UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Pharmacy 277

Studies on cysteine-rich peptides from *Nemertea* and *Violaceae*

Proteomic and transcriptomic discovery and characterization

ERIK JACOBSSON

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2019

ISSN 1651-6192
ISBN 978-91-513-0719-0
urn:nbn:se:uu:diva-390885

Dissertation presented at Uppsala University to be publicly examined in BMC A1:107a, Husargatan 3, Uppsala, Friday, 20 September 2019 at 09:15 for the degree of Doctor of Philosophy (Faculty of Pharmacy). The examination will be conducted in English. Faculty examiner: Professor Espen Hansen (UiT-The Arctic University of Norway, Marbio).

Abstract

Jacobsson, E. 2019. Studies on cysteine-rich peptides from *Nemertea* and *Violaceae*. Proteomic and transcriptomic discovery and characterization. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Pharmacy* 277. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-0719-0.

The overall aims of the projects included in this thesis were to discover, synthesize and characterize disulphide-stabilized peptides from marine worms (*Nemertea* sp.) and plants (*Viola* sp.).

One of the main outcomes of this thesis is the discovery of a new family of highly active cysteine-rich toxins, alpha nemertides, from nemertean worms (paper **II)**. Functional characterization and production routes of nemertides were further explored (papers **II-III**). In addition, 12 new cyclotides from the bog violet were discovered (paper **I)**. Finally, transcriptomes and mucus of the Antarctic nemertean *Parborlasia corrugatus* were investigated for toxin content (paper **IV)**.

In paper **I** wild-type leaf and callus tissue of the endangered bog violet, *V. uliginosa*, were analyzed using transcriptomics and LC-MS, resulting in the discovery of 12 new cyclotides (i.e. cysteine-rich cyclic peptides). In addition, cyclotide expression under different cell-growth conditions was monitored.

In paper **II** the discovery and initial characterization of a new family of highly active peptides, the alpha nemertides, from the epidermal mucus of the world's longest animal; *Lineus longissimus* is described. The most abundant alpha nemertide, alpha-1, was extracted in minute amounts, prompting the use solid phase peptide synthesis (SPPS) for further characterization. The tertiary structure of alpha-1 was elucidated and revealed an inhibitory cystine knot (ICK) framework. The knotted core-structure is similar to the cyclic cystine knot (CCK) motif, found in the cyclotides described in paper **I**.

In manuscript **III**, the production route established in paper **II** was used to produce nemertides alpha 1-7. These were tested *in vivo* in an *Artemia* microwell assay as well as on an extended panel of voltage-gated sodium channels ($Na_V1.1$ – 1.8 and $BgNa_V1$). All seven alpha nemertides were highly active in the *in vivo Artemia* assay with EC_{50} values in the sub to low μM range. The alpha nemertides were also active in the Na_Vs tested. However, differences in the activity profiles were observed, indicating an opportunity for future optimization of alpha nemertides to reach higher specificity to certain Na_V subtypes.

In manuscript **IV,** the exploration of nemertide toxins was extended to include the Antarctic *P. corrugatus*. Resulting findings include a set of cysteine-rich peptides, some similar to the nemertides previously discovered in paper **II**. Two purified peptides and one fraction were evaluated for their membranolytic activity.

Keywords: Peptide toxin, cystine knot, nemertide, cyclotide, nemertea.

Erik Jacobsson, Department of Medicinal Chemistry, Division of Pharmacognosy, Box 574, Uppsala University, SE-75123 Uppsala, Sweden. Department of Medicinal Chemistry, Farmakognosi, Box 574, Uppsala University, SE-751 23 Uppsala, Sweden.

ISSN 1651-6192
ISBN 978-91-513-0719-0
urn:nbn:se:uu:diva-390885 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-390885)

Damn it, it wasn't quite fresh enough!

-Herbert West,
Reanimator

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I Slazak, B., **Jacobsson, E.**, Göransson, U and Kuta, E. (2015) Exogenous plant hormones and cyclotide expression in *Viola uliginosa* (Violaceae). *Phytochemistry*, 117, pp. 527–36

II **Jacobsson, E.**, Andersson H. S., Strand, M., Peigneur, S., Eriksson, C., Lodén, H., Shariatgorji, M., Andrén P.E., Lebbe, E. K. M., Rosengren, K. J., Tytgat, J. and Göransson, U. (2018) Peptide ion channel toxins from the bootlace worm, the longest animal on Earth. Scientific reports, 8(1), p. 4596

III **Jacobsson, E.**, Andersson, H. S., Laborde, Q., Strand, M., Peigneur, S., Tytgat, J. and Göransson, U. Functional characterization of the nemertide alpha family of peptide toxins (Manuscript)

IV **Jacobsson, E.**, Strömstedt, A. A., Andersson H. S., Avila, C. and Göransson, U. Peptide toxins from the Antarctica: the nemertean predator and scavenger *Parborlasia corrugatus* (Manuscript)

Additional Publications

- Göransson, U., **Jacobsson, E.**, Strand, M., and Andersson, H. S. (2019) The toxins of nemertean worms. *Toxins*, 11(2):120
- Park, S., Yoo, K., Marcussen, T., Backlund, A., **Jacobsson, E.**, Rosengren, K. J., Doo, I. and Göransson, U. (2017) Cyclotide evolution: insights from the analyses of their precursor sequences, structures and distribution in Violets (Viola). *Frontiers in plant science*, 8, p. 2058
- Strand, M., Hedström, M., Seth, H., McEvoy, E. G., **Jacobsson, E.**, Göransson, U., Andersson, H. S. and Sundberg, P. (2016). The bacterial (*Vibrio Alginolyticus*) production of tetrodotoxin in the ribbon worm *Lineus longissimus* - Just a false positive?. *Marine Drugs* 14 (4): 1–11
- Göransson U., Rosengren J., **Jacobsson E.**, Andersson H. and Strand M. (2018) Nemertea-derived bioactive compounds (Patent) WO 2018/004433 A.

Contents

Abbreviations

aa	Amino acid
ASW	Artificial sea water
CRP	Cysteine-rich peptide
ESI	Electrospray ionization
HPLC	High performance liquid chromatography
LC	Liquid chromatography
MALDI	Matrix assisted laser desorption ionization
MALDI-IMS	MALDI imaging
MS	Mass spectrometry
ORF	Open reading frame
PTM	Post-translational modification
SEC	Size exclusion chromatography
SPPS	Solid phase peptide synthesis
TTX	Tetrodotoxin
UPLC	Ultra performance liquid chromatography
VGSC	Voltage-gated sodium channel

1 Introduction

Peptides and proteins are the bearers of the functionality encoded in our genomes. They can have a wide variety of forms, and have almost no boundaries to what functionality they convey in terms of activity; some have structural functions, some are enzymes with specific tasks, some make up muscles, some are used to protect a host from attacks from other organisms, yet others are utilized to catch prey. The discovery, characterization, and production of peptides exhibiting the latter two functionalities, defense and attack-related peptides, are the focus of this thesis. Many such peptides/proteins share a feature: they contain disulphide bonds.

1.1 Stabilized peptides for protection and attack

Disulphide-stabilized peptides, or cysteine rich-peptides (CRPs) are, as the name implies, peptides that contain cysteine residues. The cysteine residues form disulphide bonds that stabilize the peptide structure. These disulphide bonds convey tertiary and quaternary structure and, may be organized in several configurations, for example in a knot[1] or ladder[2] framework. Furthermore, the peptides may be classified based on their cysteine patterns found in the primary sequence. The latter is often the case in the cone snail (*Conus* sp.) derived conotoxins[3].

A variety of CRP classes with different activities has been identified, including protease inhibitors[4], antimicrobial peptides[5], and toxins[6–8]. Some classes of CRPs from both animals and plants can be viewed as natural libraries of variable functional peptides, containing variable sequences yet containing conserved cysteine patterns[9,10].

The majority of research focused on peptide toxins have traditionally been limited to a few taxa, including cone snails[11] , snakes[12] and arachnids[13]. The development of proteomic- and transcriptomics techniques has however reduced the amount of sample required (and cost for analysis) for identification of new peptides and proteins. This reduction opens up for more extensive investigations into previously neglected organisms[14].

1.1.1 Cysteine frameworks and patterns

The cysteine pattern, the number and spacing of cysteine residues in the primary sequence, has been used for peptide classification. It is also a feature that can be utilized in the bioinformatic exploration of genetic data i.e. by using simple prosite patterns[15] or regular expressions[16]. Cysteines that are connected through disulphide bonds in certain connectivities are referred to as cystine frameworks. One such cystine framework commonly found in both peptide toxins and plant defense peptides is the cystine knot.

The cystine knot is made up by three disulphide bonds, where two disulphides form a macrocycle together with the peptide backbone and the third disulphide is threaded through the macrocycle. If the peptide backbone is head to tail cyclic, as in cyclotides, the framework is called a cyclic cystine knot, CCK, Figure 1.

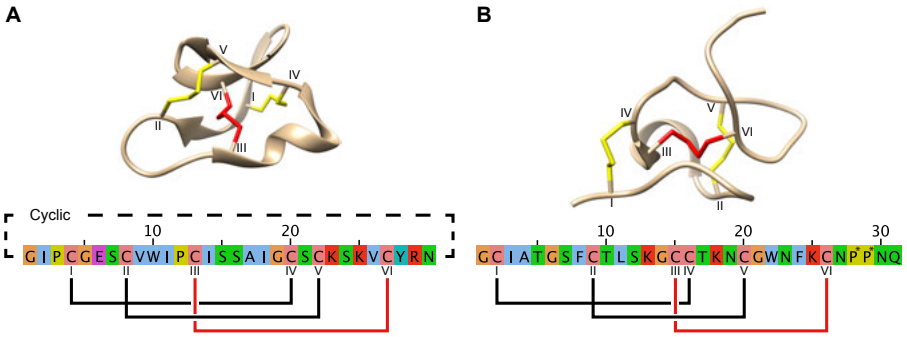

Figure 1. Cystine knots. A: Cyclic cystine knot, exemplified by the cyclotide cycloviolacin 02 (PDB 2KNM). The cyclic backbone is visualized by a dotted line. B: Inhibitory cystine knot, ICK, present in the neurotoxic nemertide alpha-1 (PDB 6ENA). P*: hydroxyproline. Cysteine residues are represented by roman numerals. The cystine knot framework is marked in bars, with the macrocycle-piercing disulphide (III-VI) marked in red.

The cystine knot framework is characterized by remarkable stability against heat and enzymatic proteolysis. These characteristics together with the variability in sequences have been utilized in various grafting projects, especially the CCK framework[17,18].

In the field of cysteine-rich conotoxins, the cysteine pattern may be used to classify the peptides. In the database of conotoxins, ConoServer[19], there are currently 26 different defined cysteine patterns containing 4 to 10 cysteine residues.

1.2 Discovery of cysteine rich peptides

In this section the procedures commonly used for the isolation of CRPs, from discovery to functional activity, are described.

1.2.1 Observation of activity

Discovery of peptides and other natural products usually starts with a biological or ecological observation. For example, the birth-accelerating herb 'Kalata-kalata' (*Oldenlandia affinis*) used in traditional medicine in Zaire (Congo) was found to contain polypeptides responsible for the activity[20]. These peptides were later discovered to be cyclotides.

In the case of toxins, the observations might be painful, often lethal, to the organism it is tested on, e.g. spiders preying on insects, cone snails hunting fish/worms/molluscs or the complications following snakebites. Such observations are not only of academic interest: some have led to new biopesticides[21,22], medicines [6,23] and biotechnological tools[24].

Peptides used as biopesticides include a conventional spray of recombinant spider toxin (ion channel active ω/κ-hexatoxin-Hv1a Spear®[21]) and a cyclotide-containing extract (*Clitoria ternatea*; Sero-X®[22]).

There is currently one conotoxin, Prialt® (ziconotide, ω-MVIIA), which is registered as a medicinal drug, for severe pain[25]. Chlorotoxin, a toxin initially extracted from a scorpion, *Leiurus quinquestriatus*[26], is explored as a visualization tool to diagnose and illuminate cancer tissue for neurosurgeons[24].

1.2.2 Extraction and isolation of peptides

When an observation has been made, the next step is to amass enough raw material for extraction and isolation of individual or groups of compounds. Extraction and isolation procedures depend on the matrix from which the compounds will be extracted and the characteristics of the compounds.

The peptides isolated in the work described in this thesis are all water soluble, and extraction was performed in water-acetonitrile solutions. The subsequent isolation was performed by the use of reverse-phase high performance liquid chromatography (RPHPLC-UV), in some cases in connection with size exclusion chromatography (SEC). These techniques are commonly used in natural product isolation, but the solubility and size of peptides distinguishes them from small molecules.

1.2.3 Identification, reduction, alkylation and enzymatic cleavage of peptides

The sequencing and identification of peptides is often performed using LC-MS or LC-MS/MS. LC-MS might be sufficient in cases where the molecular weight and presence of the peptide in the extract is already known, while MS/MS is needed for sequencing of unknown peptides (*de-novo* sequencing).

To simplify interpretation of the MS/MS spectra, peptides and proteins are cleaved into shorter stretches of amino acids[27]. The cleavage is performed by the use of enzymes; in the context of this thesis usually trypsin, chymotrypsin and endoproteinase GluC. In the case of disulphide containing peptides, the disulphides must be reduced and alkylated prior to enzymatic cleavage and MS/MS sequencing[28].

Reduction removes the S-S bond, and subsequent alkylation blocks the re-formation of the disulphide, thus keeping the peptide available for enzymatic digestion. The procedure also provides information about the number of cysteine residues, by specifically adding a certain mass (for iodoacetamide, IAM; 58 Da per cysteine residue involved in a disulphide bond) to each alkylated residue[28].

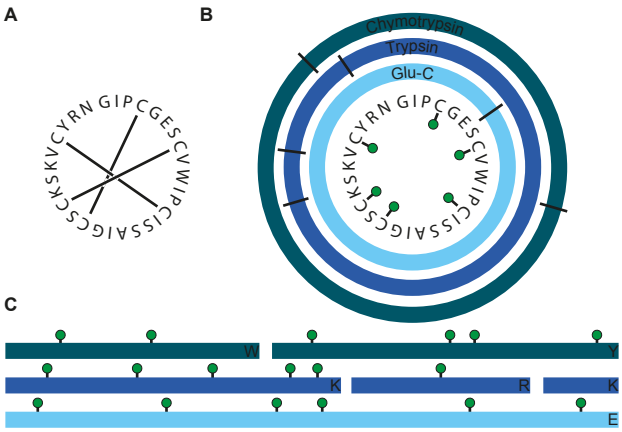

Figure 2. Schematic drawing of reduction, alkylation and enzymatic cleavage of cycloviolacin O2. A: Representation of CyO2 with intact disulphides (black bars). B: middle circle; reduced and alkylated Cys residues. Light blue circle; Glu-C cleavage site (black bar). Dark blue circle; Trypsin cleavage sites. Outer, green-blue circle; Chymotrypsin cleavage sites. C: Hypothetical enzymatic fragments from cleavage of CyO2, color scheme as in B.

1.2.4 Mass spectrometry based peptide de-novo sequencing

The amino acid sequence of the enzymatically cleaved peptides are determined by the use of MS/MS. Mild ionization by electrospray is advantageous for peptides and proteins. The most common fragments found in collision-

induced fragmentation of peptides, i.e. y, b and sometimes a ions, are displayed in figure 3. The mass differences between peaks in the spectrum correspond to amino acids and are used to assign the positions of amino acids in the sequence. For reliable sequencing both the b- and y-ion series should be nearly complete. Furthermore, the direction of the sequence (N-terminal to C-terminal; b-series or C to N-terminal; y-series) has to be determined.[29]

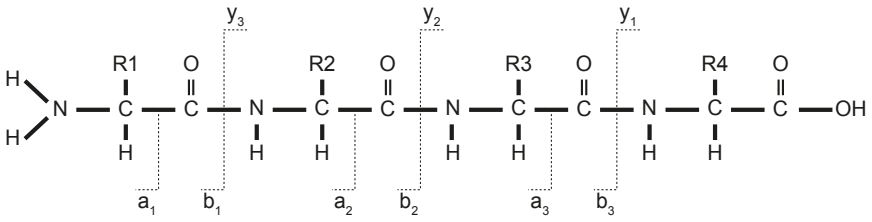

Figure 3. Fragmentation of peptides with y, b and a ion-types displayed. R 1-4; sidechains for the amino acids 1-4.

Following this strategy, short stretches of sequences or the full-length fragment may be deduced, and post translational modifications, PTMs, can be detected. A limitation is, however, that isobaric (Ile/Leu) or near-isobaric amino acids (Gln/Lys) are difficult to differentiate from each other[30]. One way to simplify the sequencing is the parallel use of genetic data, e.g. transcriptomes.

1.2.5 Transcriptome-aided sequencing of short peptides

Ambiguities (e.g. Ile/Leu) in short MS/MS derived sequences can be resolved by the use of genetic data i.e. transcriptome sequencing[31]. The transcriptome (the transcribed RNA) can readily be extracted by commercially available kits. The sequencing is performed using next generation sequencing, NGS. This is a technique in which all expressed nucleotide sequences are sequenced in a massive parallel manner often using a continuous "sequence by synthesis" principle[32].

The transcriptomes in this thesis were sequenced using RNA-seq Illuminia platform[33] and assembled with the Trinity[34] software.

With the assembled transcriptome at hand it is possible to utilize it in the sequencing from MS/MS experiments. For convenience, the nucleotides are translated into protein sequences in all possible reading frames, ORFs. To search for short stretches of peptide sequences a regular expression based approach (prosite pattern[15] and fuzzpro/tran[35]) is used, for exact sequence hits. Retrieved sequences from the search is then evaluated for cysteine residues and common processing sites to deduce the mature sequence from the precursor.

When the full precursor sequence or long stretches thereof has been identified, a basic local alignment search tool[36] (BLAST) is used to find sequences similar to the query (e.g. homologous sequences).

1.2.6 Laboratory scale production of peptides

If the sequenced peptide cannot be isolated in sufficient amounts for biological and structural studies, the peptide may be synthesized using solid phase peptide synthesis, SPPS[37].In SPPS the growing peptide chain is assembled on a solid support (resin beads) in a reaction vessel with a filter in the bottom connected to a draining tube. The synthesis is carried out from the C-terminal amino acid (i.e. the elongation takes place on the amino group) through iterative cycles of deprotection and coupling of the amino acids. When assembly is complete, the peptide is cleaved from the solid phase, the resin, and simultaneously deprotected from side-chain protecting groups.

The process can be automated, and the coupling times may be shortened by the application of microwaves[38]. In this thesis, fmoc-SPPS[39] was used to synthesize alpha nemertides (paper II and III). The SPPS workflow is displayed in figure 4.

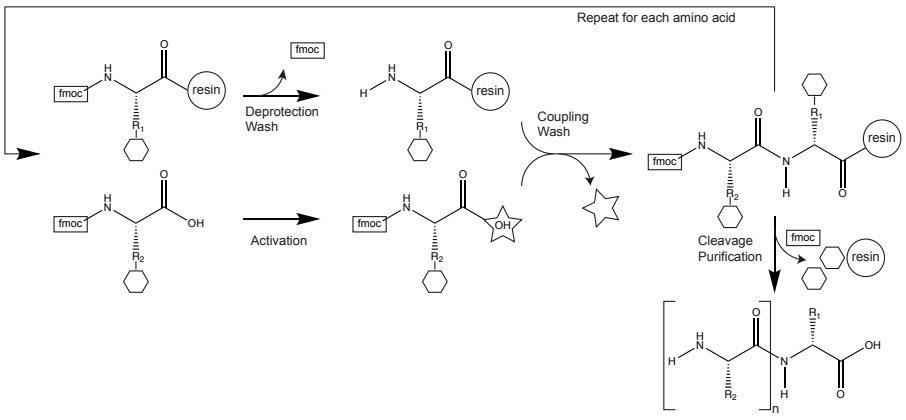

Figure 4. Solid phase peptide synthesis workflow. R; amino acid side chain. Hexagon; side chain protection group. Fmoc (fluorenylmethyloxycarbonyl): N-terminal protecting group

SPPS is suitable for lab-scale production of peptides and while it is possible to produce larger batches, also other means of production, i.e. cellular expression systems, are possible[40].

1.2.7 Folding of cysteine-rich peptides

After synthesis and purification of the peptide, the correct disulphide connectivity has to be formed to obtain the native product. The larger the number of

cysteines the more complex the folding might be (i.e. a higher number of possible connectivities – a peptide six cysteins can theoretically form 15 different three disulphide isomers).

In vivo the formation of disulphides is enzymatically assisted, mainly by protein disulphide isomerases, PDIs[41]. The PDIs are generally localized to the ER lumen, and have dual functionality; to form the initial disulphide bridges and to re-shuffle the bonds into the final, functional form[41,42].

The *in vitro* formation of correct disulphide pairing can be troublesome and a several different folding conditions often have to be tested to achieve correct folding[43]. The complex disulphide framework of the ICK peptides in this thesis has been generated by oxidative folding in a system containing reduced and oxidized forms of glutathione, and DMSO (or isopropanol), figure 5A. After purification the synthesized, folded peptide must be shown to exhibit the same activity as, and be identical to the native form. The latter can be accomplished by a co-injection experiment, figure 5B.

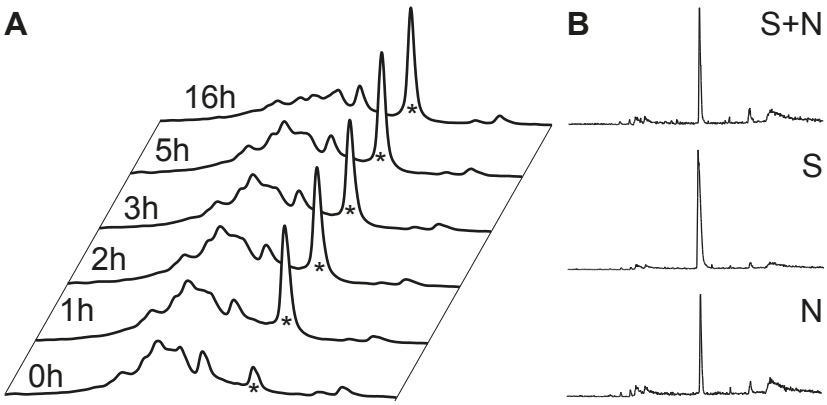

Figure 5. Folding and co-injection. A; folding of nemertide alpha-1 monitored over 16 h, by HPLC-UV. * denotates the folded product. B; co-injection (S+N) in UPLC-QToF of synthetic (S) and native (N) peptide.

1.2.8 Assays and functionality of compounds

With enough material gathered from extraction or synthesis, the compounds can be tested in various functional assays/bioassays for determination of function. The assay is often chosen in relation to the initial observation that sparked the interest in the organism studied. For example, if the observation was a clear area in a petri dish surrounding a mold colony, a starting point would be to test against microorganisms; if a snake bite is found to cause hemolysis, a test for the membranolytic properties of the venom (or compounds) would be advised.

The growing body of sequence-activity (or structure-activity) data the amassed in databases has allowed the development of tools for activity prediction[44,45]. These tools may give an idea of which assays to use as starting points for further exploration.

A schematic flowchart of the process from observation to assay is displayed in figure 6.

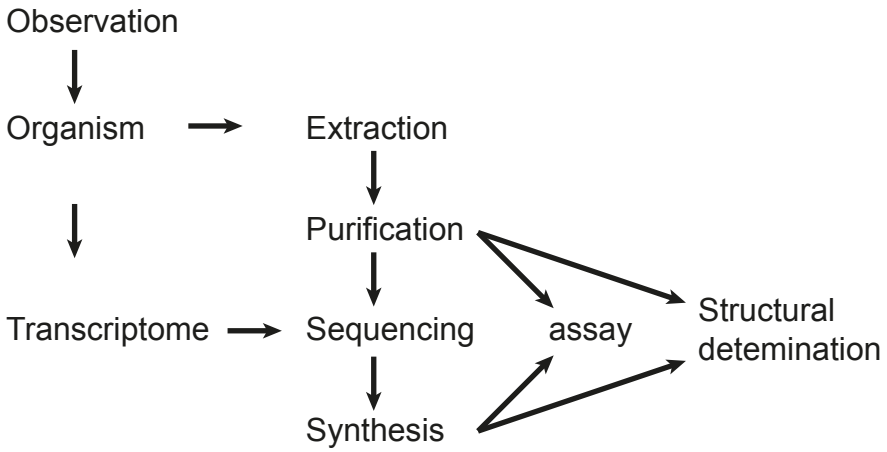

Figure 6. Simplified flowchart of the process from observation to assay and structural determination.

2 Aims

The overall aim of the projects included in the thesis has been to discover, synthesize and characterize disulphide-stabilized peptides from marine worms (*Nemertea* sp.) and plants (*Viola* sp.).

This thesis is divided into two parts related to discovery of cysteine rich peptides. In the first part I investigated the production of cyclotides in endangered bog violet, *Viola uliginosa*. In addition, a suspension cell culture was established and evaluated for cyclotide production *in vitro*.

In the second part, the peptide toxin presence in epidermal mucus and transcriptomes of nemertean worms was investigated, and routes for synthesis and folding of alpha nemertides were established. Furthermore, characterization of alpha nemertides with regard to toxicity *in vivo* on *Carcinus maenas*, *Blatella dubia,* and *Artemia salina* was performed. Finally, I investigated the activity of alpha nemertides on a set of voltage gated sodium channels from insects and vertebrate origin.

The specific aims for this thesis were as follows:

- To map cyclotides in *V. uliginosa* using transcriptomics and UPLC-QToF, and to generate a suspension cell culture suitable for production and biochemical investigations, **paper I**.
- To investigate nemertean peptide toxins from the epidermal mucus of *Lineus longissimus*, **paper II**.
- To produce and characterize alpha nemertides from nemertean worms; **paper II** and **manuscript III**.
- To perform a preliminary characterization of the proteins/peptides from mucus and transcriptomes of the Antarctic nemertean worm *Parborlasia corrugatus*, **manuscript IV**.

3 Cyclotides from *Viola uliginosa*: callus tissue and wild type plant (Paper I)

3.1 Cyclotides

The cyclotides are a well-studied group of peptides[22,46–48], which have been found in species belonging to *Rubiaceae*, *Violaceae*, *Cucurbitaceae, Fabaceae* and *Solanaceae*[49]. Cyclotide-like sequences have also been found in *Poaceae*[50].

The cyclotides are head to tail CCK peptides (figure 1A) of approximately 30 amino acids, known for their resistance to heat and enzymatic degradation[51]. These properties together with the highly modifiable loops between cysteines have made them suitable for grafting experiments[52].

Furthermore, the intrinsic activities of cyclotides include membrane disruptive-[53], antimicrobial-[54], and insecticidal[55] functions. These natural functions implicate them as host-defense compounds, and they are commercially utilized in a cyclotide containing extract from the Fabaceaean *Clitoria ternatea* marketed as an bio-pesticide[22] (Sero-X®).

Transcriptomic and proteomic in-depth analysis of the number of cyclotides in one species, *V. tricolor,* totaled 164[10], and the total number of cyclotides in *Violaceae* has been estimated to 25000-150.000[10,56]. Furthermore, the number of unique cyclotide sequences predicted in *Rubiaceae* is 10.000-50.000[57].

The cyclotides are divided into two main sub-groups: the bracelet and the möbius. The möbius cyclotides harbor a cis-proline in loop 5, conceptually twisting the backbone 180 degres[58]. From a production viewpoint, the möbius type is readily synthesized and folded into its native fold, while the bracelet subgroup that makes up roughly two thirds of the known sequences are more cumbersome to fold *in vitro*[59,60].

3.1.1 Cyclotide expression in *Viola uliginosa*

Viola uliginosa is an endangered species of violets distributed in the Baltic sea region. The species has been subjected to micropropagation experiments, which laid the foundation for *Viola* suspension cultures to be used in bio-reactor based production of cyclotides[61].

In **paper I** the cyclotide content from *Viola uliginosa* was explored using transcriptomic- and mass spectrometric techniques. Furthermore, the utility of suspension cell cultures for studies on the effect of growth modulators on the bio-production of cyclotides was investigated.

Two transcriptomes from *V. uliginosa*, one for callus tissue and one from a leaf of wild type plant collected nearby Uppsala, Sweden, were analyzed in conjunction with UPLC-QToF for cyclotides. These transcriptomes were among the first RNA-seq libraries of cyclotide containing plant made available in public databases.

The combined data revealed in total 20 different cyclotides of which 12 were new sequences (viul A-L). Four of the sequences were only found in transcriptomic data (viul A, D, J and cycloviolin D), while one (cyO13) cyclotide was found on peptide level only, figure 7. CyO13 was the most abundant cyclotide in the extracts.

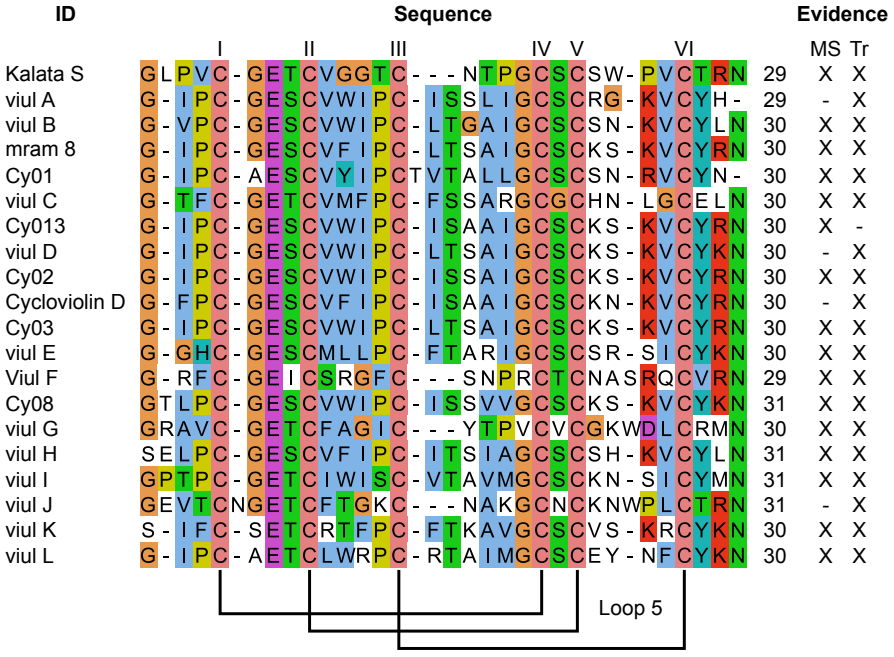

Figure 7. Cyclotide sequences found in *V. uliginosa*. ClustalW alignment of the 20 cyclotides found in combined data from transcriptomes and UPLC-MS and MS/MS. Cysteine residues are indicated in roman numerals. Disulphide connectivity in black bars. Evidence column; MS: found on peptide level by mass spectroscopy; Tr present in transcriptomes.

In addition to the new cyclotide sequences found, the presence of components of the cyclotide processing machinery, i.e. possible enzymes with the ability to cyclize peptides, such as the asparginyl endoprotease (AEP) butelase-1 discovered in *C. ternateae*[62], was investigated. A potential sequence in the *V.*

uliginosa transcriptomes with similarity to the full-length butelase-1, including the proposed active site was identified.

3.1.2 Generation of *V. uliginosa* suspension cell-culture

A series of explants from petiole, leaf and root was inoculated with varying concentrations of auxins (2,4-dichlorophenoxyacetic acid (2,4-D) and 1-naphthaleneacetic acid (NAA)) and cytokins (kinetin (KIN) and thidiazuron (TDZ)) in Murashige and Skoog (M&S) medium[63].

The two combinations with the highest variance in cyclotide expression patterns (M&S supplemented with 1 mg/L TDZ and M&S supplemented with 2 mg/L KIN and 2 mg/L 2.4-D), as judged by UPLC-QToF analyses, were selected for suspension culture generation.

The suspension culture derived from callus tissue supplemented with TDZ alone exhibited slow increase in biomass and displayed signs of organogenesis i.e. root like structures. The culture obtained from explants supplemented with the combination of cytokine and auxin yielded a suspension with a comparably higher biomass doubling rate (three days). The yield of cyclotides from this suspension culture peaked after 10 days, with a production of CyO13 reaching above 4 mg/ml dry mass, twice the yield from wild type leaves and four times more than what can be extracted from roots. No organogenesis was observed for this culture.

Key results and observations from **paper I**:

- 12 new cyclotides (Viul A-L) were discovered and a total of 20 cyclotide sequences were identified on RNA and peptide level.
- A high cyclotide yielding suspension cell culture was established from *V. uliginosa*, opening for large scale production of natural cyclotides and biochemical investigation into cyclotide production machinery.

4 Bioactive peptides from nemertean worms (papers II-IV)

There are approximately 1350 valid species of nemertean worms[64]. Most of the nemerteans are found in marine environments, some are found in freshwater and a few are terrestrial. The size distribution among the nemertean species is spectacular; some are of millimeter length, while the taxon also contains the longest animal in the world, *Lineus longissimus,* with lengths up to 55 m, figure 8A.

Most of the nemertean worms are predators using a eversible proboscis to catch prey[65]. The proboscis can comprise up to two thirds of the body length, and in hoplonemertean nemerteans (formerly; enopla), it is armed with a stylet. The nemertean worms that lack the stylet are organized in the pilidiophora and palaenemertean classes (formerly; anopla) figure 8B[66].

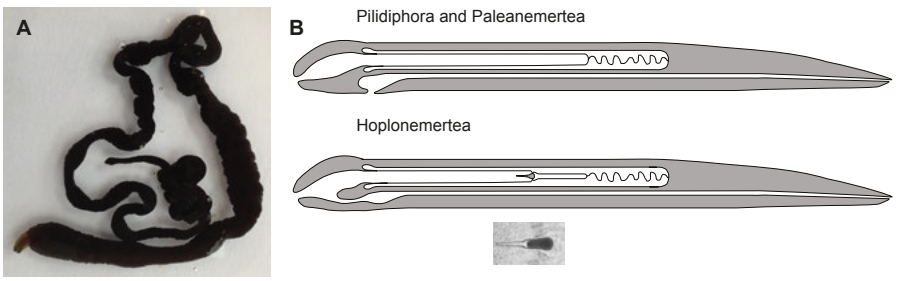

Figure 8. L. longissimus and internal layout of armed and unarmed nemerteans respectively. A: *L. longissimus* in the lab. B: top; the outline of the unarmed pilidiophora and paleonemertea classes (formerly anopla). Middle; outline of the armed hoplonemertea class (formerly enopla). Bottom; example of a stiletto from *Amphiporus lactifloreus*.

The nemertean worms are seldomly preyed upon by other organisms[67]. This feature is attributed to their ability to excrete large amounts of mucus (hence the Swedish common name “slemmask”; mucus worm). The mucus has been investigated for toxicity in some species prior to this study, revealing both neurotoxic[68] (neurotoxin B I-IV) and membrane disruptive activities[69,70] (cytolysins A I-II and parborlysins) by small proteins.

Small molecular toxins have also been extracted from nemertean worms, including the fugu toxin tetrodotoxin (TTX)[71] and anabaseine[72], figure 9.

Derivates of anabaseine (especially the DMXBA derivate[73]) have been investigated for the treatment of Alzheimer disease.

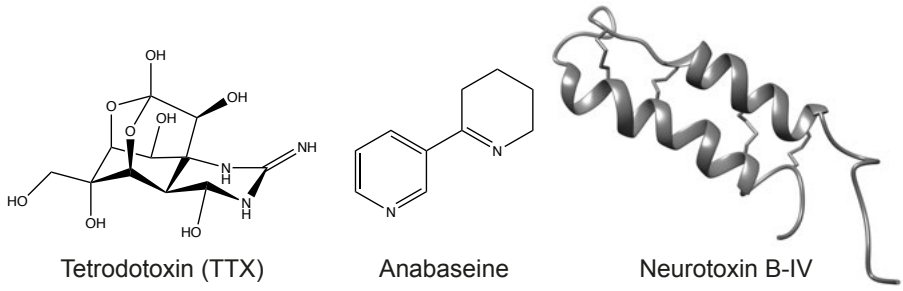

Figure 9. Structures of toxins from nemertean worms. Tetrodotoxin (TTX), anabasein and neurotoxin B-IV (PDB: 1VIB).

4.1.1 Definitions: venoms, poisons, toxins and toxungens.

Several classification schemes for toxic compounds have been proposed, with slightly different meanings. In this thesis the definitions comply with the definitions described by Nelsen *et al.*[74].

Toxins, the actual compounds are parts of venoms, poisons and toxungens. Venoms are mixtures of toxins that are injected into the prey or aggressor through specialized contraptions, such as spines or teeth, connected to a venom gland. Poisons are toxins that are not administrated through a specialized organ: the poisons are localized to certain parts of the body (e.g. fugu liver), or distributed throughout the tissues. The poisons rely on the aggressor to ingest it to be effective. Toxungens, like poisons are also not delivered to the prey/aggressor through a wound. It is however excreted and delivered to an aggressor as a response to a threat; e.g. spitting cobras[74] and excretions from toads[75].

The toxins present in all types; venoms, poisons and toxungens, may be produced by the organism, sequestered from the environment, or produced by symbiotic microorganisms. Furthermore, an animal may use more than one of the classes.

4.2 Discovery of alpha nemertides from *Lineus longissimus* and initial characterization of nemertide alpha-1 (Paper II)

L. longissmus, the longest animal in the world, can be found in the waters on the northern hemisphere. The organism is found from depth of approximately

10 m and deeper in Swedish waters while it may be found in the intertidal zone by the coasts of United Kingdom.

L. longissimus has been proposed to harbor bacteria which produce TTX[76], and that was the starting point for the investigation into the species. No TTX could however be found in initial experiments, but rather compounds of much higher molecular weights[77].

4.2.1 *L. longissimus* specimen and mucus collection

The living *L. longissimus* specimen used in this study was collected in the Koster fjord on the west coast of Sweden at a depth of approximately 35 m and identified by Dr. Malin Strand, Swedish University of Agricultural Science. The specimen was reared in a small aquarium in artificial sea-water (ASW), and mucus was collected weekly.

Mucus was collected by massaging the worm with a Pasteur pipette in a small container containing ASW. The mucus was lyophilized prior to further analysis.

A series of tissue samples was collected from the worm and stored either in RNAlater at -20° C or flash frozen in liquid nitrogen and stored at -80° C, before shipping for transcriptomic analysis. Flash frozen tissue sections were also used for MALDI-IMS.

4.2.2 Purification strategy for peptides and proteins from mucus

The collected mucus contained large amounts of salt from the sea water. To remove the bulk amount of salt, a gravity-based size exclusion chromatography (SEC) column was used, producing two fractions from the lyophilized mucus: one high molecular weight and one low molecular weight fraction. The high molecular weight fraction, containing putative peptides/proteins was further fractionated using HPLC-UV, revealing three major peaks. UPLC-QToF identified the main peaks to be two compounds at 3.3 kDa (henceforth named nemertide alpha-1 and -2) and one at 6.4 kDa (nemertide beta-1), figure 10. Reduction and alkylation using DTT and IAM respectively revealed that the smaller alpha type of peptides each contain six cysteines, whereas the beta nemertide contains eight cysteines.

4.2.3 MALDI-IMS display epidermal localization of the peptides

Transversal sections of flash frozen *L. longissimus* were subjected to MALDI-IMS analysis to explore the micro localization of the three main peptides seen in HPLC and UPLC-QToF experiments. All three peptides were found to be

localized close to the outer epidermal layer, and were not found in conjunction to the proboscis, figure 10B.

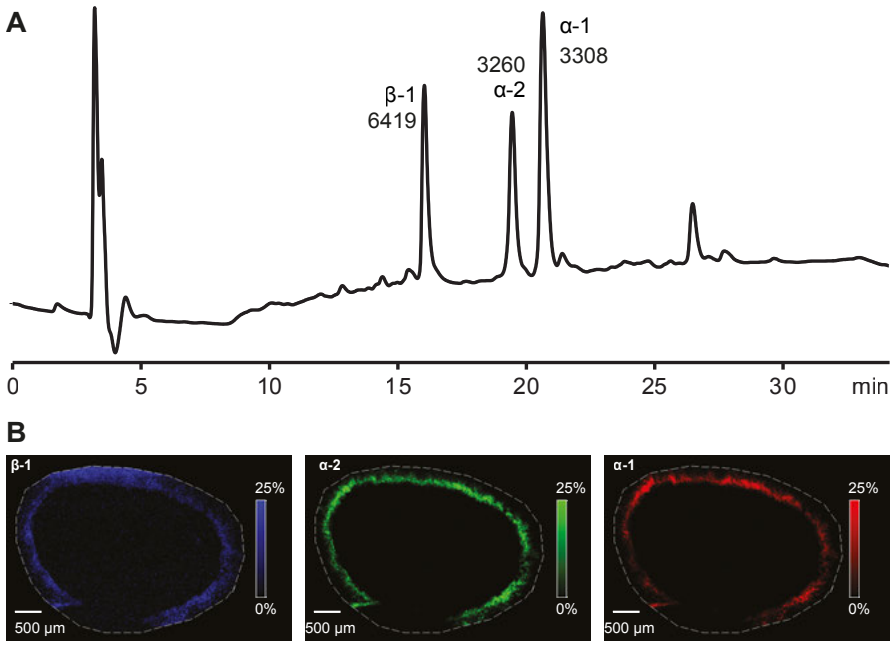

Figure 10. HPLC-UV and MALDI-IMS. A: HPLC-UV of high molecular SEC fraction displaying the three major peaks (beta-1, alpha-2 and alpha-1) with masses. B: MALDI-IMS at a resolution of 15 µm of extracted masses for beta-1, alpha-2 and alpha-1 in transversal sections of *L. longissimus*.

The epidermal localization of the peptides indicates a function as part of a host-defense-system, in analogy to the localization of amphibian toxin-glands[78]

4.2.4 Sequencing of peptides using mass spectrometry and transcriptomic data.

Enzymatic cleavage using trypsin and chymotrypsin revealed similar cleavage patterns for the two alpha nemertides; Glu-C did not produce any detectable products whereas trypsin and chymotrypsin gave rise to two fragments each, out of which one was identical per enzyme. The sequences could however not be fully elucidated by mass spectrometry alone, figure 11A.

RNA was extracted from different cross sections of the worm and sent to Macrogen™ for next generation sequencing using the Illuminia 2000 platform, and subsequent assembly by Trinity. Regular expression searches in the translated transcriptome yielded a sequence corresponding to the fragments found in MS/MS experiments for alpha-1.

However, the experimental mass of nemertide alpha-1 and -2 did not match the sequence found in the transcriptome, suggesting the presence of posttranslational modifications, PTMs. PTMs are often found in peptide toxins and include disulphide bridges, C-terminal amidation and the hydroxylation of proline into hydroxyproline (Hyp). The presence of Hyp residues was later confirmed by co-injection of native and synthetic alpha-1.

The precursor and conoprec[19] predicted processing sites are displayed in figure 11B. Nemertide alpha-2 could not be identified in the transcriptome produced in **paper II**, but the full sequence could however be identified in publicly available *L. longissimus* transcriptomes (NCBI accessions: SRX565180-SRX565176; released to the public by PopPhyl)[79,80].

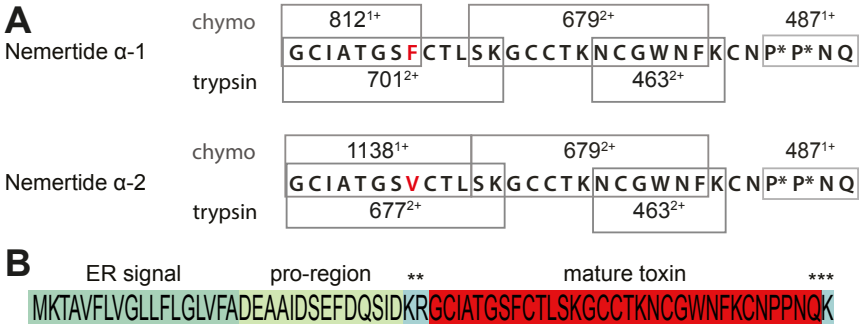

Figure 11. Enzymatic fragments of nemertide alpha-1 and -2, and precursor sequence for alpha-1 found in *L longissimus* transcriptome. A: ion map displaying the cleavage products from chymotrypsin and trypsin. The Phe to Val change between alpha-1 and -2 is displayed in red. The 487^{1+} ion corresponds to the C-terminal part of the toxins and was not fully sequenced. Hydroxyproline is displayed as P*. B: precursor sequence of nemertide alpha-1 found in the *L. longissimus* transcriptome. ER signal, pro-region and mature toxin are marked. ** premature toxin cleavage site. *** post sequence cleavage site.

4.2.5 Synthesis, folding and solution NMR structure of alpha-1

With the sequence at hand, nemertide alpha-1 was synthesized and folded (figure 5) in sufficient amounts for structural elucidation using solution NMR spectroscopy. The analysis revealed a compact structure with an ICK motif at the core, and a single short alpha helix in between Cys II and III (loop 2). In addition, an aromatic patch comprising Phe 8, Thr 22 and Phe 24 could be identified at one side of the surface, figure 12.

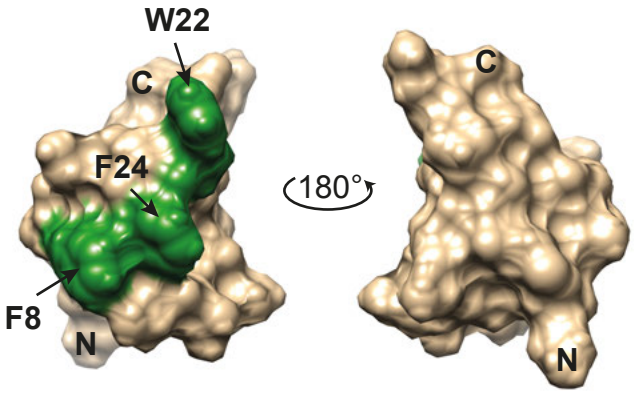

Figure 12. Surface representation of nemertide alpha-1. The aromatic residues Phe-8, Trp-22 and Phe-24 are displayed in green. Aromatic residues are only found on only one side of the molecule.

4.2.6 Available transcriptomic data from nemertean worms

From the start of the nemertean toxin project to the publication of **paper II**, the amount of transcriptomic data had increased from null to 17 transcriptomes, including our own, in public databases (i.e. NCBI SRA). In **paper II** all of them were screened using the BLAST+ suite and the Fuzz-pro/tran/nuc (as implemented in EMBOSS[35]) programs for the presence of alpha nemertides. The search resulted in five additional full-length and one partial toxin sequence (alpha 3-8) with high similarity to alpha-1 and -2, figure 13. All alpha nemertides were found to be concentrated to one taxon (the Lineidean family) of nemertean worms.

In addition to the nemertide alphas found, also sequences similar to nemertide beta-1 and the parborlysins/cytolysin A -toxins were detected in the transcriptomes.

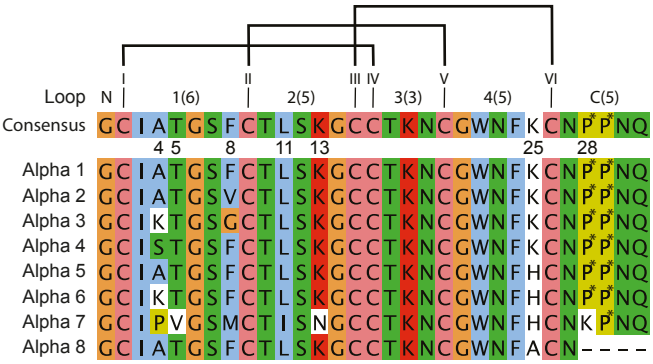

Figure 13. Alignment of alpha nemertides 1-8. ClustalW alignment of alpha 1-8 found in the transcriptomic data (except alpha-1 and -2, which was also detected on the mucus from *L. longissimus*). The cystine framework is displayed in bars, cysteine residues in roman numerals and positions with variability in numbers. The alpha-8 sequence is probably partial but does however contain the full pre-toxin sequence (not shown). P*; hydroxyproline as inferred by the presence in alpha-1 and -2. The consensus sequence is identical to alpha-1.

4.2.7 Activity of alpha nemertides

In **paper II** the bioactivity of the prototypic nemertide alpha-1 in shore crab (*Carcinus maenas)*, cockroaches (*Blaptica dubia*), and in electrophysiological experiments on a small set of voltage-gated sodium channels (VGSC) was investigated.

Nemertide alpha-1 was highly active when injected into both crabs and cockroaches: paralysis and death followed doses of 0.3 pmol/g (1 µg/kg) in crabs, while cockroaches were paralyzed at 2.1 pmol/g (7.1 µg/kg). These values are in the same range, and lower, relative to the most potent insecticidal spider peptide toxins on their target prey (10-100 pmol/g)[81–83].

Electrophysiological experiments were performed on three arthropod VGSCs (*Drosophila melanogaster*: $DmNa_V$ 1, *Blatella germanica:* $BgNa_V$ 1, and *Varroa destructor:* $VdNa_V$ 1) and five vertebrate VGSCs ($Na_V1.1$, 1.4, 1.5, 1.6 and 1.8). EC_{50} values were determined for $BgNa_V1$ and $Na_V1.6$ with activity in the low nanomolar range for $BgNa_V1$ and two orders of magnitude higher for $Na_V1.6$ (figure 14), indicating a possible use of alpha-1 as an insecticide.

The mode of interaction was proposed to be through delaying fast inactivation via interaction with neurotoxin binding site three[84]. Furthermore, pulsing experiments displayed that the interaction is independent on whether the channel is open or closed.

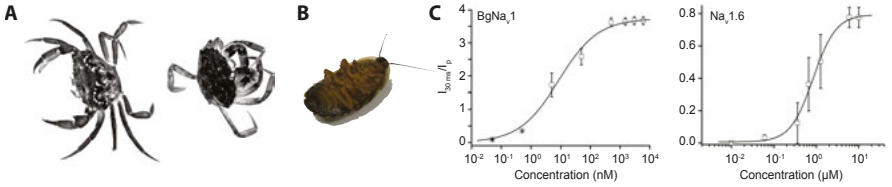

Figure 14. Activity profile of nemertide alpha-1. A: left control (injected with ASW) *Carcinus maenas*. Right: *C maenas* injected with alpha-1. B: *B. dubia* injected with alpha-1. C: dose response curves; left: BgNaV1 (EC_{50} 8.6 nM) right: NaV1.6 (EC_{50} 0.8 μM).

Key results and observations from **paper II**

- A new family of ICK peptides, the alpha nemertides, was discovered in the mucus of *L. longissimus*.
- Transcriptomic analyses of all available nemertean transcriptomes suggest that the presence of alpha nemertides is restricted to the Lineidean family of nemertean worms.
- A SPPS based route for production of nemertide alpha-1 was established and the oxidative folding into the native structure was achieved.
- Nemertide alpha-1 displays high potency *in vivo* activity in *C. maenas* and *B. dubia*.
- Electrophysiological measurements show that Alpha-1 is selective in favor for insect VGSCs over vertebrate ditto.
- A solution-NMR structure of nemertide alpha-1 was obtained. The structure reveals an aromatic patch localized to one side of the molecule.

4.3 Characterization of nemertide alpha 1-7 in *Artemia* assay and VGSCs (paper III)

In **paper III** the exploration of the alpha nemertides discovered in **paper II** was continued. The family of the alpha nemertides was characterized in *Artemia* assay and electrophysiological measurements in vertebrate (Na_VS 1.1-1.8) and insect ($BgNa_V1$) VGSCs. In addition, the folding procedure was modified to avoid peptide aggregation. Nemertide alpha-8 was excluded from the study since the C-terminal part was not fully known.

4.3.1 Synthesis and folding of alpha nemertides

In **paper III**, the alpha nemertides (3-6) were assembled on 2-chlorotrityl resin instead of the high swelling HMPA-resin used in **paper II**. The change was prompted by high maintenance load: repeatedly clogged tubing of the automated peptide synthesizer. The change to 2-chlorotrityl resin did not alter the yields to any noticeable extent. The synthesis of alpha-7 was troublesome, probably due to oxidation of Met8, and the peptide could not be produced in adequate amounts. This peptide was instead purchased in reduced form from a commercial vendor.

The initial folding condition used in **paper II** was prone to aggregation of peptide, resulting in low yields of folded product. In **paper III** the folding procedure was modified to minimize aggregation of peptide, and includes a thorough pre mix of the reduced peptides and a switch from isopropanol to DMSO[85] in the folding buffer. No aggregation could be seen during the folding using the improved protocol. Alpha-7 did however again prove to be problematic; only low sub-milligram amounts of folded toxin could be produced in pure form.

4.3.2 *Artemia* assay divides alpha nemertides into two groups.

To explore possible differences in toxicity among the alpha nemertides, a microwell *Artemia* assay developed by Solis *et al.*[86] was employed. The assay was performed in a microwell (96 well) format, and required only sub-milligram amounts of peptide. Another advantage over the *C. maenas* assay used in **paper II**, is that the availability of the organisms is season-independent.

All alpha nemertides tested were found to be highly potent with EC_{50} values in the μM range. A pattern was observed where alpha-1, -4 and -5 exhibited high activity with EC_{50} values in 0.4-0.5 μM, while alpha-2, -3 and -6 were approximately one order of magnitude less potent (EC_{50} 3.1-5.2 μM), figure 15. Alpha-7 was omitted from the *Artemia* microwell assay due to its low availability.

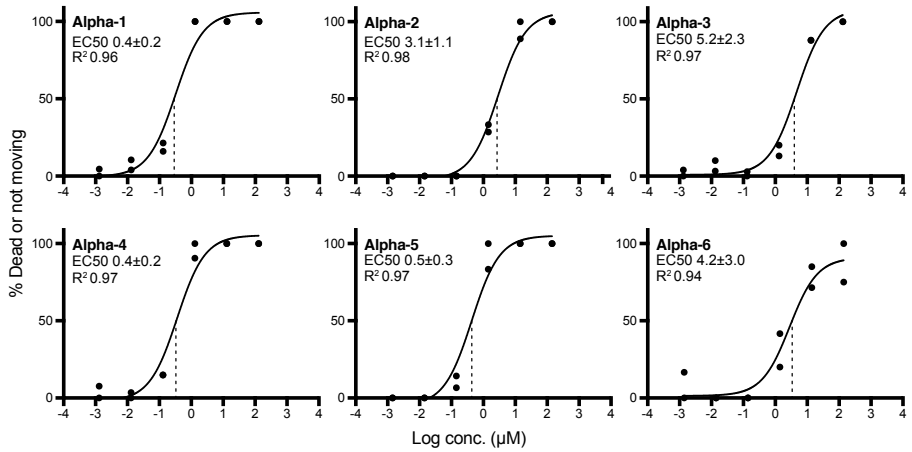

Figure 15. Artemia toxicity of nemertide alpha 1-6. Dose response curves, with EC_{50} values for alpha 1-6 in *Artemia* assay. All experiments were performed in duplicate. All data points represent individual wells and the lines are the fitted nonlinear regressions. Dotted line marks the EC_{50}.

4.3.3 Electrophysiology displays Na_V subgroup specificity for alpha nemertides

The alpha nemertides (1-7) were assayed against $BgNa_V1$ and Na_Vs 1.1-1.8 expressed in *X. laevis* oocytes. The resulting data was normalized to the results from $BgNa_V1$ for each toxin, table 1, and compared for specificity.

Table 1. *Normalized (to BgNaV1) activity for nemertide alpha 1-7 in Na_V 1.1-1.8.*

alpha	BgNaV1 (*)	Nav1.1	Nav1.2	Nav1.3	Nav1.4	Nav1.5	Nav1.6	Nav1.7	Nav1.8
1	1 (3.3)	14.4	41.8	15.7	16.9	16.1	28.0	8.9	n.a.
2	1 (33.5)	1.4	1.1	1.5	13.2	1.7	15.6	14.9	n.a.
3	1 (37.5)	1.3	1.4	1.4	1.5	1.1	1.0	1.0	n.a.
4	1 (4.3)	8.3	12.1	1.2	1.3	2.5	11.1	7.3	n.a.
5	1 (3.0)	13.1	20.0	1.2	2.0	17.0	8.6	9.4	n.a.
6	1 (1.0)	3.0	9.3	40.6	17.8	82.8	14.0	37.4	n.a.
7	1 (3.7)	18.1	5.3	17.9	85.3	16.4	15.5	13.6	n.a.

(*) normalized to the activity on $BgNa_V1$ for alpha-6. n.a.: not active.

Alpha-1 had the highest specificity towards the insect channel compared to the other channels tested, but in fact it was not the most potent peptide in the assay; alpha-6 (EC_{50}: 2.6 nM) was approximately three times more potent and alpha-5 (EC_{50}: 7.8nM) was approximately equipotent with alpha-1 (EC_{50}: 8.6 nM) on $BgNa_V1$.

The grouping obtained in the *Artemia* assay; that alpha-1, -4 and -5 were one order of magnitude more potent than the other toxins, was partly true also in the electrophysiology experiments with the exception of alpha-6. The

discrepancy of activity in the *Artemia* assay contra the electrophysical assays for alpha-6 conveys the importance of employing the combination of *in vivo* and *in vitro* assays.

4.3.4 Sequence variation among alpha nemertides, and presence in different species.

The alpha nemertides differ in only seven out of 31 positions, with alpha-7 being the most divergent with all seven positions differing from the consensus sequence. Most of the variations are found in the N-terminal part, while positions 14-24 are fully conserved. The aromatic patch (Phe 8, Trp 22 and Phe 24) is conserved with exception of position 8. Furthermore, most of the variation (i.e. positions 4, 5, 8, 25 and 28) is localized to the same side of the molecule.

Alpha nemertides are expressed in pairs (*L. ruber* expresses three: alpha-1, 2 and 7) on RNA level. The pairs consist of one highly potent toxin and one less active toxin, as judged from *Artemia* assay, figure 16.

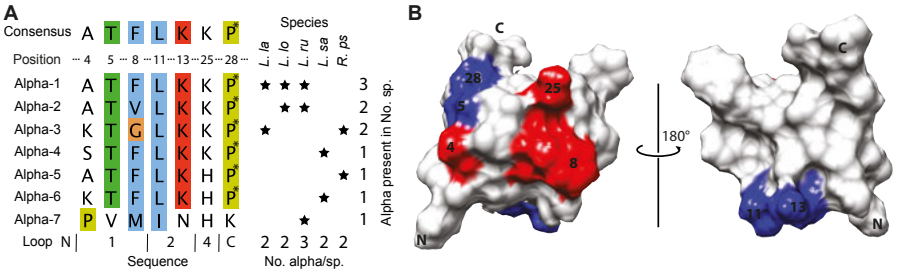

Figure 16. Sequence and structural variation of alpha nemertides. A: alignment of the variable positions, with consensus sequence, and loop numbers marked. L. la; *Lineus lacteus*, L. lo, *L. longissimus*, L. ru, *L. ruber*, L. sa *L. sanguens* and R. ps, *Rieseriellus pseudolacteus*. B: surface representation of alpha-1 (RSCB id: 6ENA) with the variable positions marked in red (variable positions in alpha 2-6; 4, 8 and 25) and blue (variable positions only found in alpha-7; 5, 11, 13 and 28). Compare to figure 12 for the aromatic patch.

Key results and observations from **paper III**:

- Alpha nemertides are expressed at least in pairs in the species investigated.
- Folding conditions for alpha nemertides were modified to reduce aggregation.
- Alpha nemertides show variable specificity profiles in VGSCs ($Na_V1.1$-1.8, and $BgNa_V1$).

- Small changes in the sequence can have impact on the selectivity and potency of alpha nemertides as displayed by *Artemia* and VGSC assays.

4.4 Characterization of epidermal mucus from the Antarctic nemertean *Parborlasia corrugatus* (Paper IV).

In **Paper IV** the mucus of *P. corrugatus*, one of the top predators in the Antarctic benthic fauna, was investigated. The worm can reach lengths up to three meters and a diameter of approximately two centimeters. *P corrugatus* is common in shallow waters, but have also been found at 3500 m depth[87].

Both crude body extracts[88] and mucus preparations[69] of *P. corrugatus* have previously been investigated for toxicity. The crude body extract was found to be toxic to Antarctic sea urchin spermatozoa, and to have deterrent effects in two fish species[88].

In a study by Berne *et al*. the mucus was found to contain small (≈10 kDa) hemolytic proteins; parborlysins[69]. Later investigations revealed seven parborlysin isoforms through PCR[89] and subsequent sequencing. Parborlysins contain six cysteines, and are to a high extent comprised of alpha helices, as determined by homology to another nemertean protein; cytolysin A-III. The authors also presented a putative structure for one of the isoforms[89]. However, the structure must be considered somewhat flawed; only one disulphide bridge was accounted for.

In **paper IV** the content of the mucus from *P. corrugatus* collected during the Diversity and structure of Antarctic benthic communities (DISTANCOM) expedition, by Prof. Conxita Ávila, University of Barcelona was investigated. The exploration of the mucus was focused on peptides and small proteins using UPLC-QToF, transcriptomics and bioassays.

4.4.1 Initial screening for cysteine-rich peptides in the *P. corrugatus* mucus

Resuspended lyophilized mucus was analyzed using UPLC-QToF, revealing a complicated mass-chromatogram containing a plethora of peaks. The peaks were found to be in the mass ranges of the known classes of peptides from nemertean toxins: alpha nemertides, beta nemertides and parborlysins/cytolysins. In addition, the profile revealed a peak that eluted well before the putative parborlysins, with a mass of 8751 Da.

To explore if the peptides and small proteins detected contain disulphide bridges, an aliquot of the sample was reduced and alkylated and re-analyzed. A the comparison of the two samples showed four highly abundant peptides

with three disulphide bonds each and one peptide containing four disulphides, figure 17.

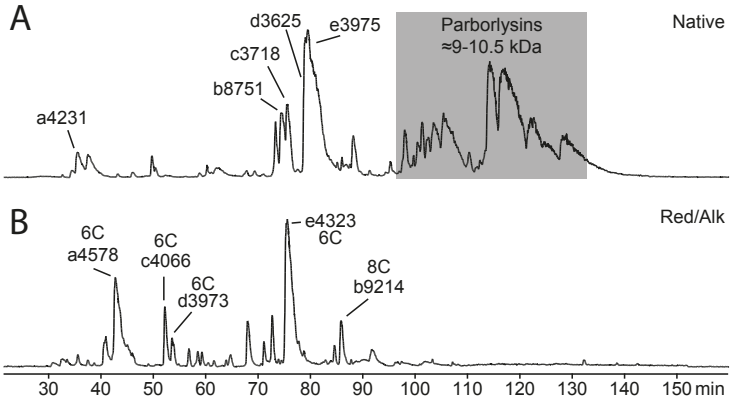

Figure 17. Mass chromatograms of native and alkylated mucus from *P. corrugatus*. A: Native mucus. Putative parborlysins are boxed in gray. B: alkylated and SEC filtered sample. Lower case letters followed by numbers mark the identified peptides and their masses. Numbers combined with capital C denotate the number of Cys residues engaged in disulphide bonds, i.e. the native a4231 corresponds to the alkylated a4578 6C.

4.4.2 Isolation and partial sequencing of 3624 and 8750 Da peptides.

Two peptides, 3624 Da (6C) and 8750 Da (8C) were isolated using HPLC-UV from lyophilized mucus and partly sequenced using enzymatic cleavage and subsequent UPLC-QToF MS/MS analysis.

The alkylated and trypsinated 3624 Da peptide gave rise to only one fragment (both 2^+ and 3^+) suitable for sequencing by MS/MS; a stretch of 15 amino acids and an N-terminal part where no deducible ions could be sequenced.

For the 8750 Da protein, fragments from tryptic, Glu-C and chymotryptic cleavage were identified amounting to a total of 52 amino acid positions. None of the isolated peptides could be fully sequenced using MS/MS alone, transcriptome sequencing was used next.

4.4.3 Isolation of pure parborlysin isoforms is problematic

The purification of individual parborlysin isoforms proved problematic, and as a result a fraction of mixed peptides were used for functional assays. The problem of yielding pure parborlysins has been highlighted already by Berne *et al.*[69]. They were also unable to procure pure peptides even after exhaustive attempts using different chromatographic techniques including size exclusion, ion exchange and RP-HPLC.

In the experimental work of this thesis, not even UPLC-QToF with a gradient over 180 minutes could separate the isoforms completely, figure 17.

4.4.4 Transcriptomes

A transcriptome from a transversal section of a *P. corrugatus* specimen stored in RNAlater was acquired. This sample was sent for next generation sequencing (RNA-seq, Illumina 2500 Hiseq, Trinity assembly), and returned a 67.148 contigs (43298283 bases), which could be translated into 968918 ORFs of a minimum length of 100 amino acids.

There is also a transcriptome available from 2014 (NCBI id: SRX731467)[90,91] in the public databases. This transcriptome was assembled using the implementation of Trinity at the official Galaxy[92] server, resulting in 8174 contigs comprised of a total of 2298549 bases. However, only 1393 ORFs of a length above 100 amino acids could be identified.

4.4.5 Identification of 3624 and 8750 Da peptides in the transcriptomes

The MS/MS sequenced part of the 3624 Da peptide was matched to two contigs in the transcriptomes, albeit that they do not match completely, figure 18 A. The lack of fragments from the C-terminal part of the peptide may be explained if the sequence is highly similar to the hits also in that region. The fragments identified from the enzymatic cleavage of the 8750 Da peptide also match a contig in the transcriptomes, figure 18B.

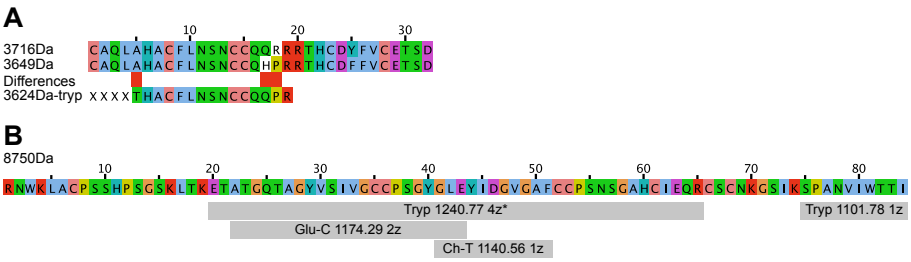

Figure 18. Identification of contigs matching the partly sequenced 3624 and 8750 Da peptides. A: tryptic fragment from 3624 Da (bottom), together with the sequences found in the transcriptomes. Differences are highlighted in red. X; unidentified amino acid(s). B: Contig matching the enzymatic fragments from the 8750 Da small protein. * not fully sequenced; matched to a high abundance of ions from predicted spectra.

The 8750 Da peptide contains a cysteine pattern (C-CC-CC-C-C-C) not previously found in the conotoxin database Conoserver[19], or in the arachnoid toxin database Arachnoserver[13]. However, in a general search in the NCBI non

redundant database, an uncharacterized protein (NCBI accession: XP_012564781.1) from the fresh-water polyp *Hydra vulgaris*, sharing the cysteine pattern, was found.

4.4.6 Activity in *Artemia* and membrane permeability assays of extracts, fractions and isolated peptides

The toxicity of five extracts and fractions (Body extract, mucus, mucus LMW fraction, mucus HMW fraction and a fraction of parborlysins) from *P. corrugatus* were tested in *Artemia* microwell assay.

Quantification of the extract and fractions for *Artemia* was performed both by weight, except for the parborlysin fraction which was performed by UV-absorbance measurements at 280 nm (using an average extinction coefficient for all known parborlysin and cytolysin sequences). The toxicity observed in the *Artemia* microwell assay is shown in table 2.

Table 2. *Toxicity of P. corrugatus extracts and fractions in Artemia microwell assay*

Fraction	mg/ml	Relative toxicity (%)
Body	10	100
	1	6.25
Mucus[a]	12.35	0
LMW	21.4	0
HMW	0.28	100
Parborlysins	1.0*	100
	0.1*	7.1

In control wells, with only MQ-water (75 nauplii, 13 wells) the toxicity was 1.3%. *: as measured at 280 nm. Experiments were performed in duplicate.

The membrane disruptive activity of parborlysins, 8750 Da and 3624 Da peptides was measured in a membrane permeabilization assay[93]. In short, *E. coli* polar lipid film extracts were re-suspended together with carboxyfluorescein, a fluorescent dye which is self-quenching at high concentrations. The resulting suspension was extruded through a membrane to form liposomes with dye trapped inside. The purified liposomes were then incubated with sample peptides, and the membrane leakage was measured by fluorescence as the dye leaks out to the surrounding media.

The parborlysin fraction displayed high membrane disruptive activity, with an EC_{50} of 0.07μM. The purified 8750 Da and the 3624 Da peptide also exhibited potent activity in this assay with EC_{50} values of 0.9 and 1.8 μM respectively. The benchmark bee venom derived toxin melittin was used as a reference and disrupts membranes with an EC_{50} of 0.3 μM, figure 19.

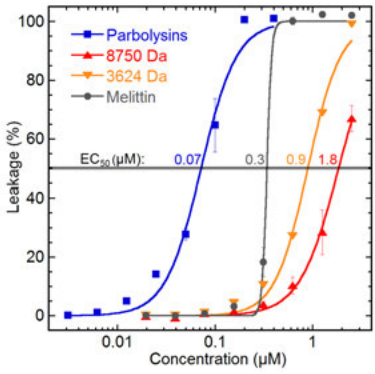

Figure 19. Membrane permeabilization induced by mucus peptides from *P. corrugatus*. Each data point is the mean of four replicates, with standard deviation marked in bars, except for the 3624 Da peptide where only one replicate was possible. Absolute EC50 values are highlighted by a vertical line.

4.4.7 Transcriptome mining for nemertides in P. corrugatus transcriptomes.

The available *P. corrugatus* transcriptomes were mined for nemertide alpha, beta/neurotoxin B and cytolysins/parborlysins using NCBI BLAST+[36,94] and MEME[95] suites. The combined results were manually curated after alignment, and passed through the conoprec[19] webservice or signalP[96] for prediction of mature toxins. In total 16 contigs with similarity to known full-length nemertides were identified on the RNA level: ten alpha-like, four beta/B-neurotoxin like and two new parborlysins/A-cytolysin like contigs, figure 20. In addition to the contigs mentioned above, five sequences similar to the 3624 Da peptide and one additional contig similar to the 8750 Da peptide, were identified.

None of the previously described parborlysins[69,89] were found on either peptide level or RNA level. The reason for this discrepancy is unclear but suggests that this sample either represents a different population, or merely that individual variation is significant when it comes to parborlysins.

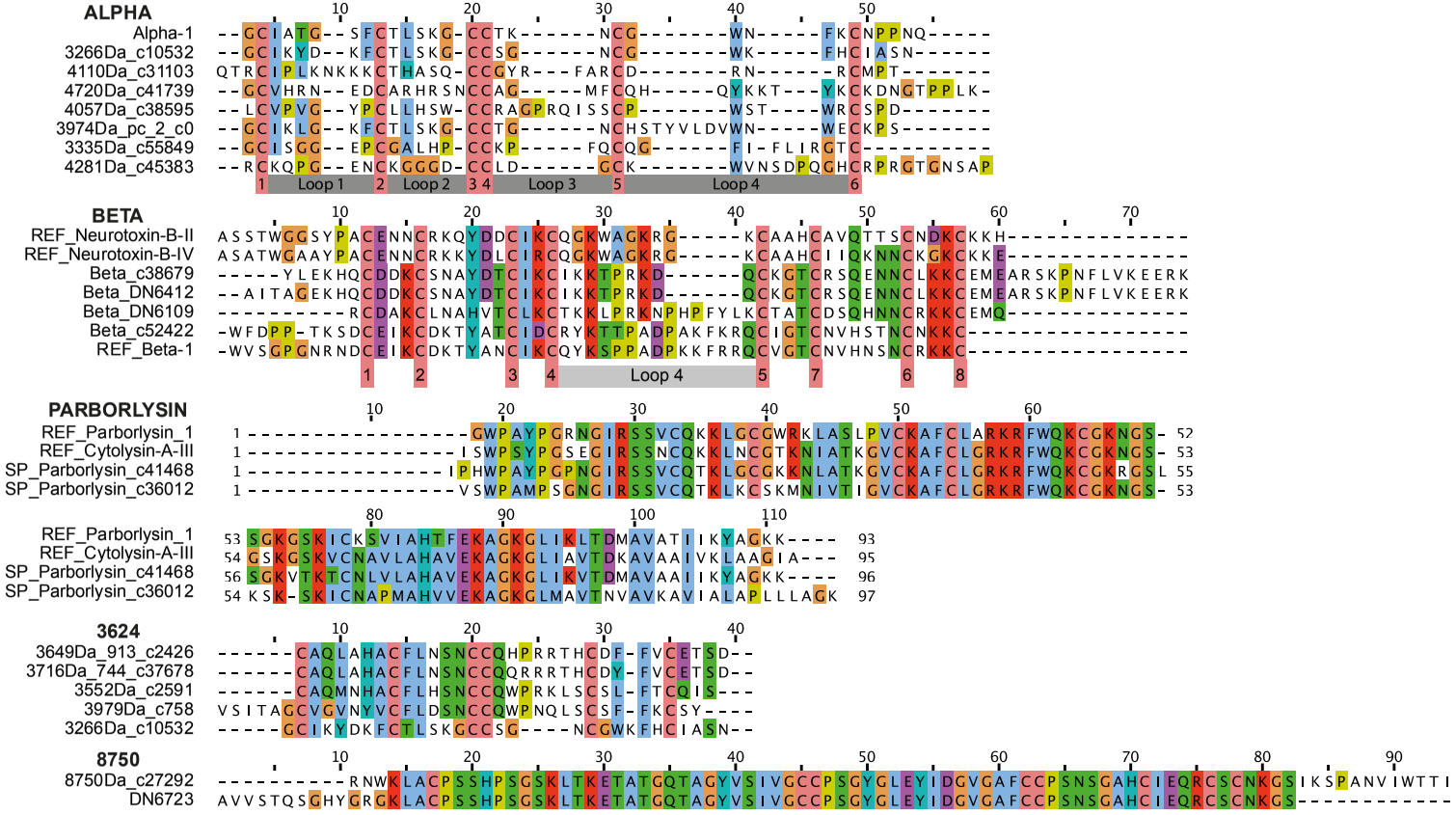

Figure 20. Full length nemertide sequences, including sequences those similar to 3624 and 8750 Da peptides, found in *P. corrugatus* transcriptomes. REF: selected reference (query) sequences

4.4.8 *In silico* functionality prediction of contigs found in the *P. corrugatus* transcriptomes using CSPred

The possible functionalities of the contigs found in section 4.3.6 were predicted using the CSPred online service. CSPred is a prediction model specialized for the of functionality prediction of disulphide-stabilized peptides into five categories from primary structure: ion channels, antimicrobial (AMP), acetylcholine inhibitor, serine protease inhibitor and hemolytic activity[44].

Parborlysin and cytolysin A like sequences scored high (100%) in the AMP category, and slightly lower in the related hemolytic activity probability (34-74%). These values are in concordance with our experimental data on membranes, and earlier studies on the hemolytic activity of parborlysins[69].

The 8750 Da peptide was predicted to exhibit AMP activity with a probability score of 83% and scored low probabilities in the other categories (1-6%). Among the sequences with similarity to the 3624 Da peptide, no category scored a higher probability than 57% (ion channel activity).

Key results and observations from **paper IV**:

- The mucus of *P. corrugatus* contains a complex mixture of peptides and proteins including four peptides containing six cysteines, and one small protein harboring eight cysteines detected by UHPLC-QToF
- One six cysteine peptide (3624 Da) partly sequenced by MS/MS, share its cysteine pattern with nemertide alpha-1.
- One eight cysteine protein (8750 Da) was fully sequenced by the combination of MS/MS and transcriptomic data; revealing a unique cysteine pattern.
- Transcriptomic data revealed a small set of contigs with similarity to known nemertides.
- A fraction of putative parborlysins shows extraordinary membrane disruptive activity in membrane permeabilization assay, while 3624 Da and 8750 Da peptides display potent membrane disruptive activity.

5 Discussion and future perspectives

5.1 Pharmacognosy

The majority of the experimental work described in this thesis have been conducted in the only division for pharmacognosy in Sweden (currently research group for pharmacognosy). The term 'pharmacognosy' [Greek; Pharmakon, drug; gnosis, knowledge], has historically encompassed the investigation into all things relating to drugs from natural sources. This includes identification of raw material by its taste and smell, as well as quality determination and detection of adulterations.

An updated definition was proposed by Larsson *et al.*, which also include the *in silico* approach in the pharmacognostic toolbox[97]. Today, pharmacognosy is a multi-disciplinary subject connected to natural products: the connections between organism(s), biological activity, chemical structure and big data[97]. The projects in this thesis span all the areas of the newer definition.

The overall aims of the projects included in the thesis were to discover, synthesize and characterize cysteine-rich peptides from marine worms (*Nemertea* sp.) and plants (*Viola* sp.).

5.2 Cyclotides and suspension cell culture from *V. uliginosa,* Paper I

Paper I differs from **papers II-IV** in that the target molecules, cyclotides, are already well-studied: five review-papers have been published in the last two years only (2017-2019)[22,46,98–100] covering different aspects of cyclotides. My part of the work in **paper I** was focused on the techniques used. For one of the first times, generation sequencing was used to study cyclotides. Previously generated genetic data specific for cyclotide prospecting have utilized expressed sequence tags, EST[101], and cDNA libraries[102].

In **paper I** a total of 20 cyclotides was identified. This number is in lieu with previous findings by Burman *et al*. where a large screen of different *Violaceae* species revealed 1-25 individual cyclotides per plant[56]. In-depth analysis of transcriptomic and two flavors of mass spectrometric techniques revealed 164 cyclotides in *V. tricolor*[10]. Also in venomous animals large numbers of sequences can be found in a single species e.g. cone snails[103], while

only 2-3 alpha nemertides were found per species on transcriptomic level (**paper II-III**).

One advantage of transcriptomes is that they are versatile and can be used for different purposes. The *V. uliginosa* transcriptomes were also used to screen for plausible cyclotide processing enzymes, i.e. enzymes similar to butelase-1 from *C. ternatea*[62]. One sequence with a pairwise identity of approximately 70% was identified. The cyclotide precursors differ between *C. ternatea* and *Viola* sp. The *C. ternatea* precursors also encode for plant albumin-1[104], and it is not impossible that albumin processing enzymes are involved in the cyclization also of the cyclotides. This is in fact the case in another class of head to tail cyclic peptides: the sunflower trypsin inhibitors, SFTI[105].

Investigation into the newly sequenced genome of *V. pubescens* (NCBI accession: PRJNA379317) will hopefully soon shed light on the matter of cyclotide processing and related gene clusters in *Viola* sp.

In **paper I**, a suspension cell culture from *V. uliginosa* was proposed as a high yielding model system for production of native cyclotides, as well as for the study of biochemical processes involved in *Viola* cells.

Perhaps cell cultures also can be utilized for recombinant expression of cystine knotted peptides, e.g. alpha nemertides. Tobacco plants have, in analogous manner, been modified to produce a spider toxin as an insecticide[106], and carrot cell suspension culture has successfully been used to express a protein (Glucocerebrosidase, ELELYSO®) for replacement therapy in the rare Gaucher's disease[107]. ELELYSO® was FDA-approved in 2012[108]. The advantage of using a native cyclotide producing cell culture is the presence therein of a complete processing machinery for knotted peptides.

5.3 Nemertides, papers II-IV

5.3.1 Discovery and posttranslational modifications

The discovery of a new family of ICK nemertides was instigated by the absence of another potent toxin, TTX, in the mucus of *L. longissimus*[77]. Upon size exclusion and UPLC-QToF analysis it was evident that the mucus contained a small set of peptides and small proteins.

The most abundant of the 3 kDa peptides, later named nemertide alpha-1, was isolated and partly sequenced by means of MS/MS. The full sequence could however not be elucidated by MS/MS alone, and a transcriptome was procured.

A contig harboring the sequenced fragments could be found, but the restriction sites for the mature toxin were still unclear. No truncation of the sequence corresponded to the experimentally determined mass. This

discrepancy was later explained by occurrence of PTMs, namely hydroxylation of prolines. These results can be seen as a warning against the use of only one technique in sequence determination.

The frequency of PTMs present in sequence databases have probably been underestimated, since most of the depositories consist of data from large scale nucleotide sequencing efforts, ignoring PTMs. This is a problem for short peptides in proteomic experiments, where identification may be based on a small number, or even single, fragments from proteolytic digests.

PTMs (other than disulphides) can influence a peptide's activity and folding properties. The effect of hydroxylation of proline into hydroxyproline has been investigated in a diverse group of conotoxins[109]. The effects were found to be hard to predict: hydroxylation led to a loss of activity for some conotoxins, while no difference was displayed in others. The same inconsistency could be seen in folding trials[109].

Hyp residues have been identified in the nemertean neurotoxin B-II and B-IV[110], prior to alpha-1. The effect of the modification in the B-neurotoxins did however not shift the activity[110,111].

Initial experiments on alpha-1 without hydroxylated prolines does not indicate any discernable differences in activity (VGSCs and *Artemia*), or in folding (data not shown). If these results can be verified, the production of potent alpha nemertides in bioreactors may be possible.

5.3.2 Bioinformatic considerations from short peptides and stretches of amino acids

Since the advent (and drop in price) of large scale parallel systems for sequencing of transcriptomes and genomes the amount of data in public repositories have exploded; in 2017 the annual increase of sequenced genomes was close to 50%[112]. Even though attempts in parallelizing the sequencing of proteins are underway, the rate of protein sequencing is limited to data sets produced in mostly mass-spectrometry based proteomic experiments[113]. The annual increase of experimentally determined protein structures (NMR and crystallography) deposited was 9% as of 2017 with a total of ≈130000 entries[112].

The most common methodology for homology identification of sequence data, the BLAST suit[36], is in its default settings optimized for proteins and nucleotide sequences rather than short peptides (e.g. toxins) and short stretches of amino acids obtained by MS/MS de novo sequencing. This phenomenon may be illustrated in the case of attempts to identify toxin sequences in nemertean genetic data using bioinformatics alone[80,114]; the relatively short alpha nemertides were not identified, illustrating the need for chemical evidence of peptide presence. Sequences related to the longer parborlysins/Acytolysins were however found along with other protein toxin sequences[80,114,115].

In this thesis, regular expression (regex) based patterns were used for initial identification of short stretches from MS/MS data (**papers I-II and IV**). Regex patterns can find literal stretches of sequences in texts (e.g. transcriptomic data) and are highly customized using e.g. wildcard characters and letter classes. Prosite utilizes a flavor of regular expression patterns optimized for amino acid sequences. The prosite patterns are also used for rapid classification of proteins and peptides[15].

With only limited information from e.g. reduction and alkylation resulting in the number of Cys residues, and some short amino acid sequences from MS/MS, patterns can be constructed to filter out possible hits from a genetic database. In this case the procedure involves two steps: 1: Find all sequences containing the sequenced MS/MS fragment. 2: In the sequences from step 1, find all sequences containing the number of Cys residues present. The resulting sequences are then compared to LC-MS data of enzymatically cleaved fragments. The steps 1 and 2 can be automated, and a similar approach has been employed in the search for cyclotides from genetic data[16].

5.3.3 Notes on the production of alpha nemertides

In **paper II** and **III** the synthesis of alpha nemertides were performed by SPPS and subsequent oxidative folding yielding toxins in amounts enough for structural and functional characterization. It should however be noted that only two alpha nemertides, alpha-1 and -2, have been observed on peptide level.

This highlights some aspects of using genetic data: it is possible to find interesting sequences, and synthesize them without ever having seen them, or know if they are actually expressed on peptide level. On the other hand, it is not possible to absolutely confirm the correct folding or the presence of PTMs.

Homology to the experimentally determined peptides, both in activity and folding behavior, are strong indications that the correct folding was achieved also for alpha 3-7.

5.3.4 Biological functions of nemertides and mucus: hunting or defense or both?

The localization of alpha-1, -2 and beta-1 in the mucus and epidermal layer of *L. longissimus* displayed in MALDI-IMS experiments (**paper II**), indicates a defensive function of the peptides. However, an observation suggests that the peptides might have an alternative function: in a feeding session, where *Artemia salina* nauplii were added to a petri dish containing artificial sea-water and living *L. longissimus* specimen, the worm circled the container, leaving a

mucus trail in which the artemias were immobilized. The worm did not eat the immobilized nauplii in the timeframe of the experiment, but it suggests a hunting function of the mucus.

Nemerteans also express membrane disruptive compounds; parborlysins and cytolysins. In fact all investigated species harboring alpha nemertides also express beta- and parborlysin-like nemertides[115]. In **paper IV**, the mucus of the Antarctic *P. corrugatus* display a more complex mixture of compounds than that of *L. longissimus* (**paper II**). In addition to alpha-, beta- and parborlysin/A-cytolysin -like toxins in larger numbers, *P. corrugatus* also secrete a new type of nemertide (8750 Da) with a unique cysteine pattern.

The mixture of membrane disruptive- and neurotoxins could be part of a system where the membrane disruptive peptides facilitates the penetration of neurotoxins, analogous to frog skin secretions[116]. The components of nemertean mucus should be assayed together in assays in a controlled manner to investigate possible synergistic effects.

5.3.5 Voltage-gated sodium channels as targets for alpha nemertides

Voltage-gated sodium channels are central in signal transmission in excitable cells. Channelopathies in VGSCs include several diseases including epilepsy[117], pain disorders[118] and cardiac syndromes[119] and are targets for pharmacological intervention.

The VGSCs have three main states: closed, opened and inactivated. They consists of one large, pore-forming, alpha- and 1-4 beta-subunits[120,121]. The pore-forming alpha unit is made up by four domains (DI-IV) consisting of six transmembrane segments (S1-6) each, where S1-4 are responsible for voltage sensing and S5 and S6 forms the ion pore. A linker-loop connecting DIII and DIV is employed for the inactivation of the channel[122], figure 21.

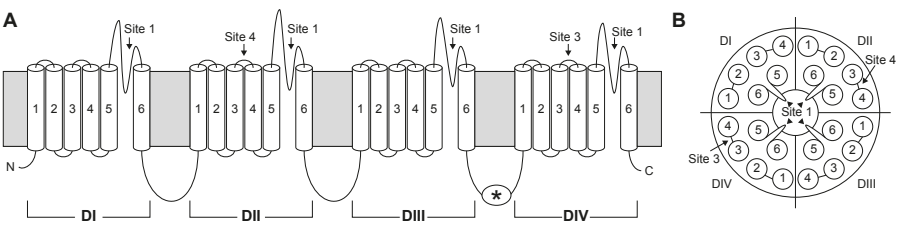

Figure 21. Voltage-gated sodium channel organization and toxin binding sites. A: Schematic VGSC in linear format, transmembrane segments numbered 1-6. Domains D I-IV. Peptide toxin binding sites are marked with arrows. * the loop between DIII and DIV involved in inactivation. B: schematic top view organization of VGSC modified from Mattei and Legros[123], numbers denotates the segments (circles). Lines between segments; extracellular loops. Sites 1, 3, and 4 annotated with arrows.

Toxins interact with channels in mainly six different binding sites, site 1-6. Peptide toxins have been found to bind to sites 1, 3, 4 and 6[124], figure 21.The interactions of toxins with the different sites are distinguishable from each other through electrophysiological experiments.

Site 1 is located in the pore (in the loops between S5 and S6 of all four domains) and binding results in a nearly complete blockage of ion current. TTX is perhaps the best studied pore-blocking toxin binding to site 1. Site 3, the most relevant site in this thesis, is the proposed binding site of alpha nemertides. The site is located on the extracellular loop between S3 and S4 in DIV. Toxins interacting with site 3 exhibit a characteristic slowing down of fast inactivation by a conveying conformational change in S4 of DIV[124]. Site 4 is located in DII, between S3 and S4. Binding to site 4 renders constant activation of the channel[124,125]. The localization of site 6 is still elusive, but binding to it results in a slow-down of the inactivation of the channel.

The vertebrate VGSC isoforms ($Na_V1.1$-8) display different expression patterns linked to a series of channelopathies (table 3).

Voltage gated sodium channels are classical targets of cysteine rich toxins of both terrestrial[126] and marine[123] origin, and have been used to characterize the different isoforms. In fact, a scorpion toxin was instrumental in the discovery of the alpha and beta subunits of VGSC through crosslinking experiments[120].

In nemertean research, the sodium channels were first, inconclusively, indicated as the targets of peptide toxins for neurotoxin B-IV[127,128]. The decision to investigate the VGSCs as targets for alpha nemertides in **paper II** and **III**, was partly based on the effects observed in Strand *et al.*[77] and partly based on the fact that many cysteine rich peptides from marine origin interact with ion-channels[123]. In **paper II** and **III** a total of eleven VGSC-isoforms have been evaluated for nemertide alpha activity, and EC_{50} values were recorded for eight of them.

Table 3. *VGSC isoforms tested in this thesis; their source, localization and EC_{50} values for alpha-1*

Na_V isoform	**Source**	**Localization and channelopathy[125]**	**EC_{50} alpha-1 (nM)**
$BgNa_V1$	Insect		8.6 ± 2.9
$VdNa_V1$	Insect		-
$DgNa_V1$	Insect		-
1.1	Rat	CNS, epilepsy	124.1 ± 28.7
1.2	Rat	CNS, epilepsy	359.6 ± 89.8
1.3	Rat	CNS, epilepsy	135.4 ± 76.3
1.4	Rat	skeletal musscle, paralysis	145.5 ± 57.5
1.5	Human	Heart, arrythmia	138.3 ± 25.5
1.6	Mouse	CNS, epilepsy, movement	240.4 ± 22.3
1.7	Rat	PNS, pain	76.5 ± 33.9
1.8	Rat	PNS, pain	n.a.

CNS: central nervous system, PNS: peripheral nervous system, -: EC_{50} not determined, n.a.: not active at 6 μM alpha-1.

The spectacular activity on the VGSCs indicate that they are the main target of alpha nemertides. A screen of other ion channels should however be considered as other marine peptide toxins (e.g. sea anemone APETx) have overlapping activity profiles in sodium and potassium channels[129].

5.3.6 Structure and sequence activity relationships in alpha nemertides

In **paper II** the transcriptomes of 17 nemertean species were examined for alpha nemertides, revealing seven full length sequences distributed in five species in the Lineidean taxon. The maximum number of alpha nemertides in any of the species were three (*L. ruber,* figure 16).

The low number of alpha nemertides per species can be compared to the complex venoms of cone snails with hundreds of individual toxins per species[103,130]. The snails are generally ambush hunters and use a harpoon loaded with venom in their hunting strategy. The harpoon is connected to a venom gland expressing a plethora of peptides. How the remarkable abundance of expressed peptides in the glands have evolved is still unclear. However, multiple duplication events of conotoxin genes, maintaining the cystine framework, has been suggested as the mechanism for diversification[131,132]. The need for rapid divergence of toxins in cone snails probably also derives from biological factors like a sedentary lifestyle.

The division of natural alpha nemertides into two groups based on the potency in *Artemia* microwell assay was attributed to the variation in sequence. A small amino acid (Ala or Ser) in position 4 in combination with a Phe in

position 8 was important for high potency, while a charged amino acid, Lys, in position four decreased the activity.

In the VGSC experiments (**paper III**) the relationships were unclear. A preference for the insect channel $BgNa_V1$ in most of the toxins was however identified. Alpha-6, which has low potency in *Artemia* was highly active in $BgNa_V1$, indicating the importance of testing substances both *in vivo* and *in vitro*.

A more thorough investigation using point mutations is required to evaluate the specific relevance of individual amino acids and positions in the nemertide alpha sequence. Initial data from an alanine scan (data not shown) of alpha-1 indicate that the amino acids in the aromatic patch (Phe8, Trp22 and Phe24, figure 12) are important for insect channel selectivity.

5.3.7 Alpha nemertides as possible insecticides

The selectivity profiles of the natural alpha nemertides 1-7 in VGSCs (table 1) in combination with the *in vivo Artemia* microwell assay results (figure 15) suggest an application of these toxins as insecticides. Nemertide alpha-1, the toxin which had the most favorable profile with regard to insect VGSC selectivity was at least 8.9 times more active on $BgNa_V1$ than on the vertebrate channels. Also alpha-6 should be considered with its favorable selectivity ratio between $BgNa_V1$ and the vertebrate cardiac $Na_V1.5$. Alpha-6 is however less active than alpha-1 in the *Artemia* microwell assay, probably due to bioavailability issues.

Factors other than the *in vitro* safety profile on VGSCs that should be considered before use as bioinsecticides include: toxicity to beneficial insects (e.g. pollinators)[133], bioaccumulation, degradation and possible to administer to target insects[21]. Furthermore *in vivo* activity studies in vertebrates should be assayed for topical, oral and parenteral administration.

5.3.8 Peptides from *P. corrugatus* exhibit potent membrane disruptive activity

In **paper IV** the mucus of the Antarctic nemertean *P. corrugatus* was investigated using similar techniques as in **paper II**: HPLC-UV, reduction and alkylation, UPLC-QToF, MS/MS and transcriptome analysis. Furthermore, two purified peptides, and a fraction of parborlysins were assayed for their membrane disruptive activity on resuspended *E. coli* model membranes.

All three samples were found to be highly potent in the assay, with spectacular activity for the parborlysin fraction. The parborlysin fraction was found to exhibit an EC_{50} of 0.07μM, lower than the benchmark compound

melittin, and the human antimicrobial peptide LL37[93]. In a previous study by Berne *et al.* on a heterogenous fraction of parborlysins, hemolytic activity was displayed[69]. These results suggest a broad membrane disruptive activity profile, and that the parborlysins may take part in the host-defense system against microorganisms and predators.

The nemertean membrane disruptive peptides may facilitate the penetration of neurotoxins present in the mucus. However, this hypothesis is yet to be explored.

5.3.9 *In silico* functionality prediction of nemertides

In **paper IV** a *in silico* prediction server, CSPred[44], was used to predict functionality of the 8750 Da peptide. The 8750 Da peptide harbors a unique eight cysteine pattern (C-CC-CC-C-C-C), and was predicted to belong to the antimicrobial peptide (AMP) group. Indeed, the peptide proved to be active in the membrane permeabilization assay with an EC_{50} of 1.8μM.

The parborlysins and B-neurotoxins were correctly classified into their respective categories. The alpha nemertides were classified mainly as ion-channel active, but they were also to some extent, predicted to have AMP-activity, prompting the peptides to be tested in antimicrobial- or membrane permeabilization assays.

5.3.10 Repurposing of transcriptomic data

The publicly available transcriptomes used in this thesis work (**paper II and IV**) were originally produced for other purposes, mainly for phylogenetic studies[79,134]. With the rapid growing body of such data available from a more and more diverse sets of organisms, the opportunity to study different aspects (e.g. host-defense or attack related sequences) from strange and interesting organisms increase by the day[112].

One challenge will probably be to select species or families, and to obtain material for sequence and structural confirmation on the peptide level. The selection of species to investigate will largely depend on homology searches for proteins with queries from previously known sequences i.e. BLAST. For short peptides a pattern/motif based approach should be used, at least in combination with the homology based techniques.

Also, literature survey should not be neglected as a tool for selection of species to study: There are several groups of organisms that have been largely neglected from a toxinology viewpoint e.g. venomous mollusks other than cone snails and venomous crustacean species[14].

6 Concluding remarks

The work presented in this thesis spans all parts of contemporary pharmacognosy relating to cysteine rich peptides: from source organism to peptide sequence determination, functional characterization and an initial elucidation of structure activity relationships. Furthermore, big data in the form of transcriptomes was utilized to expand the knowledge and sequence diversity of both cyclotides and nemertean toxins, as well as for sequence determination. Finally, routes for production of natural occurring cyclotides in plant cell suspension culture and *in vitro* production of bioactive alpha nemertides using SPPS were explored.

The main outcomes from my thesis work are:

Identification of cyclotides from *V. uliginosa*, and generation of a suspension cell culture, **paper I.**

- 12 new cyclotide sequences from *V. uliginosa* was discovered using transcriptomic and mass-spectrometric techniques.
- A cell-suspension culture, that can be used for production of native cyclotides and/or in investigations into the biochemistry of cyclotides, was generated

The discovery and characterization of a new family of highly potent toxins, alpha nemertides, from nemertean worms **paper II-III.**

- Alpha nemertides comprise a highly conserved ICK family of toxins consisting of eight members, alpha 1-8, which are restricted to one clade of nemertean worms (Liniedea).
- The archetype alpha nemertide, alpha-1, exhibited spectacular toxicity towards cockroach (*B. dubia*, paralytic dose 2.1 pmol/g injection) and crustacean species (*C. maenas* 0.3 pmol/g injection).
- Alpha nemertides act on VGSCs, with a general selectivity for insect channels. The mechanism of action was proposed to be inhibition of fast inactivation, conveyed by toxins binding to site 3.

In **paper IV** the exploration of nemertean toxins was continued with chemical investigation of the mucus as well as investigation of transcriptomes of the Antarctic nemertean *P. corrugatus*.

- A polypeptide harboring an unique cysteine pattern comprised of eight cysteines was discovered and found to exhibit membrane disruptive activity.
- A heterogenous fraction of parborlysins was exhibit potent membrane disruptive activity (EC_{50} 0.07μM).
- Transcriptomic analysis revealed a set of putative toxins with similarity to all types of nemertides (alpha, beta/B-neurotoxin and parborlysin/A-cytolysins) previously described.

7 Populärvetenskaplig sammanfattning

Peptider och proteiner är viktiga beståndsdelar i alla celler och innehar otaliga funktioner. De utgör till exempel muskler och vävnad och är involverade i processer såsom ämnesomsättning, immunförsvar och syretransport. Förutom att förmedla viktiga funktioner inom en organism kan de även användas till jakt efter föda eller till försvar mot mikroorganismer eller predatorer. Kägelsnäckor, spindlar och ormar är exempel på djur som använder peptider till jakt medan försvarpeptider ofta återfinns både växtriket och i djurriket. Dessa peptider har ofta det gemensamt att de innehåller aminosyran cystein, varför de kallas cysteinrika peptider. Disulfidbryggor mellan cysteinerna i dessa peptider medför struktur och stabilitet som särskiljer dem från andra proteiner/peptider.

Den här avhandlingen handlar om olika aspekter av dessa cysteinrika, stabila peptider: hur man upptäcker dem, hur man kan tillverka dem i laboratoriet samt hur man utvärderar deras effekter i biologiska system.

I **delarbete I** undersöktes sumpviolen (*Viola uliginosa*), för sitt innehåll av en speciell grupp av cysteinrika peptider: cyklotider. Cyklotider är cykliska och har tre disulfidbryggor organiserade i en sorts knut, en s.k. ’cystine knot’. Dessa egenskaper gör cyklotiderna mycket motståndskraftiga mot värme och enzymatisk nedbrytning. Vidare är cyklotider en del av violens försvar mot mikroorganismer, främst genom att förstöra mikroorganismernas cellmembran. Ett cyklotidinnehållande extrakt (Sero-X®) finns på marknaden som ett naturligt bekämpningsmedel, och en klinisk fas I-studie av en syntetisk cyklotidanalog för behandling av multipel skleros, MS, har precis avslutats med lyckat resultat.

Med hjälp av genetiska data (transkriptom) och kemisk analys baserad på mass-spektrometri kunde 12 nya cyklotider identifieras och deras sekvenser kunde utrönas. Utöver de nya cyklotiderna hittades även åtta sekvenser som redan var kända.

I **delarbete I** undersöktes även om det går att använda cellkulturer i suspension, dvs odla enskilda eller små aggregat av violceller i en näringslösning, detta för att kunna tillverka de naturligt förekommande cyklotiderna på biologisk väg utan att behöva skörda vilda bestånd av violer. Experimentet lyckades och en suspensioncellkultur med förmåga att producera cyklotider med ett dubbelt så högt utbyte jämfört med vad man kan erhålla från vilda växter etablerades.

Cellkulturen kan även användas till att studera hur violer fungerar, speciellt med fokus på hur cyklotider tillverkas naturligt: vilka enzym är inblandade? Går det att öka tillverkningen med hjälp av t.ex. hormoner? Kan man genmodifiera cellkulturen för att kunna tillverka andra typer av peptider och proteiner? Dessa frågeställningar, och många fler skulle kunna besvaras genom att studera detta system.

I de följande delarbetena (**II-IV**) flyttades fokus från växter till marina djur, närmare bestämt slemmaskar (*Nemertea* sp.) Det finns circa 1350 olika arter av slemmaskar globalt och de olika arterna kan variera enormt i längd. De minsta lever bland sandkorn, medan världens längsta djur, långsnöret (*Lineus longissimus*) med sitt världsrekord på 55 m, också ingår i släktet.

I **delarbete II** undersöktes just hur långsnöret försvarar sin potentiellt extrema kroppslängd mot predatorer. Tidigare har man trott att masken försvarar sig genom att utsöndra ett gift, tetrodotoxin (TTX, känt från den japanska blåsfisken, fugu) i sitt slem. I initiala försök kunde dock inget TTX påvisas, varför fokus flyttades till att försöka ta reda på vilka andra typer av försvarsmolekyler som kunde tänkas finnas i slemmet och peptider kunde påvisas. Peptiderna isolerades och analyserades på samma sätt som i delarbete I; med hjälp av genetisk data och mass-spektrometri, och uppvisade en struktur liknande den som finns i cyklotider med undantaget att slemmask-peptiden inte är cyklisk. Förutom cystinknuten, som finns i båda typerna av peptider, saknas likheter i sekvens. Då endast mycket små mängder av peptiderna kunde renas fram, tillverkades peptiderna med hjälp av fast-fas peptid-syntes, och undersöktes för funktionalitet.

En peptid, alpha-1, var extremt potent mot både krabbor och kackerlackor, och är ett av de giftigaste peptid-toxinerna som hittats: (1g räcker till att döda ca. 140 ton kackerlackor eller ca 1000 ton krabbor, om peptiden injiceras).

Försök på spänningsberoende natrium-jonkanaler, viktiga för nervsignaler, visade att toxinet uppvisar selektivitet för insektsjonkanaler över däggdjurskanaler, dvs de är inte lika giftiga för däggdjur som för insekter. Detta medför att peptiderna kanske kan användas som insektsgifter.

I **delarbete III**, fortsatte undersökningarna kring alpha nemertiderna som upptäcktes i **delarbete II**. Med hjälp av genetiska data från publika databaser från flera arter av slemmaskar kunde sju toxin-liknande sekvenser upptäckas, syntetiseras och testas funktionellt. Den här gången testades nemertiderna på hela små kräftdjur, salträkor (*Artemia salina*, också kända som sea monkeys, eller instant life). Även i det här försöket var peptiderna väldigt aktiva, men skillnader i aktivitet kunde urskiljas mellan de olika gifterna.

Alpha nemertiderna testades också på ett större urval än tidigare av jonkanaler från däggdjur. I de här testerna visade det sig att några av toxinerna inte skiljer på jonkanaler från ryggradsdjur och insekter, medan alpha-1 som isolerats först (i **delarbete II**) fortfarande uppvisade selektivitet även när man jämför mellan andra jonkanaler. Fortsatta studier behövs dock, och har redan

påbörjats, för att undersöka vilka delar av toxinerna som är viktiga för selektivitet och aktivitet på jonkanalerna.

I **delarbete IV** tog undersökningen av toxiner från slemmaskar en lite annan riktning; slemmaskar från Antarktis (Parborlasia corrugatus) undersöktes med avseende på slemmets beståndsdelar. *P. corrugatus* har tidigare undersökts med avseende på hur proteiner (parborlysiner) i slemmet förstör röda blodkroppar, s k. hemolytisk aktivitet.

I delarbete IV undersöktes slem från *P. corrugatus* med hjälp av massspektrometri och genetiska data. Det framkom att masken inte bara har hemolytiska proteiner i sitt slem, utan även mindre peptider. Totalt kunde fyra stycken peptider urskiljas att ha tre cystinbryggor, och en peptid visade sig ha fyra cystinbryggor. Två peptider renades fram, men endast för en av dem kunde hela aminosyrasekvensen utläsas. Det visade sig att den sekvenserade peptiden har ett nytt och unikt cysteinmönster.

Även proteiner av liknande storlek som de hemolytiska proteinerna som tidigare beskrivits (parborlysiner) från samma art upptäcktes. Det visade sig dock att det var svårt att få fram rena fraktioner av just dessa proteiner.

För att undersöka huruvida de isolerade peptiderna och den mindre rena hypotetiskt hemolytiska fraktionen kan förstöra membran testades detta. Det visade sig att de båda peptiderna hade hög effekt på bakteriemembran, men också att den hemolytiska fraktionen är extremt aktiv på dessa membran.

Sammanfattningsvis kan sägas att i den här avhandlingen har flera cysteinrika peptider och proteiner identifierats och sekvenserats, främst med hjälp av mass-spektrometri i kombination med genetiska data. Vidare så har en grupp av extremt potenta jonkanalstoxiner upptäckts, tillverkats, och testats i flera olika biologiska system inklusive salträka, kackerlacka och krabba. Utöver detta har nya cyklotider upptäckts i sumpviol och en suspensions-cellkultur upprättats som kan användas dels för framställning av naturliga cyklotider samt för vidare studier av violers biologi och cyklotiders bioprocessning.

Acknowledgements

I would like to express my gratitude to the following people and animals:

My main supervisor, Ulf Göransson for giving me the opportunity to do my PhD-work in your group, and for fantastic support during the course of my thesis work. From the first tube of lyophilized mucus to the very end you have been inspiring and helpful

I would also like to thank my co-supervisor Håkan Andersson. Your calming Kalmar-aura, positive pessimism and fun discussions have been very valuable.

Adam, you have always been there to rescue students in need, I miss your omni-present laughter in the laboratory, but not your never-ending fascination for Thai-Break food. Anders and Lars, for introducing me to the exciting world of pharmacognosy. Thank you for letting me be part of the global pharmacy course and for long night discussions after the traditional division Christmas dinners. Paco, for your optimism, and all 'bon appetites' I have received since you arrived at the research group. Christina for support during my time as a teacher, and for the delicious truffles Hesham and Sunithi, you always bring smiles to the lab, thank you for not being Swedish

For my fellow PhD students, Taj: you have been there since the UGSBR days, it has been quite a ride… thank you for your hard work maintaining the synthesizer, nearly always with a smile. Astrid: Thank you for adding a refreshing element to the social life of the corridor, and for being a such a good friend. Karin, where does your energy come from? You seem to absorb it up like a… mushroom. Elisabet, you are almost still present..oknytt? Somehow I miss the backstreet boys. Blazej, Thank you for all discussions on everything large, and mostly small things, good times and for showing me cold Krakow. Momo , for showing me Taiwan, thank you for all the fun and high scores in the Taiwanese rating system.. Luke, you know who got first price at JNPC, right?

I would also like to thank the following former members of our research group: Stefan, Sungkyu, and Sohaib. Thank you. To the exam workers involved in my thesis work; Julia, Carl-Magnus, Abbe, thank

you for your company and efforts and good luck in the future. I would also like to thank the administrative personnel at the department, PhD-related formalia would be even worse.

I would like to express my gratitude to Co-authors and to Apotekarsocieteten which enabled me to present my work abroad and connect with other researchers in the field.

A special thanks to the ravens, crows, opossums, raccoons and skunks on various live cams that have kept me company late nights in front of the computer in my office.

Lina, thank you for all support and great advice since I moved to Uppsala, for letting me be a part of your and Adams family. Sten, for sharing your views on everything, I would not have been who I am without your influence as a big brother. Mor och Far, for all support I've got through life, I would not have managed this without you.

Camilla for always being there for me.

References

1. Rosengren KJ, Daly NL, Plan MR, Waine C, Craik DJ. Twists, knots, and rings in proteins: Structural definition of the cyclotide framework. J Biol Chem. 2003;278(10):8606–16.
2. Conibear AC, Rosengren KJ, Harvey PJ, Craik DJ. Structural characterization of the cyclic cystine ladder motif of θ-defensins. Biochemistry. 2012 Dec 4;51(48):9718–26.
3. Kaas Q, Westermann J-C, Craik DJ. Conopeptide characterization and classifications: an analysis using ConoServer. Toxicon. 2010 Jul;55(8):1491–509.
4. Luckett S, Garcia RS, Barker JJ, Konarev A V., Shewry PR, Clarke AR, et al. High-resolution structure of a potent, cyclic proteinase inhibitor from sunflower seeds. J Mol Biol. 1999 Jul 9;290(2):525–33.
5. Tam JP, Wang S, Wong KH, Tan WL. Antimicrobial Peptides from Plants. Pharmaceuticals (Basel). 2015 Nov 16;8(4):711–57.
6. Pennington MW, Czerwinski A, Norton RS. Peptide therapeutics from venom: Current status and potential. Bioorganic Med Chem. 2018;26(10):2738–58.
7. Walker AA, Dobson J, Jin J, Robinson SD, Herzig V, Vetter I, et al. Buzz Kill: Function and Proteomic Composition of Venom from the Giant Assassin Fly Dolopus genitalis (Diptera: Asilidae). Toxins (Basel). 2018;10(11).
8. Robinson SD, Li Q, Lu A, Bandyopadhyay PK, Yandell M, Olivera BM, et al. The Venom Repertoire of Conus gloriamaris (Chemnitz, 1777), the Glory of the Sea. Mar Drugs. 2017 May 20;15(5):1–20.
9. Sollod BL, Wilson D, Zhaxybayeva O, Gogarten JP, Drinkwater R, King GF. Were arachnids the first to use combinatorial peptide libraries? Peptides. 2005 Jan;26(1):131–9.
10. Hellinger R, Koehbach J, Soltis DE, Carpenter EJ, Wong GK, Gruber CW. Peptidomics of Circular Cysteine-Rich Plant Peptides: Analysis of the Diversity of Cyclotides from Viola tricolor by Transcriptome and Proteome Mining. J Proteome Res. 2015 Nov 6;14(11):4851–62.
11. Gao B, Peng C, Yang J, Yi Y, Zhang J, Shi Q. Cone Snails: A Big Store of Conotoxins for Novel Drug Discovery. Toxins (Basel). 2017;9(12):1–17.
12. Utkin YN. Last decade update for three-finger toxins: Newly emerging structures and biological activities. World J Biol Chem. 2019 Jan 7;10(1):17–27.
13. Herzig V, Wood DL a, Newell F, Chaumeil P-A, Kaas Q, Binford GJ, et al. ArachnoServer 2.0, an updated online resource for spider toxin sequences and structures. Nucleic Acids Res. 2011 Jan;39(Database issue):D653-7.
14. von Reumont B, Campbell L, Jenner R. Quo Vadis Venomics? A Roadmap to Neglected Venomous Invertebrates. Toxins (Basel).

2014;6(12):3488–551.
15. Sigrist CJ a, Cerutti L, Hulo N, Gattiker A, Falquet L, Pagni M, et al. PROSITE: a documented database using patterns and profiles as motif descriptors. Brief Bioinform. 2002 Sep;3(3):265–74.
16. Zhang J, Hua Z, Huang Z, Chen Q, Long Q, Craik DJ, et al. Two Blast-independent tools, CyPerl and CyExcel, for harvesting hundreds of novel cyclotides and analogues from plant genomes and protein databases. Planta. 2015 Apr;241(4):929–40.
17. Gunasekera S, Foley FM, Clark RJ, Sando L, Fabri LJ, Craik DJ, et al. Engineering stabilized vascular endothelial growth factor-A antagonists: synthesis, structural characterization, and bioactivity of grafted analogues of cyclotides. J Med Chem. 2008 Dec 25;51(24):7697–704.
18. Poth AG, Huang Y-H, Le TT, Kan M-W, Craik DJ. Pharmacokinetic characterization of kalata B1 and related therapeutics built on the cyclotide scaffold. Int J Pharm. 2019 Jun 30;565:437–46.
19. Kaas Q, Yu R, Jin A-H, Dutertre S, Craik DJ. ConoServer: updated content, knowledge, and discovery tools in the conopeptide database. Nucleic Acids Res. 2012 Jan;40(Database issue):D325-30.
20. Gran L. Oxytocic principles of Oldenlandia affinis. Vol. 36, Lloydia. 1973. p. 174–8.
21. King GF. Tying pest insects in knots: the deployment of spider-venom-derived knottins as bioinsecticides. Pest Manag Sci. 2019 Apr 25;(February).
22. Oguis GK, Gilding EK, Jackson MA, Craik DJ. Butterfly Pea (Clitoria ternatea), a Cyclotide-Bearing Plant With Applications in Agriculture and Medicine. Front Plant Sci. 2019;10(May):645.
23. Holford M, Daly M, King GF, Norton RS. Venoms to the rescue. Science. 2018 Aug 31;361(6405):842–4.
24. Dardevet L, Rani D, Aziz TA El, Bazin I, Sabatier J-M, Fadl M, et al. Chlorotoxin: a helpful natural scorpion peptide to diagnose glioma and fight tumor invasion. Toxins (Basel). 2015 Mar 27;7(4):1079–101.
25. Schmidtko A, Lötsch J, Freynhagen R, Geisslinger G. Ziconotide for treatment of severe chronic pain. Lancet. 2010 May;375(9725):1569–77.
26. DeBin JA, Maggio JE, Strichartz GR. Purification and characterization of chlorotoxin, a chloride channel ligand from the venom of the scorpion. Am J Physiol. 1993 Feb;264(2 Pt 1):C361-9.
27. Riviere LR, Tempst P. Enzymatic digestion of proteins in solution. Curr Protoc Protein Sci. 2001 May;Chapter 11(1995):Unit 11.1.
28. Crankshaw MW, Grant GA. Modification of cysteine. Curr Protoc Protein Sci. 2001 May;Chapter 15(1):Unit15.1.
29. Steen H, Mann M. The ABC's (and XYZ's) of peptide sequencing. Nat Rev Mol Cell Biol. 2004 Sep;5(9):699–711.
30. Papayannopoulos IA. The interpretation of collision-induced dissociation tandem mass spectra of peptides. Mass Spectrom Rev.

1995 Jan;14(1):49–73.
31. Manzoni C, Kia DA, Vandrovcova J, Hardy J, Wood NW, Lewis PA, et al. Genome, transcriptome and proteome: the rise of omics data and their integration in biomedical sciences. Brief Bioinform. 2018;19(2):286–302.
32. Liu L, Li Y, Li S, Hu N, He Y, Pong R, et al. Comparison of next-generation sequencing systems. J Biomed Biotechnol. 2012;2012:251364.
33. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, et al. Accurate whole human genome sequencing using reversible terminator chemistry. Nature. 2008 Nov 6;456(7218):53–9.
34. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 2011;29(7):644–52.
35. Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. 2000 Jun;16(6):276–7.
36. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990 Oct 5;215(3):403–10.
37. Merrifield RB. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. J Am Chem Soc. 1963 Jul;85(14):2149–54.
38. Pedersen SL, Tofteng AP, Malik L, Jensen KJ. Microwave heating in solid-phase peptide synthesis. Chem Soc Rev. 2012 Mar 7;41(5):1826–44.
39. Carpino LA, Han GY. The 9-Fluorenylmethoxycarbonyl Function, a New Base-Sensitive Amino-Protecting Group. J Am Chem Soc. 1970;92(19):5748–9.
40. Yang F, Liu S, Zhang Y, Qin C, Xu L, Li W, et al. Expression of recombinant α-toxin BmKM9 from scorpion Buthus martensii Karsch and its functional characterization on sodium channels. Peptides. 2018;99(September 2017):153–60.
41. Wilkinson B, Gilbert HF. Protein disulfide isomerase. Biochim Biophys Acta. 2004 Jun 1;1699(1–2):35–44.
42. Wang Z-Q, Han Y-H, Shao X-X, Chi C-W, Guo Z-Y. Molecular cloning, expression and characterization of protein disulfide isomerase from Conus marmoreus. FEBS J. 2007 Sep 1;274(18):4778–87.
43. Wu X, Wu Y, Zhu F, Yang Q, Wu Q, Zhangsun D, et al. Optimal cleavage and oxidative folding of α-conotoxin TxIB as a therapeutic candidate peptide. Mar Drugs. 2013 Sep 17;11(9):3537–53.
44. Islam SMA, Kearney CM, Baker EJ. Assigning biological function using hidden signatures in cystine-stabilized peptide sequences. Sci Rep. 2018 Jun 13;8(1):9049.
45. Liang L, Ma C, Du T, Zhao Y, Zhao X, Liu M, et al. Bioactivity-explorer: a web application for interactive visualization and exploration of bioactivity data. J Cheminform. 2019 Jul 10;11(1):47.
46. Camarero JA, Campbell MJ. The Potential of the Cyclotide Scaffold for Drug Development. Biomedicines. 2019 Apr 19;7(2):31.

47. Burman R, Gunasekera S, Strömstedt AA, Göransson U. Chemistry and biology of cyclotides: Circular plant peptides outside the box. J Nat Prod. 2014;77(3):724–36.
48. Craik DJ, Malik U. Cyclotide biosynthesis. Curr Opin Chem Biol. 2013 Aug;17(4):546–54.
49. Gerlach SL, Göransson U, Kaas Q, Craik DJ, Mondal D, Gruber CW. A systematic approach to document cyclotide distribution in plant species from genomic, transcriptomic, and peptidomic analysis. Biopolymers. 2013 Sep;100(5):433–7.
50. Mulvenna JP, Mylne JS, Bharathi R, Burton RA, Shirley NJ, Fincher GB, et al. Discovery of cyclotide-like protein sequences in graminaceous crop plants: ancestral precursors of circular proteins? Plant Cell. 2006 Sep;18(9):2134–44.
51. Colgrave ML, Craik DJ. Thermal, chemical, and enzymatic stability of the cyclotide kalata B1: the importance of the cyclic cystine knot. Biochemistry. 2004 May 25;43(20):5965–75.
52. Craik DJ, Du J. Cyclotides as drug design scaffolds. Curr Opin Chem Biol. 2017 Jun;38:8–16.
53. Burman R, Strömstedt A a, Malmsten M, Göransson U. Cyclotide-membrane interactions: Defining factors of membrane binding, depletion and disruption. Biochim Biophys Acta. 2011 Nov;1808(11):2665–73.
54. Tam JP, Lu YA, Yang JL, Chiu KW. An unusual structural motif of antimicrobial peptides containing end-to-end macrocycle and cystine-knot disulfides. Proc Natl Acad Sci U S A. 1999;96(16):8913–8.
55. Jennings C, West J, Waine C, Craik D, Anderson M. Biosynthesis and insecticidal properties of plant cyclotides: the cyclic knotted proteins from Oldenlandia affinis. Proc Natl Acad Sci U S A. 2001;98(19):10614–9.
56. Burman R, Yeshak MY, Larsson S, Craik DJ, Rosengren KJ, Göransson U. Distribution of circular proteins in plants: large-scale mapping of cyclotides in the Violaceae. Front Plant Sci. 2015;6(October):855.
57. Gruber CW, Elliott AG, Ireland DC, Delprete PG, Dessein S, Göransson U, et al. Distribution and evolution of circular miniproteins in flowering plants. Plant Cell. 2008;20(9):2471–83.
58. Craik DJ, Daly NL, Bond T, Waine C. Plant cyclotides: A unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. J Mol Biol. 1999;294:1327–36.
59. Weidmann J, Craik DJ. Discovery, structure, function, and applications of cyclotides: circular proteins from plants. J Exp Bot. 2016;67(16):4801–12.
60. Leta Aboye T, Clark RJ, Craik DJ, Göransson U. Ultra-stable peptide scaffolds for protein engineering-synthesis and folding of the circular cystine knotted cyclotide cycloviolacin O2. Chembiochem. 2008 Jan 4;9(1):103–13.
61. Slazak B, Sliwinska E, Saługa M, Ronikier M, Bujak J, Słomka A, et

al. Micropropagation of Viola uliginosa (Violaceae) for endangered species conservation and for somaclonal variation-enhanced cyclotide biosynthesis. Plant Cell, Tissue Organ Cult. 2015 Jan 23;120(1):179–90.
62. Nguyen GKT, Wang S, Qiu Y, Hemu X, Lian Y, Tam JP. Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. Nat Chem Biol. 2014 Jul 20;10(July).
63. Murashige T, Skoog F. A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. Physiol Plant. 1962 Jul;15(3):473–97.
64. Horton T, Kroh A, Ahyong S, Bailly N, Boyko CB, Brandão SN, et al. World Register of Marine Species (WoRMS). WoRMS Editorial Board; 2019.
65. Thiel M, Kruse I. Status of the Nemertea as predators in marine ecosystems. Hydrobiologia. 2001;456:21–32.
66. Strand M, Norenburg J, Alfaya JE, Ángel Fernández-Álvarez F, Andersson HS, Andrade SCS, et al. Nemertean taxonomy-Implementing changes in the higher ranks, dismissing Anopla and Enopla. Zool Scr. 2019 Jan;48(1):118–9.
67. McDermott J. Status of the Nemertea as prey in marine ecosystems. Hydrobiologia. 2001;456:7–20.
68. Kem WR. Purification and characterization of a new family of polypeptide neurotoxins from the heteronemertine Cerebratulus lacteus (Leidy). J Biol Chem. 1976 Jul 25;251(14):4184–92.
69. Berne S, Sepčić K, Križaj I, Kem WR, McClintock JB, Turk T. Isolation and characterisation of a cytolytic protein from mucus secretions of the Antarctic heteronemertine Parborlasia corrugatus. Toxicon. 2003 Mar;41(4):483–91.
70. Kem WR, Blumenthal KM. Purification and characterization of the cytotoxic Cerebratulus A toxins. J Biol Chem. 1978 Aug 25;253(16):5752–7.
71. Miyazawa K, Higashiyama M, Ito K, Noguchi T, Arakawa O, Shida Y, et al. Tetrodotoxin in two species of ribbon worm (Nemertini), Lineus fuscoviridis and Tubulanus punctatus. Toxicon. 1988 Jan;26(9):867–74.
72. Coates RM, Kem WR, Abbott BC. Isolation and structure of a hoplonemertine toxin. Toxicon. 1971 Jan;9(1):15–22.
73. Hunter BE, de Fiebre CM, Papke RL, Kem WR, Meyer EM. A novel nicotinic agonist facilitates induction of long-term potentiation in the rat hippocampus. Neurosci Lett. 1994 Feb 28;168(1–2):130–4.
74. Nelsen DR, Nisani Z, Cooper AM, Fox GA, Gren ECK, Corbit AG, et al. Poisons, toxungens, and venoms: Redefining and classifying toxic biological secretions and the organisms that employ them. Biol Rev. 2014;89(2):450–65.
75. Yang H-L, Shen Z-Q, Liu X, Kong Y. Two novel antimicrobial peptides from skin venoms of spadefoot toad Megophrys minor. Chin J Nat Med. 2016 Apr;14(4):294–8.

76. Carroll S, McEvoy EG, Gibson R. The production of tetrodotoxin-like substances by nemertean worms in conjunction with bacteria. J Exp Mar Bio Ecol. 2003 Mar;288(1):51–63.
77. Strand M, Hedstr??m M, Seth H, McEvoy EG, Jacobsson E, G??ransson U, et al. The bacterial (Vibrio alginolyticus) production of tetrodotoxin in the ribbon worm Lineus longissimus - Just a false positive? Mar Drugs. 2016;14(4):1–11.
78. Mailho-Fontana PL, Antoniazzi MM, Sciani JM, Pimenta DC, Barbaro KC, Jared C. Morphological and biochemical characterization of the cutaneous poison glands in toads (Rhinella marina group) from different environments. Front Zool. 2018;15(1):46.
79. Romiguier J, Gayral P, Ballenghien M, Bernard A, Cahais V, Chenuil A, et al. Comparative population genomics in animals uncovers the determinants of genetic diversity. Nature. 2014 Nov 13;515(7526):261–3.
80. Whelan N V, Kocot KM, Santos SR, Halanych KM. Nemertean Toxin Genes Revealed through Transcriptome Sequencing. Genome Biol Evol. 2014 Jan;6(12):3314–25.
81. Skinner WS, Dennis PA, Li JP, Quistad GB. Identification of insecticidal peptides from venom of the trap-door spider, Aptostichus schlingeri (Ctenizidae). Toxicon. 1992;30(9):1043–50.
82. Figueiredo SG, Garcia ME, Valentim AC, Cordeiro MN, Diniz CR, Richardson M. Purification and amino acid sequence of the insecticidal neurotoxin Tx4(6-1) from the venom of the “armed” spider Phoneutria nigriventer (Keys). Toxicon. 1995 Jan;33(1):83–93.
83. Windley MJ, Herzig V, Dziemborowicz SA, Hardy MC, King GF, Nicholson GM. Spider-venom peptides as bioinsecticides. Toxins (Basel). 2012;4(3):191–227.
84. Deuis JR, Mueller A, Israel MR, Vetter I. The pharmacology of voltage-gated sodium channel activators. Neuropharmacology. 2017;127:87–108.
85. Tam JP, Wu CR, Liu W, Zhang JW. Disulfide Bond Formation in Peptides by Dimethyl-Sulfoxide - Scope and Applications. J Am Chem Soc. 1991;113(17):6657–62.
86. Solis PN, Wright CW, Anderson MM, Gupta MP, Phillipson JD. A Microwell Cytotoxicity Assay using Artemia salina (Brine Shrimp). Planta Med. 1993;59(July):250–2.
87. Gibson R. Antarctic nemerteans: The anatomy, distribution and biology of Parborlasia corrugatus (McIntosh, 1876) (heteronemertea, lineidae). In 1983. p. 289–316.
88. Heine JN, McClintock JB, Slattery M, Weston J. Energetic composition, biomass, and chemical defense in the common antarctic nemertean Parborlasia corrugatus McIntosh. J Exp Mar Bio Ecol. 1991;153(1):15–25.
89. Butala M, Šega D, Tomc B, Podlesek Z, Kem WR, Küpper FC, et al.

Recombinant expression and predicted structure of parborlysin, a cytolytic protein from the Antarctic heteronemertine Parborlasia corrugatus. Toxicon. 2015;
90. Kocot KM. A combined approach toward resolving the phylogeny of Mollusca. Auburn University; 2013.
91. Halanych KM, Kocot KM. Repurposed transcriptomic data facilitate discovery of innate immunity toll-like receptor (TLR) genes across lophotrochozoa. Biol Bull. 2014;227(2):201–9.
92. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Cech M, et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res. 2018 Jul 2;46(W1):W537–44.
93. Strömstedt AA, Kristiansen PE, Gunasekera S, Grob N, Skjeldal L, Göransson U. Selective membrane disruption by the cyclotide kalata B7: Complex ions and essential functional groups in the phosphatidylethanolamine binding pocket. Biochim Biophys Acta - Biomembr. 2016;1858(6):1317–27.
94. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009 Dec 15;10:421.
95. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009 Jul;37(Web Server issue):W202-8.
96. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol. 2019;37(4):420–3.
97. Larsson S, Backlund A, Bohlin L. Reappraising a decade old explanatory model for pharmacognosy. Phytochem Lett. 2008 Nov;1(3):131–4.
98. Camarero JA. Cyclotides, a versatile ultrastable micro-protein scaffold for biotechnological applications. Bioorg Med Chem Lett. 2017;27(23):5089–99.
99. Troeira Henriques S, Craik DJ. Cyclotide Structure and Function: The Role of Membrane Binding and Permeation. Biochemistry. 2017 Feb 7;56(5):669–82.
100. Zhang R-Y, Thapa P, Espiritu MJ, Menon V, Bingham J-P. From nature to creation: Going around in circles, the art of peptide cyclization. Bioorg Med Chem. 2018;26(6):1135–50.
101. Qin Q, McCallum EJ, Kaas Q, Suda J, Saska I, Craik DJ, et al. Identification of candidates for cyclotide biosynthesis and cyclisation by expressed sequence tag analysis of Oldenlandia affinis. BMC Genomics. 2010 Feb 16;11(1):111.
102. Burman R, Gruber CW, Rizzardi K, Herrmann A, Craik DJ, Gupta MP, et al. Cyclotide proteins and precursors from the genus Gloeospermum: Filling a blank spot in the cyclotide map of Violaceae. Phytochemistry. 2010;71(1):13–20.

103. Peng C, Yao G, Gao B-M, Fan C-X, Bian C, Wang J, et al. High-throughput identification of novel conotoxins from the Chinese tubular cone snail (Conus betulinus) by multi-transcriptome sequencing. Gigascience. 2016;5(1):17.
104. Nguyen GKT, Zhang S, Nguyen NTK, Nguyen PQT, Chiu MS, Hardjojo A, et al. Discovery and characterization of novel cyclotides originated from chimeric precursors consisting of albumin-1 chain a and cyclotide domains in the Fabaceae family. J Biol Chem. 2011 Jul 8;286(27):24275–87.
105. Mylne JS, Colgrave ML, Daly NL, Chanson AH, Elliott AG, McCallum EJ, et al. Albumins and their processing machinery are hijacked for cyclic peptides in sunflower. Nat Chem Biol. 2011 May;7(5):257–9.
106. Khan SA, Zafar Y, Briddon RW, Malik KA, Mukhtar Z. Spider venom toxin protects plants from insect attack. Transgenic Res. 2006 Jun;15(3):349–57.
107. Donini M, Marusic C. Current state-of-the-art in plant-based antibody production systems. Biotechnol Lett. 2019 Mar;41(3):335–46.
108. Tekoah Y, Shulman A, Kizhner T, Ruderfer I, Fux L, Nataf Y, et al. Large-scale production of pharmaceutical proteins in plant cell culture-the Protalix experience. Plant Biotechnol J. 2015 Oct;13(8):1199–208.
109. Lopez-Vera E, Walewska A, Skalicky JJ, Olivera BM, Bulaj G. Role of hydroxyprolines in the in vitro oxidative folding and biological activity of conotoxins. Biochemistry. 2008 Feb 12;47(6):1741–51.
110. Blumenthal KM, Keim PS, Heinrikson RL, Kem WR. Structure and action of heteronemertine polypeptide toxins. Amino acid sequence of Cerebratulus lacteus toxin B-II and revised structure of toxin B-IV. J Biol Chem. 1981;256(17):9063–7.
111. Howell ML, Blumenthal KM. Mutagenesis of Cerebratulus lacteus neurotoxin B-IV identifies NH2-terminal sequences important for biological activity. J Biol Chem. 1991 Jul 15;266(20):12884–8.
112. NCBI Resource Coordinators. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 2018;46(D1):D8–13.
113. Swaminathan J, Boulgakov AA, Hernandez ET, Bardo AM, Bachman JL, Marotta J, et al. Highly parallel single-molecule identification of proteins in zeptomole-scale mixtures. Nat Biotechnol. 2018 Oct 22;36(11):1076–91.
114. Luo YJ, Kanda M, Koyanagi R, Hisata K, Akiyama T, Sakamoto H, et al. Nemertean and phoronid genomes reveal lophotrochozoan evolution and the origin of bilaterian heads. Nat Ecol Evol. 2018;2(1):141–51.
115. Göransson U, Jacobsson E, Strand M, Andersson HS. The Toxins of Nemertean Worms. Toxins (Basel). 2019 Feb 15;11(2).
116. Raaymakers C, Verbrugghe E, Hernot S, Hellebuyck T, Betti C, Peleman C, et al. Antimicrobial peptides in frog poisons constitute a

molecular toxin delivery system against predators. Nat Commun. 2017;8(1):1495.
117. Kaplan DI, Isom LL, Petrou S. Role of Sodium Channels in Epilepsy. Cold Spring Harb Perspect Med. 2016;6(6):1–18.
118. Bennett DLH, Woods CG. Painful and painless channelopathies. Lancet Neurol. 2014 Jun;13(6):587–99.
119. Chockalingam P, Wilde A. The multifaceted cardiac sodium channel and its clinical implications. Heart. 2012 Sep;98(17):1318–24.
120. Beneski DA, Catterall WA. Covalent labeling of protein components of the sodium channel with a photoactivable derivative of scorpion toxin. Proc Natl Acad Sci U S A. 1980 Jan;77(1):639–43.
121. Billen B, Bosmans F, Tytgat J. Animal peptides targeting voltage-activated sodium channels. Curr Pharm Des. 2008;14(24):2492–502.
122. de Lera Ruiz M, Kraus RL. Voltage-Gated Sodium Channels: Structure, Function, Pharmacology, and Clinical Indications. J Med Chem. 2015 Sep 24;58(18):7093–118.
123. Mattei C, Legros C. The voltage-gated sodium channel: a major target of marine neurotoxins. Toxicon. 2014 Dec;91:84–95.
124. Stevens M, Peigneur S, Tytgat J. Neurotoxins and their binding areas on voltage-gated sodium channels. Front Pharmacol. 2011;2(November):71.
125. Israel MR, Tay B, Deuis JR, Vetter I. Sodium Channels and Venom Peptide Pharmacology. In: Advances in Pharmacology. 1st ed. Elsevier Inc.; 2017. p. 67–116.
126. Gilchrist J, Olivera BM, Bosmans F. Animal Toxins Influence Voltage-Gated Sodium Channel Function. In: Current Opinion in Pharmacology. 2014. p. 203–29.
127. Toth GP, Blumenthal KM. Structure and action of heteronemertine polypeptide toxins Binding of cerebratulus lacteus toxin B-IV to axon membrane vesicles. Biochim Biophys Acta - Biomembr. 1983 Jul;732(1):160–9.
128. Lieberman DL, Blumenthal KM. Structure and action of heteronemertine polypeptide toxins. Specific cross-linking of Cerebratulus lacteus toxin B-IV to lobster axon membrane vesicles. Biochim Biophys Acta. 1986 Feb 13;855(1):41–8.
129. Peigneur S, Béress L, Möller C, Marí F, Forssmann W-G, Tytgat J. A natural point mutation changes both target selectivity and mechanism of action of sea anemone toxins. FASEB J. 2012 Dec;26(12):5141–51.
130. Safavi-Hemami H, Siero W a, Gorasia DG, Young ND, Macmillan D, Williamson N a, et al. Specialisation of the venom gland proteome in predatory cone snails reveals functional diversification of the conotoxin biosynthetic pathway. J Proteome Res. 2011 Sep 2;10(9):3904–19.
131. Mebs D. Toxicity in animals. Trends in evolution? Toxicon. 2001 Jan;39(1):87–96.
132. Duda TF, Remigio EA. Variation and evolution of toxin gene

expression patterns of six closely related venomous marine snails. Mol Ecol. 2008 Jun;17(12):3018–32.
133. Nakasu EYT, Williamson SM, Edwards MG, Fitches EC, Gatehouse JA, Wright GA, et al. Novel biopesticide based on a spider venom peptide shows no adverse effects on honeybees. Proceedings Biol Sci. 2014 Jul 22;281(1787).
134. Andrade SCS, Montenegro H, Strand M, Schwartz ML, Kajihara H, Norenburg JL, et al. A transcriptomic approach to ribbon worm systematics (nemertea): resolving the pilidiophora problem. Mol Biol Evol. 2014 Dec 27;31(12):3206–15.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Pharmacy 277

Editor: The Dean of the Faculty of Pharmacy

A doctoral dissertation from the Faculty of Pharmacy, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Pharmacy. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Pharmacy".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-390885

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2019