THE UNIVERSITY OF RHODE ISLAND
COLLEGE OF THE ENVIRONMENT AND LIFE SCIENCES

Reliable Data Sharing on Antimicrobial Peptide Efficacy

Lenore M. Martin[a], Katelyn Raimond[a], Kimberly Kluglein[a], Husam Bahra[a], Atalia Rodrigues[a], Matthew Daily[b], Emily Costello[c], Archana Chittoor[b], and Joseph Janiszewski[b] of the *Departments of Cell and Molecular Biology*[a], *Computer Science and Statistics*[b], *and Biomedical and Pharmaceutical Sciences*[c] *at The University of Rhode Island, Kingston, RI 02881 USA*

Abstract

Reliable and consistent experimental data may be shared and compared describing the effectiveness of Antimicrobial Peptides (AMPs) against well-defined pathogens using software tools we are developing and sharing on the antimicrobial peptide editable (AMPed.uri.edu) database platform. Antimicrobial resistance is a growing problem in treatment of bacterial infections. To fight back against resistance, our team tested six synthetic antimicrobial peptides (AMPS) at various concentrations against 2 strains of bacteria known to cause infection, *Escherichia coli* and *Staphylococcus aureus*. In the 96-well plate format, we used two controls: NaOH and chloramphenicol in Müeller-Hinton broth assays. To analyze the growth or inhibition of the bacteria after treatment with AMPs, serial dilutions of each AMP were prepared using the "Hancock diluent"[1] consisting of 0.01 % acetic acid and 0.2% BSA. We monitored growth in real time using the optical density at 630 nm but also determined the inhibition of bacterial growth more accurately by plating different dilutions of bacterial cultures and counting the colony forming units (CFUs). Based on these results, we concluded that among the six AMPs tested, peptides LM4-15 and CP-29A caused the most potent inhibition in the growth of these 2 bacterial species, and verified the sequences of LM4-15 and CP-29A by comparing the theoretically calculated molecular masses to those obtained using matrix-assisted laser desorption mass spectrometry (MALDI). Future experiments will explore the activity of these peptides against clinical isolates of antibiotic-resistant *Staphylococcus aureus* at Rhode Island Hospital.

[1]http://cmdr.ubc.ca/bobh/method/modified-mic-method-for-cationic-antimicrobial-peptides/ last accessed on 2019-05-17

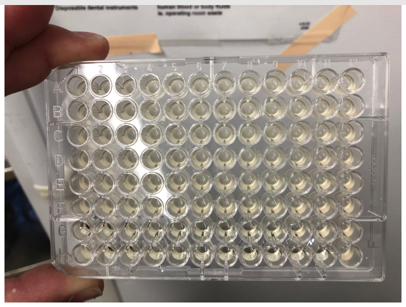
Computational Objectives

1) *Development of algorithms for compositional analysis or "MassSpec Inversion"*

To improve reproducibility across different labs, some simple investigations into the computational feasibility of inferring the composition of a synthetic antimicrobial peptide from its mass spectrum were performed. Specifically, given the masses of the 20 different amino acids that can make up these peptides, and comparing all possible AA compositions with the measured mass of a peptide obtained by MALDI-MS or ES-MS, can we figure out the most probable composition of that peptide.

1) *Facilitating 96-Well Plate MIC Analysis*

The standard way of testing an antimicrobial's potency is to measure the MIC in a broth dilution growth curve analysis. This measures the growth of bacteria in a 96-well plate to see the minimum concentration of antimicrobial peptide that starts inhibiting the growth of a bacterium. The purpose of this software was to facilitate data transfer for Martin Labs and their collaborators worldwide, which currently scans the 96-well-plates with instrumentation running SoftMax® Pro Version 15.



Additional Pleurocidin Analogs Synthesized by SPPS

Code	N-Term AA Sequence C-Term	# AA
LM4-12	H-GWGSFFKKAAHVGKHVGKAALTHYL-NH2	[25 AA]
LM4-13	H-WGWGSFFKKAAHVGKHVGKAALTHYL-NH2	[26 AA]
LM4-14	H-GWGWGSFFKKAAHVGKHVGKAALHYL-NH2	[27 AA]
LM4-15	H-RGWGSFFKKAAHVGKHVGKAALTHYL-NH2	[26 AA]
LM4-16	Ac-WGWGSFFKKAAHVGKHVGKAALTHYL-NH2	[26 AA]

Designed Peptides Tested: Pleurocidin, a small, amphipathic α-helical peptide, 25 -residues long, was originally isolated from mucous secretions of winter flounder, and exhibits moderately potent but broad-spectrum antibacterial activity against a variety of both gram-negative and gram-positive human pathogens. Many of the genes coding for Pleurocidin and related flatfish AMPs have been mapped and sequenced. The pleurocidin gene expressed from a TATA-box promoter in *Pleuronectes americanus* has four exons. The 5'- and 3'-untranslated regions of the mature transcript flank a 3-domain propeptide consisting of 22-residue *N*-terminal signal sequence, the 25-residue sequence of active pleurocidin, and a 21-residue *C*-terminal sequence. We designed a series of analogs of pleurocidin to test our hypotheses about which regions of the peptide are necessary for optimal activity. We also explored whether we could increase the potency of pleurocidin against *Escherichia coli* and *Staphylococcus aureus* by altering the hydrophobicity of the N-terminus. CP-29 was synthesized and used to compare our results with those of Hancock *et al.*

Algorithm #1: The machine learning algorithm is a "MassSpec inversion". This is a special form of the classical Knapsack problem in computer science which, simply stated, asks given a set of items each with a volume and a value, what is the maximum value you can fit into a knapsack of a given total volume. In this case, the value is the weight of the peptide and the task is to get exactly the volume of the knapsack as the value. This problem is known to be NP complete and thus can only be definitively solved with a brute force search, which grows exponentially with the size of the problem, in this case, the number of acids in the peptide, which is upper bounded by the ratio of the total weight of the peptide to the lightest amino acid type.

Here we are saved by the fact that the number of acids in the peptides of interest is limited to about 24. Still, the number of potential combinations of 24 acids is approximately 19^24 or or 4x10^30 combinations. This amount of search is only borderline tractable on the architectures we have available today so the algorithm developed here has been heavily optimized (and will continue to be optimized).

The algorithm implemented is a recursive algorithm where each depth of the recursion corresponds to an amino acid type. The recursive function takes these arguments: the current weight, the target weight, the amino acid type, and a vector of counts of each type already assigned. The function cycles through all possible counts of the type of amino acid at his recursion depth, updates the weight to reflect that number, and compares it to the target. If it is over, it returns to the caller. If it is equal, it saves the composition to a file and returns. If it is less, it invokes itself for the next amino acid type. The number of operations of each test is thus quite small, and the memory footprint is negligible.

Algorithm #2: The SSEAPS statistical analysis package is tasked with uploading 96 well plate data and do automated analysis of that data prior to submitting it to the AMPED databases. The R script first loads the file from the user and separates each plate into its own data frame (matrix). It finds the average of the blanks/broth in column one from the first six rows for each plate and subtracts this value from each peptide well (rows 1-6, columns 2-11), antibiotic wells (row 8, columns 2-11). Then it checks to see if any of the data for the 6 data points of each peptide at each concentration are within 3 standard deviations of the mean. After removing any points, the 6 (or less) data points from each peptide at each concentration are averaged and added to a new data frame. There is one data frame for each peptide that will have 10 columns, one column for each concentration, and as many rows as there are time points. There is also a data frame for the antibiotic row which will also have 10 columns and as many rows as time points and one other data frame for the bacteria that only has 1 column with the number of rows equal to time points. After this process is complete the program plots the growth for each concentration at the first time point after 1440 minutes. Specifically, it computes mean and variance across 10 wells for each concentration and forms two types of plots. First is the total growth over the entire period for each of the concentrations tested. This is compared to a reference antibiotic, plotted below the first two.

Results and Conclusions

Peptide Length	Serial Algorithm	Parallel Algorithm
6	0.170	0.852
7	0.758	2.975
8	3.005	5.418
9	11.088	10.151
10	36.320	16.156
11	106.701	32.041

Table 1: Peptide Compostion Run Times

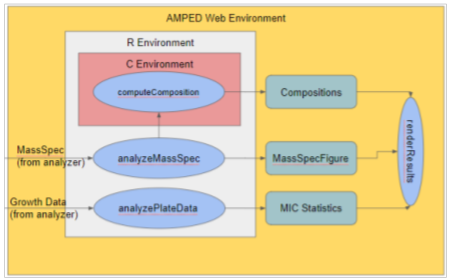
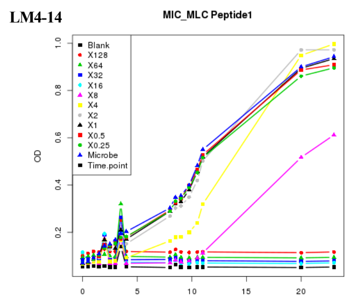


Figure 3: Data Flow and Software Environment

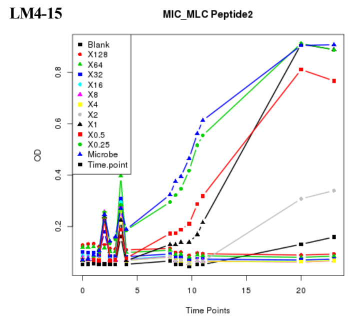
Results and Conclusions (cont)

Figure 2: Growth Curve Data

LM4-14



LM4-15



SSEAPs was designed with the AMPed website and Martin Labs in mind. Both the MIC and mass spectrometer programs work well, though the mass spectrometer script does not run quite as quickly as the authors would like. However, both programs are far superior to the versions currently found on the AMPed website. This is especially in the case of the mass spectrometer program, as the link currently redirects to the poorly-functioning MIC calculator. The project involved changing the current coding in the AMPed website to implement the two tools.

The goals set out by the project proposal were all achieved – the ideal goal was the creation and implementation of both the MIC and mass spectrometry, which were both accomplished. The team worked well together to ensure that these projects received equal time and effort, with each member contributing according to their own areas of expertise. Overall, the project was a success.

References

Cole et al. *Characterization of Fish Antimicrobial Peptide: Gene Expression, Subcellular Localization, and Spectrum of Activity.* Antimicrobial Agents and Chemotherapy (2000), pp 2039-2045.

Hancock et al. *Antimicrobial and Host-Defense Peptides as New Anti-infective Therapeutic Strategies.* Nature Biotechnology, 12, (2006), pp 1551-1557.

Acknowledgements

We gratefully acknowledge encouragement and financial support from the Office of Undergraduate Research and Innovation at URI to Kaitlyn, Kimberly, Hussam, and Atalia. URI data science Professors Noah Daniels and Christopher Hemme provided invaluable encouragement and critiques for the algorithm development program in CMB/CSC/STA 522/BPS542.