data for each identified doubly charged ion. More recent advances in mass spectrometric hardware combine the best attributes of several technologies with MALDI sample excitation and delivery of ions into an ion-trap configuration that permits subsequent MS^n analyses. The problems of analysis of the peptidome of a hitherto unstudied species skin

secretion begin to arise at this point. The conventional downstream bioinformatic applications that involve matching these generated MS/MS spectra with peptides or domains of proteins are generally of little use at this stage as they contain limited relevant entries and of course, no entries for entirely novel peptides. Peptides such as bradykinins, which are well represented in amphibian skin secretions and online databases, terminate (like typical tryptic peptides) and often initiate with arginyl residues that favor the production of good fragmentation spectra with often successful hits by this method. However, the vast majority of amphibian skin peptides do not terminate in arginyl residues and often contain multiple prolyl residues, blocked N-terminals, and amidated C-terminals rendering the acquisition of appropriate MS/MS spectra and their interpretation very difficult. Thus, while conventional contemporary "proteomic/peptidomic" hardware can readily generate MS/MS spectra of amphibian skin secretion peptide libraries in a high throughput mode, their quality, interpretation, and subsequent unequivocal primary structural assignments are fraught with difficulties for the reasons specified. However, at this point, we will proceed with those for which primary structures have been established by this method and will deal later with those that have not or cannot be structurally elucidated. The primary structures of novel peptides can be unequivocally confirmed by molecular cloning of the biosynthetic precursor. Using the assigned primary structure, a suitable degenerate primer is designed and used to amplify the corresponding transcript from a skin-secretion-derived cDNA library by 3′-RACE. When such product(s) are cloned and sequenced and found to be correct, then a gene-specific antisense primer (usually to a region of the 3′-noncoding region) can be synthesized to effect 5′-RACE and thus total open-reading frame deduction. The assigned primary structure of the novel peptide should be located as a contiguous sequence located between known propeptide convertase processing sites [6–10,13,15,18,36,38–41]. However, caution should be exercised at this point to account for new and hitherto undescribed processing sites and potential splice variation/heterogeneity of encoding mRNAs, both of which our research group has already encountered [6–10,13,15,18,36,38–41]. Our experience with the "so-called" *de novo* sequencing has been mixed. Software program claiming to perform this black art are in a constant state of development but a research program such as ours, where many months of work could be carried out on a synthetic replicate of a *de novo* sequenced peptide that proves to be wrong, could be catastrophic in terms of wasted resource in the form of time, money, and externally funded, time-limited research grant or postgraduate studentship. We prefer to leave this approach to the seasoned gamblers in the research community until such times when we are confident in the assertions and claims of the software engineers. To summarize, novel peptides structurally assigned by Edman degradation can have structures confirmed by molecular cloning of biosynthetic precursor prior to chemical synthesis for the purpose of deduction of biological activity. In some instances, the model to test

bioactivity of our novel peptide can be selected on the basis of structural similarities to other known members of a structurally homologous peptide "family," while in many instances, the *a priori* selected range of bioassays available within one's own research group dictated by specific interests represents a first port of call. It is here that a wide spectrum of interested collaborators and their particular bioassays proves invaluable as many are amenable to adding a few unknowns to their routine systems with the potential of a major payoff as coauthors on the publication of the discovery of a new biological entity. A word of warning, however, is that the true discovery game at the frontier of human knowledge, such as the establishment of the discrete bioactivity of a novel peptide, may occur serendipitously, or after a hunch/vision/inspired thought/dream or may take more years than the allotted span of the discoverer. Figures 1.1–1.5 and Table 1.1, reprinted with kind permission from Elsevier from Reference 17, illustrate the data obtained via a peptidomic approach using information previously generated by a bioassay approach.

The genomic approach has been alluded to in the previous approaches, but its power is much greater than simply designing degenerate primers to novel peptide primary structures and cloning homologous biosynthetic precursors. The discovery by our research group that amphibian granular gland transcriptomes persist in stimulated skin secretions was a milestone in this particular research field and opened up new possibilities in high throughput peptidome analysis and characterization within this model [6–22,38–41]. Accumulated data from several

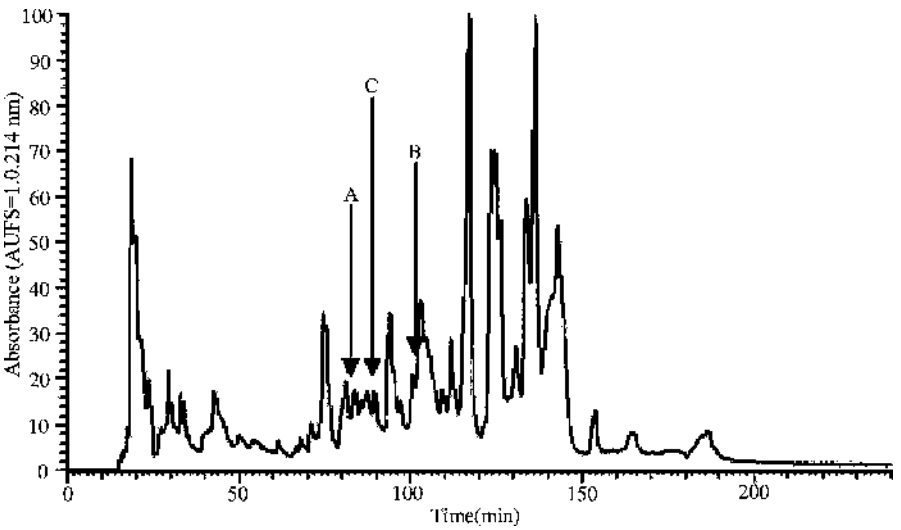

FIGURE 1.1 Reversed-phase HPLC chromatogram of *Phyllomedusa sauvagei* skin secretion. Fractions containing (**a**) bradykinins, (**b**) phyllokinins and (**c**) (Val^1, Thr^6)-bradykinin are indicated by arrows. Reprinted from Reference 17 with permission from Elsevier.

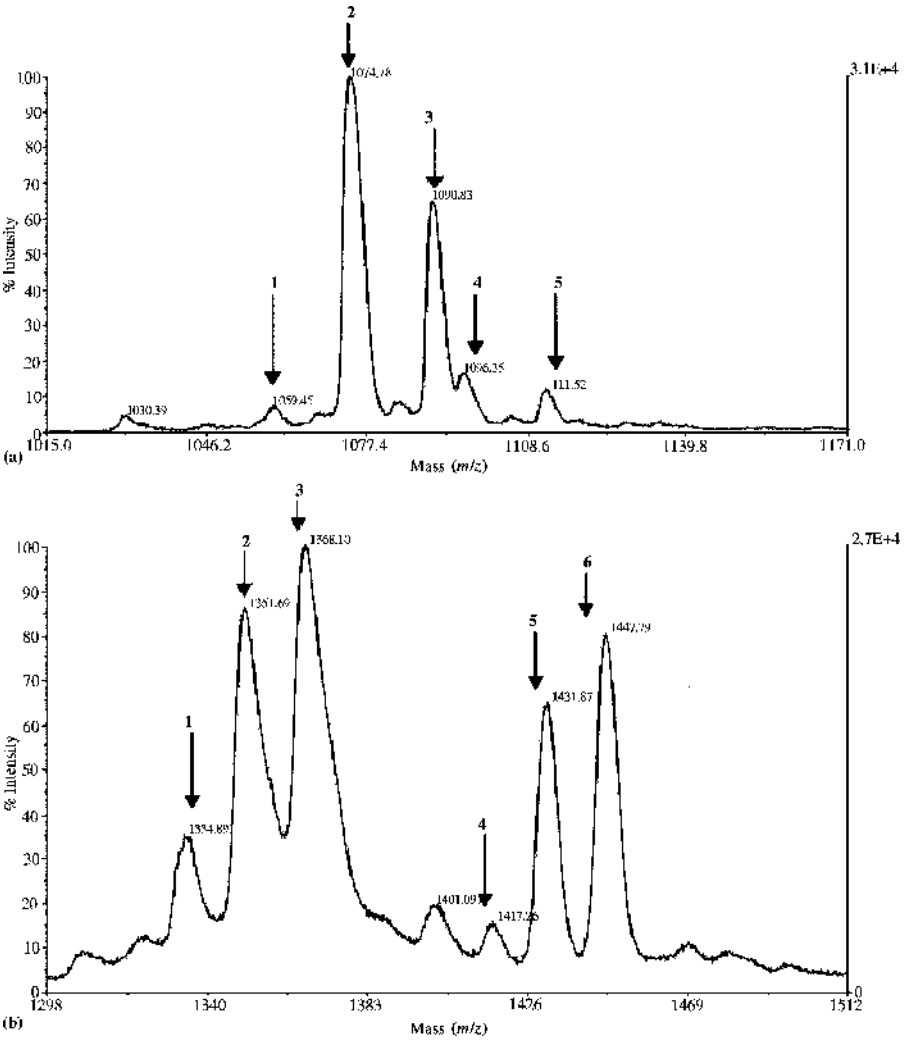

FIGURE 1.2 MALDI-TOF mass spectrograms of fraction 83 (**a**) containing 1-bradykinin, 2-(Thr^6)-bradykinin, 3-(Hyp^3, Thr^6)-bradykinin, 4-(Thr^6)-bradykinin sodium adduct, 5-(Thr^6)-bradykinin potassium adduct and fraction 100; (**b**) containing 1-phyllokinin, 2-(Thr^6)-phyllokinin, 3-(Hyp^3, Thr^6)-phyllokinin. 4-sulfated phyllokinin, 5-sulfated (Thr^6)-phyllokinin, 6-sulfated (Hyp^3, Thr^6)-phyllokinin. Reprinted from Reference 17 with permission from Elsevier.

research groups in the field has established that the organization of amphibian skin peptide biosynthetic precursors is both highly ordered and highly conserved between different taxa [6–10,13,15,18,36,38–41]. In general, the hydrophobic amino acid residue rich signal peptide (these are secreted products) is followed by a region rich in acidic amino acid residues, a typical dibasic (-Lys-Arg-)

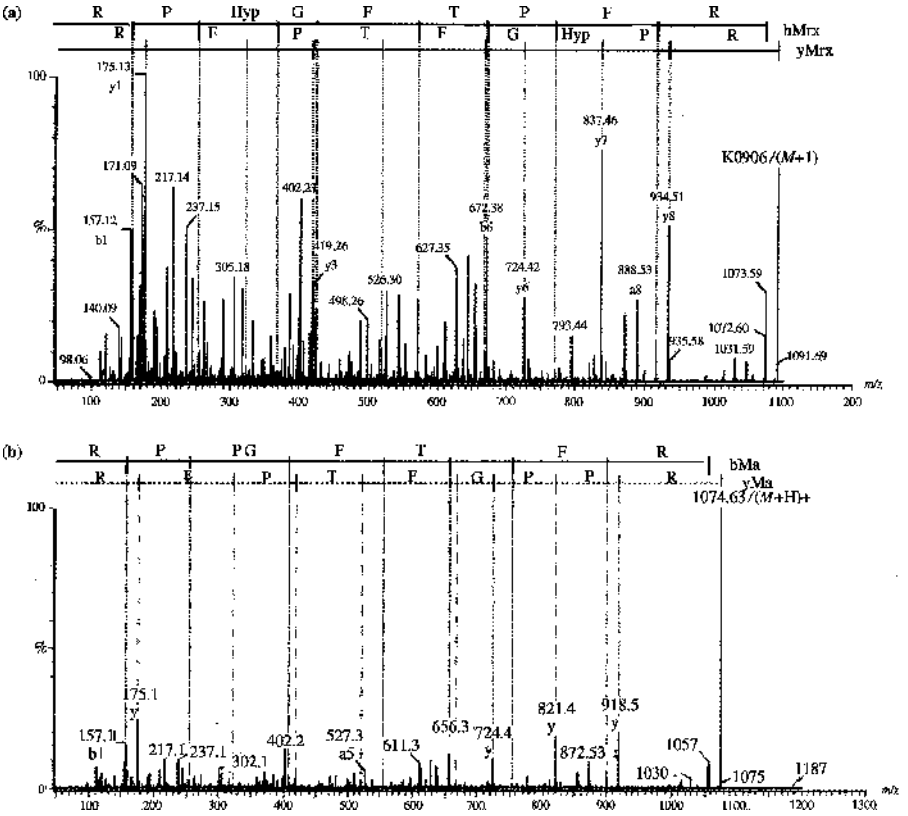

FIGURE 1.3 QTOF MS/MS fragmentation spectra and associated b- and y-ion sequence assignations of (**a**) (Hyp^3, Thr^6)-bradykinin and (**b**) (Thr^6)-bradykinin. Reprinted from Reference 17 with permission from Elsevier.

TABLE 1.1 Bradykinins and Phyllokinins Identified in the Skin Secretion of *P. Sauvagei.*

Bradykinin	**Arg**-Pro-**Pro**-Gly-Phe-**Ser**-Pro-Phe-Arg
(Thr^6)-bradykinin	Arg-**Pro**-Pro-**Gly**-**Phe**-Thr-**Pro**-**Phe**-**Arg**
(Val^1, Thr^6)-bradykinin	Val-**Pro**-Pro-**Gly**-**Phe**-Thr-**Pro**-**Phe**-**Arg**
(Hyp^3, Thr^6)-bradykinin	Arg-**Pro**-Hyp-**Gly**-**Phe**-Thr-**Pro**-**Phe**-**Arg**
Phyllokinin	Arg-**Pro**-Pro-**Gly**-**Phe**-Ser-**Pro**-**Phe**-**Arg**-Ile-Tyr
(Thr^6)-phyllokinin	Arg-**Pro**-Pro-**Gly**-**Phe**-Thr-**Pro**-**Phe**-**Arg**-Ile-Tyr
(Hyp^3, Thr^6)-phyllokinin	Arg-**Pro**-Hyp-**Gly**-**Phe**-Thr-**Pro**-**Phe**-**Arg**-Ile-Tyr
(Thr^6)-phyllokinin sulfated	Arg-**Pro**-Pro-**Gly**-**Phe**-Thr-**Pro**-**Phe**-**Arg**-Ile-Tyr (SO_3H)
(Hyp^3, Thr^6)-phyllokinin sulfated	Arg-**Pro**-Hyp-**Gly**-**Phe**-Thr-**Pro**-**Phe**-**Arg**-Ile-Tyr (SO_3H)

Fully conserved residues in bold type.
Reprinted from Reference 17 with permission from Elsevier.

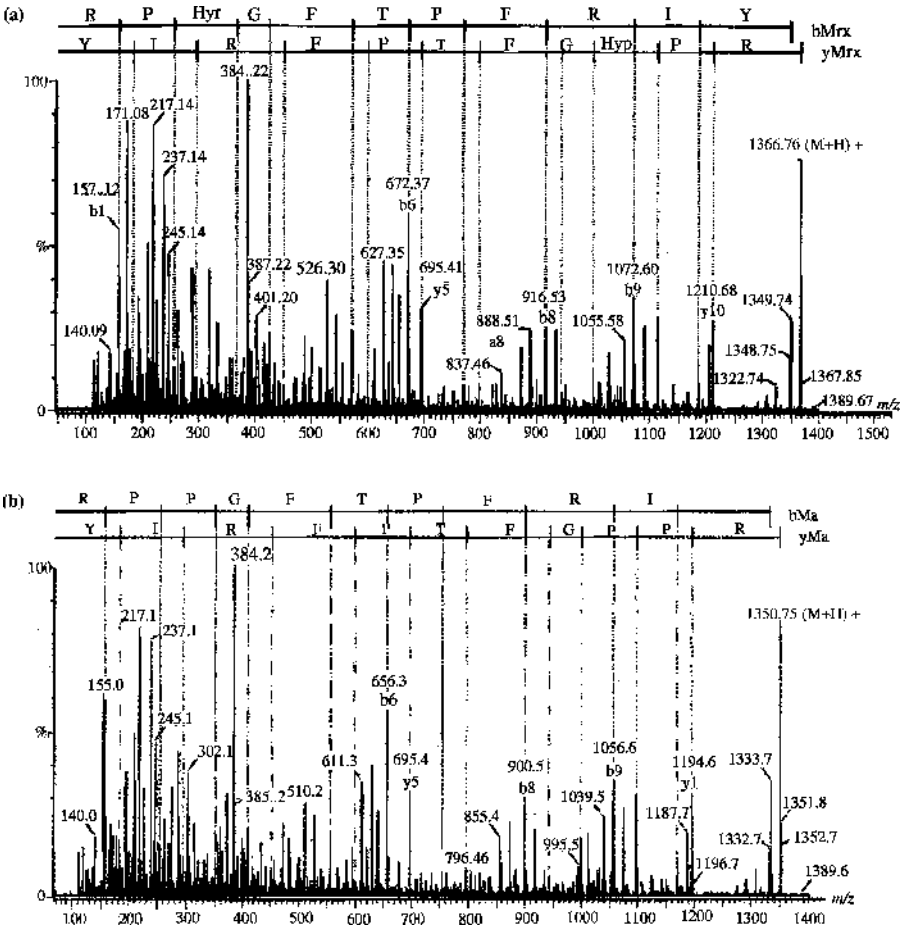

FIGURE 1.4 QTOF MS/MS fragmentation spectra and associated b- and y-ion sequence assignations of (**a**) (Hyp^3, Thr^6)-phyllokinin and (**b**) (Thr^6)-phyllokinin. Reprinted from Reference 17 with permission from Elsevier.

propeptide convertase processing sites and finally a single copy of bioactive peptide [6,7,9,10,15,18,36,38–41]. This is of course a general scheme, and many variations exist such as processing sites within the acidic domain, C-terminal, extensions following the mature bioactive peptide, glycine amide donors flanking the C-terminus, and tandem repeat sequences found in the biosynthetic precursors of skin opiate peptides and bradykinins [21,30]. However, alignment of either nucleic acid sequences or open-reading frame amino acid sequences of cloned biosynthetic precursor transcripts from many species of amphibians, reveals regions of both highly conserved nucleic acid and amino acid sequence specifically within signal peptide domains but also nucleic acid sequence domains within both 3′- and 5′-nontranslated regions. This has permitted design of general PCR primer

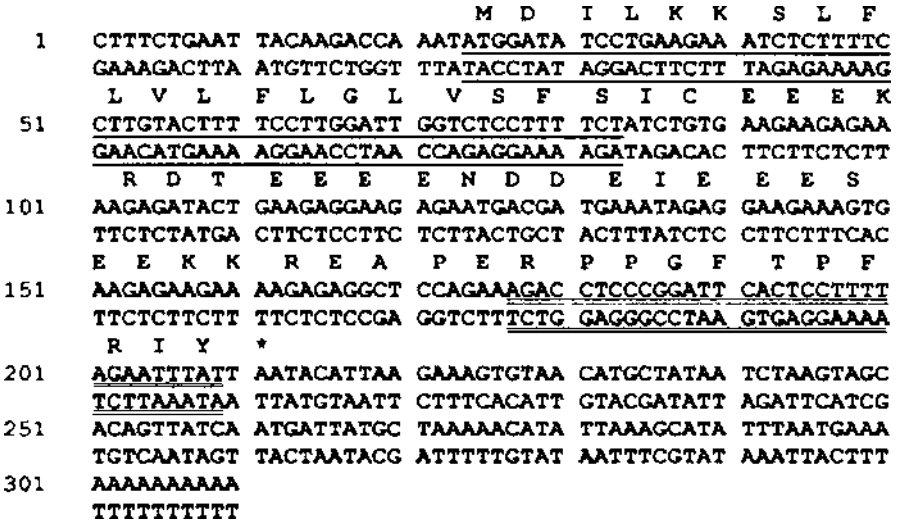

FIGURE 1.5 Nucleotide sequence of full-length cDNA-(both strands) encoding (Thr^6) bradykinin/phyllokinin from *P. sauvagei* skin secretion. The putative signal peptide is single-underlined and the (Thr^6)-bradykinin/phyllokinin sequence, located at the C-terminus of the open-reading frame, is double-underlined. The stop codon is indicated by an asterisk. Reprinted from Reference 17 with permission from Elsevier.

sets that are either class, family, or genus specific within the amphibians. Thus, armed with a skin-secretion-derived cDNA library from an unstudied species and a general primer set, one can "shotgun" amplify and clone secretion peptide transcripts in the total absence of any specific peptide primary structural data. This technique has been both tried and tested as evidenced by many published studies [6,7,9,10,15,18,38–41]. In some instances, as many as 23 different peptide transcripts have been cloned in the first pass of sequencing just 50 clones. With this information, additional primers can be designed to conserve nucleotide sequence domains within the particular species under study that, when used in repeat PCR experiments, result in additional novel transcripts being cloned. The putative primary structures and molecular masses of encoded peptides can then be deduced tentatively from precursor open-reading frames, and interrogation of archive mass spectrometric data sets can confirm the presence of an isobaric peptide in chromatographic fractions of the same sample of skin secretion. The peptide can then either be subjected to automated Edman degradation or preferably by MS/MS fragmentation using the primary structure as an internal reference for unequivocal identification. When large data sets of deduced peptides arise from the shotgun cloning approach, then these can be used to construct an in-house customized FASTA formatted database for rapid confirmation structures following reference to mass spectrometric data—the two approaches thus integrate extremely well in this respect. Alternatively, entire open-reading frame structural data can be used as a database, and this serves, following

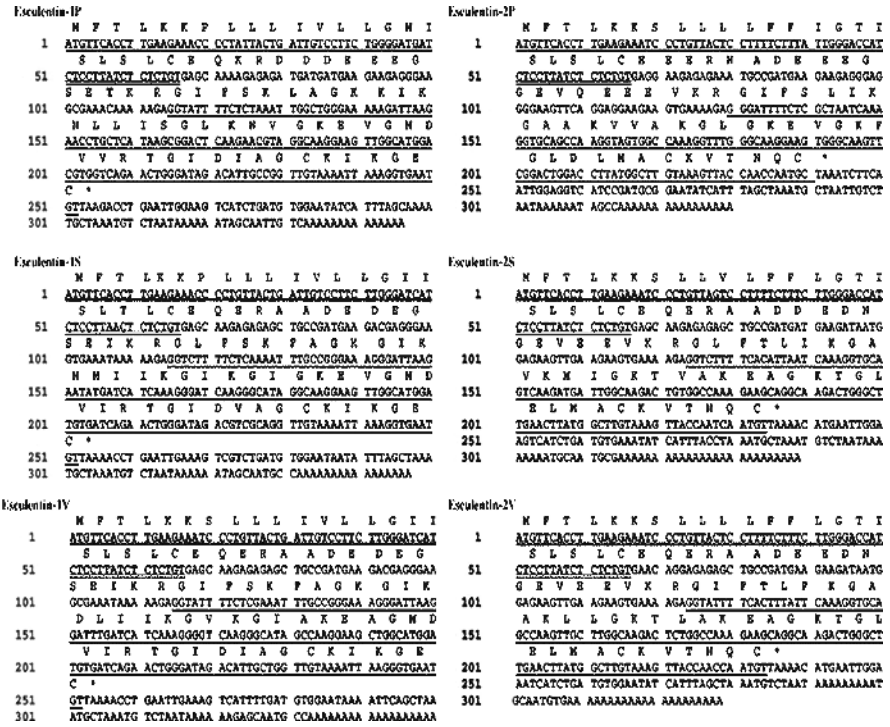

FIGURE 1.6 Translated nucleic acid sequences of cloned cDNAs encoding frog skin antimicrobial peptides—esculentins 1P, 1S, 1V, 2P, 2S, and 2V. Putative signal peptides are double-underlined, mature peptides are single-underlined, and stop codons are indicated by asterisks. Reprinted from Reference 9 with permission from Elsevier.

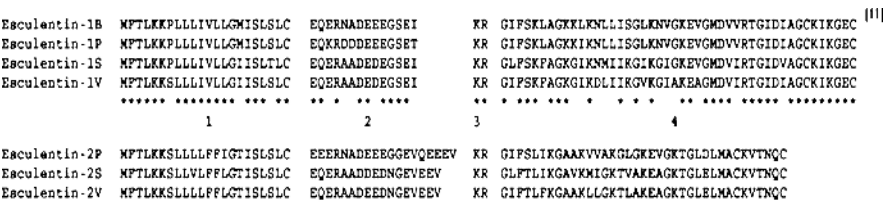

FIGURE 1.7 Domain architecture of preproesculentins. (1) Putative signal peptide, (2) acidic "spacer" peptide, (3) KR (Lys-Arg) propeptide convertase processing site, and (4) Mature peptide. Fully conserved amino acid residues within each group are indicated by asterisks. Esculentin 1B from *R. esculenta*—the only cloned esculentin precursor for comparison. Reprinted from Reference 9 with permission from Elsevier.

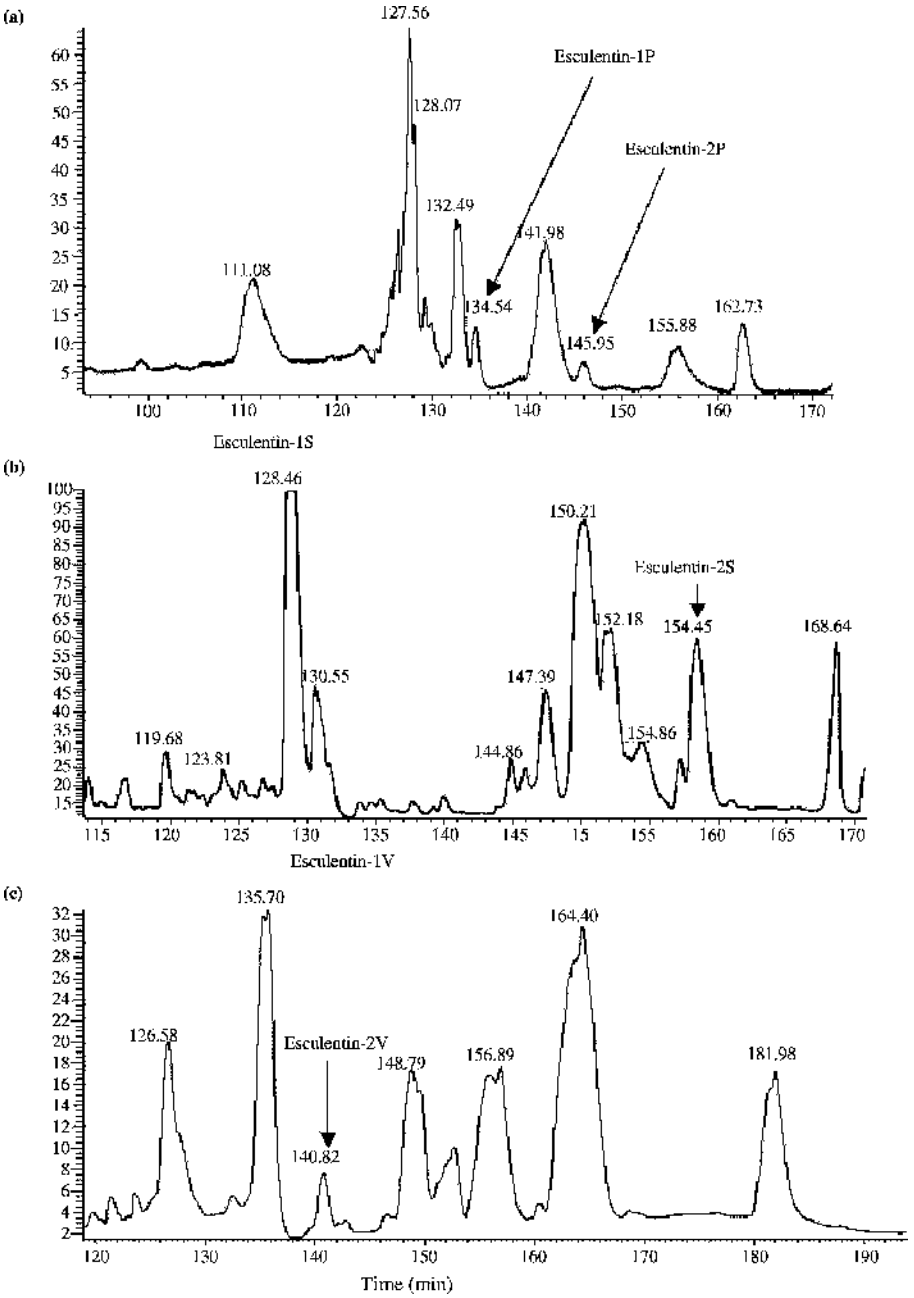

FIGURE 1.8 Regions of reverse phase HPLC chromatograms of respective frog skin secretions, (**a**) *R. plancyi fukienensis,* (**b**) *R. schmackeri,* and (**c**) *R. versabilis,* indicating the retention times (arrows) of individual esculentins 1 and 2. Reprinted from Reference 9 with permission from Elsevier.

		Observed mass	Calculated Mass
Esculentin-1P	GIFSKLAGKKIKNLLISGLKNVGKEVGMDVVRTGIDIAGCKIKGEC	4800.21	4798.79
Esculentin-1S	GLFSKFAGKGIKNMIIKGIKGIGKEVGMDVIRTGIDVAGCKIKGEC	4778.36	4777.79
Esculentin-1V	GIFSKFAGKGIKDLIIKGVKGIAKEAGMDVIRTGIDIAGCKIKGEC	4747.97	4746.71
Esculentin-2P	GIFSLIKGAAKVVAKGLGKEVGKFGLDLMA---------CKVTNQC	3792.12	3791.59
Esculentin-2S	GLFTLIKGAVKMIGKTVAKEAGKTGLELMA---------CKVTNQC	3850.05	3849.69
Esculentin-2V	GIFTLFKGAAKLLGKTLAKEAGKTGLELMA---------CKVTNQC	3852.19	3851.64

FIGURE 1.9 Primary structures of esculentins established by a combination of automated Edman degradation and mass spectrometry. The observed and calculated masses of mature peptides from respective secretion fractions are indicated. Reprinted from Reference 9 with permission from Elsevier.

```
Esculentin-1A   GIFSKLAGKKIKNLLISGLKNVGKEVGMDVVRTGIDIAGCKIKGEC
Esculentin-1B   GIFSKLAGKKLKNLLISGLKNVGKEVGMDVVRTGIDIAGCKIKGEC

Esculentin-1P   GIFSKLAGKKIKNLLISGLKNVGKEVGMDVVRTGIDIAGCKIKGEC
Esculentin-1S   GLFSKFAGKGIKNMIIKGIKGIGKEVGMDVIRTGIDVAGCKIKGEC
Esculentin-1V   GIFSKFAGKGIKDLIIKGVKGIAKEAGMDVIRTGIDIAGCKIKGEC
                * *** *** **   * * *   ** **** ***** *********

Esculentin-2B   GIFSLVKGAAKLAGKGLAKEGGKFGLELIACKIAKQC

Esculentin-2P   GIFSLIKGAAKVVAKGLGKEVGKFGLDLMACKVTNQC
Esculentin-2S   GLFTLIKGAVKMIGKTVAKEAGKTGLELMACKVTNQC
Esculentin-2V   GIFTLFKGAAKLLGKTLAKEAGKTGLELMACKVTNQC
                * * * *** *   *   ** ** ** * ***   **
```

FIGURE 1.10 Comparison of amino acid sequences of esculentins identified in the present study with those previously identified in *R. esculenta* (1A, 1B, and 2B). Identical amino acid residues in each group of mature peptides are indicated by asterisks. Reprinted from Reference [9] with permission from Elsevier.

data-dependent acquisition experiments on the electrospray mass spectrometer, to identify additional peptide entities such as degradation fragments, extended or incompletely processed peptides, those arising from splice variants, and other peptides arising from different domains of the biosynthetic precursors. Peptides in chromatographic fractions or mass spectrometer data sets not identified using this approach can then be focused upon on a one-to-one basis for primary structural characterization. In our hands and within our model system, we have found this genomic approach to be a very powerful tool in the structural elucidation of the peptidome. Figures 1.6–1.10, reprinted with kind permission from Elsevier from Reference 9, illustrate the data obtained via a genomic approach.

1.6 LIMITATIONS OF EACH APPROACH

The bioassay approach has the unique advantage of indicating a biological activity/role for the novel bioactive peptide at the beginning of study. This in itself can be employed as a means of monitoring purification to homogeneity

for structural analysis that can be carried out by MS/MS fragmentation or by chemical means. Bioassay can give no indication of structure or purity of the active and in the past has been very time-consuming and requiring sacrifice of many experimental animals or at best, the acquisition and use of recent post-mortem tissues such as from an abattoir. *The peptidomic approach* gives no indication of bioactivity, except that inferred by structural similarity to known actives of related structure, but can define primary structure and posttranslational modifications if a variety of technologies are employed to these ends. However, *de novo* sequencing with 100% certainty is not currently possible—a factor that may severely limit this approach in species such as most amphibians, for which little, if any, peptides structures are present in online databases.

The genomic approach, with a limited amount of information on amphibian peptide cDNA structures, has proven to be a very rapid and powerful means of accessing *de novo* structural information on the skin peptidome. This can provide both deduced primary structures and molecular masses of novel peptides, both of which can be confirmed by mass spectrometric hardware. However, unusual propeptide convertase cleavages can occur that sabotage this approach in some instances, and posttranslational modifications are not apparent. However, certain equivocal or near-isobaric amino acid residues can be resolved.

It is thus clear that while each approach has advantages over the others, the disadvantages of each are very effectively cancelled out when all three are integrated and run in parallel in elucidation of the nature, structure, and function of the amphibian defensive skin secretion peptidome. The holistic approach to elucidate the structure and function of the secretory peptidome of the amphibian skin granular gland model described here thus gravitates around the integrated deployment of a powerful array of technological weaponry that the investigator, armed with such, can direct to the generation of a considerable body of novel data toward the understanding of this complex and intriguing natural chemical defense system. Many of the nuances unraveled here in the form of technological modifications, integrations, and developments could have significant relevance to applications in understanding and elucidating the functional genomics of other peptidome models.

1.7 CLOSING THOUGHTS

The sequencing of the human genome was a true milestone in science. In parallel with this and since then, there has been a huge number of sequencing projects initiated in a vast array of organisms including viruses, bacteria, fungi, plants, and both invertebrate and vertebrate animals. Some of the claims made by genetic researchers prior to this milestone event were quite bizarre, but few can equal the assertion that following possession of the DNA blueprint for the human life form “we would know what it is to be human.” Well, I think not.

How many of us know human identical twins that are identifiable not by their appearance but by their personality? The latter truly reflects what it is to be human, and this is invariably different even in humans of identical genotype. What the sequencing of the human genome taught us, and it is only recently that a near-complete edit has been forthcoming, is that we are not much more complex with respect to numbers of genes than a nematode or fruit fly. What a wake-up call! The so-called junk DNA is now thought to play a pivotal role in the regulation and multiplexing integration of temporal gene expression that provides the basis for virtually infinitesimal combinations that in turn provides the mathematics for complexity. So "the stone that the builders rejected has become the cornerstone." Genes, and even their definition is under scrutiny and the subject of debate, are essentially inert codes that perform functions and engineer structures through their encoded proteins, and it is these that occur in a greater array than genetic complexity through transcriptional splice variation and posttranslational modification, protease cleavage, and ultimate spatial location. Proteomics, as a discipline, is specifically concerned with defining the protein expression of a given cell type or tissue or globally, the protein expression of an organism. A whole technology has grown up around achievement of this end involving the use of suitable extraction media for various protein groups followed by 2D electrophoresis, stained spot identification and comparative analysis, trypsinization of proteins in spots, and the identification of protein tryptic fragments and hence parent proteins, by MS/MS fragmentation followed by database trawling using the huge amount of translated sequence information available for many organisms but particularly for humans. So proteomics, when minimized in a technical sense, is nothing more than peptide generation, fragmentation sequencing, and ultimate parent protein identification. Peptidomics essentially cuts out the middleman and focuses on those peptides that are already present as biologically relevant entities in cell/tissue extracts or body fluids. In the case of the majority of regulatory peptides, these are synthesized as part of larger precursor proteins intracellularly, but are processed from such by specialized site-directed proteases (prohormone/propeptide convertases) and often posttranslationally modified in a highly regulated manner. The resultant bioactive entities are usually stored in cytoplasmic granules prior to release by exocytosis. Other peptides, such as bradykinins and angiotensins, are stored in precursor form being released on demand by specific proteases, and these peptides are found in free form in body fluids. There is thus a plethora of bioactive peptides that conventional proteomic preparatory technologies are incapable of resolving—these are the raw material for peptidomics—the cell, tissue, fluid, or organism-specific complement of bioactive peptides. While this chapter has focused on a relatively unusual model system—the amphibian skin peptidome—much of our experience in problem solving in this system is directly applicable to many others. It is our expectation that the methods described here, and particularly their integration, may be helpful to other kindred researchers in their quest to inventory the myriad of specific peptidomes present in other life forms.

REFERENCES

1. Alves RR, Rosa IL. Why study the use of animal products in traditional medicines? J Ethnobiol Ethnomed 2005;30:1–5.
2. Anastasi A, Bertaccini G, Erspamer V. Pharmacological data on phyllokinin (bradykinyl-isoleucyl-tyrosine-*O*-sulphate) and bradykinyl-isoleucyl-tyrosine. Br J Pharmacol 1966;27:479–485.
3. Bevins BL, Zasloff M. Peptides from frog skin. Ann Rev Biochem 1990;59: 395–414.
4. Bick RJ, Poindexter BJ, Sweney RR, Dasgupta A. Effects of Chan Su, a traditional Chinese medicine, on the calcium transients of isolated cardiomyocytes: cardiotoxicity due to more than Na, K-ATPase blocking. Life Sci 2002;72:699–709.
5. Clarke BT. The natural history of amphibian skin secretions, their normal functioning and potential medical applications. Biol Rev Camb Philos Soc 1997;72: 365–379.
6. 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;27:2129–2136.
7. 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;27:1738–1744.
8. Chen T, Zhou M, Walker B, Harriot P, Mori K, Miyazato M, Kangawa K, Shaw C. Structural and functional analogs of the novel mammalian neuropeptide, neuromedin S (NmS), in the dermal venoms of Eurasian bombinid toads. Biochem Biophys Res Commun 2006;345:377–384.
9. 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;87:638–644.
10. Chen T, Li L, Zhou M, Rao P, Walker B, Shaw C. Amphibian skin peptides and their corresponding cDNAs from single lyophilized secretion samples: identification of novel brevinins from three species of Chinese frogs. Peptides 2006;27:42–48.
11. Chen T, Reid CN, Walker B, Zhou M, Shaw C. Kassinakinin S: a novel histamine-releasing heptadecapeptide from African running frog (*Kassina senegalensis*) skin secretion. Biochem Biophys Res Commun 2005;337:474–480.
12. Chen T, Gagliardo G, Walker B, Zhou M, Shaw C. Partial structure of the phylloxin gene from the giant monkey frog, *Phyllomedusa bicolor:* parallel cloning of precursor cDNA and genomic DNA from lyophilized skin secretion. Peptides 2005;26:2624–2628.
13. Chen T, Walker B, Zhou M, Shaw C. Dermatoxin and phylloxin from the waxy monkey frog, *Phyllomedusa sauvagei:* cloning of precursor cDNAs and structural characterization from lyophilized skin secretion. Regul Pept 2005;129:103–108.
14. Chen T, Xue Y, Zhou M, Shaw C. Molecular cloning of mRNA from toad granular gland secretion and lyophilized skin: identification of Bo8—a novel prokineticin from *Bombina orientalis*. Peptides 2005;26:377–383.
15. Chen T, Scott C, Tang L, Zhou M, Shaw C. The structural organization of aurein precursor cDNAs from the skin secretion of the Australian green and golden bell frog, *Litoria aurea*. Regul Pept 2005;128:75–83.

16. Chen T, Orr DF, O'Rourke M, McGlynn C, Bjourson AJ, McClean S, Hirst DG, Rao P, Shaw C. *Pachymedusa dacnicolor* tryptophyllin-1 (PdT-1): structural characterization, pharmacological activity and cloning of precursor cDNA. Regul Pept 2004;117:25–32.
17. Chen T, Shaw C. Cloning of the (Thr^6)-phyllokinin precursor from *Phyllomedusa sauvagei* skin confirms a non-consensus tyrosine *O*-sulfation motif. Peptides 2003;24:1123–1130.
18. 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;116:147–154.
19. Chen T, O'Rourke M, Orr DF, Coulter DJM, Hirst DG, Rao P, Shaw C. Kinestatin: a novel bradykinin B_2 receptor antagonist peptide from the skin secretion of the Chinese toad, *Bombina maxima*. Regul Pept 2003;116:139–146.
20. Chen T, Shaw C. Identification and molecular cloning of novel trypsin inhibitors from the dermal venom of the Oriental fire-bellied toad (*Bombina orientalis*) and the European yellow-bellied toad (*Bombina variegata*). Peptides 2003;24:873–880.
21. Chen T, Bjourson AJ, McClean S, Orr DF, O'Kane EJ, Rao P, Shaw C. Cloning of maximakinin precursor cDNAs from Chinese toad, *Bombina maxima,* venom. Peptides 2003;24:853–861.
22. Chen T, Farragher S, Bjourson AJ, Orr DF, Rao P, Shaw C. Granular gland transcriptomes in stimulated amphibian skin secretions. Biochem J 2003;371: 125–130.
23. Daly JW, Myers CW, Whittaker N. Further classification of skin alkaloids from neotropical poison frogs (Dendrobatidae), with a general survey of toxic/noxious substances in the amphibia. Toxicon 1987;25:1023–1095.
24. Daly JW, Spoande TF. Amphibian alkaloids: chemistry, pharmacology and biology. In: Pelletier SW, editor. Alkaloids: Chemical and Biological Perspectives. Volume 4. New York: J. Wiley and Sons, Inc; 1986. p1–274.
25. Daly JW, Garraffo HM, Spande TF, Jaramillo C, Rand S. Dietary source for skin alkaloids of poison frogs (Dendrobatidae)? J Chem Ecol 1994;20:943–955.
26. Erspamer V. Bioactive secretions of the integument. In: Heatwole H, Barthalmus, editors. Amphibian Biology, Volume 1. The Integument. Chipping Norton: Surrey Beatty & Sons; 1994. p 179–350.
27. Gehrke CW. In: Wixom RL, Bayer E, editors. Chromatography—A Century of Discovery 1900–2000. Elsevier Science; 2001.
28. Hofrichter R, editor. Amphibians in the cultural hertitage of peoples around the world. In: Amphibians: The World of Frogs, Toads, Salamanders and Newts. Buffalo, NY: Firefly Books; 2000.
29. Lazarus LH, Atilla M. The toad, ugly and venomous, wears yet a precious jewel in his skin. Prog Neurobiol 1993;41:473–507.
30. Richter K, Egger R, Negri L, Corsi R, Severini C, Kreil G. cDNAs encoding [D-Ala2]deltorphin precursors from skin of Phyllomedusa bicolor also contain genetic information for three dermorphin-related opioid peptides. Proc Natl Acad Sci U S A 1990;87:4836–4839.
31. Summers K. Convergent evolution of bright coloration and toxicity in frogs. Proc Natl Acad Sci U S A 2003;100:12533–12534.

32. Santos JC, Coloma LA, Cannatella DC. From the cover: multiple, recurring origins of aposematism and diet specialization in poison frogs. Proc Natl Acad Sci U S A 2003;100:12792–12797.
33. Stebbins RC, Cohen NW. Contributions of Amphibians to Human Welfare. In: Stebbins RC, Cohen, Natural History of Amphibians. Princeton: Princeton University Press; 1955.
34. Toledo RC, Jared C. Cutaneous granular glands and amphibian venoms. Comp Biochem Physiol A 1995;111:1–29.
35. Tyler MJ, Stone DJM, Bowie JH. A novel method for the release and collection of dermal, glandular secretions from the skin of frogs. J Pharmacol Toxicol Meth 1992;28:199–200.
36. Vanhoye D, Bruston F, Nicolas P, Amiche M. Antimicrobial peptides from hylid and ranin frogs originated from a 150-million-year-old ancestral precursor with a conserved signal peptide but a hypermutable antimicrobial domain. Eur J Biochem 2003;270:2068–2081.
37. Zasloff M. Magainins, a class of antimicrobial peptides from Xenopus skin: iso lation, characterization of two active forms, and partial cDNA sequence of a precursor. Proc Natl Acad Sci U S A 1987;84:5449–5453.
38. Zhou M, Chen T, Walker B, Shaw C. Lividins: novel antimicrobial peptide homologs from the skin secretion of the Chinese large odorous frog, *Rana* (*Odorrana*) *livida*: identification by "shotgun" cDNA cloning and sequence analysis. Peptides 2006;27:2118–2123.
39. Zhou M, Chen T, Walker B, Shaw C. Pelophylaxins: novel antimicrobial peptide homologs from the skin secretion of the Fukien gold-striped pond frog, *Pelophylax plancyi fukienensis:* identification by "shotgun" cDNA cloning and sequence analysis. Peptides 2006;27:36–41.
40. Zhou M, Chen T, Walker B, Shaw C. Novel frenatins from the skin of the Australasian giant white-lipped tree frog, *Litoria infrafrenata*: cloning of precursor cDNAs and identification in defensive skin secretion. Peptides 2005;26:2445–2451.
41. Zhou M, Liu Y, Chen T, Fang X, Walker B, Shaw C. Components of the peptidome and transcriptome persist in lin wa pi: the dried skin of the Heilongjiang brown frog (*Rana amurensis*) as used in traditional Chinese medicine. Peptides 2006;27:2688–2694.