BIOPHYSICAL CHARACTERIZATION AND MASS SPECTRAL IDENTIFICATION OF CATIONIC ANTIMICROBIAL PEPTIDES

by

Melanie Lynn Juba

A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Chemistry and Biochemistry

Committee:

_____	Dr. Barney Bishop, Dissertation Director
_____	Dr. Monique van Hoek, Committee Member
_____	Dr. Robert Honeychuck, Committee Member
_____	Dr. Paul Cooper, Committee Member
_____	Dr. John Schreifels, Department Chair
_____	Dr. Donna M. Fox, Associate Dean, Office of Student Affairs & Special Programs, College of Science
_____	Dr. Peggy Agouris, Dean, College of Science
Date: _____	Spring Semester 2014 George Mason University Fairfax, VA

Biophysical Characterization and Mass Spectral Identification of Cationic Antimicrobial Peptides

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

Melanie Lynn Juba
Bachelors of Science
George Mason University, 2010

Director: Barney Bishop, Professor
Department of Chemistry and Biochemistry

Spring Semester 2014
George Mason University
Fairfax, VA

DEDICATION

This is dedicated to my amazing parents Robert and Suzanne Juba, who supported me in every way throughout this entire process.

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my advisor, Dr. Barney Bishop, for his invaluable help and guidance throughout my entire graduate career. His constant encouragement and ideas have inspired me and made me a better scientist. I would also like to thank my Ph.D. committee members, Dr. Monique van Hoek, Dr. Robert Honeychuck and Dr. Paul Cooper for their helpful suggestions along the process and their willingness to participate in my research. I thank Dr. Paul Russo for his advice and mentorship during my final years of graduate school. I am thankful for the members of the Bishop Lab and van Hoek Lab that have helped me along the way and provided a creative and collaborative atmosphere. I would like to personally thank and give credit to Dr. Susan Gillmor for performing the confocal microscopy experiments, Scott Dean for performing antimicrobial assays, Stephanie Barksdale for carrying out plasma challenging, harvests and performance assays, Megan Devine for performing particle synthesis, harvest and elution and Carlos Rodriguez for performing antimicrobial assays. I would like to thank Elissa Williams for her help and use of the FESEM. I gratefully acknowledge the Defense Threat Reduction Agency (DTRA) HDTRA1-12-C-0039 for their support. Finally, I would like to thank my parents, my fiancé and my friends for their love and support throughout my graduate program.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

LIST OF ABBREVIATIONS AND SYMBOLS

2, 2, 2-Trifluoroacetic Acid ..... TFA
2, 2, 2-Trifluoroethanol ..... TFE
Cardiolipin CL
Cationic Antimicrobial Peptide ..... CAMP
Circular Dichroism ..... CD
Collision Induced Dissociation ..... CID
Discriminant Analysis ..... DA
Dynamic Light Scattering ..... DLS
Electron Transfer Dissociation ..... ETD
Electrospray Ionization ..... ESI
Formic Acid ..... FA
Giant Unilamellar Vesicle ..... GUV
Liquid Chromatography LC
Mass Spectrometry ..... MS
Mean Residue Ellipticity ..... θ
Multilamellar Vesicle ..... MLV
Naja atra Cathelicidin ..... NA-CATH
Peptide-to-Lipid Ratio ..... P/L
Phosphatidylcholine ..... PC
Phosphatidylethanol ..... PE
Phosphatidylglycerol ..... PG
Random Forest ..... RF
Scanning Electron Microscopy ..... SEM
Small Unilamellar Vesicle ..... SUV
Support Vector Machine ..... SVM
Tandem Mass Spectrometry ..... MS/MS

ABSTRACT

BIOPHYSICAL CHARACTERIZATION AND MASS SPECTRAL IDENTIFICATION OF CATIONIC ANTIMICROBIAL PEPTIDES

Melanie Lynn Juba, Ph.D.

George Mason University, 2014

Dissertation Director: Dr. Barney Bishop

Cationic antimicrobial peptides (CAMPs) are a highly sequence and structurally diverse group of peptides that exert antibacterial, antifungal and antiviral effects. They have recently been of interest due to their therapeutic potential, however the current CAMP library is limited and those that are known have mechanisms that are not fully understood. One of the main objectives of this dissertation was to investigate CAMP-membrane interactions and how it relates to their antimicrobial potency. Chapter 1 examines similarities and differences of short CAMP isomers (L-ATRA-1A and D-ATRA-1A) in their antimicrobial effectiveness and interactions with model membranes. This chapter establishes the ability of these CAMPs to exert differing antimicrobial potencies against varied bacterial strains, as well as, their ability to adopt helical structure and directly interact with model membranes. The results observed here give insights into the ability of D-isomers to be used as viable therapeutic candidates in place of their L-

counterparts. In chapter 2, the mechanism by which these truncated CAMPs and their full-length parent peptide (NA-CATH) interact with the bacterial membrane is investigated. Here the ability of these CAMPs to depolarize and disrupt the membranes of different bacteria gives insight into differences in the mechanisms each peptide employs. The results from these studies provides a starting point for developing an understanding how truncating full-length CAMPs effects the antimicrobial mechanism.

The second objective of this dissertation was to identify novel antimicrobial peptides to improve the current library of known CAMPs. Chapter 3 describes the development of a new, sample agnostic process for the identification of novel CAMPs, which was applied in the analysis of alligator plasma resulting in the identification of novel peptides that exhibit antimicrobial activity. The process established here has the potential to dramatically impact the way CAMPs and other peptides of interest, such as biomarkers, are discovered in the future. Finally, in chapter 4 the use of LC-MS/MS for *de novo* sequencing of the novel alligator peptides is further investigated. Initially, known CAMPs exhibiting varying physico-chemical properties were chosen, and mass spectrometry parameters were adjusted in order to yield successful *de novo* sequences for these peptides. Once these parameters had been established, *de novo* sequencing of alligator peptides was performed. The results from this study demonstrate the power of LC-MS/MS for *de novo* sequencing peptides from highly complex samples. Overall, this dissertation provides an improved understanding of the way CAMPs exert their antimicrobial effectiveness, as well as, an improved method for the identification and sequencing of novel peptides, including CAMPs.

CHAPTER ONE. CHARACTERIZATION AND PERFORMANCE OF SHORT CATIONIC ANTIMICROBIAL PEPTIDE ISOMERS

Cationic antimicrobial peptides (CAMPs) represent an ancient defense mechanism against invading bacteria, with peptides such as the cathelicidins being essential elements of vertebrate innate immunity. CAMPs are typically associated with broad-spectrum antimicrobial potency and limited bacterial resistance. The cathelicidin identified from the elapid snake *Naja atra* (NA-CATH) contains a semi-conserved repeated 11-residue motif (ATRA motif) with a sequence pattern consistent with formation of an amphipathic helical conformation. Short peptide amides (ATRA-1, -1A, -1P and -2) generated based on the pair of ATRA motifs in NA-CATH exhibited varied antimicrobial potencies. The small size of the ATRA peptides, coupled with their varied antimicrobial performances, make them interesting models to study the impact various physico-chemical properties have on antimicrobial performance in helical CAMPs. Accordingly, the D- and L-enantiomers of the peptide ATRA-1A, which in earlier studies had shown both good antimicrobial performance and strong helical character, were investigated in order to assess the impact peptide stereochemistry has on antimicrobial performance and interaction with chiral membranes. The ATRA-1A isomers exhibit varied potencies against four bacterial strains, and their conformational properties in the presence of mixed zwitterionic/anionic liposomes are influenced by anionic lipid content.

These studies reveal subtle differences in the properties of the peptide isomers. Differences are also seen in the abilities of the ATRA-1A isomers to induce liposome fusion/aggregation, bilayer rearrangement and lysing through turbidity studies and fluorescence microscopy. The similarities and differences in the properties of the ATRA-1A isomers could aid in efforts to develop D-peptide-based therapeutics using high-performing L-peptides as templates.

Introduction

Encoded in the genes of higher organisms, cationic antimicrobial peptides (CAMPs) are potent elements of innate immunity and represent an ancient defensive strategy against infection. The emergence of antibiotic resistance in bacteria has resulted in interest in CAMP-based therapeutics due to their broad-spectrum antimicrobial effectiveness and low incidence of bacterial resistance. These peptides exert a direct antimicrobial effect on bacteria, via a non-receptor mechanism that is believed to involve direct interaction with anionic bacterial membranes. Cellular membranes present complex and dynamic targets, and details regarding CAMP-membrane interactions and their contributions to antimicrobial potency and selectivity remain unclear.

Cathelicidins are a sequence diverse family of antimicrobial peptides that have been found in a wide range of vertebrates. Originally thought to exist exclusively in mammals, cathelicidins have since been discovered in reptiles, birds, and fish (*1–4*). These peptides are identified primarily based on the highly conserved cathelin domain, present in the C-terminal portion of the inactive cathelicidin precursor protein. The cathelin domain is proteolytically liberated as part of the cathelicidin activation and

secretion process (*5*, *6*). The active peptides are grouped into subfamilies based on shared structural features, including linear α-helical, β-hairpin, and Pro-rich / Trp-rich (*3*, *4*). The majority of cathelicidins are classified as α-helical, referring to their tendency to adopt amphipathic helical structures upon interaction with anionic bacterial membranes, which generally is essential to their antimicrobial activity (*3*, *7*).

Recently, peptide sequences from cDNA obtained from the venom glands of the elapid snakes *Naja atra*, *Bungarus fasciatius* and *Ophiophagus hannah* revealed the presence of helical cathelicidins (*5*, *6*). The *N. atra* cathelicidin (NA-CATH) contains within its sequence an imperfect repeated 11-residue sequence pattern that differ at the third and tenth positions (ATRA motif), and a series of short peptide amides (ATRA-1, ATRA-1A, ATRA-1P and ATRA-2) designed based on this motif exhibited varying degrees of antimicrobial effectiveness (*8*, *9*). The sequences of the ATRA-1 and ATRA-2 peptides are based on the two ATRA motifs present in NA-CATH, while ATRA-1A and -1P are hybrid peptides that combine elements found in both of the NA-CATH ATRA motifs (Table 1).

Table 1: Peptide Sequences with Associated Net Charge.
In the NA-CATH sequence the ATRA motifs are shaded in grey, with the residues that differ between the motifs and the ATRA peptides bolded and underlined.

Peptide	Sequence	Net Charge
NA-CATH	KR**F**KKFFKK**L**KNSVKKR**A**KKFFKK**P**KVIGVTFPF	+15
ATRA-1	KR**F**KKFFKK**L**K-NH_2	+8
ATRA-2	KR**A**KKFFKK**P**K-NH_2	+8
ATRA-1A	KR**A**KKFFKK**L**K-NH_2	+8
ATRA-1P	KR**F**KKFFKK**P**K-NH_2	+8

The peptides ATRA-1 and ATRA-1A showed good antimicrobial effectiveness, while ATRA-2 and ATRA-1P proved ineffective (*8*, *9*). The sequences of ATRA-1 and -1A are consistent with formation of amphipathic helical conformations, and both peptides exhibited high degrees of helical structure in the presence of SDS, with ATRA-1A having a higher degree of structure making it a good peptide for structure/function analysis (*8*, *9*). Short peptides, such as ATRA-1A are susceptible to proteolytic degradation, which could negatively impact their therapeutic potential. However, the D-enantiomers of these peptides are generally resistant to proteases, and this increased resistance to degradation could enhance the therapeutic utility of the D-isomers (*10*).

Historically, CAMP enantiomers have exhibited similar antimicrobial properties, but there are notable exceptions where the peptide enantiomers present very different potencies. The antimicrobial potency differences that have been observed between L- and D-peptide isomers have been largely attributed to the increased resistance to proteolytic degradation that is associated with D-peptides. In 1990, Wade et al. reported on the

synthesis and characterization of the L- and D-isomers of three naturally occurring CAMPs (cecropin A, magainin 2 amide and melittin). The results of these studies show no significant differences in the antimicrobial properties between the enantiomers included in the study (*10*). However, in a subsequent study by Vunnam et al. significant differences are seen in the antimicrobial performance between the D- and L-isomers of cecropin A/melittin chimeric peptides (*11*). Then in 2008, Lee and Lee reported that D- and L-enantiomers of a helical CAMP exhibits different antimicrobial potencies against Gram-positive and Gram-negative bacteria. However, they observe that the D-isomer is less effective than the L-peptide (*12*). The results of the above mentioned studies suggest that the factors that influence the performances of the D- and L-isomers of CAMPs are much more complex than previously thought, with some enantiomeric pairs demonstrating divergent antimicrobial properties and others showing no significant differences.

Bacterial membranes are comprised of various formulations of lipids, such as phosphatidylethanolamine (PE), phosphatidylglycerol (PG) and cardiolipin (CL); the inherent chirality that is associated with these lipids has often been neglected. However, there are notable cases where the chiral properties of membrane lipids influence their interactions with other molecules (*13, 14*). Liposomes comprised of the zwitterionic lipid phosphatidylcholine (PC) have been shown to weakly induce chirality in the conformation of an achiral probe (1,6-diphenyl-1,3,5-hexatriene), a change that is detectable by circular dichroism (*14*). Additionally, the preferential interaction between lipid membranes and the LD-isomer of the endorphin-like dipeptide kyotorphin (D-KTP)

may contribute to its superior potency relative to other isomers of kyotrophin. The authors suggest that the D-KTP in this study may be able to more effectively target and concentrate in more rigid regions of the membrane, which have been shown to keep opioid receptors in an active form (*13*). These examples suggest that the inherent chirality of membranes and membrane components can manifest in their interactions with other molecules.

Similar differences may arise in the interactions between helical CAMP stereoisomers and chiral elements in the lipid bilayer, and these differences may contribute to differences in their performances against bacteria. In order to ascertain the potential significance of peptide stereochemistry and their interactions with membranes, the behavior of the L- and D-enantiomers of the 11-residue helical CAMP ATRA-1A in the presence of membranes of varied composition have been investigated. The antimicrobial effectiveness of the ATRA-1A isomers has been assessed against two Gram-negative and two Gram-positive bacteria in order to ascertain any differences in their performances. In order to more directly determine whether CAMP stereochemistry impacts their interaction with membranes, the structural properties of the ATRA-1A isomers and their ability to disrupt and induce fusion of membranes are studied using liposomes of varied ratios of zwitterionic and anionic lipids. These studies suggest significant differences exist in the antimicrobial performances of the peptide isomers and in the way that they interact with membranes.

Materials and Methods

The L- and D-peptides used in these studies were custom synthesized by AAPPTEC, LLC (Louisville, KY) and Genscript USA, Inc. (Piscataway, NJ). The suppliers reported the purities of L-ATRA-1A and D-ATRA-1A to be 95.2% and 95.4%, respectively, based on chromatographic analysis of the purified peptides. The bacterial strains of *Escherichia coli* (*E.*) *coli* (ATCC# 25922), *Bacillus* (*B.*) *cereus* (ATCC# 11778), *Pseudomonas* (*P.*) *aerugniosa* (ATTCC# 19429) and *Staphylococcus* (*S.*) *aureus* (ATCC# 25923) used in these studies were purchased from the American Type Culture Collection (Manassas, VA). The lipids L-α-Phosphatidylglycerol (PG); L-α-Phosphatidylcholine (PC); 1, 2-dioleoyl-sn-glycero-phosphoglycerol) (DOPG) and 1,2 dioleoyl-sn-glycero-phosphocholine (DOPC) were purchased from Avanti Polar Lipids, Inc. (Alabaster, AL). From Invitrogen (Carlsbad, CA), the following dyes were used without modification: 1, 1'-dioctadecyl-3,3,3',3'-tetramethylindo-dicarbocyanine perchlorate (DiD) and Dextran, Alexa Fluor® 488; 3,000 MW, Anionic (dextran-488). From NN-labs, LLC. (Fayetteville, AR), 5 nm Magnetic Iron Oxide (Fe_3O_4) Nanocrystals in Non-Polar Solvent were used without modification.

Liposome preparation

Small unilamellar vesicles (SUVs) were prepared with varying lipid concentrations of PC and PG lipids. Lipids dissolved in chloroform were rotor-evaporated under nitrogen for 30 minutes followed by vacuum drying to remove all solvent. The lipids were re-suspended in DDI water to multilamellar vesicles (MLVs), which were then converted to SUV's using sonication for 1 hour with a Branson 1510 sonicator or by extrusion with Avanti MiniExtruder passing through a 50nm

polycarbonate filters 10 times. Size was determined using dynamic light scattering (DLS) using a Beckman Coulter N5 submicron particle size analyzer. Size data was collected using a 1 cm path-length cuvette, with an equilibrium time of 30 minutes at a temperature of 25°C and light scattering angle of 90°. The averages of 3 repetitions were used for final size determination.

Circular Dichroism

Circular dichroism (CD) spectra of the peptides were collected using a Jasco J-815 Spectropolarimeter. Samples were allowed to equilibrate at least 10 minutes at room temperature before data collection in a 1 mm path-length cuvette, at a chamber temperature 25°C. Spectra were collected from 190 to 260 nm using 0.2-nm intervals averaged over 5 scans. A peptide concentration of 125 µg/mL was used in samples containing 10 mM phosphate (pH 7.4) or 50 % 2, 2, 2 - trifluoroethanol (TFE). For samples containing liposomes, a peptide concentration of 62.5 µg/mL and a liposome concentration of 759 µg/mL in10 mM sodium phosphate buffer (pH 7.4) were used giving a peptide/ lipid ratio (P/L) of ~0.05. Boxcar smoothing was used to remove noise from signal, implemented by convolving the raw input data with a box-shaped pulse of ΣM+Mi+M-i/3.

Liposome Fusion/Aggregation Assay

These measurements were carried out in a 96-well format on a UV-Vis spectrophotometer at 436 nm over 30 minutes. The total sample volume of 200 µL was in 10 mM phosphate buffer. The lipid concentration of the sample was 759 µg/mL using varying liposome formulations including 100% PC, 80/20 PC: PG, 70/30 PC: PG and

60/40 PC: PG. The peptide concentration (L- and D-ATRA-1A) was varied from 0 to 125 µg/mL to provide P/L molar ratios ranging from 0 to ~0.1.

Antimicrobial Assay

The antimicrobial activity of ATRA-1A isomers, against *E. coli*, *B. cereus*, *S. aureus* and *P. aeruginosa* were determined. For antimicrobial assays, frozen aliquots of bacteria with known bacteria CFU/mL concentrations were thawed and mixed immediately before use. In a 96-well plate, 1×10^5 CFU per well bacteria were incubated with peptide concentrations 0 to 1000 µg/mL (in serial dilutions of 1:10 or 1:2) in a solution of sterile 10 mM phosphate buffer (pH 7.4) and incubated at 37° C (*E .coli*, *S. aureus* and *P. aeruginosa*) or 30° C (*B. cereus*) for 3 hours. Negative control wells contained bacteria with no peptide. Serial dilutions were then carried out in sterile 1× PBS and plated in triplicate on LB plates and incubated at 37° C (*E. coli*, *S. aureus* and *P. aeruginosa*) or 30° C (*B. cereus*) for 16 to 18 hours then colonies were counted. Bacterial survival at each peptide concentration was then calculated based on the percentage of survival of colonies in each experimental plate relative to the average number of colonies observed for assay cultures lacking peptide. The peptide concentration required to kill 50 % of the viable cells in the assay (EC50) was determined by plotting percent survival as a function of the log of peptide concentration (log µg/mL) and fitting the data, using GraphPad Prism (GraphPad Software Inc., San Diego, CA, USA) to Equation 1, which describes a sigmodial dose-response curve. Experiments against *S. aureus* and *P. aeruginosa* were performed by Scott Dean.

Equation 1

$$Y = \text{Bottom} + ((\text{Top}-\text{Bottom}) / (1 + 10^{[(\log EC50 - X) * \text{Hill Slope}]}))$$

Labeled Vesicle Preparation:

Formation of giant unilamellar vesicles (GUVs) via an electrochemical cell has been studied previously (*15*). GUVs were synthesized using electroformation with modified procedures adapted from earlier works (*16*, *17*). The lipids were combined in a 80:20 molar ratio of DOPC:DOPG in a chloroform and methanol solvent ratio of 9:1 (V:V). The fluorescent dye (DiD) was incorporated for an overall 0.08 mole percent compared to lipids. Iron nanoparticle (Fe_3O_4) solution was added to the lipid mixture at ~ 0.01% (v/v). The solvated lipids, nanoparticles and dye were mixed thoroughly to obtain homogeneity. Small droplets of the mixed sample (~ 2 µL) were applied to two Pt wire electrodes, each with a diameter of 1.2 mm. The lipid cake was applied as discrete droplets and allowed to evaporate. Once both electrodes were coated with approximately 10 µL of sample, the Pt wires were placed under vacuum for 2 hours to ensure complete solvent evaporation. A nonelectrolyte buffer solution of freshly prepared 2% (M/V) sucrose with 0.167 µM dextran-488 was heated to 80° C and combined with the Pt wires in a plastic cell. Multiple electrodes were then connected in parallel to a Hewlett Packard waveform generator. The sample temperature was maintained at 80° C during the entire process. The electroformation procedure began at 0.7 V with a frequency of 10 Hz. In a stepwise fashion, the voltage was increased 0.05 V every 5 minutes up to 1.4 V, where it was maintained for 3 hours. A final step of 0.6 V and 4 Hz was used to separate the vesicles from the wires. Sample cells were then removed from the oven and allowed to

cool slowly to room temperature. The effect of 4.5 hours at 80° C on DOPC vesicles via mass spectrometry was evaluated and showed no evidence of oxidation (data not shown). Experiments were performed by Dr. Susan Gillmor.

Labeled Vesicle Separation:

The iron oxide particles incorporated into the lipid bilayer, adding mass to the vesicle for separation (*18*, *19*). Small (500 µL) microcentrifuge tubes were nested into large (1500 µL) microcentrifuge tubes with a small magnetic flea at the tip of the larger centrifuge tube. The nested configuration was placed inside the centrifuge (Fisher Scientific, Pittsburg, PA). To separate, 300 µL of vesicle solution was mixed with 100 µL of unlabeled 2% sucrose solution. Microcentrifuge tubes of the mixture were placed in an Eppendorf microcentrifuge (Fisher Scientific, Pittsburg, PA) and were spun for 10 minutes at 13K rpm. The top 200 µL of the solution was removed and 200 µL of unlabeled solution is added and mixed. This process was repeated three times. In the microscopy imaging, the vesicles exhibited a higher signal and concentration from dextran-488 than the exterior solution, providing a means to track vesicle lysing. Experiments were performed by Dr. Susan Gillmor.

Vesicle Lysing:

For imaging and lysing, 135 µL of separated vesicles were placed in MicroWell 96-Well Optical-Bottom Plates (VWR, Radnor, PA) and were left to settle for 2 – 4 hours. After imaging to identify a high density of vesicles within a viewing window (143 µm x 143 µm), 15 µL of 1 mg/mL of L or D- isomer peptide was added to the well for final concentration of 100 µg/mL, unless otherwise noted. Vesicle lysing behavior was

documented by collecting a time series of images (every 30 seconds for 35 minutes) over the same window frame immediately after adding the peptide. Experiments were performed by Dr. Susan Gillmor.

Microscopy:

An inverted Zeiss LSM 510 confocal miscroscope was used with a 63x 1.2 NA water objective to image the vesicles. The dye (DiD) was excited using a HeNe 633 nm laser and the image was obtained by collecting the emissions from 650 nm – 750 nm. For dextran-488, imaging was conducted using an argon 488 nm laser line, collecting emission from 515-750 nm. Experiments were performed by Dr. Susan Gillmor.

Results and Discussion

Antimicrobial Performance

The antimicrobial effectiveness of the ATRA-1A peptide enantiomers have been assessed against a panel of Gram-positive and Gram-negative bacteria comprised of *E. coli*, *B. cereus*, *S. aureus*, and *P. aeruginosa*, and their performances against these microbes are compared. The results of these experiments are presented in Figure 1 A-D and the calculated EC50 values for the peptides are provided in Table 2.

Table 2: Antimicrobial Performance Data of ATRA-1A Isomers.
Antimicrobial activity is expressed in terms of EC50 (μg/mL) values and corresponding 95% confidence interval ranges for *S. aureus*, *B. cereus*, *P. aeruginosa* and *E. coli*. Data for *S. aureus* and *P. aeruginosa* courtesy of Scott Dean.

Antimicrobial Performance of ATRA-1A Peptide Isomers				
Bacterium	Antimicrobial Activity (EC_{50}, μg/mL)			
	L-ATRA-1A	95% CI Range	D-ATRA-1A	95% CI Range
S. aureus	1.9	1.3 to 2.8	26.1	19.8 to 34.4
B. cereus	72.9	64.8 to 82.0	2.3	2.1 to 2.6
P. aeruginosa	14.2	10.4 to 19.4	9.6	6.5 to 14.2
E. coli	4.3	4.0 to 4.6	1.4	1.1 to 1.5

The two peptide enantiomers show varied performances against the two Gram-positive bacteria tested, *S. aureus* and *B. cereus*. Against the Gram-positive bacterium *S. aureus*, L-ATRA-1A proves to be very effective with an EC50 of 1.9 μg/ml, while the D-peptide is significantly less potent, with an EC50 of 25.0 μg/ml (Figure 1A). L-ATRA-1A, with an EC50 of 1.9 μg/ml, is nearly 14 times more effective against *S. aureus* than the D-peptide (Table 2). In contrast, against *B. cereus* D-ATRA-1A has an EC50 of 2.3 μg/ml, which is over thirty times more effective than the L-peptide, which demonstrates an EC50 of 72.9 μg/ml (Figure 1B). The EC50 values for L- and D-ATRA-1A against *P. aeruginosa* (Figure 1C), a Gram-negative bacterium, are 14.2 μg/ml and 9.6 μg/ml, respectively, with significant overlap in the 95% confidence intervals (L-ATRA-1A: 10.4 to 19.4 μg/ml and D-ATRA-1A: 6.5 to 14.2 μg/ml), suggesting comparable effectiveness of the two peptide isomers against *P. aeruginosa*. Against *E. coli*, a Gram-negative bacterium, D-ATRA-1A exhibits an EC50 of 1.4 μg/mL, which is approximately three times more effective than the L-isomer with an EC50 of 4.3 μg/mL (Figure 1D). Although the difference in *E. coli* EC50 values express a small absolute value difference,

the 3-fold increase that D-ATRA-1A has over L-ATRA-1A is significant based on the 95% confidence interval and an independent t-test.

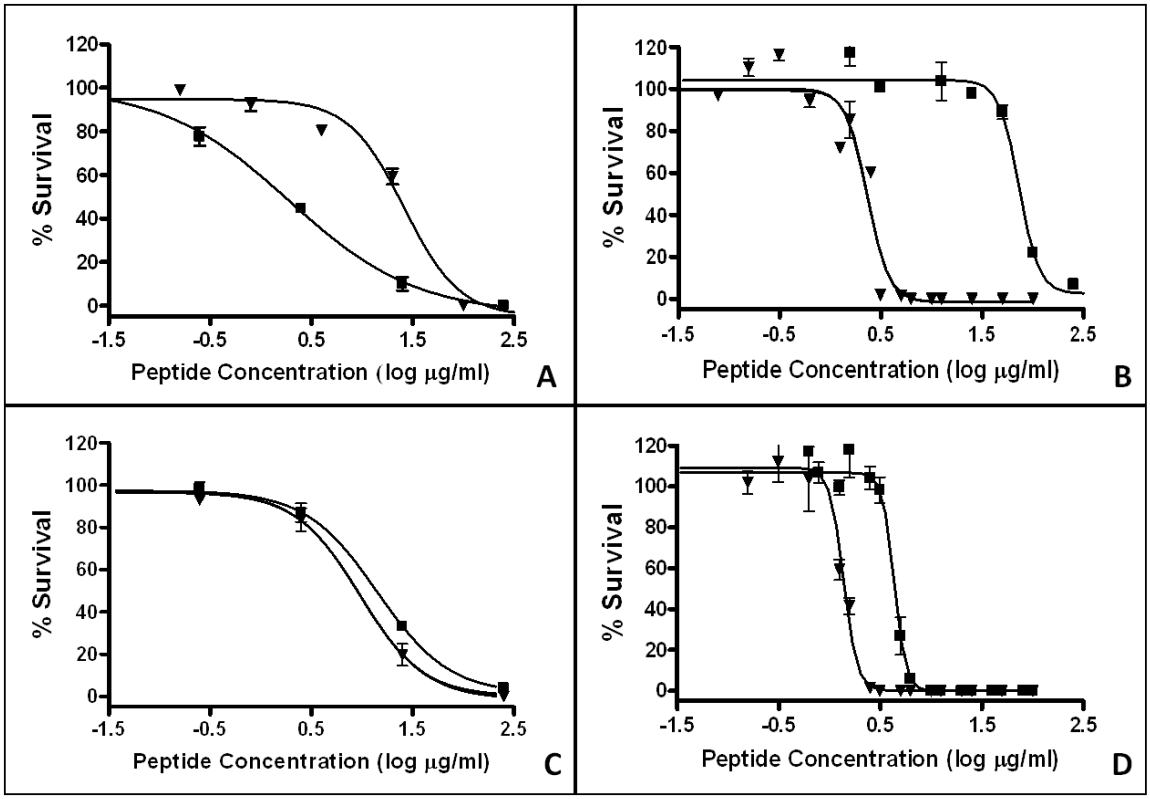

Figure 1: Antimicrobial Performance Dose-Response Curves.
Antimicrobial performances of L-ATRA-1A (represented in filled squares = ■) and D-ATRA-1A (represented in filled triangles = ▲) against *S. aureus* (A), *B. cereus* (B), *P. aeruginosa* (C) and *E. coli* (D). Data are fit to Equation 1, a standard equation for a dose-response relationship, in order to obtain EC50 values. Data for *S. aureus* and *P. aeruginosa* courtesy of Scott Dean.

Historically, differences in performance between the D- and L-enantiomers of antimicrobial peptides have been largely attributed to the enhanced protease resistance of the D-peptides (*10*); however the results for the ATRA-1A peptide isomers suggest a

more complex relationship between peptide stereochemistry and performance. The distinct differences in the performances of the L- and D-isomers (the D-isomer is more effective against *E. coli* and *B. cereus*, less effective against *S. aureus*, and equally effective against *P. aeruginosa*) are not easily attributed solely to protease resistance typically associated with D-peptides. If protease resistance were the only differentiating factor, D-ATRA-1A would be expected to exhibit antimicrobial activities that were superior or comparable to those of the L-peptide in the antimicrobial performance assays.

If bacterial membrane compositions are taken into account, then the ratio of lipids may influence the differing efficacy of the L- and D-isomers. Bacterial membranes are comprised of zwitterionic and anionic lipids, with lipid composition varying between Gram-positive and Gram-negative bacteria, as well as, between bacteria within the same class. The membranes of Gram-negative *E. coli* contain an 80:20 zwitterionic/anionic lipid ratio, whereas *P. aeruoginosa* have a 60:40 zwitterionic/anionic lipid ratio (*20*). The membranes of Gram-positive *S. aureus* contain only anionic lipids, while *B. cereus* membranes have a 40:60 zwitterionic/anionic lipid ratio (*20*).

Liposomes and membrane formulation

Direct interactions between CAMPs and bacterial membranes are believed to be key factors in their antimicrobial mechanism, and these membranes contain chiral elements. In the case of ATRA-1A enantiomers, differences in the way that the peptide stereoisomers interact with membrane lipids, which have defined stereochemistry, may contribute to the observed difference in their antimicrobial performances. Similar to the conformational chiral biasing observed for the achiral probe 1,6-diphenyl-1,3,5-

hexatriene when interacting with PC membranes, differences in the interactions between peptide isomers and the chiral elements in lipid membranes may manifest in the structural properties of the peptides when they interact with anionic membranes and their ability to induce membrane fusion/lysis (*14*). In the case of CAMPs, these interactions are likely affected by the membrane lipid composition.

Liposomes consisting of varied ratios of zwitterionic phosphatidylcholine (PC) and anionic phosphatidylglycerol (PG) have been used as simplified models for studying CAMP-membrane interactions and how they are affected by the membrane anionic character and lipid formulation. In these studies, liposomes with varied PC and PG ratios are prepared to approximate different bacterial membrane lipid formulations. These liposomes are then used to study the secondary structure induced in the D- and L-ATRA-1A isomers interacting with lipid membranes and how lipid formulation impact the ability of the peptides to induce membrane fusion and aggregation.

Circular Dichroism

Helical cathelicidins, such as the NA-CATH parent peptide, assume a helical conformation in the presence of anionic lipid membranes (*8*, *21*). The sequence of the ATRA-1A peptide, which is based on the ATRA motifs present in NA-CATH, is consistent with formation of an amphipathic helix. That being the case, the L-isomer would be expected to adopt a right-handed helical conformation and the D-isomer a left-handed one (*7*, *22*). Circular dichroism (CD) provides a means of assessing changes in conformational changes in the peptide backbone (*23–25*). In earlier studies, L-ATRA-1A at a concentration of 200 µg/mL has shown helical character in aqueous buffer containing

90 mM sodium dodecyl sulfate (P/L ~0.002), an anionic surfactant frequently used in CD studies to simulate an anionic membrane environment (*8*). In the present study, the effect environmental conditions have on the conformational properties of the L- and D-isomers of ATRA-1A are experimentally determined using CD.

Characteristic changes in the differences in absorption of circularly polarized light by the backbone amide bonds result in secondary structures, such as α-helix, β-sheet and random coil, being associated with unique CD spectral properties (*23*). For instance, α-helices present CD spectra with strong negative band peaks at 222 nm and 208 nm, corresponding to the $n \rightarrow \pi^*$ and the parallel component of the split $\pi \rightarrow \pi^*$ electronic transitions, respectively (*23*, *26*). For a peptide with 100% α-helix structure, the ratio of the signals at 222 nm and 208 nm, adjusted for concentration and the number of residues (mean residue ellipticity, [θ]), has been experimentally determined to be unity, R = 1 (*23*, *26*, *27*). Other helical conformations exhibit spectra that deviate from that of α-helices and can be detected by CD. For example, theoretical calculations and experimental CD spectral data collected for known 3_{10}-helix forming peptides reveal the presence of a weak shoulder at 222 nm and a strong band at 207 nm with an R = 0.4, and these features are considered indicative of 3_{10}-helical structure in a peptide (*26*–*30*).

In order to assess the basic structural properties of the peptide enantiomers, CD spectra have been collected for L- and D-ATRA-1A in phosphate and in 50 % TFE, which is commonly used to promote helical structure in peptides with helical tendencies (*26*). In 10 mM phosphate buffer (pH 7.4), both of the ATRA-1A peptide isomers present spectra consistent with random coil and no significant helical character (Figure 2A). In

contrast, both peptide isomers present spectra consistent with a high degree of helical structure in 50 % TFE (Figure 2B). The n → π*: π → π* peak ratio of each peptide isomer in 50 % TFE is shown in Table 2.

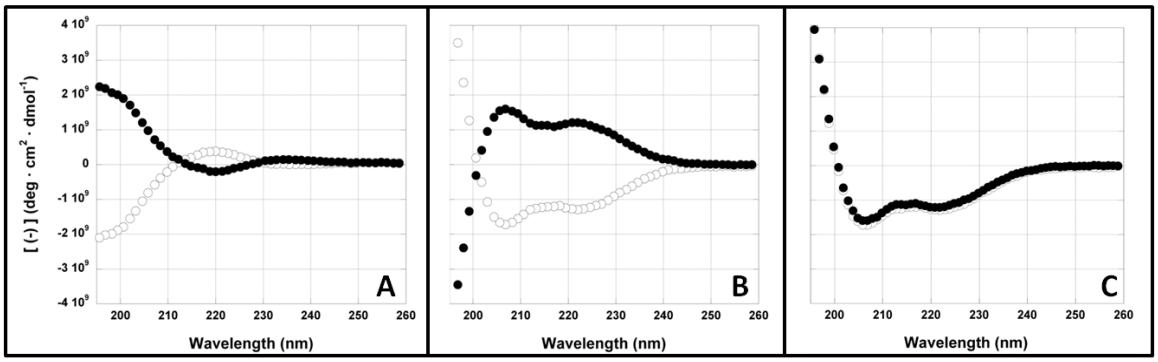

Figure 2: CD Spectra in Phosphate and TFE.
Circular dichroism spectra of the ATRA-1A isomers in 10 mM phosphate buffer (pH = 7.4) (A), 50 % trifluoroethanol (B) and circular dichroism spectra of L-ATRA-1A and corresponding inverted D-ATRA-1A spectra in 50 % TFE (C). L-ATRA-1A is represented in open circle (○) and D-ATRA-1A represented in closed circle (●).

The negative peak wavelengths for L-ATRA-1A were 222.2 nm and 206.0 nm giving a ratio of 0.75, while D-ATRA-1A exhibits positive peaks at 221.4 nm and 206.2 nm, giving a ratio of 0.75 (Table 3). The spectra for L- and D-ATRA-1A are near mirror images of each other, which would be expected, because aqueous TFE provides an achiral helix-promoting environment, which should not differentiate between the two stereoisomers. The mirror-image relationship between the spectra collected for D- and L-isomers in 50 % TFE is clearly illustrated by plotting the inverted spectra of the D-peptide with that of the L-isomer (Figure 2C). Both peptide isomers demonstrate a slight

blue shift in the $\pi \rightarrow \pi^*$ transition from that characteristic of an α-helix structure, as well as $n \rightarrow \pi^*$: $\pi \rightarrow \pi^*$ peak ratios of R = 0.75. This deviation from R =1 for α-helix could result from the N and C terminal residues being poorly defined, which may have a significant impact on the observed CD spectra due to the small size of the peptides.

Circular dichroism spectra have been collected for both L- and D-ATRA-1A in phosphate buffered solutions containing liposomes consisting of varied ratios of zwiterionic PC and anionic PG lipids in order to study the structural properties of the peptide isomers when interacting with membranes. Unlike TFE, the lipids comprising PC/PG liposomes contain defined stereocenters and present a chiral environment, which may impact their interactions with the peptide enantiomers. Moreover, varying the PC/PG ratios of the liposomes provides a means of determining the impact membrane anionic character has on the structural properties of the peptides. Spectra have been collected for the ATRA-1A enantiomers in the presence of liposomes consisting of 100% PC, 100% PG and varied combinations of PC/PG, shown in Figures 3A-E.

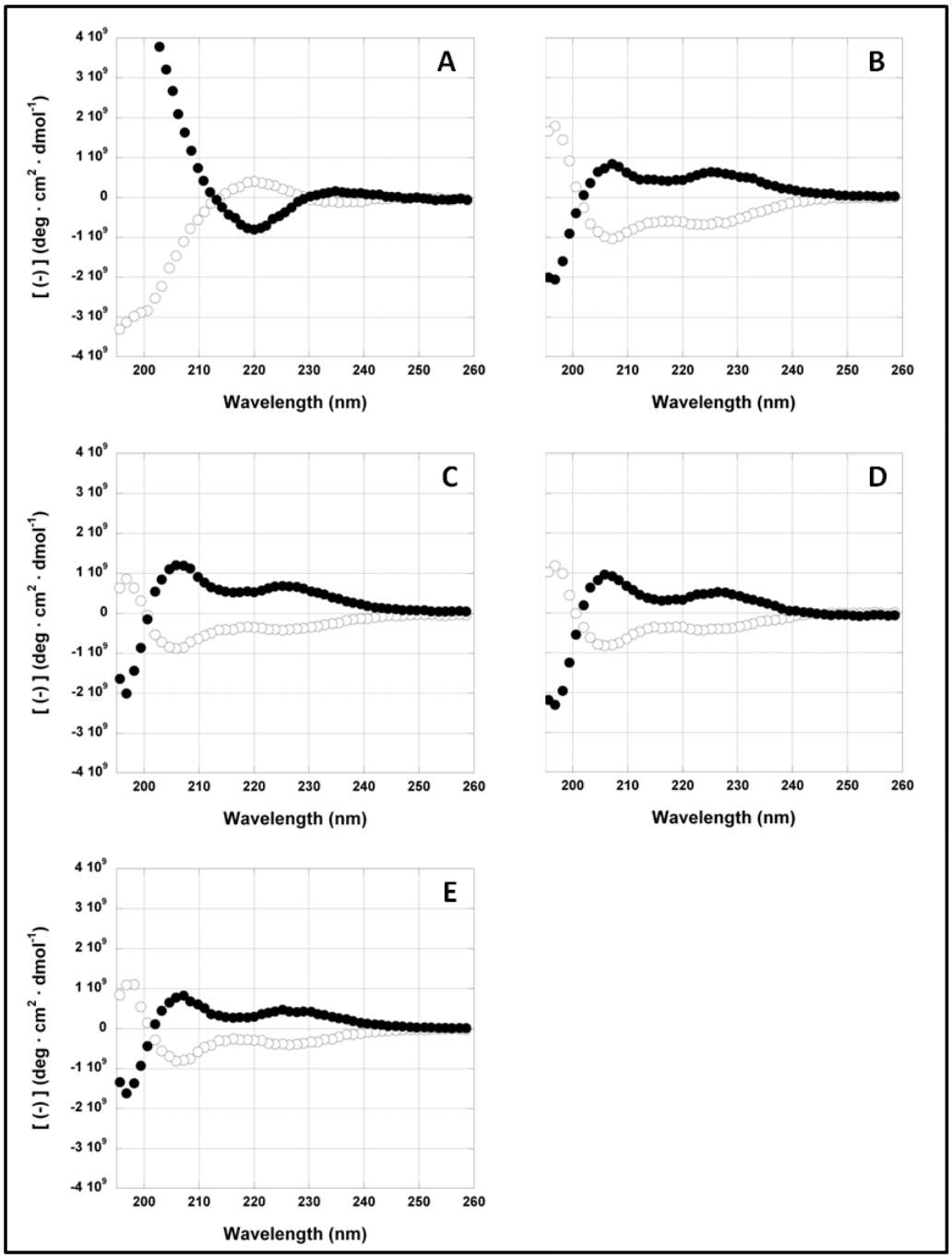

Figure 3: CD Spectra in PC and Varied PC/PG Liposome Formulations.
Circular dichroism spectra of ATRA-1A isomers with liposomes of varying anionic concentration (100% PC [A], 80:20 PC/PG [B], 70:30 PC/PG [C], 60:40 PC/PG [D] and 100% PG [E]) and 10 mM phosphate buffer (pH = 7.4). L-ATRA-1A is represented in open circles (○) and D-ATRA-1A represented in closed circles (●).

In the presence of 100% PC liposomes (Figure 3A), a fully neutral vesicle, both ATRA-1A peptide isomers do not exhibit defined bands at 222 nm or 208 nm and are consistent with random coil structure. The spectra are similar to those collected for the peptides in 10 mM phosphate buffer (pH 7.4). These results suggest that either no significant interaction occurs between the peptide isomers and the neutral membranes or any interactions that do occur do not result in significant helical structure in the peptides.

As anionic lipid, in the form of PG, is introduced into the liposome formulations the spectra of the isomers change dramatically from those involving neutral PC liposomes. The spectra for the peptide enantiomers in the presence of liposomes containing PG suggest a shift in the secondary structures of the peptide enantiomers from random coil to helical (Figures 3B-E). In aqueous buffer containing liposomes with 80:20 PC/PG, seen in Figure 3B, both ATRA-1A isomers exhibit defined peaks in the 222 nm and 208 nm regions, indicating the presence of helical structure. Distinct negative band peaks at 224.2 nm and 206.0 nm are present in the spectrum of L-ATRA-1A, and strong positive band peaks at 225.4 nm and 207.0 nm are seen in that of the D-isomer. Subtle differences occur in the spectra of the peptide isomers under these conditions, which are clearly illustrated when the inverse spectra for D-ATRA-1A is overlaid with that of the L-isomer (Figure 4A). While both peptide isomers exhibit a slight red shift in the $n \rightarrow \pi^*$ transition from the α-helical ideal of 222 nm and a blue shift in the $\pi \rightarrow \pi^*$ transition from the ideal of 208 nm (Table 3), the L-ATRA-1A $n \rightarrow \pi^*$: $\pi \rightarrow \pi^*$ peak ratio is R = 0.66 and that of D-ATRA-1A is R = 0.74. Both represent significant deviations from the ideal peak ratio R = 1 that is associated with an all α-helical peptide. The shifts in peak

wavelengths and peak ratios observed for the ATRA-1A isomers suggest that the nature and degree of structure that they adopt upon interaction with the 80:20 PC/PG liposomes differ and may reflect contributions of different helix types.

Table 3: CD Signal Analysis of ATRA-1A Isomers.
CD spectra peak ratios for ATRA-1A isomers. The signal intensities (MRE = deg• cm^2 • $dmol^{-1}$) for each peak corresponding to the n → π* and π → π* electronic transitions were obtained for both isomers in various environments, including liposomes of varied PC/PG formulations. These peak ratios for the peptide isomers were then determined in order to further highlight conformational differences that may arise as environmental conditions changed.

	50% Trifluoroethanol		**80:20 PC/PG Liposomes**		**70:30 PC/PG Liposomes**		**60:40 PC/PG Liposomes**		**100% PG Liposomes**	
	Wavelength (nm)	Signal Intensity (MRE)	Wavelength (nm)	Signal Intensity (MRE)	Wavelength (nm)	Signal Intensity (MRE)	Wavelength (nm)	Signal Intensity (MRE)	Wavelength (nm)	Signal Intensity (MRE)
	D-ATRA-1A		**D-ATRA-1A**		**D-ATRA-1A**		**D-ATRA-1A**		**D-ATRA-1A**	
n → π*	221.4	1.21E+09	225.4	6.42E+08	224.6	6.87E+08	226.4	5.27E+08	224.6	4.86E+08
π → π*	206.2	1.61E+09	207	8.72E+08	205.8	1.25E+09	205.8	9.98E+08	206.6	8.45E+08
Ratio		0.75		0.74		0.55		0.53		0.58
	L-ATRA-1A		**L-ATRA-1A**		**L-ATRA-1A**		**L-ATRA-1A**		**L-ATRA-1A**	
n → π*	222.2	-1.30E+09	224.2	-6.99E+08	225	-4.29E+08	223.8	-4.43E+08	226.6	-4.16E+08
π → π*	206	-1.72E+09	206	-1.06E+09	205.2	-8.94E+08	205.4	-8.44E+08	205.8	-8.23E+08
Ratio		0.75		0.66		0.48		0.53		0.51

Increasing the liposome anionic lipid content to 70:30 PC/PG (Figure 3C), results in shifts in the peptide isomer spectra, relative to the spectra collected with 80:20 PC/PG liposomes. While the n → π* and π → π* transition wavelengths for both isomers are similar to those observed for 80:20 PC/PG liposomes, the peak ratios are smaller (Table 3). L-ATRA-1A has a weak shoulder at 225.0 nm and stronger band at 205.2 nm with a peak ratio of R = 0.48. This ratio is very close to the R = 0.4 of a 3_{10}-helix proposed by Toniolo et al. in 1996, suggesting the possible presence of 3_{10}-helical character in the

peptide (*27*). The D-ATRA-1A spectrum exhibits a similar shift in peak ratio, R = 0.55, indicating a change in secondary structure compared to the structure seen in 80:20 PC/PG liposome-peptide interaction. Moreover, the peak ratios for the two enantiomers differ significantly from each other, suggesting conformational differences exist between the peptide isomers in the presence of the 70:30 PC/PG membranes.

When the anionic lipid content is increased to 40% of total lipid, 60:40 PC/PG liposomes, another shift in the CD spectra of the two ATRA-1A enantiomers is observed (Figure 3D). L-ATRA-1A exhibits negative maxima at 223.8 nm and 205.4 nm with a ratio of R = 0.53, while D-ATRA-1A shows maxima at 226.4 nm and 205.8 nm with a ratio of R = 0.53 (Table 3). The spectra for the L-ATRA-1A isomer in the presence of 60:40 PC/PG liposomes present a slight blue shift in the $n \rightarrow \pi^*$ transition from that observed for 70:30 PC/PG liposomes (223.8 nm), while that of D-ATRA-1A exhibits a red shift to 226.4 nm (Figure 3D and Table 3). Neither isomer shows a significant wavelength shift in the $\pi \rightarrow \pi^*$ transition. While the peak ratio in the L-ATRA-1A spectra increases to R = 0.53, and the peak ratio for the D-isomer (R = 0.53) was very similar to that exhibited in the presence of 70:30 PC/PG liposomes. The large increase in peak ratio for L-ATRA-1A likely reflects a change in the peptide conformation. The peak ratios for both enantiomers in the presence of 60:40 PC/PG liposomes are identical, suggesting that the secondary structures of the two isomers are near mirror images of each other, with the mirror image relationship between the spectra becoming evident when the spectra for D-ATRA-1A is inverted and overlaid on the spectra for the L-isomer (Figure 4B).

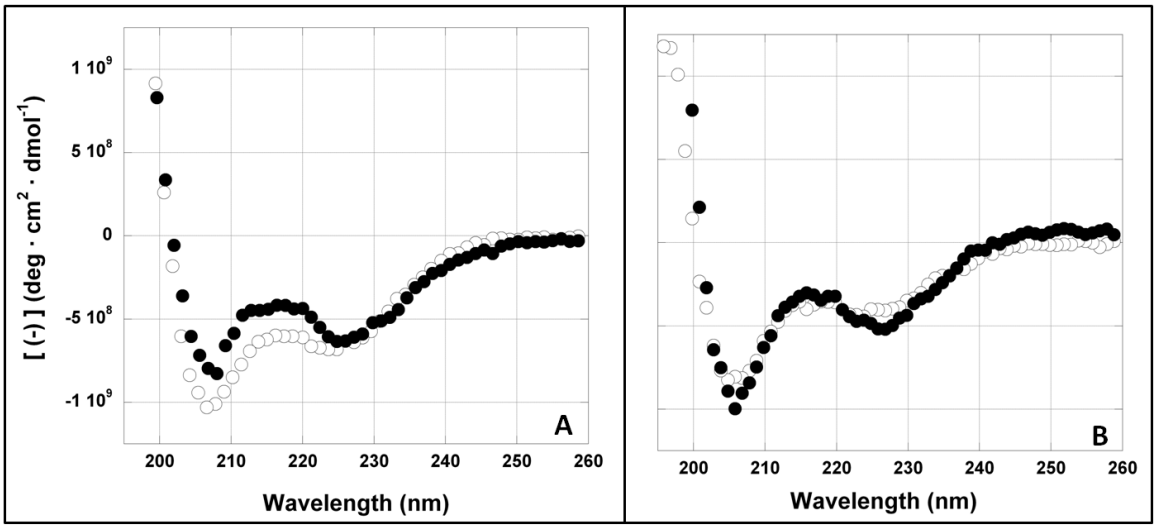

Figure 4: Inverse CD Spectra of ATRA-1A Isomers in the Presence of Liposomes. Circular dichroism spectra of L-ATRA-1A and corresponding inverted D-ATRA-1A spectra in the presence of 80:20 PC/PG (A) and 60:40 PC/PG (B) and 10 mM phosphate buffer (pH = 7.4). L-ATRA-1A is represented in open circles (○) and D-ATRA-1A represented in closed circles (●).

Liposomes consisting of 100% PG are used to simulate membranes composed entirely of anionic lipids, and further changes are evident in the CD spectra for both ATRA-1A isomers collected under these conditions (Figure 3E). The spectrum for the L-ATRA-1A isomer in the presence of 100% PG liposomes presents a slight red shift in the n → π* transition from that observed with 60:40 PC/PG liposomes (226.6 nm), while that of D-ATRA-1A exhibited a slight blue shift to 224.6 nm (Table 3). L-ATRA-1A showed no significant shift in the wavelength of the π → π* transition, while D-ATRA-1A showed a slight red shift to 206.6 nm. The peak ratio in the L-ATRA-1A spectrum, R = 0.51, is similar to the peak ratio observed in the 60:40 PC/PG spectrum (R = 0.53).

However, the peak ratio for the D-isomer (R = 0.58) increases slightly from the ratio R = 0.53 observed in spectra collected using 60:40 PC/PG liposomes. The increase in peak ratio for D-ATRA-1A suggests a change the peptide conformational properties. The changes in the secondary structure properties of the enantiomers that occurred as the anionic lipid content of the liposomes are increased suggest that the degree and nature of the helical conformation of the peptide enantiomers are differentially influenced by the anionic lipid content in these simplified model membranes. In addition to the potential conformational flexibility of residues at the N- and C-termini, more detailed analysis of the structural properties of the peptide isomers by CD may be complicated by the presence of 2 phenylalanine residues in the 11-residue ATRA-1A sequence. The L_a and B low energy transitions of the $\pi \rightarrow \pi^*$ transition of the aromatic ring of phenylalanine (208 nm and 188 nm, respectively) overlap with the corresponding transition for the peptide amide bond (208 nm) (*31*).

Peptide-induced Liposome Fusion and Aggregation

Membrane aggregation is not considered a spontaneous process and usually requires stimulus from an aggregate-inducing agent, such as cationic antimicrobial peptides (*32*). The ability of CAMPs to induce fusion is a direct result of their intimate interaction with membranes. Because the fusion and aggregation of liposomes is associated with increased turbidity, monitoring the turbidity of peptide-liposome solutions provides a means of studying peptide-induced membrane fusion and aggregation (*32*, *33*). By varying the CAMP concentration and holding liposome

concentration constant, it is possible to elucidate peptide concentration dependent membrane fusion behavior.

Here, the ability of the ATRA peptide isomers to induce fusion of liposomes of varied lipid composition (100% PC, 80:20, 70:30, 60:40 PC/PC and 100% PG) has been evaluated by incubating peptides in buffers containing liposomes for thirty minutes and monitoring the absorbance at 436 nm for each solution.

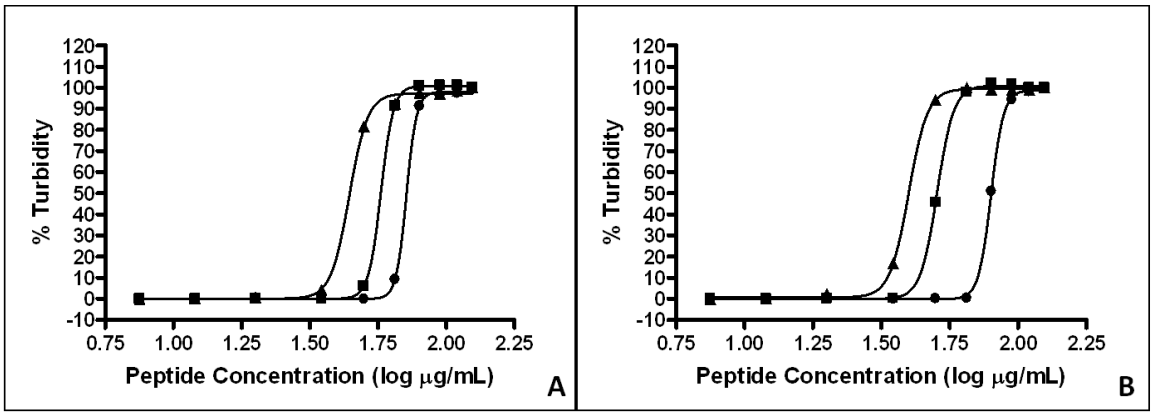

Figure 5: Peptide-Induced Turbidity Dose-Response Curves.
Peptide induced turbidity data for L-ATRA-1A (A) and D-ATRA-1A (B) with liposomes containing 80:20 (▲), 70:30 (■) and 60:40 (●) PC/PG lipid ratios. Data are fit to a standard dose-response equation in order to obtain EC50 values, corresponding to the peptide concentration required to achieve half-maximal induced turbidity.

The resulting data is then used to generate dose-response plots correlating percent maximum turbidity with log of peptide concentration. In the absence of peptide, liposome suspensions for all lipid formulations tested are stable and no change in turbidity is observed over the 30-minute incubation time. Even at the highest peptide concentrations tested (125 μg/ml), introduction of neither L- nor D-ATRA-1A to solutions containing

neutral zwitterionic liposomes, 100% PC, results in a change in turbidity, indicating that no detectable induced fusion/aggregation is occurring (data not shown). Percent maximum turbidity for both ATRA-1A peptide isomers is plotted as a function of peptide concentration (range of 0 to 125 µg/ml) following incubation with anionic PC/PG liposomes (Figure 5), and the resulting data is fit to a standard equation, describing a dose-response relationship, to generate EC50 values for L-ATRA-1A with each liposome formulation (Table 4).

Table 4: Peptide-Induced Turbidity of ATRA-1A Isomers.
Peptide induced turbidity data for ATRA-1A isomers. Peptide induced turbidity is expressed in terms of EC50 (µg/mL) values and corresponding 95% confidence interval ranges for liposome formulations containing 80:20, 70:30 and 60:40 PC/PC.

Induced Turbidity of ATRA-1A Peptide Isomers				
Liposome Formulation	**Turbidity (EC_{50}, µg/mL)**			
	L-ATRA-1A	95% CI Range	D-ATRA-1A	95% CI Range
80:20 PC/PG	44.2	42.5 to 45.9	39.9	39.0 to 40.7
70:30 PC/PG	57.7	56.6 to 58.9	50.7	47.8 to 53.7
60:40 PC/PG	71.5	70.7 to 72.4	79.8	78.5 to 81.1

Maximum turbidity for each liposome formulation corresponds to the highest turbidity attainable for each peptide isomer, at which point turbidity remained unchanged over the higher range of peptide concentrations tested. The EC50 value for L-ATRA-1A and 80:20 PC/PG liposomes is determined to be 44.2 µg/mL, for 70:30 PC/PG liposomes the L-ATRA-1A EC50 is found to be 57.7 µg/mL, and with 60:40 PC/PG liposomes an

EC50 of 71.5 µg/mL is noted (Table 4). The EC50 values for D-ATRA-1A against 80:20, 70:30 and 60:40 PC/PG liposomes are determined to be 39.9 µg/mL, 50.7 µg/mL and 79.8 µg/mL, respectively (Table 4). Similar experiments, performed with liposomes consisting of 100% PG, exhibit complex behavior and no clear turbidity maximum is observed (Data not shown). The correlations between anionic lipid content in PC/PG liposomes and their susceptibilities to D- and L-ATRA-1A induced fusion/aggregation are evident in the observed relationship between EC50 values and PG content for each peptide isomer, with EC50 values increasing as PG content increases (Figure 6).

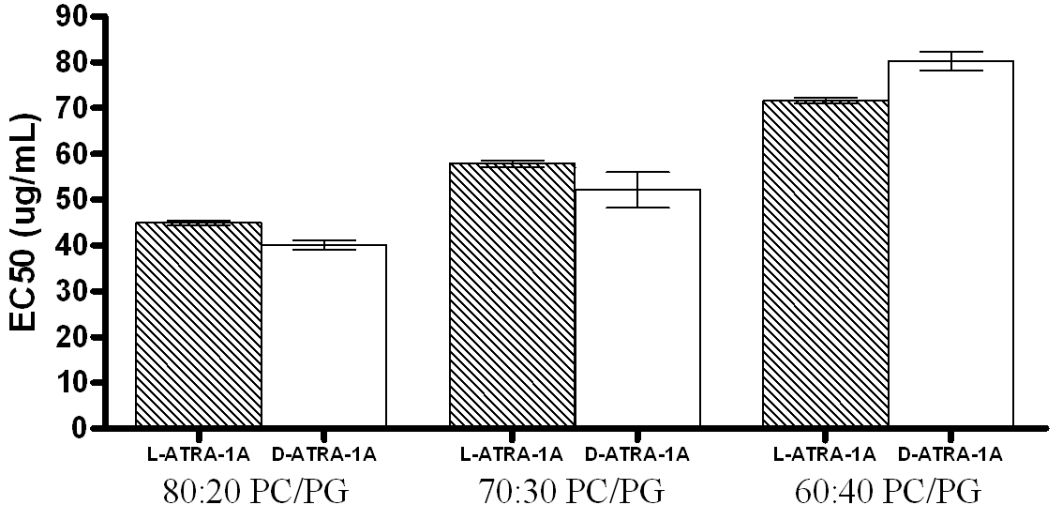

Figure 6: Peptide-Induced Turbidity EC_{50} Comparison of ATRA-1A Isomers. L-ATRA-1A (hatched) and D-ATRA-1A (solid) fusion/aggregation in liposome formulations 80:20, 70:30 and 60:40 PC/PC. The EC50 values, corresponding to half-maximal induced turbidity, are plotted and standard deviations based on duplicate experiments are provided.

These trends likely reflect the fact that as the anionic lipid content is increased, these anionic lipids need to be desolvated and the electrostatic repulsion between juxtaposed liposomes must be overcome before fusion/aggregation can occur (*32*). Similar results were observed for the 32 amino acid murine CAMP, Cryptdin-4, with liposomes of varied zwiterionic/anionic ratios (*32*).

The data from these fusion/aggregation studies suggest there may be differences in the abilities of the two ATRA-1A isomers to induce membrane fusion. The differences in the EC50 values for the ATRA-1A isomers are small, but each experiment was performed in duplicate and the values are outside of the 95% confidence interval calculated from the dose-response curve fit. In these studies, the D-isomer appears to be more effective at inducing fusion/aggregation in liposomes containing 80:20 and 70:30 PC/PG (Figure 6). While at a higher anionic ratio of 60:40 PC/PG, L-ATRA-1A appears more effective at inducing membrane fusion (Figure 6). These results and the structural properties of the ATRA-1A isomers, seen in the CD liposome studies show subtle differences in the way that D- and L-ATRA-1A interact with the model membranes.

Microscopy

Vesicles with a lipophilic dye (DiD) incorporated into their bilayers and filled with buffer containing fluorescently tagged dextran have been used to visualize the activity of the peptide isomers on lipid membranes. The bright internal contents of the vesicles highlight leaking and lysing caused by L-ATRA-1A or D-ATRA-1A. In these studies done by our collaborator, Dr. Susan Gillmor at George Washington University, each peptide isomer is added to microwells filled with vesicles. Time sequence images

record snapshots of the vesicles at regular time intervals revealing the effect of the peptides on the vesicles (Figure 7).

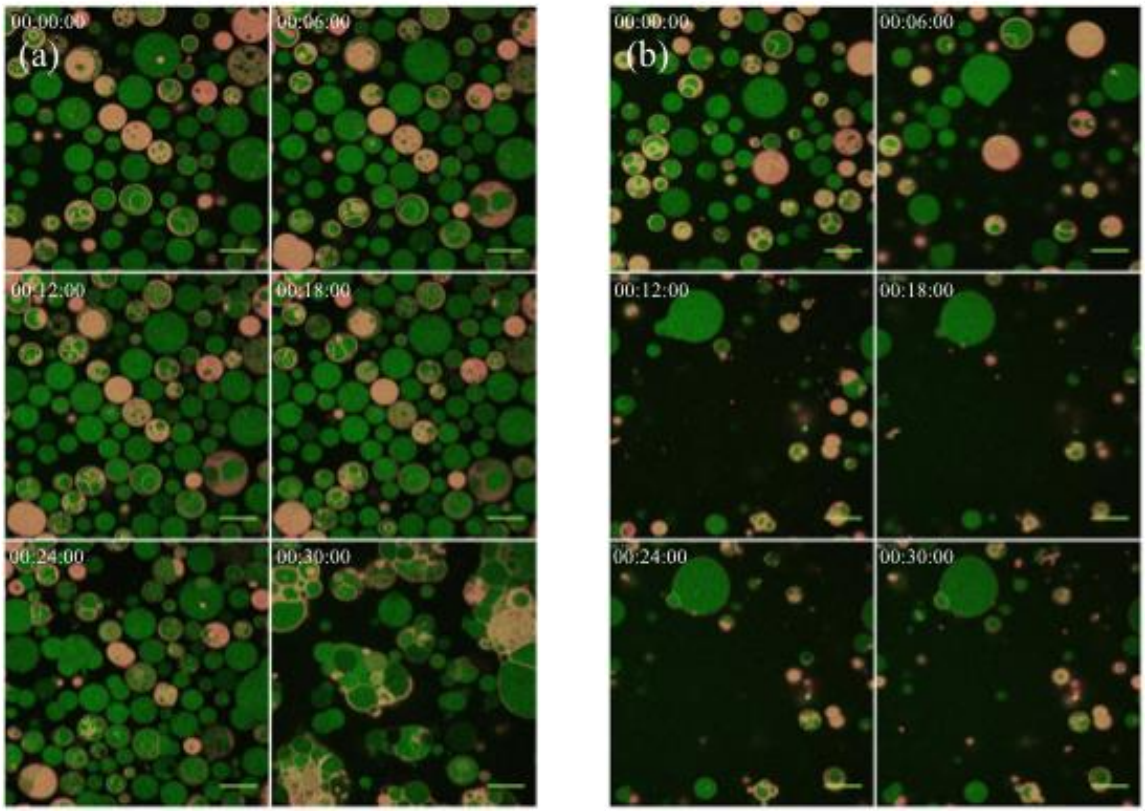

Figure 7: ATRA-1A Isomer Interaction with Vesicles.
Peptide interaction with vesicles. (a) L-ATRA-1A is added to a microwell filled with lipid vesicles of 80:20 DOPC:DOPG (100 µg/mL). The images above show snapshots of peptide-vesicle activity every 6 min from t = 0 min to t = 30 min. The onset of activity begins at 23 min (see movie S1) and results in vesicle fusion, bilayer rearrangement, lysing and leakage of tagged solution. (b) D-ATRA-1A (100 µg/mL) is added to a microwell of identical content. The snapshots of peptide-vesicle activity every 6 min from t = 0 min to t = 30 min reveal an onset of activity at 3 min (see movie S2). The resulting interaction shows vesicle lysing, leakage of tagged solution and vesicle fusion. All scale bars correspond to 20 µm. Microscopy images courtesy of Dr. Susan Gillmor.

Differences in lag time from activity onset with the addition of each peptide isomer are observed (see movies S1 & S2). However, the sample size in microscopy is limited to the imaging window, and the relative number of vesicles that the average peptide encounters before entering the viewing field is unknown. Therefore, caution is taken in drawing any conclusion on peptide activity due to lag time. It is expected to be different from well to well on the microwell plate. Nevertheless, from the onset of the peptide activity, L-ATRA-1A and D-ATRA-1A exhibit similar duration times of interaction from initial lysing to final vesicle state

The membrane lysing behavior of the peptides is confirmed using fluorescence microscopy and vesicles containing PG lipids. Furthermore, visualization of lysing in this manner reveals multiple behaviors of the ATRA-1A peptide isomers. First, the complete release of vesicle contents and bilayer destabilization (Figure 7 (a) and (b), movies S1-3) is observed when ATRA-1A peptide isomers are present. It is inferred that lysis proceeds via many pores in close proximity or a few large pores. In either case, these pores are below the optical resolution of the microscope (less that 500 nm). The images show lysing, indicating a high localized concentration of peptides in a small region of the vesicle bilayer, which suggests peptide clustering and cooperative behavior. Furthermore, movie S3 captures an excellent example of uneven dye distribution. It is clearly seen that DiD (lipophilic dye in the bilayer) on the surface of the vesicle clusters into bright regions following introduction of peptide. These vesicles are single phase and the dye in a single-phase vesicle is evenly dispersed, so there should be no bright spots or clustering. The presence of D-ATRA-1A induces uneven dye distribution, which suggests

clustering of the anionic DOPG and the cationic peptide. This is not a widespread phenomenon nor is it exclusive to the D isomer. The cationic peptide interactions with anionic lipids are non-specific, and similar clustering behavior with the L isomer is expected.

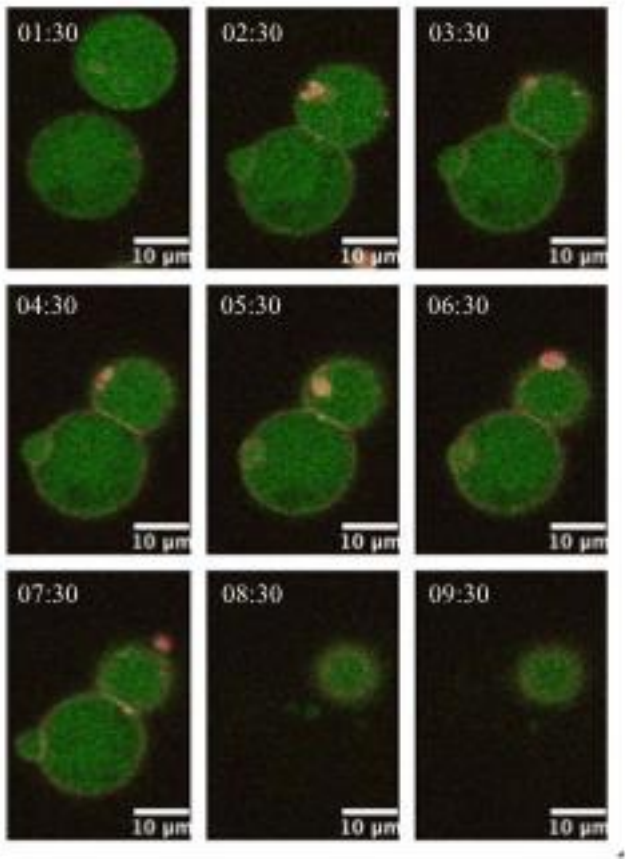

Figure 8: D-ATRA-1A Vesicle Fusion.
Vesicle fusion in the presence of D-ATRA-1A. A series of time sequence microscopy images highlights a vesicle fusion event (300 µg/mL, D-ATRA-1A). In less than 60 sec, two vesicles merge and form a stable, single bilayer junction. After 7:30 min, the lower portion of the vesicle ruptures, leaving the top section intact. Microscopy images courtesy of Dr. Susan Gillmor.

Second, several vesicles fade over time compared to their bright neighbors. Potentially, multiple, dispersed pores allow solution to exchange across the bilayer allowing tagged dextran to escape from the vesicle interior. In movie S1 (L-ATRA-1A), two compartments of the central, merged vesicles lose signal from t = 27 minutes to t = 35 minutes. Unlike lysing, the membrane remains stable, suggesting that the pores and peptide are dispersed throughout the membrane. The formation of stable pores, as the microscopy data suggest, indicate that the peptide exhibits a long dwell time in the bilayer.

Finally, membrane rearrangements, including budding and fusion events, are evident for both ATRA-1A isomer behaviors. In Figure 7 (a) & (b), movie S1 (L-ATRA-1A) and Figure 8 (D-ATRA-1A), adjacent vesicles fuse, merging two bilayers into a single layer. This activity suggests cooperative behavior between peptides in adjacent bilayers. Moreover, the newly formed vesicle is stable, suggesting that the pores and peptides are dispersed and exhibit a long peptide dwell time in the bilayer.

A variety of outcomes including but not limited to lysing, leaking, fusion, clustering and bilayer rearrangement have been documented through direct observation of both peptides interacting with vesicles using microscopy. The multiple activities associated with both ATRA-1A enantiomers suggest that the same peptide isomer may induce different effects, and local peptide concentration determines the formation of many stable pores in close proximity or disperse pores, leading to leaking or lysing. Peptide dwell time in the bilayer seems to be an important factor influencing clustering between the cationic peptide and anionic lipid (see movie S3), similar to clustering

behavior in other systems, such as Cryptdin-4 (*32*), GM1 and the cholera toxin B (*34*, *35*) and antigens to the B cell receptor (*36–38*). The occurrence of vesicle fusion and lipid rearrangement, instead of limited disruptive behavior, also depends on peptide-membrane interactions over time. While each peptide exhibits all behaviors described above, L-isomer displays a higher degree of fusion and bilayer rearrangement compared to the D-isomer. In contrast, the D-enantiomer shows more lysing activity.

Conclusion

The potency differences exhibited by the ATRA-1A stereoisomers against the four evaluated bacterial strains are too complex for them to be easily attributed solely to the increased protease resistance typically associated with D-peptides. Circular dichroism spectra collected for the ATRA-1A isomers in the presence of mixed zwitterionic/anionic liposomes provide evidence that the conformational properties of the peptide isomers are influenced by membrane anionic lipid content. In future studies, nuclear magnetic resonance would provide a powerful and complimentary tool for assessing the conformational properties of the ATRA-1A isomers in the presence of liposomes in greater detail. Moreover, the spectra suggest that significant differences arise in the conformations of the two isomers in the presence of 80:20 and 70:30 PC/PG liposomes, which are lipid formulations similar to those found in bacterial membranes. Similarly, subtle differences exist in the ability of the ATRA-1A isomers to induce liposome fusion/aggregation and how liposome formulation impacts their activity. Fluorescence microscopy of liposomes treated with the ATRA-1A isomers indicates that L-isomers favor fusion and bilayer rearrangement, while D-isomers exhibit a high number of lysing

events to disrupt 80:20 DOPC/DOPG liposomes. Overall, lysing, fusion, leaking, clustering and bilayer rearrangements are documented for both peptide isomers. Combined, the results of the microscopy and turbidity studies reveal subtle differences in the way that the ATRA-1A enantiomers, and perhaps other helical CAMPs, interact with membranes. These differences may contribute to divergences in their antimicrobial properties. However, the data from these studies do not reveal a clear correlation between the distinct antimicrobial activities of the ATRA-1A peptide stereoisomers and their interactions with the chiral membrane. Further study is required to determine how and if the observed differences in their interactions with membranes relate to CAMP antimicrobial properties. Because D-peptides provide a route to improving proteolytic stability, better understanding how CAMP stereochemistry impacts performance may be important in efforts to develop CAMP-based therapeutics.

CHAPTER TWO. MECHANISTIC STUDY OF CATIONIC ANTIMICROBIAL PEPTIDES: NA-CATH, L-ATRA-1A AND D-ATRA-1A

Cationic antimicrobial peptides (CAMPs) are important elements of innate immunity in higher organisms, representing an ancient defense mechanism against pathogenic bacteria. These peptides exhibit broad-spectrum antimicrobial activities, utilizing mechanisms that involve targeting bacterial membranes. Recently, a 34-residue CAMP was identified in cDNA from the venom gland of the Chinese cobra (*Naja atra*), and an 11-residue truncated peptide, ATRA-1A, was generated based on a semi-conserved 11-residue pattern observed in its sequence. While the antimicrobial and biophysical properties of stereoisomers of ATRA-1A have been studied, their modes of action remain unclear. Studying the ability of full-length NA-CATH and the ATRA-1A isomers to induce bacterial membrane depolarization and cause more general membrane disruption will provide insights into fundamental differences and similarities in the ways that they attack membranes. Such information would further facilitate the development of shortened peptides based on larger naturally occurring antimicrobial peptides for potential therapeutic applications.

Introduction

Cationic antimicrobial peptides (CAMPs) are pervasive in nature and represent an evolutionarily ancient mechanism for defending against invading microorganisms. These peptides exhibit broad spectrum antimicrobial effectiveness and important elements of

innate immunity, which provide the first line of defense against infection. Despite their widespread use they exhibit limited bacterial resistance. These qualities provide CAMPs an advantage over conventional therapeutics for fighting infections. Although CAMPs offer great potential as the basis for a new class of antibiotics, many of the details of the mechanisms by which they exert their antimicrobial effects remain unclear. Greater understanding of the relationship between CAMP physico-chemical properties and antimicrobial action is needed in order to realize their therapeutic potential.

CAMPs have been shown to interact with bacterial membranes and in many cases induce membrane disruption. However, these interactions appear to be complex and the correlations between peptide physico-chemical properties, membrane composition and modes of action are poorly understood. CAMPs are usually amphipathic peptides presenting discreet cationic and hydrophobic surfaces. The spatial partitioning of these surfaces allows favorable electrostatic interaction with negatively charged lipid head groups on the outer surface of bacterial membranes and insertion into the hydrophobic interior of the bilayer, leading ultimately to membrane disruption (*39*, *40*). Widely accepted membrane disruption mechanisms range from the "barrel-stave" model to the "carpet model". In the "barrel-stave" model, amphipathic helical peptides insert into the membrane, forming peptide lined structures with large central pores (*41–43*). A similar proposed model is the "toroidal pore" where amphipathic helical peptides insert into the lipid membrane and form less defined transient supramolecular pores (*41*, *42*, *44*). In the "carpet model," the peptides gather and concentrate at the membrane surface, interacting with the anionic lipid head groups, until the peptide concentration threshold is reached.

This results in distortions in the lipid bilayer curvature and formation of transient gaps in the membrane (*41–43*).

CAMPs can be grouped into families based on multiple factors, such as evolutionary relationships and conserved sequence patterns and structural elements. Cathelicidins are a sequence diverse family of vertebrate antimicrobial peptides that are identified based on the highly conserved cathelin domain present in the precursor protein (*1–4*). Recently, the sequence of a 34-residue helical cathelicidin, NA-CATH, was identified in cDNA from the venom glands of the elapid snake, Naja atra (*5*, *6*). Analysis of the NA-CATH sequence revealed a semi-conserved 11-residue repeated sequence pattern. The 11-residue peptide amide, ATRA-1A, was designed based on this pattern. The D-isomer of ATRA-1A was also generated because D-peptide isomers are generally more resistant to proteases than the corresponding L-peptides, and this increased resistance to proteolytic degradation could enhance their therapeutic utility. In previous studies, significant differences were observed in the antimicrobial activity, structural interactions and disruption induced in model membranes by the ATRA-1A peptide isomers (*45*). The present study focuses on key aspects of peptide-induced membrane disruption and antimicrobial kinetics, with the aim of elucidating more clearly similarities and differences in the mechanisms employed by full-length NA-CATH and the ATRA-1A isomers.

Material and Methods

The peptides used in these studies were custom synthesized by AAPPTEC, LLC (Louisville, KY). The supplier reported purities of NA-CATH, L-ATRA-1A and D-

ATRA-1A were 95 %, 95.2 % and 95.4 %, respectively, based on HPLC analysis of the purified peptides. The bacterial strains of *Escherichia* (*E.*) *coli* (ATCC# 25922) and *Bacillus* (*B.*) *cereus* (ATCC# 11778) used in these studies were purchased from the American Type Culture Collection (Manassas, VA). 3, 5-Dipropylthiacarbocyanine ($diSC_3$-(5)) was purchased from AnaSpec (Fremont, CA). SYTOX Green was purchased from Invitrogen (Carlsbad, CA). Mueller Hinton Broth (MHB) was purchased from Becton Dickinson and Company (Sparks, MD). Phosphate buffered saline (PBS) was purchased from Corning-cellgro (Manassas, VA). Resazurin, sodium salt is purchased from Sigma-Aldrich (St. Louis, MO). A SpectraMax Gemini EM is used for all experiments utilizing a plate-reading fluorimeter (Molecular Devices, Sunnyvale, CA).

Antimicrobial Activity

The antimicrobial performances of NA-CATH, L-ATRA-1A and D-ATRA-1A are determined using a resurzarin-based assay (*46*, *47*). Frozen enumerated bacterial aliquots are thawed on ice and gently mixed. For each strain, bacteria are diluted to 2×10^6 CFU/mL in sterile 10 mM sodium phosphate (pH 7.4) solution and added in 50 µL aliquots to the wells of a black 96-well microtiter plate containing serially diluted peptide (50 µL), dissolved in the same phosphate buffer. In these assays control wells are prepared containing bacteria with no peptide. The microtiter plate is incubated for 3 hours at 37° C (*E. coli*) or 30° C (*B. cereus*). After three hours, 100 µL of PBS solution with dissolved resazurin and MHB is added to the wells of the microtiter plate. The amounts of resazurin and MHB used are dependent on bacterial strain, with the final resazurin/ MHB concentrations being 100 µM resazurin/ 0.2% (wt/vol) for *E. coli* and 12.5 µM

resazurin/ 0.05% (wt/vol) for *B. cereus*. Following addition of resazurin/MHB buffer, the plate is immediately placed in a plate-reading fluorimeter for incubation overnight at either 37° C (*E. coli*) or 30° C (*B. cereus*) while monitoring fluorescence for each well (530 nm_{ex}/ 590 nm_{em}). These antimicrobial measurements are performed in triplicate in order to provide statistical significance.

Fluorescence data for each well is compiled and the onset time of half maximal fluorescence ($T_{0.5}$) is determined. For both *E. coli* and *B. cereus*, standard curves are generated using serially diluted bacterial suspensions (~10^6 CFU/mL - 10^3 CFU/mL) in the absence of CAMPs. Observed $T_{0.5}$ values are plotted against initial CFU counts, which are verified by plating on MHB agar plates, and analysis of the data by linear regression produces Equation 2 for *E. coli* and Equation 3 for *B. cereus*.

Equation 2

$$\log (CFU/mL_{E.\ coli}) = -0.0002(T_{0.5}) + 9.3144$$

Equation 3

$$\log (CFU/mL_{B.\ cereus}) = -0.0002(T_{0.5}) + 6.4755$$

Equations 2 and 3 are then used to interpolate the surviving bacterial concentration (CFU/mL) based on $T_{0.5}$ values determine following incubation with varied peptide concentrations. These values can be used to calculate bacterial survival (%) relative to cultures incubated in the absence of peptide. The peptide concentration required to kill 50 % of the viable cells in the assay (EC50) is determined by plotting survival as a function of the log of peptide concentration (log mg/mL) and fitting the

data, using GraphPad Prism X5 (Graph-Pad Software, San Diego, CA) to Equation 4, which describes a sigmodial dose-response curve. Experiments against NA-CATH and all assays in high salt conditions performed by Carlos Rodriguez.

Equation 4

$$Y = \text{Bottom} + ((\text{Top}-\text{Bottom}) / (1 + 10^{[(\log EC50 - X) * \text{Hill Slope}]}))$$

Antimicrobial Kinetics

The antimicrobial kinetics of these peptides were determined at concentrations of 200, 2 and 0.2 µg/mL and at time intervals of 0.5, 2, 4, 10 and 20 minutes of NA-CATH, L-ATRA-1A and D-ATRA-1A are determined. In a 96-well plate, 55 µL of a 400 µg/mL peptide solution in 10 mM phosphate buffer (pH 7.4) is added to 5 wells (1 well for each time point) and then serially diluted by taking 5 µL of peptide into 50 µL of 10 mM phosphate buffer (pH 7.4). Frozen aliquots of either *E. coli* or *B. cereus* are thawed on ice and then diluted in10 mM phosphate buffer (pH 7.4) to a concentration of 1×10^5 CFU/mL. Using a multi-channel pipette, 50 µL of bacterial stock was added to each well containing peptide solution. Additionally, aliquots of bacterial stock are added to wells containing 50 µL of 10 mM phosphate buffer (pH 7.4) as a survival reference. The prepared plate is allowed to incubate at room temperature with samples being collected at 0.5, 2, 4, 10 and 20 minutes. Samples collected for each time point are diluted 10-fold by taking 50 µL from the 96-well plate into 450 µL of 1× PBS and then serially diluted using the same buffer. The surviving bacteria are plated in triplicate and incubated for 18

hours at 37° C (*E. coli*) or 30° C (*B. cereus*). The percentage of the bacterial killing relative to the positive control was determined for each peptide concentration at each time point. The percent killing for each peptide concentration was plotted as a function of time. The current protocol is also performed using 10 mM phosphate buffer (pH 7.4) containing 100 mM KCl.

Membrane Depolarization Assay

Cytoplasmic membrane depolarization was determined using the membrane potential-sensitive cyanine dye diSC_3-(5). In this method, frozen aliquots of enumerated bacteria (*E. coli* or *B. cereus*) are thawed on ice and washed 3 times with buffer (5 mM HEPES with 20 mM Glucose, pH 7.4). Following washing, the pelleted bacteria are re-suspended in HEPES buffer (5 mM HEPES, pH 7.4, 20 mM Glucose) containing either, 10 or 100 mM KCl. A 96-well plate was prepared where wells are charged with 360 µL of bacterial suspension (2×10^7 CFU/mL) and 4.19 µL of diSC_3-(5) (200nM) for a total volume of 364.19 µL. The bacteria are incubated at room temperature and fluorescence is monitored (622 nm_{ex}/ 670 nm_{em}) until diSC_3-(5) maximal uptake is obtained. Maximal diSC_3-(5) uptake is indicated by a baseline in fluorescence due to self-quenching as the dye concentrates in the cell membrane. Peptide (NA-CATH, L- or D-ATRA-1A) is added at varied concentrations (200 – 2 µg/mL) in 20 µL aliquots and the fluorescence increase due to induced depolarization of the cytoplasmic membrane is recorded. A negative control of bacteria and diSC_3-(5) is used as a background. As a positive control, complete collapse of the membrane potential is attained with Valinomycin, a potassium ionophore. Measurements are performed in triplicate for each condition and each peptide. The

peptide-induced fluorescence is baseline subtracted for each peptide and the maximal relative fluorescence units (RFU) are plotted as a function of peptide concentration.

SYTOX Green Uptake Assay

Induced membrane permeabilization caused by NA-CATH, L- or D-ATRA-1 is monitored via the fluorescence increase that occurs when the cationic cyanine dye SYTOX Green intercalates DNA. SYTOX Green is impermeant to living cells, yet can easily penetrate compromised membranes (*40*, *48–50*). In this method, frozen aliquots of enumerated bacteria (*E. coli* or *B. cereus*) are thawed on ice and washed 3 times with buffer (5 mM HEPES with 20 mM Glucose, pH 7.4). Following the final wash, the bacteria are pelleted and then re-suspended in 1 mL of HEPES buffer (5 mM HEPES with 20 mM Glucose, pH 7.4) containing either 10 or 100 mM KCl. Aliquots of re-suspended bacteria are further diluted with their respective buffers to a concentration of 4 $\times 10^7$ CFU/mL in 1mL. The diluted cells are then charged with 1µL of 5 mM SYTOX Green and incubated for 15 minutes in the dark at room temperature. Following incubation, bacteria-SYTOX suspension is added in 100 µL aliquots to the wells of a 96-well plate and the fluorescence of each well monitored in order to establish baseline fluorescence. After 5 minutes of baseline fluorescence collection, 100 µL aliquots of peptide solutions, with concentrations ranging from 200 – 2 µg/mL, are added to each well and the increase in SYTOX Green fluorescence is measured (485 nm_{ex}/ 520 nm_{em}) for 40 minutes. As a negative control, fluorescence data was collected for bacteria suspended in buffer containing SYTOX Green in the absence of peptide. As a positive control, meletin is added to bacterial cells to achieve complete lysis and maximum

fluorescence. Measurements are performed in triplicate for each condition and each peptide. The peptide-induced fluorescence is baseline subtracted for each peptide and the maximal relative fluorescence units (RFU) are plotted as a function of peptide concentration.

Scanning Electron Microscopy

Enumerated frozen aliquots of either *E. coli* or *B. cereus* are thawed on ice and then diluted to a concentration of 2×10^8 CFU/mL with 10 mM phosphate buffer, pH 7.4. Stock solutions are prepared for NA-CATH, L-ATRA-1A and D-ATRA-1A (100 µg/mL) in 10 mM phosphate buffer, pH 7.4. Bacterial suspension and peptide solutions (50 µL each) are added to the wells of a 96-well plate and incubated at room temperature for 20 minutes. Controls consisting of bacteria alone suspended in buffer, as well as, peptide alone suspended in buffer are similarly prepared. Following incubation, bacterial-peptide and control solutions are filtered on to a 0.22 µm membrane filter that have been pretreated with 0.1% poly-L-lysine in order to improve cell adhesion to the filter (*51*). The retentates on the membrane surface are fixed for 2 hours with 2.5% glutaraldehyde, dehydrated with graded ethanol series and then critical point dried. An 8 nm layer of Au/Pd alloy is sputtered on the samples to avoid charging in the microscope. Samples are imaged in high resolution mode with an upper detector using a Hitachi-4700 FESEM at an accelerating voltage of 10 keV.

Results and Discussion

Antimicrobial Effectiveness and Kinetics

Ionic strength in media has historically been shown to impact the antimicrobial activity of CAMPs (*52–55*). For many CAMPs, high salt conditions impedes their ability

to kill bacteria, however the degree to which the peptides are affected can vary significantly. Here, the antimicrobial properties and bactericidal kinetics of full-length NA-CATH and the ATRA-1A isomers have been assessed against representative Gram-negative and Gram-positive bacteria, *E. coli* and *B. cereus* under low and high salt conditions.

In order to assess the degree to which salt conditions impact their antimicrobial effectiveness, the performances of NA-CATH, L-ATRA-1A and D-ATRA-1A were evaluated against *E. coli*, a model Gram-negative bacterium, and *B. cereus*, a model Gram-positive bacterium, under both low (10 mM phosphate buffer, pH 7.4) and high (100 mM KCl in 10 mM phosphate buffer, pH 7.4) salt conditions. The results of these assays indicate that the half-maximal effective concentration (EC50) values for the peptides are dependent on both the nature of the bacteria being tested and whether assays were performed in low or high salt conditions (Table 5).

Table 5: Antimicrobial Performance of NA-CATH and ATRA-1A Isomers under High and Low Salt Conditions.

Antimicrobial activity is expressed in terms of EC50 (μg/mL) values and corresponding 95% confidence interval ranges for *E. coli* and *B. cereus.*

*Values were obtained using plating assay and have been previously published (13). Data for NA-CATH and all high salt experiments courtesy of Carlos Rodriguez.

Bacterium	Antimicrobial Activity (EC50, μg/mL)											
	NA-CATH				L-ATRA-1A				D-ATRA-1A			
	High Salt	95% CI Range	Low Salt	95% CI Range	High Salt	95% CI Range	Low Salt	95% CI Range	High Salt	95% CI Range	Low Salt	95% CI Range
E.coli	0.024	0.015 to 0.040	0.023	0.022 to 0.025	~9.7	very wide	4.3*	4.0 to 4.6	~7.1	very wide	1.4*	1.1 to 1.5
B.cereus	0.60	0.49 to 0.72	0.35	0.33 to 0.40	N/A	N/A	73*	65 to 82	4.3	3.8 to 4.9	2.3*	2.1 to 2.6

Against the Gram-negative bacterium, *E. coli*, the EC50 values in low salt for NA-CATH, L-ATRA-1A and D-ATRA-1A were 0.023 µg/mL, 4.3 µg/mL and 1.4 µg/mL, respectively. In high salt conditions NA-CATH had an EC50 of 0.024 µg/mL, while L-ATRA-1A and D-ATRA-1A had EC50s of ~9.7 µg/mL and ~7.1µg/mL, respectively. When tested against *B. cereus*, a Gram-positive bacterium, the EC50 values for low salt conditions were found to be 0.35 µg/mL, 72.9 µg/mL, and 2.3 µg/mL for NA-CATH, L-ATRA-1A and D-ATRA-1A, respectively. In high salt conditions, the EC50 value found for NA-CATH was 0.60 µg/mL and that of D-ATRA-1A was 4.3 µg/mL. L-ATRA-1A was completely ineffective against *B. cereus* under these conditions.

In addition to antimicrobial potency, the rate with which the peptides exerted their bactericidal effect was evaluated under both low and high salt conditions (Figure 9 (*E. coli*) and Figure 10 (*B. cereus*)).

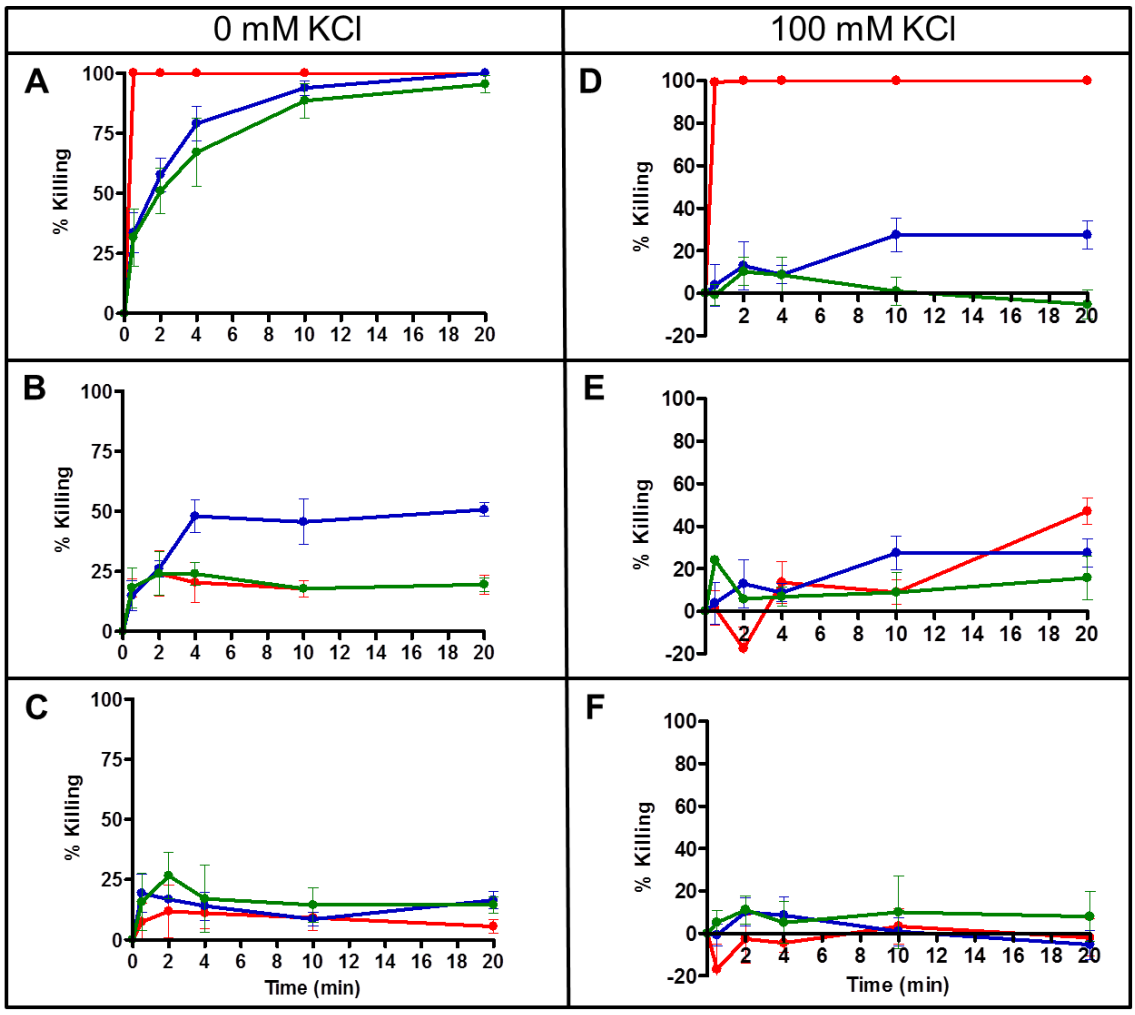

Figure 9: Antimicrobial Kinetics of NA-CATH and the ATRA-1A Isomers against *E. coli*.
The killing kinetics of NA-CATH (red), L-ATRA-1A (green) and D-ATRA-1A (blue) were evaluated against *E. coli* in low salt conditions and high salt conditions at peptide concentrations of 200 µg/mL (A and D), 2 µg/mL (B and E) and 0.2 µg/mL (C and F).

Here, initial killing kinetics was established by monitoring bacterial survival as a function of time for the first 20 minutes following introduction of peptide to the bacterial culture. Under low salt conditions, the peptides exhibited differences in their antimicrobial kinetics against both bacteria. At a concentration of 200 µg/mL, full-length

NA-CATH achieved complete killing in less than 30 seconds against both *E. coli* (Figure 9A) *B. cereus* (Figure 10A)., After 20 minutes, NA-CATH displayed ~ 25 % killing at a concentration of 2 µg/mL and ~ 5 % killing at 0.2 µg/mL for both *E. coli* (Figure 9; B and C) and *B. cereus* (Figure 10; B and C). Under the same conditions, L-ATRA-1A exhibited 100 % killing against *E. coli* at 200 µg/mL, while presenting ~ 20 % killing at both 2 µg/mL and 0.2 µg/mL (Figure 9; A-C) after 20 minutes. Against *B. cereus*, ~ 90 % killing was observed for L-ATRA-1A at 200 µg/mL after 20 minutes, while ~ 20 % killing was achieved at 2 µg/mL and ~ 10 % killing at 0.2 µg/mL in the same timeframe (Figure 10; A-C). While, D-ATRA-1A exhibited 100 % killing at 200 µg/mL, ~ 50 % killing at 2 µg/mL and ~ 20 % killing at 0.2 µg/mL (Figure 9; A-C) against *E. coli* after 20 minutes. Against the Gram-positive bacterium *B. cereus*, D-ATRA-1A exhibited 100 % and ~ 50 % killing after 20 minutes at peptide concentrations of 200 µg/mL and 2 µg/mL, respectively (Figure 10; A and B), however diminished killing (~ 5 %) occurred at 0.2 µg/mL (Figure 10C).

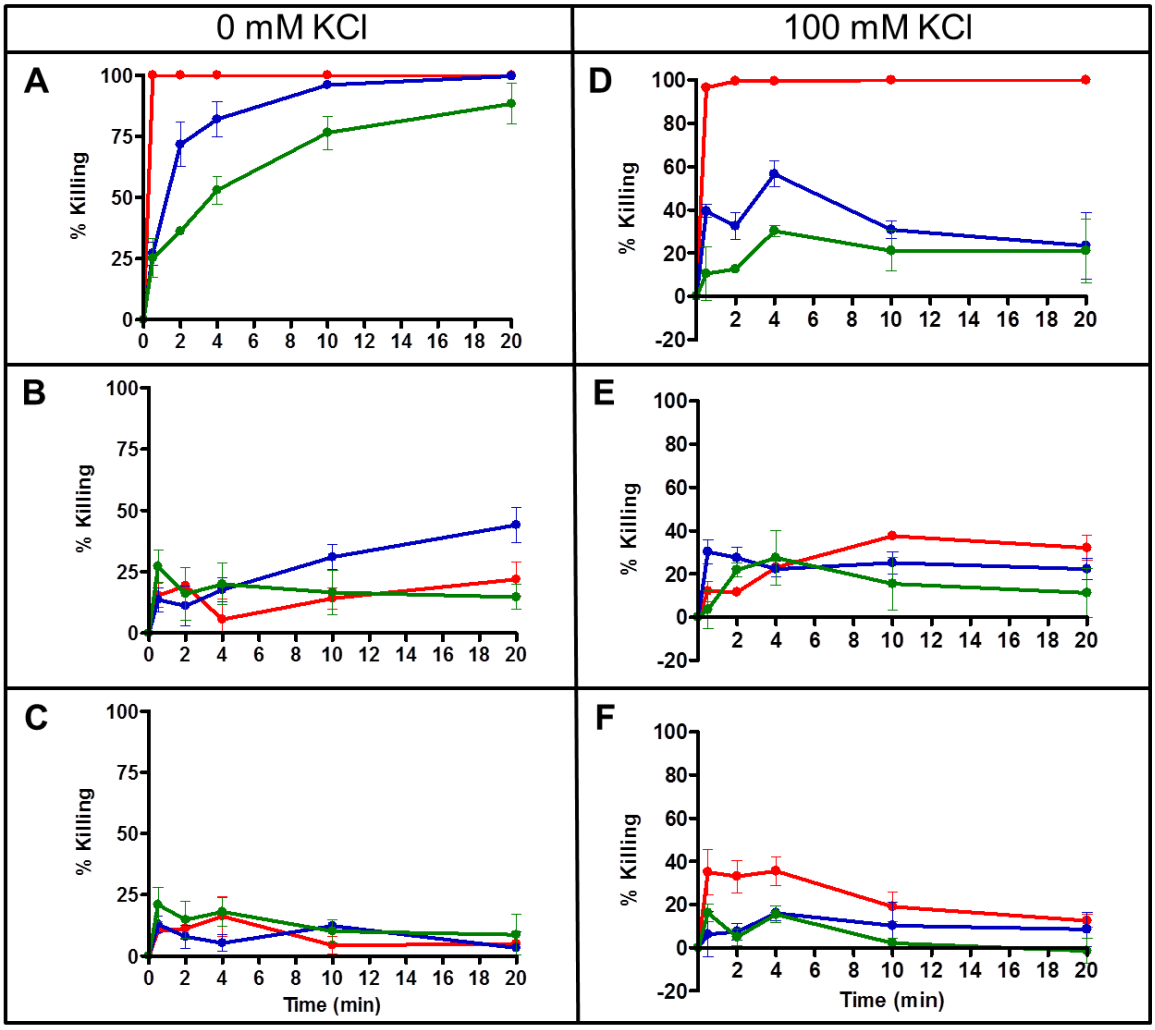

Figure 10: Antimicrobial Kinetics of NA-CATH and the ATRA-1A Isomers against *B. cereus*.
The killing kinetics of NA-CATH (red), L-ATRA-1A (green) and D-ATRA-1A (blue) were evaluated against *B. cereus* in low salt conditions and high salt conditions at peptide concentrations of 200 µg/mL (A and D), 2 µg/mL (B and E) and 0.2 µg/mL (C and F).

Under high salt conditions, the peptides exhibit changes in their killing kinetics against both *E. coli* (Figure 9D-F) and *B. cereus* (Figure 10D-F). At a concentration of 200 µg/mL, NA-CATH kills 100 % of bacteria within 30 seconds for both *E. coli* and *B. cereus*, which is consistent with the kinetics observed under low salt conditions (Figure

9D and 10D). At lower concentrations, the kinetics are similar to those observed in low salt, with ~ 40 % killing occurring after 20 minutes at a peptide concentration of 2 µg/mL and ~ 0 % killing at 0.2 µg/mL of peptide for *E. coli* (Figure 9E and F), and ~ 30 % killing at 2 µg/mL of peptide and ~ 10 % killing at 0.2 µg/mL for *B. cereus* (Figure 10E and F). The killing kinetics for L-ATRA-1A against both *E. coli* and *B. cereus* were significantly reduced under high salt conditions relative to the killing kinetics observed in low salt medium. Against *E. coli*, L-ATRA-1A exhibited ~ 25 %, ~ 20 %, and ~ 10 % killing after 20 minutes at peptide concentrations of 200 µg/mL, 2 µg/mL and 0.2 µg/mL, respectively (Figure 9D-F). After 20 minute incubation with *B. cereus*, L-ATRA-1A provided killing of ~ 25 % at 200 µg/mL of peptide, ~ 15 % at 2 µg/mL and ~ 0 % at 0.2 µg/mL (Figure 10D-F). Against *E. coli*, D-ATRA-1A exhibited slower killing kinetics in high salt relative to that observed in low salt conditions. Incubation of D-ATRA-1A with *E. coli* for 20 minutes realized killing of ~ 50 %, ~ 25 % and ~ 0 % at peptide concentrations of 200 µg/mL, 2 µg/mL and 0.2 µg/mL, respectively (Figure 9D-F). Against *B. cereus*, D-ATRA-1A killed ~ 25 % at concentrations of 200 µg/mL and 2 µg/mL, while at 0.2 µg/mL ~ 10 % killing was achieved (Figure 10D-F).

Correlating the EC50 values and killing kinetics of each peptide can provide insights into their modes of action. Against *E. coli*, NA-CATH is unable to achieve 50 % killing at peptide concentrations 10 and 100 times its EC50 in low salt buffer (0.023 µg/mL) or in high salt conditions (0.024 µg/mL). Similar results were exhibited by NA-CATH against the Gram-positive bacterium *B. cereus*, where at ~ 6 times the EC50 value noted in low salt conditions (0.35 µg/mL) and ~ 3 times the EC50 value from high salt

(0.60 μg/mL), half-maximal effectiveness was not achieved. The kinetics of antibacterial effectiveness have been observed for only the initial 20 minutes following addition of peptide, and the data indicate that NA-CATH's mechanism of action requires incubation for longer than 20 minutes in order to exert its full bactericidal effectiveness. Under low-salt conditions the killing kinetics of L-ATRA-1A show a quarter of the population surviving at half the EC50 against *E. coli* (4.1 μg/mL) after 20 minutes and only reaching above 50 % survival at 3 times the EC50 value against *B. cereus* (72 μg/mL) in the same time period. In high salt conditions, the killing kinetics for L-ATRA-1A showed the peptide had very limited antimicrobial activity against *B. cereus* at any concentration. Against *E. coli*, L-ATRA-1A failed to attain 50 % killing even at concentrations 20 times its EC50 (9.7 μg/mL). In contrast, D-ATRA-1A's killing kinetics showed 50 % killing against *E. coli* at a peptide concentration under low salt conditions within 20 minutes equal to its EC50 in the same conditions (1.4 μg/mL). While in high salt conditions, the D-peptide failed to attain 50 % killing inside of the observed 20 minute window even at a peptide concentration 28 times its EC50 (7.1 μg/mL) under the same conditions.

The decreases in killing activity within the 20 minute observation window of the kinetics studies for the peptides in high salt conditions are most likely due to direct interaction of the salt ions with the peptide's hydrophilic residues and the lipid-bilayer head groups (*54*). Hofmeister effects, ionic screening and specific ion binding between the solution ions and both peptide side-chains and lipid head groups all likely contribute to the observed reductions in antimicrobial effectiveness (*56–59*). For example, these interactions could result in a reduction in CAMP helicity, and thus amphipathic character,

which would greatly affect their ability to directly interact with bacterial membranes and reduce or negate their antimicrobial effectiveness.

Bacterial Membrane Disruption

Cationic antimicrobial peptides in general have been shown to be capable of causing varying degrees of disruption in bacterial membranes, ranging from transient gaps, large pores and micellization. Accordingly, the extent to which NA-CATH and the ATRA-1A isomers interact with bacterial membranes is ascertained using assays designed to monitor membrane depolarization and permeabilization. Depolarization of bacterial membranes is evaluated using diSC_3-(5), a membrane potential sensitive dye, which concentrates within the lipid bilayer resulting in the dye self-quenching (*40, 50, 55, 60, 61*). If peptides depolarize the membrane, the potential dissipates, and diSC_3-(5) is released into solution causing an increase in fluorescence, which is directly proportional to the degree of membrane potential reduction. Uptake of SYTOX Green, a nucleic acid stain, is used to detect greater degrees of peptide-induced membrane permeabilization. SYTOX Green is impermeant to the membranes of healthy cells, but can penetrate disrupted cell membranes and bind to nucleic acids (*49*). If the bacterial cell membrane integrity is compromised by pores large enough to allow passage or cell lysis, influx of the dye and subsequent binding to DNA causes a > 500-fold increase in fluorescence (*40, 48–50*). Scanning electron microscopy SEM allows for visualization of morphological changes in bacterial membranes (*48, 62–64*). Therefore, SEM is used to probe the ability of the CAMPs to induce gross changes in membrane morphology, including blebbing, aggregation, pore formation and cell lysis.

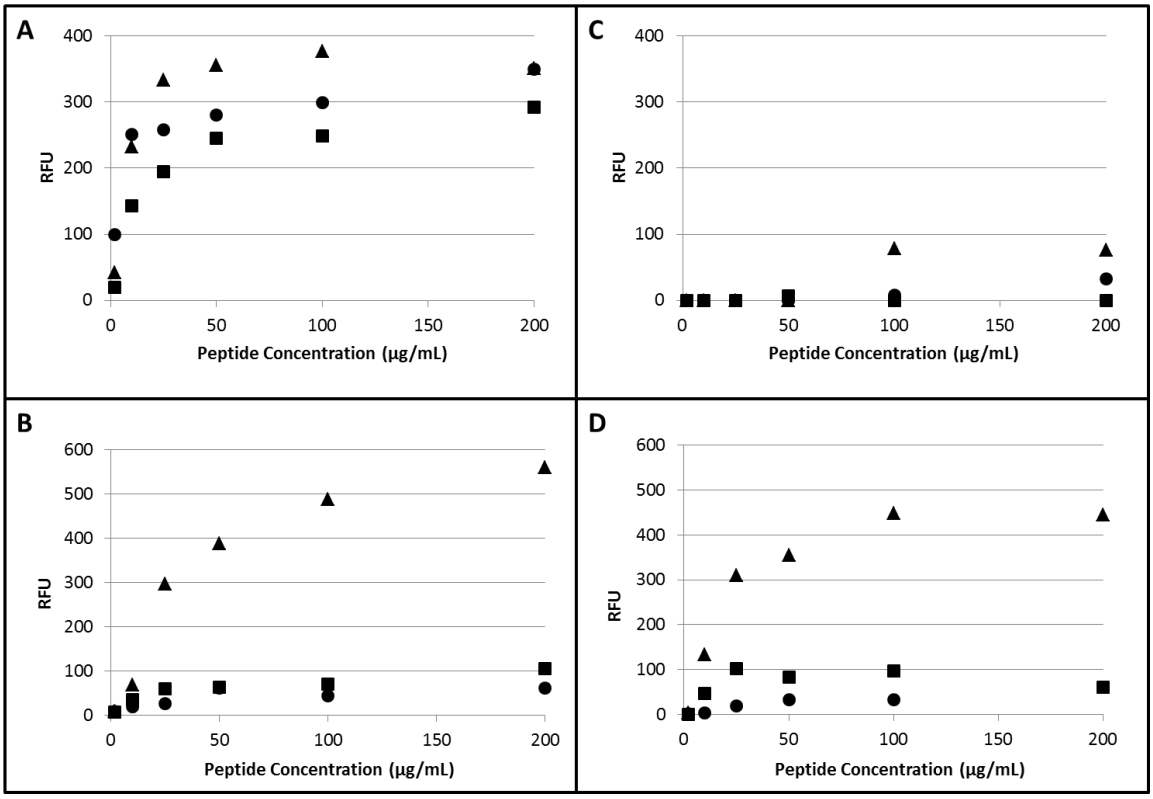

Figure 11: Effect of Peptide Concentration on Membrane Depolarization.
Using diSC_3-(5) fluorescence to monitor the effect of peptide concentration on induced membrane depolarization for NA-CATH (▲), L-ATRA-1A (●) and D-ATRA-1A (■) in low salt conditions against *E. coli* (A) and *B. cereus* (B) and high salt conditions *E. coli* (C) and *B. cereus* (D).

The results of the diSC_3-(5) depolarization studies indicate that all three peptides differentially dissipate membrane potential for *E. coli* and *B. cereus* depending on peptide concentration and salt conditions. In low salt conditions NA-CATH is able to depolarize both *E. coli* (Figure 11A) and *B. cereus* (Figure 11B) cells at the lowest concentration of 2 µg/mL with depolarization increasing as the peptide concentration increases. In high salt conditions against *E. coli* cells by NA-CATH depolarization is not detected below

100 µg/mL, with depolarization at 100 and 200 µg/mL being considerably lower than those seen in low salt conditions (Figure 11C). In contrast, NA-CATH causes depolarization of *B. cereus* cells at peptide concentration of 2 µg/mL under high salt conditions with depolarization increasing as the peptide concentration increases up to 100 µg/mL (Figure 11D). However, depolarization seems to remain unchanged as the peptide concentration rises from 100 µg/mL to 200 µg/mL. Under low salt conditions, both L- and D-ATRA-1A exhibit slight depolarization of *E. coli* (Figure 11A) at peptide concentrations ranging from 50 µg/mL and 200 µg/mL, yet depolarization of *B. cereus* cells is observed at a peptide concentration as low as 2 µg/mL, increasing as the peptide concentration rises to 200 µg/mL (Figure 11B). In high salt conditions, L-ATRA-1A exhibits no depolarization against *E. coli* (Figure 11C) and very little depolarization against *B. cereus* (Figure 11D). D-ATRA-1A in high salt conditions similarly shows no depolarization against *E. coli* (Figure 11C), however exhibits increased depolarization against *B. cereus* compared to L-ATRA-1A (Figure 11D).

All three peptides exhibit higher degrees of depolarization in a low salt environment than under high salt conditions. In low salt conditions, L-ATRA-1A and D-ATRA-1A display higher degrees of depolarization with *B. cereus* than with *E. coli*, while NA-CATH shows similar depolarization with both bacterial strains. In high salt conditions, all three peptides show effective depolarization with *B. cereus*, while only NA-CATH exhibits depolarization with *E. coli*. Membrane depolarization data for the three peptides is consistent with their antimicrobial effectiveness and killing kinetics.

To further investigate the extent of membrane disruption induced by NA-CATH and the ATRA-1A isomers, membrane permeabilization is detected by monitoring SYTOX Green fluorescence. In low salt conditions, NA-CATH is able to permeabilize the membranes of both *E. coli* (Figure 12A) and *B. cereus* (Figure 12B) at peptide concentrations as low as 12.5 μg/mL. Similar results are seen in high salt conditions for NA-CATH and its ability to permeabilize the membranes of both *E. coli* (Figure 12C) and *B. cereus* (Figure 12D). In contrast, L-ATRA-1A is able to only slightly permeabilize the membranes of *E. coli* (Figure 12A) and *B. cereus* (Figure 12B) in low salt conditions. However, L-ATRA-1A is unable to permeabilize either bacterium in high salt conditions (Figure 12C-D). Similarly, D-ATRA-1A is able to slightly permeabilize the membranes of *E. coli* (Figure 12A) and *B. cereus* (Figure 12B) in low salt conditions, yet in high salt conditions no permeabilization is observed for either bacterium (Figure 12C-D).

In the SYTOX Green uptake studies, NA-CATH induces membrane disruption in both bacterial strains in both high and low salt environments, which is consistent with the peptide's performance in diSC_3-(5) depolarization studies. However, L-ATRA-1A and D-ATRA-1A display only slight SYTOX fluorescence, even under conditions where they exhibit significant depolarization.

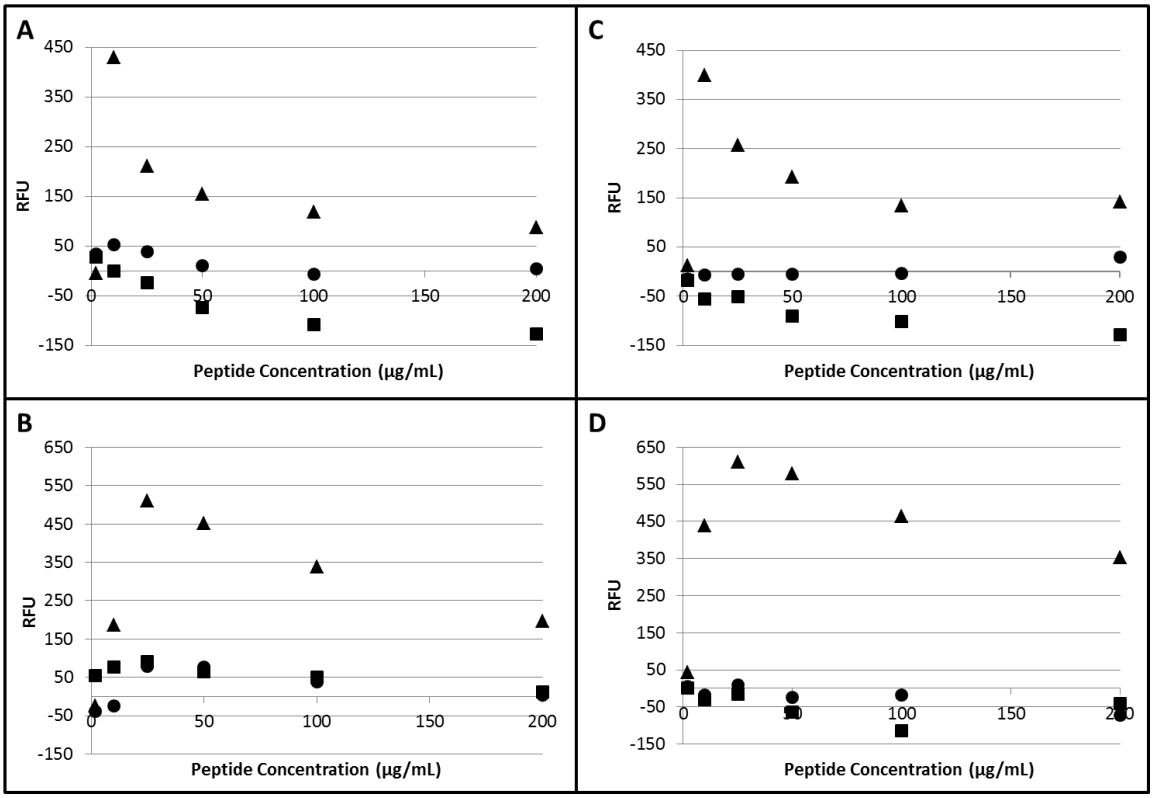

Figure 12: Effect of Peptide Concentration on Membrane Permeabilization. Evaluation of peptide-induced membrane permeabilization based on SYTOX Green fluorescence using varied concentrations of NA-CATH (▲), L-ATRA-1A (●) and D-ATRA-1A (■) in low salt conditions against *E. coli* (A) and *B. cereus* (B) and high salt conditions *E. coli* (C) and *B. cereus* (D).

Greater understanding of the ability of CAMPs to inflict substantial changes in bacterial membranes can be achieved by using SEM to visualize morphological changes in the membranes of bacteria that have been treated with NA-CATH and the ATRA-1A isomers. The exposure of *E. coli* and *B. cereus* to 50 µg/mL of peptide for 20 min causes notable alterations in cell morphology. When untreated *E. coli* cells are prepared in low ionic strength buffer the SEM images show intact cells with a corrugated morphology, typical of this strain (Figure 13I; A and B) (*62*). However, *E. coli* exposed to peptide

present many morphological abnormalities. After incubation with L-ATRA-1A, the morphology of the *E. coli* cells change drastically, displaying blebbing on the cell surface and some cells exhibiting intracellular leakage (Figure 13I; C and D). Similar morphological changes are seen in *E. coli* treated with D-ATRA-1A, however the extent of surface blebbing significantly increases (Figure 13I; E and F). *E. coli* exposed to NA-CATH exhibit large blebs along the surface, cellular leakage, as well as, cell lysis (Figure 13I; G and H). Untreated *B. cereus* prepared in low ionic strength buffer in the absence of peptide appear normal, exhibiting smooth cell morphology and minimal aggregation (Figure 13II, A and B). Minor morphological changes are observed in *B. cereus* exposed to L-ATRA-1A with roughening of the cell surface and micro-blebbing. Additionally, these cells exhibit increased aggregation (Figure 13II; C and D). Treatment of *B. cereus* with D-ATRA-1A results in cellular aggregation throughout the entire sample, as well as, roughening and micro-blebbing of the membrane. Pore formation is observed in some cells (Figure 13II; E and F). *B. cereus* cells exhibit the largest extent of aggregation, along with blebbing along the surface, cellular leakage, as well as, pore formation leading to cell lysis following exposure to NA-CATH (Figure 13II; G and H).

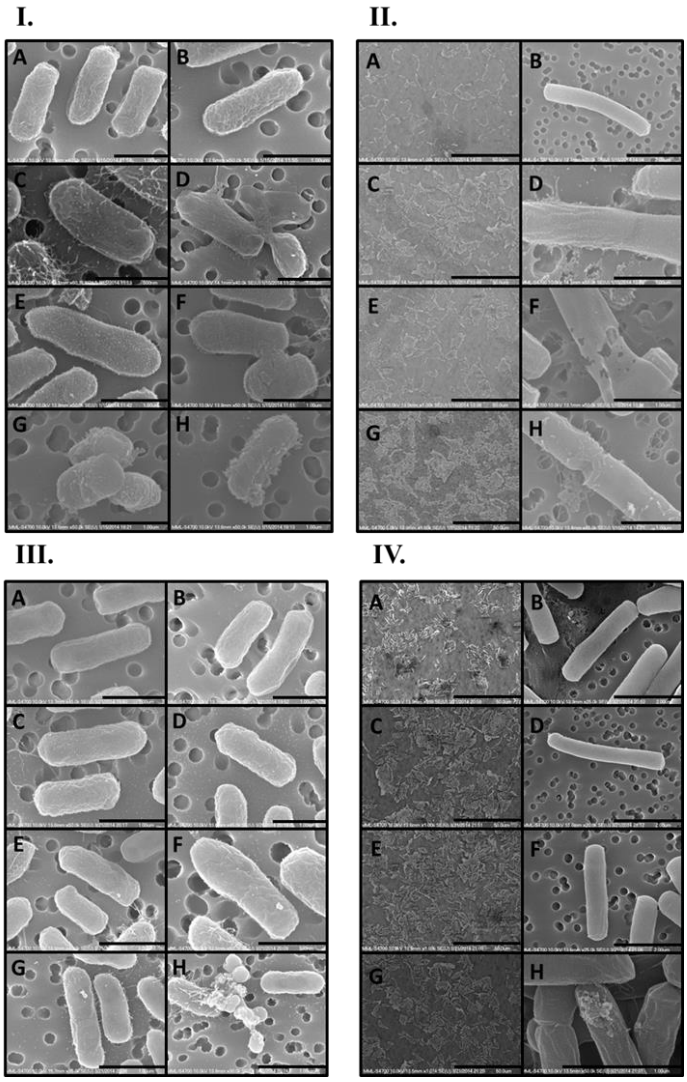

Figure 13: SEM Micrographs of Bacteria Following Peptide Exposure.
I. *E. coli* untreated under low-salt conditions (A and B), and following incubation with L-ATRA-1A (C and D), D-ATRA-1A (E and F) and NA-CATH (G and H). In **I** A-H the scale bar is equal to 1µm. **II**. *B. cereus* untreated under low-salt conditions (A and B), and following incubation with L-ATRA-1A (C and D), D-ATRA-1A (E and F) and NA-CATH (G and H). In **II** the scale bar is equal to 50 µm (A, C, E, and G), 2 µm (B) and 1 µm (D, F, and H). **III**. *E. coli* untreated under high-salt conditions (A and B), and following incubation with L-ATRA-1A (C and D), D-ATRA-1A (E and F) and NA-CATH (G and H). In **III** A-H the scale bar is equal to 1µm. **IV**. *B. cereus* untreated under high-salt conditions (A and B), and following incubation with L-ATRA-1A (C and D), D-ATRA-1A (E and F) and NA-CATH (G and H). In **IV** the scale bar is equal to 50 µm (A, C, E, and G) and 2 µm (B, D, F, and H).

Under high salt conditions, the peptides exhibit changes in the observed morphology changes in both *E. coli* (Figure 13 III) and *B. cereus* (Figure 13 IV). When untreated *E. coli* cells are prepared in high ionic strength buffer the SEM images show intact cells with a smoother surface then seen in low salt buffer (Figure 13 III; A and B). L-ATRA-1A and D-ATRA-1A exhibit similar morphological changes when incubated with *E. coli* cells, both displaying roughening and blebbing on the cell surface (Figure 13 III; C, D, E and F). In high salt conditions, *E. coli* exposed to NA-CATH is still able to exhibit blebbing along the surface, cellular leakage, as well as, cell lysis (Figure 13 III; G and H). Untreated *B. cereus* prepared in high salt exhibit smooth cell morphology and increase aggregation from that seen in low salt buffer (Figure 13 IV, A and B). No morphological changes are observed in *B. cereus* exposed to L-ATRA-1A or D-ATRA-1A, however these cells do exhibit increased aggregation (Figure 13 IV; C, D, E and F). Treatment of *B. cereus* with NA-CATH results in cellular aggregation, as well as, roughening and micro-blebbing along the surface and cell lysis (Figure 13 IV; G and H).

For both *E. coli* and *B. cereus*, the morphological changes caused by NA-CATH are more severe than those seen with either of the ATRA-1A isomers in both low and high salt conditions. Additionally, the ATRA-1A isomers appear to have uniform effect across the entire bacterial populations in both *E. coli* and *B. cereus*. However, NA-CATH is less consistent in its affect, with some cells manifesting severe detrimental morphological changes and others appearing less affected. The diSC_3-(5) and SYTOX Green experiments reflect bulk properties and do not distinguish the effect the peptides have on individual cells. By contrast, SEM allows examination of individual cells within

the population. Inconsistency between peptide affects by NA-CATH versus the ATRA-1A isomers would be missed without SEM data and is critical to understanding these peptides' mechanisms. While L-ATRA-1A shows significant depolarization, it does not show permeabilization with SYTOX Green. However, the SEM data shows mainly roughening and blebbing of the cell membranes, suggesting these morphological changes are not indicative of permeabilization. In the case of D-ATRA-1A, depolarization, slight permeabilization and greater degrees of blebbing with observed pore formations are exhibited, suggesting these slight increases in gross morphological changes could be linked with permeabilization. While NA-CATH exhibits strong depolarization and permeabilization in the diSC_3-(5) and SYTOX Green assays, the SEM data reveals that the peptide has a more dramatic, but less consistent impact on cell morphology.

Conclusion

While the exact mechanisms of action employed by these peptides cannot be extracted solely from the data reported here, it does provide insights into the means by which they exert their antimicrobial effects and suggests differences exist in the ways that each peptide interacts with bacterial membranes. In these studies, NA-CATH exhibits a very high potency in both high and low salt environments with the ability to rapidly cause bacterial membrane disruption at low peptide concentrations, which is consistent with the cooperative formation of pores or large gaps leading to cell death. In the case of the shorter ATRA-1A isomers, D-ATRA-1A appears to be more potent than its L-counterpart, however the data reported here suggests that their interactions with bacterial membranes are similar. Unlike full-length NA-CATH, the ATRA-1A isomers associate

and concentrate on the outer surface of the bacterial membrane, causing small transient gaps and the loss of critical membrane potential, intracellular leakage and cell death. These results are consistent with our earlier observation that exposure to NA-CATH resulted in the complete lysis of anionic liposomes, while incubation with the ATRA-1A peptides appeared to cause leakage, fusion and aggregation of the liposomes (*45*). The results of the studies reported here suggests NA-CATH and the ATRA-1A peptides employ different mechanisms of action, with that employed by NA-CATH resembling pore formation and those employed by the ATRA-1A isomers being more consistent with a carpet model. The differences observed in the behaviors of the full-length parent peptide, NA-CATH, and the truncated ATRA-1A isomers demonstrate how altering the length and charge of an antimicrobial peptide can dramatically impact the antimicrobial effectiveness and the mechanism of action employed by the peptide. Such mechanistic factors must be considered in future efforts to identify minimal antimicrobial units within larger naturally occurring peptides.

CHAPTER THREE. A BIOPROSPETCTING APPROACH TO ANTIMICROBIAL PEPTIDE DISCOVERY

Cationic antimicrobial peptides and their therapeutic potential have garnered growing interest because of the proliferation of bacterial resistance. However, the discovery of new antimicrobial peptides has proven challenging due to the limitations associated with conventional biochemical purification and difficulties in predicting active peptides from genomic sequences, if known. We have developed a novel approach for the discovery of new antimicrobial peptides, one that capitalizes on their fundamental and conserved physico-chemical properties. This robust, sample-agnostic process employs functionalized hydrogel microparticles to harvest cationic antimicrobial peptides from biological samples, followed by *de novo* sequencing of captured peptides, eliminating the need to isolate individual peptides. Based on their net charges and hydrophobicities, select peptide sequences are chemically synthesized, and their antibacterial properties assessed. Here, we report the implementation of this process to identify multiple novel antibacterial peptides from Alligator mississippiensis plasma.

Introduction

There has been a growing interest in cationic antimicrobial peptides (CAMPs) as a potential source of new therapeutics to address the growing problem of bacterial antibiotic resistance (*65*, *66*). Nature provides a prescreened library of peptides that has

been selected over millions of years of evolution for their ability to defend against infection under physiological conditions. The American alligator (*Alligator mississippiensis*) and other crocodilians are evolutionarily ancient animals whose plasma and leukocyte extracts have been shown to exhibit potent antimicrobial activity (*67–69*). This antimicrobial potency is likely to be attributable at least in part to the presence of CAMPs in the plasma and extracts. These peptides have been shown to be capable of exerting an antimicrobial effect, and they figure prominently in innate immunity of vertebrates and other higher organisms.

The discovery and identification of novel CAMPs has proven challenging using conventional proteomics tools. Methods used to fractionate and isolate peptides are labor-intensive, can result in sample and activity loss, and are unable to detect low-abundance peptides. To address these limitations, prior efforts to identify crocodilian antimicrobial peptides have resorted to using very large sample volumes (*69, 70*), which can be problematic if the animals are endangered or sample size is limiting. Further complicating matters, the high sequence and structural diversity of CAMPs presents an impediment to traditional mass spectrometry methods, which employ proteolytic digestion and database searches to facilitate peptide sequence determination. Subjecting samples to proteolytic digestion in this manner destroys information regarding the original native, intact peptide sequences. To overcome these challenges, Dr. Bishop and Dr. van Hoek have employed a multidisciplinary strategy that draws from protein biophysics, peptide chemistry, nanomaterials, advanced mass spectrometry techniques, and microbiology.

We report here development of a novel and versatile bioprospecting approach to antimicrobial peptide discovery, which builds upon recent advances in proteomics and biomarker discovery (Figure 14) (*71*, *72*). It utilizes a novel approach for extracting peptides, including CAMPs, from very small sample volumes (e.g. 100 µL) followed by analysis of the harvested peptides using advanced mass spectrometry techniques (*de novo* peptide sequencing) to identify CAMPs that may be present.

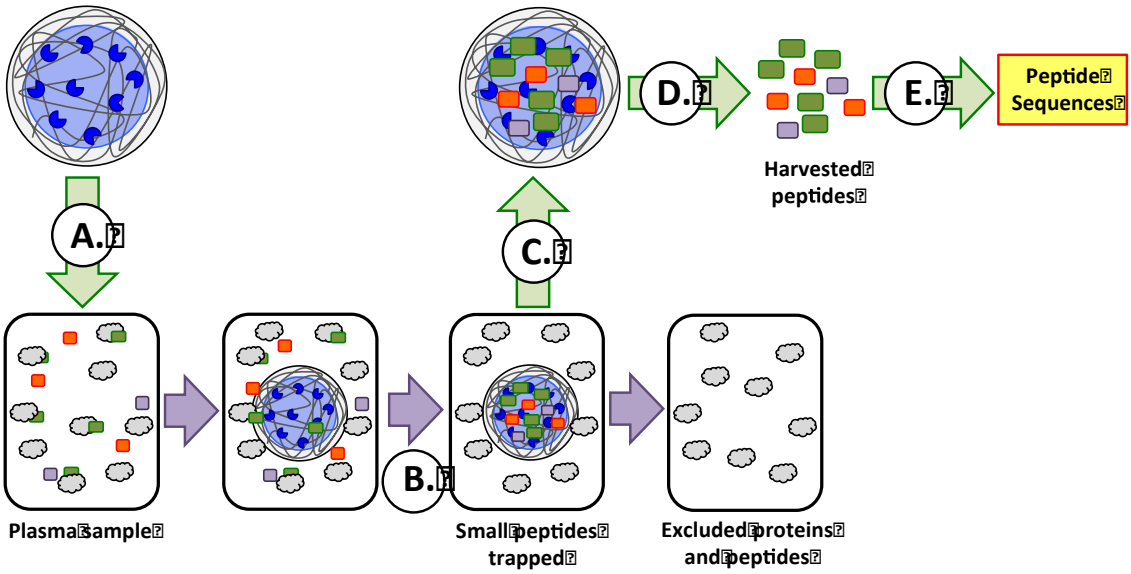

Figure 14: Bioprospecting Approach to CAMP discovery.
(A) Hydrogel microparticles are introduced into the plasma sample, and (B) the particles capture small cationic peptides present in the sample, while excluding high molecular weight proteins. (C) The particles are then recovered, (D) captured low molecular weight peptides are eluted from the particles and (E) analyzed by high-resolution MS/MS. Figure courtesy of Dr. Barney Bishop (George Mason University).

The process designed by Dr. Bishop and Dr. van Hoek employs custom-made functionalized hydrogel microparticles to harvest CAMPs in their native form from

biological samples, agnostic to source, based on their physico-chemical properties. Mass spectral analysis of the harvested intact peptides using an Orbitrap Elite mass spectrometer (Thermo Scientific) equipped with electron transfer dissociation (ETD) is used to determine their sequences in a *de novo* manner. The sequences are compared to available genomic and proteomic information in order to confirm, complete and correct the *de novo* peptide sequences. Additionally, all sequences are ultimately manually verified, especially those for which no genomic information is available. From these peptide sequences, likely CAMPs are predicted using a combination of rational analysis and web-based CAMP predictor algorithms (*73–75*). High probability CAMP candidates are then synthesized and evaluated for activity. We have applied this process to plasma from the American alligator, leading to the identification of five novel peptides that exhibit antimicrobial activity, $APOC1_{64\text{-}88}$, $APOC1_{67\text{-}89}$, $A1P_{394\text{-}428}$, $FGG_{398\text{-}413}$ and $FGG_{401\text{-}413}$.

Materials and Methods

The peptides used in these studies were custom synthesized by ChinaPeptides Company (Shanghai, China) and had purities of ≥ 95 %, based on chromatographic analysis of the purified peptides. Synthetic peptides were verified on a Thermo LTQ mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA). The bacterial strains of *Escherichia coli* (ATCC 25922), *Bacillus cereus* (ATCC 11778), *Pseudomonas aeruginosa* (ATCC 19429), and *Staphylococcus aureus* (ATCC 25923) used in these studies were purchased from the American Type Culture Collection (Manassas, VA). Resazurin, sodium salt is purchased from Sigma-Aldrich (St. Louis, MO). N-

Isopropylacrylamide (NIPAm), N, N′-Methylenebisacrylamide (BIS), Acrylic acid (AAc), 2-Acrylamido-2-methylpropane sulfonic acid (AMPS), Methyl Acrylate (MA), Lithium hydroxide (LiOH) and potassium persulfate (KPS) are all purchased from Sigma-Aldrich (St. Louis, MO). Mueller Hinton Broth (MHB) was purchased from Becton Dickinson and Company (Sparks, MD). Phosphate buffered saline (PBS) was purchased from Corning-cellgro (Manassas, VA) Alligator blood was acquired from St. Augustine's Alligator Farm (St. Augustine, FL). All protocols involving the alligators were approved by the GMU IACUC.

Particle Synthesis

The p-NIPAm-based particles are synthesized using one-pot free radical precipitation polymerization following previously published protocols (*71*). Particles incorporating AAc and AMPS are synthesized as follows: NIPAm (2.98 g, 26.28 mmol), BIS (111.4 mg, 0.72 mmol), AAc (370 µL, 5.4 mmol), and AMPS (746.1 mg, 3.69 mmol) are dissolved in 120 mL H_2O. The reaction is heated to 72 – 78° C with stirring while degassing with N_2. Once the reaction has stabilized at 77° C, the polymerization is initiated with the addition of KPS (24 mg, 8.88 µmol), and allowed to continue for three hours at 77° C under N_2. The reaction is allowed to cool and the resulting particle suspension is dialyzed against water at room temperature for three days, with the dialyzed particles lyophilized and ready for use in harvesting. Core-shell particles incorporating AAc are synthesized using a similar approach, with NIPAm (1.08 g, 9.54 mmol), BIS (55.5 mg, 0.36 mmol) and MA (734 µL, 8.10 mmol) as the initial monomer feed dissolved in 60 mL H_2O. The shell is introduced three hours after initiation, with the

addition of a new combination of feed monomers, NIPAm (2.0 g, 17.64 mmol) and BIS (55.5 mg, 0.36 mmol) in 60 mL H_2O. The reaction is allowed to continue with stirring another 3 hours under N_2 at 74° C. Particles are dialyzed to remove unreacted monomer and byproducts. The core-shell MA particles are saponified using lithium hydroxide in aqueous methanol to convert the MA units to AAc. The hydrated diameters of the particles are determined using dynamic light scattering at a scattering angle of 90°. The AAc/AMPS particles were determined to be 591.9 ± 78.6 nm in diameter and the core-shell AAc particles 1290 ± 214 nm. The particles are combined in a 50:50 mixture by weight for use in harvesting. Particle synthesis was performed by Megan Devine.

Harvest and Elution

Alligator plasma (100 µL) from ionomycin stimulated blood (1 µM, 30', 30° C) is diluted into 1.6 mL of Hydrogel particles (40 mg) suspended in 10 mM Tris-Cl buffer (particle suspension = pH 5), for a final volume of ~ 1.7 mL. After incubating approximately 18 hours at room temperature, the plasma–particle harvest mixture is centrifuged at 16.1×10^3 rcf to pellet the particles, and the pelleted particles are re-suspended in 10 mM Tris-Cl buffer (pH 7.4). This centrifugation and re-suspension process is repeated at least two times to ensure removal of excluded proteins and peptides. Following the final wash with Tris-Cl buffer, the pelleted particles are suspended in an elution solution of 1:1 trifluoroethanol (TFE): 0.1% TFA in water. The particles are gently agitated for one hour at room temperature before pelleting (as described above). The supernatant layer, containing eluted captured peptides, is set aside for later use. To ensure all peptides had been removed from the particle interior, the

elution process is repeated three more times with 20’ incubations. All elution supernatants are combined and dried via speed vacuum before de-salting by Zip-Tip for mass spectrometry analysis. Plasma challenging was performed by Stephanie Barksdale. Harvesting and elution was performed by Stephanie Barksdale and Megan Devine.

LC-MS/MS

Particle eluate is analyzed by high-sensitivity nanospray LC–MS/MS with an LTQ-Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) equipped with an EASY-nLC 1000 HPLC system (Thermo Fisher Scientific, Waltham, MA, USA). The reversed-phase LC column is a PepMap 50 μm i.d. × 15 cm long with 3 μm, 100 Å pore size, C18 resin (Thermo Fisher Scientific, Waltham, MA, USA). The mobile phase is a gradient prepared from 0.1 % aqueous formic acid (mobile phase component A) and 0.1 % formic acid in acetonitrile (mobile phase component B). After sample injection, the column is washed for 5 min with A; the peptides are eluted by using a linear gradient from 0 to 50 % B over either 45 min or 2 hours and ramping to 100 % B for an additional 2 min; the flow rate is 300 nL/min. The LTQ-Orbitrap Elite is operated in a data-dependent mode in which each full MS scan (120,000 resolving power) is followed by five MS/MS scans (120,000 resolving power) in which the five most abundant molecular ions are dynamically selected and fragmented by electron transfer dissociation (ETD) using fluoranthene as the electron transfer reagent. “FT master scan preview mode”, “Charge state screening”, “Monoisotopic precursor selection”, and “Charge state rejection” were enabled so that only the ≥ 3+ ions are selected and fragmented by ETD.

Tandem mass spectra were imported directly as .RAW files and analyzed by PEAKS *de novo* sequencing software version 6 (Bioinformatics Solutions Inc., Waterloo, ON Canada). PEAKS first performs a *de novo* sequence analysis using the ETD MS/MS data. Mass tolerance for precursor ions was 10 ppm and mass tolerance for fragment ions was 0.05 Da. Data were analyzed with no enzyme specificity, along with oxidation (+15.9949 Da) on methionine as a variable post translation modification. Confident *de novo* peptide identifications were achieved by filtering Average Local Confidence (ALC) to ≥ 30 %. Sequence tags from the confident *de novo* sequences are searched against 2 separate databases. The first is an expressed sequence tag (EST) database obtained by searching the EST database at NCBI (http://www.ncbi.nlm.nih.gov) for all known alligator EST sequences. A total of 5469 alligator EST sequences are found from a number of sources, including the Adult American Alligator Testis Library (University of Florida, Department of Zoology, Gainesville, FL), the Juvenile American Alligator Liver Library (NIBB, Japan), and the Adult American Alligator Liver Library (University of Florida, Department of Zoology, Gainesville, FL). The second database was an Alligator mississippiensis transcriptome obtained from the International Crocodilian Genome Working Group (www.crocgenomes.org) (*76*). A 1 % false discovery rate (FDR) was used as a cut-off value for reporting peptide spectrum matches (PSM) from either database. Peptides of interest, both those that are sequenced from the databases and those that have only a *de novo* sequence and thus no database equivalent, are all manually verified. For *de novo*-only sequences, only leucine (L) was denoted since it is indistinguishable from isoleucine (I) by ETD fragmentation.

CAMP prediction

Verified sequences are input into web-based CAMP prediction sites (*CAMP* database, *AntiBP2* and *APD2* (*73–75*) where each peptide is scored and predicted to have antimicrobial activity or not. Furthermore, the physico-chemical properties (length, molecular weight, nominal solution charge, pI and hydrophobicity) of all verified sequences are calculated and sorted. Peptide sequences that show good correlation to physico-chemical properties associated with known CAMPs are selected for synthesis to be evaluated for antimicrobial activity, regardless of whether the prediction sites suggest that they will have antimicrobial activity.

Resazurin Assay

Antibacterial activity of CAMPs is assessed against selected bacterial strains using the redox indicator resazurin. When incubated with metabolically active cells, the blue resazurin is converted to the highly fluorescent pink resorufin, as a result of chemical reduction by live cells. Fluorometric detection of the rate of resazurin conversion to resorufin at $530_{ex}/590_{em}$ allows quantification of bacterial survival following exposure to antibacterial compounds, such as CAMPs (*46*, *47*). It has been confirmed that the time that bacterial cultures require to achieve specified fluorescence intensities correlates inversely to the initial bacterial concentration (*46*, *47*). Results obtained using resazurin-based assays are comparable to those determined using classical dilution-plating assays for evaluating bacterial viability (*46*, *47*).

Frozen enumerated bacterial aliquots were thawed on ice and mixed. For each strain, bacteria are diluted to 2×10^6 CFU/mL in sterile 10 mM sodium phosphate (pH 7.4) and added in 50 µL aliquots to the wells of a 96-well black microtiter plate (Greiner

Bio-One 655201) containing 50 µL volumes of serially diluted CAMP, dissolved in the same phosphate buffer. Control wells contain bacteria with no peptide. The microtiter plate is incubated for 3 hours at 30° C (*B. cereus*) or 37° C for other strains. After three hours, 100 µL of PBS solution with dissolved resazurin and MHB is added to each well. The amounts of resazurin and MHB that is added is bacterial strain dependent, with the final resazurin (µM)/ MHB (wt/vol) concentrations being 100 µM/ 0.2 % for *E. coli*, 12.5 µM/ 0.05 % for *B. cereus*, 25 µM/ 2.2 % for *P. aerugoinsa*, and 50 µM/ 2.2 % for *S. aureus*. Following addition of resazurin/MHB buffer, the plate is immediately placed in either a SpectraMax Gemini EM plate-reading fluorimeter (*E. coli* and *B. cereus*) or a TeCan Safire 2 fluorimeter (*S. aureus* and *P. aeruginosa*) for incubation overnight at either 30° C (*B. cereus*) or 37° C (other strains) while monitoring fluorescence for each well.

Fluorescence data is collected from each well during the monitoring period using equations compiled by microplate data software (SoftMax Pro 4.5 or Magellen 6). Onset time of half maximal fluorescence ($T_{0.5}$) is used for quantifying *E. coli* and *B. cereus* concentrations. Onset time of 20,000 RFU (T_{20000}) is used for quantifying *S. aureus* and *P. aeruginosa* concentrations. Standard curves were generated in preliminary experiments using serially diluted bacterial suspensions (~10^6 CFU/mL - 10^3 CFU/mL) without CAMPs. Observed $T_{0.5}$ and T_{20000} values are plotted against initial CFU counts that had been determined by plating on MHB agar plates, and the relationships analyzed by linear regression, affording the following equations:

Equation 5
$\log (CFU/mL_{E.\ coli}) = -0.0002(T_{0.5}) + 9.3144$
Equation 6
$\log (CFU/mL_{B.\ cereus}) = -0.0002(T_{0.5}) + 6.4755$
Equation 7
$\log (CFU_{P.\ aeuroginosa}) = (-84806 + T_{20000})/-9956$
Equation 8
$\log (CFU_{S.\ aureus}) = (-107970 + T_{20000})/-13117$

These linear regression equations are used to interpolate survival following incubation of bacteria with CAMPs, with the CFU/mL and CFU for each well being determined based on their respective $T_{0.5}$ and T_{20000} values. Correlating bacterial CFU values for wells containing peptide with control wells containing no CAMPs it is possible to determine bacterial survival for wells containing CAMPs. Experiments against *E. coli* and *B. cereus* were performed by Carlos Rodriguez. Experiments against S. aureus and *P. aeuroginosa* were performed by Stephanie Barksdale.

Statistical Analysis

Antibacterial measurements are performed in triplicate. Bacterial survival results generated for each CAMP are fit to a variable-slope sigmoidal regression model to reveal bacterial survival curves using Prism 5 (GraphPad Software, Inc). Best-fit values generated for the survival curve-fit parameter log (EC50) are used as performance criteria. Log (EC50) represents the log of the peptide concentration (PC) that causes a halfway response between Smin and Smax, the minimal and maximal survival values, respectively, where Hill slope (HS) is the parameter used to quantify the steepness of the transition slopes in sigmoidal survival curves.

Equation 9

$$\text{Bacterial Survival} = S_{min} + \frac{(S_{max} - S_{min})}{(1+10^{((\log(EC50)-(\log(PC)))*HS))})}$$

Antilogs of the log (EC50) values, the EC50 values, are tabulated, and 95% confidence intervals (CI) are presented to demonstrate overlap and statistical significance. This data is presented in Table 6B and in graphical format in Figure 16 in the report. Statistical analysis was performed by both Carlos Rodriguez and Stephanie Barksdale.

Results and Discussion

Harvesting and Elution

Prior to harvesting, alligator blood is treated with ionomycin (a calcium ionophore) to stimulate cellular peptide release into the plasma. Ionomycin has been demonstrated to trigger the release of hCAP18 (the human cathelicidin LL-37 precursor) from neutrophil granules (*77*). Following stimulation, the plasma and cells are separated, and particle harvesting is performed from 100 µL of stimulated plasma. The particles are washed to remove excluded peptides and proteins, with the captured peptides remaining trapped in the interior of the microparticles. The trapped peptides are then eluted from the particles and desalted for mass spectrometry.

Hydrogel microparticles based on cross-linked N-isopropylacrylamide copolymer frameworks are central to the CAMP discovery process (*71*, *72*). Harvesting is performed using a 50:50 combination of two types of particles, one incorporating acrylic acid as its

affinity bait and the other combining acrylic acid and 2-acrylamido-2-methyl-propanesulfonic acid as baits. These particles enable multidimensional separation of targeted peptides from other proteins and peptides present in the samples. Negatively charged acidic groups, such as carboxylic acids and sulfonic acids, provide affinity baits for the capture of cationic peptides and proteins. At the same time, the cross-linking of the polymer scaffold excludes larger peptides and proteins, while allowing low molecular weight peptides access to affinity baits residing in the particle interior. Thus, the particles simultaneously combine elements of cation exchange and size-exclusion chromatography when capturing peptides and proteins from complex biological samples, favoring peptides with physico-chemical properties similar to those of CAMPs.

Mass Spectrometry and *de novo* Sequencing

The second step in the bioprospecting process is the identification and sequencing of potential CAMPs. The sequences of captured native intact peptides, including potential CAMPs, are elucidated using an Orbitrap Elite mass spectrometer equipped with ETD fragmentation, which has been shown to be ideally suited for fragmenting large, highly charged peptides (*78*, *79*). When combined with the high sensitivity, resolution and mass accuracy of the Orbitrap, ETD can be used for the *de novo* sequencing of full-length functional peptides. Here, ETD spectra are analyzed by PEAKS software to sequence peptides in a *de novo* manner. PEAKS then uses sequence tags from the *de novo* sequences to search an American alligator transcriptome database (*76*). However, not all of the *de novo* peptide sequences are represented in this database. Peptides of interest, both those that are confirmed from the database and those that have no database

equivalent, are all manually verified. To illustrate how peptide sequences can be derived *de novo* from ETD mass spectra, a representative spectrum is presented in Figure 15.

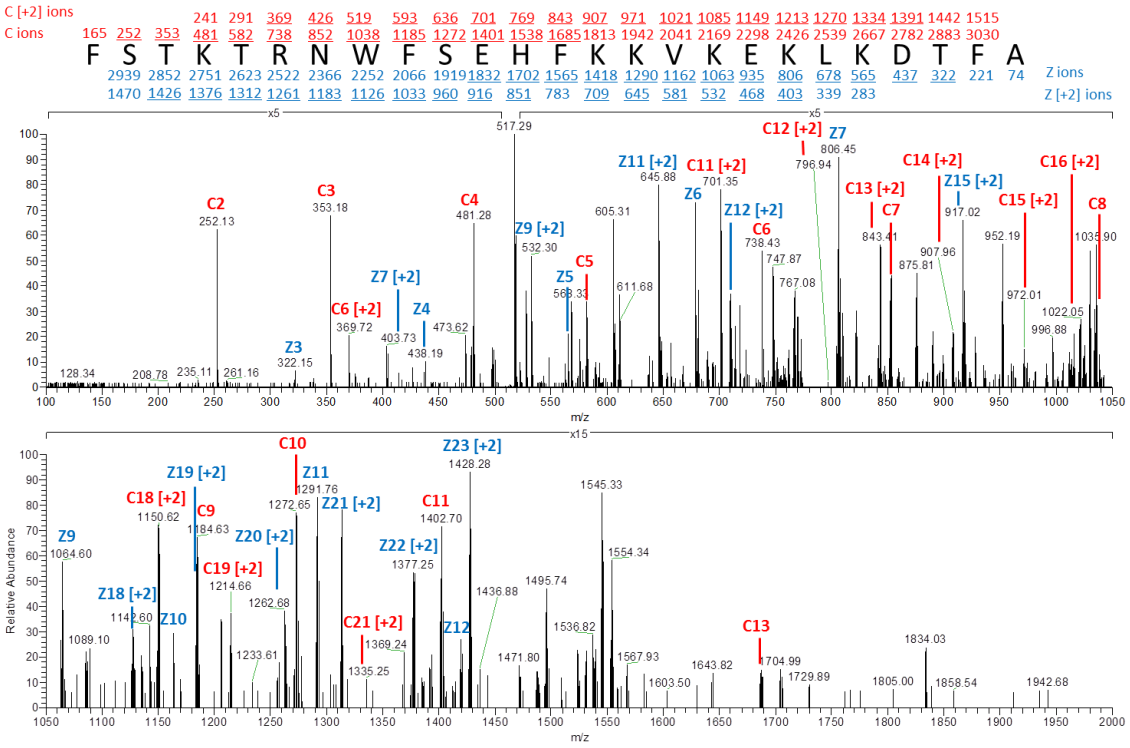

Figure 15: ETD Mass Spectrum for APOC1$_{64\text{-}88}$.
ETD mass spectrum recorded for the 25-residue peptide on the $(M+5H)^{+5}$ ion at m/z 621.33 (MW 3103.57 Da). Observed singly and doubly charged c (red) and z (blue) ions are indicated on the peptide sequence and are labeled in the spectrum. (Ions present in the spectrum are underlined.)

The bioprospecting CAMP-discovery process has led to the capture and identification of more than 570 peptides from 100 µL of alligator plasma. Using a combination of rational peptide sequence assessment based on known CAMP physico-chemical properties and web-based CAMP prediction algorithms (*73–75*), 45 potential

CAMPs were identified. Out of the potential CAMP sequences, we have chosen eight peptides to be synthesized for further evaluation. These peptides and their theoretical physico-chemical properties are presented in Table 6A.

Table 6: Activity and Sequence of Novel Alligator CAMPs.
A. Physico-chemical Properties. The physico-chemical properties for eight novel alligator CAMPs identified via the process. The peptide name is determined based on the parent protein with the amino acid sequence numbers in the subscript. B: Antibacterial Performance Data for Alligator CAMPs. Antibacterial activities against *E. coli, B. cereus, P. aeruginosa* and *S. aureus* are expressed in terms of EC50 (µg/mL) values with corresponding 95% confidence interval (CI) range. LL-37 is a human CAMP that is used as a standard for assessing antibacterial performance (*21*). NA= no activity. Data for *E. coli* and *B. cereus* courtesy of Carlos Rodriquez. Data for *S. aureus* and *P. aeruginosa* courtesy of Stephanie Barksdale.

*These peptides were *de novo* identified.
‡Hydrophobicity was calculated using the per-residue hydrophobicity scale determined by George Rose *et al.*(*80*).

A.

Peptide	Sequence	Length	Molecular Weight (Da)	Nominal Solution Charge (pH = 7)	pI	Hydrophobicity‡
$APOC1_{64-88}$	FSTKTRNWFSEHFKKVKEKLKDTFA	25	3103.57	4	10.00	-1.17
$APOC1_{67-89}$	KTRNWFSEHFKKVKEKLKDTFA	22	2766.49	4	10.00	-1.36
$FGG_{398-413}$	YSLKKTSMKIIPFTRL	16	1926.39	4	10.46	-0.05
$FGG_{401-413}$	KKTSMKIIPFTRL	13	1562.92	4	11.26	-0.19
$A1P_{394-428}$	PPPVIKFNRPFLMWIVERDTRSILFMGKIVNPKAP	35	4106.28	4	11.00	0.02
AVTG2LP*	LQTKLKKLLGLESVF	15	1717.11	2	9.70	0.36
ASAP130LP*	PPGASPRKKPRKQ	13	1445.85	5	12.02	-2.31
$NOTS_{17-38}$	VERIPLVRFKSIKKQLHERGDL	22	2656.17	3	10.27	-0.62

B.

Peptide	*E.coli*		*B.cereus*		*P.aeruginosa*		*S.aureus*	
	EC50 (µg/mL)	95% CI	EC50 (µg/mL)	95% CI	EC50 (µg/mL)	95% CI	EC50 (µg/mL)	95% CI
LL-37	0.0480	0.0346 to 0.0664	0.168	0.141 to 0.200	4.63	2.60 to 8.24	4.57	3.37 to 6.21
$APOC1_{64-88}$	0.770	0.518 to 1.14	0.983	0.895 to 1.08	7.64	4.37 to 13.4	27.7	12.0 to 63.6
$APOC1_{67-89}$	0.555	0.263 to 1.17	0.770	0.663 to 0.895	4.68	3.49 to 6.27	30.8	24.5 to 38.6
$A1P_{394-428}$	0.483	0.234 to 0.996	3.77	1.26 to 11.3	28.9	25.1 to 33.3	9.85	6.43 to 15.1
$FGG_{398-413}$	0.828	0.406 to 1.69	23.3	19.3 to 28.2	24.6	19.0 to 31.8	55.1	19.3 to 158
$FGG_{401-413}$	0.521	0.319 to 0.766	39.8	wide	32.2	26.4 to 39.2	166	very wide
AVTG2LP*	NA	NA	NA	NA	NA	NA	NA	NA
ASAP130LP*	NA	NA	NA	NA	NA	NA	233	very wide
$NOTS_{17-38}$	NA	NA	NA	NA	NA	NA	946	very wide

Antibacterial Evaluation

The third step in the process is evaluation of the antimicrobial effectiveness of the newly identified peptides. The synthetic peptides are tested against a panel of Gram-positive and Gram-negative bacteria, both pathogenic and non-pathogenic. These bacteria include *Bacillus cereus*, *Staphylococcus aureus*, *Escherichia coli*, and *Pseudomonas aeruginosa*. Antimicrobial assays designed to determine the half-maximal effective concentrations (EC50) are performed using a high-throughput resazurin metabolic indicator assay (Table 6B) (*46*, *47*).

The identified alligator CAMP candidates exhibit a variety of amino acid lengths, sequences, hydrophobicities, and charges (Table 6A). Based on CAMP prediction

algorithms, ASAP130LP, $NOTS_{17-38}$, and AVTG2LP were predicted likely to be effective antimicrobial peptides. However, none of these peptides exhibited significant antibacterial activity against the panel of bacteria (Table 6B), revealing limitations in the utility of currently available CAMP prediction models (Table 7) (*73–75*).

Table 7: CAMP prediction results.

Using 3 different web-based CAMP prediction applications (*CAMP database, AntiBP2 and APD2*) each peptide was scored and given a prediction of whether it would have antimicrobial activity (AMP) or not (Non-AMP). *CAMP* database uses three different algorithms, Support Vector Machine (SVM), Random Forest (RF) and Discriminate Analysis (DA), to calculate antimicrobial predictions (*74*). *AntiBP2* uses SVM to determine CAMP predictions (*73*). *APD2* uses a qualitative determination based on the probability of the input sequence to be antimicrobial based on known antimicrobial peptides (*75*).
**AntiBP2* requires the sequence length be at least 15 amino acids to give a prediction score.

Peptide	CAMP database Prediction Score			AntiBP2 Prediction Score	APD2 Prediction Probability
	SVM	RF	DA	SVM	Qualitative
$APOC1_{64-88}$	0.894 : Non-AMP	0.728 : Non-AMP	0.667 : Non-AMP	-0.210 : Non-AMP	+
$APOC1_{67-89}$	0.598 : Non-AMP	0.692 : Non-AMP	0.352 : Non-AMP	-0.052 : Non-AMP	+
$FGG_{398-413}$	0.508 : AMP	0.656 : Non-AMP	-0.384 : AMP	-0.172 : Non-AMP	-
$FGG_{401-413}$	0.732 : AMP	0.514 : AMP	-1.296 : AMP	ND*	-
$A1P_{394-428}$	0.935 : Non-AMP	0.838 : Non-AMP	0.363 : Non-AMP	-0.241 : Non-AMP	+
AVTG2LP*	0.821 : AMP	0.386 : Non-AMP	0.877 : AMP	0.223 : AMP	+
ASAP130LP*	0.157 : Non-AMP	0.4495 : Non-AMP	0.077 : Non-AMP	ND*	+
$NOTS_{17-38}$	0.757 : AMP	0.6 : AMP	-0.165 : Non-AMP	0.618 : AMP	+

Of the eight synthesized peptides, five show significant antibacterial activity against the bacterial panel, based on EC50 values (Figure 16): $APOC1_{64-88}$, $APOC1_{67-89}$, $A1P_{394-428}$, $FGG_{398-413}$ and $FGG_{401-413}$.

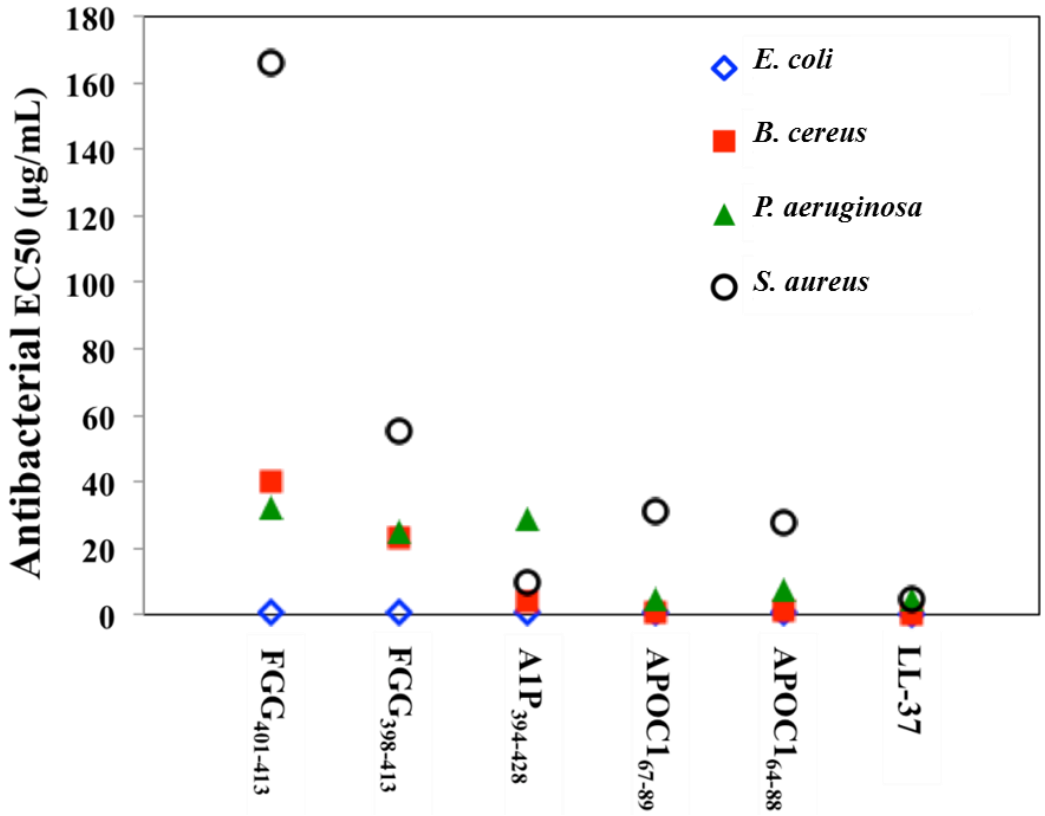

Figure 16: Potencies of LL-37 and Five Novel Alligator CAMPs.
Comparison of the antibacterial effectiveness of each peptide against *E. coli, B. cereus, P. aeruginosa* and *S. aureus*, expressed in terms of EC50 (µg/mL) values. Data for *E. coli* and *B. cereus* courtesy of Carlos Rodriquez. Data for *S. aureus* and *P. aeruginosa* courtesy of Stephanie Barksdale.

Two fragments of apolipoprotein C, $APOC1_{64\text{-}88}$ (25aa) and $APOC1_{67\text{-}89}$ (22aa), are highly homologous and share a nominal net charge of +4 at physiological pH. While neither peptide was predicted to be strongly antimicrobial by the CAMP prediction algorithms tested, $APOC1_{64\text{-}88}$ and $APOC1_{67\text{-}89}$ exerted significant antimicrobial activity against non-pathogenic *E. coli*, *B. cereus* and *P. aeruginosa*, but were not as effective against *S. aureus*. Two peptides derived from fibrinogen, $FGG_{398\text{-}413}$ (16aa) and $FGG_{401\text{-}413}$ (11aa), both carry a nominal charge of +4 at physiological pH. $FGG_{401\text{-}413}$ was

predicted by all but one of the algorithms to have antimicrobial activity, while only 2 out of 5 algorithms tested predicted $FGG_{398-413}$ to be a CAMP. Interestingly, neither peptide was found to have strong antimicrobial activity except against non-pathogenic *E. coli*. $A1P_{394-428}$, a fragment of the alpha-1-antiproteinase, is a 35-residue peptide with a predicted +4 charge at neutral pH. Although none of the algorithms predicted it to be antimicrobial, $A1P_{394-428}$ showed good activity against *E. coli*, moderate activity against *B. cereus* and *S. aureus*, and poor activity against *P. aeruginosa*.

Conclusion

We have developed a new and effective method for identifying novel and potentially useful antimicrobial peptides. The ability to harvest, sequence and evaluate novel peptides from small quantities of plasma in a high-throughput process has the potential to revolutionize the way CAMPs are identified. In this process, hydrogel microparticles harvest functional CAMPs based on their physico-chemical properties. Coupled with subsequent mass spectral analysis of the intact captured peptides, this process eliminates current labor-intensive, low-yield processes associated with conventional approaches for CAMP identification. Using this process and only 100 µL of alligator plasma, we have successfully identified and sequenced five new peptides that exhibit antimicrobial activity against Gram-positive and/or Gram-negative bacteria. Although this CAMP discovery process has only been used to analyze samples of alligator plasma to date, the relatively small sample volume requirement and the fact that the process is sample agnostic make it applicable to a broad spectrum of animals that were previously thought inaccessible, such as organisms of smaller body mass or

endangered species. This will allow analysis of the peptidomes in some of the world's most remarkable species, to dramatically expand the current CAMP library and potentially unlock the key to overcoming antibiotic resistance via the discovery of new antimicrobial peptides. Beyond CAMP discovery, we envision the bioprospecting approach being applied to mining peptidomes for diagnostic biomarkers that would be missed using conventional proteomic methods.

CHAPTER FOUR. *DE NOVO* SEQUENCING OF CATIONIC ANTIMICROBIAL PEPTIDES (CAMPS) BY ELECTRON-TRANSFER DISSOCIATION (ETD)

The identification and sequencing of novel CAMPs has proven challenging due to the limitations associated with traditional proteomics methods and difficulties sequencing peptides present in complex bimolecular mixtures. We present here a process for *de novo* sequencing novel CAMPs using tandem mass spectrometry equipped with electron-transfer dissociation (ETD). This process was initially evaluated and verified using known CAMPs with varying physico-chemical properties. The effective parameters were then applied in the analysis of a complex mixture of peptides harvested from American alligator plasma. Here, we report the successful *de novo* sequencing process for CAMPs that has led to the identification of over 600 peptides and the discovery of 5 novel CAMPs, from American alligator plasma.

Introduction.

Cationic antimicrobial peptides (CAMPs) are produced by nearly all living organisms and are an essential part of the innate immune defense against invading pathogens in higher organisms (*66*, *81*, *82*). CAMPs tend to be low molecular weight peptides that are both highly cationic and amphipathic in nature. These physico-chemical properties allow CAMPs to directly interact with pathogens in a non-receptor mediated pathway and exert broad spectrum effectiveness. Although these peptides are both used

pervasively in nature and evolutionarily ancient, limited bacterial resistance has been observed (*42*). Thus, CAMPs have the potential to revolutionize current therapeutics, with their unique ability to exert direct broad spectrum antimicrobial, antiviral and antifungal properties (*39*). The discovery and cataloging of these remarkable peptides could unlock the key to overcoming antibiotic resistance. Unfortunately, the peptide diversity provided by current CAMP libraries are limited due to the inefficient and labor-intensive approaches currently used for CAMP discovery.

The approaches that have been used to discover and identify native CAMPs from biological samples have proven slow and low-yielding. Current proteomic methods for CAMP discovery usually require large sample volumes (frequently upwards of 1 L), involve time consuming HPLC or electrophoretic fractionation, and rely on enzymatic digestions coupled with collision induced dissociation (CID) mass spectrometry to determine peptide sequences (*69*, *83*). Large sample volumes and HPLC fractionation can lead to loss of low abundance peptides such as CAMPs. Enzymatic digestion of samples is problematic since information regarding the native form of peptides can be lost, which could result in incorrect peptide sequences and erroneous antimicrobial performance data. In addition, being that CAMPs are known to contain an abundance of lysine and arginine residues, enzymatic digestion with commonly used trypsin can produce very small peptide fragments that are too hydrophilic to be retained on an HPLC column. While CAMPs are comparatively small peptides relative to other proteins and peptides present in plasma and other biological environments, they are large when compared to the peptide fragments generated by proteolytic digestion for analysis by mass spectrometry.

Unfortunately, CID fragmentation efficiency drops off significantly for peptides with charge states of +4 or greater. Therefore, this type of fragmentation is not well suited for larger, more highly charged peptides such as intact CAMPs (*79*). To overcome these issues a new approach to CAMP discovery must be taken.

Recent advances in protein mass spectrometry have greatly improved mass accuracy resolving power, sensitivity limits, and data acquisition speed. Another recent advance comes in the form of new peptide fragmentation chemistries. One of these advances, electron-transfer dissociation (ETD), is of particular relevance as it allows for the efficient fragmentation of larger, more highly charged peptides (*78*, *79*, *84*). This is accomplished by the transfer of an electron from a radical anion to a protonated peptide, resulting in the fragmentation of the peptide along the Cα-N bonds. The resulting peptide fragments produce a complementary c and z-type ion series, as opposed to the typical b and y-type ion series generated by CID (Figure 17).

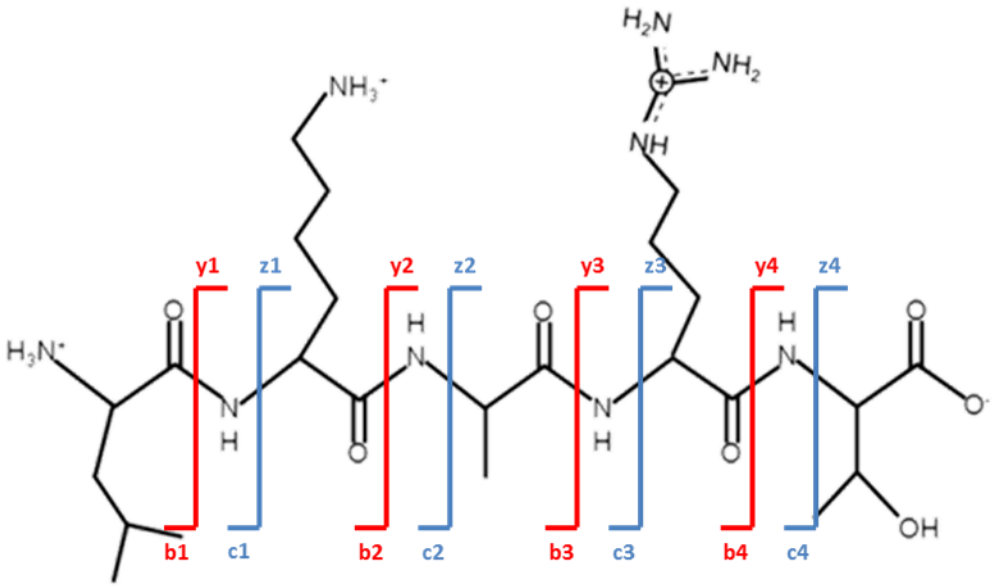

Figure 17: CID vs. ETD Peptide Fragmentation.
Peptide fragmentation produces complementary ion series, CID results in b and y-type ion series (red), while ETD results in c and z-type ion series (blue).

By combining ETD with high resolution, and high mass accuracy, it becomes possible to sequence the larger, more highly charged CAMPs in a *de novo* manner. This negates the need for enzymatic digestion and allows for identification of intact, full-length native peptides.

Recently, we developed a new and effective method for CAMP identification that allows for the rapid extraction and analysis of the native, functional peptidome. This method uses microparticle harvesting of intact, functional peptides from biological samples coupled with analysis of the harvested peptides using ETD mass spectrometry. The microparticles preferentially harvest CAMPs and CAMP-like peptides based on their physico-chemical properties. Incorporation of anionic affinity baits in the forms of carboxylic and sulfonic acids within the particle matrix electrostatically complement the

positive charges associated with CAMPs. The cross-linking of the polymer framework and inclusion of an inert shell both help to exclude larger proteins/peptides from interacting effectively with the anionic baits. Initial analysis of the harvested peptides by mass spectrometry reveals a complex mixture of peptides. *De novo* sequencing of this highly complex mixture is able to be accomplished with the assistance of PEAKS, a *de novo* sequencing software package. PEAKS can import and work from raw MS/MS data, allowing rapid determination of peptide sequences in a *de novo* manner that can then be manually verified. In this study, we compare the ability of various mass spectrometers equipped with ETD to *de novo* sequence full-length, functional CAMPs. Using known CAMPs we are able to establish and optimize parameters for *de novo* sequencing, which we in turn implanted in the *de novo* sequencing of multiple novel CAMPs from American alligator plasma.

Materials and Methods.

Peptide Harvest and Elution.

Alligator plasma (100 µL) from ionomycin stimulated blood (1 µM, 30', 30° C) is diluted into 1.6 mL of Hydrogel particles (40 mg) suspended in 10 mM Tris-Cl buffer (particle suspension = pH 5), for a final volume of ~ 1.7 mL. After incubating approximately 18 hours at room temperature, the plasma–particle harvest mixture is centrifuged at 16.1×10^3 rcf to pellet the particles, and the pelleted particles are re-suspended in 10 mM Tris-Cl buffer (pH 7.4). This centrifugation and re-suspension process is repeated at least two times to ensure removal of excluded proteins and peptides. Following the final wash with Tris-Cl buffer, the pelleted particles are

suspended in an elution solution of 1:1 trifluoroethanol (TFE): 0.1% TFA in water. The particles are gently agitated for one hour at room temperature before pelleting (as described above). The supernatant layer, containing eluted captured peptides, is set aside for later use. To ensure all peptides had been removed from the particle interior, the elution process is repeated three more times with 20' incubations. All elution supernatants are combined and dried via speed vacuum before de-salting by Zip-Tip for mass spectrometry analysis.

Chromatography.

The LTQ-ETD uses a reversed-phase manually packed 75 μm i.d. × 10 cm long with 5 μm, 200 Å pore size, C18 resin LC column (Michrom Bioresources, Auburn, CA). The mobile phase is a gradient prepared from 0.1 % aqueous formic acid (mobile phase component A) and 0.1 % formic acid in acetonitrile (mobile phase component B). After sample injection, the column is washed for 10 min with A; the peptides are eluted by using a linear gradient from 0 to 50 % B over 45 min and ramping to 100 % B for an additional 2 min; the flow rate is 300 nL/min.

The LTQ-Orbitrap Elite uses a reversed-phase PepMap 50 μm i.d. × 15 cm long with 3 μm, 100 Å pore size, C18 resin LC column (Thermo Fisher Scientific, Waltham, MA, USA). The mobile phase is a gradient prepared from 0.1 % aqueous formic acid (mobile phase component A) and 0.1 % formic acid in acetonitrile (mobile phase component B). After sample injection, the column is washed for 5 min with A; the peptides are eluted by using a linear gradient from 0 to 50 % B over either 45 min or 2 hours and ramping to 100 % B for an additional 2 min; the flow rate is 300 nL/min.

LC-MS/MS.

Particle eluate is analyzed on two different instruments, the LTQ-ETD and the LTQ-Orbitrap Elite mass spectrometers (Thermo Fisher Scientific, Waltham, MA, USA). The LTQ-ETD uses nanospray LC-MS/MS equipped with an Agilent 1100 HPLC system (Agilent Technologies, Santa Clara, CA). The LTQ-ETD is operated in a data-dependent mode in which each full MS scan is followed by five MS/MS scans in which the five most abundant molecular ions are dynamically selected and fragmented by electron transfer dissociation (ETD) using fluoranthene as the electron transfer reagent.

The LTQ-Orbitrap Elite uses high-sensitivity nanospray LC–MS/MS equipped with an EASY-nLC 1000 HPLC system (Thermo Fisher Scientific, Waltham, MA, USA). The LTQ-Orbitrap Elite is operated in a data-dependent mode in which each full MS scan (120,000 resolving power) is followed by five MS/MS scans (120,000 resolving power) in which the five most abundant molecular ions are dynamically selected and fragmented by electron transfer dissociation (ETD) using fluoranthene as the electron transfer reagent. "FT master scan preview mode", "Charge state screening", "Monoisotopic precursor selection", and "Charge state rejection" were enabled so that only the ≥ 3+ ions are selected and fragmented by ETD.

Spectra Analysis.

Tandem mass spectra were imported directly as .RAW files and analyzed by PEAKS *de novo* sequencing software version 6 (Bioinformatics Solutions Inc., Waterloo, ON Canada). PEAKS first performs a *de novo* sequence analysis using the ETD MS/MS data. Mass tolerance for precursor ions was 20 ppm (LTQ-ETD) or 10 ppm (LTQ-Orbitrap) and mass tolerance for fragment ions was 0.5 Da (LTQ-ETD) or 0.05 Da (LTQ-

Orbitrap). Data were analyzed with no enzyme specificity, along with oxidation (+15.9949 Da) on methionine as a variable post translation modification. Confident *de novo* peptide identifications were achieved by filtering Average Local Confidence (ALC) to ≥ 30 %. Sequence tags from the confident *de novo* sequences are searched against 2 separate databases. The first is an expressed sequence tag (EST) database obtained by searching the EST database at NCBI (http://www.ncbi.nlm.nih.gov) for all known alligator EST sequences. A total of 5469 alligator EST sequences are found from a number of sources, including the Adult American Alligator Testis Library (University of Florida, Department of Zoology, Gainesville, FL), the Juvenile American Alligator Liver Library (NIBB, Japan), and the Adult American Alligator Liver Library (University of Florida, Department of Zoology, Gainesville, FL). The second database was an Alligator mississippiensis transcriptome obtained from the International Crocodilian Genome Working Group (www.crocgenomes.org) (*76*). A 1 % false discovery rate (FDR) was used as a cut-off value for reporting peptide spectrum matches (PSM) from either database. Peptides of interest, both those that are sequenced from the databases and those that have only a *de novo* sequence and thus no database equivalent, are all manually verified. For *de novo*-only sequences, only isoleucine (L) was denoted since it is indistinguishable from isoleucine (I) by ETD fragmentation.

Comparative Spectra Analysis.

Mass spectra were directly imported as .RAW files and analyzed by SIEVE comparative spectra analysis software (Thermo Scientific, Waltham, MA, USA). A proteomics comparison was done using non-differential single class analysis of the raw

data. The parameters were set for full m/z range (300 – 2000), full retention time range, full frame time width (2.5 min) and m/z width (10 ppm). Once raw data parameters are set the output parameters of 20,000 max frames, a peak intensity threshold of 200,000 and max charge state of +8 are set.

CAMP prediction.

Verified sequences are input into web-based CAMP prediction sites (*CAMP* database, *AntiBP2* and *APD2* (*73–75*) where each peptide is scored and predicted to have antimicrobial activity or not. Furthermore, the physico-chemical properties (length, molecular weight, nominal solution charge, pI and hydrophobicity) of all verified sequences are calculated and sorted. Peptide sequences that show good correlation to physico-chemical properties associated with known CAMPs are selected for synthesis to be evaluated for antimicrobial activity, regardless of whether the prediction sites suggest that they will have antimicrobial activity.

Results and Discussion.

De novo Sequencing of Known CAMPs

As the *de novo* sequencing of CAMPs using mass spectrometry has only recently become an area of interest, the best methods for achieving this objective have yet to be established. Therefore, initial studies focused on three known CAMPs (SMAP-29, Buforin and Indolicidin) in order to establish and verify a workflow and parameters for successful *de novo* sequencing CAMPs. These three CAMPs vary in their length, molecular weight, net charge, pI and hydrophobicity (Table 8).

Table 8: Known CAMP Properties.
The sequences, length, MW, charge, pI and hydrophobicity of each known CAMP used for validating *de novo* sequencing methods.

Peptide	Actual Sequence	Length (res)	Molecular Weight (Da)	Net Charge	pI	Hydrophobicity
SMAP-29	RGLRRLGRKIAHGVKKYGPTVLRIIRIAG	29	3254.03	9	12.31	-0.21
Buforin	TRSSRAGLQFPVGRVHRLLRK	21	2433.43	6	12.60	-0.64
Indolicidin	ILPWKWPWWPWRR-NH2	13	1906.03	3	12.01	-1.07

The three known model CAMPs were selected for *de novo* sequencing method development based on their varied physico-chemical properties to ensure that different classes of CAMPs were represented.

Initial sequencing of the 3 known CAMPs was performed using a Thermo LTQ mass spectrometer equipped with ETD (LTQ-ETD). The LTQ-ETD is a low resolution instrument, and therefore, it is not possible to accurately determine the charge states of the highly charged precursor ions of the peptides, or of the highly charged ETD fragment ions (Figures 18-20).

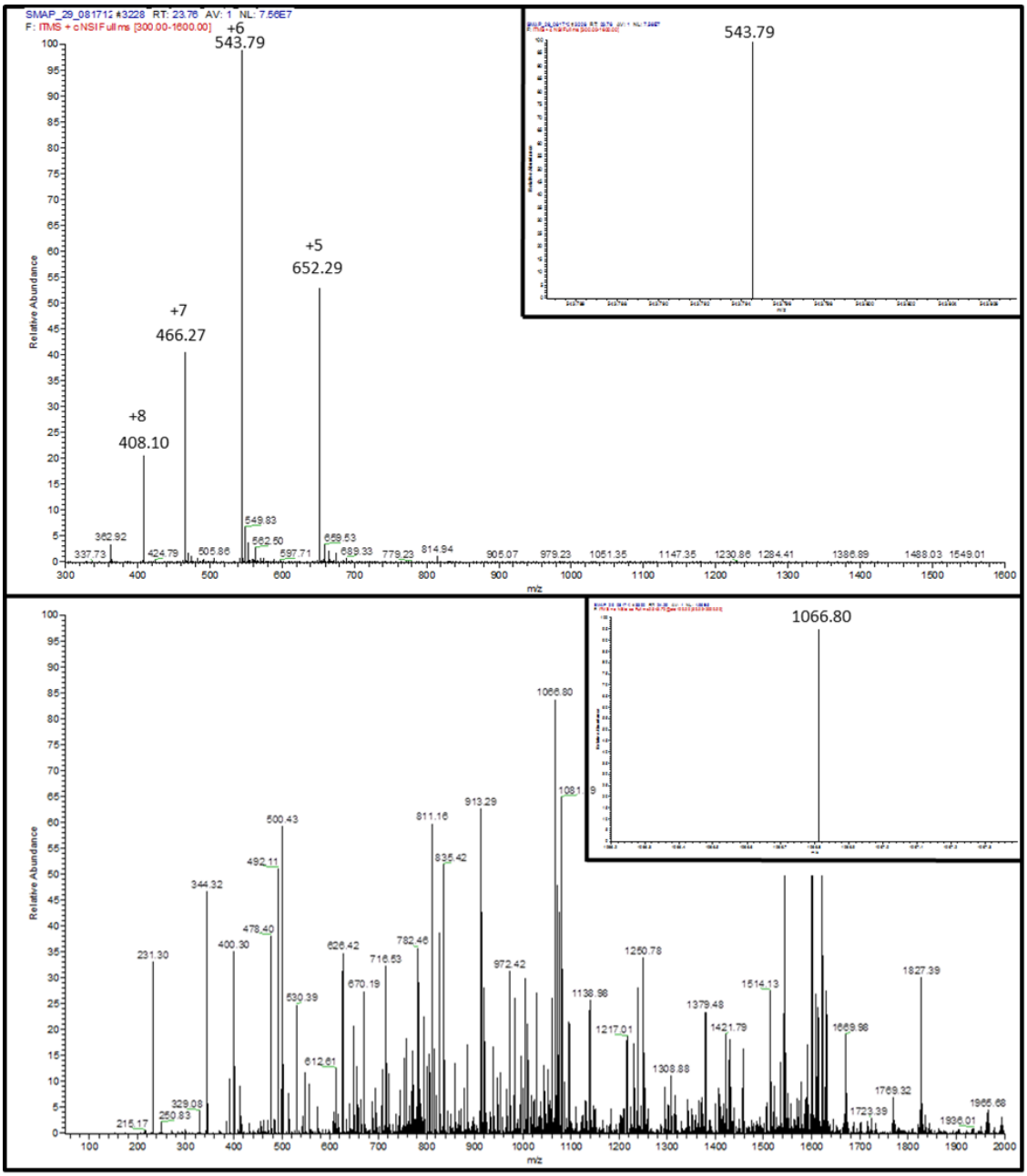

Figure 18: MS and MS/MS Spectra of SMAP-29 on LTQ-ETD.
The MS spectrum is presented with identified charge states and an inset showing the isotopic distribution for the $(M+6H)^{+6}$ ion at m/z 543.79 (top pane). The ETD spectrum for the 543.79 precursor ion is presented with an inset showing the isotopic distribution of a selected fragment ion (bottom pane).

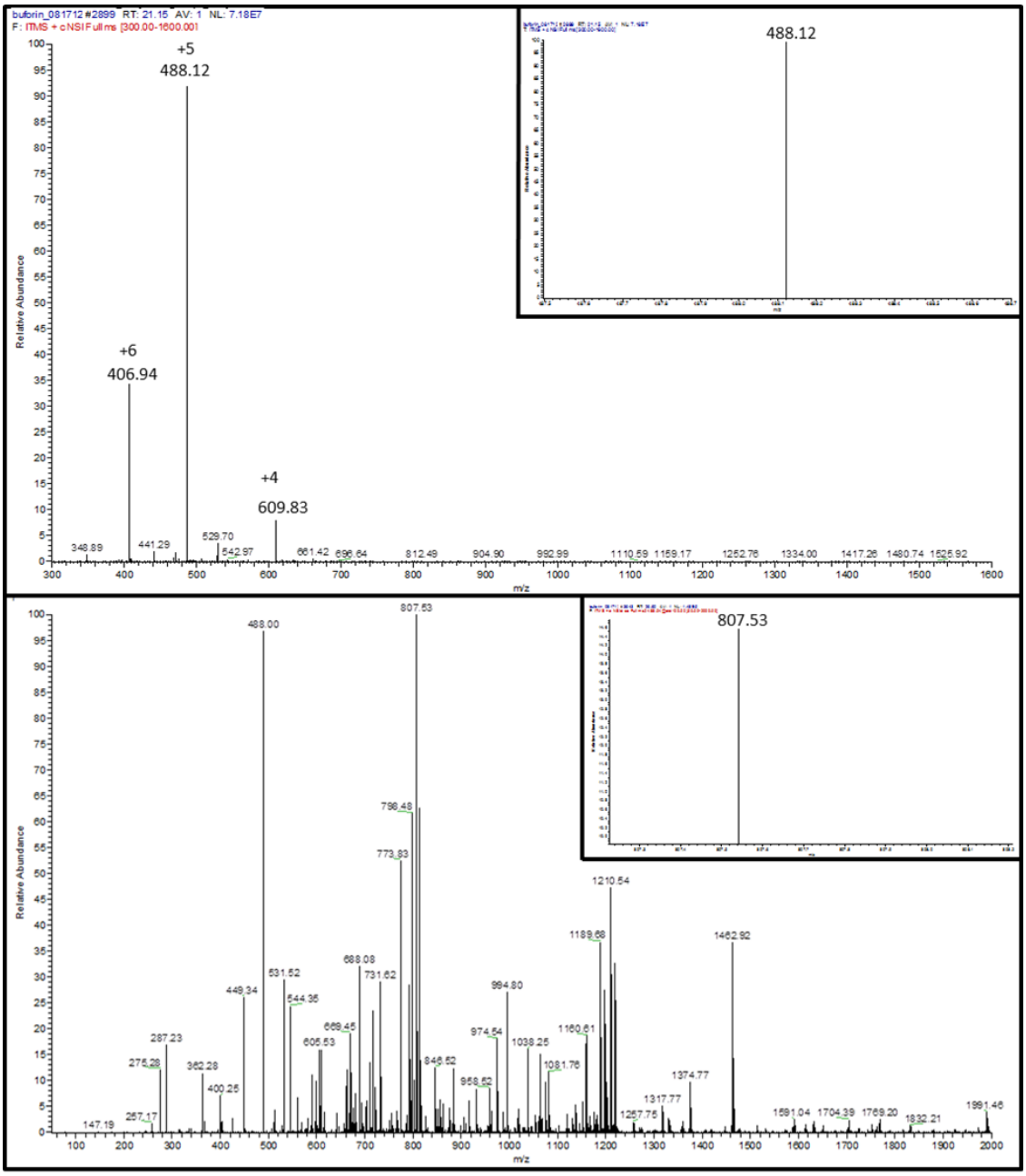

Figure 19: MS and MS/MS Spectra of Buforin on LTQ-ETD.
The MS spectrum is presented with identified charge states and an inset showing the isotopic distribution for the $(M+H)^{+5}$ ion at m/z 488.12 (top pane). The ETD spectrum for the 488.12 precursor ion is presented with an inset showing the isotopic distribution of a selected fragment ion (bottom pane).

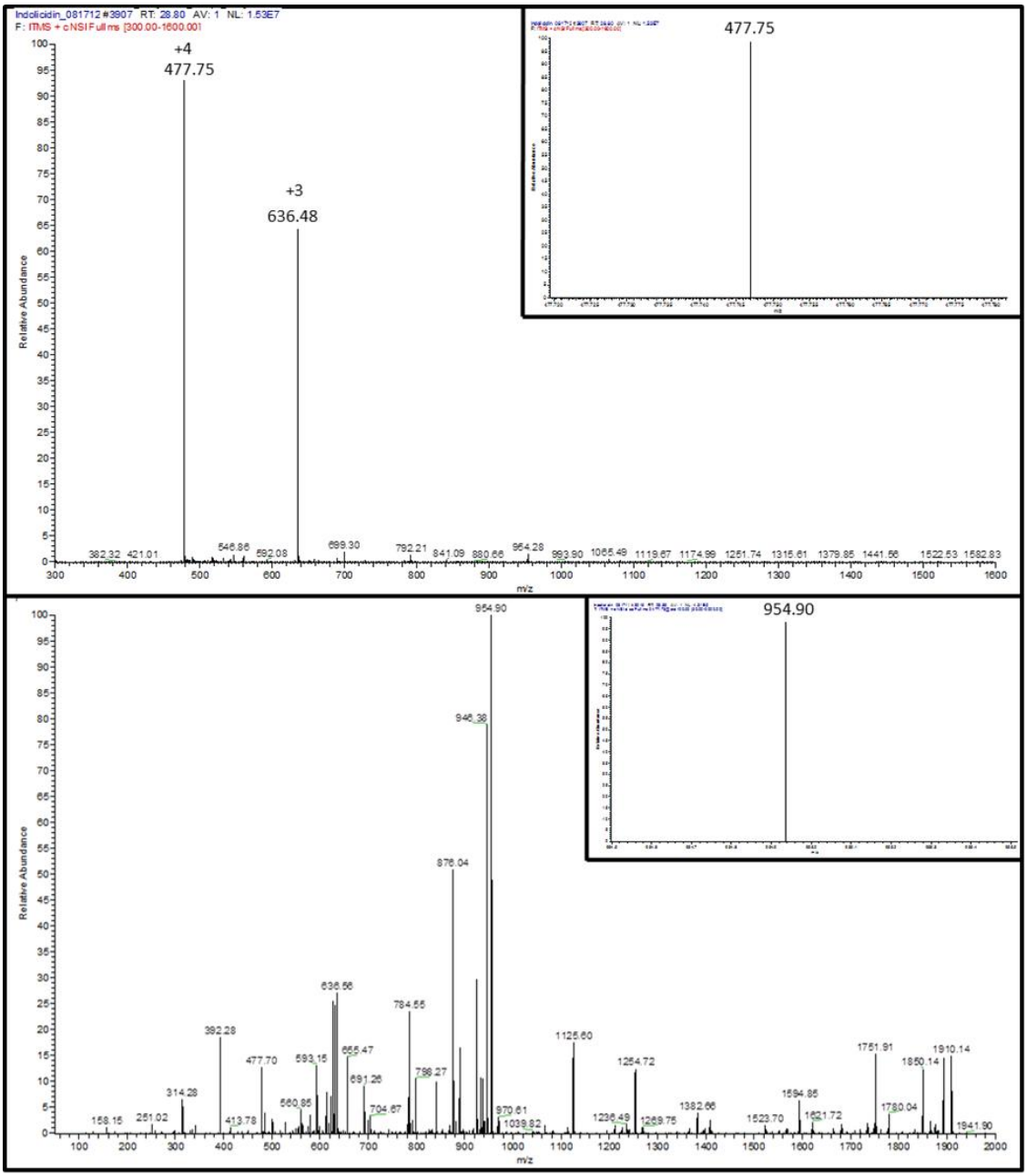

Figure 20: MS and MS/MS Spectra of Indolicidin on LTQ-ETD.
The MS spectrum is presented with identified charge states and an inset showing the isotopic distribution for the $(M+4H)^{+4}$ ion at m/z 477.75 (top pane). The ETD spectrum for the 477.75 precursor ion is presented with an inset showing the isotopic distribution of a selected fragment ion (bottom pane).

Being that the LTQ-ETD is also a low accuracy instrument, it is not possible to determine an accurate mass of these peptides. When using this data in PEAKS, the software is unable to produce a *de novo* sequence at all due to the inability to determine the correct precursor charge state, and thus, the correct accurate mass. After manually adjusting the charge states for each precursor ion, PEAKS still produces low confidence *de novo* sequences due to the low-mass accuracy of the spectra (Table 9).

Table 9: Comparison of PEAKS Identification with LTQ-ETD vs. Orbitrap-ETD Data.
The 3 known CAMPs were run on both Thermo LTQ-ETD (low resolution) and Orbitrap-ETD (high resolution) instruments and put through PEAKS *de novo* sequencing software. The *de novo* sequencing found by PEAKS is compared to the actual CAMP sequence and the identity correct is recorded.
*Precursor charge state could not be determined by PEAKS, so correct charge state was manually corrected to obtain a *de novo* sequence.

SMAP-29		% Correct
Actual Sequence	RGLRRLGRKIAHGVKKYGPTVLRIIRIAG	
LTQ-ETD	RGLRRLGSLAFGMCGSSALKLFHLWLRLV*	24.1
Orbitrap-ETD	RGLRLRGRRRHGVKKYPGTVLRLLARVA	82.3
Buforin		
Actual Sequence	TRSSRAGLQFPVGRVHRLLRK	
LTQ-ETD	TRSSRAGLKSKKGSSSNLRAAHH*	38.1
Orbitrap-ETD	TRSSRAGLQKTNGRVHRLLRK	85.7
Indolicidin		
Actual Sequence	ILPWKWPWWPWRR-NH2	
LTQ-ETD	LPLWKEAERELWPLK(-.98)*	38.4
Orbitrap-ETD	LPLWKEGPWQRWRR(-.98)	76.9

PEAKS determined the sequence of SMAP-29 with 24% accuracy and those of Buforin and Indolicidin with 38% correct sequence identity. The low sequence accuracy

is most likely due to the fact that the LTQ-ETD provides low-mass accuracy for both MS and MS/MS ions. The reliability of a *de novo* determined peptide sequence improves with increased accuracy in mass measurement since the exact mass contains information regarding elemental composition and without the monoisotopic mass poor sequence identification was observed.

Since analyzing highly charged peptides with low mass accuracy and resolution has been shown to be detrimental in *de novo* sequencing these CAMPs, the next step was analyzing them using an LTQ-Orbitrap Elite, which offers increased resolution and higher mass accuracy. This increased mass accuracy and resolution will allow correct determination of the monoisotopic masses and charge states for both the precursor and fragment ions (Figure 21-23). The MS spectrum of SMAP-29 (Figure 21) shows an isotopic distribution of 0.17 Da for the $(M+6H)^{+6}$ ion at m/z 543.35 confirming the charge state of +6.

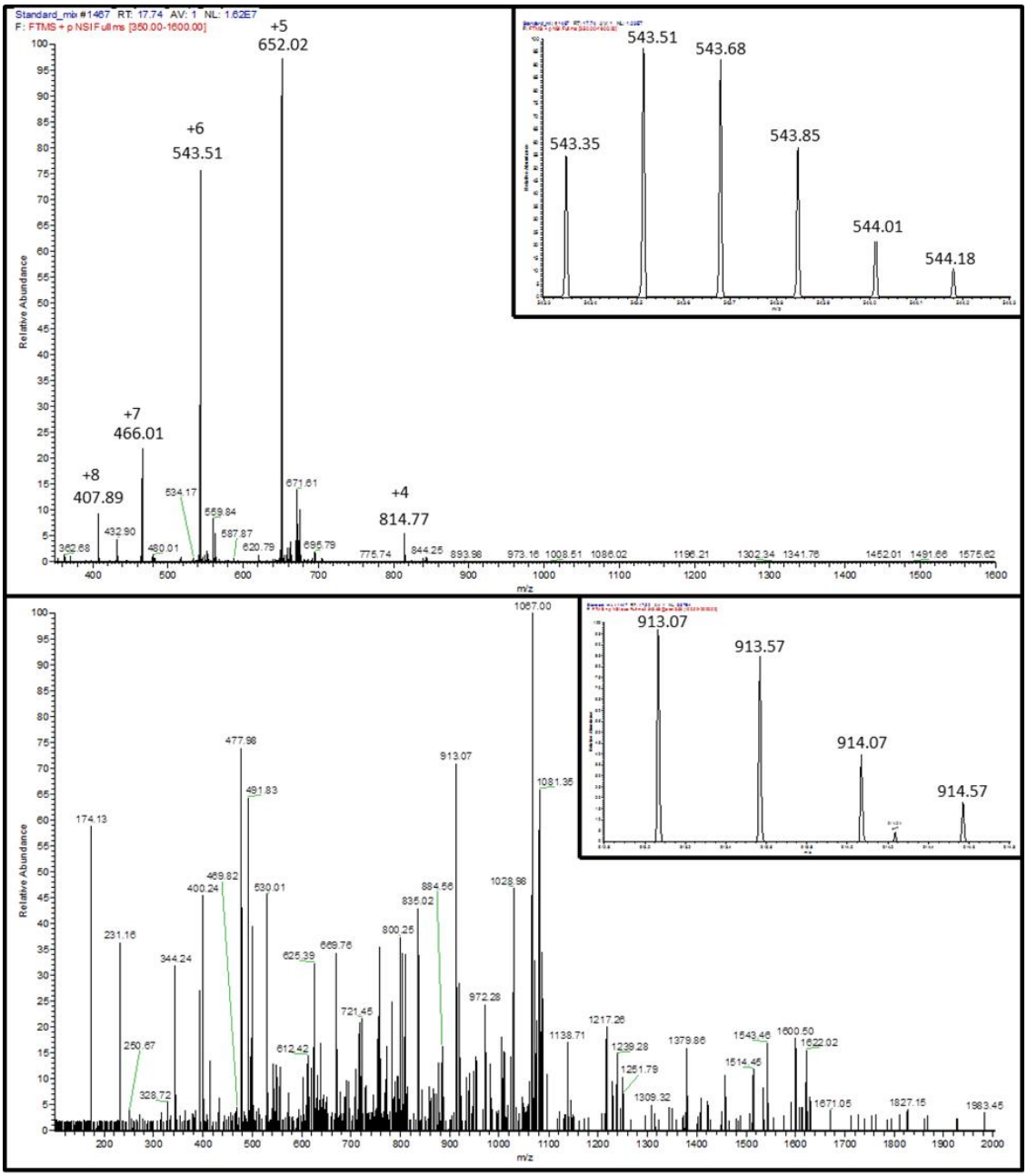

Figure 21: MS and MS/MS Spectra of SMAP-29 on Orbitrap-ETD.
The MS spectrum is presented with identified charge states and an inset showing the isotopic distribution for the $(M+6H)^{+6}$ ion at m/z 543.51 (top pane). The ETD spectrum for the 543.51 precursor ion is presented with an inset showing the isotopic distribution of a selected fragment ion (bottom pane).

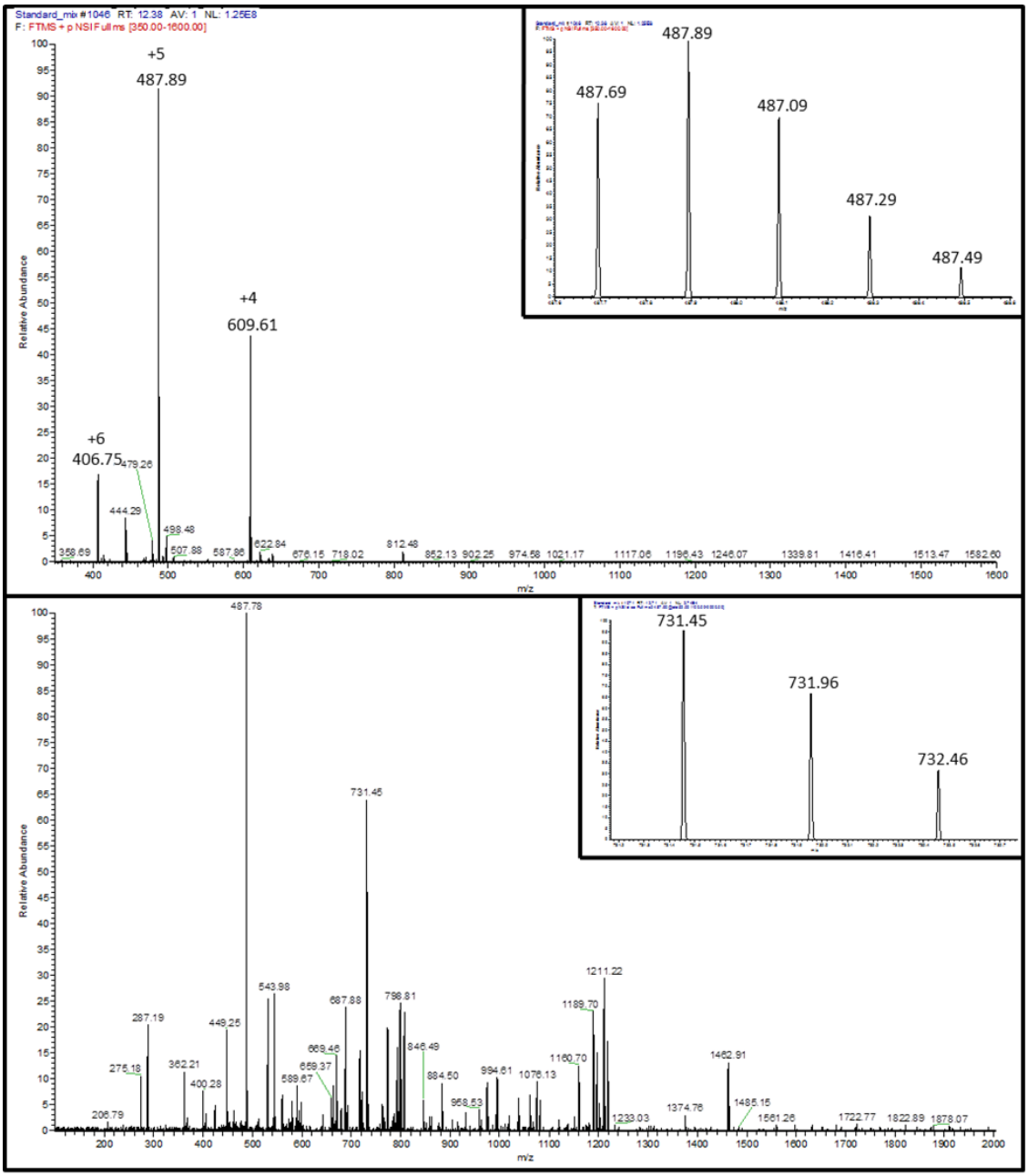

Figure 22: MS and MS/MS Spectra of Buforin on Orbitrap-ETD.
The MS spectrum is presented with identified charge states and an inset showing the isotopic distribution for the $(M+5H)^{+5}$ ion at m/z 487.89 (top pane). The ETD spectrum for the 487.89 precursor ion is presented with an inset showing the isotopic distribution of a selected fragment ion (bottom pane).

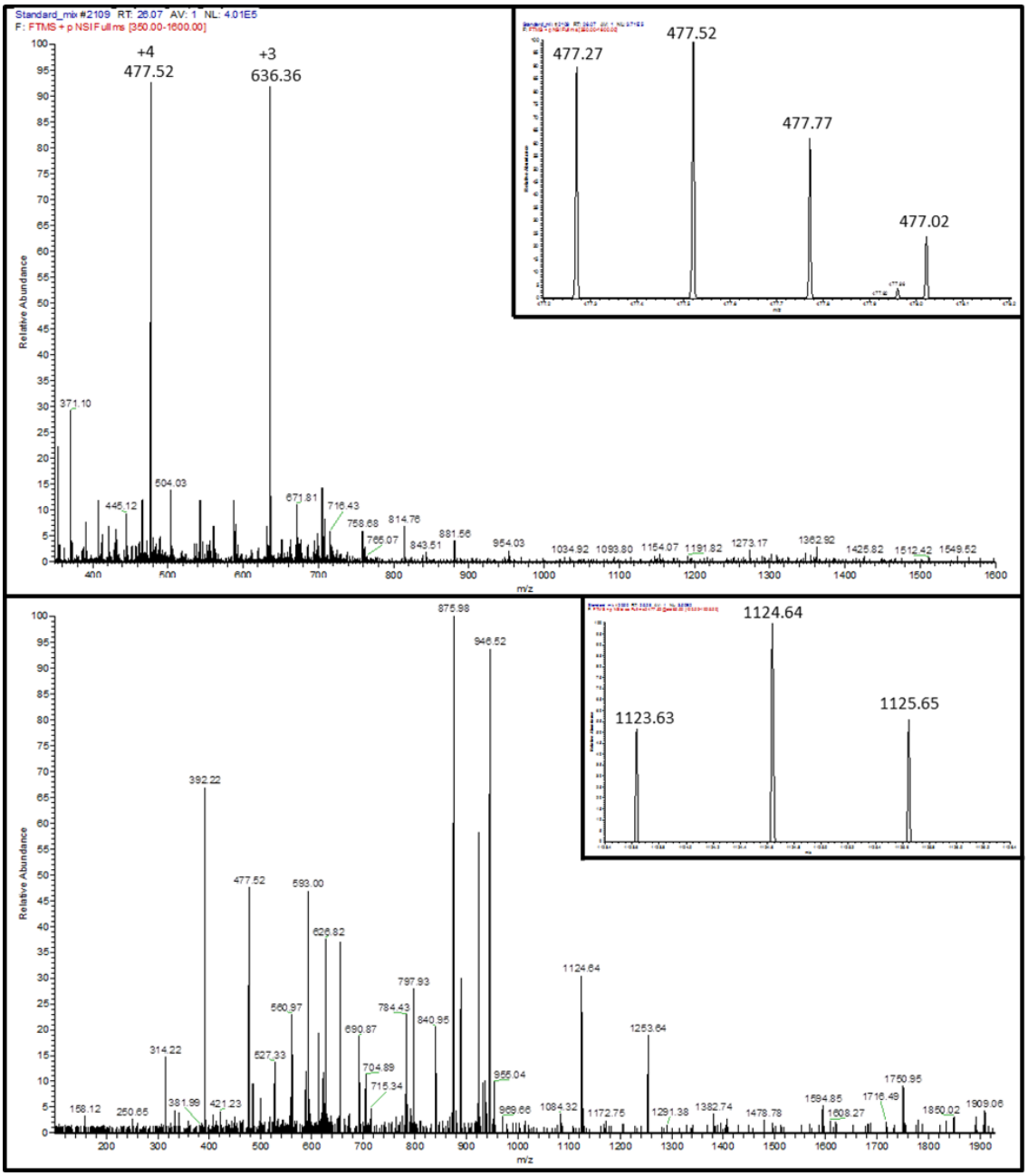

Figure 23: MS and MS/MS Spectra of Indolicidin on Orbitrap-ETD.
The MS spectrum is presented with identified charge states and an inset showing the isotopic distribution for the $(M+4H)^{+4}$ ion at m/z 477.52 (top pane). The ETD spectrum for the 477.52 precursor ion is presented with an inset showing the isotopic distribution of a selected fragment ion (bottom pane).

The MS spectrum of Buforin, shown in Figure 22, has an isotopic distribution of 0.20 Da for the $(M+5H)^{+5}$ ion at m/z 487.69 allowing for the correct precursor charge state of +5 to be determined. Indolicidins' $(M+4H)^{+4}$ ion at m/z 477.52 shows an isotopic distribution of 0.25 Da identifying its +4 charge state (Figure 23). The ETD spectra for each of these peptides also have the required isotopic distribution of each fragment ion to determine charge state. The sequencing ability of the PEAKS software is greatly enhanced when analyzing the higher resolution and accuracy spectra generated using the LTQ-Orbitrap Elite (Table 9). The software *de novo* generated sequences with 82, 85 and 77 % correct sequence for SMAP-29, Buforin and Indolicidin, respectively. The ability to establish the exact masses and charges of the peptide ions greatly enhanced PEAKS ability to correctly *de novo* determine the peptide sequences from the raw MS/MS data. Although, the PEAKS-generated sequences were not 100 % correct, the discrepancies with the actual correct peptide sequences are easily detected after manual *de novo* sequencing from the raw data.

De novo Sequencing of Novel Alligator CAMPs

After identifying the appropriate instrument and parameters required for the *de novo* sequencing of known CAMPs, the process can be applied to identify novel CAMPs from alligator plasma. This process begins with the harvesting of CAMPs and CAMP-like peptides from 100 µL of alligator plasma using microparticles, followed by the elution of the captured peptides from the particles. The eluents, containing intact native peptides, are then, de-salted and dried. The dried peptides are dissolved in 0.1 % formic acid (FA), and the complex mixture is loaded directly onto a reverse-phase C18 LC

column. The intact peptides are eluted with a linear acetonitrile/0.1 % FA gradient, mass separated in the LTQ, and then analyzed in the orbitrap. Following initial analysis, data-dependent analysis was used to identify the top 5 parent peptide ions, which are then selected and fragmented by ETD using fluoranthene followed by analysis in the orbitrap to get high resolution spectra of sufficient quality for sequencing.

Due to the complex nature of the sample, both 1 and 3 MS/MS microscans were performed to compare the quality of the resulting spectra. Increasing the number of microscans can be used to increase the signal-to-noise ratio in the spectra, but also increases the scan cycle time. After evaluating the effect of different microscan collections, it was determined that increasing the number of microscans from 1 to 3 did not significantly affect the quality of the resulting spectra. However, the amount of data and number of MS/MS spectra acquired significantly decreased as the number of microscans collected was increased from 1 to 3 (Table 10).

Table 10: SIEVE Comparison of 1 vs. 3 Microscans.
The MS and MS/MS data of 1 and 3 microscans were compared to identify changes in data collection based on total ion count, total peptide count and total MS/MS spectra collected for peptide sequencing.

Number of Microscans	Total Number of Ions	Total Number of Ions with Charge 3-8	Total Number of Peptides with Charge 3-8	Total Number of MS/MSSpectra	Total Number of Peptide MS/MS Spectra with Charge 3-8
1	20000	4585	~ 4100	2900	569
3	17213	3553	~ 3000	1137	328

This resulted in the total number of captured peptide ions decreasing by approximately 15 %, while the total number of MS/MS scans decreased by over 60 %. Using 3 microscans also resulted in fewer identifications of actual peptide MS/MS scans (328 scans) compared to the number identified using 1 microscan (564 scans), resulting in fewer peptides to be sequenced. These results indicated that the time lapse required for 3 microscan was too long for efficient capturing of the peptide diversity in the complex sample and that 1 microscan was sufficient to obtain quality spectra.

Once the data had been collected, *de novo* assembly of the peptide sequences could be performed using the spectra. To begin the sequencing process, PEAKS *de novo* sequencing software was used to generate initial *de novo* sequences. The raw data files are directly uploaded to PEAKS and then run through a data refine, which allows for the selection of peaks based on mass only or mass and charge, as well as, filters for the scan data. The mass only data refine was used on the raw data for the harvested alligator peptides since the high resolution orbitrap data allows for assignment of correct precursor mass. Additionally, a quality filter of 0.65 was applied to ensure only quality MS/MS spectra are used for *de novo* sequencing. Following the data refine, *de novo* sequencing is performed based on the filtered MS/MS data with the error tolerance for the monoisotopic precursor ion set to 20 ppm (LTQ-ETD) or 10 ppm (Orbitrap-ETD) and the fragment ion set to 0.5 Da (LTQ-ETD) or 0.05 Da (Orbitrap-ETD). The PEAKS software is then able to assign a *de novo* sequence for each MS/MS scan that falls within the set parameters, and the resulting sequences are presented with associated average local confidence (ALC) scores. The ALC score gives an indication of the confidence that each

amino acid given within the *de novo* generated sequence is correct, thus the higher the score the higher the probability of a correct peptide sequence. PEAKS then uses sequence tags from the *de novo* sequences and searches against an American alligator EST database and American alligator transcriptome database. The database search generates sequences that align with any genes found and allows for identification of potentially wrong amino acids or differentiation between leucine and isoleucine that were found in the *de novo* sequences.

After PEAKS has generated both *de novo* and database aligned sequences for the alligator peptides, the next step is to predict which sequences will likely have antimicrobial properties. The sequences are submitted for analysis by three different CAMP prediction websites (*APD2*, *CAMP* database and *AntiBP2*) and their respective prediction models. Each prediction algorithm employs its own set of algorithms using various known CAMP properties to assess the uses classification algorithms and machine learning algorithms based antimicrobial potential for each sequence.

- *APD2*: Performs a residue analysis of the input peptide sequence, by using known principles for antimicrobial peptides to determine whether the sequence has the potential to be antimicrobial (*75*). Additionally, it performs alignments between the input sequence and the known sequences in the database.

- *CAMP* database: Uses machine learning algorithms based on experimentally validated sequences for CAMP prediction. It employs support vector machine (SVM), random forest (RF) and/or discriminant analysis (DA). SVM performs pattern recognition and regression based on higher dimensional non-linear

transformation. RF uses an ensemble of trees for classification and regression. DA uses linear combinations of independent variables to predict the group membership for each dependent variable (*74*).

- *AntiBP2*: This program uses classification and learning algorithms based on known CAMP sequences for CAMP prediction based on both N- and C-termini approaches. It employs Quantitative Matrix (QM), Artificial Neural Network (ANN) and Support Vector Machines (SVM). QM uses a matrix with the propensity of each residue at each position in the sequence to be antimicrobial. ANN performs pattern recognition and regression. SVM implements pattern recognition and regression based on higher dimensional non-linear transformation based on amino acid composition between antibacterial and non-antibacterial peptides (*73*).

The prediction results for all *de novo* sequenced alligator peptides are tabulated in Appendix, Table 14. This represents only half of the process used to identify sequences that may correspond to peptides with antimicrobial properties. In addition to analysis with the prediction algorithms, the sequences are sorted and their physico-chemical properties (length, MW, charge, hydrophobicity and pI) calculated (Appendix, Table 15). The charge is the nominal solution charge at a neutral pH. The hydrophobicity was calculated using the per-residue hydrophobicity scale determined by George Rose *et al* (*80*). The analysis of harvests from chapter 3 and subsequent harvest analysis have led to the identification of 691 peptides. In addition to the CAMP prediction scores, peptides are rationally selected as potential CAMPs based on our knowledge of CAMPs and the

properties associated with them. This combination of computational and rational analysis has yielded 45 potential CAMPs (Table 11).

Table 11: Calculate Physico-Chemical properties of Potential CAMPs.
The parent pepide, sequences, length, MW, charge, pI and hydrophobicity of each potential-CAMP was tabulated.

Source Protein	Peptide Sequence	Length (res)	Molecular Weight (Da)	Net Charge	pI	Hydrophobicity
Alpha-2-macroglobulin precursor	FVLKSFAQARRY	12	1485.75	3	11.00	-0.11
Apolipoprotein A-I precursor	RESIKPYTESIKTHL	15	1802.06	1	8.50	-1.04
Apolipoprotein A-I precursor	VKDLSRQKLEL	11	1328.57	1	8.56	-0.73
Apolipoprotein B-I	KSRVNRMKQNL	11	1372.77	4	12.02	-1.65
Apolipoprotein B-I	KSRVNRMKQNL	11	1372.77	4	12.02	-1.65
Apolipoprotein C-I precursor	EHFKKVKEKLK	11	1413.73	3	9.83	-1.72
Apolipoprotein C-I precursor	FKKVKEKLKDTFA	13	1581.92	3	9.83	-0.91
Apolipoprotein C-I precursor	KVFADNIGEKTKA	13	1420.63	1	8.50	-0.63
Apolipoprotein C-I precursor	EHFKKVKEKLKDTFA	15	1848.18	2	9.41	-1.23
Apolipoprotein C-I precursor	FSEHFKKVKEKLKDTFA	16	2082.43	2	9.40	-0.97
Apolipoprotein C-I precursor	KTRNWFSEHFKKVKEKLKDTFA	22	2766.49	4	10.00	-1.36
Apolipoprotein C-I precursor	STKTRNWFSEHFKKVKEKLKDTFA	24	2954.57	4	10.00	-1.31
Apolipoprotein C-I precursor	FSTKTRNWFSEHFKKVKEKLKDTFA	25	3103.57	4	10.00	-1.17
Apolipoprotein E precursor	ALRDQGQRLREQL	13	1582.78	1	9.56	-1.4
Compliment 3	NKGKIVQAGRQLRQAGQNL	19	2078.18	4	12.02	-1.01
Compliment 3	ILNKGKIVQAGRQLRQAGQNL	21	2304.35	4	12.02	-0.51
Dermicidin	PGLARQAPKPRKQ	13	1445.86	4	12.02	-1.66
Dystonin	DRLEELREFANFDFDIWRKKYMRWMNHKKSRVMDFFRRI	39	5224.63	4	10.14	-1.11
F2 Prothrombin precursor	MHLKKPVAFSNF	12	1418.72	2	10.00	0.03
Fibrinogen	YSLKKTSMKIIPFTRL	16	1926.39	4	10.46	-0.05
Fibrinogen	YSLKKTSMKIIPFTRL	16	1926.39	4	10.46	-0.05
Fibrinogen	KKTSMKIIPFTRL	13	1562.92	4	11.26	-0.19
Fibrinogen	KKTSMKIIPFTRL	13	1562.98	4	11.26	-0.19
Fibrinogen	WLGNEKIHL	9	1109.29	0	6.75	-0.37
Fibrinogen	WKGSWYSLK	9	1154.33	2	9.70	-1.01
Fibrinogen	YSLKKMSMKIRPFFPQ	16	2000.07	4	10.46	-0.51
Fibrinogen	KKMSMKIRPFFPQ	13	1636.89	4	11.26	-0.75
Glutathione peroxidase 3 precursor	PALKYVRPGGGFAPNFQL	18	1932.26	2	10.01	-0.09
Hemoglobin subunit alpha	DMSHNSAQIRAHGKKVFSAL	20	2197.5	2	9.99	-0.47
Hemoglobin subunit alpha	NSAQIRAHGKKVFSAL	16	1727	3	11.27	-0.24
Hemoglobin subunit epsilon	ASFGEAVKHLDNIKGHFANL	20	2168.44	0	6.96	-0.13
Hypothetical Protein	KRTFTPSQAG	10	1092.22	2	11.00	-1.15
Immunoglobulin	KFIQKSVQKQPG	12	1387.65	3	10.30	-1.13
Immunoglobulin	KFTQRSIQKTAG	12	1364.57	3	11.17	-1.07
Immunoglobulin	KFSQRSVQKSPGN	13	1462.63	3	11.17	-1.55
Immunoglobulin	ALPMKFIQKSVQKQPG	16	1800.19	3	10.30	-0.48
NOTS1	VERIPLVRFKSIKKQLHERGDL	22	2656.17	3	10.27	-0.62
SAP 130	PPGASPRKKPRKQ	13	1445.85	5	12.02	-2.31
SERPIN Alpha-1-antiproteinase	PPPVIKFNRPFLMWIVERDTRSILFMGKIVNPKAP	35	4106.28	4	11.00	0.02
Titin	YRFGKELVQSRKYR	14	1829	4	10.43	-1.52
Transferrin	EQQTRFGR	8	1022.1	1	9.70	-2.23
Vitellogenin	LQTKLKKLLGLESVF	15	1717.11	2	9.70	0.36
Vitellogenin	KVPRVKEHSKGK	12	1391.83	4	10.46	-1.77
Unknown	HFPLRSKYNRLTK	13	1658.93	4	11.10	-1.35
Unknown	TPVFPGRRRGSTNLRASPG	19	2025.09	4	12.48	-0.94

Of the identified 45 potential alligator CAMPs, 21 have been synthesized and tested for antibacterial potency against Gram-positive and Gram-negative bacteria. These

studies led to the identification of 5 active CAMPs, $APOC1_{64-88}$, $APOC1_{67-89}$, $FGG_{401-413}$, $FGG_{398-413}$ and $A1P_{394-428}$ (Table 12).

Table 12: Antibacterial Performance Data for Alligator CAMPs.
Antibacterial activities against *E. coli, B. cereus, P. aeruginosa* and *S. aureus* are expressed in terms of EC50 (µg/mL) values with corresponding 95% confidence interval (CI) range. LL-37 is a human CAMP that is used as a standard for assessing antibacterial performance (*21*). Data for *E. coli* and *B. cereus* courtesy of Carlos Rodriquez. Data for *S. aureus* and *P. aeruginosa* courtesy of Stephanie Barksdale. Data referenced in Chapter 3.

Peptide	*E.coli*		*B.cereus*		*P.aeruginosa*		*S.aureus*	
	EC50 (µg/mL)	**95% CI**	**EC50 (µg/mL)**	**95% CI**	**EC50 (µg/mL)**	**95% CI**	**EC50 (µg/mL)**	**95% CI**
LL-37	0.0480	0.0346 to 0.0664	0.168	0.141 to 0.200	4.63	2.60 to 8.24	4.57	3.37 to 6.21
$APOC1_{64-88}$	0.770	0.518 to 1.14	0.983	0.895 to 1.08	7.64	4.37 to 13.4	27.7	12.0 to 63.6
$APOC1_{67-89}$	0.555	0.263 to 1.17	0.770	0.663 to 0.895	4.68	3.49 to 6.27	30.8	24.5 to 38.6
$A1P_{394-428}$	0.483	0.234 to 0.996	3.77	1.26 to 11.3	28.9	25.1 to 33.3	9.85	6.43 to 15.1
$FGG_{398-413}$	0.828	0.406 to 1.69	23.3	19.3 to 28.2	24.6	19.0 to 31.8	55.1	19.3 to 158
$FGG_{401-413}$	0.521	0.319 to 0.766	39.8	wide	32.2	26.4 to 39.2	> 100	very wide

The 5 active CAMPs showed varying sequences and spectral quality (Figure 24). Comparing the PEAKS *de novo* assigned sequences to the actual sequences, as determined based on alignment with the alligator databases and manual verification, it is evident that PEAKS *de novo* sequencing works very well with some peptides and not as well with others (Table 13).

Table 13: Comparison of PEAKS *de novo* Sequences vs. PEAKS Database Sequences.
The *de novo* sequences found by PEAKS are compared to the manually verified CAMP sequences and the identity correct is recorded.

Peptide	Actual Sequence	PEAKS *de novo* Sequence	% Identity
$APOC1_{64-88}$	FSTKTRNWFSEHFKKVKEKLKDTFA	PHTKTRNFWSEHFKKVKEKLKDTFA	92.0
$APOC1_{67-89}$	KTRNWFSEHFKKVKEKLKDTFA	TKRKM(+15.99)HPPGLPVMPVPGPHVRTRY	13.6
$FGG_{398-413}$	YSLKKTSMKIIPFTRL	SYLKKTSMKLPLFTRL	100
$FGG_{401-413}$	KKTSMKIIPFTRL	KKTSMKLLPFTLR	100
$A1P_{394-428}$	PPPVIKFNRPFLMWIVERDTRSILFMGKIVNPKAP	YKVLQTVGMFKWGVSKRGRQSVGSLLGMFKLVVLTHG	31.4

$APOC1_{64-88}$, $FGG_{398-413}$ and $FGG_{401-413}$ show very good PEAKS *de novo* sequence correlation to the actual peptide sequence. The PEAKS *de novo* sequences obtained for $APOC1_{64-88}$ gave 92 % correct identity, while both $FGG_{398-413}$ and $FGG_{401-413}$ gave 100 % correct amino acid identification. Looking at the ETD spectra for these peptides it is evident that the quality of the spectra is high and all three peptides have nearly complete c and z-ion series (Figure 24A, C and D).

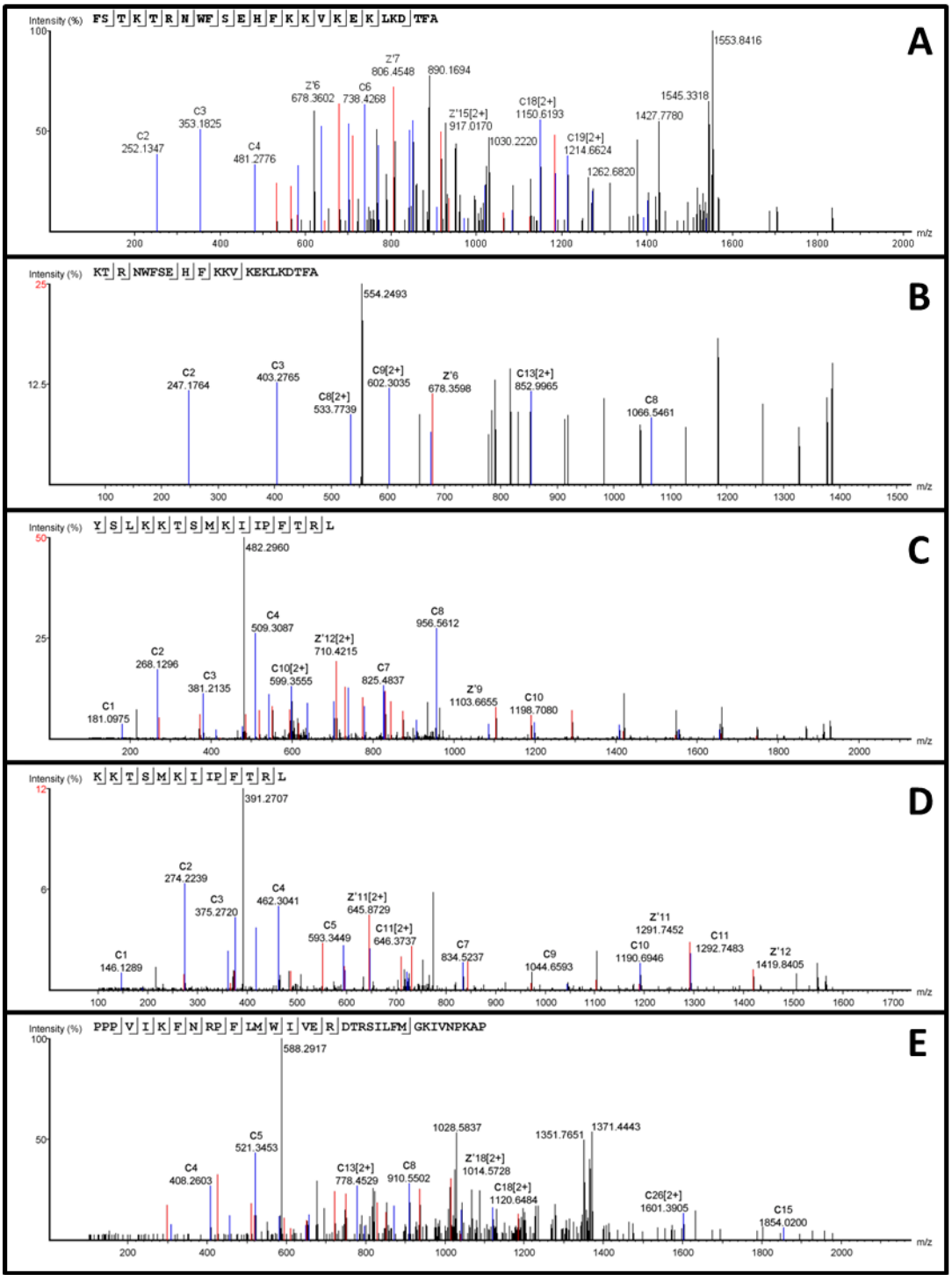

Figure 24: ETD spectra of 5 *de novo* Sequenced Alligator CAMPs.
The ETD spectra for $APOC1_{64\text{-}88}$ (A), $APOC1_{67\text{-}89}$ (B), $FGG_{401\text{-}413}$ (C), $FGG_{398\text{-}413}$ (D) and $A1P_{394\text{-}428}$ (E) are presented. Observed singly and doubly charged c (blue) and z (red)

ions are indicated on the spectra. The underlined ions in the sequence indicate ions that are present in the spectrum.

With $A1P_{394-428}$ PEAKS gave a 31 % correct *de novo* sequence compared to the actual sequence. $A1P_{394-428}$ is a larger, high molecular weight peptide and examination of the ETD spectra (Figure 24E) reveals few high molecular weight fragment ions or doubly charged ions, which causes the c-terminal amino acids to be hard to identify. However, using the same spectra when searched against the database allowed for proper identification of the sequence. The PEAKS *de novo* sequence of $APOC1_{67-89}$ was significantly off with a correct identity of 13 % of the amino acids. The quality of the ETD spectra for $APOC1_{67-89}$ is very low, with very few fragmentation ions to use to assign correct amino acids in a *de novo* manner (Figure 24B). Yet, when the spectrum is searched against the PEAKS database a correct sequence identity is found. The ability of PEAKS to assign peptide sequences in a *de novo* manner based on MS/MS spectra of varying quality, identify sequence tags for database alignment, and then use this information to correct and complete the predicted peptide sequence is a critical tool in the *de novo* sequencing of these complex mixtures. However, manual verification of the PEAKS-determined sequences is still essential, in order to assure confidence in the predicted sequences.

Conclusion.

In this study we were able to develop and verify a process to *de novo* sequence novel CAMPs based on MS/MS spectra. Known CAMPs were used to optimize the proper instrumentation and fragmentation techniques and parameters required for

effective *de novo* sequencing of highly charged, high molecular weight peptides. Once the parameters for *de novo* sequencing of CAMPs were identified, they were applied in the identification of CAMPs from alligator plasma. In the process of identifying novel CAMP sequences from alligator plasma we used the assistance of PEAKS *de novo* sequencing software that greatly enhanced the throughput and sensitivity of *de novo* sequencing the complex mixture. PEAKS gave the ability of taking both high and low quality spectra and determining a correct sequence between *de novo* and database searches that would have been missed with strictly manual sequencing. The ability to rapidly and precisely *de novo* sequences, particularly CAMPs, allows for the identification of numerous species and greatly increase the CAMP library for future therapeutic potential.

APPENDIX

Table 14: CAMP Prediction for All *de novo* Sequenced Peptides.
The results of CAMP prediction using Support Vector Machine (SVM), Random Forest (RF) and Discriminant Analysis (DA) are tabulated for all 691 *de novo* sequenced peptides. The colors relate to which algorithms predicted that sequence to be antimicrobial. SVM only (red), RF only (yellow), DA only (blue), SVM and RF (orange), SVM and DA (purple), RF and DA (green), all three algorithms (grey). Peptide sequence can be found with corresponding Sequence ID number in Table 15.

	SVM		RF		DA	
Seq. ID	**Class**	**AMP Probability**	**Class**	**AMP Probability**	**Class**	**AMP Probability**
1	AMP	0.504	AMP	0.5095	AMP	0.669
2	AMP	0.821	Non-AMP	0.386	AMP	0.877
3	AMP	0.841	Non-AMP	0.4635	AMP	0.699
4	Non-AMP	0.026	Non-AMP	0.351	Non-AMP	0.476
5	Non-AMP	0.181	Non-AMP	0.268	Non-AMP	0.047
6	AMP	0.864	Non-AMP	0.3595	Non-AMP	0.365
7	AMP	0.828	Non-AMP	0.4675	Non-AMP	0.122
8	Non-AMP	0.381	Non-AMP	0.4265	AMP	0.918
9	Non-AMP	0.171	AMP	0.5145	Non-AMP	0.447
10	Non-AMP	0.174	Non-AMP	0.288	Non-AMP	0.047
11	Non-AMP	0.122	Non-AMP	0.337	Non-AMP	0.272
12	Non-AMP	0.42	AMP	0.5105	AMP	0.617
13	Non-AMP	0.482	Non-AMP	0.405	Non-AMP	0.466
14	Non-AMP	0.42	AMP	0.5105	AMP	0.617
15	Non-AMP	0.304	Non-AMP	0.1985	Non-AMP	0.034

16	Non-AMP	0.254	AMP	0.5365	AMP	0.576
17	Non-AMP	0.174	Non-AMP	0.288	Non-AMP	0.047
18	AMP	0.584	Non-AMP	0.289	Non-AMP	0.026
19	AMP	0.997	Non-AMP	0.408	Non-AMP	0.218
20	Non-AMP	0.269	AMP	0.53	AMP	0.521
21	AMP	0.783	Non-AMP	0.4565	AMP	0.749
22	Non-AMP	0.269	AMP	0.53	AMP	0.521
23	AMP	0.642	Non-AMP	0.39	Non-AMP	0.13
24	AMP	0.486	Non-AMP	0.3315	Non-AMP	0.128
25	Non-AMP	0.393	AMP	0.5085	AMP	0.708
26	Non-AMP	0.24	Non-AMP	0.2585	Non-AMP	0.015
27	AMP	0.605	AMP	0.591	Non-AMP	0.289
28	AMP	0.899	Non-AMP	0.3395	Non-AMP	0.254
29	Non-AMP	0.264	Non-AMP	0.261	Non-AMP	0.172
30	Non-AMP	0.333	AMP	0.503	Non-AMP	0.3
31	Non-AMP	0.324	Non-AMP	0.34	AMP	0.812
32	Non-AMP	0.418	Non-AMP	0.2375	Non-AMP	0.174
33	Non-AMP	0.113	Non-AMP	0.4155	Non-AMP	0.371
34	Non-AMP	0.319	Non-AMP	0.475	Non-AMP	0.049
35	Non-AMP	0.238	Non-AMP	0.2555	Non-AMP	0.31
36	Non-AMP	0.379	Non-AMP	0.294	Non-AMP	0.036
37	AMP	0.584	Non-AMP	0.289	Non-AMP	0.026
38	AMP	0.728	AMP	0.5665	Non-AMP	0.484
39	Non-AMP	0.481	Non-AMP	0.4195	AMP	0.801
40	Non-AMP	0.369	Non-AMP	0.3885	Non-AMP	0.296
41	AMP	0.828	Non-AMP	0.4675	Non-AMP	0.122
42	AMP	0.982	AMP	0.586	AMP	0.925
43	AMP	0.496	Non-AMP	0.2535	Non-AMP	0.317
44	AMP	0.513	Non-AMP	0.2735	Non-AMP	0.23

45	Non-AMP	0.392	Non-AMP	0.3365	Non-AMP	0.019
46	Non-AMP	0.462	Non-AMP	0.461	Non-AMP	0.065
47	AMP	0.917	AMP	0.714	AMP	0.647
48	Non-AMP	0.436	Non-AMP	0.3555	AMP	0.657
49	AMP	0.697	AMP	0.5775	AMP	0.803
50	Non-AMP	0.016	Non-AMP	0.368	Non-AMP	0.067
51	Non-AMP	0.003	Non-AMP	0.212	Non-AMP	0.029
52	AMP	0.504	AMP	0.5095	AMP	0.669
53	Non-AMP	0.464	AMP	0.7355	Non-AMP	0
54	AMP	0.488	Non-AMP	0.2995	Non-AMP	0.374
55	Non-AMP	0.416	Non-AMP	0.348	AMP	0.598
56	AMP	0.908	AMP	0.75	AMP	0.556
57	Non-AMP	0.079	AMP	0.543	Non-AMP	0.132
58	AMP	0.726	Non-AMP	0.4935	Non-AMP	0.056
59	Non-AMP	0.134	AMP	0.6235	AMP	0.926
60	Non-AMP	0.076	AMP	0.6285	AMP	0.8
61	Non-AMP	0.474	AMP	0.893	AMP	0.875
62	AMP	0.568	Non-AMP	0.449	Non-AMP	0.427
63	Non-AMP	0.04	AMP	0.6815	AMP	1
64	Non-AMP	0.297	Non-AMP	0.023	Non-AMP	0.002
65	Non-AMP	0.441	Non-AMP	0.367	Non-AMP	0.176
66	AMP	0.647	Non-AMP	0.445	Non-AMP	0.135
67	AMP	0.89	Non-AMP	0.3665	AMP	0.506
68	AMP	0.863	Non-AMP	0.386	Non-AMP	0.466
69	AMP	0.605	Non-AMP	0.335	Non-AMP	0.22
70	AMP	0.527	AMP	0.817	AMP	0.876
71	AMP	0.851	AMP	0.598	AMP	0.917
72	Non-AMP	0.429	Non-AMP	0.375	AMP	0.672
73	AMP	0.815	AMP	0.8115	AMP	0.908
74	AMP	0.548	Non-AMP	0.27	Non-AMP	0.004
75	Non-	0.334	Non-	0.464	Non-	0.087

	AMP		AMP		AMP	
76	AMP	0.863	AMP	0.6425	Non-AMP	0.371
77	AMP	0.96	Non-AMP	0.489	AMP	0.754
78	Non-AMP	0.375	Non-AMP	0.458	Non-AMP	0.214
79	Non-AMP	0.166	Non-AMP	0.192	Non-AMP	0.116
80	AMP	0.953	AMP	0.975	AMP	0.966
81	Non-AMP	0.272	Non-AMP	0.315	Non-AMP	0.014
82	AMP	0.924	AMP	0.96	AMP	0.864
83	Non-AMP	0.105	Non-AMP	0.2715	Non-AMP	0.145
84	Non-AMP	0.32	Non-AMP	0.343	Non-AMP	0.035
85	Non-AMP	0.12	Non-AMP	0.2715	Non-AMP	0.03
86	AMP	0.624	AMP	0.5475	AMP	0.557
87	AMP	0.794	Non-AMP	0.4495	AMP	0.918
88	Non-AMP	0.191	Non-AMP	0.4345	Non-AMP	0.133
89	AMP	0.877	Non-AMP	0.3505	Non-AMP	0.275
90	AMP	0.925	AMP	0.8885	AMP	0.852
91	AMP	0.833	AMP	0.526	Non-AMP	0.347
92	Non-AMP	0.157	AMP	0.6565	AMP	0.889
93	AMP	0.694	AMP	0.845	AMP	0.871
94	AMP	1	AMP	0.5785	AMP	0.593
95	AMP	0.796	AMP	0.652	AMP	0.94
96	AMP	0.954	Non-AMP	0.3695	Non-AMP	0.007
97	Non-AMP	0.312	AMP	0.522	Non-AMP	0.037
98	Non-AMP	0.371	Non-AMP	0.3805	Non-AMP	0.088
99	AMP	0.981	AMP	0.6255	AMP	0.977
100	AMP	0.958	Non-AMP	0.462	Non-AMP	0.17
101	Non-AMP	0.089	Non-AMP	0.436	AMP	0.711
102	AMP	0.768	Non-AMP	0.4695	Non-AMP	0.076
103	AMP	0.994	AMP	0.6435	AMP	0.982
104	AMP	0.892	AMP	0.774	Non-AMP	0.349
105	AMP	0.788	AMP	0.7735	AMP	0.886
106	AMP	0.635	AMP	0.653	Non-	0.322

					AMP	
107	AMP	0.825	AMP	0.5745	AMP	0.984
108	AMP	0.723	Non-AMP	0.489	Non-AMP	0.037
109	AMP	0.898	AMP	0.6425	AMP	0.716
110	AMP	0.814	AMP	0.558	AMP	0.951
111	AMP	0.983	AMP	0.954	AMP	1
112	AMP	0.491	Non-AMP	0.2445	Non-AMP	0.096
113	AMP	0.66	AMP	0.893	AMP	0.667
114	AMP	0.999	AMP	0.7695	AMP	0.988
115	AMP	0.754	AMP	0.5885	AMP	0.666
116	Non-AMP	0.187	Non-AMP	0.313	Non-AMP	0.044
117	AMP	0.785	AMP	0.7175	AMP	0.867
118	Non-AMP	0.304	Non-AMP	0.4895	Non-AMP	0.181
119	AMP	0.638	AMP	0.8885	AMP	0.928
120	AMP	0.617	AMP	0.795	AMP	0.731
121	AMP	0.67	AMP	0.7365	AMP	0.938
122	AMP	0.518	Non-AMP	0.396	Non-AMP	0.261
123	AMP	0.894	AMP	0.7325	AMP	0.876
124	Non-AMP	0.33	AMP	0.5845	AMP	0.957
125	Non-AMP	0.184	Non-AMP	0.409	Non-AMP	0.012
126	AMP	0.744	AMP	0.8535	AMP	0.907
127	Non-AMP	0.279	Non-AMP	0.479	Non-AMP	0.29
128	AMP	0.782	AMP	0.517	AMP	0.902
129	Non-AMP	0.124	AMP	0.594	Non-AMP	0.106
130	AMP	1	AMP	0.676	AMP	0.999
131	AMP	0.614	Non-AMP	0.229	Non-AMP	0.132
132	AMP	0.88	AMP	0.547	AMP	0.959
133	AMP	0.695	AMP	0.8245	Non-AMP	0.135
134	AMP	0.722	Non-AMP	0.4485	AMP	0.786
135	AMP	0.59	AMP	0.7085	AMP	0.883
136	Non-AMP	0.216	Non-AMP	0.4435	AMP	0.795
137	Non-AMP	0.165	Non-AMP	0.25	AMP	0.713
138	AMP	0.945	AMP	0.6105	AMP	0.903
139	AMP	0.896	AMP	0.5505	AMP	0.873

140	Non-AMP	0.007	AMP	0.555	AMP	0.982
141	AMP	0.659	AMP	0.783	AMP	0.991
142	Non-AMP	0.402	AMP	0.5395	Non-AMP	0.494
143	AMP	0.813	AMP	0.7335	AMP	0.922
144	AMP	0.998	AMP	0.5475	AMP	0.89
145	AMP	0.919	AMP	0.5265	AMP	1
146	AMP	0.959	AMP	0.877	AMP	0.784
147	AMP	0.969	AMP	0.6335	AMP	0.987
148	AMP	0.757	AMP	0.8615	AMP	0.681
149	Non-AMP	0.146	Non-AMP	0.32	Non-AMP	0.082
150	AMP	0.891	AMP	0.84	AMP	0.913
151	AMP	0.542	AMP	0.5955	AMP	0.582
152	AMP	0.894	AMP	0.7325	AMP	0.876
153	Non-AMP	0.29	Non-AMP	0.413	AMP	0.947
154	AMP	0.992	AMP	0.848	AMP	0.95
155	AMP	0.994	AMP	0.625	AMP	0.997
156	AMP	0.542	AMP	0.533	Non-AMP	0.244
157	AMP	0.511	AMP	0.617	AMP	0.999
158	AMP	0.99	AMP	0.823	AMP	0.992
159	AMP	0.679	AMP	0.554	AMP	0.512
160	AMP	0.58	AMP	0.6415	AMP	0.77
161	AMP	0.977	AMP	0.7245	AMP	0.9
162	AMP	0.983	AMP	0.646	AMP	0.981
163	Non-AMP	0.021	Non-AMP	0.3195	AMP	0.601
164	AMP	0.694	Non-AMP	0.584	Non-AMP	0.103
165	Non-AMP	0.849	Non-AMP	0.968	Non-AMP	1.268
166	Non-AMP	0.924	Non-AMP	0.968	Non-AMP	1.097
167	Non-AMP	0.879	Non-AMP	0.776	Non-AMP	-0.087
168	AMP	0.663	Non-AMP	0.658	Non-AMP	0.024
169	Non-AMP	0.65	Non-AMP	0.916	AMP	-0.571
170	AMP	0.757	AMP	0.6	Non-AMP	-0.165
171	Non-AMP	0.895	Non-AMP	0.802	Non-AMP	-0.145
172	Non-AMP	0.891	Non-AMP	0.962	Non-AMP	1.198

173	Non-AMP	0.761	Non-AMP	0.964	Non-AMP	0.55
174	Non-AMP	0.805	Non-AMP	0.768	Non-AMP	0.647
175	Non-AMP	0.9	Non-AMP	0.95	Non-AMP	0.733
176	Non-AMP	0.901	Non-AMP	0.702	Non-AMP	0.33
177	Non-AMP	0.969	Non-AMP	0.87	Non-AMP	0.151
178	AMP	0.797	AMP	0.702	AMP	-1.048
179	Non-AMP	0.675	Non-AMP	0.742	Non-AMP	0.531
180	Non-AMP	0.882	Non-AMP	0.808	Non-AMP	1.232
181	Non-AMP	0.983	Non-AMP	0.988	Non-AMP	1.427
182	Non-AMP	0.894	Non-AMP	0.728	Non-AMP	0.667
183	Non-AMP	0.71	Non-AMP	0.962	Non-AMP	0.064
184	Non-AMP	0.967	Non-AMP	0.996	Non-AMP	0.75
185	Non-AMP	0.838	Non-AMP	0.94	Non-AMP	0.798
186	Non-AMP	0.955	Non-AMP	0.996	Non-AMP	1.493
187	Non-AMP	0.937	Non-AMP	0.848	Non-AMP	1.119
188	AMP	0.508	Non-AMP	0.656	AMP	-0.384
189	Non-AMP	0.954	Non-AMP	0.998	Non-AMP	0.197
190	Non-AMP	0.92	Non-AMP	0.838	Non-AMP	1.129
191	Non-AMP	0.734	Non-AMP	0.718	Non-AMP	0.534
192	Non-AMP	0.986	Non-AMP	0.964	Non-AMP	1.553
193	Non-AMP	0.912	Non-AMP	0.754	Non-AMP	1.192
194	Non-AMP	0.883	Non-AMP	0.944	Non-AMP	2.03
195	Non-AMP	0.964	Non-AMP	0.998	Non-AMP	1.328
196	Non-AMP	0.983	Non-AMP	0.988	Non-AMP	1.427
197	Non-AMP	0.977	Non-AMP	0.996	Non-AMP	1.391
198	Non-AMP	0.882	Non-AMP	0.966	Non-AMP	1.263
199	Non-AMP	0.996	Non-AMP	0.916	Non-AMP	0.997
200	Non-AMP	0.71	Non-AMP	0.95	Non-AMP	0.306
201	Non-AMP	0.879	Non-AMP	0.972	Non-AMP	0.321

202	Non-AMP	0.955	Non-AMP	0.996	Non-AMP	1.493
203	AMP	0.528	AMP	0.51	AMP	-0.318
204	Non-AMP	0.924	Non-AMP	0.968	Non-AMP	1.097
205	Non-AMP	0.98	Non-AMP	0.974	Non-AMP	1.213
206	Non-AMP	0.674	Non-AMP	0.904	AMP	-0.54
207	Non-AMP	0.979	Non-AMP	0.958	Non-AMP	0.298
208	Non-AMP	0.945	Non-AMP	0.956	Non-AMP	1.233
209	Non-AMP	0.776	Non-AMP	0.838	Non-AMP	0.283
210	Non-AMP	0.944	Non-AMP	0.932	Non-AMP	1.315
211	Non-AMP	0.779	Non-AMP	0.96	Non-AMP	1.073
212	Non-AMP	0.969	Non-AMP	0.962	Non-AMP	1.304
213	Non-AMP	0.971	Non-AMP	0.988	Non-AMP	0.966
214	Non-AMP	0.854	Non-AMP	0.964	Non-AMP	0.569
215	Non-AMP	0.937	Non-AMP	0.966	Non-AMP	1.194
216	Non-AMP	0.976	Non-AMP	0.9	Non-AMP	1.902
217	AMP	0.683	Non-AMP	0.582	AMP	-0.367
218	Non-AMP	0.953	Non-AMP	0.998	Non-AMP	1.859
219	Non-AMP	0.945	Non-AMP	0.948	Non-AMP	1.157
220	Non-AMP	0.706	Non-AMP	0.674	Non-AMP	0.303
221	Non-AMP	0.925	Non-AMP	0.962	Non-AMP	1.493
222	Non-AMP	0.887	Non-AMP	0.826	Non-AMP	0.13
223	Non-AMP	0.779	Non-AMP	0.96	Non-AMP	1.073
224	AMP	0.508	Non-AMP	0.656	AMP	-0.384
225	Non-AMP	0.85	Non-AMP	0.962	Non-AMP	0.658
226	Non-AMP	0.718	Non-AMP	0.764	AMP	-0.262
227	Non-AMP	0.912	Non-AMP	0.986	Non-AMP	1.148
228	Non-AMP	0.77	Non-AMP	0.97	Non-AMP	0.522
229	Non-AMP	0.85	Non-AMP	0.95	Non-AMP	0.337
230	AMP	0.601	AMP	0.838	Non-AMP	0.062

231	Non-AMP	0.85	Non-AMP	0.786	Non-AMP	0.489
232	Non-AMP	0.85	Non-AMP	0.786	Non-AMP	0.489
233	Non-AMP	0.912	Non-AMP	0.754	Non-AMP	1.192
234	Non-AMP	0.947	Non-AMP	0.776	Non-AMP	0.509
235	Non-AMP	0.798	Non-AMP	0.802	Non-AMP	0.401
236	Non-AMP	0.649	Non-AMP	0.964	Non-AMP	0.152
237	Non-AMP	0.967	Non-AMP	0.99	Non-AMP	1.424
238	Non-AMP	0.926	Non-AMP	0.99	Non-AMP	0.635
239	AMP	0.732	AMP	0.514	AMP	-1.296
240	Non-AMP	0.719	Non-AMP	0.872	Non-AMP	0.516
241	Non-AMP	0.959	Non-AMP	0.978	Non-AMP	0.355
242	Non-AMP	0.769	Non-AMP	0.918	Non-AMP	0.77
243	Non-AMP	0.991	Non-AMP	0.922	Non-AMP	1.951
244	AMP	0.524	AMP	0.59	AMP	-0.625
245	Non-AMP	0.912	Non-AMP	0.95	Non-AMP	0.648
246	Non-AMP	0.776	Non-AMP	0.838	Non-AMP	0.283
247	Non-AMP	0.923	Non-AMP	0.94	Non-AMP	1.281
248	Non-AMP	0.899	Non-AMP	0.876	Non-AMP	0.604
249	Non-AMP	0.745	Non-AMP	0.936	Non-AMP	0.501
250	Non-AMP	0.95	Non-AMP	0.962	Non-AMP	1.183
251	AMP	0.745	Non-AMP	0.828	AMP	-1.023
252	Non-AMP	0.962	Non-AMP	0.846	Non-AMP	0.168
253	Non-AMP	0.872	Non-AMP	0.946	Non-AMP	0.987
254	Non-AMP	0.988	Non-AMP	0.782	Non-AMP	0.255
255	Non-AMP	0.806	AMP	0.552	AMP	-1.01
256	Non-AMP	0.908	Non-AMP	0.926	Non-AMP	0.293
257	Non-AMP	0.703	Non-AMP	0.672	Non-AMP	-0.15
258	Non-AMP	0.942	Non-AMP	0.798	Non-AMP	1.103
259	Non-AMP	0.972	Non-AMP	0.936	Non-AMP	1.317

260	Non-AMP	0.957	Non-AMP	0.874	Non-AMP	0.81
261	Non-AMP	0.985	Non-AMP	0.986	Non-AMP	1.744
262	Non-AMP	0.962	Non-AMP	0.974	Non-AMP	1.803
263	Non-AMP	0.959	Non-AMP	1	Non-AMP	0.811
264	Non-AMP	0.985	Non-AMP	0.986	Non-AMP	1.744
265	Non-AMP	0.926	Non-AMP	0.85	Non-AMP	-0.193
266	Non-AMP	0.888	Non-AMP	0.986	Non-AMP	0.759
267	Non-AMP	0.776	Non-AMP	0.838	Non-AMP	0.283
268	Non-AMP	0.977	Non-AMP	0.93	Non-AMP	1.502
269	Non-AMP	0.911	Non-AMP	0.832	Non-AMP	0.586
270	Non-AMP	0.941	Non-AMP	0.95	Non-AMP	1.202
271	Non-AMP	0.972	Non-AMP	0.97	Non-AMP	0.684
272	Non-AMP	0.964	Non-AMP	0.956	Non-AMP	1.861
273	Non-AMP	0.884	Non-AMP	0.602	AMP	-0.363
274	Non-AMP	0.799	Non-AMP	0.982	Non-AMP	0.93
275	Non-AMP	0.989	Non-AMP	1	Non-AMP	1.687
276	Non-AMP	0.911	Non-AMP	0.832	Non-AMP	0.586
277	Non-AMP	0.934	Non-AMP	0.976	Non-AMP	1.124
278	Non-AMP	0.977	Non-AMP	0.93	Non-AMP	1.502
279	Non-AMP	0.937	Non-AMP	0.966	Non-AMP	1.194
280	Non-AMP	0.969	Non-AMP	0.98	Non-AMP	1.155
281	Non-AMP	0.988	Non-AMP	0.82	Non-AMP	1.681
282	AMP	0.704	Non-AMP	0.616	Non-AMP	0.079
283	Non-AMP	0.96	Non-AMP	0.982	Non-AMP	0.447
284	Non-AMP	0.875	Non-AMP	0.734	Non-AMP	0.393
285	Non-AMP	0.945	Non-AMP	0.928	Non-AMP	0.981
286	Non-AMP	0.754	Non-AMP	0.906	Non-AMP	0.027
287	Non-AMP	0.938	Non-AMP	0.7	Non-AMP	0.474
288	Non-	0.944	Non-	0.822	Non-	0.415

	AMP		AMP		AMP	
289	Non-AMP	0.823	Non-AMP	0.672	Non-AMP	0.1
290	Non-AMP	0.993	Non-AMP	0.842	Non-AMP	0.586
291	Non-AMP	0.97	Non-AMP	0.926	Non-AMP	1.321
292	AMP	0.908	AMP	0.776	AMP	-1.571
293	Non-AMP	0.691	Non-AMP	0.982	Non-AMP	0.401
294	Non-AMP	0.988	Non-AMP	0.82	Non-AMP	1.681
295	Non-AMP	0.966	Non-AMP	0.842	Non-AMP	1.848
296	Non-AMP	0.961	Non-AMP	0.892	Non-AMP	0.608
297	Non-AMP	0.979	Non-AMP	0.996	Non-AMP	1.168
298	Non-AMP	0.988	Non-AMP	0.918	Non-AMP	1.777
299	Non-AMP	0.808	Non-AMP	0.942	Non-AMP	1.374
300	Non-AMP	0.881	Non-AMP	0.978	Non-AMP	0.583
301	Non-AMP	0.874	Non-AMP	0.926	Non-AMP	2.051
302	Non-AMP	0.935	Non-AMP	0.838	Non-AMP	0.363
303	Non-AMP	0.702	Non-AMP	0.694	Non-AMP	-0.187
304	Non-AMP	0.796	Non-AMP	0.658	AMP	-0.304
305	Non-AMP	0.962	Non-AMP	0.926	Non-AMP	1.182
306	Non-AMP	0.808	Non-AMP	0.944	Non-AMP	1.363
307	Non-AMP	0.917	Non-AMP	0.96	Non-AMP	0.998
308	Non-AMP	0.728	Non-AMP	0.816	AMP	-0.801
309	Non-AMP	0.866	Non-AMP	0.884	Non-AMP	0.433
310	AMP	0.51	Non-AMP	0.588	Non-AMP	0.117
311	Non-AMP	0.866	Non-AMP	0.816	Non-AMP	1.052
312	Non-AMP	0.873	Non-AMP	0.97	Non-AMP	1.331
313	Non-AMP	0.972	Non-AMP	0.926	Non-AMP	1.156
314	Non-AMP	0.675	Non-AMP	0.588	Non-AMP	0.255
315	Non-AMP	0.979	Non-AMP	0.992	Non-AMP	1.321
316	Non-AMP	0.839	AMP	0.628	Non-AMP	-0.101

317	Non-AMP	0.984	Non-AMP	0.838	Non-AMP	0.444
318	AMP	0.732	AMP	0.514	AMP	-1.296
319	Non-AMP	0.85	Non-AMP	0.844	Non-AMP	0.577
320	Non-AMP	0.955	Non-AMP	0.934	Non-AMP	0.705
321	AMP	0.519	Non-AMP	0.584	AMP	-0.474
322	Non-AMP	0.973	Non-AMP	0.966	Non-AMP	1.026
323	Non-AMP	0.913	Non-AMP	0.992	Non-AMP	1.019
324	Non-AMP	0.808	Non-AMP	0.942	Non-AMP	1.374
325	Non-AMP	0.944	Non-AMP	0.994	Non-AMP	1.489
326	Non-AMP	0.962	Non-AMP	0.958	Non-AMP	1.129
327	Non-AMP	0.973	Non-AMP	0.896	Non-AMP	0.976
328	Non-AMP	0.918	Non-AMP	0.92	Non-AMP	1.124
329	Non-AMP	0.596	Non-AMP	0.74	Non-AMP	0.15
330	Non-AMP	0.965	Non-AMP	0.564	Non-AMP	-0.073
331	AMP	0.704	Non-AMP	0.616	Non-AMP	0.079
332	Non-AMP	0.994	Non-AMP	0.998	Non-AMP	1.543
333	Non-AMP	0.931	Non-AMP	0.782	Non-AMP	1.147
334	Non-AMP	0.74	Non-AMP	0.884	Non-AMP	0.644
335	Non-AMP	0.778	Non-AMP	0.956	Non-AMP	0.5
336	Non-AMP	0.926	Non-AMP	0.876	Non-AMP	0.822
337	Non-AMP	0.891	Non-AMP	0.962	Non-AMP	1.198
338	Non-AMP	0.981	Non-AMP	0.916	Non-AMP	1.522
339	Non-AMP	0.97	Non-AMP	0.998	Non-AMP	1.225
340	Non-AMP	0.888	Non-AMP	0.96	Non-AMP	0.223
341	Non-AMP	0.924	Non-AMP	0.914	Non-AMP	0.908
342	Non-AMP	0.98	Non-AMP	0.97	Non-AMP	1.094
343	Non-AMP	0.799	Non-AMP	0.982	Non-AMP	0.93
344	Non-AMP	0.914	Non-AMP	0.962	Non-AMP	1.416
345	Non-AMP	0.959	Non-AMP	0.86	Non-AMP	1.51

346	Non-AMP	0.924	Non-AMP	0.946	Non-AMP	0.515
347	AMP	0.593	Non-AMP	0.52	AMP	-0.629
348	Non-AMP	0.53	Non-AMP	0.69	Non-AMP	0.278
349	Non-AMP	0.964	Non-AMP	0.774	Non-AMP	0.644
350	Non-AMP	0.732	Non-AMP	0.536	AMP	-0.664
351	Non-AMP	0.821	Non-AMP	0.968	Non-AMP	1.073
352	Non-AMP	0.896	Non-AMP	0.964	Non-AMP	1.151
353	Non-AMP	0.857	Non-AMP	0.98	Non-AMP	1.036
354	Non-AMP	0.976	Non-AMP	0.79	Non-AMP	1.389
355	AMP	0.83	AMP	0.818	AMP	-1.247
356	Non-AMP	0.981	Non-AMP	0.916	Non-AMP	1.522
357	Non-AMP	0.975	Non-AMP	0.552	Non-AMP	1.233
358	Non-AMP	0.879	Non-AMP	0.936	Non-AMP	0.453
359	Non-AMP	0.85	Non-AMP	0.648	Non-AMP	0.126
360	Non-AMP	0.934	Non-AMP	0.946	Non-AMP	1.123
361	Non-AMP	0.868	Non-AMP	0.938	Non-AMP	0.894
362	Non-AMP	0.99	AMP	0.878	Non-AMP	0.704
363	Non-AMP	0.835	Non-AMP	0.992	Non-AMP	0.317
364	AMP	0.538	Non-AMP	0.942	Non-AMP	0.2
365	AMP	0.509	Non-AMP	0.732	AMP	-0.54
366	Non-AMP	0.868	Non-AMP	0.95	Non-AMP	1.538
367	Non-AMP	0.935	Non-AMP	0.838	Non-AMP	0.363
368	Non-AMP	0.833	Non-AMP	0.952	Non-AMP	1.174
369	Non-AMP	0.959	Non-AMP	0.86	Non-AMP	1.51
370	Non-AMP	0.842	Non-AMP	0.896	Non-AMP	0.326
371	Non-AMP	0.655	Non-AMP	0.888	Non-AMP	1.088
372	Non-AMP	0.898	Non-AMP	0.942	Non-AMP	1.349
373	Non-AMP	0.885	Non-AMP	0.758	Non-AMP	0.469
374	Non-AMP	0.873	Non-AMP	0.97	Non-AMP	1.331

375	Non-AMP	0.937	Non-AMP	0.788	Non-AMP	1.865
376	Non-AMP	0.675	Non-AMP	0.742	Non-AMP	0.531
377	Non-AMP	0.966	Non-AMP	0.842	Non-AMP	1.848
378	Non-AMP	0.807	AMP	0.526	Non-AMP	-0.106
379	Non-AMP	0.717	Non-AMP	0.714	AMP	-0.856
380	Non-AMP	0.932	Non-AMP	0.984	Non-AMP	0.926
381	Non-AMP	0.933	Non-AMP	0.992	Non-AMP	0.801
382	Non-AMP	0.992	Non-AMP	0.91	Non-AMP	0.209
383	Non-AMP	0.99	Non-AMP	0.892	Non-AMP	0.185
384	Non-AMP	0.991	Non-AMP	0.922	Non-AMP	1.951
385	Non-AMP	0.976	Non-AMP	0.79	Non-AMP	1.389
386	Non-AMP	0.969	Non-AMP	0.988	Non-AMP	1.319
387	Non-AMP	0.695	Non-AMP	0.904	Non-AMP	1.82
388	Non-AMP	0.874	Non-AMP	0.926	Non-AMP	2.051
389	Non-AMP	0.948	Non-AMP	0.988	Non-AMP	1.225
390	AMP	0.793	Non-AMP	0.588	AMP	-0.487
391	Non-AMP	0.983	Non-AMP	0.546	Non-AMP	-0.091
392	AMP	0.856	AMP	0.656	AMP	-1.341
393	Non-AMP	0.64	Non-AMP	0.59	AMP	-0.607
394	Non-AMP	0.883	Non-AMP	0.944	Non-AMP	2.03
395	Non-AMP	0.852	AMP	0.64	AMP	-0.299
396	Non-AMP	0.949	Non-AMP	1	Non-AMP	0.767
397	Non-AMP	0.894	Non-AMP	0.942	Non-AMP	1.895
398	Non-AMP	0.699	AMP	0.528	Non-AMP	0.517
399	Non-AMP	0.901	Non-AMP	0.942	Non-AMP	0.858
400	Non-AMP	0.952	Non-AMP	0.988	Non-AMP	0.951
401	Non-AMP	0.584	AMP	0.65	Non-AMP	1.516
402	Non-AMP	0.653	Non-AMP	0.594	Non-AMP	-0.088
403	Non-AMP	0.953	AMP	0.556	Non-AMP	0.161

404	Non-AMP	0.943	Non-AMP	0.752	Non-AMP	1.074
405	Non-AMP	0.965	Non-AMP	0.918	Non-AMP	1.113
406	Non-AMP	0.853	Non-AMP	0.778	Non-AMP	1.085
407	Non-AMP	0.68	Non-AMP	0.9	Non-AMP	1.191
408	Non-AMP	0.959	Non-AMP	0.928	Non-AMP	0.774
409	Non-AMP	0.918	Non-AMP	0.994	Non-AMP	0.463
410	Non-AMP	0.993	Non-AMP	0.986	Non-AMP	1.446
411	Non-AMP	0.878	Non-AMP	0.962	Non-AMP	1.181
412	Non-AMP	0.868	Non-AMP	0.938	Non-AMP	0.894
413	Non-AMP	0.534	AMP	0.502	AMP	-0.768
414	Non-AMP	0.949	Non-AMP	1	Non-AMP	0.767
415	AMP	0.639	AMP	0.522	AMP	-0.703
416	AMP	0.761	AMP	0.582	AMP	-0.763
417	Non-AMP	0.97	Non-AMP	0.926	Non-AMP	1.321
418	Non-AMP	0.728	Non-AMP	0.816	AMP	-0.801
419	AMP	0.82	AMP	0.648	AMP	-0.481
420	Non-AMP	0.992	Non-AMP	0.772	Non-AMP	1.586
421	Non-AMP	0.994	Non-AMP	0.986	Non-AMP	1.83
422	Non-AMP	0.834	Non-AMP	0.62	Non-AMP	0.462
423	Non-AMP	0.653	Non-AMP	0.526	AMP	-0.487
424	Non-AMP	0.571	Non-AMP	0.77	Non-AMP	1.232
425	Non-AMP	0.769	Non-AMP	0.942	Non-AMP	0.871
426	Non-AMP	0.763	Non-AMP	0.614	Non-AMP	0.075
427	Non-AMP	0.999	Non-AMP	0.998	Non-AMP	2.238
428	Non-AMP	0.841	Non-AMP	0.96	Non-AMP	1.885
429	Non-AMP	0.808	AMP	0.504	Non-AMP	-0.144
430	Non-AMP	0.917	AMP	0.792	Non-AMP	-0.041
431	AMP	0.81	AMP	0.732	AMP	-1.952
432	AMP	0.745	AMP	0.614	AMP	-0.922
433	Non-AMP	0.834	Non-AMP	0.988	Non-AMP	0.137

434	Non-AMP	0.98	Non-AMP	0.902	Non-AMP	1.789
435	Non-AMP	0.521	AMP	0.904	Non-AMP	0.261
436	Non-AMP	0.995	Non-AMP	0.992	Non-AMP	1.947
437	AMP	0.724	Non-AMP	0.646	Non-AMP	0.825
438	Non-AMP	0.957	Non-AMP	0.944	Non-AMP	1.059
439	AMP	0.5	AMP	0.862	Non-AMP	0.79
440	Non-AMP	0.842	Non-AMP	0.978	Non-AMP	0.699
441	Non-AMP	0.799	Non-AMP	0.854	Non-AMP	-0.104
442	Non-AMP	0.994	Non-AMP	0.986	Non-AMP	1.83
443	Non-AMP	0.882	Non-AMP	0.884	Non-AMP	0.271
444	Non-AMP	0.967	Non-AMP	0.986	Non-AMP	1.229
445	Non-AMP	0.973	Non-AMP	0.992	Non-AMP	1.154
446	AMP	0.803	Non-AMP	0.782	AMP	-0.493
447	Non-AMP	0.81	Non-AMP	0.776	Non-AMP	0.838
448	Non-AMP	0.98	Non-AMP	0.902	Non-AMP	1.789
449	Non-AMP	0.655	Non-AMP	0.888	Non-AMP	1.088
450	Non-AMP	0.725	Non-AMP	0.59	AMP	-0.88
451	Non-AMP	0.974	Non-AMP	0.624	Non-AMP	1.601
452	Non-AMP	0.841	Non-AMP	0.96	Non-AMP	1.885
453	Non-AMP	0.964	Non-AMP	0.976	Non-AMP	1.196
454	Non-AMP	0.547	Non-AMP	0.804	Non-AMP	1.239
455	Non-AMP	0.894	Non-AMP	0.942	Non-AMP	1.895
456	Non-AMP	0.997	Non-AMP	0.962	Non-AMP	0.896
457	Non-AMP	0.943	Non-AMP	0.994	Non-AMP	0.597
458	Non-AMP	0.738	AMP	0.62	Non-AMP	0.266
459	Non-AMP	0.998	Non-AMP	0.97	Non-AMP	2.274
460	AMP	0.845	AMP	0.738	AMP	-1.699
461	Non-AMP	0.935	Non-AMP	0.96	Non-AMP	0.564
462	Non-AMP	0.932	Non-AMP	0.928	Non-AMP	0.308

463	Non-AMP	0.976	Non-AMP	0.832	Non-AMP	1.624
464	Non-AMP	0.989	Non-AMP	0.9	Non-AMP	0.928
465	AMP	0.801	Non-AMP	0.558	AMP	-1.609
466	Non-AMP	0.72	Non-AMP	0.662	AMP	-1.343
467	AMP	0.531	AMP	0.592	AMP	-0.653
468	Non-AMP	0.96	Non-AMP	0.984	Non-AMP	0.973
469	Non-AMP	0.882	Non-AMP	0.966	Non-AMP	1.263
470	Non-AMP	0.591	AMP	0.772	Non-AMP	0.935
471	Non-AMP	0.994	Non-AMP	0.612	AMP	-1.77
472	AMP	0.915	AMP	0.678	AMP	-1.325
473	AMP	0.793	Non-AMP	0.588	AMP	-0.487
474	Non-AMP	0.766	Non-AMP	0.678	Non-AMP	-0.138
475	Non-AMP	0.965	AMP	0.558	Non-AMP	-0.073
476	Non-AMP	0.988	Non-AMP	0.918	Non-AMP	1.777
477	Non-AMP	0.975	Non-AMP	0.984	Non-AMP	1.685
478	AMP	0.841	Non-AMP	0.63	Non-AMP	-0.067
479	Non-AMP	0.969	Non-AMP	0.998	Non-AMP	1.746
480	Non-AMP	0.989	Non-AMP	1	Non-AMP	1.687
481	Non-AMP	0.935	Non-AMP	0.79	Non-AMP	0.658
482	Non-AMP	0.965	Non-AMP	1	Non-AMP	1.235
483	AMP	0.57	Non-AMP	0.616	Non-AMP	0.114
484	Non-AMP	0.983	Non-AMP	0.998	Non-AMP	1.571
485	Non-AMP	0.68	Non-AMP	0.732	Non-AMP	0.406
486	Non-AMP	0.924	Non-AMP	0.894	Non-AMP	0.81
487	Non-AMP	0.81	Non-AMP	0.96	Non-AMP	1.228
488	Non-AMP	0.878	Non-AMP	0.988	Non-AMP	0.709
489	AMP	0.889	Non-AMP	0.644	Non-AMP	1.017
490	Non-AMP	0.963	AMP	0.622	AMP	-1.312
491	Non-AMP	0.924	Non-AMP	0.504	AMP	-0.38

492	AMP	0.63	Non-AMP	0.748	Non-AMP	0.863
493	Non-AMP	0.754	Non-AMP	0.948	Non-AMP	0.995
494	Non-AMP	0.68	Non-AMP	0.9	Non-AMP	1.191
495	Non-AMP	0.982	Non-AMP	0.978	Non-AMP	0.964
496	Non-AMP	0.853	Non-AMP	0.904	Non-AMP	0.524
497	AMP	0.777	Non-AMP	0.57	AMP	-1.015
498	Non-AMP	0.964	Non-AMP	0.962	Non-AMP	1.301
499	Non-AMP	0.878	Non-AMP	0.988	Non-AMP	0.709
500	Non-AMP	0.748	Non-AMP	0.782	Non-AMP	-0.011
501	Non-AMP	0.717	Non-AMP	0.714	AMP	-0.856
502	Non-AMP	0.779	Non-AMP	0.544	Non-AMP	0.453
503	Non-AMP	0.96	Non-AMP	0.952	Non-AMP	1.222
504	Non-AMP	0.939	AMP	0.532	Non-AMP	-0.008
505	Non-AMP	0.949	Non-AMP	0.848	Non-AMP	0.571
506	Non-AMP	0.563	Non-AMP	0.548	AMP	-0.931
507	Non-AMP	0.858	Non-AMP	0.784	Non-AMP	0.461
508	Non-AMP	0.965	Non-AMP	0.96	Non-AMP	1.559
509	AMP	0.715	Non-AMP	0.672	AMP	-0.7
510	Non-AMP	0.967	Non-AMP	0.786	Non-AMP	1.592
511	Non-AMP	0.986	Non-AMP	0.742	AMP	-0.288
512	Non-AMP	0.999	Non-AMP	0.992	Non-AMP	2.458
513	Non-AMP	0.693	Non-AMP	0.79	Non-AMP	1.083
514	Non-AMP	0.895	Non-AMP	0.89	Non-AMP	0.719
515	Non-AMP	0.815	Non-AMP	0.868	Non-AMP	0.365
516	Non-AMP	0.983	Non-AMP	0.584	Non-AMP	0.324
517	Non-AMP	0.944	Non-AMP	0.94	Non-AMP	0.738
518	Non-AMP	0.922	Non-AMP	0.926	Non-AMP	0.913
519	Non-AMP	0.97	Non-AMP	0.626	AMP	-0.262
520	Non-	0.97	Non-	0.882	Non-	1.305

	AMP		AMP		AMP	
521	AMP	0.807	Non-AMP	0.576	AMP	-0.949
522	AMP	0.817	AMP	0.758	Non-AMP	0.449
523	Non-AMP	0.929	Non-AMP	0.966	Non-AMP	0.976
524	Non-AMP	0.911	Non-AMP	0.9	Non-AMP	0.746
525	Non-AMP	0.896	Non-AMP	0.964	Non-AMP	1.151
526	Non-AMP	0.864	AMP	0.766	AMP	-0.32
527	Non-AMP	0.84	Non-AMP	0.832	Non-AMP	0.821
528	Non-AMP	0.693	Non-AMP	0.79	Non-AMP	1.083
529	Non-AMP	0.905	Non-AMP	0.938	Non-AMP	0.022
530	Non-AMP	0.954	Non-AMP	0.916	Non-AMP	1.142
531	Non-AMP	0.988	Non-AMP	0.946	Non-AMP	1.114
532	Non-AMP	0.98	Non-AMP	0.978	Non-AMP	1.1
533	Non-AMP	0.916	Non-AMP	0.934	Non-AMP	0.717
534	Non-AMP	0.99	Non-AMP	0.992	Non-AMP	1.415
535	Non-AMP	0.916	Non-AMP	0.988	Non-AMP	0.884
536	AMP	0.763	AMP	0.668	AMP	-0.958
537	Non-AMP	0.965	Non-AMP	0.96	Non-AMP	1.559
538	Non-AMP	0.878	Non-AMP	0.94	Non-AMP	1.267
539	Non-AMP	0.927	Non-AMP	0.888	Non-AMP	0.681
540	Non-AMP	0.975	Non-AMP	0.97	Non-AMP	0.917
541	Non-AMP	0.965	Non-AMP	0.852	Non-AMP	1.932
542	Non-AMP	0.988	Non-AMP	0.978	Non-AMP	1.076
543	Non-AMP	0.9	Non-AMP	0.884	Non-AMP	0.724
544	Non-AMP	0.997	Non-AMP	0.978	Non-AMP	0.485
545	Non-AMP	0.832	Non-AMP	0.94	Non-AMP	1.731
546	Non-AMP	0.833	Non-AMP	0.952	Non-AMP	1.174
547	Non-AMP	0.979	AMP	0.64	Non-AMP	0.304
548	Non-AMP	0.962	Non-AMP	0.826	Non-AMP	0.636

549	Non-AMP	0.967	Non-AMP	0.86	Non-AMP	0.265
550	AMP	0.92	Non-AMP	0.606	Non-AMP	-0.061
551	Non-AMP	0.926	Non-AMP	0.876	Non-AMP	0.822
552	Non-AMP	0.947	Non-AMP	0.908	Non-AMP	1.513
553	Non-AMP	0.64	Non-AMP	0.952	Non-AMP	0.443
554	Non-AMP	0.776	AMP	0.512	AMP	-0.368
555	Non-AMP	0.868	Non-AMP	0.95	Non-AMP	1.538
556	Non-AMP	0.975	Non-AMP	0.984	Non-AMP	1.685
557	Non-AMP	0.938	Non-AMP	0.7	Non-AMP	0.474
558	Non-AMP	0.933	Non-AMP	0.88	Non-AMP	0.737
559	Non-AMP	0.7	AMP	0.708	Non-AMP	1.369
560	Non-AMP	0.885	Non-AMP	0.994	Non-AMP	1.734
561	Non-AMP	0.969	Non-AMP	1	Non-AMP	1.926
562	Non-AMP	0.901	Non-AMP	0.814	Non-AMP	0.79
563	Non-AMP	0.987	Non-AMP	0.97	Non-AMP	1.985
564	Non-AMP	0.91	Non-AMP	0.992	Non-AMP	0.864
565	Non-AMP	0.598	Non-AMP	0.692	Non-AMP	0.352
566	Non-AMP	0.986	Non-AMP	0.986	Non-AMP	1.324
567	Non-AMP	0.992	Non-AMP	0.996	Non-AMP	2.034
568	Non-AMP	0.995	AMP	0.6	AMP	-0.476
569	Non-AMP	0.896	Non-AMP	0.812	Non-AMP	0.829
570	Non-AMP	0.829	Non-AMP	0.794	Non-AMP	0.732
571	Non-AMP	0.976	Non-AMP	0.944	Non-AMP	0.044
572	Non-AMP	0.987	Non-AMP	0.986	Non-AMP	1.692
573	Non-AMP	0.798	Non-AMP	0.722	Non-AMP	0.499
574	Non-AMP	0.832	Non-AMP	0.94	Non-AMP	1.731
575	Non-AMP	0.95	Non-AMP	0.626	Non-AMP	0.348
576	Non-AMP	0.764	Non-AMP	0.96	Non-AMP	0.281
577	Non-	0.882	Non-	0.624	Non-	0.068

	AMP		AMP		AMP	
578	Non-AMP	0.716	Non-AMP	0.818	Non-AMP	0.122
579	Non-AMP	0.926	Non-AMP	0.792	Non-AMP	1.212
580	AMP	0.52	AMP	0.5	Non-AMP	-0.086
581	Non-AMP	0.68	Non-AMP	0.608	Non-AMP	-0.084
582	Non-AMP	0.863	Non-AMP	0.888	Non-AMP	0.66
583	Non-AMP	0.826	AMP	0.544	AMP	-0.723
584	Non-AMP	0.9	Non-AMP	0.528	Non-AMP	0.71
585	Non-AMP	0.995	Non-AMP	0.996	Non-AMP	1.772
586	Non-AMP	0.987	Non-AMP	0.982	Non-AMP	0.641
587	Non-AMP	0.974	Non-AMP	0.624	Non-AMP	1.601
588	Non-AMP	0.963	Non-AMP	0.656	AMP	-0.374
589	Non-AMP	0.995	Non-AMP	0.996	Non-AMP	1.772
590	Non-AMP	0.682	Non-AMP	0.732	Non-AMP	0.732
591	Non-AMP	0.97	Non-AMP	0.998	Non-AMP	0.866
592	Non-AMP	0.994	Non-AMP	0.93	AMP	-0.262
593	Non-AMP	0.962	Non-AMP	0.908	Non-AMP	0.168
594	Non-AMP	0.879	Non-AMP	0.64	Non-AMP	-0.246
595	Non-AMP	0.886	Non-AMP	0.738	Non-AMP	1.384
596	Non-AMP	0.875	Non-AMP	0.96	Non-AMP	1.049
597	Non-AMP	0.966	Non-AMP	1	Non-AMP	1.042
598	Non-AMP	0.986	Non-AMP	0.998	Non-AMP	2.078
599	Non-AMP	0.987	Non-AMP	0.964	Non-AMP	0.842
600	AMP	0.524	Non-AMP	0.684	Non-AMP	-0.041
601	Non-AMP	0.911	Non-AMP	0.9	Non-AMP	0.746
602	Non-AMP	0.986	Non-AMP	0.992	Non-AMP	1.455
603	Non-AMP	0.772	Non-AMP	0.658	Non-AMP	0.344
604	Non-AMP	0.951	Non-AMP	0.93	Non-AMP	0.94
605	Non-AMP	0.913	Non-AMP	0.982	Non-AMP	1.01

606	Non-AMP	0.9	Non-AMP	0.976	Non-AMP	1.65
607	Non-AMP	0.861	Non-AMP	0.908	Non-AMP	2.271
608	Non-AMP	0.841	Non-AMP	0.952	Non-AMP	1.272
609	Non-AMP	0.935	Non-AMP	0.838	Non-AMP	0.363
610	Non-AMP	0.921	Non-AMP	0.958	Non-AMP	0.235
611	Non-AMP	0.987	Non-AMP	0.986	Non-AMP	1.692
612	Non-AMP	0.747	AMP	0.628	Non-AMP	1.489
613	Non-AMP	0.966	Non-AMP	1	Non-AMP	1.042
614	Non-AMP	0.976	Non-AMP	0.9	Non-AMP	1.902
615	AMP	0.938	AMP	0.702	AMP	-1.296
616	Non-AMP	0.912	Non-AMP	1	Non-AMP	1.533
617	Non-AMP	0.623	Non-AMP	0.582	Non-AMP	-0.031
618	Non-AMP	0.969	Non-AMP	0.988	Non-AMP	1.706
619	Non-AMP	0.955	Non-AMP	0.982	Non-AMP	0.597
620	Non-AMP	0.999	Non-AMP	0.992	Non-AMP	2.458
621	Non-AMP	0.965	Non-AMP	0.938	Non-AMP	1.5
622	Non-AMP	0.964	Non-AMP	0.99	Non-AMP	1.03
623	Non-AMP	0.983	Non-AMP	0.978	Non-AMP	1.439
624	Non-AMP	0.878	Non-AMP	0.888	Non-AMP	0.874
625	Non-AMP	0.641	Non-AMP	0.504	AMP	-0.825
626	Non-AMP	0.71	Non-AMP	0.606	Non-AMP	0.788
627	Non-AMP	0.991	Non-AMP	0.992	Non-AMP	1.741
628	Non-AMP	0.925	Non-AMP	0.962	Non-AMP	1.493
629	Non-AMP	0.898	Non-AMP	0.942	Non-AMP	1.349
630	Non-AMP	0.908	Non-AMP	0.986	Non-AMP	1.004
631	Non-AMP	0.799	Non-AMP	0.978	Non-AMP	0.433
632	Non-AMP	0.907	Non-AMP	0.954	Non-AMP	1.02
633	Non-AMP	0.934	Non-AMP	0.664	Non-AMP	0.331
634	Non-AMP	0.904	Non-AMP	0.788	Non-AMP	1.457

635	Non-AMP	0.912	Non-AMP	0.694	Non-AMP	0.756
636	Non-AMP	0.598	Non-AMP	0.782	Non-AMP	-0.238
637	Non-AMP	0.999	AMP	0.712	Non-AMP	0.669
638	Non-AMP	0.803	AMP	0.522	Non-AMP	-0.061
639	Non-AMP	0.801	Non-AMP	0.842	Non-AMP	0.223
640	AMP	0.892	AMP	0.544	AMP	-0.617
641	Non-AMP	0.884	Non-AMP	0.848	Non-AMP	1.042
642	Non-AMP	0.599	Non-AMP	0.54	AMP	-0.853
643	Non-AMP	0.974	Non-AMP	0.842	Non-AMP	0.776
644	Non-AMP	0.763	Non-AMP	0.742	Non-AMP	0.543
645	Non-AMP	0.942	Non-AMP	0.944	Non-AMP	0.639
646	Non-AMP	0.992	Non-AMP	0.996	Non-AMP	2.034
647	Non-AMP	0.829	Non-AMP	0.886	Non-AMP	0.539
648	Non-AMP	0.942	Non-AMP	0.904	Non-AMP	0.927
649	Non-AMP	0.971	Non-AMP	0.854	Non-AMP	1.615
650	Non-AMP	0.743	AMP	0.59	AMP	-0.481
651	Non-AMP	0.998	Non-AMP	0.73	AMP	-0.658
652	Non-AMP	0.966	Non-AMP	0.978	Non-AMP	1.141
653	Non-AMP	0.95	Non-AMP	0.748	Non-AMP	1.234
654	Non-AMP	0.966	Non-AMP	0.848	Non-AMP	1.311
655	Non-AMP	0.991	Non-AMP	0.998	Non-AMP	1.902
656	Non-AMP	0.968	Non-AMP	0.88	Non-AMP	1.126
657	Non-AMP	0.943	Non-AMP	0.996	Non-AMP	0.564
658	AMP	0.742	AMP	0.74	AMP	-1.141
659	Non-AMP	0.555	AMP	0.678	AMP	-0.567
660	Non-AMP	0.967	Non-AMP	0.958	Non-AMP	0.964
661	AMP	0.591	Non-AMP	0.634	Non-AMP	1.21
662	AMP	0.646	AMP	0.506	AMP	-1.032
663	Non-AMP	0.945	Non-AMP	0.962	Non-AMP	1.073

664	Non-AMP	0.86	AMP	0.5	Non-AMP	1.872
665	Non-AMP	0.925	Non-AMP	0.888	Non-AMP	0.97
666	Non-AMP	0.918	Non-AMP	0.726	AMP	-0.341
667	Non-AMP	0.541	Non-AMP	0.61	Non-AMP	-0.034
668	Non-AMP	0.944	Non-AMP	0.986	Non-AMP	0.567
669	Non-AMP	0.551	Non-AMP	0.586	Non-AMP	0.143
670	Non-AMP	0.992	AMP	0.582	Non-AMP	1.066
671	Non-AMP	0.999	Non-AMP	1	Non-AMP	2.546
672	Non-AMP	0.909	Non-AMP	0.816	Non-AMP	1.237
673	AMP	0.76	AMP	0.748	Non-AMP	0.679
674	Non-AMP	0.93	Non-AMP	0.63	Non-AMP	0.367
675	Non-AMP	0.977	Non-AMP	0.986	Non-AMP	1.253
676	Non-AMP	0.954	Non-AMP	0.77	Non-AMP	1.293
677	Non-AMP	0.727	Non-AMP	0.716	Non-AMP	0.747
678	Non-AMP	0.642	AMP	0.57	AMP	-0.335
679	Non-AMP	0.991	Non-AMP	0.992	Non-AMP	1.741
680	Non-AMP	0.945	Non-AMP	0.948	Non-AMP	1.624
681	AMP	0.987	AMP	0.754	AMP	-2.088
682	Non-AMP	0.992	Non-AMP	0.708	Non-AMP	1.431
683	Non-AMP	0.889	Non-AMP	0.636	AMP	-0.774
684	Non-AMP	0.87	Non-AMP	0.944	Non-AMP	0.517
685	AMP	0.561	AMP	0.544	AMP	-0.437
686	Non-AMP	0.915	Non-AMP	0.608	Non-AMP	0.467
687	Non-AMP	0.978	Non-AMP	0.97	Non-AMP	1.317
688	Non-AMP	0.953	Non-AMP	0.778	Non-AMP	1.97
689	Non-AMP	0.993	Non-AMP	0.992	Non-AMP	1.947
690	AMP	0.821	Non-AMP	0.386	AMP	0.877
691	Non-AMP	0.157	Non-AMP	0.4495	Non-AMP	0.077

Table 15: Physico-chemical Properties of all *de novo* Sequenced Peptides.
The length, molecular weight, charge, hydrophobicity and pI were determined for all 691 *de novo* sequenced peptides.

Seq. ID	Peptide Sequence	Length (res)	Molecular Weight (Da)	Net Charge	Hydrophobicity	pI
1	LDRLPLVRFKVTKKQLHERGDL	22	2662.18	3	-0.68	10.27
2	LQTKLKKLLGLESVF	15	1717.13	2	0.36	9.70
3	ALKALGNVGHPASL	14	1347.58	1	0.51	8.80
4	VERLPLVRF	9	1128.38	1	0.52	9.57
5	KVFADNLGEKTKA	13	1420.63	1	-0.68	8.50
6	GGSTFMVKSNKR	12	1311.52	3	-0.83	11.17
7	FKKVKEKLKDTFA	13	1581.92	3	-0.91	9.83
8	KLKKLLGLESVF	12	1374.73	2	0.48	9.70
9	DLKLKGVWQWW	11	1458.73	1	-0.55	8.59
10	KFTQRSLQKTQ	11	1364.57	3	-1.67	11.17
11	VERLPLVRFK	10	1256.56	2	0.08	10.83
12	KKTSMKLPLFTLR	13	1562.98	4	-0.3	11.26
13	FVLKSFAQARRY	12	1485.75	3	-0.11	11.00
14	KKTSMKLLPFTLR	13	1562.98	4	-0.3	11.26
15	KGGSTFETKSKLTN	14	1497.67	2	-1.19	9.70
16	KSRVNRMKQAR	11	1373.64	5	-1.93	12.31
17	KFTQRSLQKTQ	11	1364.57	3	-1.67	11.17
18	KRTFTPSQQ	9	1092.22	2	-1.82	11.00
19	QEQTRFGRE	9	1150.22	0	-2.37	6.14
20	KSRVNRFQQAR	11	1389.58	4	-1.81	12.30
21	LAKALGRGGHPASI	14	1347.58	2	0.16	11.00
22	KSRVNRFQQAR	11	1389.58	4	-1.81	12.30
23	VKDLSRQKNKL	11	1328.58	3	-1.43	10.29
24	GAKLEQFKENVKVF	14	1636.91	1	-0.46	8.50
25	KKTSMKLLPFTRL	13	1562.98	4	-0.3	11.26
26	ERSLFNKTQ	9	1122.25	1	-1.53	8.85
27	PALKYVRPGGGFAPNFLQ	18	1932.26	2	-0.09	10.01
28	NPEAKPNTKFGQKTMY	16	1854.11	2	-1.59	9.53
29	HAAFQKLVRQVAAALAAEYH	20	2194.52	1	0.24	8.60
30	KKTSMKLTKQTRL	13	1562.93	5	-1.31	11.33
31	KFLQKSVQKQPG	12	1387.65	3	-1.18	10.30
32	AAGFNKTLRKHANELL	16	1783.06	2	-0.47	9.99
33	NVAPKVRRLL	10	1165.45	3	-0.02	12.01
34	AYEHFKKVKEKLKDRY	16	2082.43	3	-1.73	9.60
35	KFSQRSVQKPSNG	13	1462.63	3	-1.55	11.17
36	RLVQLRLENSNA	12	1412.61	1	-0.53	9.60
37	RKTFTPSQQ	9	1092.22	2	-1.82	11.00
38	KHASKKNKKLS	11	1268.53	5	-2.02	10.60
39	ASFGEAVKHLDNLKGHFANI	20	2168.44	0	-0.13	6.96
40	EHFKKVKEKLK	11	1413.73	3	-1.72	9.83
41	FKKVKEKLKDTFA	13	1581.92	3	-0.91	9.83
42	LQTKLKLKLMQTKK	14	1701.19	5	-0.77	10.60
43	APVKNHTGSPSNKP	14	1433.59	2	-1.39	10.00

44	APVKNHTGPSSGGKP	15	1433.59	2	-1.12	10.00
45	VEKLKGVWKEQN	12	1457.69	1	-1.23	8.47
46	APLKNVPKVR	10	1121.39	3	-0.5	11.17
47	GWLFLQKSVRKPQG	14	1643.95	3	-0.63	11.17
48	LDRLPLVRFKSLKKQLHERGDEI	23	2791.29	2	-0.79	9.69
49	ALLDKFLQKSVQQKPG	16	1800.13	2	-0.52	9.70
50	TPWALLLRPKA	11	1265.56	2	0.16	11.00
51	QTYETKSLKTN	11	1312.44	1	-1.7	8.50
52	LDRLPLVRFKVTKKQLHERGDL	22	2662.18	3	-0.68	10.27
53	VLLPVLVTVV	10	1051.38	0	3.01	5.49
54	GAPFLGSCKQEKLASGRQ	18	1877.15	2	-0.59	9.31
55	LDRLPLVRFKSLKKQLHEGRDLEVF	25	3037.6	2	-0.48	9.69
56	GWLFLQKSVRKGPQ	14	1643.95	3	-0.63	11.17
57	HKALVLHTLAPL	12	1312.62	1	0.87	8.76
58	DDRHHYDEKRKM	12	1629.77	0	-3.05	6.92
59	RKLLLRPR	8	1051.35	4	-0.95	12.30
60	YKSRGSKRLLHKHV	14	1709.03	5	-1.33	11.17
61	VVPCLLGLLVAVAKAP	16	1563.02	1	2.03	8.19
62	KHALPPKKSPRVS	13	1444.74	4	-1.23	11.26
63	KHSKYKKSNKKL	12	1488.8	6	-2.43	10.40
64	MRHVCVGCDKPAFANTTGHKSAAFHTPPFPPATGT NGPYPMFPVCVPGVW	50	5298.14	2	-0.06	8.64
65	GVDLNRRRWFQK	12	1574.81	3	-1.53	11.71
66	HKASKKPVRSPPL	13	1444.74	4	-1.23	11.26
67	GAMLFLQKSVRKPQG	15	1660.01	3	-0.28	11.17
68	RPPGFPVMR	9	1056.29	2	-0.59	12.00
69	ARSNSGHNGPRYKLEQFKENVKVF	24	2806.14	3	-1.28	9.99
70	WPTVALVLGPKA	12	1251.53	1	0.88	8.75
71	KRFPAKKSPVRS	12	1400.69	5	-1.39	12.02
72	ASFGEAVKHLDNLKGHFNAL	20	2168.44	0	-0.16	6.96
73	FGLLVAPPRVLAS	13	1339.64	1	1.33	9.75
74	ARLDQGRQLREQL	13	1582.78	1	-1.4	9.56
75	HKRPVAPPSLKRAF	14	1603.93	4	-0.8	12.02
76	ARLFLMKRLKNKSS	14	1692.1	5	-0.56	12.02
77	LLTNTPKLKKAF	12	1373.7	3	-0.18	10.30
78	HKAVVFRAPPSALVR	15	1647.99	3	0.3	12.01
79	RLVQLRQNDSI	11	1341.53	1	-0.68	9.60
80	LWPLKRRWSHVLRL	14	1860.28	4	-0.39	12.30
81	PQKAPLPQKLARVS	14	1532.85	3	-0.68	11.17
82	AVLNSAKKPSVLPVRAF	17	1797.17	3	0.46	11.17
83	DKNKFSQRSVQKPSGGG	17	1820	3	-1.67	10.29
84	KNAQEKYGKGCWSQATR	17	1955.18	3	-1.68	9.63
85	KLRTFVYFG	9	1130.36	2	0.31	9.99
86	RKKQWGCHAGPGGHDSSGCQLMGWAKFPHVA HYVPWLRHTMVTWRAWHRCGFAAHSGNMAGCSKAKAKK	69	7698.92	10	-0.62	10.26
87	WHSNKTNHKKPGASGH	16	1785.94	3	-2.01	10.30
88	FRTCARKTVT	10	1182.41	3	-0.37	10.86
89	GGKAQEKYGKGDMNKSQL	18	1939.17	2	-1.63	9.40
90	LAKALGNRSFASLKRLM	17	1876.29	4	0.18	12.02
91	ALPKKVPNTLRVL	13	1448.82	3	0.15	11.17

92	DLVGWMCCCCCRAR	14	1619	1	0.74	7.83
93	KYVLSNKATLGCKKGNYGTPPFKGKKYLDGKVLGGPKPA	39	4124.9	8	-0.73	10.01
94	WLKKKFMHR	9	1273.61	4	-1.31	11.26
95	FGRKLSPPSALKKS	14	1515.82	4	-0.71	11.26
96	NQQGPKEEQQVHMR	14	1708.87	0	-2.29	6.76
97	RKCQWARQ	8	1075.26	3	-2.06	10.86
98	CLSTRTFWKL	10	1254.41	2	0.14	9.51
99	ARRRYLSHKKLGRRNRR	17	2223.62	9	-2.3	12.40
100	AQNRRTARAQKRLSQ	15	1784.01	5	-1.88	12.48
101	VHRRQMRRPPG	11	1389.65	4	-2.02	12.48
102	QAPRALKKPL	10	1121.39	3	-0.78	11.17
103	RLKKRRLFWQEFRRRKKK	18	2560.14	10	-2.29	12.31
104	RWVLSHQKLKPKK	13	1648.03	5	-1.41	11.33
105	YFQKHALKKKRCHPWWWWW	19	2715.22	5	-1.39	10.19
106	TFSVVKALFVVRPP	14	1559.92	2	1.06	11.00
107	FRRANSKRRRVPAKV	15	1841.2	7	-1.43	12.60
108	MECQHQQCSCCYKKRLLKLNRLEKKMNDAHHKRPLMV GLERTRLMAKKLKLRGDCCGDAADYCSGGCP	68	7813.3	8	-0.74	9.28
109	AVKHSKKNEKFHTRHKKG	18	2160.51	6	-2.09	10.68
110	RKKHLSAKKRKLVLFMHR	18	2276.87	8	-1.01	12.32
111	AFKLFKKKKCVGPGKPAFAAVRAHGKVQCGVWPKHL GRTTRLPGGGGKGVGVVPTVAKALV	61	6263.6	13	1.02	11.29
112	LQGMKTPNQLGKQ	13	1442.7	2	-1.18	10.00
113	LAPCVVVALVLVAKGP	16	1548.99	1	2.05	8.22
114	RWLKSRRWKKKTK	13	1801.22	8	-2.5	12.32
115	LLKKLVLKVLLKLMECLKCWVMQNRNSGFFPDNARQ AREVGGQQREADATWATGFCCMGSEDCCGCCC	68	7608.98	2	-0.04	8.15
116	VRHTCVVQCQL	11	1285.55	1	0.55	8.05
117	LWPKVVKHSKSRR	13	1620.96	5	-1.22	12.02
118	HTPVRKKTKPRVAF	14	1665.02	5	-1.11	12.02
119	WVVPWALVPVAVKA	14	1534.91	1	1.52	8.75
120	LAPLWSVALVPTLRPAK	17	1832.26	2	0.79	11.00
121	LQREMVRGAWANCHKVKNANHLTCRNAGRVS KNCKAG	37	4122.78	7	-0.8	10.47
122	AKHGLPVPVPPI	12	1224.51	1	0.38	8.80
123	ALTHKVALALVVTV	14	1434.79	1	1.79	8.80
124	KWWHWTWLKWEKM	13	1945.32	2	-1.38	9.70
125	RRRRKYWWCLWYFMMYEEYKWYY	23	3507.11	4	-1.36	9.58
126	LAKALGNVGHPASLKCCSKLSSSMRRKAKAK	31	3241.93	8	-0.33	10.67
127	HNTRLKKHLKNKS	13	1603.89	5	-2.11	11.33
128	APWLKGPAPPLVVKS	15	1559.92	2	0.22	10.00
129	TNQPNVKFWCLGMLGLKRKSTSQNWYRKRWFHRN KCWRWQEWSWVWRWKM	50	6645.77	11	-1.26	11.43
130	RKRKRLHRHYKKKHKH	16	2236.71	10	-3.23	12.03
131	ERYEWLYSCSNHDRKGYSRSG	21	2593.77	1	-1.77	8.25
132	WPRRRMCYQWWYKKKKKPWWLLLRRRRRHYV	31	4461.43	13	-1.74	11.81
133	WKCSRKEKQVWARKWSYNTFGTNWFCK RAVWQKQELKSFLS	41	5170.99	8	-1.05	10.24
134	YYPYFKRLLRSRLLRKRKMMW	21	2905.61	8	-0.93	11.49
135	LRESLGDKSKCVKRHKLSCKSKQSASKVKG	30	3316.93	8	-1.1	10.24
136	EPWFWWNHERRARPSDWRRKGWQW	24	3353.72	3	-2.18	11.42
137	QTKMDTKKAFLFHKSN	16	1924.25	3	-1.15	10.00

138	SKADHAWKKKRDRRLKRKK	19	2435.91	9	-2.62	11.61
139	WRRRRRLLMYMWW	13	2009.47	5	-1.16	12.18
140	WFNYWWYWHKK	11	1743.99	2	-1.63	9.53
141	KWWLWAAHKHRTFKFWW	17	2414.85	4	-0.87	11.26
142	KHRRRVKGP	9	1133.37	5	-2.48	12.31
143	LKKRHAVLRAKK	12	1447.84	6	-1.03	12.03
144	AKKKKFKKKRSRVRME	16	2048.57	9	-2.15	11.77
145	KKKKKKRRRRRWSYMKS	17	2350.91	12	-3.04	12.19
146	LVRNVVARMVKYYWKRKV	18	2308.86	6	-0.21	11.12
147	LLRRYLEAVRRSKKRRLKSFVLERRQL	27	3471.21	9	-0.94	12.01
148	GKWVWKVYLCNWRDYRLVYGWLVTVHPRW NCLKCLKCKRTT	41	5173.22	8	-0.34	9.88
149	QEQTRSQCDERGFKEKTNG	19	2241.38	0	-2.24	6.25
150	TWPAKVFALMKKERQYWWSNVTHCPGKTARKGRTC NWPHPKKWVCFGGGGGKKMLKAFVKAGWR	64	7420.83	14	-0.75	10.67
151	KPKASVRPKKSS	12	1312.58	5	-1.64	11.33
152	ALTHKVALALVVTV	14	1434.79	1	1.79	8.80
153	CCTWRCCWMT	10	1292.59	1	0.42	7.89
154	WWYFLGLYRKRLHYKKWRR	19	2756.3	7	-1.35	11.10
155	WRKHRKVNFRKRRKREL	17	2393.88	9	-2.52	12.31
156	EKRFYCSWHQKQSGCKPL	18	2225.57	3	-1.34	9.31
157	FKRSVMRKKWKKRRRRVS	18	2433.01	11	-1.99	12.70
158	LLKWLRRASKHRRLKR	16	2117.62	8	-1.38	12.61
159	HLLLLWWWKYRYYFKRR	17	2528.05	5	-0.77	10.55
160	WKKSALPPRTR	11	1339.61	4	-1.53	12.02
161	WFRFKPFFKFGRYK	14	1954.35	5	-0.78	11.17
162	AFKKHHKRLKKQFRS	15	1939.34	7	-1.87	12.04
163	RRFFYRQHLMWRHQM	15	2192.6	4	-1.36	12.00
164	VERIPLVRFKSIKKQLHERGDLEVF	25	3037.6	2	-0.4	9.69
165	ADNIGEKTKAALQELHDSE	19	2069.21	-3	-1.05	4.67
166	GEPTLMQKLEQFKENVKVFADNIGEKTKAALQELHD	36	4100.65	-2	-0.76	5.17
167	DMSHNSAQIRAHGKKVFSAL	20	2197.5	2	-0.47	9.99
168	VERIPLVRFKSIKKQLHERGDLE	23	2791.29	2	-0.74	9.69
169	HLKTDNINILNGLKHFSYL	19	2240.59	1	-0.23	8.51
170	VERIPLVRFKSIKKQLHERGDL	22	2662.17	3	-0.62	10.27
171	DMSHNSAQIRAHGKKVF	17	1926.18	2	-0.84	9.99
172	ARSEPGEPTLMQKLEQFKENVKVF	24	2806.23	0	-0.85	6.33
173	LRESIKPYTESIKTHLLNL	19	2255.64	1	-0.41	8.50
174	STKTRNWFSEHFKKVKEKLKDTFA	24	2956.39	4	-1.31	10.00
175	ADNIGEKTKAALQELHD	17	1853.02	-2	-0.92	4.90
176	ATPVKIRIENSNAFLSR	17	1916.21	2	-0.23	10.84
177	FQKGDVNGEKEQKVYTF	17	2017.22	0	-1.26	6.18
178	ILNKGKIVQAGRQLRQAGQNL	21	2305.71	4	-0.51	12.02
179	RAAASQATGKYQEMKAKTQQL	21	2309.62	3	-1.09	10.00
180	TQRSPKVQVYTRYPLGSKESNF	22	2585.9	3	-1.17	9.99
181	PQTRLDRFKDMLNVY	15	1896.19	1	-0.93	9.00
182	FSTKTRNWFSEHFKKVKEKLKDTFA	25	3103.57	4	-1.15	10.00
183	RESIKPYTESIKTHLLNL	18	2142.48	1	-0.64	8.50
184	YQLLTWEQANTAVKGVLDKVHSTGVEKLRDIYDKSVD	37	4220.75	-1	-0.5	5.61
185	LDRKVAQTDMTLRHIVSQF	19	2258.62	1	-0.28	8.75

186	SSMEHEIGPGQANEDAQGTGHAR	23	2379.46	-3	-1.27	4.80
187	DSEFSTKTRNWFSEHFKKVKEKLKDTFA	28	3434.85	2	-1.3	9.31
188	YSLKKTSMKIIPFTRL	16	1926.39	4	-0.05	10.46
189	REKVTPLVQDLRESIKPYTESIKTHLLNL	29	3421.98	1	-0.58	8.44
190	SEFSTKTRNWFSEHFKKVKEKLKDTFA	27	3319.77	3	-1.22	9.70
191	FSEHFKKVKEKLKDTFA	17	2082.43	2	-0.97	9.40
192	AKRSFQEGSASPYDLKE	17	1913.07	0	-1.28	6.35
193	LAKENYMQQVTQKIKD	16	1937.24	1	-1.16	8.43
194	PEDVDTPMGEEASVDVRGHRPLD	23	2521.7	-5	-1.03	4.26
195	TESIKTHLLNLFQEARKTLS	20	2329.68	1	-0.45	8.29
196	PQTRLDRFKDMLNVY	15	1896.19	1	-0.93	9.00
197	ASEHLDAFQRYLEELKRTFTPSQAG	25	2895.18	-1	-0.84	5.53
198	EPGEPTLMQKLEQFKENVKVF	21	2491.88	-1	-0.81	4.95
199	FTPFIQKATKPLNTQQEEELHRQ	23	2784.12	0	-1.27	6.76
200	IKPYTESIKTHLLNL	15	1770.1	1	-0.18	8.51
201	NIGEKTKAALQELHD	15	1666.85	-1	-0.93	5.54
202	SSMEHEIGPGQANEDAQGTGHAR	23	2379.46	-3	-1.27	4.80
203	FKKVKEKLKDTFA	13	1581.91	3	-0.91	9.83
204	GEPTLMQKLEQFKENVKVFADNIGEKTKAALQELHD	36	4100.65	-2	-0.76	5.17
205	LAQQGRQWASEHLDAFQRYLEELKRTFTPSQAG	33	3863.26	0	-0.96	6.77
206	HLKTDNINILNGLKHFSY	18	2127.43	1	-0.46	8.51
207	REKVTPLVQDLRESIKPYT	19	2272.63	1	-0.89	8.50
208	VEEGSTFEDEGGVARGPRLSERAQ	24	2576.72	-3	-1.04	4.42
209	YSLKKMSMKIRPFFPQ	16	2001.48	4	-0.51	10.46
210	SGLAQQGRQWASEHL	15	1667.8	0	-0.92	6.47
211	GEPTLMQKLEQFKENVKVF	19	2265.65	0	-0.63	6.23
212	VEEGSTFEDEGGVARGPRLSER	22	2377.51	-3	-1.05	4.42
213	STHGHWDQISKVEGH	15	1717.81	-1	-1.29	6.20
214	DLRESIKPYTESIKTHLLNL	20	2370.73	0	-0.56	6.76
215	EPGEPTLMQKLEQFKENVKVFADNIGEKTKAALQELHD	38	4326.89	-3	-0.86	4.91
216	TDLAKENYMQQVTQKIKD	18	2153.43	0	-1.26	5.89
217	AYRTQGGSKNYLHL	14	1607.79	2	-1.01	9.70
218	SSMEHEIGPGQANEDAQGTGHA	22	2223.27	-4	-1.12	4.39
219	SAREAEHKSEIAHRFSD	17	1970.09	-1	-1.35	6.00
220	QKLEQFKENVKVF	13	1636.91	1	-0.88	8.50
221	EPGEPTLMQKLEQFKENVKVFAD	23	2678.05	-2	-0.81	4.71
222	PVKIRIENSNAFLSR	15	1744.02	2	-0.33	10.84
223	GEPTLMQKLEQFKENVKVF	19	2265.65	0	-0.63	6.23
224	YSLKKTSMKIIPFTRL	16	1926.39	4	-0.05	10.46
225	VVFAIHHPSALSPEIHASLDKFL	23	2528.93	-1	0.6	6.26
226	VAEAYKTQGGTKNYLHL	17	1893.13	1	-0.64	8.41
227	AEYPKTLRKHANELLDRKVAQTDMTL	26	3042.5	1	-0.89	8.48
228	ADNIGEKTKAALQELH	16	1737.93	-1	-0.76	5.45
229	SIKPYTESIKTHLLNL	16	1857.18	1	-0.22	8.24
230	ALKAIGNVGHPASLKRIMKFIPGYTTSAADL	31	3241.84	3	0.25	10.00
231	MQKLEQFKENVKVF	14	1768.1	1	-0.68	8.25
232	MQKLEQFKENVKVF	14	1768.1	1	-0.68	8.25
233	LAKENYMQQVTQKIKD	16	1937.24	1	-1.16	8.43

234	LAQQGRQWASEHL	13	1523.67	0	-0.97	6.75
235	HVQWSRKNQKPVNPKPLEL	19	2298.67	3	-1.44	10.29
236	RESIKPYTESIKTHLLN	17	2029.32	1	-0.9	8.50
237	AKRSFQEGSASPYDLKEVL	19	2125.36	0	-0.73	6.22
238	LREKVTPLVQDLRESIKPYTESIKTHLLNL	30	3535.14	1	-0.43	8.44
239	KKTSMKIIPFTRL	13	1562.97	4	-0.19	11.26
240	AEYPKTLRKHANEL	14	1669.9	1	-1.31	8.55
241	REKVTPLVQDLRESIKPYTESIKTHL	26	3081.56	1	-0.8	8.44
242	LRESIKPYTESIKTHL	16	1915.22	1	-0.74	8.50
243	TQKMQKLRESLPYTSNVRDQ	20	2422.74	2	-1.5	9.69
244	PALKYVRPGGGFAPNFQL	18	1932.25	2	-0.09	10.01
245	ADNIGEKTKAALQELHDSEFSTKTRNWFSEH FKKVKEKLKDTFA	44	5154.77	1	-1.1	8.32
246	YSLKKMSMKIRPFFPQ	16	2001.48	4	-0.51	10.46
247	KTKAALQELHDSEF	14	1616.79	-1	-0.89	5.45
248	RESIKPYTESIKTHLLNLFQEARKTLS	27	3203.69	2	-0.74	9.52
249	FKVGPETDKFRLTYG	15	1758	1	-0.72	8.50
250	VEEGSTFEDEGGVARGPRLSERA	23	2448.59	-3	-0.93	4.42
251	NVKVFADNIGEKTKA	15	1633.86	1	-0.5	8.50
252	PVLHVPQDLRTLKL	14	1628.97	1	0.08	9.18
253	EPGEPTLMQKLEQFKENVKVFADNIGEKTKAALQ	34	3832.38	-1	-0.77	5.11
254	AQQGRQWASEHL	12	1410.51	0	-1.37	6.79
255	KKVKEKLKDTFA	12	1434.74	3	-1.22	9.83
256	KVHSTGVEKLRDIYDKSVD	19	2189.45	0	-0.87	6.76
257	AVQNFNKRSERHFLYAL	17	2093.37	2	-0.66	9.99
258	FSTKTRNWFSEH	12	1539.67	1	-1.41	8.76
259	QRYLEELKRTFTPSQAG	17	2024.26	1	-1.19	8.59
260	FQRYLEELKRTFT	13	1730.98	1	-0.99	8.59
261	RYFWQHDDPQTRLDRFKDMLNVY	23	3044.39	0	-1.37	6.75
262	MSGLAQQGRQWASEHLD	17	1914.08	-1	-0.91	5.44
263	YQLLTWEQANTAVKGVLDKVHSTGVEKLRD	30	3399.85	0	-0.47	6.77
264	RYFWQHDDPQTRLDRFKDMLNVY	23	3044.39	0	-1.37	6.75
265	RKYNGDKADGNQFAL	15	1696.84	1	-1.45	8.50
266	YQLLTWEQANTAVKGVLDKVH	21	2413.76	0	-0.23	6.75
267	YSLKKMSMKIRPFFPQ	16	2001.48	4	-0.51	10.46
268	TQKMQKLRESLPYTSNVRDQAVQHLSNL	28	3285.72	2	-0.98	9.69
269	KIKPMKDSTVLPHFKAGD	18	2012.39	2	-0.66	9.53
270	HDSEFSTKTRNWFSEHFKKVKEKLKDTFA	29	3572	2	-1.37	9.31
271	WASEHLDAFQRY	12	1522.64	-1	-0.92	5.32
272	TQKIKDTLASFDM	13	1497.72	0	-0.44	5.63
273	REKLVPVVSNVKDLSRQKLEL	21	2450.9	2	-0.45	9.70
274	GEPTLMQKLEQFKENVKVFADNIGEKTKA	29	3293.78	0	-0.8	6.31
275	MSGLAQQGRQWASEHLDAFQRYLEELKRTFTPSQAG	36	4138.58	0	-0.86	6.52
276	KIKPMKDSTVLPHFKAGD	18	2012.39	2	-0.66	9.53
277	RSEPGEPTLMQKLEQFKENVKVF	23	2735.15	0	-0.97	6.29
278	TQKMQKLRESLPYTSNVRDQAVQHLSNL	28	3285.72	2	-0.98	9.69
279	EPGEPTLMQKLEQFKENVKVFADNIGEKTKAALQELHD	38	4326.89	-3	-0.86	4.91
280	RQWASEHLDAFQRYLEELKRTFTPSQAG	28	3365.71	0	-1.07	6.77
281	PQTRLDRFKDML	12	1519.78	1	-1.12	9.17

282	DQPAKNPTKFGQKTMY	16	1854.11	2	-1.59	9.53
283	REKVTPLVQDLRESIKPYTESIKTHLLNLFQEARKTLS	38	4483.19	2	-0.66	9.40
284	PRRATPVKIRIENSNAF	17	1969.27	3	-0.76	11.71
285	VEEGSTFEDEGGVARGPR	18	1891.97	-3	-1.01	4.25
286	DRKVAQTDMTLRHIVSQF	18	2145.46	1	-0.51	8.75
287	FRESLMTGFAAKY	13	1520.76	1	-0.02	8.59
288	TIRNEAHKYQLSL	13	1572.78	1	-0.85	8.29
289	TQRSPKVQVYTRYPLG	16	1893.17	3	-1.01	10.28
290	AQQGRQWASEHLD	13	1525.6	-1	-1.53	5.46
291	DDPQTRLDRFKDMLNVY	17	2126.37	-1	-1.24	4.58
292	FVLKSFAQARRY	12	1485.75	3	-0.11	11.00
293	AEYPKTLRKHANELL	15	1783.06	1	-0.97	8.55
294	PQTRLDRFKDML	12	1519.78	1	-1.12	9.17
295	MQQVTQKIKDTL	12	1432.69	1	-0.73	8.34
296	AQQGRQWASEHLDAFQRY	18	2191.35	0	-1.37	6.80
297	GRQWASEHLDAFQRYLEELKRTFTPSQAG	29	3422.76	0	-1.04	6.77
298	TQKMQKLRESLPYTSNVRD	19	2294.61	2	-1.39	9.69
299	EPGEPTLMQKLEQFKENVK	19	2245.57	-1	-1.26	4.95
300	VERIPLVRFKSIKKQLHERGDLEVFWSNHQPDIF	34	4162.81	1	-0.58	8.48
301	EPGEPTLMQKLEQFKEN	17	2018.27	-2	-1.43	4.49
302	PPPVIKFNRPFLMWIVERDTRSILFMGKIVNPKAP	35	4109.01	4	0.02	11.00
303	AVQNFNKRSERHFL	14	1745.96	2	-1.11	10.84
304	KLEQFKENVKVF	12	1508.78	1	-0.66	8.50
305	KTKAALQELHDSE	13	1469.61	-1	-1.18	5.58
306	ADNIGEKTKAALQELHDSEF	20	2216.39	-3	-0.86	4.50
307	FTQLVETGKKAAEQ	14	1549.74	0	-0.66	6.14
308	DIFLNPKIHPQVRMLAA	17	1963.37	1	0.22	8.75
309	EEASVDVRGHRPLD	14	1579.69	-2	-1.07	4.90
310	HFKKVKEKLKDTFA	14	1719.06	3	-1.07	9.83
311	SEFSTKTRNWFSEHFKKVKEKLKD	24	3000.4	3	-1.54	9.70
312	GEPTLMQKLEQFKENVKVFAD	21	2451.82	-1	-0.65	5.01
313	LSEDQNQQARAQRHHSPGNN	20	2287.35	0	-2.16	6.92
314	EHFKKVKEKLKDTFA	15	1848.17	2	-1.23	9.41
315	SELHAHSLRVD	11	1263.37	-1	-0.54	5.96
316	FASDDKGRLYAFRKGYY	17	2057.29	2	-0.96	9.40
317	EAEHKSEIAHRF	12	1453.57	-1	-1.27	6.03
318	KKTSMKIIPFTRL	13	1562.97	4	-0.19	11.26
319	VLSMEDKSNVKAI	13	1433.68	0	0.04	6.04
320	TPVKIRIENSNAFLS	15	1688.94	1	-0.08	8.41
321	RYTTLHTFKNILL	13	1619.92	2	0.02	9.99
322	FQRYLEELKRTFTPSQAG	18	2171.44	1	-0.97	8.59
323	LGLHHEAKKDPSLQIRDVEAL	21	2369.7	-1	-0.55	6.02
324	EPGEPTLMQKLEQFKENVK	19	2245.57	-1	-1.26	4.95
325	GIESSDKAFTSHNAMQF	17	1870.02	-1	-0.53	5.32
326	ELHDSEFSTKTRNWFSEHFKKVKEKLKDTFA	31	3814.27	1	-1.27	8.44
327	VSDVKSGLKDSPPK	14	1456.66	1	-0.89	8.47
328	EAEHKSEIAHRFSD	14	1655.74	-2	-1.39	5.43
329	FKSIKKQLHERGDL	14	1698.98	2	-1.16	9.70

330	AVSAVLTSKYR	11	1194.39	2	0.35	9.99
331	DQPAKNPTKFGQKTMY	16	1854.11	2	-1.59	9.53
332	VTQKIKDTLASFDMSGLAQQGRQWASEHLDAFQRYL EELKRTFTPSQAG	49	5586.24	0	-0.7	6.74
333	FSQAVQRNPSREL	13	1531.69	1	-1.05	9.60
334	ADNIGEKTKAALQEL	15	1600.79	-1	-0.59	4.68
335	ESIKPYTESIKTHLLNL	17	1986.29	0	-0.41	6.85
336	PTLMQKLEQFKENVKVF	17	2079.48	1	-0.47	8.90
337	ARSEPGEPTLMQKLEQFKENVKVF	24	2806.23	0	-0.85	6.33
338	TQKMQKLRESLPYTSNVRDQAVQHL	25	2971.38	2	-1.08	9.69
339	HLDAFQRYLEELKRTFTPSQAG	22	2607.91	0	-0.84	6.76
340	EEASVDVRGHRPLDK	15	1707.86	-1	-1.26	5.49
341	LRESIKPYTESIKTHLLNLFQEARKTLS	28	3316.85	2	-0.58	9.52
342	GLAQQGRQWASEHLDAFQRYLEELKRTFTPSQAG	34	3920.31	0	-0.94	6.77
343	GEPTLMQKLEQFKENVKVFADNIGEKTKA	29	3293.78	0	-0.8	6.31
344	MLGFSQAVQRNPSREL	16	1833.09	1	-0.52	9.35
345	DDPQTRLDRFKDML	14	1749.96	-1	-1.46	4.58
346	STFKAHFLGHIAEPFEVG	18	1987.24	-1	0.18	5.95
347	FIRPLKAKEQFL	12	1489.82	2	-0.12	9.99
348	LEQFKENVKVF	11	1380.6	0	-0.36	6.14
349	TESIKTHLLNL	11	1268.47	0	-0.04	6.41
350	KSRVNRMKQNL	11	1373.64	4	-1.65	12.02
351	VRDQAVQHLSNLREKVTPLVQDL	23	2659.04	0	-0.43	6.73
352	PGEPTLMQKLEQFKENVKVF	20	2362.77	0	-0.68	6.64
353	ADNIGEKTKAALQELHDSEFSTKTRNWFS	29	3324.61	-1	-1.04	5.56
354	TQKMQKLRESLPY	13	1621.91	2	-1.36	9.70
355	ILNKGKIVQAGRQL	14	1537.87	3	-0.07	11.17
356	TQKMQKLRESLPYTSNVRDQAVQHL	25	2971.38	2	-1.08	9.69
357	PQTRLDRFKD	10	1275.43	1	-1.91	9.17
358	ATPVKIRIENSNAFL	15	1672.94	1	0.09	8.79
359	KLRKTLAPYKEEL	13	1588.91	2	-1.05	9.53
360	VEEGSTFEDEGGVARGPRLSERAQA	25	2647.8	-3	-0.92	4.42
361	VLSMEDKSNVK	11	1249.44	0	-0.53	6.04
362	GRQWASEHLD	10	1198.26	-1	-1.47	5.45
363	AVTASELKSGDNLPVNFHLKTDNINILNGLKHFSYL	36	3984.52	0	-0.12	6.96
364	RESIKPYTESIKTHL	15	1802.06	1	-1.04	8.50
365	MHLKKPVAFSNF	12	1418.72	2	0.03	10.00
366	EPGEPTLMQKLEQFK	15	1775.05	-1	-1.15	4.79
367	PPPVIKFNRPFLMWIVERDTRSILFMGKIVNPKAP	35	4109.01	4	0.02	11.00
368	MLGFSQAVQRNPSRELVL	18	2045.38	1	-0.02	9.35
369	DDPQTRLDRFKDML	14	1749.96	-1	-1.46	4.58
370	VRDQAVQHLSNL	12	1379.53	0	-0.39	6.71
371	GEPTLMQKLEQFKENVKV	18	2118.47	0	-0.82	6.23
372	EPGEPTLMQKLEQFKENVKVFADNIGE	27	3091.48	-3	-0.8	4.57
373	LNLFQEARKTLS	12	1419.64	1	-0.37	8.75
374	GEPTLMQKLEQFKENVKVFAD	21	2451.82	-1	-0.65	5.01
375	LNTQQEEELHRQ	12	1524.61	-2	-2.11	4.75
376	RAAASQATGKYQEMKAKTQQL	21	2309.62	3	-1.09	10.00
377	MQQVTQKIKDTL	12	1432.69	1	-0.73	8.34

378	IRPLKAKEQFL	11	1342.64	2	-0.38	9.99
379	DIFLNPKIHPQVRM	14	1708.05	1	-0.26	8.75
380	KLAVETDPSPITAKAGD	17	1712.91	-1	-0.36	4.78
381	PSDKPLASGYVKIFGQEFL	19	2096.41	0	-0.08	6.48
382	REKVTPLVQDLRE	13	1582.82	0	-1.02	6.30
383	KTKAALQELHD	11	1253.42	0	-1	6.76
384	TQKMQKLRESLPYTSNVRDQ	20	2422.74	2	-1.5	9.69
385	TQKMQKLRESLPY	13	1621.91	2	-1.36	9.70
386	GEPQGSKFVHPTYGKETH	18	1999.17	0	-1.42	6.92
387	LRHRIEGEELTEL	13	1594.78	-2	-0.88	4.91
388	EPGEPTLMQKLEQFKEN	17	2018.27	-2	-1.43	4.49
389	ASEHLDAFQRYLEELKRTF	19	2353.62	-1	-0.83	5.53
390	KKMSMKIRPFFPQ	13	1638.06	4	-0.75	11.26
391	KTKAALQELH	10	1138.33	1	-0.75	8.60
392	VERIPLVRFK	10	1256.55	2	0.15	10.83
393	FIRPLKAKEQF	11	1376.66	2	-0.47	9.99
394	PEDVDTPMGEEASVDVRGHRPLD	23	2521.7	-5	-1.03	4.26
395	EQFKENVKVF	10	1267.44	0	-0.78	6.24
396	STFKAHFLGHIAEPFEVGMRAERLQEIL	28	3227.73	-1	-0.01	5.99
397	EPGEPTLMQKLEQFKE	16	1904.16	-2	-1.3	4.69
398	FQEARKTLS	9	1079.22	1	-0.94	8.75
399	VEEGSTFEDEGGVARGPRL	19	2005.13	-3	-0.76	4.25
400	AFQRYLEELKRTFTPSQAG	19	2242.52	1	-0.82	8.63
401	NTQQEEELHRQ	11	1411.45	-2	-2.65	4.75
402	YGPEGKNPSLRKFINNL	17	1947.22	2	-1.03	9.70
403	QGRQWASEHL	10	1211.3	0	-1.47	6.75
404	PNRYRPEGLPEKY	13	1618.81	1	-2.11	8.90
405	AQQGRQWASEHLDAFQRYLEELKRTFT	27	3309.65	0	-1.15	6.81
406	DDPQTRLDRFKD	12	1505.6	-1	-2.18	4.78
407	GEPTLMQKLEQFKENVK	17	2019.34	0	-1.11	6.23
408	PVVSNVKDLSRQKLEL	16	1825.13	1	-0.34	9.00
409	FGHQGFFPDSTSKAL	15	1638.8	0	-0.32	6.74
410	SGLAQQGRQWASEHLDAFQRYLEELKRTFTPSQAG	35	4007.39	0	-0.94	6.49
411	EPGEPTLMQKLEQFKENVKVFA	22	2562.96	-1	-0.69	4.95
412	VLSMEDKSNVK	11	1249.44	0	-0.53	6.04
413	VLKSFAQARRYM	12	1469.76	3	-0.18	11.00
414	STFKAHFLGHIAEPFEVGMRAERLQEIL	28	3227.73	-1	-0.01	5.99
415	KSFAQARRYM	10	1257.47	3	-1.02	11.00
416	EHFKKVKEKLK	11	1413.72	3	-1.72	9.83
417	DDPQTRLDRFKDMLNVY	17	2126.37	-1	-1.24	4.58
418	DIFLNPKIHPQVRMLAA	17	1963.37	1	0.22	8.75
419	ARRPPGFTPFRSL	13	1501.75	3	-0.69	12.30
420	YMQQVTQKIKD	11	1381.61	1	-1.2	8.50
421	TLASFDMSGLAQQGRQWASEHLDAFQRY LEELKRTFTPSQAG	42	4773.27	-1	-0.66	5.53
422	RDQGQRLREQL	11	1398.54	1	-2.16	9.51
423	KVKEKLKDTFA	11	1306.56	2	-0.97	9.53
424	STKTRNWFSEH	11	1392.49	1	-1.79	8.49
425	FSQAVQRNPSRELVL	15	1743.98	1	-0.37	9.60

426	RGHVDKLRKTLAPYKEEL	18	2153.51	2	-1.17	9.52
427	LHEIYPQTSPLQTAEEGKD	19	2156.33	-3	-1.12	4.56
428	GEPTLMQKLEQFKEN	15	1792.03	-1	-1.28	4.79
429	HGLEVPLRRS	10	1163.34	1	-0.67	9.61
430	IKDFSEHFR	9	1178.31	0	-1.03	6.75
431	VKHLPKVYRL	10	1252.56	3	-0.24	10.29
432	NKGKIVQAGRQLRQAGQNL	19	2079.39	4	-1.01	12.02
433	FVNQHPHEGRQVLERKNVL	19	2300.61	1	-1.02	8.76
434	TQKMQKLRESLPYTSN	16	1924.2	2	-1.42	9.70
435	RYFWQHDD	8	1166.21	-1	-2.2	5.39
436	VQYKEGFGHLSPDDQTEF	18	2097.22	-3	-1.09	4.31
437	VKDLSRQKLEL	11	1328.57	1	-0.73	8.56
438	LRFNPVSGDVPAHRYPLDSRDY	22	2574.84	0	-0.89	6.75
439	EIKDFSEHF	9	1151.24	-2	-0.92	4.65
440	IKESMDNNIPSAIRVL	16	1800.1	0	-0.03	6.07
441	EEASVDVRGHRP	12	1351.44	-1	-1.28	5.45
442	TLASFDMSGLAQQGRQWASEHLDAFQRYLEELK RTFTPSQAG	42	4773.27	-1	-0.66	5.53
443	SLRTNKEKEKHPNLVN	16	1907.16	2	-1.76	9.70
444	QQGRQWASEHLDAFQRYLEELKRTFTPSQAG	31	3679.02	0	-1.2	6.77
445	ALQELHDSEFSTKTRNWFSEHFKKVKEKLKDTFA	34	4126.64	1	-1.1	8.39
446	FADNIGEKTKA	11	1193.32	0	-0.77	6.07
447	SEHFKKVKEKLKDTFA	16	1935.25	2	-1.21	9.40
448	TQKMQKLRESLPYTSN	16	1924.2	2	-1.42	9.70
449	GEPTLMQKLEQFKENVKV	18	2118.47	0	-0.82	6.23
450	GTREGGRRPRLLG	13	1424.62	3	-1.37	12.00
451	MQQVTQKIKD	10	1218.43	1	-1.19	8.35
452	GEPTLMQKLEQFKEN	15	1792.03	-1	-1.28	4.79
453	LDAFQRYLEELKRTFTPSQAG	21	2470.76	0	-0.73	6.18
454	TKTSKYFKPGMPFEL	15	1774.1	2	-0.73	9.52
455	EPGEPTLMQKLEQFKE	16	1904.16	-2	-1.3	4.69
456	PVLHVPQDLRT	11	1274.48	0	-0.24	7.17
457	AVQHLSNLREKVTPLVQDLRESIKPYTESIK THLLNL	37	4284.97	1	-0.38	8.49
458	EIKDFSEHFR	10	1307.42	-1	-1.28	5.45
459	STRYITHTTDD	11	1309.35	-1	-1.37	5.37
460	ALKAIGNVGHPASL	14	1347.58	1	0.56	8.80
461	NIGEKTKAALQELHDSEFSTKTRNWFSEHFKK VKEKLKDTFA	42	4968.6	2	-1.12	9.17
462	SHGLEVPLRRS	11	1250.42	1	-0.68	9.34
463	ENYMQQVTQKIKD	13	1624.83	0	-1.55	6.29
464	DDGDKHPSKVEPTAE	15	1624.68	-3	-1.84	4.61
465	FVGDKYYRVNL	11	1373.57	1	-0.31	8.50
466	KSIDGANVKHL	11	1181.35	1	-0.45	8.60
467	SSRPAYRRWIL	11	1404.63	3	-0.8	11.71
468	DAFQRYLEELKRTFTPSQAG	20	2357.6	0	-0.96	6.18
469	EPGEPTLMQKLEQFKENVKVF	21	2491.88	-1	-0.81	4.95
470	AEVDQYREKL	10	1250.37	-1	-1.39	4.68
471	EDEGGVARGPR	11	1142.19	-1	-1.48	4.68
472	VERIPLVRF	9	1128.38	1	0.6	9.57

473	KKMSMKIRPFFPQ	13	1638.06	4	-0.75	11.26
474	ADNIGEKTKAAL	12	1230.38	0	-0.48	6.11
475	AVEKVKKILD	10	1142.4	1	-0.02	8.54
476	TQKMQKLRESLPYTSNVRD	19	2294.61	2	-1.39	9.69
477	FDMSGLAQQGRQWASEHL	18	2061.26	-1	-0.7	5.32
478	DQPAKNPTKFGQKTM	15	1690.93	2	-1.61	9.70
479	TQKIKDTLASFDMSGL	16	1755.01	0	-0.19	5.63
480	MSGLAQQGRQWASEHLDAFQRYLEELKRTFTPSQAG	36	4138.58	0	-0.86	6.52
481	ATPVKIRIENSNA	13	1412.6	1	-0.4	8.79
482	EHLDAFQRYLEELKRTFTPSQAG	23	2737.02	-1	-0.96	5.53
483	VLRAAATSLRTIDA	14	1457.69	1	0.63	9.57
484	SEHLDAFQRYLEELKRTFTPSQAG	24	2824.1	-1	-0.95	5.51
485	VQRNPSRELVL	11	1310.51	1	-0.54	9.57
486	EAPAPAKTEVSVKENKAKE	19	2026.27	0	-1.17	6.47
487	VELFTQLVETGKKAAEQ	17	1891.15	-1	-0.28	4.79
488	AEYPKTLRKHANELLDRKVAQTDMTLRHIVSQF	33	3910.51	2	-0.72	9.52
489	ALRDQGQRLREQL	13	1582.78	1	-1.4	9.56
490	KEKLKDTFA	9	1079.26	1	-1.22	8.50
491	PTKEIVQRL	9	1083.29	1	-0.58	9.18
492	SIGEGQQQAGGVKQVGDV	18	1756.89	-1	-0.49	4.37
493	GEPTLMQKLEQFKENVKVFA	20	2336.73	0	-0.51	6.23
494	GEPTLMQKLEQFKENVK	17	2019.34	0	-1.11	6.23
495	TSNVRDQAVQHLSNL	15	1681.82	0	-0.65	6.41
496	VQDLRESIKPY	11	1347.53	0	-0.92	6.04
497	DQPAKNPTKFGQK	13	1458.63	2	-1.95	9.70
498	LEELKRTFTPSQAG	14	1576.77	0	-0.78	6.14
499	AEYPKTLRKHANELLDRKVAQTDMTLRHIVSQF	33	3910.51	2	-0.72	9.52
500	ADNIGEKTKAALQ	13	1358.51	0	-0.71	6.11
501	DIFLNPKIHPQVRM	14	1708.05	1	-0.26	8.75
502	LFQEARKTLS	10	1192.38	1	-0.47	8.75
503	VEEGSTFEDEGGVARGPRLSERAQAS	26	2734.87	-3	-0.92	4.42
504	IRPLKAKEQF	10	1229.48	2	-0.8	9.99
505	VPVVSNVKDLSRQKLEL	17	1924.27	1	-0.08	8.56
506	VALPGAHPYAAALRL	15	1519.81	1	0.8	8.73
507	NLFQEARKTLS	11	1306.48	1	-0.75	8.75
508	PKDWPEWNFLNKMQQLEL	18	2316.66	-1	-1.18	4.68
509	VKVFADNIGEKTKA	14	1519.76	1	-0.29	8.47
510	SQAVQRNPSREL	12	1384.51	1	-1.37	9.31
511	REKVTPLVQDLR	12	1453.7	1	-0.81	8.75
512	SLMLHEIYPQTSPLQTAEEGKDPDKGP	27	2982.31	-3	-1.01	4.50
513	TKTPKYFKPGMPFEL	15	1784.14	2	-0.79	9.52
514	ATPVKIRIENSNAFLS	16	1760.02	1	0.04	8.79
515	VLSMEDKSNVKAIWG	15	1676.94	0	-0.05	6.04
516	ATPVKIRIEN	10	1140.34	1	-0.27	8.79
517	PSDKPLASGYVKIFGQE	17	1836.07	0	-0.48	6.66
518	RFNPVSGDVPAHRYPLDSRDY	21	2461.68	0	-1.11	6.75
519	TREGGRRPRLLG	12	1367.57	3	-1.45	12.00
520	SGLAQQGRQWASEHLDAFQRY	21	2448.64	0	-1.05	6.47

521	SHNSAQIRAHGKKVF	15	1679.9	3	-0.84	11.17
522	LEQFKENVK	9	1134.29	0	-1.22	6.14
523	RYLEELKRTFTPSQAG	16	1896.13	1	-1.04	8.59
524	GEPTLMQKLEQFKENVKVFADNIGEKTKAAL QELHDSEFSTKTRNWFSEHFKKVKEKLKDTFA	63	7402.41	1	-0.96	8.21
525	PGEPTLMQKLEQFKENVKVF	20	2362.77	0	-0.68	6.64
526	ESIKTHLLNL	10	1167.37	0	0.03	6.85
527	QLSVQQIIENLQKRINT	17	2025.33	1	-0.54	8.75
528	TKTPKYFKPGMPFEL	15	1784.14	2	-0.79	9.52
529	IKEVGNKDFPEARAY	15	1736.94	0	-0.97	6.18
530	MQKNADKAYQVAVKQVDEIDL	21	2406.73	-1	-0.61	4.68
531	DGDKHPSKVEPTAE	14	1509.59	-2	-1.72	4.95
532	KLAVETDPSPITAKAGDD	18	1828	-2	-0.54	4.42
533	HVDKESYDNIQRQL	14	1744.88	-1	-1.59	5.38
534	WASEHLDAFQRYLEELKRTFTPSQAG	26	3081.39	-1	-0.84	5.53
535	STFKAHFLGHIAEPFEVGM	19	2118.43	-1	0.27	5.95
536	ARRPPGFTPFR	11	1301.51	3	-1.09	12.30
537	PKDWPEWNFLNKMQQLEL	18	2316.66	-1	-1.18	4.68
538	ATKPLNTQQEEELHRQ	16	1922.08	-1	-1.86	5.50
539	SVQQIIENLQKRINT	15	1784.04	1	-0.63	8.46
540	EAEHKSEIAHRFSDLKETTFKEVALITFAQYLQK	34	4009.53	-1	-0.57	6.07
541	MQQVTQKIKDTLAS	14	1590.85	1	-0.56	8.34
542	KIKPMKDSTVLPHFKAGDPKDWPEWNFL	28	3325.87	1	-0.88	8.39
543	EELKRTFTPSQAG	13	1463.61	0	-1.13	6.24
544	KPADLPPIAEKYIQD	15	1697.94	-1	-0.77	4.78
545	GEPTLMQKLEQFKE	14	1677.93	-1	-1.12	5.01
546	MLGFSQAVQRNPSRELVL	18	2045.38	1	-0.02	9.35
547	GRQWASEHL	9	1083.17	0	-1.24	6.75
548	TPVKIRIENSNA	12	1341.53	1	-0.58	8.41
549	SVDVRGHRPLD	11	1250.38	0	-0.89	6.48
550	KRTFTPSQAG	10	1092.22	2	-1.15	11.00
551	PTLMQKLEQFKENVKVF	17	2079.48	1	-0.47	8.90
552	DDPQTRLDRFKDMLN	15	1864.06	-1	-1.59	4.58
553	VDKLRKTLAPYKEEL	15	1803.13	1	-0.86	8.40
554	TLRHIVSQF	9	1100.28	1	0.29	9.44
555	EPGEPTLMQKLEQFK	15	1775.05	-1	-1.15	4.79
556	FDMSGLAQQGRQWASEHL	18	2061.26	-1	-0.7	5.32
557	FRESLMTGFAAKY	13	1520.76	1	-0.02	8.59
558	AFQRYLEELKRTFT	14	1802.06	1	-0.79	8.63
559	LRHRIEGEEL	10	1251.4	-1	-1.1	5.50
560	MEHEIGPGQANEDAQGTGHA	20	2049.11	-4	-1.15	4.39
561	TQKIKDTLASFDMSG	15	1641.85	0	-0.46	5.63
562	EPGEPTLMQKLEQFKENVKVFADNIGEKTKAALQ ELHDSEFSTKTRNWFSEHFKKVKEKLKDTFA	65	7628.64	0	-1.01	7.03
563	ARSEPGEPTLMQKLEQF	17	1961.22	-1	-0.99	4.79
564	KSIKKQLHERGDLEVFWSNHQPDIF	25	3052.44	0	-1.01	6.93
565	KTRNWFSEHFKKVKEKLKDTFA	22	2768.21	4	-1.36	10.00
566	WASEHLDAFQRYLEE	15	1894.03	-3	-0.95	4.60
567	TLASFDMSGLAQQGRQWASEHLD	23	2548.77	-2	-0.52	4.74
568	REKVTPLVQD	10	1184.35	0	-0.9	6.19

569	ELKRTFTPSQAG	12	1334.49	1	-0.93	8.85
570	SLEFQRKLQEKA	12	1476.69	1	-1.24	8.31
571	REKVTPLVQDLRESIKPY	18	2171.52	1	-0.91	8.50
572	DMSGLAQQGRQWASEHLDAFQRYLEELK RTFTPSQAG	37	4253.67	-1	-0.93	5.56
573	SIKPYTESIKTHL	13	1516.75	1	-0.58	8.24
574	GEPTLMQKLEQFKE	14	1677.93	-1	-1.12	5.01
575	IKVPSKDLKKD	11	1270.53	2	-1.14	9.53
576	KSIKKQLHERGDLEVF	16	1927.23	1	-0.97	8.50
577	FKQSYPIGKLEAF	13	1527.78	1	-0.25	8.50
578	VEKVSSYGVFPHYSL	15	1711.93	0	0.11	6.72
579	SQAVQRNPSRELVL	14	1596.8	1	-0.6	9.31
580	EHFKKVKEKLKD	12	1528.81	2	-1.87	9.41
581	VLSMEDKSNVKAIWGKA	17	1876.2	1	-0.17	8.47
582	RVLSDWKALPSDKPLASGY	19	2103.4	1	-0.44	8.50
583	QFKENVKVF	9	1138.33	1	-0.48	8.59
584	VQRNPSREL	9	1098.22	1	-1.54	9.57
585	SFDMSGLAQQGRQWASEHLDAFQRYLEELK RTFTPSQAG	39	4487.93	-1	-0.83	5.53
586	REKVTPLVQDLRESIKPYTE	20	2401.74	0	-1.03	6.35
587	MQQVTQKIKD	10	1218.43	1	-1.19	8.35
588	KENVKVFADNIGE	13	1462.62	-1	-0.63	4.94
589	SFDMSGLAQQGRQWASEHLDAFQRYLEELK RTFTPSQAG	39	4487.93	-1	-0.83	5.53
590	LKRTFTPSQAG	11	1205.38	2	-0.7	11.00
591	YQLLTWEQANTAVKGVLDKVHSTGVEKLRDIYD	33	3791.27	-1	-0.44	5.59
592	REKVTPLVQDL	11	1297.51	0	-0.47	6.07
593	EEAQERMRGHVDKLR	15	1854.07	0	-1.79	6.87
594	ERVGLDRRAEAINL	14	1611.82	0	-0.57	6.28
595	ASQATGKYQEMKAKTQQL	18	2011.28	2	-1.22	9.53
596	EPGEPTLMQKLEQFKENVKVFADNIGEKTKA	31	3520.01	-1	-0.91	5.11
597	VLSMEDKSNVKAIWGKASGHLEEYGAEALERMF	33	3697.19	-2	-0.35	5.06
598	PYYHPRAPSAEVEMTA	16	1819.02	-1	-0.76	5.41
599	PDKGPRENLGPGLE	14	1478.62	-1	-1.49	4.94
600	EELTQKMQKLR	11	1403.66	1	-1.59	8.69
601	GEPTLMQKLEQFKENVKVFADNIGEKTKAALQ ELHDSEFSTKTRNWFSEHFKKVKEKLKDTFA	63	7402.41	1	-0.96	8.21
602	ESLPYTSNVRDQAVQHLSNL	20	2271.47	-1	-0.66	5.32
603	AVQRNPSRELVL	12	1381.59	1	-0.34	9.64
604	SIKPYTESIKTHLLNLFQEARKTLS	25	2918.38	2	-0.48	9.52
605	AIHHPSALSPEIHA	14	1479.65	-1	0.02	6.26
606	SEPGEPTLMQKLEQFKENVKVF	22	2578.96	-1	-0.81	4.94
607	EESEDTDLAKENYMQQVTQKIKD	23	2742.95	-4	-1.63	4.31
608	EPGEPTLMQKLEQFKENVKV	20	2344.71	-1	-0.99	4.95
609	PPPVIKFNRPFLMWIVERDTRSILFMGKIVNPKAP	35	4109.01	4	0.02	11.00
610	KVHSTGVEKLRDIYD	15	1759.98	0	-0.83	6.76
611	DMSGLAQQGRQWASEHLDAFQRYLEELKRTFTPSQAG	37	4253.67	-1	-0.93	5.56
612	MQKLEQFKEN	10	1294.48	0	-1.68	5.90
613	VLSMEDKSNVKAIWGKASGHLEEYGAEALERMF	33	3697.19	-2	-0.35	5.06
614	TDLAKENYMQQVTQKIKD	18	2153.43	0	-1.26	5.89
615	RPPGFTPFR	9	1074.25	2	-1.03	12.00

616	KTKAALQELHDSEFST	16	1804.97	-1	-0.88	5.45
617	VLRAAATSLRTID	13	1386.61	1	0.54	9.57
618	TQKIKDTLASFDMSGLAQQGRQWASEHL	28	3147.51	0	-0.7	6.42
619	LTWEQANTAVKGVLDKVHSTGVEKLRDIYDKSVD	34	3816.28	-1	-0.51	5.61
620	SLMLHEIYPQTSPLQTAEEGKDPDKGP	27	2982.31	-3	-1.01	4.50
621	PDKGPRENLGPGLENED	17	1836.93	-3	-1.84	4.34
622	AKRSFQEGSASPYDLKEVLRVLSDWKALPSD KPLASGY	38	4210.76	1	-0.58	8.38
623	TQKMQKLRESLPYTSNVRDQAVQHLSNL REKVTPLVQDL	39	4565.23	2	-0.84	9.52
624	VEKIKHQESLVM	12	1440.72	0	-0.31	6.73
625	REGKLENGYRKSR	13	1592.77	3	-2.38	10.27
626	DIKARAEEREIK	12	1457.65	0	-1.52	6.26
627	LAKENYMQQVTQKIKDTLASFDMSGLAQQG RQWASEHLDAFQRYLEELKRTFTPSQAG	58	6692.5	0	-0.79	6.78
628	EPGEPTLMQKLEQFKENVKVFAD	23	2678.05	-2	-0.81	4.71
629	EPGEPTLMQKLEQFKENVKVFADNIGE	27	3091.48	-3	-0.8	4.57
630	ARSEPGEPTLMQKLEQFKENVKVFADNI GEKTKAALQELHD	41	4641.23	-2	-0.88	5.23
631	FIHNLRSEHNLP	12	1476.65	0	-0.74	6.92
632	EEVMQEAIDLKKVNM	15	1777.08	-2	-0.43	4.41
633	DRLEELREFANFDFDIWRKKYMRWMNH KKSRVMDFFRRI	39	5212.06	4	-1.11	10.14
634	LSDQIQRARQLEEERR	16	2027.22	0	-1.84	6.26
635	IQEQKLLQRLLDDRK	15	1896.22	1	-1.21	8.59
636	VEKLNQEPFKKN	12	1473.69	1	-1.67	8.47
637	ETNQRGKDLL	10	1173.29	0	-1.59	6.17
638	YRFGKELVQSRKYR	14	1830.12	4	-1.52	10.43
639	LQKQLLFAEFQKQHEHLTRQHEVQLQKHLKQQ	32	4020.61	2	-1.29	9.53
640	RPPGFTPFRSL	11	1274.48	2	-0.57	12.00
641	GKKKEMMEKWEKHWEWL	17	2303.72	1	-1.89	8.39
642	AHDPGRYYRA	10	1205.29	1	-1.67	8.64
643	TPVKIRIENSN	11	1270.45	1	-0.8	8.41
644	VRVSSYISWIERTIANN	17	2008.26	1	-0.06	8.72
645	YWIDGRVPEQVSKML	15	1821.12	0	-0.35	6.07
646	TLASFDMSGLAQQGRQWASEHLD	23	2548.77	-2	-0.52	4.74
647	DPLREQKDLAFAQAYLNRV	19	2247.54	0	-0.68	6.12
648	SCKMVLKEYITFKFKNESAINRRESTDL	28	3351.88	2	-0.65	9.03
649	ELEEKQVTMIQEK	13	1604.83	-2	-1.16	4.49
650	NLPWIEIQTKVGTRHWRQCKSRWLSV	26	3222.76	4	-0.72	10.92
651	EKEEARRKKEFLEKMEKA	18	2279.64	1	-2.08	8.47
652	AQQGRQWASEHLDAFQRYLEELKRTFTPSQAG	32	3750.1	0	-1.11	6.81
653	KMSEKQLKQENN	12	1476.66	1	-2.32	8.50
654	LIDYYESQINQMKKELRRY	19	2490.86	1	-1.26	8.38
655	TDLAKENYMQQVTQKIKDTLASFDMSGLA QQGRQWASEHLDAFQRYLEELKRTFTPSQAG	60	6908.69	-1	-0.84	5.62
656	QYMNELFSKGYREIKQ	16	2034.31	1	-1.29	8.43
657	VRDQAVQHLSNLREKVTPLVQDLRESIKP YTESIKTHLLNL	41	4783.51	1	-0.52	8.42
658	WKGSWYSLK	9	1154.33	2	-1.01	9.70
659	VEKINAAIYRPPS	13	1457.69	1	-0.3	8.56
660	IGSANHKESKITLFE	15	1673.88	0	-0.45	6.77
661	RLDRERMERERLERERMHIEQERRREQ	27	3723.17	1	-2.62	9.25

662	HFPLRSKYNRLTK	13	1659.95	4	-1.35	11.10
663	AQQGRQWASEHLDAFQRYLEEL	22	2675.9	-2	-1.09	4.83
664	SRIMEKTLS	9	1064.26	1	-0.44	8.46
665	IKPYTESIKTHLLNLFQEARKTLS	24	2831.31	2	-0.47	9.53
666	TPVFPGRRRGSTNLRASPG	19	2026.29	4	-0.94	12.48
667	FQSIREKICQKTQ	13	1608.87	2	-1.04	9.31
668	GWLDKNKDPLNETVVA	16	1799.01	-1	-0.69	4.56
669	EKRSLRWTRVNRDYTIYDTR	20	2628.93	3	-1.72	10.25
670	CLTHSQERK	9	1101.24	1	-1.53	8.23
671	SLMLHEIYPQTSPLQTAEEGKD	22	2487.76	-3	-0.75	4.56
672	QFSKNRVDLQTQ	12	1463.61	1	-1.38	8.75
673	LQEALHLIDQ	10	1179.33	-2	0.05	4.35
674	RVHTDGSVWRYVRASASYTPY	21	2471.71	2	-0.7	9.69
675	DTIQECIKSKYAPLSYFEEKEQNFEAVVKEL	31	3680.14	-3	-0.66	4.55
676	WQLEDLRQRYEQ	12	1663.81	-1	-2.05	4.68
677	LMLEVKKEAQLVLLN	15	1741.16	0	0.62	6.14
678	FKQEWLKKFWF	11	1586.9	2	-0.75	9.70
679	LAKENYMQQVTQKIKDTLASFDMSGLAQ QGRQWASEHLDAFQRYLEELKRTFTPSQAG	58	6692.5	0	-0.79	6.78
680	NQSLQKEMERVHVDNK	16	1955.17	0	-1.67	6.76
681	NIRLKIRQLPL	11	1250.55	3	-0.1	12.01
682	HKTHSQRTPGTRER	14	1690.84	3	-2.55	11.71
683	RIAIKEIPEKDIR	13	1580.89	1	-0.7	8.59
684	DEYIRKIQKRLEEDTFA	17	2154.4	-1	-1.32	4.94
685	LFQEREHVLRL	11	1439.68	0	-0.39	6.76
686	MGKIVNPTEK	10	1116.34	1	-0.69	8.35
687	PLRKEPEIITVTL	13	1508.82	0	0.06	6.57
688	ASRQQQQQQQQQQQQ	16	1998.06	1	-3.06	9.80
689	TLDIPVELEEQTMGKYNWATTPTTFK	26	3014.39	-2	-0.62	4.41
690	LQTKLKKLLGLESVF	15	1717.12	2	9.7	9.70
691	PPGASPRKKPRKQ	13	1446.71	5	-2.31	12.02

REFERENCES

1. E. Kosciuczuk *et al.*, Cathelicidins: family of antimicrobial peptides. A review., *Mol. Biol. Rep.* , 1–14 (2012).

2. L. Tomasinsig, M. Zanetti, The cathelicidins--structure, function and evolution., *Curr. protein Pept. Sci.* **6**, 23–34 (2005).

3. R. Bals, J. M. Wilson, Cathelicidins--a family of multifunctional antimicrobial peptides., *Cell. Mol. life Sci. C.* **60**, 711–720 (2003).

4. M. Zanetti, Cathelicidins, multifunctional peptides of the innate immunity., *J. Leukoc. Biol.* **75**, 39–48 (2004).

5. Y. Wang *et al.*, J. El Khoury, Ed. Snake Cathelicidin from Bungarus fasciatus Is a Potent Peptide Antibiotics, *PLoS One* **3**, 9 (2008).

6. H. Zhao *et al.*, Identification and characterization of novel reptile cathelicidins from elapid snakes, *Peptides* **29**, 1685–1691 (2008).

7. A. Giangaspero, L. Sandri, A. Tossi, Amphipathic alpha helical antimicrobial peptides., *Fed. Eur. Biochem. Soc. J.* **268**, 5589–5600 (2001).

8. F. A. De Latour, L. S. Amer, E. A. Papanstasiou, B. M. Bishop, M. L. Van Hoek, Antimicrobial activity of the Naja atra cathelicidin and related small peptides., *Biochem. Biophys. Res. Commun.* **396**, 825–830 (2010).

9. L. S. Amer, B. M. Bishop, M. L. Van Hoek, Antimicrobial and antibiofilm activity of cathelicidins and short, synthetic peptides against Francisella., *Biochem. Biophys. Res. Commun.* **396**, 246–251 (2010).

10. D. Wade *et al.*, All-D amino acid-containing channel-forming antibiotic peptides., *Proc. Natl. Acad. Sci. U. S. A.* **87**, 4761–4765 (1990).

11. S. Vunnam, P. Juvvadi, K. S. Rotondi, R. B. Merrifield, Synthesis and study of normal, enantio, retro, and retroenantio isomers of cecropin A-melittin hybrids, their end group effects and selective enzyme inactivation., *J. Pept. Res. Off. J. Am. Pept. Soc.* **51**, 38–44 (1998).

12. J. Lee, D. G. Lee, Structure-antimicrobial activity relationship between pleurocidin and its enantiomer, *Exp. Mol. Med.* **40**, 370–376 (2008).

13. S. C. D. N. Lopes, A. Fedorov, M. A. R. B. Castanho, Chiral recognition of D-kyotorphin by lipidic membranes: relevance toward improved analgesic efficiency., *ChemMedChem* **1**, 723–728 (2006).

14. H. Nakagawa, Y. Kobori, M. Yoshida, Koh-ichi, Chiral recognition by single bilayered phosphatidylcholine vesicles using [5]thiaheterohelicene as a probe , *Chem. Commun.* , 2692–2693 (2001).

15. M. I. Angelova, D. S. Dimitrov, Liposome electroformation, *Faraday Discuss. Chem. Soc.* **81**, 303 (1986).

16. T. G. D'Onofrio, C. W. Binns, E. H. Muth, C. D. Keating, P. S. Weiss, Controlling and Measuring Local Composition and Properties in Lipid Bilayer Membranes, *J. Biol. Phys.* **28**, 605–617 (2002).

17. S. D. Gillmor, P. S. Weiss, Dimpled vesicles: the interplay between energetics and transient pores., *J. Phys. Chem. B* **112**, 13629–13634 (2008).

18. S. Nappini, T. Al Kayal, D. Berti, B. Nordèn, P. Baglioni, Magnetically Triggered Release From Giant Unilamellar Vesicles: Visualization By Means Of Confocal Microscopy, *J. Phys. Chem. Lett.* **2**, 713–718 (2011).

19. E. Amstad *et al.*, Triggered release from liposomes through magnetic actuation of iron oxide nanoparticle containing membranes., *Nano Lett.* **11**, 1664–1670 (2011).

20. R. M. Epand, S. Rotem, A. Mor, B. Berno, R. F. Epand, Bacterial membranes as predictors of antimicrobial potency., *J. Am. Chem. Soc.* **130**, 14346–52 (2008).

21. S. N. Dean, B. M. Bishop, M. L. Van Hoek, Natural and synthetic cathelicidin peptides with anti-microbial and anti-biofilm activity against Staphylococcus aureus., *BMC Microbiol.* **11**, 114 (2011).

22. R. Bessalle, A. Kapitkovsky, A. Gorea, I. Shalit, M. Fridkin, All-D-magainin: chirality, antimicrobial activity and proteolytic resistance., *FEBS Lett.* **274**, 151–155 (1990).

23. S. M. Kelly, T. J. Jess, N. C. Price, How to study proteins by circular dichroism, *Biochim. Biophys. Acta* **1751**, 119–139 (2005).

24. J. Yao, G. Zhang, C.-H. Lei, Conformational Transformation Exhibited by the Peptide Extracted from Crystalline Region of Bombyx mori Silk Fibroin in Solid and Solution States, *Chinese J. Chem.* , 705–710 (2006).

25. M. C. Manning, R. W. Woody, Theoretical CD studies of polypeptide helices: examination of important electronic and geometric factors., *Biopolymers* **31**, 569–586 (1991).

26. F. Formaggio, C. Toniolo, Electronic and vibrational signatures of peptide helical structures: A tribute to Anton Mario Tamburro, *Chirality* **22**, 30–39 (2010).

27. C. Toniolo, A. Polese, F. Formaggio, M. Crisma, J. Kamphuis, Circular Dichroism Spectrum of a Peptide 3 (10) -Helix, *J. Am. Chem. Soc.* **118**, 2744–2745 (1996).

28. Z. Biron *et al.*, A monomeric 3(10)-helix is formed in water by a 13-residue peptide representing the neutralizing determinant of HIV-1 on gp41., *Biochemistry* **41**, 12687–96 (2002).

29. L. G. J. Hammarstrom, T. J. Gauthier, R. P. Hammer, M. L. McLaughlin, Amphipathic control of the 3(10)-/α-helix equilibrium in synthetic peptides, *J. Pept. Res.* **58**, 108–116 (2001).

30. S. Miick, G. Martinez, W. Fiori, A. Todd, Short alanine-based peptides may form 310-helices and not α-helices in aqueous solution, *Lett. To Nat.* **359**, 653–655 (1992).

31. N. Sreerama *et al.*, Tyrosine, Phenylalanine, and Disulfide Contributions to the Circular Dichroism of Proteins: Circular Dichroism Spectra of Wild-Type and Mutant Bovine Pancreatic Trypsin Inhibitor, *Biochemistry* **38**, 10814–10822 (1999).

32. J. E. Cummings, T. K. Vanderlick, Aggregation and hemi-fusion of anionic vesicles induced by the antimicrobial peptide cryptdin-4., *Biochim. Biophys. Acta* **1768**, 1796–804 (2007).

33. T. Kiyota, S. Lee, G. Sugihara, Design and synthesis of amphiphilic alpha-helical model peptides with systematically varied hydrophobic-hydrophilic balance and their interaction with lipid- and bio-membranes., *Biochemistry* **35**, 13196–204 (1996).

34. A. Hammond *et al.*, Crosslinking a lipid raft component triggers liquid ordered-liquid disordered phase separation in model plasma membranes, *Proc. Natl. Acad. Sci.* **102**, 6320–6325 (2005).

35. D. Lingwood, J. Ries, P. Schwille, K. Simons, Plasma membranes are poised for activation of raft phase coalescence at physiological temerature, *Proc. Natl. Acad. Sci.* **105**, 10005–10010 (2008).

36. S. Pierce, W. Liu, The tipping points in the initiation of B cell signalling: how small changes make big differences, *Nat. Rev. Immunol.* **10**, 767–777 (2010).

37. P. Tolar, Inside the microcluster: antigen receptor signalling viewed with molecular imaging tools, *Immunology* **133**, 271–277 (2011).

38. P. Tolar, J. Hanna, P. Krueger, S. Pierce, The Constant Region of the Membrane Immunoglobulin Mediates B Cell-Receptor Clustering and Signaling in Response to Membrane Antigens, *Immunity* **30**, 44–55 (2009).

39. K. A. Brogden, Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria?, *Nat. Rev. Microbiol.* **3**, 238–250 (2005).

40. R. Rathinakumar, W. F. Walkenhorst, W. C. Wimley, Broad-spectrum antimicrobial peptides by rational combinatorial design and high-throughput screening: the importance of interfacial activity., *J. Am. Chem. Soc.* **131**, 7609–17 (2009).

41. V. Teixeira, M. J. Feio, M. Bastos, Role of lipids in the interaction of antimicrobial peptides with membranes., *Prog. Lipid Res.* **51**, 149–77 (2012).

42. M. Yeaman, N. Yount, Mechanisms of Antimicrobial Peptide Action and Resistance, *Pharmacol. Rev.* **55**, 27–55 (2003).

43. Y. Shai, Mode of action of membrane active antimicrobial peptides, *Biopolym. (Peptide Sci.* **66**, 236–248 (2002).

44. D. Sengupta, H. Leontiadou, A. E. Mark, S. J. Marrink, Toroidal pores formed by antimicrobial peptides show significant disorder, *Biochim. Biophys. Acta - Biomembr.* **1778**, 2308–2317 (2008).

45. M. Juba, D. Porter, S. Dean, S. Gillmor, B. Bishop, Characterization and performance of short cationic antimicrobial peptide isomers., *Biopolymers* **100**, 387–401 (2013).

46. M. U. Shiloh, J. Ruan, C. Nathan, Evaluation of bacterial survival and phagocyte function with a fluorescence-based microplate assay., *Infect. Immun.* **65**, 3193–3198 (1997).

47. A. Mariscal, R. M. Lopez-Gigosos, M. Carnero-Varo, J. Fernandez-Crehuet, Fluorescent assay based on resazurin for detection of activity of disinfectants against bacterial biofilm., *Appl. Microbiol. Biotechnol.* **82**, 773–783 (2009).

48. M. L. Mangoni *et al.*, Effects of the antimicrobial peptide temporin L on cell morphology, membrane permeability and viability of Escherichia coli., *Biochem. J.* **380**, 859–65 (2004).

49. B. L. Roth, M. Poot, S. T. Yue, P. J. Millard, Bacterial viability and antibiotic susceptibility testing with SYTOX green nucleic acid stain., *Appl. Environ. Microbiol.* **63**, 2421–2431 (1997).

50. J. Y. Kim *et al.*, Novel Antibacterial Activity of β2-Microglobulin in Human Amniotic Fluid, *PLoS One* **7** (2012).

51. M. Kalab, A. Yang, D. Chabot, Conventional scanning electron microscopy of bacteria, *Infocus Mag.* (2008) (available at http://www.rms.org.uk/Resources/Royal Microscopical Society/infocus/Images/KALAB ARTICLE.pdf).

52. D. M. E. Bowdish *et al.*, Impact of LL-37 on anti-infective immunity., *J. Leukoc. Biol.* **77**, 451–459 (2005).

53. M. J. Goldman *et al.*, Human beta-defensin-1 is a salt-sensitive antibiotic in lung that is inactivated in cystic fibrosis., *Cell* **88**, 553–560 (1997).

54. J. P. Tam, Y.-A. Lu, J.-L. Yang, Correlations of cationic charges with salt sensitivity and microbial specificity of cystine-stabilized beta -strand antimicrobial peptides., *J. Biol. Chem.* **277**, 50450–50456 (2002).

55. M. Wu, E. Maier, R. Benz, R. E. Hancock, Mechanism of interaction of different classes of cationic antimicrobial peptides with planar bilayers and with the cytoplasmic membrane of Escherichia coli., *Biochemistry* **38**, 7235–42 (1999).

56. K. Xiong, E. K. Asciutto, J. D. Madura, S. a Asher, Salt dependence of an alpha-helical peptide folding energy landscapes., *Biochemistry* **48**, 10818–26 (2009).

57. J. S. Smith, J. M. Scholtz, Energetics of polar side-chain interactions in helical peptides: salt effects on ion pairs and hydrogen bonds., *Biochemistry* **37**, 33–40 (1998).

58. B. V Sweet, A. K. Schwemm, D. M. Parsons, Review of the processes for FDA oversight of drugs, medical devices, and combination products., *J. Manag. care Pharm. JMCP* **17**, 40–50 (2011).

59. C. Friedrich, M. G. Scott, N. Karunaratne, R. E. W. Hancock, Salt-Resistant Alpha-Helical Cationic Antimicrobial Peptides, *Antimicrob. Agents Chemother.* **43**, 1542–1548 (1999).

60. P. J. Sims, a S. Waggoner, C. H. Wang, J. F. Hoffman, Studies on the mechanism by which cyanine dyes measure membrane potential in red blood cells and phosphatidylcholine vesicles., *Biochemistry* **13**, 3315–30 (1974).

61. Y. Wang, Z. Zhou, J. Zhu, Y. Tang, T. Canady, Dark antimicrobial mechanisms of cationic phenylene ethynylene polymers and oligomers against Escherichia coli, *Polymers (Basel).* **1**, 3–6 (2011).

62. M. Hartmann *et al.*, Damage of the bacterial cell envelope by antimicrobial peptides gramicidin S and PGLa as revealed by transmission and scanning electron microscopy., *Antimicrob. Agents Chemother.* **54**, 3132–42 (2010).

63. S. Klayraung, S. Okonogi, Antibacterial and antioxidant activities of acid and bile resistant strains of Lactobacillus fermentum isolated from miang, *Brazilian J. Microbiol.* , 757–766 (2009).

64. E. R. Fischer, B. T. Hansen, V. Nair, F. H. Hoyt, D. W. Dorward, *Scanning electron microscopy.* (2012; http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3352184&tool=pmcentrez&rendertype=abstract), p. Unit 2B.2.

65. R. E. W. Hancock, H.-G. Sahl, Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies., *Nat. Biotechnol.* **24**, 1551–1557 (2006).

66. M. Zasloff, Antimicrobial peptides of multicellular organisms., *Nature* **415**, 389–395 (2002).

67. M. E. Merchant *et al.*, Broad spectrum antimicrobial activity of leukocyte extracts from the American alligator (Alligator mississippiensis)., *Vet. Immunol. Immunopathol.* **110**, 221–228 (2006).

68. M. E. Merchant, C. Roche, R. M. Elsey, J. Prudhomme, Antibacterial properties of serum from the American alligator (Alligator mississippiensis)*Comp. Biochem. Physiol. Part B Biochem. Mol. Biol.* **136**, 505–513 (2003).

69. L. N. F. Darville, M. E. Merchant, A. Hasan, K. K. Murray, Proteome analysis of the leukocytes from the American alligator (Alligator mississippiensis) using mass spectrometry., *Comp. Biochem. Physiol. Part D. Genomics Proteomics* **5**, 308–316 (2010).

70. S. Pata *et al.*, Characterization of the novel antibacterial peptide Leucrocin from crocodile (Crocodylus siamensis) white blood cell extracts., *Dev. Comp. Immunol.* **35**, 545–553 (2011).

71. A. Luchini *et al.*, Smart hydrogel particles: biomarker harvesting: one-step affinity purification, size exclusion, and protection against degradation., *Nano Lett.* **8**, 350–361 (2008).

72. C. Longo *et al.*, Core-shell hydrogel particles harvest, concentrate and preserve labile low abundance biomarkers., *PLoS One* **4**, e4763 (2009).

73. S. Lata, B. Sharma, G. Raghava, Analysis and prediction of antibacterial peptides, *BMC Bioinformatics* **10**, 1–10 (2007).

74. S. Thomas, S. Karnik, R. S. Barai, V. K. Jayaraman, S. Idicula-Thomas, CAMP: a useful resource for research on antimicrobial peptides., *Nucleic Acids Res.* **38**, D774–80 (2010).

75. G. Wang, X. Li, Z. Wang, APD2: the updated antimicrobial peptide database and its application in peptide design., *Nucleic Acids Res.* **37**, D933–D937 (2009).

76. J. a St John *et al.*, Sequencing three crocodilian genomes to illuminate the evolution of archosaurs and amniotes., *Genome Biol.* **13**, 415 (2012).

77. O. Sørensen, K. Arnljots, J. B. Cowland, D. F. Bainton, N. Borregaard, The human antibacterial cathelicidin, hCAP-18, is synthesized in myelocytes and metamyelocytes and localized to specific granules in neutrophils., *Blood* **90**, 2796–2803 (1997).

78. J. E. P. Syka, J. J. Coon, M. J. Schroeder, J. Shabanowitz, D. F. Hunt, Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry, *Proc. Natl. Acad. Sci. U. S. A.* **101**, 9528–9533 (2004).

79. L. M. Mikesh *et al.*, The utility of ETD mass spectrometry in proteomic analysis., *Biochim. Biophys. Acta* **1764**, 1811–1822 (2006).

80. G. Rose, A. Geselowitz, G. Lesser, Hydrophobicity of amino acid residues in globular proteins, *Science (80-.).* **229**, 834–838 (1985).

81. R. E. Hancock, M. G. Scott, The role of antimicrobial peptides in animal defenses., *Proc. Natl. Acad. Sci. U. S. A.* **97**, 8856–61 (2000).

82. W. Van T Hof, E. C. Veerman, E. J. Helmerhorst, A. V Amerongen, Antimicrobial peptides: properties and applicability., *Biol. Chem.* **382**, 597–619 (2001).

83. C. Stegemann *et al.*, Isolation, purification and de novo sequencing of TBD-1, the first beta-defensin from leukocytes of reptiles., *Proteomics* **9**, 1364–73 (2009).

84. N. D. Udeshi, P. D. Compton, J. Shabanowitz, D. F. Hunt, K. L. Rose, Methods for analyzing peptides and proteins on a chromatographic timescale by electron-transfer dissociation mass spectrometry., *Nat. Protoc.* **3**, 1709–1717 (2008).