# The Production, Design, and Application of Antimicrobial Peptides

by

Christopher Loose

B.S.E., Chemical Engineering

Princeton University, 2002

Submitted to the Department of Chemical Engineering
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Chemical Engineering

at the

Massachusetts Institute of Technology

24 April 2007
[June 2007]



Signature of Author .....
Christopher Loose
Department of Chemical Engineering
24 April, 2007

Certified by .....
Gregory Stephanopoulos
Professor of Chemical Engineering
Thesis Supervisor

Accepted by .....
William Deen
Professor of Chemical Engineering
Chairman, Committee for Graduate Students

-- Blank Page --

# The Production, Design, and Application of Antimicrobial Peptides

by

Christopher Loose

Submitted to the Department of Chemical Engineering on April 24th, 2007
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Chemical Engineering

## Abstract

With the spread of drug-resistant bacteria, existing antibiotics are losing their potency. Antimicrobial peptides (AmPs) represent an exciting class of drug candidates, particularly because their mechanism of action is unlikely to induce drug resistance. If resistance to AmPs were also slower to emerge in the clinic, they would have longer useful lifetimes than existing antibiotics. Nevertheless, a number of limitations exist for AmPs in the clinic. The high cost of peptide manufacture requires that highly potent sequences are created. Additionally, AmP selectivity must be improved if effective systemic doses are to be given without hemolytic activity or other toxicity. Improved high-throughput methods for AmP design or discovery could enable the achievement of both of these goals. To this end, we developed an approach based on the discovery of semi-conserved motifs across natural AmPs, which we demonstrated are associated with antimicrobial activity. Additionally, we created novel AmP formulations that may bypass some of these clinical limitations.

In order to evaluate AmP design approaches, a high-throughput production and assay platform was created using *in vitro* translation. This technology may produce peptides that would be toxic to recombinant hosts and synthesize peptides of arbitrary length. The cost per peptide was minimized through a series of process improvements. First, we created methods to construct oligonucleotides that mimicked our motif-based design of AmPs. This approach allowed the reuse of primers for many peptides, reducing cost and enabling the study of pattern synergy. Additionally, we found peptide translation was enhanced by co-translating a fusion partner in frame with the AmP. The AmP could be freed from the fusion partner after translation using enterokinase digestion. Further, we increased yield 3-fold by optimizing the length of fusion partner. The partner was made as short as possible to limit the translational resources required to synthesize the fusion partner, while being long enough to ensure stability from proteases. The solubility of the fusion partner-AmP construct was also improved through the selection of a highly soluble partner of the optimal length. Finally, we developed a purification scheme to ensure that the *in vitro* translation extract would not impact measurement of antimicrobial activity.

We also developed and evaluated the design of AmPs using semi-conserved motifs. We used a database of over 500 natural AmPs as a training set for pattern

discovery. The resulting motifs were exhaustively recombined to create all 20 amino acid sequences that were entirely covered by these patterns. These sequences were clustered, and 42 diverse members selected for characterization using representative Gram negative and Gram positive bacteria. Approximately 50% of the designed AmPs were active against at least one of the bacteria at 256 ug/ml. Control peptides were created in which the amino acids in the designed peptides were rearranged such that they were not homologous to any antimicrobial motifs. Thus, these controls had the same bulk physiochemical properties frequently associated with antimicrobial activity as the designed sequences, but we hypothesized they would not be active because they did not match the antimicrobial motifs. In fact, only 5% of the control sequences had activity at 256 ug/ml, indicating that the antimicrobial motifs give a 10-fold enrichment in activity. Further, two highly active designed peptides had MICs of 16 ug/ml against *Bacillus cereus* and 64 ug/ml against *Escherichia coli.* Additionally, AmPs active against *B. cereus* were all active against the hospital pathogen *Staphylococcus aureus*, and the bioterror agent, *Bacillus anthracis*.

Our motif-based design may be most effective as the first stage of a two-stage design tool. In the first stage, highly diverse leads with novel profiles are created and evaluated. Promising leads could then be optimized using a variety of techniques. By creating just 44 variants of one lead, we designed an AmP with broad spectrum activity that had MICs of 16 ug/ml against *E. coli* and 8 ug/ml against *B. cereus* and 4 ug/ml against *S. aureus*. Another approach to build on our design tool would be to incorporate activity and toxicity characteristics of members of the training set into the design or scoring of new sequences. In order to begin assembling this data using a standardized method, a representative set of 100 natural, linear AmPs was chosen through clustering. Their antimicrobial activity against *E. coli, S. epidermidis,* and *S. aureus* were evaluated, along with hemolytic activity. When further supplemented, this information may enable an improved scoring metric to be created. Additionally, we systematically demonstrated that amidating the c-terminus of natural AmPs improves both antimicrobial activity and therapeutic index.

Finally, we recognized that AmP's mechanism of action would allow activity to be retained when they are permanently tethered to medical device surfaces. Unlike existing coatings which rely on the slow release of silver or other antibiotics, a permanently tethered approach could have a longer lifetime and reduced systemic toxicity concerns. A versatile chemistry was developed to create immobilized AmP coatings. These formulations had broad spectrum antimicrobial activity without significant hemolytic activity. Further, the coatings were effective through multiple bacterial challenges. The combination of the AmP design tool along with localized formulations represent a significant advance in the process of moving AmPs to the clinic to combat drug-resistant infections.

Thesis Supervisors:
Gregory Stephanopoulos, Bayer Professor of Chemical Engineering
Robert Langer, Institute Professor

# Acknowledgments

I am deeply grateful for the support of the following people:

- Mom and Dad for encouraging me on my quest to become an enormous nerd.
- My wife, Krista, for unwavering support and sharing my love of peptides.
- Greg Stephanopoulos for your support and vision of the potential impact of this work.
- Robert Langer for encouraging me to think big and guiding me through the entrepreneurial process.
- Joel Moxley for talking me into this crazy grad school idea and for years of friendship and pitbull-like loyalty.
- Geoff Loose for support, loyalty, and Black and Tans.
- Kyle Jensen for your enthusiasm and persistence in all of this work.
- David Lucchino for driving the SteriCoat mini-van.
- Michael Hencke for going that inch.
- W. Shannan O'Shaughnessy, Lino Ferreira, and Trevor Squire for instilling deadly properties to otherwise harmless polymers.
- Dane Wittrup for probing questions and guidance along the way.
- Tanguy Chau for your endless enthusiasm for killing bacteria.
- Curt Fischer, Keith Tyo, Mark Styczynski, Jose Aleman, and Hal Alper for fun and support in the office and lab.
- Brian Mickus for your friendship, spirited debate, and always doing the dishes.
- Michael Jewett, Kim Woodrow, and Jim Swartz for sharing your time and expertise in *in vitro* translation.
- John and Fannie Hertz Foundation for financial support and giving me the flexibility to pursue my ideas.

# Contents

*Abstract* ..... *3*

*Acknowledgments* ..... *5*

*List of Figures* ..... *13*

*List of Tables* ..... *16*

*1. Introduction* ..... *17*

**1.1 Motivation** ..... **17**

**1.2 Objectives and approach** ..... **19**

**1.3 Thesis organization** ..... **21**

*2. Antimicrobial Peptide Overview* ..... *22*

**2.1 Antimicrobial peptides in nature** ..... **22**

2.1.1 AmPs in innate immunity ..... 23

2.1.2 Classes of AmPs ..... 24

2.1.3 Mechanism of action ..... 26

2.1.4 Antibiotic resistance ..... 29

**2.2 Screening for AmPs** ..... **29**

2.2.1 Screening of natural organisms ..... 30

2.2.2 Unguided approaches to produce AmPs ..... 31

**2.3 Rational Design and Optimization of AmPs** ..... **32**

2.3.1 Mutations around natural AmPs ..... 33

2.3.2 Computational approaches to AmP design ..... 34

**2.4 AmPs as antibiotics** ..... **36**

2.4.1 AmPs in clinical development ..... 36

2.4.2 Current difficulties in development ..... 39

**2.5 Summary** ..... **41**

***3. In vitro Translation of AmPs*** ..... ***42***

**3.1 Approaches to peptide synthesis** ..... **42**

**3.2 Overview of *in vitro* translation** ..... **44**

**3.3 Desired properties for high-throughput screen** ..... **45**

**3.4 Combinatorial DNA design and synthesis** ..... **48**

3.4.1 Initial methods for DNA construction ..... 48

3.4.2 Drawbacks with prior methods ..... 51

3.4.3 Combinatorial DNA construction ..... 51

3.4.3.1 Combinatorial oligonucleotide design ..... 52

3.4.3.2 Combinatorial oligonucleotide synthesis ..... 56

3.4.3.3 Improved combinatorial design using ligation ..... 60

3.4.4 Summary ..... 62

**3.5 Initial AmP translation studies** ..... **63**

3.5.1 Methods for in vitro translation in the Roche system ..... 63

3.5.2 Initial in vitro translation results ..... 64

**3.6 Multimer AmPs** ..... **67**

**3.7 Fusion partner length optimization** ..... **70**

3.7.1 Background on fusion partners ..... 70

3.7.2 Optimizing length of GFP fusion partner ..... 72

3.7.2.1 Template construction for truncated GFP ..... 72

3.7.2.2 Effect of translated product length of yield ..... 77

3.7.3 Effect of translation system on optimal length ..... 79

3.7.3.1 Methods for translation in the KC6 and NMR5 extracts ..... 79

3.7.3.2 Optimal fusion partners in alternate *in vitro* systems ..... 81

3.7.4 Effect of fusion partner sequence on optimal length ..... 81

3.7.4.1 Methods for creating and translating truncated luciferase ..... 81

3.7.4.2 Optimal translated product length for truncated luciferase ..... 82

3.7.5 Translating with a desired AmP in frame ..... 84

3.7.5.1 Yield with AmP translated in frame ..... 84

3.7.5.2 Digestion from fusion partner ..... 86

**3.8 Exploration of protease activity ..... 88**

3.8.1 Translated product length defines optimal yield ..... 89

3.8.1.1 Methods for varying amino acid length independently of mRNA length ..... 90

3.8.1.2 Results of varying amino acid length independently of mRNA length ..... 92

3.8.2 Protease inhibitor studies ..... 92

3.8.2.1 Methods for investigating effect of protease inhibitors on degradation rates ..... 92

3.8.2.2 Methods for investigating effect of protease inhibitors on degradation rates ..... 95

3.8.3 Application of fusion partner optimization ..... 96

**3.9 Improving solubility ..... 97**

3.9.1 Measuring solubility of existing templates ..... 97

3.9.2 Selection of appropriate partner ..... 99

3.9.2.1 Methods to create and translate a Trx fusion partner ..... 99

3.9.2.2 Results of initial Trx translations ..... 102

3.9.3 Creating high-yield Trx translations from low DNA template concentrations ..... 102

3.9.3.1 Methods for creating Trx on GFP upstream region ..... 103

3.9.3.2 Effect of upstream region on Trx yield at low DNA concentrations ..... 106

**3.10 Developing purification scheme ..... 106**

3.10.1 Protein precipitation to remove toxic in vitro translation buffer component ..... 107

3.10.2 Engineering toxicity out of in vitro system ..... 109

3.10.3 High-throughput isolation of AmPs ..... 111

3.10.3.1 His-tag purification ..... 111

3.10.3.2 Dialysis to remove imidazole ..... 112

3.10.4 Activity of purified AmP ..... 114

**3.11 Summary ..... 116**

***4. AmP Design Based on Pattern Discovery ..... 118***

**4.1 Pattern-based AmP design ..... 118**

4.1.1 Patterns in natural AmPs ..... 119

4.1.2 Elucidation of grammars ..... 123

4.1.2.1 *Database selection* ..... 124

4.1.2.2 Teiresias ..... 124

4.1.2.3 Application of Teiresias to AmPs ..... 127

4.1.2.4 Filtering for selectivity ..... 128

4.1.3 Designing novel AmPs from grammars ..... 130

4.1.3.1 *In silico* evolution ..... 131

4.1.3.2 Combinatorial enumeration ..... 135

4.1.3.3 Clustering and scoring of designs ..... 138

4.1.3.4 Designing negative controls ..... 140

**4.2 Experimental validation ..... 141**

4.2.1 Peptide synthesis ..... 141

4.2.1.1 *Recombinant production* ..... 142

4.2.1.2 *In vitro* translation ..... 142

4.2.1.3 Chemical synthesis ..... 143

4.2.2 Antimicrobial assays ..... 147

4.2.2.1 Plate counting ..... 148

4.2.2.2 Fluorescent reporting ..... 149

4.2.2.3 Microdilution assay ..... 154

4.2.3 MICs of natural AmPs ..... 158

4.2.4 MICs of control sequences ..... 160

4.2.5 MICs of designed sequences ..... 163

4.2.6 MICs of shuffled sequences ..... 165

4.2.7 Killing curves and bactericidal measurements ..... 168

**4.3 High value targets ..... 171**

4.3.1 Gram positive targets ..... 171

4.3.2 Antimicrobial activity against S. aureus and B. anthracis ..... 173

**4.4 Diversity of designed sequences ..... 173**

**4.5 Summary ..... 174**

***5. Refining AmP Design ..... 176***

**5.1 Optimization of activity ..... 176**

5.1.1 Helical projection based optimization ..... 177

5.1.2 Antimicrobial activity of optimized sequences ..... 180

**5.2 Evaluating mammalian toxicity ..... 181**

5.2.1 Methods for assessing hemolytic activity ..... 181

**5.3 Characterizing natural AmPs ..... 182**

5.3.1 Screening for antimicrobial activity ..... 183

5.3.2 Antimicrobial activity of natural AmPs ..... 185

**5.4 Correlation of antimicrobial activity and toxicity ..... 186**

5.4.1 Screening for hemolytic activity ..... 186

5.4.2 Correlation of hemolytic and antimicrobial activity ..... 188

5.4.3 Distribution of therapeutic indices for natural AmPs ..... 188

**5.5 Effect of c-terminal amidation ..... 190**

5.5.1 Effect of amidation of representative AmPs ..... 190

5.5.2 Effect of amidation on a broader set of natural AmPs ..... 192

5.5.3 Effect of amidation on hemolytic activity ..... 196

**5.6 Summary ..... 198**

***6. Medical Device Applications ..... 199***

**6.1 Infections associated with medical devices ..... 199**

6.1.1 Pathways for medical device infection ..... 200

6.1.2 Products associated with infection ..... 202

6.1.3 Biofilm formation ..... 204

**6.2 Existing antimicrobial technology ..... 205**

6.2.1 Coatings on the market and in development ..... 205

6.2.2 Shortcomings of slow-release coatings ..... 206

6.2.2.1 Drug resistance ..... 207

6.2.2.2 Systemic Toxicity and Hypersensitivity ..... 207

**6.3 AmPs as device coatings ..... 208**

6.3.1 Immobilized AmPs retain activity ..... 208

6.3.1.1 Methods to create permanently tethered AmPs on resin beads ..... 209

6.3.1.2 Antimicrobial activity of AmPs synthesized on resin beads ..... 210

6.3.2 Chemistries for attaching AmPs to device surfaces ..... 212

6.3.2.1 Coupling AmPs via sulfo-GMBS ..... 212

6.3.2.2 Antimicrobial activity of AmPs attached via sulfo-GMBS ..... 214

6.3.3 Reattaching AmPs to membranes ..... 214

6.3.3.1 Methods to attach AmPs to a cellulose membrane ..... 215

6.3.3.2 Antimicrobial activity of AmP functionalized membranes ..... 215

6.3.4 Killing through multiple bacterial challenges ..... 216

6.3.5 Confirmation that activity results from non-leached material ..... 217

**6.4 Improved device coatings** ..... **218**

6.4.1 Effect of orientation on specific activity ..... 219

6.4.1.1 Methods to generated AmP surfaces without orientation ..... 220

6.4.1.2 Methods to generate AmP surfaces without orientation ..... 221

6.4.1.3 Activity of oriented AmP at lower concentrations ..... 222

6.4.2 Hydrogel formula ..... 222

6.4.2.1 Production of hydrogels ..... 223

6.4.2.2 Activity of hydrogels ..... 224

6.4.3 Hemolytic activity ..... 224

**6.5 Commercial potential** ..... **225**

***7. Conclusions and recommendations*** ..... ***227***

**7.1 Summary and conclusions** ..... **227**

**7.2 Recommendations and future studies** ..... **230**

7.2.1 Iterative optimization of AmPs ..... 230

7.2.2 In vivo activity and toxicity ..... 233

***8. Appendix*** ..... ***234***

**8.1 Sequences for *in vitro* translation studies** ..... **234**

***9. References*** ..... ***248***

# List of Figures

2.1: Structures of representative AmPs of different structural classes. ..... 25

2.2: Alpha-helical AmP structure and mechanism of action. ..... 27

3.1: Semi-continuous in vitro translation. ..... 47

3.2: Proof-of-concept scheme to create DNA templates for in vitro translation of AmPs. ..... 50

3.3: Combinatorial DNA design method mimicking pattern-based AmP design approach. ..... 55

3.4: Combinatorial DNA synthesis method mimicking pattern-based AmP design approach. ..... 58

3.5: Example oligonuleotides designed and synthesized using combinatorial approach. 59

3.6: Alternative combinatorial DNA synthesis method requiring a shorter conserved region. ..... 61

3.7: Effect of DNA concentration on GFP yield ..... 66

3.8: Construction strategy for multimer AmPs. ..... 68

3.9: DNA gel showing multimer AmPs. ..... 69

3.10: Diagram and gel showing DNA coding for truncated versions of GFP. ..... 74

3.11: Procedure and primers to create a truncated GFP fusion partner. ..... 76

3.12: Dependence of molar yield on translated product length for truncated GFP. ..... 78

3.13: Molar yield of truncated GFP in alternate in vitro translation mixtures ..... 80

3.14: Dependence of molar yield on translated product length for truncated Luciferase. 83

3.15: Dependence of molar yield on translated product length for truncated GFP with AmP attached. ..... 85

3.16: Protein gel showing AmP digested from multiple fusion partners. ..... 87

3.17: Dependence of molar yield on translated product length for GFP-AmP fusion with stop codon between GFP and AmP. ..... 91

3.18: Degradation profile for stable and unstable fusion partners with and without protease inhibitors. ..... 94

3.19: Solubility of truncated GFP-AmP fusions. ..... 98

3.20: Effect of DNA concentration on molar yield of Thioredoxin. ..... 101

3.21: Molar yield of thioredoxin versus DNA concentration for both a Pet32a and a Roche upstream region. ..... 105

3.22: Recovery of protein following various protein precipitation procedures for purification. ..... 108

3.23: In vitro translation yields using non-toxic reaction buffers. ..... 110

3.24: Effect of imidazole on MIC determination for cecropin. ..... 113

3.25: Synthesis and purification of cecropin (CP) with resulting activity. ..... 115

4.1: Highly conserved motif from cecropins. ..... 120

4.2: Representation of a grammar based search space. ..... 122

4.3: Sample results of a TEIRESIAS search of a sample text. ..... 125

4.4: A generalized flowsheet for pattern discovery using TEIRESIAS ..... 126

4.5: Schematic of the discovery and filtering process for antimicrobial motifs. ..... 129

4.6: In silico evolution procedure for designing AmPs. ..... 133

4.7: Example designed sequence covered by antimicrobial motifs. ..... 137

4.8: A schematic of the in silico peptide design strategy..... 139

4.9: Schematic for Fmoc synthesis of peptides..... 145

4.10: Fluorescence versus cell density for luciferase containing E. coli with various antibiotics..... 152

4.11: Schematic of a microdilution assay..... 156

4.12: Schematic of MIC results using a microdilution assay..... 157

4.13: Dose response curves for peptides against B. cereus and E. coli. ..... 169

4.14: Plating for viability following incubation with a lead design..... 170

5.1: Helical projection of lead peptide used for optimization..... 178

5.2: Distribution of antimicrobial activity for natural AmPs..... 184

5.3: Hemolytic activity versus antimicrobial activity for natural AmPs. ..... 187

5.4: Therapeutic index distribution for weak and strong natural AmPs. ..... 189

5.5: Effect of amidation on the MIC against E. coli. ..... 193

5.6: Effect of amidation on the MIC against S. aureus..... 194

5.7: Effect of amidation on the MIC against S. epidermidis. ..... 195

6.1: Paths for infection of medical devices..... 201

6.2: Diagram of polymer resin beads with immobilized AmPs..... 209

6.3: Reduction in bacterial viability due to incubation with AmP covered resin beads. 211

7.1: Proposed method for discovery and evaluation of antimicrobial peptides..... 232

# List of Tables

2.1: Antimicrobial peptides currently in development. ..... 38

2.2: Peptide therapeutics approved or in clinical trials. ..... 40

3.1: Effect of template, buffer, and agitation on GFP yield ..... 66

3.2: Summary of improvement to a high-throughput in vitro translation and assay scheme for AmPs. ..... 117

4.1: Molecular weight of selected peptides ..... 147

4.2: Representative MICs for natural AmPs. ..... 159

4.3: Minimum inhibitory concentrations for natural peptides. ..... 162

4.4: Minimum inhibitory concentrations for designed peptides. ..... 164

4.5: Minimum inhibitory concentrations for shuffled peptides. ..... 166

4.6: Summary of activity of designed, shuffled, and natural AmPs. ..... 167

4.7: Minimal inhibitory concentration of designed and shuffled sequences against S. aureus and B. anthracis. ..... 172

5.1: Minimum inhibitory concentrations for variants of a lead peptide. ..... 179

5.2: Effect of C-terminal amidation on antimicrobial activity for representative peptides. ..... 191

5.3: Improvement in antimicrobial activity and therapeutic index from amidation. ..... 197

6.1: Market sizes and infection costs for medical devices. ..... 203

# Chapter 1

# Introduction

## 1.1 Motivation

The discovery of penicillin in 1928 marked the beginning of the antibiotic era, in which an array of life-saving drugs have been developed. Nevertheless, no new classes of antibiotics were produced between 1962 and 2000, with only one created since (Coates, Hu et al. 2002). In fact, many researchers suggest that the number of intracellular targets for antibiotic discovery are being exhausted (Schmid 2006). Improved methods are needed for the design or discovery of antibiotics with novel mechanisms of action.

The emergence of drug-resistant bacteria has heightened the urgency of antibiotic discovery. The incidence of methicillin-resistant *Staphylococcus aureus* (MRSA) infections, frequently responsible for hospital infections, increased from 32% to 45% over just 3 years starting in 1994 (Lowy 2003). These infections raise hospital costs and increase patient mortality. For example, in a study of orthopedic surgery patients, hospital time increased from 11 days to 88 days if a patient carried MRSA (Tai, Nirvani et al. 2004). Additionally, MRSA is now spreading in community settings (Hisata, Kuwahara-Arai et al. 2005). The threat from hospital infection increases as strains emerge which are resistant to the antibiotics that are our last line of defense.

With the waning efficacy of existing antibiotics, there is a critical need for new drugs that kill bacteria through alternate mechanisms. Antimicrobial peptides (AmPs) are a particularly attractive class of molecules because they work through a unique membrane-targeting approach. Further, because this killing mechanism non-specifically targets bacterial membranes, drug resistance is unlikely (Boman 2003). For this reason, if AmPs with suitable clinical properties were developed, they would represent a more sustainable approach to battling infection, with a longer useful lifetime than traditional antibiotics. Significant challenges remain in the commercialization of AmPs, with improvements required in antimicrobial activity, safety, and stability. Novel design approaches, coupled with innovations in delivery or formulation, will be critical for clinical success.

Antimicrobial peptide discovery has largely been carried out by screening organisms for new sequences, or optimizing existing natural peptides (Hilpert, Volkmer-Engert et al. 2005; Mygind, Fischer et al. 2005). These approaches only search a narrow portion of sequence space that is tightly constrained around the set of natural AmPs. They ignore the overwhelming majority of the space comprised of active antimicrobial sequences, which may contain AmPs that have superior properties. We sought to create a method to efficiently search this broader space. In order to accomplish this, we created a design tool that would discover semi-conserved motifs across natural AmPs. We then used these motifs to design novel peptides with limited homology to known sequences. We sought to optimize leads generated through this search and develop applications.

## 1.2 Objectives and approach

At the outset of this work, two objectives were proposed:

1. **Create an in vitro translation production scheme for AmPs.** Due to the high cost of chemical synthesis of peptides at the start of this work, we sought to create an automatable, high-throughput production and assay system for AmPs. Our goal was to develop a robust system that would produce sufficient material to test antimicrobial activity against 2 bacteria at up to 16 ug/ml. In order to accomplish this, a platform using *in vitro* translation was created. Through a series of improvements including creating a combinatorial DNA design approach, optimizing co-translated fusion partners, and improving solubility, the system became cost-effective.

2. **Develop a pattern-based design approach for AmPs.** In order to explore a region of sequence space that is diverse relative to the set of natural AmPs, we desired a design method that used semi-conserved rules derived from natural sequences. We sought to evaluate these designs against a series of clinically-relevant bacteria. To accomplish this, a pattern discovery tool was applied to the set of natural AmPs. The patterns discovered were used to design new sequences through an exhaustive combinatorial method. Predicted sequences were validated against representative gram-negative and gram-positive bacteria before testing against *Staphylococcus aureus* and *Bacillus anthracis*.

Leads were optimized to demonstrate the utility of the pattern-based approach as the first stage of a two-stage design tool.

Since this work commenced, two additional objectives have been developed:

3. **Enable improved design through database characterization.** For the designs carried out in Objective 2, we used a database of natural AmPs as a training set. This database had no information about the activity spectrum or toxicity of these AmPs, limiting the activity and selectivity of our designs. To provide more informative data, we clustered the set of linear natural AmPs and selected approximately 100 for characterization. Their antimicrobial activities against 3 bacteria and hemolytic activity were evaluated. Further, we demonstrated that amidating the C-terminus of AmPs systematically improved antimicrobial activities and therapeutic indices.

4. **Develop medical device applications for AmPs.** Through interactions with members of the Boston hospital community, we found that catheters and implants are frequently associated with hospital infections. Unlike existing antimicrobials which must be released to reach their targets inside of bacteria, AmPs target the bacterial membrane. Thus, AmPs could be active when permanently tethered on a medical device surface. We demonstrated the broad spectrum activity of AmPs permanently attached to medical devices surfaces through a flexible chemistry. Additionally, the surface was shown to

kill through repeated bacterial challenges and was not significantly hemolytic. The specific activity of the surface was enhanced by controlling the orientation of peptides during immobilization. A team was assembled and a company formed to exploit the opportunities created by this research.

## 1.3 Thesis organization

An overview of antimicrobial peptides, along with their role in innate immunity, is given in Chapter 2. Additionally, existing methods for discovering and designing AmPs are presented, along with their clinical application. Chapter 3 describes a high-throughput method for synthesizing and assaying AmPs. The development and optimization of a platform based on *in vitro* translation is given. A pattern-based approach to design AmPs with significant diversity relative to natural AmPs is detailed in Chapter 4. Further, Chapter 5 demonstrates that this design tool may serve as the first stage of a two-stage design tool, in which heuristic methods are used to optimize lead sequences in the second stage. Also, the groundwork is laid for improvement of the design tools by characterizing the antimicrobial activity and toxicity of natural AmPs, which could be used in an enhanced scoring system during design. Chapter 6 describes applications of AmPs in medical device coatings, with preliminary formulations given. Finally, Chapter 7 suggests future work for improving AmP activity through an iterative design approach.

# Chapter 2

# Antimicrobial Peptide Overview

Although the first antimicrobial peptide, nicin, was discovered in 1928, the majority of our knowledge about the diversity and mechanisms of action of AmPs has emerged in the last 30 years (McPhee, Scott et al. 2005). Antimicrobial peptides represent an exciting class of antibiotic candidates, particularly because their method of attack is unlikely to induce drug-resistant bacteria. First, the role of AmPs in innate immunity is described. Next, the classes of AmPs and their mechanism of action are presented. Finally, an overview of existing methods for AmP discovery and challenges remaining in the development of AmPs as antibiotics are given.

## 2.1 Antimicrobial peptides in nature

AmPs are ubiquitous among multicellular eukaryotes and are found in diverse contexts including frog skin, scorpion venom, and human sweat. AmPs act as a first line of defense against bacterial invaders. In the following sections, the natural properties of AmPs are described to give background on their clinical potential. These properties include the AmP's native role in immunity, structures, mechanism of action, and susceptibility to drug-resistance.

### 2.1.1 AmPs in innate immunity

AmPs play a variety of protective roles in host defense. First, AmP may directly kill microbial invaders (Papagianni 2003). This defense mechanism is particularly important in protecting the crypts of the intestine from infection (Lehrer 2004). Recent studies show that some AmPs have activity against pathogens that are resistant to traditional antibiotics including penicillin, tetracycline, and vancomycin (Tiozzo, Rocco et al. 1998). AmPs may also be useful in treating certain cancers (Ellerby, Arap et al. 1999; Kim, Kim et al. 2003) or preventing viral infection (Zhang, Yu et al. 2002; Owen, Rudolph et al. 2004).

A common technique to study the roles of AmPs in immunity is to examine organisms deficient in AmP activity. The loss of AmP function has been reported to lead to a number of disorders in animals and humans. For instance, the inactivation of AmP protection in the altered lung ion concentrations in cystic fibrosis patients leads to increased bacterial infection (Goldman, Anderson et al. 1997). Additionally, mice deficient in cathelicidins suffer necrotic skin infections (Nizet, Ohtake et al. 2001). The lack of the human cathelicidin has also been implicated in Morbus Kostmann syndrome, which typically leads to lethal infections in childhood (Putsep, Carlsson et al. 2002). Susceptibility created by the breakdown of AmP defenses implicates their role in the protection of a variety of tissues.

Recently, the opportunity to use antimicrobial peptides to modulate the immune system has been investigated (Finlay and Hancock 2004; Brown and Hancock 2006). For

example, AmPs can lead to histamine and cytokine release (Befus, Mowat et al. 1999; Zaiou 2007). Additionally, natural peptides may play a role in lipopolysaccharide neutralization, preventing sepsis (Ciornei, Sigurdardottir et al. 2005). Neuropeptides have also been suspected to be both message transmitters for immune response and directly antimicrobial agents (Metz-Boutigue, Kieffer et al. 2003; Brogden, Guthmiller et al. 2005). As research continues, it is clear that short cationic peptides play a variety of roles in host defense, with direct killing representing just one mode of action. This has lead to the concept that AmP may be used clinically as immune adjuvants or treatments for sepsis.

### *2.1.2 Classes of AmPs*

Natural AmPs occur in a great variety of plants, animals, and bacteria. Because of the sequence diversity that exists within AmPs, they are often classified based on structure, which is often linked to evolutionary relatedness. Broadly, they can be classified as alpha-helical, beta-sheet, extended, looped, single-disulfide, multiple-disulfide, and rich in unusual amino acids, such as proline or histadine (Andreu and Rivas 1998; Jenssen, Hamill et al. 2006). Natural AmPs may also contain alterations including glycosylation, disulfide bonds, and unnatural amino acids. The spectra of antimicrobial activity overlap heavily between classes, with peptides from disparate classes playing similar roles in host defense. A number of classes have been investigated as leads for drug design, with the ease of manufacture and stability being key criteria that

differentiate the classes. Structures of representative antimicrobial peptides from a number of these classes are given in Figure 2.1.

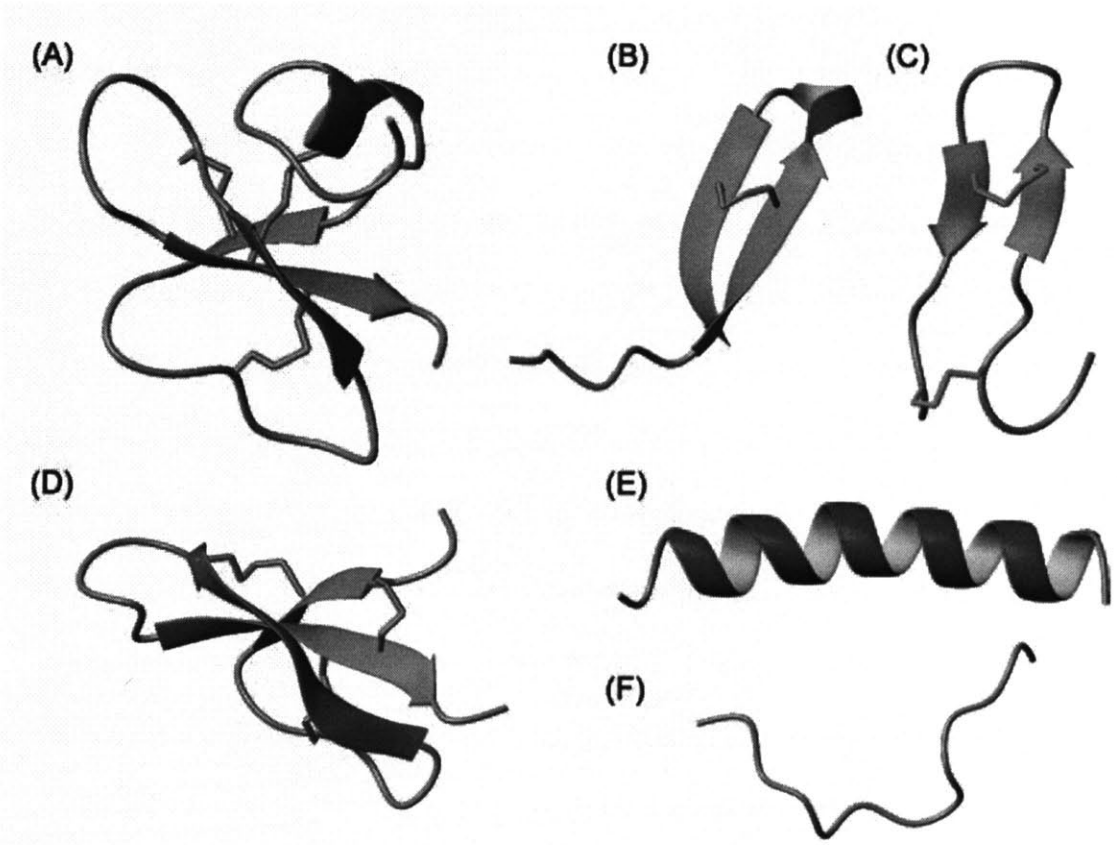



**Figure 2.1: Structures of representative AmPs of different structural classes**. A. Multiple-disulfide, human beta-defensin-2; B. Looped, thanatin; C. Beta-sheet, polyphemusin; D. Multiple-disulfide, rabbit kidney defensin-1; E. Alpha-helical, magainin-2; F. Extended, indolicidin. Figures reproduced from (Jenssen, Hamill et al. 2006).

### 2.1.3 Mechanism of action

The many disease–relevant behaviors of antimicrobial peptides are a consequence of their ability to broadly distinguish eukaryotic cells from pathogenic invaders. In general, AmPs have a net positive charge and an amphipathic 3–D structure that gives the peptides an electrostatic affinity to the outer–leaflet of the microbial membrane (Giangaspero, Sandri et al. 2001). This affinity leads to binding, disruption of the membrane, and ultimately microbial cell death (Shai 2002). Studies with model membranes, coupled with computational simulations have yielded insights on subtle differences in the way the disruption occurs (Drin, Cottin et al. 2003; Leontiadou, Mark et al. 2006). Model mechanisms include the barrel-stave model, in which multiple peptide form an aqueous pore, the toroidal pore model in which lipids may partially line the pore, and the carpet mechanism, in which a high local density of AmPs destabilizes the membrane without pore construction (Yeaman and Yount 2003). The mechanism of action of alpha-helical peptides is depicted in Figure 2.2.

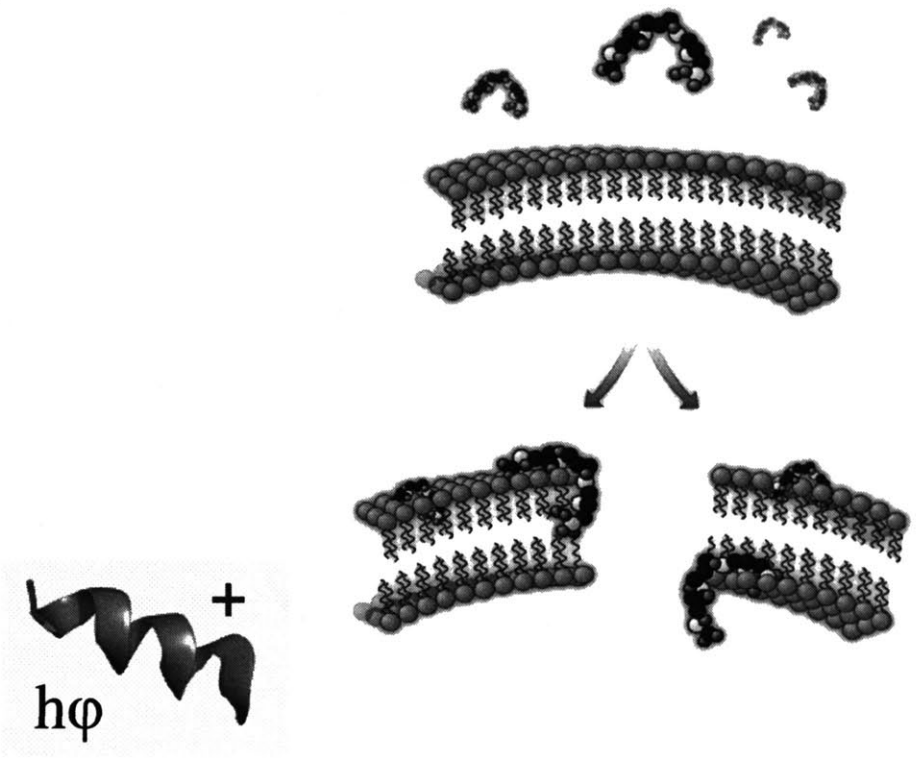



**Figure 2.2: Alpha-helical AmP structure and mechanism of action**. Illustration of an alpha-helical peptide (left) which segregates hydrophobic residues on one side of the helix and positively charged residues on the other side. Cartoon of the proposed membrane disruption mechanism of AmPs. First, the positively charged residues are attracted to the negatively charged bacterial membrane surface. Next, the hydrophobic residues allow the peptide to penetrate and destabilize the membrane. Cell death may result directly from loss of membrane integrity, or secondary internal target for the AmP may be involved in killing. Figure designed by JF Moxley.

Further analysis of the interaction of AmPs and bacteria has revealed a variety of killing mechanisms. In addition to membrane disruption, a variety of AmPs have revealed internal targets. One peptide, buforin II, kills through binding to DNA without lysing the membrane (Park, Kim et al. 1998). The lantibiotic, mersacidin, inhibits lipid II formation in biosynthesis, a similar mechanism to vancomycin (Brotz, Bierbaum et al. 1998). Additional mechanisms have been reported for individual peptides, including flocculation of intracellular contents, blockage of septum formation, inhibiting protein or DNA synthesis, or inhibiting enzyme activity (Brogden 2005). Nevertheless, the majority of killing activity is thought to arise through membrane disruption. This bactericidal attack has attractive characteristics for drug development.

Additional activities have been reported for AmPs that stem from their ability to discriminate between the cell surfaces of different classes of cells. Antifungal activity is largely thought to occur through similar mechanisms as antimicrobial activity, including membrane permeabilization and lysis (Tytler, Anantharamaiah et al. 1995; Lee, Hahm et al. 2004). Antiviral activity is less well understood than antimicrobial action, but may occur through direct envelope disruption (Jenssen, Hamill et al. 2006). For instance, indolicidin directly inactivates HIV by targeting its envelope (Robinson, McDougall et al. 1998). Among other antiviral mechanisms suggested are AmPs ability to bind heparin sulfate on mammalian cell surfaces and block viral interaction with specific cellular receptors (Tamamura, Ishihara et al. 1996; Jenssen, Andersen et al. 2004). Anti-tumour activity may result from increased surface concentration of phosphatidylserine, which AmPs may target (Utsugi, Schroit et al. 1991). The mechanism is known not to be

specific because D-amino acid peptides have equivalent efficacy (Kamysz, Okroj et al. 2003). Since membrane recognition is a critical steps in all of these mechanisms, it may be possible to adapt an AmP design tool to a variety of targets.

### *2.1.4 Antibiotic resistance*

AmPs and bacteria have been in contact for hundreds of millions of years. Detailed analysis of the modest mechanisms of resistance that have developed indicates that AmP's nonspecific attack on a range of low-specificity targets imposes a significant evolutionary cost for making changes that reduce susceptibility (Peschel and Sahl 2006). In fact, hundreds of passages at sub-inhibitory concentrations of pexiganan, an alpha-helical AmP, were required to induce resistance in a clinical resistance evolution assay (Perron, Zasloff et al. 2006). Under identical conditions, 30 passages of sub-inhibitory passages increased the MIC of an AmP only 2-4 fold, whereas the MIC of gentamicin, an aminoglycoside, increased 190 fold (Marr, Gooderham et al. 2006). If a mixture of peptides are delivered together, experts believe resistance is highly unlikely (Boman 2003). If the development of resistance to AmPs in the clinic is also slow relative to that of traditional antibiotics, a given AmP therapeutic may have a longer useful lifetime.

## 2.2 Screening for AmPs

The potential of AmPs as antibiotics has been recognized by researchers over the last 30 years. One approach to exploit this potential is through the development of

activity screens for AmPs. In the following sections, both screens of natural organisms, as well as evaluations of random amino acid sequences will be described.

### 2.2.1 Screening of natural organisms

Through the screening of natural organisms for the presence of antimicrobial peptides, a database of more than 500 natural sequences has been developed (Wang and Wang 2004). One insight that improved the effectiveness of screening was to identify organisms or tissues that are frequently exposed to pathogenic environments without becoming infected. This approach led to the identification of a series of AmPs from the skin of frogs, including the magainin class of compounds (Zasloff, Martin et al. 1988). Other groups have discovered a variety of amphibian AmPs through similar searches (Park, Jung et al. 1995; Conlon, Sonnevend et al. 2005). A related strategy has been applied to marine animals to discover other groups of antibiotics (Silphaduang and Noga 2001; Shike, Lauth et al. 2002). Additional isolation studies have revealed two classes of AmPs in humans: cathelicidin in neutrophils and defensins in Paneth cells (Zasloff 2002). Fungi are a particularly rich source for screening, and Novozymes is developing a drug candidate, Plectasin, which resulted from this screening (Mygind, Fischer et al. 2005). Nevertheless, screening organisms is a painstaking method to produce drug candidates.

### 2.2.2 Unguided approaches to produce AmPs

The enormity of sequence space for 20 amino acids long peptides (10^26 sequences) is a formidable challenge if one attempts to experimentally create AmPs from random sequences. Regardless of this, a few attempts have been made. One technique enables the inducible expression of peptides within an *E. coli* host (Walker, Roth et al. 2001). A screen is undertaken with and without an inducer of peptide expression present, and those colonies whose growth is inhibited by peptide are selected as leads. In a screen of 20,000 colonies, 21 were found that had an inhibitory effect. In this screen, however, the peptides are already within the target host cell. In order for a peptide to be effective as a therapeutic, it must be able to disrupt or transverse the cell membrane, a trait for which this screen does not select. Additionally, the peptides are over-expressed at high concentration within the cell, which may not be practically achievable through delivery.

An improved method for screening utilized a heterologous system in which AmPs are produced in one organism and excreted, with their activity measured as a zone of inhibition created for a second organism. In this work, *E. coli* containing a plasmid library is grown on solid media to produce colonies, and then transferred to a second plate containing solid media and an inducer (Raventos, Taboureau et al. 2005). The *E. coli* colonies are overlaid with a target indicator strain, and colonies are selected that produce large zones of inhibition. This approach still relies on plasmids produced by random mutagenesis, so the search space is limited to the neighborhood of the parent sequence. Further, AmPs must not be toxic to the host strain, limiting their spectrum of activity.

An alternative approach to unguided design focuses on the development of short peptides through a greedy algorithm that establishes one position at a time along the length of the peptide. In this work, all possible hexa-peptides were produced within mixtures by allowing all but the left-most amino acids to be unspecified (Houghten, Pinilla et al. 1991). To synthesize these, an equimolar mixture of all amino acids was used at each unspecified position. In a step-wise process, one additional amino acid on the left portion of the peptide was chosen by evaluating the activity of each of 20 amino acids at the position being examined, with the remaining unspecified amino acids left as mixtures. Ultimately, an "optimal" active 6-mer was produced. This unguided approach places great weight on the first few amino acids in a sequence, and would not be expected to converge to a global optimum peptide, particularly for longer sequences. In fact, no peptides identified through combinatorial approaches have reached clinical trials (Koczulla and Bals 2003). These methods have been adopted for anti-viral application and to optimize a longer known peptide of 18-amino acids (Blondelle, Takahashi et al. 1996; Boggiano, Reixach et al. 2003). These applications, however, are tightly confined to sequence space around the starting AmP.

## 2.3 Rational Design and Optimization of AmPs

As more information has been generated on natural AmPs, rational methods for AmP design have been developed. As described in the following sections, promising lead sequences may be optimized. Alternatively, computational designs may be executed using 3-dimensional templates such as an alpha-helix. Finally, methods for the

computational analysis of highly conserved clusters of sequences have been invented to improve activity.

### 2.3.1 Mutations around natural AmPs

A variety of heuristic techniques have been developed to improve the activity of natural antimicrobial sequences. One of the more common approaches is to increase the positive charge of the peptide to improve affinity for a bacterial membrane (Falla and Hancock 1997). Increasing hydrophobicity may also enhance the ability of the peptide to penetrate the lipid bilayer, and may further improve membrane affinity when positive charge is increased as well (Wu and Hancock 1999; Park, Kim et al. 2004). The addition of unnatural amino acids, fluorine atoms, and N-acylation have also been shown to strengthen activity (Wakabayashi, Matsumoto et al. 1999; Gimenez, Andreu et al. 2006). A methodical approach to improving activity was executed by systematically creating all single amino acid mutations of a known AmP sequence (Hilpert, Volkmer-Engert et al. 2005). All of these techniques produce sequences that are tightly confined to sequence space surrounding the starting AmP.

Strategies have also been created to address the developmental obstacles faced by AmPs as drug candidates. First, the cost of manufacturing peptides limits the length of peptide that may feasibly be created at a clinical scale. For this reason, researchers have sought to identify short active regions of natural AmPs by creating truncations of natural AmPs (Lupetti, Paulusma-Annema et al. 2000; Won, Jung et al. 2004; Rotem, Radzishevsky et al. 2006). This approach has been extended by substituting a lipid tail

for a portion of the antimicrobial peptide, allowing further shortening (Makovitzki, Avrahami et al. 2006). Another challenge for AmPs as systemic drugs is that high selectivity is needed if therapeutic doses are to be given. Multiple strategies, including limiting the hydrophobicity of the peptide, have been used to improve an AmP's selectivity for bacterial over mammalian cells (Yan, Li et al. 2003; Asthana, Yadav et al. 2004; Chen, Mant et al. 2005). A third challenge for development is the stability of the peptide to natural proteases. Single amino acid mutations may significantly improve peptide stability (Owens and Heutte 1997). Modifications may also include the introduction of non-natural amino acids (Ryge, Doisy et al. 2004). While these techniques may improve the clinical characteristics of an AmP marginally, these efforts do not significantly enhance the diversity of sequences available for evaluation.

### *2.3.2 Computational approaches to AmP design*

One simple computational approach to design AmPs is to select members of an antimicrobial database that have unique desirable properties and to create hybrid peptides by linking the two peptides end to end. For instance, it was discovered that while cecropin only has antimicrobial activity against gram-negative bacteria, adding a portion of melittin, which has gram-positive activity, yields sequences which are active against both (Boman, Wade et al. 1989; Andreu, Ubach et al. 1992). This strategy has been applied to a series of peptide families (Hongbiao, Baolong et al. 2005; Ferre, Badosa et al. 2006).

Additionally, computational tools have been used to identify short active regions of peptides as drug leads. Secondary structure prediction tools have been applied to known peptides, such as CAP18 to determine helical regions, which were verified to have activity (Tossi, Scocchi et al. 1994). Database scanning tools also have been used to find regions of peptides which are highly charged and likely to adopt helical structures (Zelezetsky, Pag et al. 2005). A helical propensity predictor, AGADIR, has been applied to the human AmP, LL-37, to identify shorter linear regions that display antimicrobial activity (Sigurdardottir, Andersson et al. 2006). While these tools help focus development within a set of natural peptides, the do not create significantly novel sequences.

An alternative approach to AmP design is to start with a helical projection of a peptide and place amino acids around the helix to give desired amphipathicity. In a *de novo* study, an amphipathic 12-mer composed only of arginine and valine was used as a base unit for design (Deslouches, Phadke et al. 2005). It was found that at least a 24 amino acid long helix was required for activity, with no improvement gained by using longer sequences. This approach has been extended to allow the variation of sequence, size, charge, structuring, amphipathicity, and hydrophobicity in controlled studies (Tossi, Sandri et al. 2000; Giangaspero, Sandri et al. 2001). The *de novo* design of cyclic sequences having improved stability was carried out using a similar strategy (Frecer, Ho et al. 2004). While these approaches produce interesting leads, they are limited to purely helical peptides. Many powerful, natural AmPs, however, are non-helical, and not all helical peptides are active.

More sophisticated tools have been applied to design or optimize peptides based on their structural properties or folds. Novozymes applied Quantitative Structure-

Activity Relationship (QSAR) analysis to a class of variants based on novispirin . They discovered that the volume-surface ratio and charged partial surface area were key descriptors to predict activity (Taboureau, Olsen et al. 2006). Nevertheless, this type of analysis is most appropriate within a highly conserved set of sequences. An alternative analysis was used to find conserved three dimensional structure across a diversity of sequences (Yount and Yeaman 2004). This approach was also limited to a set of disulfide containing peptides for which three dimensional structure was known. Each of the computational methods focused on a known subset of natural peptides or limited to one 3-dimensional motif.

## 2.4 AmPs as antibiotics

The broad spectrum activity of many antimicrobial peptides, coupled with their reduced susceptibility to drug resistance, make them exciting drug candidates. In the following sections, the progress to date on the development of AmPs as antibiotics is discussed. Additionally, the remaining challenges which must be overcome before AmPs enter clinical use are outlined.

### *2.4.1 AmPs in clinical development*

AmPs are under development for a variety of topical and systemic applications. The first peptide to undergo clinical trials was pexiganan, a derivative of magainin, for

the topical treatment of diabetic foot ulcers (Jacob and Zasloff 1994). This candidate performed virtually identically to the oral antibiotic standard of care and had little toxicity (Lamb and Wiseman 1998). Unfortunately, the drug was rejected due to a high baseline of improvement based on debridement of the ulcer site alone. Nevertheless, these clinical studies demonstrated the topical safety of an AmP and an efficacy that was approximately equal to oral treatment.

Since the development of pexiganan, AmPs have entered development for a variety of applications. Due to the ability of AmPs ability to bind lipopolysaccharide, sequences are being studied for septic shock treatment (Kirikae, Hirata et al. 1998; Shu, Shi et al. 2006). Additionally, AmPs have been created for the treatment of pneumonia (Steinberg, Hurst et al. 1997). AmPs are also being developed for oral indications including mucositis (Loury, Embree et al. 1999). One peptide, enfuvirtide, has been approved for treating HIV (Holguin, Faudon et al. 2007). Finally, peptides are also being studied because they have synergistic effects with traditional antibiotics (Giacometti, Cirioni et al. 2000; Hancock and Patrzykat 2002; Cirioni, Silvestri et al. 2006). This approach has the potential to extend the lifetimes of existing drugs for which drug resistance has limited their potency. A summary of peptides undergoing clinical development are shown in Table 2.1.

**Table 2 Commercial development of antimicrobial peptides of animal origin**

| Mode of use | Peptide | Company | Application | Stage |
|---|---|---|---|---|
| Topical | Pexiganan (MSI-78) | Magainin (Genaera) | Infected diabetic foot ulcers | Completed Phase III; not approved by FDA, pending additional studies |
| Topical | MBI-226 | Micrologix | Catheter infection | Phase III |
| Topical | MBI-594 | Micrologix | Acne | Phase II |
| Oral | Protegrin analogue (IB-367) | Intrabiotics | Mucositis | Phase III |
| Oral | Histatin analogue (P-113) | Demegen | Gingivitis | Phase II |
| Systemic | Heliomycin | Entomed | Antifungal | Preclinical |
| Systemic | Human lactoferricin | AM Pharma | Antibacterial | Preclinical |
| Systemic | BPI* | Xoma | Meningococcal meningitis | Phase III |

*BPI, bactericidal permeability increasing protein.

**Table 2.1: Antimicrobial peptides currently in development**. A variety of AmPs are in preclinical or clinical development for topical, oral, or systemic use. This Table is reproduced from (Zasloff 2002).

### 2.4.2 Current difficulties in development

There are a number of clinical challenges facing antimicrobial peptide development. First, all peptide therapeutics are subject to kidney clearance, which reduces half-life (Lien and Lowman 2003). Strategies to improve circulation times include the conjugation of PEG (Monkarsh, Ma et al. 1997). Peptides may also be susceptible to proteases, but this may be averted by including D-amino acids (Hamamoto, Kida et al. 2002). Peptides are also generally not stable for oral delivery, so most peptide therapeutics require IV administration. However, improvements have been made in transdermal delivery, with or without electrophoresis (Henry, McAllister et al. 1999; Kanikkannan, Singh et al. 1999). Additionally, techniques for delivery via inhalation and stabilized oral formulations are under development (Bot, Smith et al. 2001; Leone-Bay, Sato et al. 2001).

Depending on the required dose, the cost of manufacture may limit drug development (Zhang and Falla 2006). Chemical synthesis scales with the length of the peptide being produced, and depending on the dose of drug required, is not feasible for peptides longer than 35 amino acids (Latham 1999). Recombinant and transgenic approaches represent economic alternatives for large scale production. In fact, Novozymes has already developed recombinant means to produce plectasin, a promising antimicrobial peptide drug candidate (Mygind, Fischer et al. 2005). Despite these obstacles, peptide drugs have had success for a variety of applications. A sampling of peptides that are approved or in clinical trials are shown in Table 2.2.

| Drug | Company | Source | Indications | Phase[a] |
|---|---|---|---|---|
| Plenaxis™ | Praecis Pharmaceuticals (http://www.praecis.com) | Natural (GnRH) | Prostate cancer | NDA filed |
| | | | Endometriosis | II/III |
| Apan™ | Praecis Pharmaceuticals | Natural (β-amyloid) | Alzheimer's disease | I |
| Iseganan | IntraBiotics (http://www.intrabiotics.com) | Natural (protegrin) | Respiratory infection | I[a] |
| | | | Pneumonia | II/III |
| Exubra® | Inhale (http://www.nektar.com); Pfizer (http://www.pfizer.com); Aventis Pharmaceuticals (http://www.aventis.com) | Natural (insulin) | Diabetes type I and II | III |
| Oral insulin | NOBEX (http://www.nobexcorp.com); GlaxoSmithKline (http://www.gsk.com) | Natural (insulin) | Diabetes type I and II | II |
| Oratonin™ | NOBEX | Natural (calcitonin) | Osteoporosis | I |
| Oral calcitonin | Emisphere (http://www.emisphere.com); Novartis (http://www.novartis.com) | Natural (calcitonin) | Osteoporosis | II[b] |
| Oralin™ | Generex (http://www.generex.com) | Natural (insulin) | Diabetes type I and II | II/III[c] |
| HER-2/neu vaccine | Corixa (http://www.corixa.com) | Natural (HER-2/neu) | Breast and ovarian cancer | I |
| FUZEON™ | Trimeris (http://www.trimeris.com) | Natural (gp41) | AIDS | Approved |
| T-1249 | Trimeris | Natural (HIV2, SIV, gp41) | AIDS | I/II |
| Symlin® | Amylin (http://www.amylin.com) | Natural (amylin) | Diabetes type I and II | III[b] |
| Exendin-4 | Amylin | Natural (glucagon-like peptide) | Diabetes type II | II |
| Lupron depot® | Takeda Abbott Pharmaceuticals (http://www.takeda.co.jp/index-e.html) | Natural (GnRH, LH-RH) | Prostate cancer | Approved |
| | | | Endometriosis | Approved |

[a]Abbreviation: NDA, new drug application.
[b]Completed.
[c]In Canada.

**Table 2.2: Peptide therapeutics approved or in clinical trials.** Peptides are undergoing trials or approved for a variety of indications. Table reproduced from (Lien and Lowman 2003).

In addition to the challenges faced by all peptide therapeutics, specific obstacles exist for antimicrobial peptides. Antibiotics typically require large doses if used systemically in order to maintain a therapeutic concentration throughout the body. This exacerbates the cost of manufacturing obstacle for AmP development. Further, large doses imply that a highly selective molecule is used in order to avoid toxicity. For these reasons, the first successes for an AmP in the clinic may involve topical or localized applications.

## 2.5 Summary

Their unique bactericidal mechanism of action gives AmPs great promise as antibiotics. Existing approaches for screening for or optimizing natural peptides have not produced candidates which have had clinical success. The barriers to this success, however, are not overwhelming. Cost, activity, selectivity, and stability must each be improved on the margin. Design tools that enhance the diversity of sequences from which researchers may choose will be critical for finding sequences that have the appropriate clinical properties. Further, creative solutions in drug delivery or localized applications may sidestep some toxicity or stability barriers.

# Chapter 3

# *In vitro* Translation of AmPs

A number of methods exist for the production of peptides for research or commercial purposes. These approaches vary in terms of their cost, yield, and purity of peptide that can be produced. Some methods also allow the incorporation of unnatural amino acids. Additionally, certain approaches are more easily adaptable to the high-throughput production of a diversity of sequences. In this chapter, the available production methods will be reviewed. Next, the selection criteria for the high-throughput screen we are developing will be outlined. The improvements made to minimize costs and optimize yield will be discussed. Finally, work towards an automatable high-throughput platform will be described.

## 3.1 Approaches to peptide synthesis

Although a variety of approaches are available for the synthesis of peptides, there are several drawbacks for making a great diversity of sequences or sequences longer than 30 amino acids. Chemical synthesis using Fmoc chemistry gives high yields and purity (Amblard, Fehrentz et al. 2005). Nevertheless, chemical synthesis is quite costly (Latham 1999). Also, the purity of chemically synthesized peptides falls as the sequence

length increases, and labor intensive HPLC is needed to remove undesired products that are truncated or include deletions. Furthermore, the success of chemical synthesis is highly sequence dependent and the purities of some peptides are very low (Carreno, Mendez et al. 2000).

Existing *in vivo* approaches for peptide production are also not optimal for activity measurement. Phage display libraries inexpensively produce an array of peptide combinations, but they are better suited to biopanning screens than to structure or activity determination (Hoess 2001). If measurements for rationally designed peptides are desired, these screens are not appropriate. Alternatively, recombinant production can give significant yields, but require cloning and fermentation steps (Latham 1999).

More attractive for drug discovery are *in vitro* translation methods, which may quickly and cheaply produce a wide variety of peptides in parallel without cloning and fermentation (Spirin, Baranov et al. 1988; Jewett and Swartz 2004). *In vitro* translation allows the high-throughput production of many proteins or peptides, including those that would be toxic to a host cell (Sawasaki, Ogasawara et al. 2002).

## 3.2 Overview of *in vitro* translation

*In vitro* (or cell-free) translation is a process in which all of the translational machinery is purified from a cell so it can be used to create a specific protein of interest. This can be accomplished by two methods. First, one can take a crude ribosomal extract from cells, and add the appropriate amino acids and energy sources to produce a functional system. Commercial kits have been established using both *E. coli* and wheat germ. Additionally, the creators of the PURE system individually tagged and purified the components involved in translation and can custom assemble them to give an active translation mixture (Matsuura, Yanagida et al. 2007). While this system excludes proteases and other cell products, it is more expensive than extract-based systems.

*In vitro* translation systems have a number of advantages relative to recombinant production of desired proteins. First, DNA directly from a polymerase chain reaction (PCR) can be expressed without a cloning or fermentation step. Additionally, the purification of the product is simpler because cell lysis and separation of membrane from internal components is not needed. *In vitro* translation is also well adapted for the high-throughput production of custom sequences, assuming that the DNA encoding the desired product may easily be obtained (Ozawa, Headlam et al. 2004; Spirin 2004). Individual reactions may be run in separate wells of 96 well plates so the protein produced in each well is known (Busso, Kim et al. 2004; Nakano, Kawarasaki et al. 2004). Since 96 well plates may be used, an *in vitro* system may easily be integrated into an assay if a high-throughput activity screen is available (Sawasaki, Ogasawara et al. 2002). Methods have even been developed in which random mutagenesis yields a library of DNA which is

diluted to one strand per well, a PCR reaction is run *in situ*, and *in vitro* translation is carried out (Rungpragayphan, Kawarasaki et al. 2002; Rungpragayphan, Nakano et al. 2003). Finally, products that would be toxic to a microbial host may be produced in a cell-free system.

Recent advances have continued to address the shortcomings of *in vitro* translation systems. Commercial products are now available for plasmid or linear templates which may even be used for proteins with disulfide bonds (Roche Applied Sciences, 4349741001). Additionally, extracts have been developed with altered genetic codes so that unnatural amino acids can be incorporated (Josephson, Hartman et al. 2005). Additional improvements have come from optimizing the energy sources for translation or recycling energy sources. Much of the insight in this work has been to mimic the environment of the cell (Jewett and Swartz 2004). In light of these improvements, *in vitro* translation has become a viable option for the high-throughput screening of proteins or peptides.

## 3.3 Desired properties for high-throughput screen

In order to create a cost-effective and efficient high-throughput screen, a number of design criteria are desirable. First, a sufficient concentration of peptide must be synthesized to test against a range of bacteria in clinically relevant concentrations. For this application, we would like to test against 2 bacteria (one gram-positive and one gram-negative) at up to 16 ug/ml. *In vitro* translation has the potential to fulfill these

requirements, but we will have to tailor it for the production of peptides and optimize yield.

Ideally, methods for generating peptides should be suitable for automation using robotics. This requires that synthesis is carried out in parallel, preferably in a 96-well plate format. Using *in vitro* translation, individual oligos that are custom synthesized or created through error prone PCR may be added to unique wells of a 96 well plate, amplified, and used *in situ* for translation. This requires that the synthesis process is robust, without delicate parts or procedures. For this reason, a batch translation process is superior to popular semi-continuous methods in which translation is carried out within a fragile dialysis membrane so that a reservoir of energy sources and amino acids may be accessed in order to improve yield. A diagram of a semi-continuous *in vitro* translation system is given in Figure 3.1. Ultimately, a robust batch system with high yields would be desired for our platform.

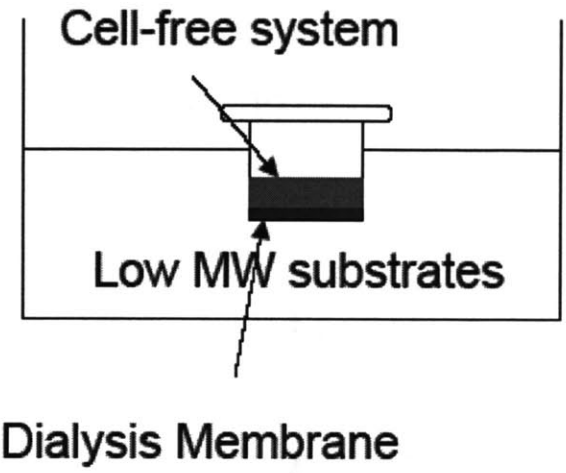



**Figure 3.1: Semi-continuous *in vitro* translation**. In semi-continuous translation, the *in vitro* extract is placed within a vial with a dialysis membrane on the bottom. This vial is floated within a larger bath containing additional amino acids and energy sources for translation. The dialysis membrane cutoff is selected such that proteins in the extract are retained, but undesired poisons that are produced during translation may exit, and fresh translation resources may enter. This system typically gives higher yields than a batch system, but is less suitable to 96 well plate automation.

## 3.4 Combinatorial DNA design and synthesis

In order to produce uniquely designed AmPs, custom oligonucleotides (oligos) are required. Despite improvements in custom oligonucleotide synthesis, ordering individual oligos for each peptides would constrain the number of sequences that may be evaluated. In the following discussions, a method to produce oligonucleotides encoding custom AmPs will be described, followed by a series of improvements that reduce the cost of this approach by mimicking our design methodology in the construction of the oligos.

### *3.4.1 Initial methods for DNA construction*

Despite significant recent progress, challenges do remain in the production of peptides by *in vitro* methods. Peptides translated on their own are typically degraded (Pilon, Yost et al. 1997). In fact, most peptides expressed naturally are produced as part of larger protein constructs, which are digested to release the free peptides after translation, or are made using non-ribosomal peptide synthesis. This knowledge has led researchers to express peptides as fusions with other proteins in order to achieve stable production (Terpe 2003). In this Thesis, any protein or protein fragment that is co-translated in frame with a peptide will be referred to as a fusion partner.

In our initial studies of antimicrobial peptide design, just a few peptides were produced, and they were only assayed at concentrations up to 10 ug/ml to look for highly

active sequences. Thus, for these studies synthesis cost and yield were not critical factors. An *in vitro* translation system was used in which the antimicrobial peptide of interest was translated in frame with Green Fluorescent Protein (GFP). The GFP fusion partner was used both to ensure stable protein accumulation and for quantification.

The DNA encoding the desired GFP-AmP constructs was created by removing the GFP with T7 upstream region from the pQB136 plasmid. A primer was used to add an enterokinase recognition site between the GFP and AmP so the AmP could be freed after translation. DNA encoding AmPs was created by designing overlapping sets of primers that code for the desired peptide. The AmP and GFP sections were then combined by a PCR reaction which takes advantage of homology designed into the two pieces. After final purification and amplification, the DNA constructs were validated by sequencing. This procedure is shown in Figure 3.2.

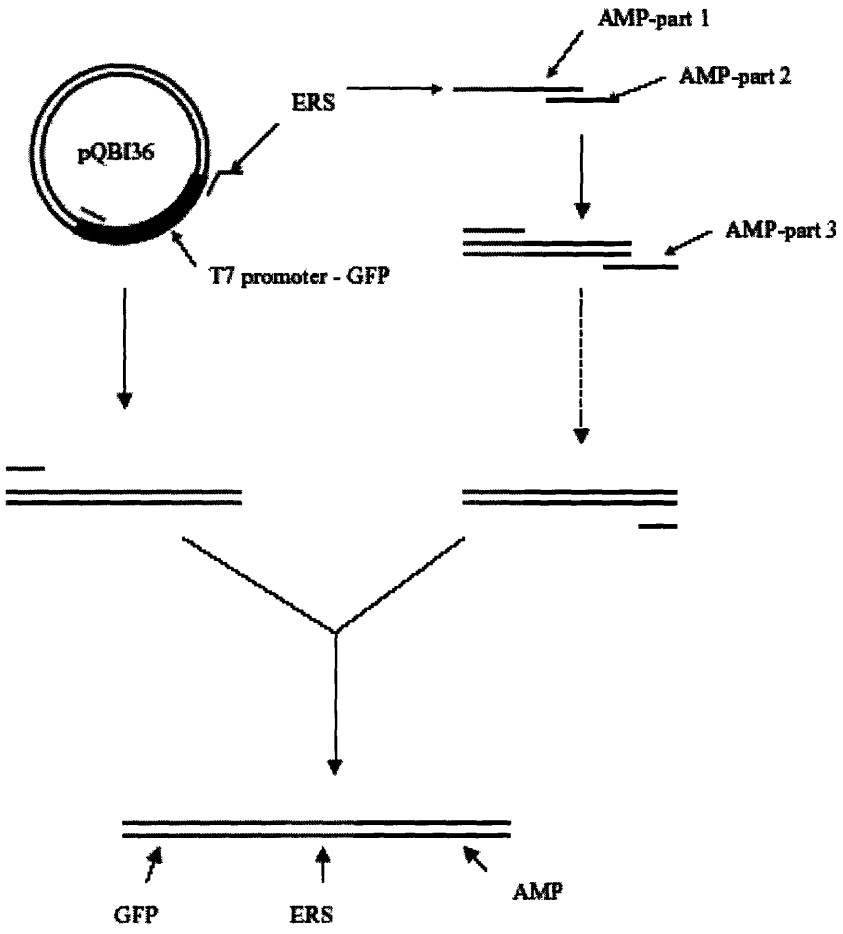



**Figure 3.2: Proof-of-concept scheme to create DNA templates for *in vitro* translation of AmPs.** A Green Fluorescent Protein (GFP) fusion partner used for quantification is taken from the pQBI36 vector with a T7 promoter region. In order to minimize cost, primers of 50 amino acids or less are linked together through a series of PCR reactions in order to create DNA coding for a unique AmP, along with an enterokinase recognition site (ERS) used to digest the AmP from the GFP following translation. The DNA segments coding for the GFP and AmP are linked with PCR using overlapping portions of the ERS to create the final template.

### 3.4.2 Drawbacks with prior methods

The above method has a number of shortcomings that limit its applicability for a high-throughput screen of antimicrobials. First, the process is costly for each AmP synthesized and would be prohibitively expensive to scale up. Multiple unique primers must be designed for each AmP produced. At the beginning of this Thesis work, each AmP would require more than two hundred dollars in primers. Additionally, overlapping regions are required for the PCR process. Primers must painstakingly be designed for each unique sequence to ensure that an appropriate melting temperature is achieved and problems will not arise from misalignment or the formation of secondary mRNA structure. It would not be practical to find individual conditions for each PCR step, which would not produce undesired side-products. Additionally, during the construction, DNA must be purified and amplified between each step.

### 3.4.3 Combinatorial DNA construction

In order to enable a more cost-effective screen, an improved method for generating oligonucleotides was desirable. We devised a more efficient strategy based on the insight that the pattern-based design method could be mimicked within DNA synthesis. In the following section, this design approach is described, followed by two sections providing experimental methods to construct these oligonucleotides.

### 3.4.3.1 Combinatorial oligonucleotide design

A critical goal for an *in vitro* translation based screen would be to reduce the cost of the oligos used. A dramatic reduction in cost could be accomplished if oligos were reused for multiple AmPs of interest. A combinatorial approach would be particularly suited to our application: designing AmPs based on patterns associated with antimicrobial activity. While a full description of the AmP design is given in Chapter 4, our approach essentially consists of linking antimicrobial patterns in unique ways to create active AmPs with novel profiles or to study the synergy of these patterns. Oligos could be designed that code for individual patterns, which could be treated as substructures, that could be combined using a variety of methods.

In order to accomplish this combinatorial production, the DNA coding for antimicrobial peptides were designed in two substructures, A and B as shown in Figure 3.3. A number of unique DNA sequences were synthesized for piece A and a different set of sequences were synthesized for piece B. The ends of the oligonucleotides for pieces A and B have overlapping and complementary codes which allow any strand A to be attached to any strand B through PCR. As mentioned above, antimicrobial patterns cover the region of overlap between the substructures. Thus, in separate wells of a 96-well plate, each A may be combined with each B to produce all possible A-B combinations. From a total of N types of A or B substructures, this technique would yield N^2 antimicrobial peptides, which would allow the exploration of a large portion of sequence space.

Our peptides were initially designed to have a final length of 22 amino acids. This is within the range of naturally occurring AmPs, which typically are 20-50 amino

acids long . Each A substructure was composed of nucleotides coding for 14 amino acids, with the first 8 amino acids being allowed to vary in each unique A. The last 6 amino acids overlap with B and were conserved between variations of A. Likewise, the first portion of each B were the complement of the overlapping region of A, and the remaining nucleotides code for 8 variable amino acids. An overlap of 6 amino acids, or 18 nucleotides, were long enough to ensure that when the PCR reaction is carried out to join substructures A and B, misalignment between A and B was unlikely. This design approach has the benefit that each unique DNA sequence could be constructed and translated in an individual well, so no separations processes would be needed.

A potentially more powerful synthesis method could be explored which would lead to a much larger set of peptides produced. Rather than creating just two substructures per peptide, 3 or 4 units could be used so that the number of combinations increases many orders of magnitude. As with the design using two substructures, patterns would cover the length of the peptide, including the regions of overlap that connect the substructures. A 3 or 4 unit methodology would allow us to accomplish our goal of creating thousands of antimicrobial sequences from a small set of peptide substructures. These peptides could also be synthesized in individual wells to eliminate the need for separations.

Additionally, this approach would allow the exploration of synergistic behavior between patterns and lead to an improved design methodology. Each of the substructures used to create the longer peptides could be synthesized and tested for activity. The substructures could be chosen to be varied in terms of their charge density, hydrophobicity, and fold. The combinatorial design approach would allow us to

investigate the cooperativity of antimicrobial sequences. This could lead to improved design if patterns with certain characteristics are discovered to complement one another. The length of the substructures could also be varied to see if there exists an optimal size of peptide for antimicrobial activity. A diagram of the combinatorial design approach used to economically create an array of AmPs of interest is shown in Figure 3.3.

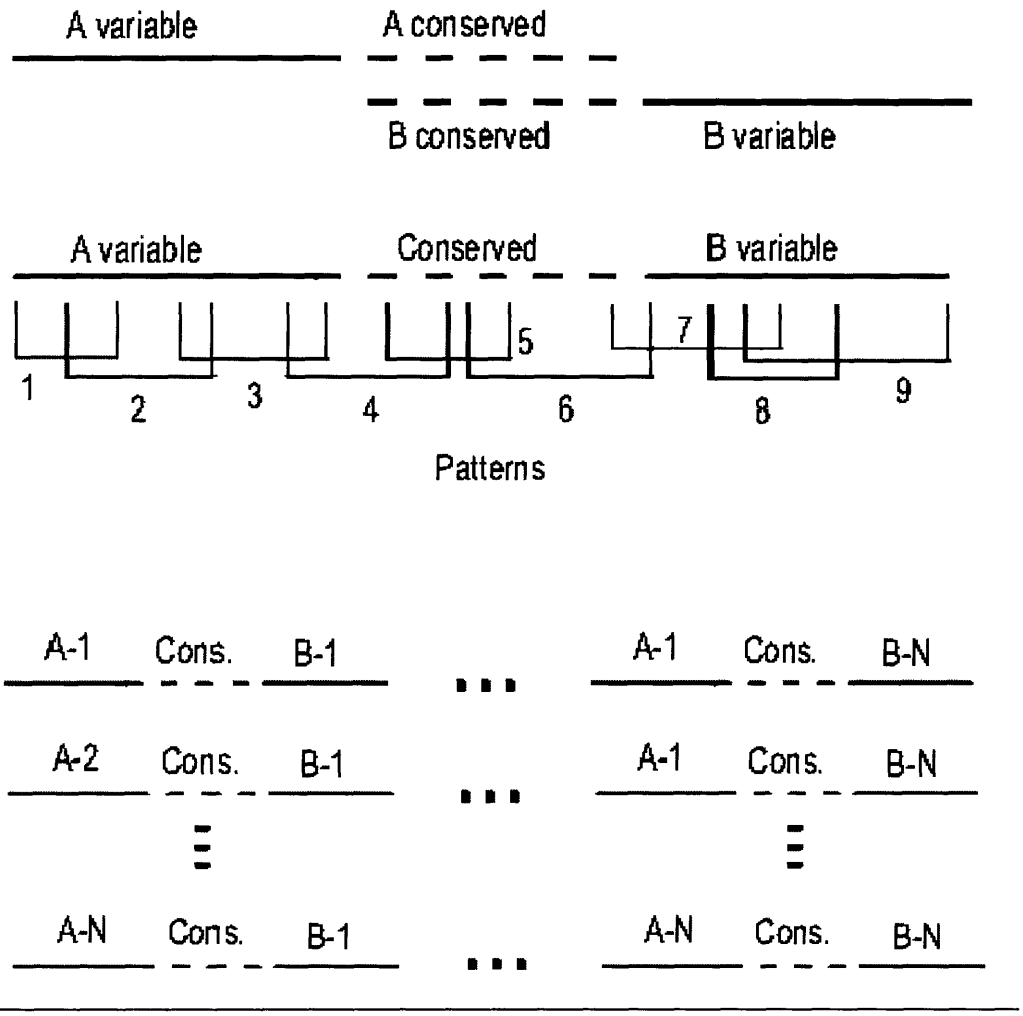



**Figure 3.3: Combinatorial DNA design method mimicking pattern-based AmP design approach**. AmPs will be designed with two distinct variable regions, A and B, both of which are entirely covered by antimicrobial patterns. A conserved region, used for linking A and B, will also be covered by patterns. N variations of section A and N variations of section B will designed, which may be linked through a conserved region to give $N^2$ unique combinations entirely covered by patterns. This approach will allow the cost-effective synthesis of a large number of sequences and analysis of the synergy of antimicrobial patterns.

### 3.4.3.2 Combinatorial oligonucleotide synthesis

The cost of synthesizing the nucleotide sequences for thousands of peptides in the range of 20-50 amino acids would be prohibitively high if each sequence were made individually. Nevertheless, most naturally occurring AmPs fall within this size range, so a method to create peptides of this length with little expense would be very valuable. In order to dramatically decrease costs for the production of novel peptides, combinatorial methods may be used to synthesize large numbers of unique peptides, which are designed to have antimicrobial activity.

As described in Section 3.4.3.1, the peptides may be designed in substructures, which could be combined to produce longer sequences covered by antimicrobial patterns. Details of the oligonucleotide synthesis procedure are shown in Figure 3.4. For peptides composed of two substructures, part A consisted of 42 nucleotides. The first 24 nucleotides coded for the 8 amino acids that vary between peptides, while 18 nucleotides overlapped with strand B and coded for 6 amino acids that were conserved in all peptides. Similarly, part B was composed of the complement to these 18 nucleotides, followed by 24 nucleotides coding for unique amino acid sequences. Antimicrobial patterns covered the entire 22 (8+6+8) amino acid sequence that were produced. Additionally, a T7 promoter, ribosome binding site, and start codon were attached on the upstream end of each of the A strands for use during transcription. A stop codon and transcription termination signal were placed downstream of each B strand. The oligonucleotides A and B were short enough that they could be synthesized inexpensively.

The substructures were combined using individual polymerase chain reactions. In separate microwells, each unique A substructure we designed was mixed with each unique B substructure and run through 30 PCR temperature cycles (95-(Tm-10)-72 ˚C, where Tm is the melting temperature of the overlapping region). An error-correcting pfu enzyme was used for the reaction to minimize the number of misalignments that occurred. The resulting DNA constructs were sequenced to ensure proper synthesis. This combinatorial approach to creating oligonucleotides will be cost effective and amenable to automation. A diagram of this process is shown in Figure 3.4.

Sample AmPs were designed using this combinatorial approach and synthesized using the above methods. A combination of 3 unique part A's were combined with 2 unique Part B's as a proof of concept. Correct synthesis was validated by DNA sequencing. The sequences used are shown in Figure 3.5.

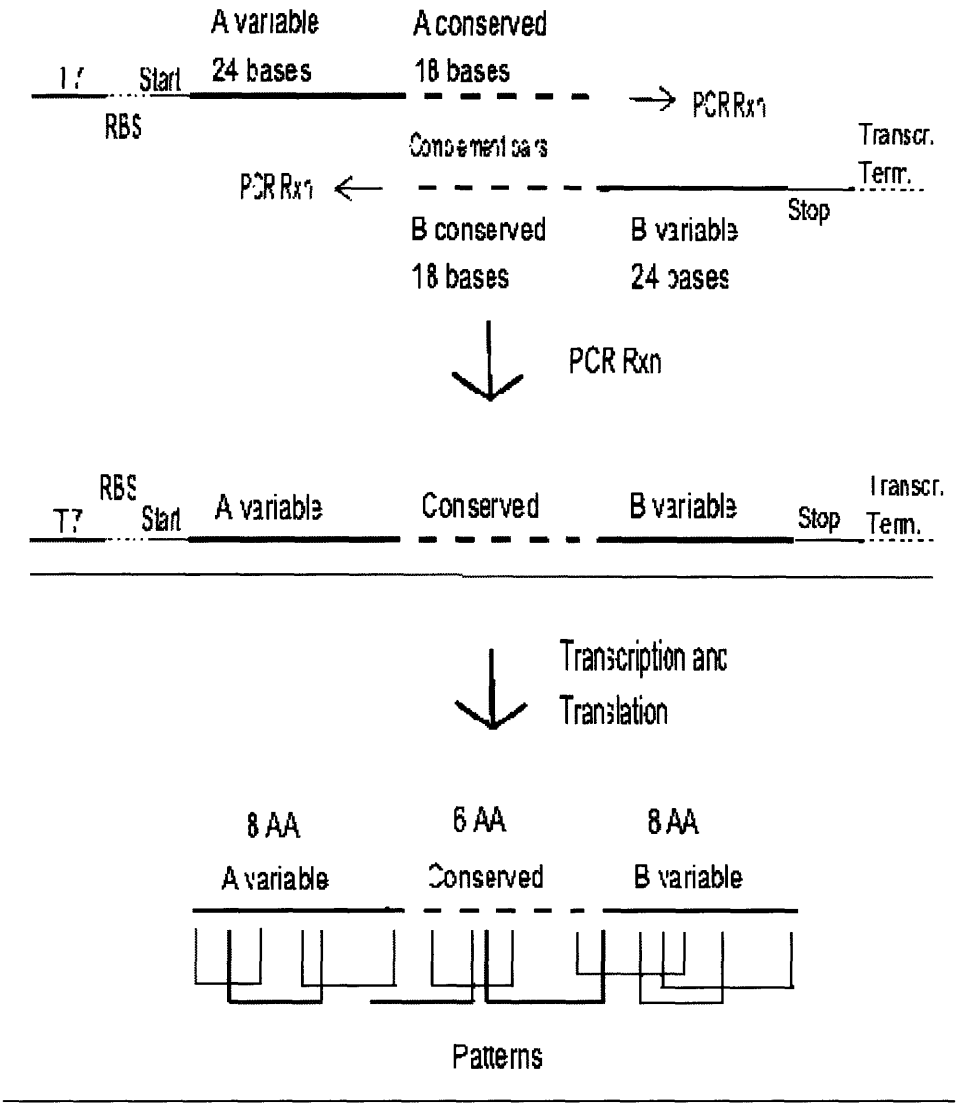



**Figure 3.4: Combinatorial DNA synthesis method mimicking pattern-based AmP design approach.** AmPs designed using the strategy shown in Figure 3.3 will be synthesized using PCR accomplished through a conserved 18 amino acid overlap region between variable chains. A T7 promoter and ribosome binding site (RBS) will be ligated upstream of all A regions, and a transcription terminator will be attached downstream of all B regions. This approach yields a cost-effective synthesis of $N^2$ unique peptides from the creation of only 2*N antimicrobial segments.

A1 – B1
Cgatcccgcgaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagaaataattttgtttaactttaagaaggaattcaggagcc
cttcaccatgacgaatttaaaagcagttattcctgtagcgggtctcgggatgcat **ATG GGT GGT CTG AAA AAA CTG GGT AAA AAA CTG GAA GGT GTT GGT AAA CGT GTT TTT AAA GCG** AGC GAA AAA GCC TTA CCG GTG GCA GTT GGG ATT AAG GCG CTG GGC *TAA TAG* CGC TCG GTT GCC GCC GGG CGT TTT TTA

A1 – B2
Cgatcccgcgaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagaaataattttgtttaactttaagaaggaattcaggagcc
cttcaccatgacgaatttaaaagcagttattcctgtagcgggtctcgggatgcat **ATG GGT GGT CTG AAA AAA CTG GGT AAA AAA CTG GAA GGT GTT GGT AAA CGT GTT TTT AAA GCG** GCG GAA CTG AAG GCA TTA CCG GTC GCG GCC *TAA TAG* CGC TCG GTT GCC GCC GGG CGT TTT TTA

A1 – B3
Cgatcccgcgaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagaaataattttgtttaactttaagaaggaattcaggagcc
cttcaccatgacgaatttaaaagcagttattcctgtagcgggtctcgggatgcat **ATG GGT GGT CTG AAA AAA CTG GGT AAA AAA CTG GAA GGT GTT GGT AAA CGT GTT TTT AAA GCG** ACC CTG GCG GAG AGC GCC GTG TTA GAA ATT *TAA TAG* CGC TCG GTT GCC GCC GGG CGT TTT TTA

A2 – B1
Cgatcccgcgaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagaaataattttgtttaactttaagaaggaattcaggagcc
cttcaccatgacgaatttaaaagcagttattcctgtagcgggtctcgggatgcat **ATG CTG GTG GAG ATT AGC GAT GAA GCG GCC GGT GTT GGT AAA CGT GTT TTT AAA GCG** AGC GAA AAA GCC TTA CCG GTG GCA GTT GGG ATT AAG GCG CTG GGC *TAA TAG* CGC TCG GTT GCC GCC GGG CGT TTT TTA

A2 – B2
Cgatcccgcgaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagaaataattttgtttaactttaagaaggaattcaggagcc
cttcaccatgacgaatttaaaagcagttattcctgtagcgggtctcgggatgcat **ATG CTG GTG GAG ATT AGC GAT GAA GCG GCC GGT GTT GGT AAA CGT GTT TTT AAA GCG** GCG GAA CTG AAG GCA TTA CCG GTC GCG GCC *TAA TAG* CGC TCG GTT GCC GCC GGG CGT TTT TTA

A2 – B3
Cgatcccgcgaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagaaataattttgtttaactttaagaaggaattcaggagcc
cttcaccatgacgaatttaaaagcagttattcctgtagcgggtctcgggatgcat **ATG CTG GTG GAG ATT AGC GAT GAA GCG GCC GGT GTT GGT AAA CGT GTT TTT AAA GCG** ACC CTG GCG GAG AGC GCC GTG TTA GAA ATT *TAA TAG* CGC TCG GTT GCC GCC GGG CGT TTT TTA

**Figure 3.5: Example oligonucleotides designed and synthesized using combinatorial approach**. A set of 2-A regions were combinatorially attached to 3 B-regions using PCR. The T7 promoter and RBS are shown in lower case, A-region in bold, B-region in unbolded, and termination sequence in underlined text. A region used for PCR is conserved in all sequences (GGTGTTGGTAAACGTGTTTTTAAAGCG). Optionally, a fusion partner could be included with the T7 upstream region.

### 3.4.3.3 Improved combinatorial design using ligation

One drawback of a PCR approach to connect the unique A and B segments is that it requires a long conserved region that limits the portion of the peptides that can be varied. Furthermore, the conserved region must be carefully designed to have an appropriate melting temperature for PCR and not to have self symmetry that could lead to secondary structure that block annealing. A superior approach would have a shorter conserved region, but would still ensure that the correct A-B construction occurred. A method to accomplish this is to directly ligate the A and B segments. For this to work, the primers for A and B were ordered along with their complementary strands. Furthermore, two of the strands were ordered with phosphorylation, which allowed the direct ligation of the double stranded A and B. A small overhang was designed on each strand to encourage the proper ligation. The ligated A-B combinations were then amplified by PCR from the ends. A T7 upstream region was added, and the final construct was again amplified by PCR. The short 4 amino acid overlap required for proper ligation allows almost complete freedom in designing the A and B sequences. A diagram of this process is shown in Figure 3.6. This approach was also executed and validated with DNA sequencing.

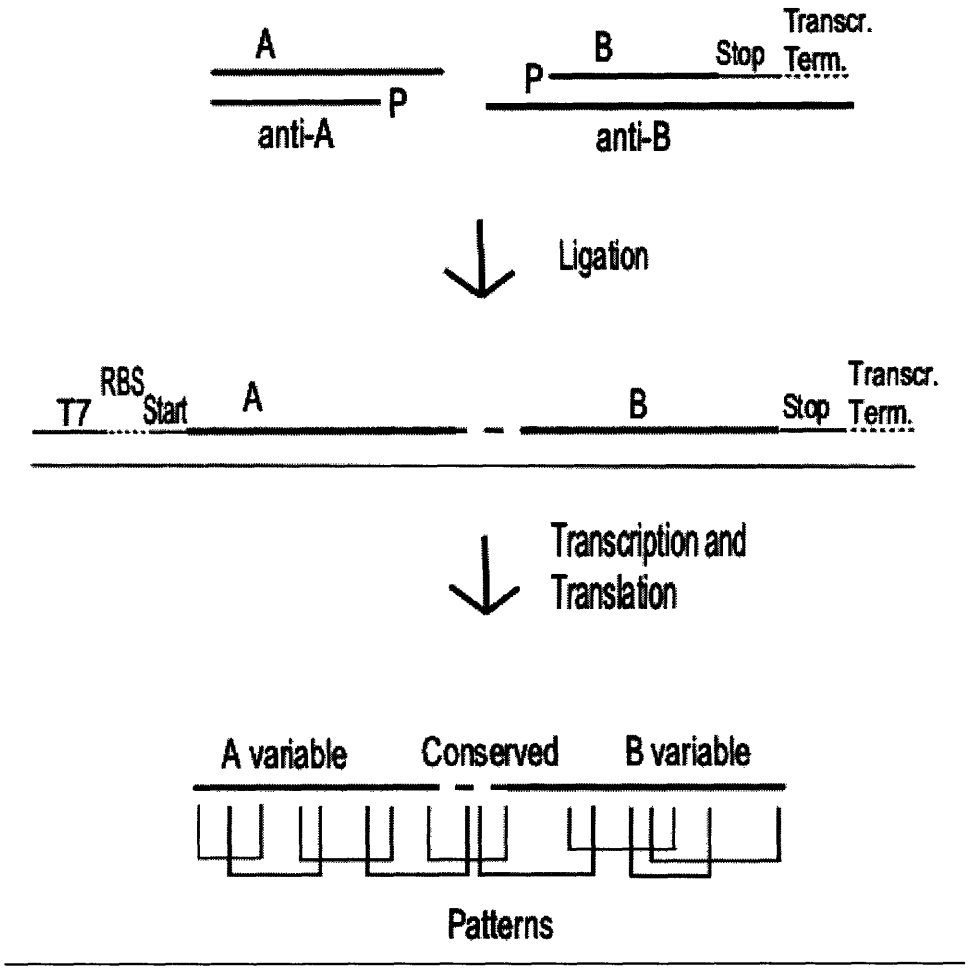



**Figure 3.6: Alternative combinatorial DNA synthesis method requiring a shorter conserved region.** A revised strategy for combining variable A and B regions for pattern based AmPs design. Rather than using PCR, which requires a longer conserved region, a 4-base pair overhang is designed so that double stranded oligos may be directly ligated after they are annealed to their complementary strands. Two of the strands were phosphorylated to enable direct ligation.

### 3.4.4 Summary

The first economic obstacle to developing a high-throughput *in vitro* translation approach to AmP discovery was the cost of custom oligonucleotides. Multiple flexible approaches were developed based on the insight that combinatorial DNA synthesis could be used to reflect the pattern-based design scheme for antimicrobial peptides.

At the time of this work, an oligo of 50 base pairs, which would be used in this work, cost ~$50, and 3 custom oligos would be required for each peptide (including overlapping segments for PCR, a transcription terminator, and an enterokinase site). For an exploration of the interaction of 10 patterns in a 10 x 10 combination, this would cost: 100 sequences x 3 primers x $50 = $15,000. For a 40 x 40 combination, the price would be $240,000. Using the combinatorial ligation approach outlined above, only 2 primers would be needed for each part A and each part B. For the 10 x 10 experiment the cost would be (2 x 10 part A's + 2 x 10 part B's) x $50 = $2,000 rather than $15,000. For the 40 x 40 experiment, which is the scale we would envision upon automating our system, the price would be $8,000 rather than $240,000. In addition to making the study affordable, the combinatorial approach allows the examination of the synergy of different classes of patterns.

## 3.5 Initial AmP translation studies

Initial *in vitro* translation optimization studies were carried out using a GFP control vector supplied with the Roche RTS100 *in vitro* translation kit (3-186-148) in order to determine parameters critical for a robust screening system. The first goal in creating a more robust system was to switch from a semi-continuous to a batch system while maintaining sufficient yield to test antimicrobial activity against 2 bacteria at up to 16 ug/ml. *In vitro* translation methods and yields in initial studies are described in the following sections.

### *3.5.1 Methods for in vitro translation in the Roche system*

The Roche RTS100 translation kit was used to translate all truncated GFP vectors because the upstream region of this vector is designed for the Roche system. The kit comes in 5 vials, all of which were reconstituted according to the kit instructions. For each 15 ul reaction, 3.6 ul of E coli Lysate, 3.0 ul of Reaction Mixture, 1.5 ul of Reaction Buffer, 1.8 ul amino acid mix, and 0.16 ul of Methionine were used. Additionally, 3.13 uCi/ml of L-[U-14C] leucine was used for protein quantification (Amersham CFB183, Buckinghamshire, England). Finally, 4 ul of template DNA at 60 ug/ml was used in all reactions to give a total DNA concentration of 16 ug/ml, unless otherwise specified. Adding additional DNA for both short and long constructs did not increase yield, indicating that the amount of DNA was never limiting. The reaction was run for 6 hrs at 30°C, as suggested in kit instructions, with 1 hr and 2 hr time points taken in some

experiments. For all translations, duplicate reactions were carried out at 15 or 45 ul, with two 5 ul samples taken from each reaction at each time point for protein quantification.

A modified version of a standard TCA precipitation was used for protein quantification (Kim, Kigawa et al. 1996). Briefly, 5 ul of sample was quenched in 100 ul of 0.1 N NaOH, with 20 ul of 0.2 mg/ml BSA as a coprecipitant. The samples were spotted on filter paper and washed 3 times with ice cold 20% TCA for 15 min. each, rinsed with EtOH, and quantified with a Liquid Scintillation Counter. In this Thesis, the terms, "accumulated product," or, "molar yield," will refer to the concentration of intact polypeptide measured with the 14C Leu quantification. When a fusion partner is linked to a target peptide in a 1:1 ratio, the molar yields of the peptide and fusion partner are identical.

### 3.5.2 Initial in vitro translation results

One primary objective for optimizing an *in vitro* system was to determine the factors limiting yield. If the limiting reagents were DNA or amino acids, they could be inexpensively supplemented to give more economical translation. First, we varied the DNA concentration used in translation to ensure it was in excess. The translation conditions described in Section 3.5.1 to translate GFP were used, with the exception that the DNA concentration was varied from 2 to 15 ug/ml. As shown in Figure 3.7, increasing DNA concentration only modestly increased yield.

Additionally, we assessed whether the elution buffer used to purify DNA from a PCR reaction impacted translation, and whether agitation affected translation. Finally,

we evaluated molar yields for linear and plasmid templates for GFP. The results of these studies are summarized in Table 3.1. None of these parameters had a significant impact on yield.

In summary, these early studies provided familiarity with the parameters important for successful translation. Further, it set a baseline yield using a control sequence. It would be desirable to increase this yield in order to drive down translation costs and increase the maximum concentration at which AmPs could be screened. Improvements to achieve these goals are discussed in the following sections.

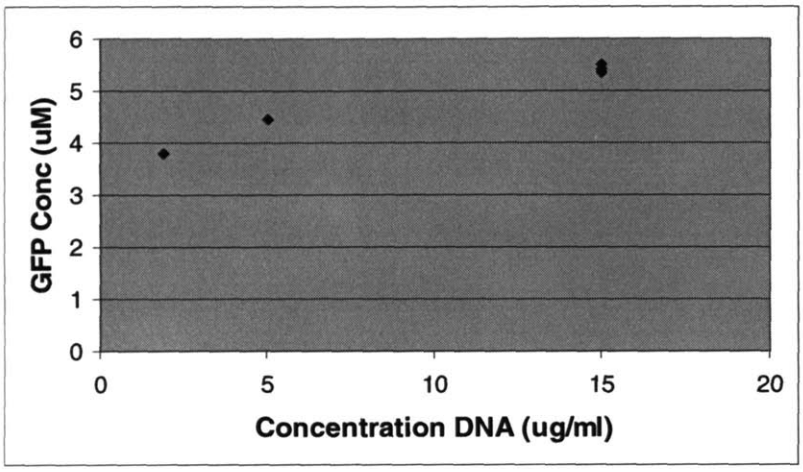



**Figure 3.7: Effect of DNA concentration on GFP yield.** Increasing DNA concentration for a GFP template above 2 ug/ml gave only a modest increase in yield when using the Roche *in vitro* translation system.

| Conditions | Molar Concentration |
|---|---|
| Plasmid GFP | 4.72 uM |
| Linear GFP in EB | 4.46 uM |
| Linear GFP in water | 4.80 uM |
| Linear GFP with Shaker | 4.98 uM |

**Table 3.1: Effect of template, buffer, and agitation on GFP yield.** *In vitro* translation was carried out using equivalent DNA concentrations of plasmid and linear DNA without a significant change in yield. Further, the yield was not affected by the presence of PCR elution buffer or agitation.

## 3.6 Multimer AmPs

After reducing primer costs by developing a combinatorial approach described in Section 3.4.3, the primary expense in the screening system is the *in vitro* translation extract. The yield generated from the translation will dictate the expense of producing the desired quantity of peptide for assay. Since we envision applying the screen to a variety of bacteria for each AmP to generate an activity profile, as well as doing *in vitro* hemolytic assays, achieving high yield is critical.

One technique used to improve peptide yield in fermentations is to translate a series of copies of the same AmP sequence in tandem (Tian, Teng et al. 2007). These strings of a repeating AmP, or multimer AmPs, can then be cleaved after translation to produce a high molar yield of peptide. In order to limit the toxicity of the AmP to the host cell, an anionic peptide may be co-expressed with the AmP in order to bind to and neutralize the AmP before the multimer is cleaved (Lee, Kim et al. 2002).

Prior researchers have sought to produce one specific AmP at high yield. In order to accomplish this, they created multimers within a plasmid. In our screen, inserting a cloning step to create and amplify a plasmid would negate the advantages of an automated *in vitro* approach. Therefore, we desired to create multimers from linear DNA. A diagram of the synthetic process is shown in Figure 3.8, in which DNA encoding the AmP has digestion sites that leave overhangs which may only reassemble with the AmP in the proper orientation relative to an upstream and downstream region. Following ligation with varying molar ratios of the AmP relative to the upstream region and termination sequence, the multimer may be amplified from the ends using PCR.

Unfortunately, the large regions of homology in the linear template created significant crossover sights, which led to the creation of multimers with fewer AmP incorporations at each amplification cycle. As shown in Figure 3.9, as more amplification cycles occur, the number of AmPs incorporated decreases, limiting the improvement in yield that could be obtained for an AmP. Alternative strategies to improve yield were explored.

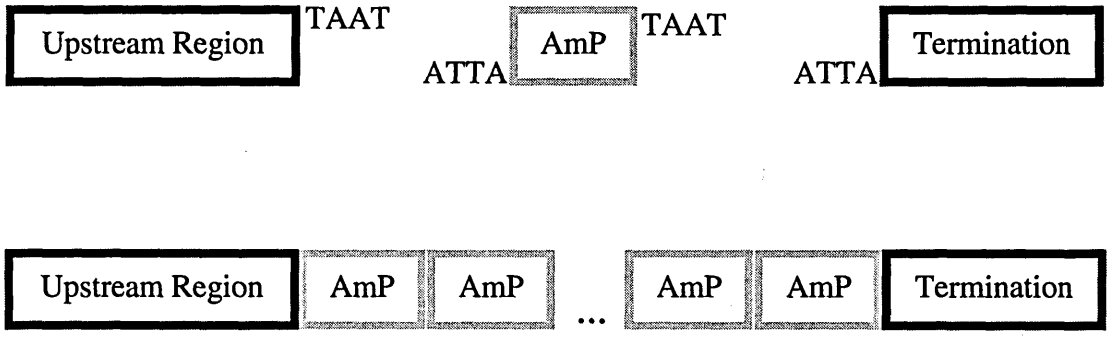



**Figure 3.8: Construction strategy for multimer AmPs.** DNA encoding an AMP is created with Bbs1 digestion sites that leave 4 base pair overhangs. The overhangs are designed such that the AmP may only be incorporated in the correct orientation. An upstream region and transcription termination sequence act as caps for this polymerization. The AmP contains protease susceptible sites at either end of the peptide so that the multimer may be digested after translation.

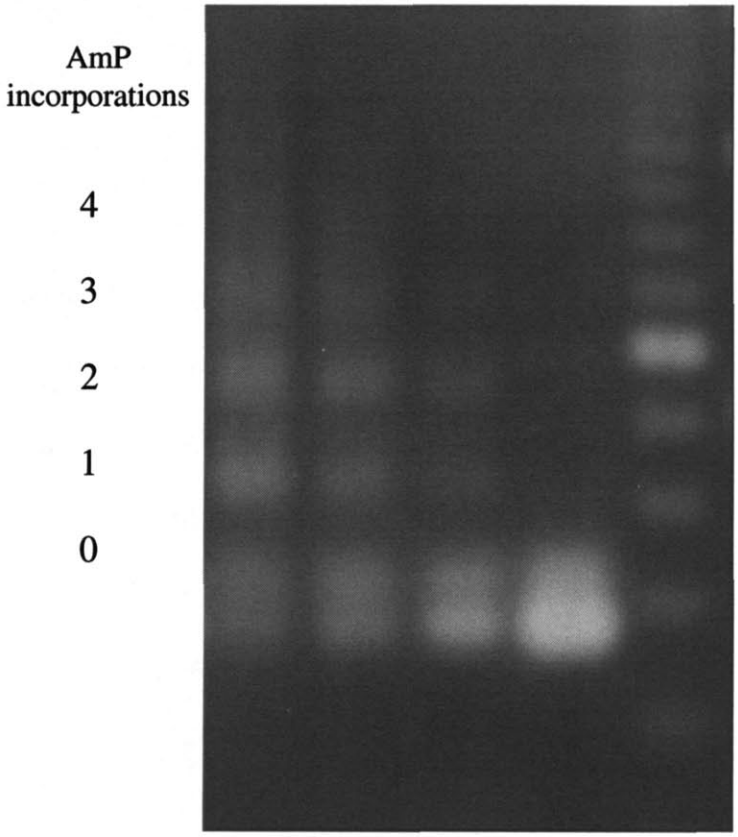



**Figure 3.9: DNA gel showing multimer AmPs.** A series of oligonucleotides were created coding for multimer AmPs. The number of amplification cycles in PCR was increased from the far left lane through the neighboring 3 lanes. With greater amplification, multimers that contained fewer AmP repetitions were created through crossovers of longer multimers. Conditions which retained large numbers of multimers as linear templates could not be found.

## 3.7 Fusion partner length optimization

As discussed in Section 3.5.2, limited yield made the high-throughput screening of designed AmPs impractical. Because AmPs were initially translated as fusions with GFP, there was a low molar AmP yield since most of the translational resources were used to make the GFP fusion partner, which was more than an order of magnitude larger. An attempt was made to increase the molar translation of AmPs per mass translated by creating multimer AmPs, but this approach was problematic for linear templates.

Another approach to increasing yield would be to make the AmP a larger fraction of the translated product by reducing the size of the fusion partner. In the following study, peptide yield was maximized for *in vitro* translation by optimizing the length of the translated product, which is composed of the fusion partner and the peptide. The DNA encoding a standard GFP fusion partner was truncated to produce fusion partners of many lengths, and the length that gave the maximum molar yield was found. This trend was generalized to 2 fusion partner sequences. Protease inhibitors were used to confirm that proteolysis was at least partially responsible for limiting accumulation of products shorter than the optimal length.

### *3.7.1 Background on fusion partners*

As described, fusion partners may be used to improve stability of translated peptides. Fusion partner may also add functionality before the peptide is cleaved. The characteristics of the fusion partner are often selected with the downstream processing of

the peptide in mind. For instance, a host of affinity tags are commonly used for one-step purification (Terpe 2003). Fusion partners may also be designed to ensure the peptide stays in solution at high expression levels (Davis, Elisee et al. 1999). Alternatively, fusions designed for the formation of insoluble inclusion bodies allow product to be isolated and resuspended (Lee, Kim et al. 2000). Additionally, fusions may neutralize the charge or activity of the designed peptide until further processing (Lee, Minn et al. 1998). Some fusion partners, such as ubiquitin, assist in proper folding of the co-expressed molecule (Pilon, Yost et al. 1997). Finally, peptides may be detected and quantified using luciferase or green fluorescent protein fusions (Su 2005).

While a great deal of attention has been paid to the functionality of fusion partners, there has been little study of how the length of fusion partners influences the peptide yields. The various functionalities mentioned previously may be achieved with fusions of different lengths. Fusion partners for purification vary widely in length from polyarginine and polyhistidine tags of 6 amino acids to maltose binding protein with 396 amino acids (Terpe 2003). Solubility partners range from the commonly used thioredoxin at 109 residues to NusA at 495 residues (Tenno, Goda et al. 2004) (Davis, Elisee et al. 1999). Insoluble partners of any length will cause the formation of inclusion bodies. Fusion partners for detection may be as small as 15 amino acids (S-tags) and as large as 250 (green fluorescent protein) or 550 amino acids (luciferase). Fusion partners that span these ranges are widely used and reported throughout the literature. Nevertheless, the effect of the length of the peptide fusion partner or total translated product on yield has not been reported.

### *3.7.2 Optimizing length of GFP fusion partner*

We desired to express S1, a 20 amino acid peptide of interest in our lab, at a high level using an *in vitro* translation system in order to characterize its activity. We found that S1 did not stably accumulate when translated on its own, or with a 10 amino acid HIS tag in the Roche Translation Kit. Therefore, we sought to translate S1 with a fusion partner to promote the stable accumulation of protein. We chose the Roche Translation Kit control vector encoding Green Fluorescent Protein (GFP) as a fusion partner.

Because the GFP is significantly longer than S1, the majority of the translation resources are used to make the GFP portion of the fusion. We sought to increase the molar yield by shortening the amino acid length of the GFP fusion partner. To this end, we created DNA encoding truncated versions of GFP, to be translated in the Roche Translation Kit. In the following sections, the methods used to create the truncations and the resulting yields are described.

#### 3.7.2.1 Template construction for truncated GFP

Linear DNA was obtained from the Roche GFP control vector (supplied in kit 3-186-148) through a PCR reaction. A primer was designed upstream of the T7 promoter site in the control vector. Downstream primers were designed to add a Bbs1 digestion site after amino acids 35, 43, 57, 70, 82, 95, 139, 203, or 250 of the GFP. This site was used to attach a transcription termination sequence for translation of the fusion partner alone, or to add DNA coding for a peptide followed by the termination sequence. The transcription termination sequence was also taken from the Roche GFP control vector.

The attachment regions added 7 amino acids to the GFP truncations, giving lengths 42, 50, 64, 77, 89, 102, 146, 210, and 257 amino acids. Figure 3.10 shows a DNA gel for a series of truncated GFP fusion partners with the transcription termination sequence attached. DNA constructs were validated by sequencing. The full DNA sequences encoding all of the truncated fusion partners created are given in the Appendix.

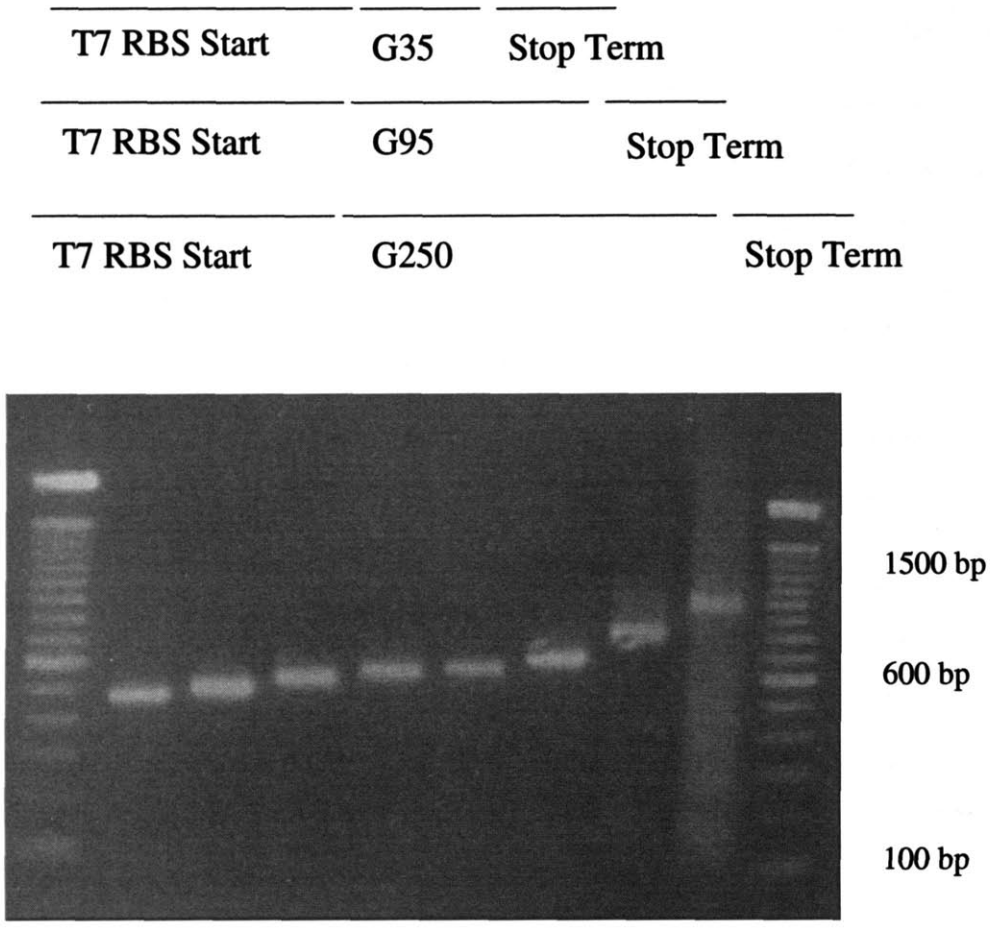



**Figure 3.10: Diagram and gel showing DNA coding for truncated versions of GFP.** A schematic of 3 truncated versions of GFP containing 35, 95, and 250 amino acids with identical upstream and termination regions is depicted. A DNA gel showing all of the truncated GFP fusion partners is also given.

For some studies, it was desired to express S1 in frame. The DNA encoding S1 was designed by reverse translating its amino acid sequence, NKVKKPLTGAHRLLFTFLFV, *in silico* using an *E. coli* codon usage table (www.entelechon.com). A Bbs1 site was added on either end to allow the DNA for S1 to be ligated in frame with DNA encoding a promoter region and a truncated GFP fusion partner on the 5' end, and a transcription termination sequence on the 3' end. The DNA sequence for S1 with the Bbs1 restriction sites is shown in Figure 1. All pieces were digested with Bbs1 (New England Biolabs, Ipswitch, MA) and ligated using T4 DNA ligase (New England Biolabs, Ipswitch, MA). Once ligated, the correctly formed construct was amplified with Forward Primer 1 from upstream of the promoter and Reverse Primer 2 from downstream of the transcription termination sequence. The addition of S1 added 29 amino acids to the GFP truncations, resulting in constructs that were 64, 72, 86, 99, 111, 124, 168, 232, and 279 amino acids. The full procedures with primers are shown in Figure 3.11. The DNA sequences with S1 fused in frame are given in the Appendix.

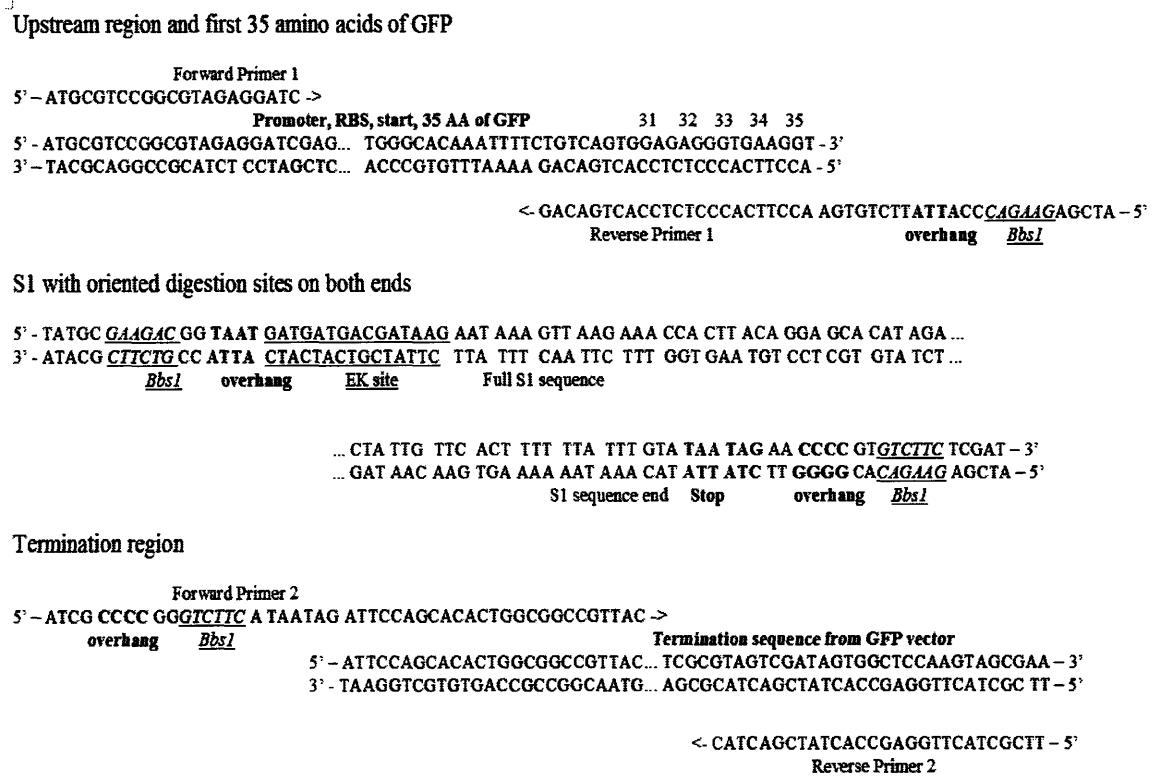



**Figure 3.11: Procedure and primers to create a truncated GFP fusion partner.**

Procedure for constructing DNA which encodes truncated versions of GFP (here 35 AA). A Bbs1 digestion site is added after 35 AA and used to attach a termination sequence, or the peptide, S1, followed by a termination sequence.

### 3.7.2.2 Effect of translated product length of yield

With a series of truncated versions of GFP created, the effect of length of the translated product on molar yield could be evaluated. A time course for translation was carried out for all truncations in the Roche translation system. The translation and quantification methods are described in Section 3.5.1. The molar yield, or accumulation of translated product, at time points of 1, 2 and 6 hours are shown in Figure 3.12.

A three-fold increase in accumulation occurs as the length of the translated GFP sequence is reduced from 257 to 102 amino acids (these lengths include 7 amino acids added from restriction sites). The time course shows that the amount of accumulated protein decreases between two and six hours after the start of translation for constructs of 77 amino acids and shorter.

Above the optimal translated product length, the molar yield of peptide falls because excessive resources are consumed to produce the fusion partners rather than the peptide of interest. If the construct is long enough to be stable, translation and accumulation will continue until the energy sources or amino acids are depleted or inhibitory agents build up in the system. For stable constructs, the total mass yield is approximately constant, but the peptide of interest comprises a smaller portion of the yield as the fusion partner length increases. The maximum molar yield of peptide occurs at the minimum translated product length for which the product is stable.

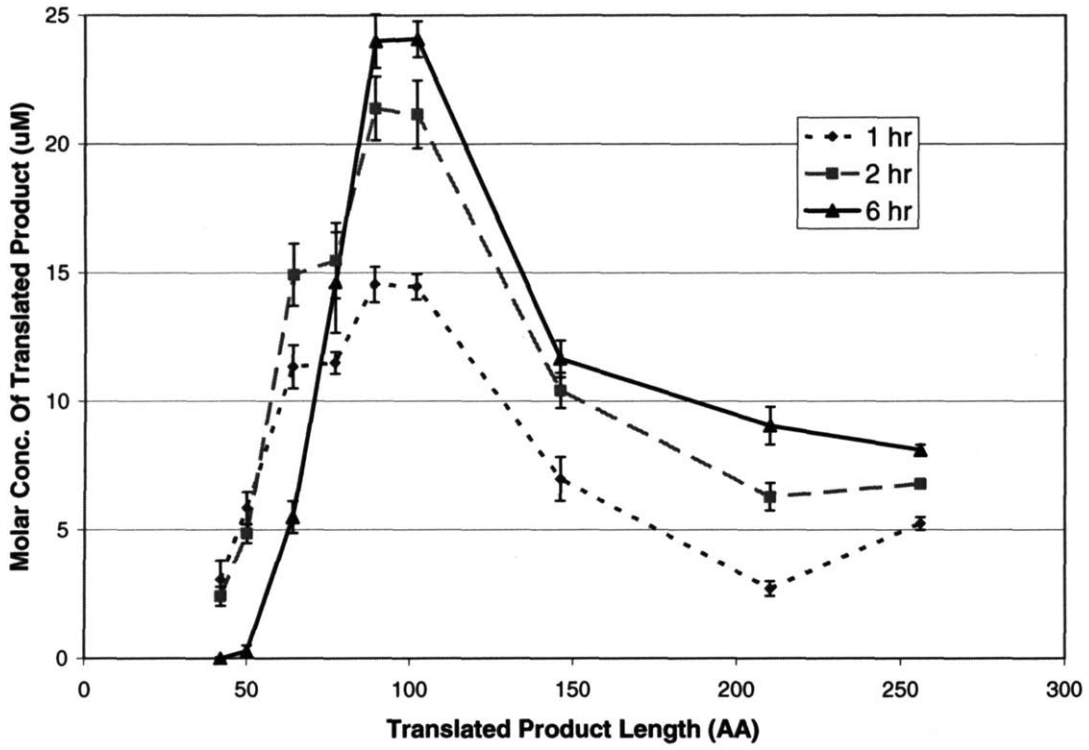



**Figure 3.12: Dependence of molar yield on translated product length for truncated GFP**. There is an optimal length for the translated product at which peptide accumulation is maximized. Truncated versions of GFP were translated in the Roche *in vitro* translation system, and protein quantification was carried out using TCA precipitation at 1, 2, and 6 hour time points. Yield peaks at roughly 100 amino acids by the 6 hr time point, when translation has been completed. Below 80 amino acids, protein accumulation falls between 2 and 6 hours, whereas at 80 amino acids or above, accumulation increases between 2 and 6 hours (n=4, error bars = SD).

### *3.7.3 Effect of translation system on optimal length*

In order to test the generality of this trend in molar yield, we studied three additional translation systems. Two alternative extract sources for *in vitro* translation, KC6 and NMR5, were used in the PANOx-SP system developed by researchers in the Swartz lab (Michel-Reydellet, Calhoun et al. 2004; Calhoun and Swartz 2005). These translation systems have the same energy source as the Roche system, but has a different extract source. KC6 and NMR5 were used to translate DNA encoding the 64, 102, and 257 amino acid versions of truncated GFP.

#### 3.7.3.1 Methods for translation in the KC6 and NMR5 extracts

The reaction mixture contains 10 mM ammonium glutamate, 20 mM magnesium glutamate, 170 mM potassium glutamate, 1.2 mM ATP, 0.86 mM each of GTP, UTP, and CTP, 34 ug/mL folinic acid, 170.6 ug/mL E. coli tRNAs, 2 mM of each of the 20 amino acids, 0.03 M phosphoenolpyruvate, 0.33 mM NAD, 0.27 mM CoA, 2.7 mM oxalic acid, 1 mM putrescine, 1.5 mM spermidine, 5.25 uM L-[U-14C] leucine, 0.1 mg/mL T7 RNA polymerase, and 24% S30 E. coli cell extract in a 57.2 mM HEPES-KOH buffer at pH 7.5 (Spirin, Baranov et al. 1988). A total DNA concentration of 16 ug/ml was used along with 1.67 uCi/ml of L-[U-14C] leucine. The reactions were run at 37°C for 4 hours as described in Swartz et al., 2004 (Swartz, Jewett et al. 2004). The resulting yields are shown in Figure 3.13.

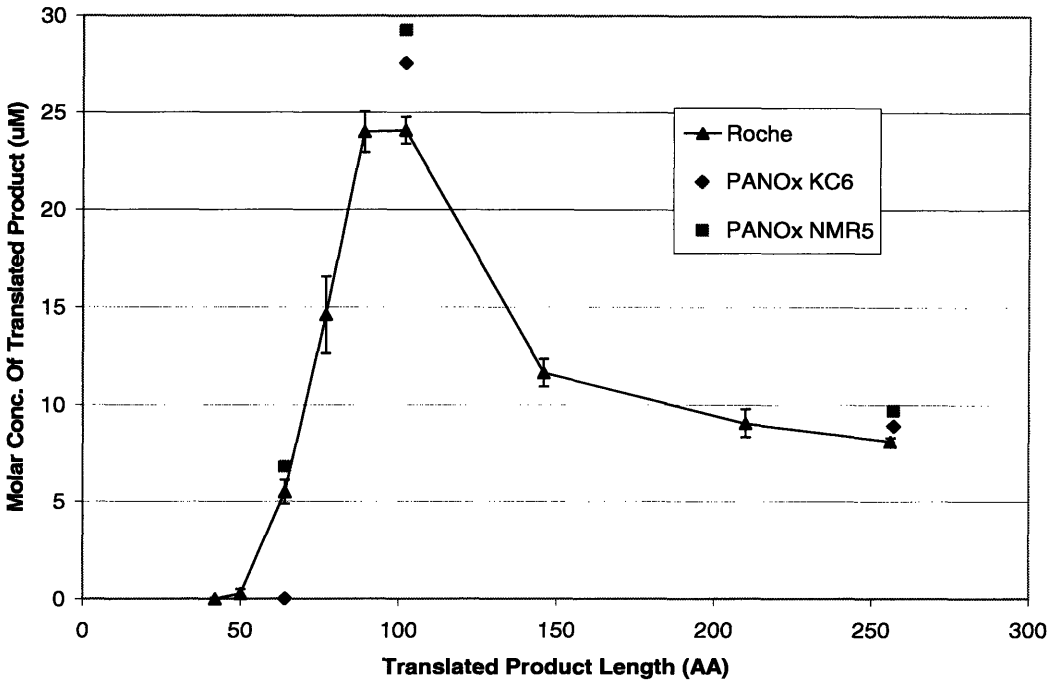



**Figure 3.13: Molar yield of truncated GFP in alternate *in vitro* translation mixtures.** Truncated versions of GFP of lengths 64, 102, and 257 amino acids were translated in the Roche, PANOx-SP KC6, and PANOx-SP NMR5 system and quantified using TCA precipitation. A similar maximum in yield occurs at the same translated product length in the PANOx-SP KC6 and NMR5 systems (n=2) as in the Roche translation system (n=4).

### 3.7.3.2 Optimal fusion partners in alternate *in vitro* systems

As seen in Figure 3.13, little accumulation occurred for the 64 amino acids construct, whereas accumulation was threefold higher for the 102 amino acid construct than for the 257 amino acid construct. The trend in these yields was similar to those obtained in the Roche system, demonstrating that the presence of an optimal translated product length occurs in multiple translations systems.

## *3.7.4 Effect of fusion partner sequence on optimal length*

The same fusion partner truncation procedure was applied to luciferase to determine whether the trend of molar yield versus translated product length was sequence-dependent. To this end, a fourth translation system, the Promega S30 for Linear Templates, was utilized because the luciferase vector was optimized for this kit.

### 3.7.4.1 Methods for creating and translating truncated luciferase

An identical procedure was used to produce truncated versions of the Promega luciferase vector pBestLuc (Supplied in L1020). The DNA sequences for the truncated versions of luciferase are included in the Supporting Information. The Promega S30 kit for Linear Extracts was used to translate the truncated Luciferase constructs because it has been designed for this kit. The reaction was composed of 4 ul S30 extract, 5.25 ul S30 Premix, and 1.5 ul amino acid mix with 16 ug/ml DNA template, and 4.17 uCi/ml of

L-[U-14C] leucine. The reactions were run at 37°C for 6 hours. The resulting yields are shown in Figure 3.14.

### 3.7.4.2 Optimal translated product length for truncated luciferase

The optimal product length does vary moderately with the identity of the fusion partner being studied. In a series of truncated luciferase templates shown in Figure 3.14, a three-fold increase in yield occurred as the translated product length was reduced to an optimal length. The maximum molar yield for truncated GFP occurs at 102 amino acids, whereas the optimum occurs at 136 amino acids for truncated Luciferase. This shift is to be expected given the hypothesis that susceptibility to proteases causes places a lower limit on the translated product length required for accumulation. Different amino acid sequences will have different propensities for protease breakdown. Additionally, the degree of tertiary folding for a given sequence length would vary with the sequence being produced. The optimal truncated GFP does not clearly correspond with a subdomain of the full GFP structure (Yang, Moss et al. 1996). This is not surprising given that the truncated protein may fold differently than when it is a part of the full GFP protein.

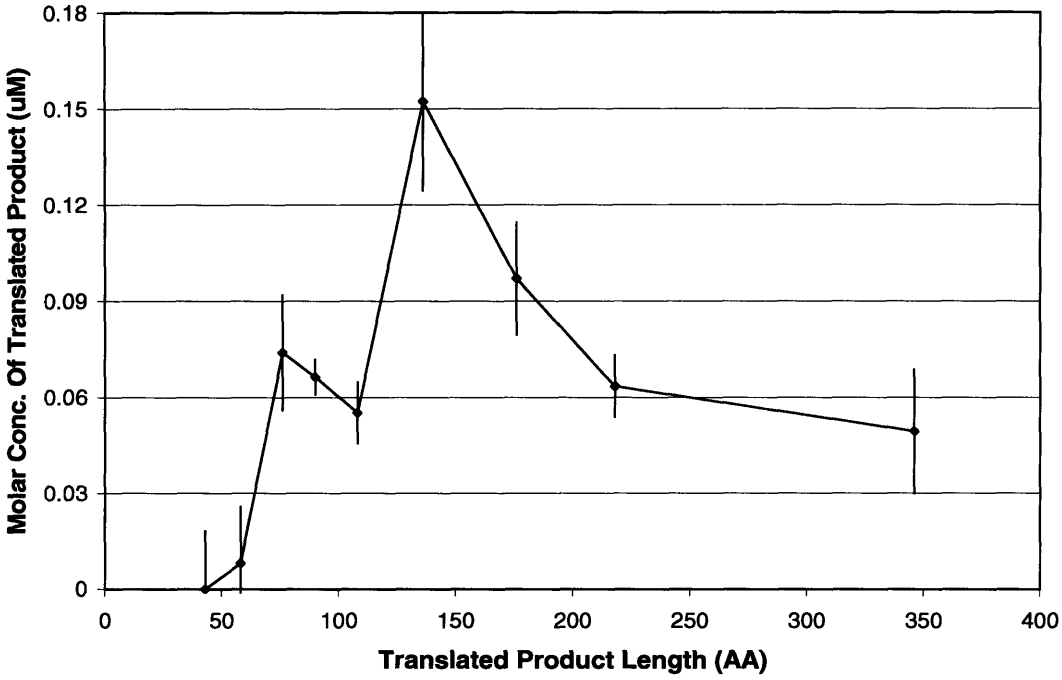



**Figure 3.14: Dependence of molar yield on translated product length for truncated Luciferase.** Truncated versions of Luciferase were translated in the Promega translation system for 6 hrs and quantified with TCA precipitation. The maximum yield for truncated sequences from Luciferase in the Promega translation system occurs when 136 amino acids are translated. Lower yields are always seen in the Promega system than in the Roche system (n=4, error bars = SD).

### *3.7.5 Translating with a desired AmP in frame*

We then sought to confirm whether the optimal translated product length changed when S1 was attached to the fusion partner. DNA encoding S1 was added in frame to the DNA encoding each of the truncated GFP fusion partners along with a small attachment region and enterokinase recognition site. The addition of S1 lengthened the translated product 29 amino acids. The effect of translated product length, composed of the fusion partner plus the peptide, on the final molar yield of product is shown in Figure 3.15.

#### 3.7.5.1 Yield with AmP translated in frame

The optimum in molar yield occurred at a similar translated product length when S1 was attached to fusion partners of different lengths. The fusion of S1 and the 70 amino acid GFP truncation gave the maximum yield. The total length of the construct was 99 amino acids, which is similar to the 102 amino acid optimal for GFP fusion partners translated without S1. This indicates that for maximum yield, a fusion partner should be selected such that the length of the peptide plus fusion partner is optimal.

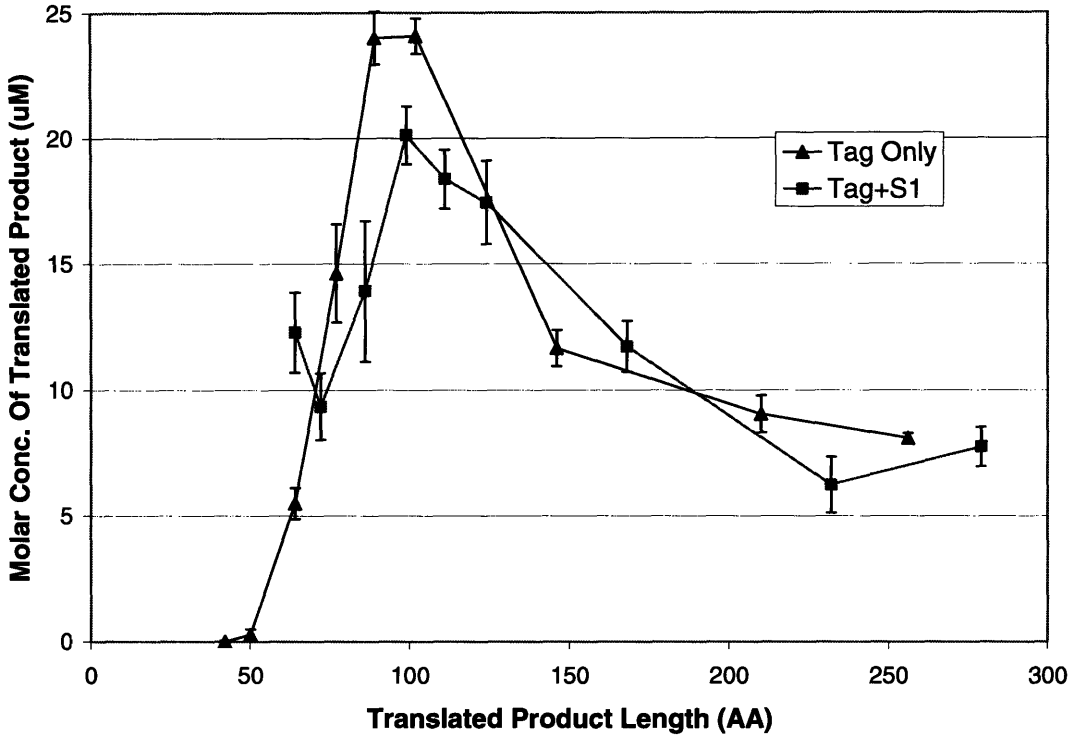



**Figure 3.15: Dependence of molar yield on translated product length for truncated GFP with AmP attached.** Translation of GFP fusion partners alone or the fusion partners plus S1 were carried out in the Roche system for 6 hrs at 30°C and quantified with TCA precipitation. The maximum in molar yield occurs at the same translated product length, which is composed of the peptide plus the fusion partner. A fusion partner of 70 amino acids plus 29 amino acids to add S1 gave maximal yield.

### 3.7.5.2 Digestion from fusion partner

After translation, the peptide, S1, must be digested from its fusion partner if its free activity is to be measured. Digestion was accomplished by adding 0.1 U/ul of enterokinase to the translation mixture and incubating at 30°C overnight without adding additional buffers (EMD Biosciences, Madison, WI). The digestions were visualized by using the Promega Transcend Detection kit (L5080) using biotinylation. The Roche Translation conditions mentioned above were used with the addition of 1.2 ul Transcend tRNA per 15 ul reaction. Samples were diluted 1:1 with Tricine Sample Buffer (Bio-Rad 161-0739) and run on 16.5% Tris/ Tricine Peptide gels (Bio-Rad 161-1179). The gel was blotted to PVDF and developed according to the Transcend Detection protocol. A sample gel with S1 digested from GFP and a truncated GFP is shown in Figure 3.16. As expected, the yield of S1 is visually greater when a truncated fusion partner is used.

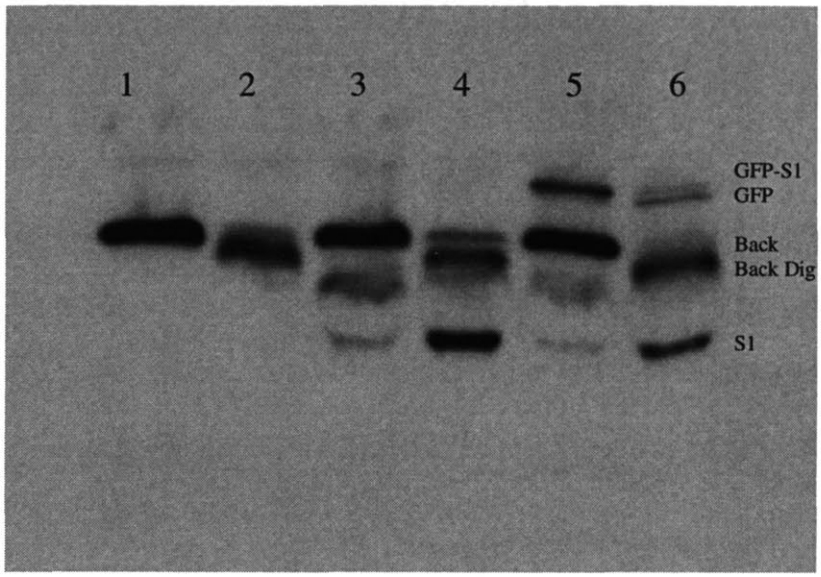



**Figure 3.16: Protein gel showing AmP digested from multiple fusion partners.** An *in vitro* translation reaction using biotinylation for detection was run for 95 and 250 amino acid fusion partners with S1 for 6 hours at 30°C. After translation, half of the samples were digested with enterokinase, separated with gel electrophoresis, transferred to PVDF and developed. Sample lanes are: 1- no DNA translated (control), background biotinylation is seen, 2 – no DNA translated, digested with EK, 3 – 95 aa fusion partner + S1 (hidden by background), 4 – 95 aa fusion partner + S1 digested, 5 – 250 aa fusion partner (GFP) + S1, 6 – 250 aa fusion partner (GFP) + S1 digested. A greater yield of S1 is seen when digested from a 95 amino acid fusion partner than from a 250 amino acid fusion partner - lane 4 versus 6.

## 3.8 Exploration of protease activity

Below the optimal length, susceptibility to proteases may reduce yields. The *in vitro* translation extracts used in this work are made from *Escherichia coli*, and thus contain dozens of proteases (Gottesman 1996). These proteases range from narrow to broad in their specificity, and may survive the extract preparation process. Genes encoding some proteases may be deleted, as was done in labs of G. Georgiou to reduce the degradation of products (Baneyx and Georgiou 1991; Jiang, Oohira et al. 2002). Cell-free systems have been designed to reduce protease activity as described in Jiang et al, 2002 (Jiang, Oohira et al. 2002). However, it may not be possible to remove many of these proteases by deleting genes in *Escherichia coli* because some deletion strains are not viable or robust (Gottesman 1996).

The enhanced degradation of fusions below 80 amino acids may be due to the fact that shorter constructs are not sufficiently long to fold in a manner that protects them from protease degradation. As shown in Figure 2, translated product appears in the first hour of the reaction but is degraded away over time for truncations shorter than 77 amino acids, as confirmed by protein precipitation using TCA. The degradation is more extensive for the shorter constructs, with complete degradation of the 42 and 50 amino acid constructs by 6 hours.

The activity of the proteases in the *in vitro* translation mixture falls over time. The rate of disappearance of the 50 amino acid construct decreases after 2 hours in Figure 6. Additionally, the S1 released from the 95 or 250 amino acid GFP constructs was stable throughout the enterokinase digestion overnight, despite the fact that S1 is

degraded in fresh *in vitro* translation mixture. The digestion was carried out in the *in vitro* translation mixture where proteases were still present, but the proteases did not display activity in the spent mixture. The decrease in protease activity may be due to wear on the proteases or the depletion of energy sources needed for activity. The loss of protease activity over time is convenient for the translation of short peptides on fusion partners since the translation can be run to completion and then the peptides may be cleaved from the fusion partner after protease activity has decreased.

### 3.8.1 Translated product length defines optimal yield

We noted that when the peptide, S1, is attached to a fusion partner, the optimal yield still occurs at the same translated product length, which is composed of the fusion partner plus the peptide. To test whether the length of the mRNA transcript or the amino acid sequence of the fusion is critical in defining the optimal fusion length, a stop codon was placed between the fusion partner and S1. Thus, the amino acid length of this construct was the same as the length of the fusion partner alone, whereas the mRNA length was equal to the length of the mRNA for the fusion partner with the peptide. If proteases were responsible for defining the optimal length, one would expect the optimum to be dependent on the length of the amino acid sequence translated, not the length of the mRNA.

### 3.8.1.1 Methods for varying amino acid length independently of mRNA length

In order to vary the translated amino acid length independently of the mRNA length, two stop codons were inserted after the DNA for the truncated GFP fusion partner but before the peptide S1. A revised Reverse Primer 1 with the 5'-ATT **CTG TGA**-3' segment replaced by 5'-ATT **CTA TTA**-3' was used in the first step of the procedure in Figure 1 to create fusions with two stop codons before S1. These constructs just had the truncated GFP translated, so they were 35, 43, 57, 70, 82, 95, 139, 203, and 250 amino acids. The DNA sequences for constructs with a stop codon before S1 are included in Supporting Information. Finally, constructs were created with S1 attached directly to the GFP upstream region or to an N-terminal HIS Tag (Roche 3-186-237). These constructs were translated and quantified using methods given in Sections 3.5.1. The resulting yields are shown in Figure 3.17.

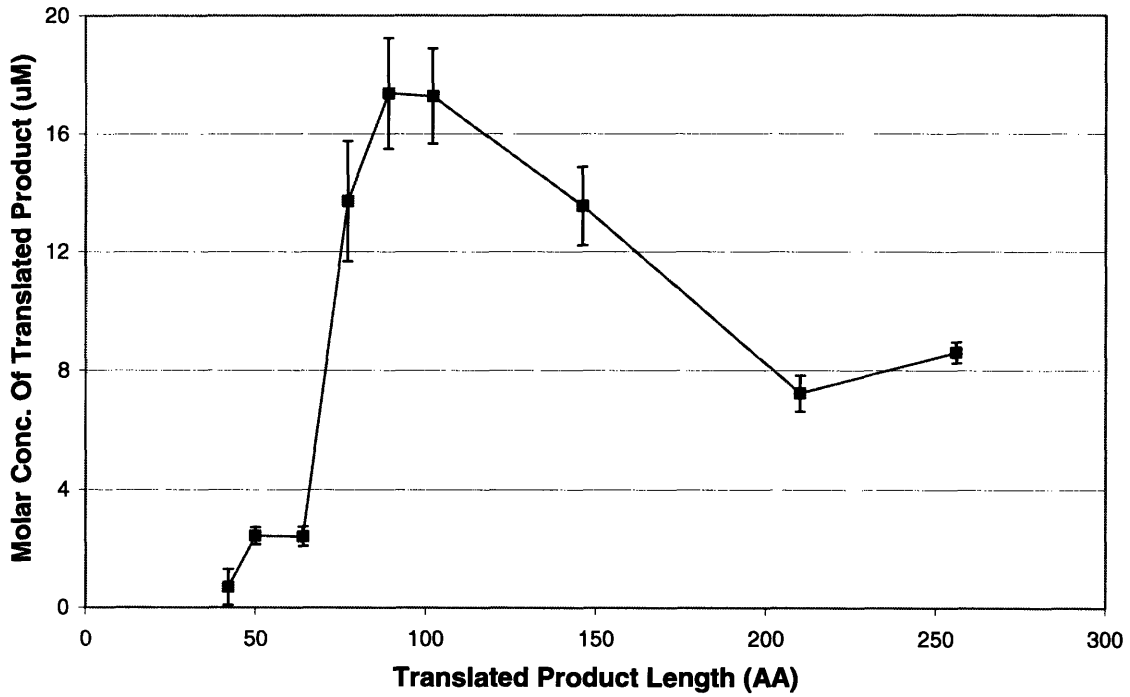



**Figure 3.17: Dependence of molar yield on translated product length for GFP-AmP fusion with stop codon between GFP and AmP.** DNA coding for fusions of truncated GFP and the peptide S1 were translated with a stop codon between the truncated GFP and the S1 for 6 hrs at 30°C in the Roche system. The mRNA was the length of the truncated GFP plus the S1, whereas the translated product was the length of the truncated GFP alone. Optimal yield with stop codon inserted after GFP fusion partner and before S1 has a maximum yield of approximately 100 amino acids, which is similar to that of the fusion partner alone. This indicates that the yield is a function of translated product length and not DNA or mRNA length (n=4, error bars = SD).

#### 3.8.1.2 Results of varying amino acid length independently of mRNA length

As shown in Figure 3.17, the optimum occurred at the same translated amino acid length as with the fusion partner alone, around 100 amino acids. This indicates that amino acid length is critical for producing optimal yield. Additionally, this suggests that the mechanisms responsible for the observed maximum in yield act post-translationally. This observation supports the hypothesis that proteases limit the accumulation of peptides below the optimal length.

### *3.8.2 Protease inhibitor studies*

To further test this hypothesis, truncated GFP constructs of 50 and 102 amino acids were translated with and without a protease inhibitor cocktail added to the Roche *in vitro* translation mixture. In order to examine the rate of degradation only, tetracycline (Tet) was added to halt translation one hour after translation was initiated.

#### 3.8.2.1 Methods for investigating effect of protease inhibitors on degradation rates

For time course experiments with protease inhibitors, GFP constructs of length 42, 50, and 102 amino acids were translated in the Roche Translation Kit with the

composition described in Section 3.5.1, except that the 1.5 ul of Reaction buffer was replaced with 1.5 ul of 10x Roche Complete Mini Protease Inhibitor Cocktail in water, or 1.5 ul water without inhibitor (11-836-153-001). After 1 hour of translation, tetracycline hydrochloride (Sigma, St Louis, MO) was added to 100 ug/ml to stop translation, so the rate of protein degradation could be quantified. Figure 3.18 shows the change in the accumulation of translated product over a 3 hour period following Tet addition.

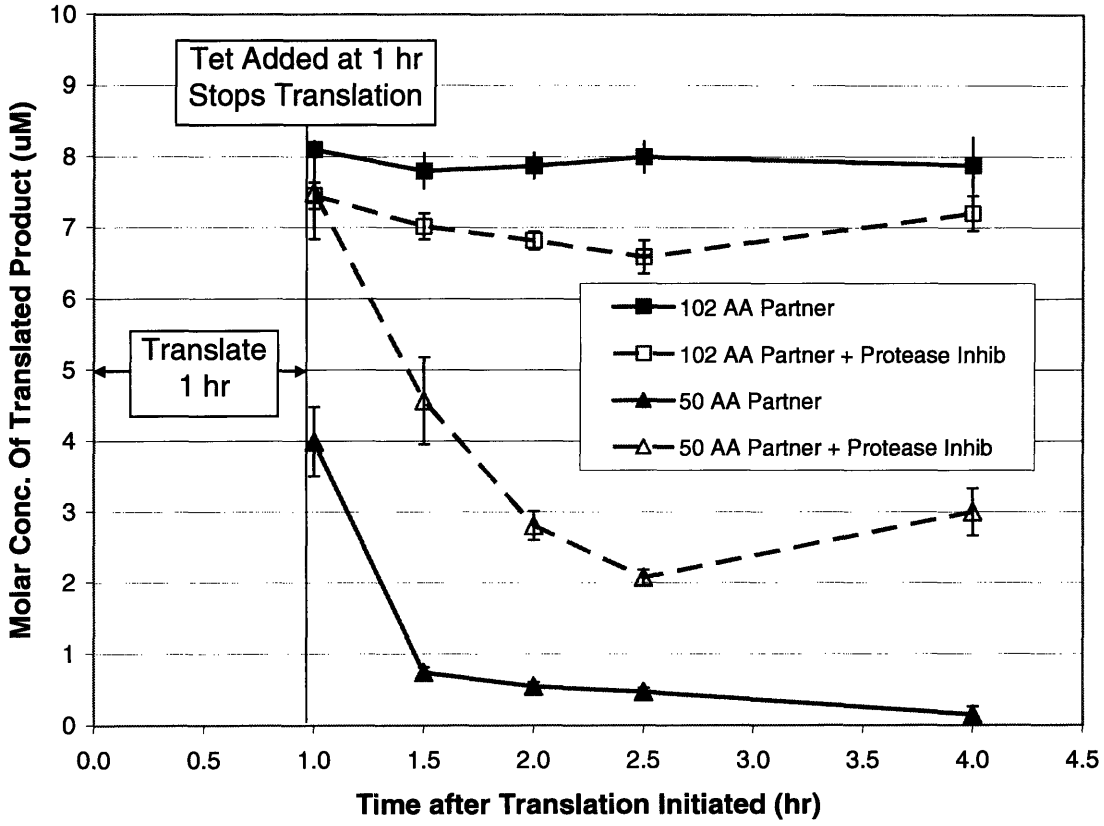



**Figure 3.18: Degradation profile for stable and unstable fusion partners with and without protease inhibitors.** Truncated GFP constructs of 50 and 102 amino acids were translated in the Roche system at 30°C with and without protease inhibitors added. After 1 hour of translation, tetracycline was added to halt translation so that protease degradation could be measured using TCA precipitation. The construct of 102 amino acids did not show susceptibility to proteases with or without inhibitors. The yield of the 50 amino acid construct was twice as high in the first hour with inhibitors present, and without inhibitors, the yield dropped below 1 uM within 30 minutes after translation ceased (n=4, error bars = SD).

### 3.8.2.2 Methods for investigating effect of protease inhibitors on degradation rates

As shown in Figure 3.18, no degradation of the 102 amino acid construct occurred with or without protease inhibitors present, indicating that it is not susceptible to proteases. While obtaining structures of each of the constructs is beyond the scope of this work, the 102 amino acid construct may be long enough to be protected from degradation by tertiary folding. However, degradation of the 50 amino acid truncation occurred and was significantly reduced by protease inhibitors. In the first hour of translation, before Tet is added in Figure 3.18, the accumulation of translated product is two times higher in the presence of protease inhibitors. This accumulation, which is the net of translation minus degradation, indicates that without protease inhibitors, more degradation occurred. This assumes that protease inhibitors do not boost translation rate in the first hour. This is a reasonable assumption because translation of the 102 amino acid construct, which does not display protease susceptibility, is unaffected by the presence of protease inhibitors. After Tet is added to translation of the 50 amino acid construct, samples without protease inhibitors fall to below 1 uM within 30 minutes, but persist longer with protease inhibitors. The fact that some degradation still occurs with protease inhibitors present suggests that the proteases are not completely inhibited.

### 3.8.3 Application of fusion partner optimization

The optimal fusion partners for production *in vitro* may not remain the optimal partners for production *in vivo*. Nevertheless, knowing the optimal length for *in vitro* translation is useful for efficiently carrying out the parallel production of peptides for a high-throughput screen. This platform would be particularly helpful for any activity assay beyond the limitations of phage display screens. The DNA for leads found in this screen may easily be cloned into a suitable host for recombinant production if a large quantity of a given protein is desired for further research. Since not all components of the protein translation and degradation machinery may function identically in *in vivo* and *in vitro* systems, the optimal fusion partner for one system will not necessarily be the optimal fusion partner in the other system.

*In vitro* translation may be useful for finding peptide structure or measuring activity for a wide variety of applications. Our results demonstrate that molar yield may be maximized by using the appropriate length fusion. A fusion partner based on GFP truncations may not be optimal for all applications. However, a fusion partner that has desirable properties for the downstream processing of a peptide be selected and its length optimized to maximize product accumulation. We further suggest that the optimal length should be approximately 100 amino acids. This knowledge is particularly important if only a short HIS tag of 6-10 amino acids is needed, so that the construct may be designed with the correct number of "buffer" codons between the fusion partner and peptide to give the optimal total translated length. Purification, quantification, solubilization, or

inclusion body formation may be accomplished at maximal yields by selecting or designing a fusion partner of the appropriate length.

## 3.9 Improving solubility

While insolubility or inclusion body formation would be desirable for some applications, we desired soluble AmP fusions that could be cleaved by enterokinase and used directly in an antimicrobial assay. Therefore we sought to assess the solubility of the various fusion partners we had created.

### *3.9.1 Measuring solubility of existing templates*

A variety of truncated GFP – AmP fusions were created and solubility evaluated. Because we hope to synthesize a great variety of AmPs, the effect of the AmP sequence on solubility was also evaluated. For these studies, the optimal truncated GFP fusion partner (70 amino acid) was fused to the AmPs: human beta defensin-2, melittin, and S1. The samples were translated using methods described in Section 3.5.1. In order to measure soluble yield, the samples were centrifuged for 15 minutes at 15,000g. The supernatant, which contains the soluble portion of the product, was removed and TCA precipitation applied as in Section 3.5.1. As shown in Figure 3.19, all of the fusions, along with the full GFP had solubilities between 20% and 60%. The economics of the platform could be improved if the translated fusion had greater solubility.

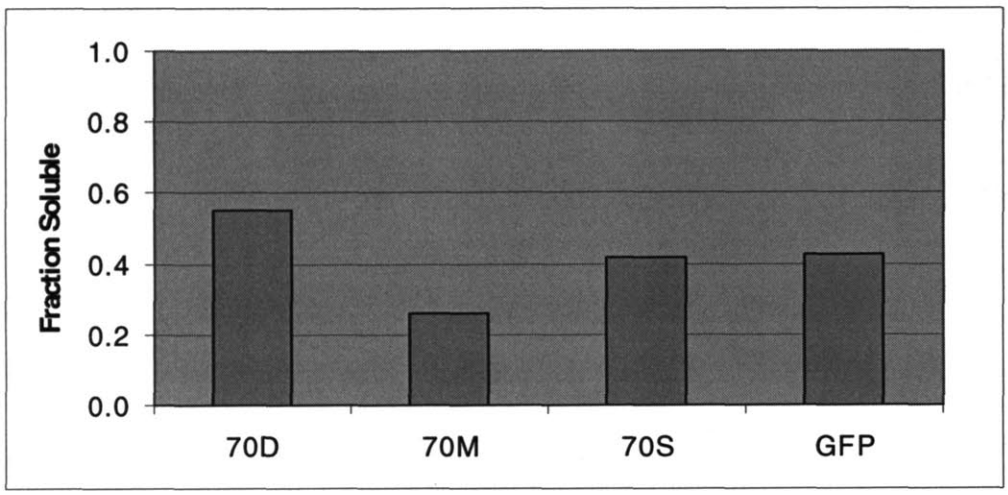



**Figure 3.19: Solubility of truncated GFP-AmP fusions**. The solubility of various AmPs on the optimal truncated GFP fusion partner vary from 20 to 60%. For our application, it would be more desirable if higher solubility were achieved.

### *3.9.2 Selection of appropriate partner*

Fusion partner optimization studies indicated that optimal *in vitro* expression could be achieved when the translated product was approximately 100 amino acids long. In a search of known fusion partners with high solubilities, thioredoxin at 110 amino acids most closely matched this criteria. In addition, recombinant thioredoxin could be purchased in order to verify if it would toxic in an antimicrobial assay after being cleaved from the AmP of interest. Methods given in 4.2.2.3 validated that thioredoxin was non-toxic at 400 ug/ml. Thus, thioredoxin is the appropriate size to give high yields and will not have to be purified after cleavage from the AmP before activity measurements.

#### 3.9.2.1 Methods to create and translate a Trx fusion partner

The *E. coli* thioredoxin construct was obtained from the pET32a vector using methods described in Section 3.7.2.1. The linear thioredoxin DNA was translating in the Roche translation system as described in Section 3.5.1. A DNA concentration of 4 ug/ml was used, which was sufficient to give maximal yield for GFP constructs of this size. When the linear thioredoxin was translated the translation yield was only 4 uM. This is significantly lower than the 16-20 uM that resulted from truncated GFP of this length. We sought to determine the limiting factor in the thioredoxin concentration that was lowering yields. The first hypothesis we sought to test was if the translation was DNA or mRNA limited. This seemed unlikely since increasing DNA concentration above 4 ug/ml did not increase translation of truncated GFP constructs. Nevertheless, if the DNA

or mRNA of the thioredoxin was less stable than that of the GFP constructs, degradation could lead DNA or mRNA concentration to be limiting. In order to test this hypothesis, a series of *in vitro* translation reactions were carried out in which the DNA concentration of thioredoxin coding template was varied. Translation and quantification methods are given in Section 3.5.1. The resulting yields are shown in Figure 3.20.

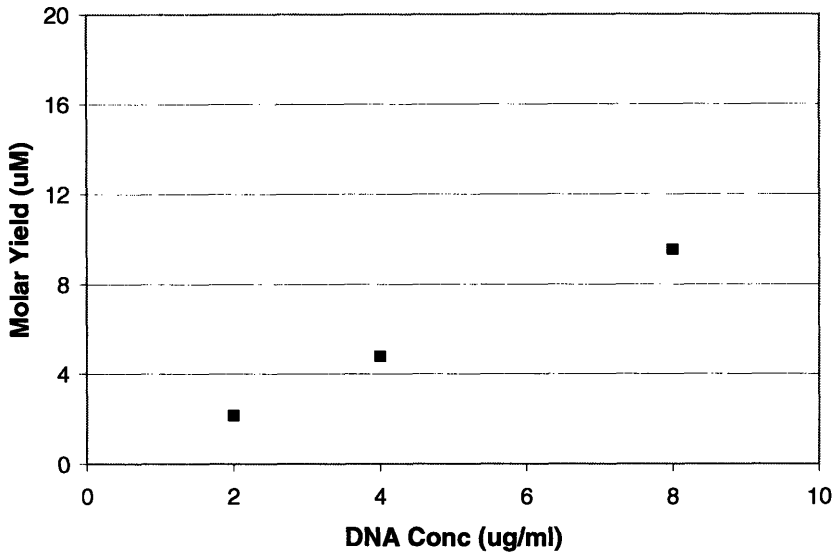



**Figure 3.20: Effect of DNA concentration on molar yield of Thioredoxin**. The yield of thioredoxin was approximately linear with DNA concentration during *in vitro* translation. Further, at a DNA concentration of 4 ug/ml the yield was only 5 ug/ml, whereas a yield of approximately 20 ug/ml would be expected for truncated GFP of this length.

#### 3.9.2.2 Results of initial Trx translations

As seen in Figure 3.20, the yield of thioredoxin was directly proportional to the concentration of DNA used. This indicated that the mRNA was unstable. If the mRNA was sufficiently stable that its rate of production by translation exceeded its rate of degradation, then the mRNA concentration would accumulate over time. This accumulation would lead to an mRNA concentration that was not directly proportional to the DNA concentration or the transcription rate. Since the translation seen in Figure 3.2 is strictly proportional to DNA concentration, accumulation of mRNA that would decouple the translation rate from DNA concentration does not appear to be occurring.

Additionally, the DNA concentration could not simply be increased to improve yield without increasing the cost of the high-throughput system. The goal of 4 ug/ml for the desired volume of *in vitro* translations was chosen based on the amount of DNA template for thioredoxin that can be produced in a single 50 ul PCR reaction. Increasing DNA concentrations would increase the cost of polymerase enzyme and limit the throughput that may be achieved with one PCR machine if multiple wells of 96 well PCR plates were necessary for each construct produced.

### *3.9.3 Creating high-yield Trx translations from low DNA template concentrations*

In order to further test the hypothesis that mRNA stability was limiting yield from the thioredoxin vector, we sought to translate the same thioredoxin vector with more

stable mRNA. mRNA stability has been linked to the sequence of the untranslated regions on either end of the mRNA construct. These regions, particularly the 5' region, protect the mRNA from degradation and turnover *in vivo* (Mignone, Gissi et al. 2002). The upstream region from the thioredoxin had been taken from the pet32a plasmid, which was designed for translation *in vivo*. This upstream sequence may not have been well designed to give high yields *in vitro*.

The hypothesis that the stability of the mRNA limited thioredoxin yield could be tested if the thioredoxin coding region were ligated to an upstream region known to give high protein yields at low DNA concentrations. As demonstrated with the shortened GFP constructs, the upstream region of Roche-GFP vector gave translation that was independent of DNA concentration down below 4 ug/ml. This indicated that the mRNA had sufficient stability that the translation yield was independent of mRNA concentration.

### 3.9.3.1 Methods for creating Trx on GFP upstream region

The upstream region of the GFP control vector, along with the first 6 amino acids after the start codon were obtained using PCR and a ligation site was added on the 3' end. This segment was cut after the start codon because the base pairs from the ribosome binding site to the start codon are known to be critical for *in vitro* translation. The thioredoxin vector was obtained from pet32a using PCR with a matching ligation site on the 5' end. After digesting, ligating, and amplifying, the resulting DNA template was verified by sequencing.

The revised thioredoxin vector was translated using the same conditions used when translating the original thioredoxin vector shown in Figure 3.20. The same DNA concentrations were used, with the expectation that a more stable mRNA would reach saturation at a lower concentration, producing higher yields of thioredoxin. The results are shown in Figure 3.21.

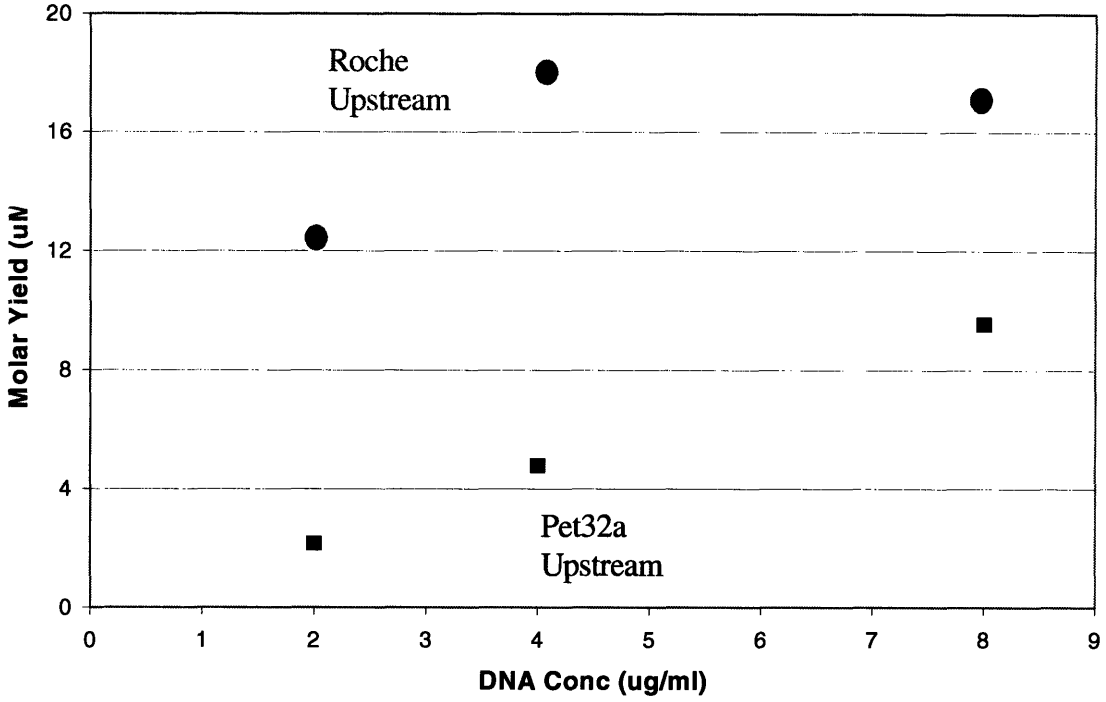



**Figure 3.21: Molar yield of thioredoxin versus DNA concentration for both a Pet32a and a Roche upstream region.** The yield of thioredoxin varies almost linearly with the DNA concentration when a pet32a upstream region is used. This indicates that the mRNA is not stably accumulating in the translation mixture. In order for a cost-effective translation to be executed, a maximal yield is desired at a DNA concentration of approximately 4 ug/ml. To address this instability, the thioredoxin vector was attached to a Roche upstream region known to produce maximal yields independent of DNA concentration down to 4 ug/ml. This solution produced thioredoxin yields equal to the maximum expected from previous fusion partner optimization studies (18 ug/ml) at the desired DNA concentration.

### *3.9.3.2 Effect of upstream region on Trx yield at low DNA concentrations*

As shown in Figure 3.21, the yield of the thioredoxin on the revised upstream region was decoupled from DNA concentration down to 4 ug/ml, as seen with the original truncated GFP constructs. Further, the yields were similar to the yields with the GFP constructs, with a yield of 18 uM for a translated product of length of approximately 110 amino acids. This yield met our goal for translation, and the low DNA concentration necessary meant that sufficient DNA could be obtained for each desired AmP from a single 50 ul PCR reaction, which reduces costs.

## 3.10 Developing purification scheme

A requirement of the high-throughput screening platform being developed is that the synthesis portion of the system does not affect the antimicrobial measurements. Interference could arise in the form of toxicity from the components of the synthesis system or side products of the synthesis. In our case, we wanted to study the effect of the *in vitro* translation mixture on a standardized antimicrobial assay. The toxicity of each component was tested using methods given in Section 4.2.2.3. Toxicity to the bacteria was seen from every component of the mixture, indicating that the reconstitution buffer, used to resuspend all components of the kit, was the toxic agent. This was confirmed by assessing the buffer on its own, and verifying that other components reconstituted in

water were non-toxic. A variety of strategies to eliminate this toxicity while retaining high peptide yield are presented below.

### *3.10.1 Protein precipitation to remove toxic in vitro translation buffer component*

A simple approach to removing the toxic agent would be to precipitate all proteins following translation, including the antimicrobial peptide still attached to a fusion partner of interest. Assuming the toxic agent remains soluble, the proteins can be resuspended for use in an antimicrobial assay. A variety of standard techniques exist for protein purification, including ethanol, acetone, and ammonium sulfate precipitation. For all experiments, standard translation conditions given in Section 3.5.1 were used. For precipitation, 90 vol% ethanol, 80 vol% acetone, or 80 vol% 4 M ammonium sulfate was added. Samples were held for 5 minutes at -20°C (ethanol, acetone) or on ice for ammonium sulfate. Mixtures were centrifuged for 15 minutes at 4°C at 15,000g, supernatant was removed, and the pellet was dried and resuspended in Hepes KOH. Fractional recovery was measured using TCA precipitation as described in Section 3.5.1. Ammonium sulfate was chosen because it was easiest to resuspend the proteins after precipitation, relative to ethanol and methanol. As shown in Figure 3.22, recoveries ranged from 60 to 75%. Further, the resuspended mixture was non-toxic, indicating the toxic component had been removed to a sufficient extent.

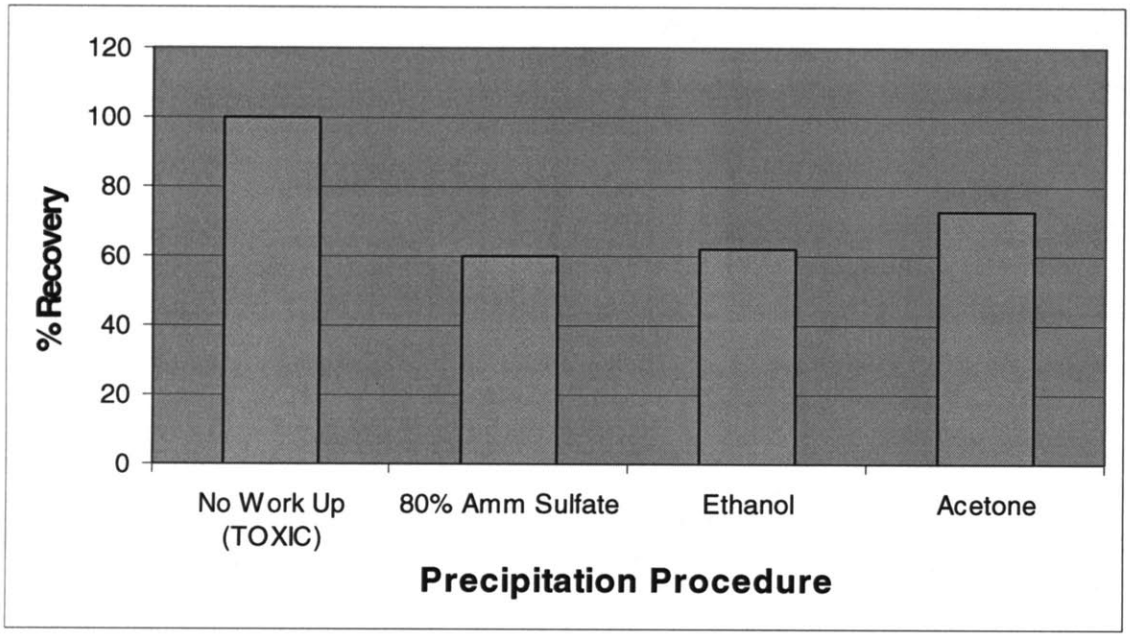



**Figure 3.22: Recovery of protein following various protein precipitation procedures for purification**. Because the *in vitro* translation mixture contains a component that kills bacteria, precipitation procedures using 80% ammonium sulfate, ethanol, and acetone were carried out. Recoveries varied from 60 to 80%, with the acetone procedure giving the smallest reduction in yield. All purified samples did not display antimicrobial activity.

### 3.10.2 Engineering toxicity out of in vitro system

A superior approach to avoiding toxicity from the *in vitro* translation system would be to develop an alternative reaction mixture without toxicity, rather than adding a protein precipitation work up. Each component of the Roche kit was tested for toxicity using an antimicrobial assay described in Section 4.2.2.3. The reconstitution buffer was determined to be the sole source of toxicity. Unfortunately, Roche would not divulge the components of this buffer. Because the Roche system was based on a process developed in Prof. James Swartz's lab in Stanford University, a series of salt solutions were evaluated that were present in his published translation systems in the hopes of creating a mixture with a high yield and no toxicity (Kim and Swartz 2001). Translations were carried out using methods given in Section 3.5.1, with the exception that the reaction buffer was replaced by pH 7.4 Hepes KOH supplemented with: 80 mM ammonium acetate and 16 mM magnesium acetate, 230 mM potassium glutamate, all three, or no supplemental salts. The resulting yields are given in Figure 3.23.

By supplementing a Hepes buffer with the appropriate complement of salts, approximately 85% of the yield was achieved. Importantly, this buffer did not display toxicity in the same antimicrobial assay for which the Roche buffer was toxic. For a small reduction in yield, this alternate buffer eliminates the need for a protein precipitation step.

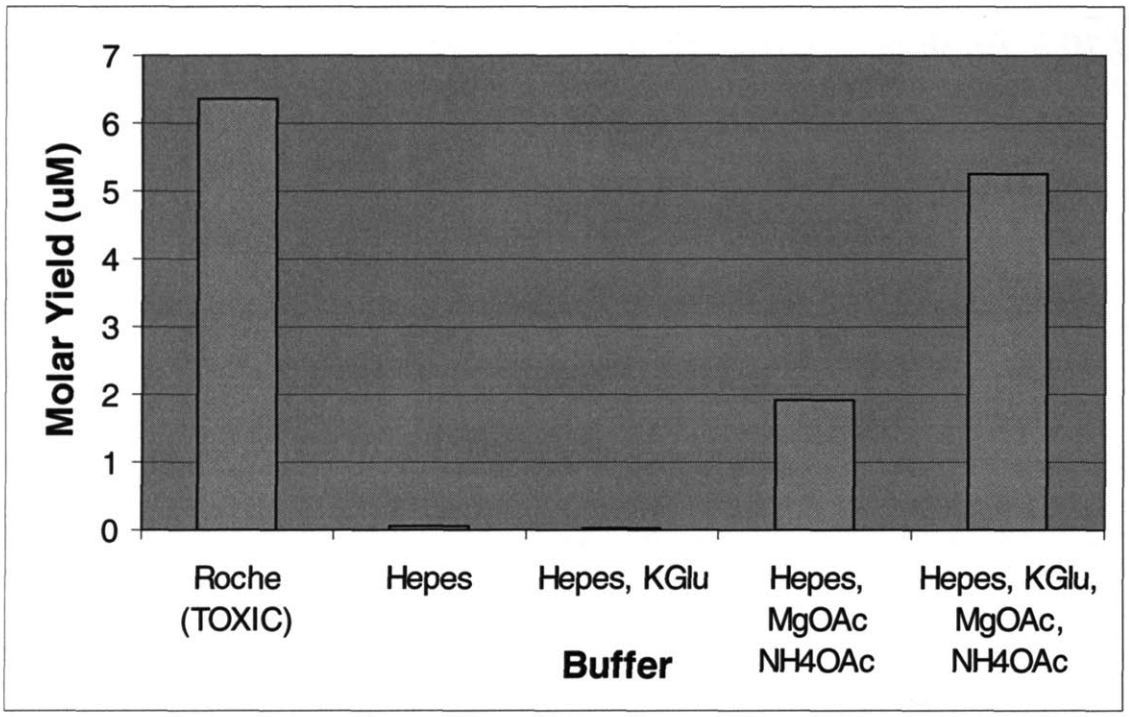



**Figure 3.23: *In vitro* translation yields using non-toxic reaction buffers.** It was determined that the Roche translation buffer was responsible for the toxicity of the *in vitro* mixture. An alternative approach was undertaken to eliminate this toxicity by designing a replacement buffer. Yields were improved by supplementing a Hepes buffer with potassium glutamate, magnesium acetate, and ammonium acetate, with the combined buffer achieving 85% yield of the Roche buffer without having any toxicity.

### *3.10.3 High-throughput isolation of AmPs*

The toxicity of the *in vitro* system was eliminated through the development of an alternate buffer for translation. This improvement enabled the testing of whether the revised *in vitro* system would significantly influence the measurements of MICs, even if the mixture was not toxic. Unfortunately, the *in vitro* translation mixture still affected antimicrobial activity measurements using methods given in Section 4.2.2.3 even after the toxic component had been removed. The cellular extract may alter the growth resources for target bacteria or may bind to or otherwise inhibit the AmP. A strategy to remove cellular components from the Fusion partner-AmP are described below.

#### 3.10.3.1 His-tag purification

A desired purification scheme for an *in vitro* AmP production scheme would be amenable to automation (preferably in 96 well plates), have little product loss, and would be very selective for the desired product. To this end, a commercial 96-well plate HIS-tag purification kit (Vivapure Cobalt Chelate) was utilized to purify the Trx-AmP. Standard kit buffers were used, with product eluted in 500 mM Imidazole. Using this approach, a recovery fraction of the Trx-AmP fusion 0.67 was achieved, with other extract components removed.

### 3.10.3.2 Dialysis to remove imidazole

While the purification scheme removed the components of the *in vitro* translation mixture which impeded accurate MIC measurements, it introduced imidazole, which may be toxic to bacteria. First, if imidazole was only mildly toxic, the elution conditions could be adjusted to reduce imidazole at the expense of a portion of the protein recovery. The MIC of imidazole was measured by using methods given in Section 4.2.2.3. Additionally, the effect of imidazole on MIC quantification was assessed by measuring the MIC of cecropin in the presence of a range of imidazole concentrations. The OD of samples after incubation is given in Figure 3.24. Imidazole shows toxicity at 30 mM and above, but has no effect on MIC determination at 10 mM or below. In order to reduce the imidazole concentration below 10 mM, 96-well dialysis could be used. Dialysis was carried out using a 2 kDa MW cutoff membrane and a pH 7.2 Hepes KOH buffer overnight.

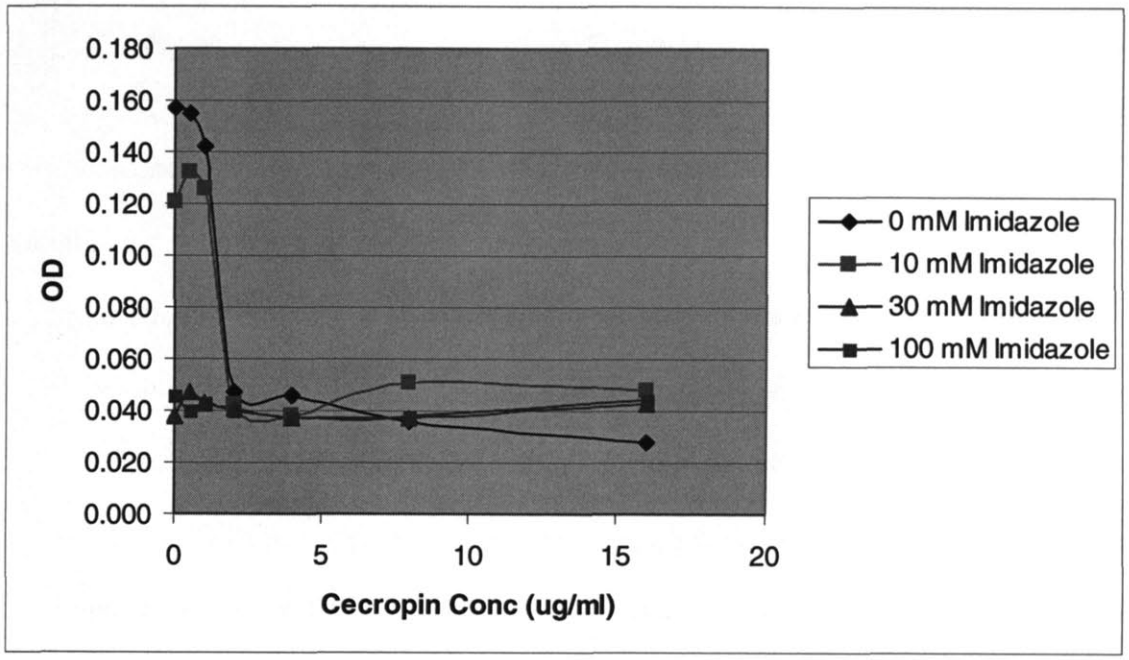



**Figure 3.24: Effect of imidazole on MIC determination for cecropin**. Bacteria was grown in the presence of various concentrations of imidazole and cecropin using methods given in Section 4.2.2.3. Imidazole was toxic at 30 mM and above. Imidazole at 10 mM and below was non-toxic, and furthermore, the presence of imidazole at 10 mM did not affect the measured MIC for cecropin. The desired result for dialysis to remove imidazole will be 10 mM or below.

### 3.10.4 Activity of purified AmP

After all components of the *in vitro* translation and purification system were removed apart from the AmP and its fusion partner, the activity of cecropin, a natural AmP, was validated versus a recombinant standard. The fusion was translated using methods in Section 3.5.1, and HIS-tag purification in Section 3.10.3.1. The eluted sample was dialyzed for 16 hours in pH 7.2 Hepes buffer in a 2 kDa MW cutoff membrane. Enterokinase was added to 0.34 units/ml for 20 hours at 30°C. To assess antimicrobial activity, the standard MIC assay was adapted to allow a higher volume fraction of the resuspended. Twelve ul of purified Trx or Trx-cecropin was added to 4 ul of *E. coli* at OD ~0.0125 in Mueller Hinton Broth (MHB) and 4 ul of 4x MHB. Samples were incubated overnight at 37°C and assayed for growth inhibition visually. The Trx and Trx-cecropin samples were tested in dilutions up to 8 ug/ml. An overview of the synthesis and assay scheme, as well as the resulting MICs are given in Figure 3.25.

Thioredoxin translated on its own did not display antimicrobial activity. When cecropin was translated in frame and digested from the thioredoxin, an MIC of 4 ug/ml was measured. This is marginally less active than the 1-2 ug/ml measured for a recombinant cecropin standard. Nevertheless, this degree of accuracy is suitable for a screen, the results of which would be verified through chemical synthesis.

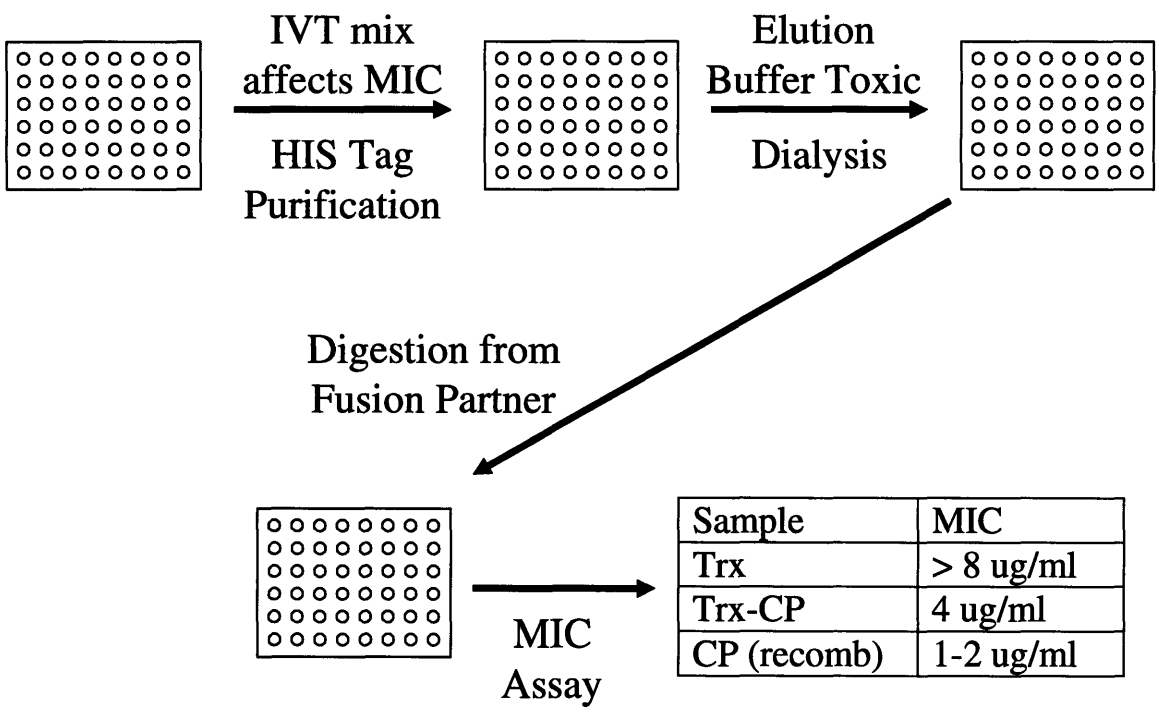



**Figure 3.25: Synthesis and purification of cecropin (CP) with resulting activity.** Following *in vitro* translation of thioredoxin-CP, HIS tag purification removes translation components that affect MIC measurement. Next, dialysis removes toxic imidazole from the HIS tag elution buffer. The AmP is digested from the fusion partner and available for MIC measurement. When assayed for antimicrobial activity, thioredoxin without cecropin did not have a measurable MIC, whereas cecropin translated with and digested from thioredoxin inhibited growth at 4 ug/ml. This is marginally higher than a recombinant cecropin standard with an MIC of 1-2 ug/ml.

## 3.11 Summary

*In vitro* translation is an attractive approach to produce a diverse set of peptides which could be toxic to host-cells in fermentations. Nevertheless, the cost of production for this system had to be reduced for this method to become feasible. This was accomplished through a series of innovations. First, we developed a cost-effective strategy for constructing oligonucleotides which mimicked a pattern-based approach to AmP design. Next, peptide yield was improved by optimized the size of a fusion partner and solubility was improved. Finally, a purification scheme was created so that the synthesis approach would not significantly impact MIC measurement. These improvements, along with their impact on cost, are described in Table 3.2. Work on fusion partner optimization was published as (Loose, Langer et al. 2007).

| Process | Old | New | Changes |
|---|---|---|---|
| Oligo Cost | $90 | $7 | Combinatorial,<br>Ligate Trans Term |
| Amplification | $12 | $3 | Lower DNA Conc. Needed<br>Improved Upstream |
| Translation | $100 | $3,<br>batch more robust | Optimal Size Tag –<br>Yield up 4x<br>Solubility up 2.5x<br>Cost down 4x per ul reaction |
| Assay | Prohibitively slow | Automatable,<br>clinically relevant | Switch to NCCLS microdilution,<br>luciferase |

**Table 3.2: Summary of improvement to a high-throughput *in vitro* translation and assay scheme for AmPs**. This table gives the cost per peptide for producing 100 peptides (10x10 combinations) at sufficient quantity to test against 2 bacteria at up to 16 ug/ml. Developing a combinatorial scheme reduces the cost of the oligonucleotides by the square root of the number of sequences tested. Changing the upstream region of thioredoxin reduced the required amount of DNA for maximal translation, reducing the cost of amplifying DNA. Translation yield was maximized by optimizing the translation partner, improving solubility, and switching to a lower cost, more robust batch system. Finally, the assay was switched from a prohibitively slow plate counting technique to a standardized microdilution assay.

# Chapter 4

# AmP Design Based on Pattern Discovery

Existing techniques for AmP design rely largely on screening organisms for new AmPs or optimizing known sequences. In order to increase the diversity of known AmP sequences, we generated a pattern-based approach in which semi-conserved rules are derived by examining the set of all natural AmPs, and the rules are used to guide novel designs. First, the methods of pattern discovery and their application to AmPs are described. Next, approaches to filter patterns for significance and reassemble them in non-homologous sequences are outlined. Finally, methods to synthesize and evaluate designed sequences are given, along with the performance of the designed agents.

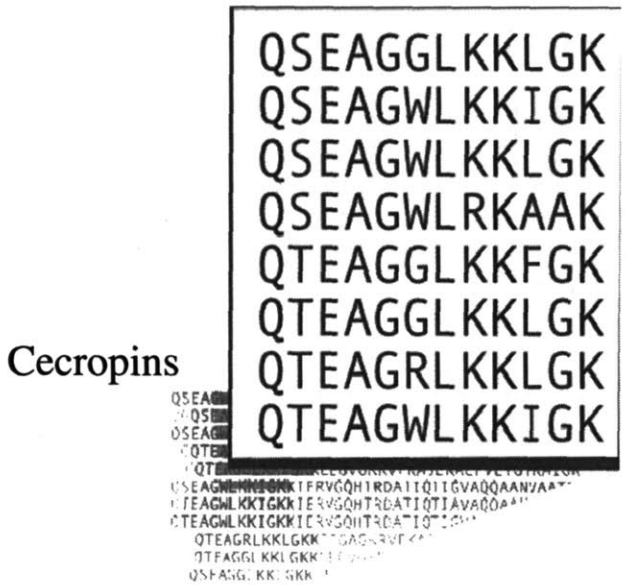
## 4.1 Pattern-based AmP design

More diverse natural AmPs could be designed if general rules about AmP sequences could be deduced. Unlike full proteins, many AmPs do not depend on tertiary structure for their activity (Boman 2003). Their activity is derived from the secondary structure of the peptides, which is tightly tied to the primary sequence. It is our

hypothesis that developing rules about the allowed primary sequence for AmPs may generate useful design parameters for novel AmPs.

### 4.1.1 Patterns in natural AmPs

Our preliminary studies of natural AmPs indicated that their amphipathic structure gives rise to a modularity among the different AmP amino acid sequences. The repeated usage of sequence modules — which may be a relic of evolutionary divergence and radiation — is reminiscent of phrases in a natural language, such as English. For example, the pattern Q.EAG.L.K..K (the "." is a "wildcard", which indicates that any amino acid will suffice at that position in the pattern) is present in over 90% of cecropins, an AmP common in insects. This example instance of a motif is given in Figure 4.1. Based on this observation we modelled the AmP sequences as a formal language — a set of sentences using words from a fixed vocabulary. In this case, the vocabulary is the set of naturally occurring amino acids, represented by their one–letter symbols. We attempted to discover rules about the arrangement of amino acids by examining the broad set of natural AmPs.




Q[ST]EAG.L[KR]K.[GA]K

**Figure 4.1: Highly conserved motif from cecropins**. Alignment of cecropins from a variety of organisms reveals a highly conserved pattern, or motif. This motif can be represented by the bottom sequence in which bracketed terms indicate a specific set of amino acids may be selected at a position, and a period indicates that any amino acid may be selected at a position. More loosely conserved motifs exist across AmP classes, and will be utilized in our design algorithm. Figure reproduced from KL Jensen.

We conjectured that the "language of AmPs" could be described by a set of regular grammars. Regular grammars are, in essence, simple rules for describing the allowed arrangements of words. These grammars, such as the cecropin pattern mentioned previously, are commonly written as regular expressions and are widely used to describe patterns in nucleotide and amino acid sequences (Falquet, Pagni et al. 2002; Searls 2002).

Previous approaches to the creation of synthetic AmPs have produced peptides that are either closely related to naturally occurring peptides or composed of only a handful of amino acids (Tiozzo, Rocco et al. 1998; Tossi, Sandri et al. 2000). Because of the great similarity between the designed peptides and native sequences, they often have similar efficacies against the same microbes as natural AmPs. In order to design peptide that target novel diseases or were much more potent than known peptides, sequences with little relation to the current database would be desired. As shown in our prior research, AmPs with little overall homology to existing peptides could be designed as long as underlying patterns for activity are retained.

In contrast to techniques in which single amino acids of existing peptides are randomly changed, our synthetic AmPs may be designed with limited homology to existing AmPs. Our peptides will populate a region of sequence space that is not occupied by naturally occurring AmPs, as shown in Figure 4.2. While peptides that we design in grammar space may not overlap perfectly with functioning AmPs, they will contain many antimicrobials that are outside of the realm of known peptides. These newly designed peptides may have activity against microbes with reduced susceptibility to existing AmPs. Further, it has been suggested that using therapeutic agents highly homologous to natural AmPs could compromise our natural defenses (Bell and Gouyon

2003). In essence, our linguistic approach is a means to rationally expand the natural sequence space without using structure-activity information or complex folding simulations. Instead, we will rely upon the ability of sequence patterns to capture the underlying function of the peptides.

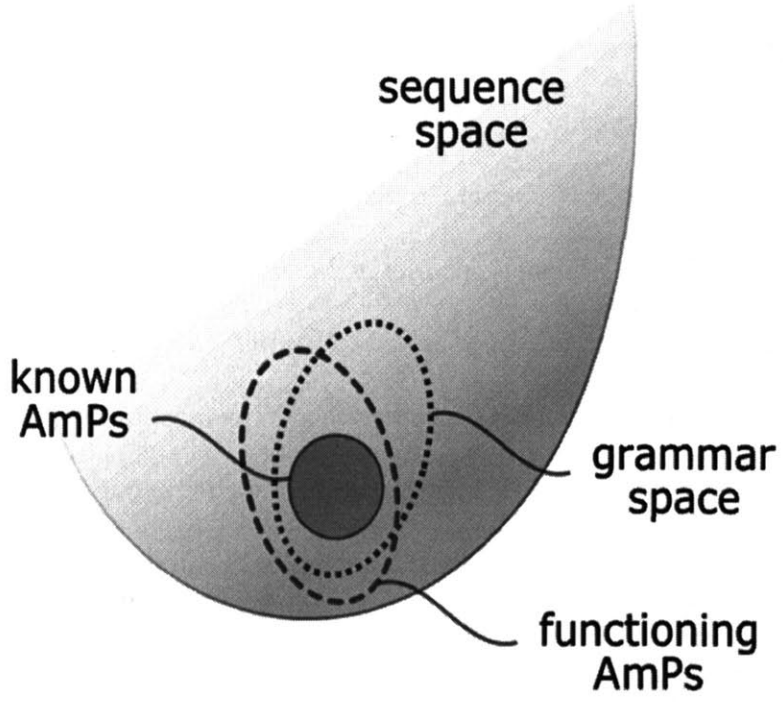



**Figure 4.2: Representation of a grammar based search space**. While sequence space for 20 amino acid long peptides (typical for AmPs) contains approximately 10^26 sequences, only hundreds of AmPs are known. Our goal is to develop a grammar based approach that generates a sequence space that overlaps well with the much larger space of all possible functioning AmPs, while eliminating most other portions of sequence space to give a high probability of predicting active sequences.

### 4.1.2 Elucidation of grammars

In order to develop grammars based on natural AmPs, two criteria must be defined:

1. **Antimicrobial Peptide Database.** A wide variety of collections of natural AmPs exist. Additionally, many peptides in the Swiss-Prot collection are annotated as having antimicrobial activity. These databases vary in size and quality. Importantly, the database should be curated to ensure that the entries are well justified and that only the antimicrobial portion of a peptide or protein is included. Most AmPs are translated within cells as part of a larger sequence including a non-antimicrobial pro-region. Because these pro-regions are often highly conserved, they would dominate pattern discovery attempts and result in designs that do not have activity.
2. **Pattern Discovery Tool.** Many tools for discovering patterns in sequential data exist. While some find only a portion of patterns meeting defined criteria within a dataset, others are exhaustive. The discovery tools also vary in their runtime for a given data set, and the degree to which users may tune the parameters for the patterns being discovered. Ideally, a discovery tool should be efficient, exhaustive, and highly adaptable.

#### 4.1.2.1 *Database selection*

After comparing the available databases for natural AmPs, the University of Nebraska Medical Center Antimicrobial Peptide Database was chosen (http://aps.unmc.edu/AP/). This database includes 526 sequences, all of which have only the active AmP included. The data may be subdivided based on length, nature of activity (antimicrobial, antifungal, viral, cancer), structure, hydrophobic percentage, or net charge.

#### 4.1.2.2 Teiresias

TEIRESIAS is an algorithm that discovers patterns in any database of sequential data including DNA and peptides. A full discussion of TEIRESIAS is available in (Rigoutsos and Floratos 1998). All patterns are discovered given the following parameters:

1. Patterns must be at least L characters long.
2. Patterns must occur at least K times in the dataset to be considered significant.
3. Patterns will have a minimum fraction of L/W non-wild card characters over any window of length W.

Example patterns discovered using these parameters are shown in Figure 4.3. After pattern discovery is executed, resulting patterns should be filtered to ensure they are statistically significant, and further data processing may be desired. A variety of methods are available for this, and a sample methodology is shown in Figure 4.4.

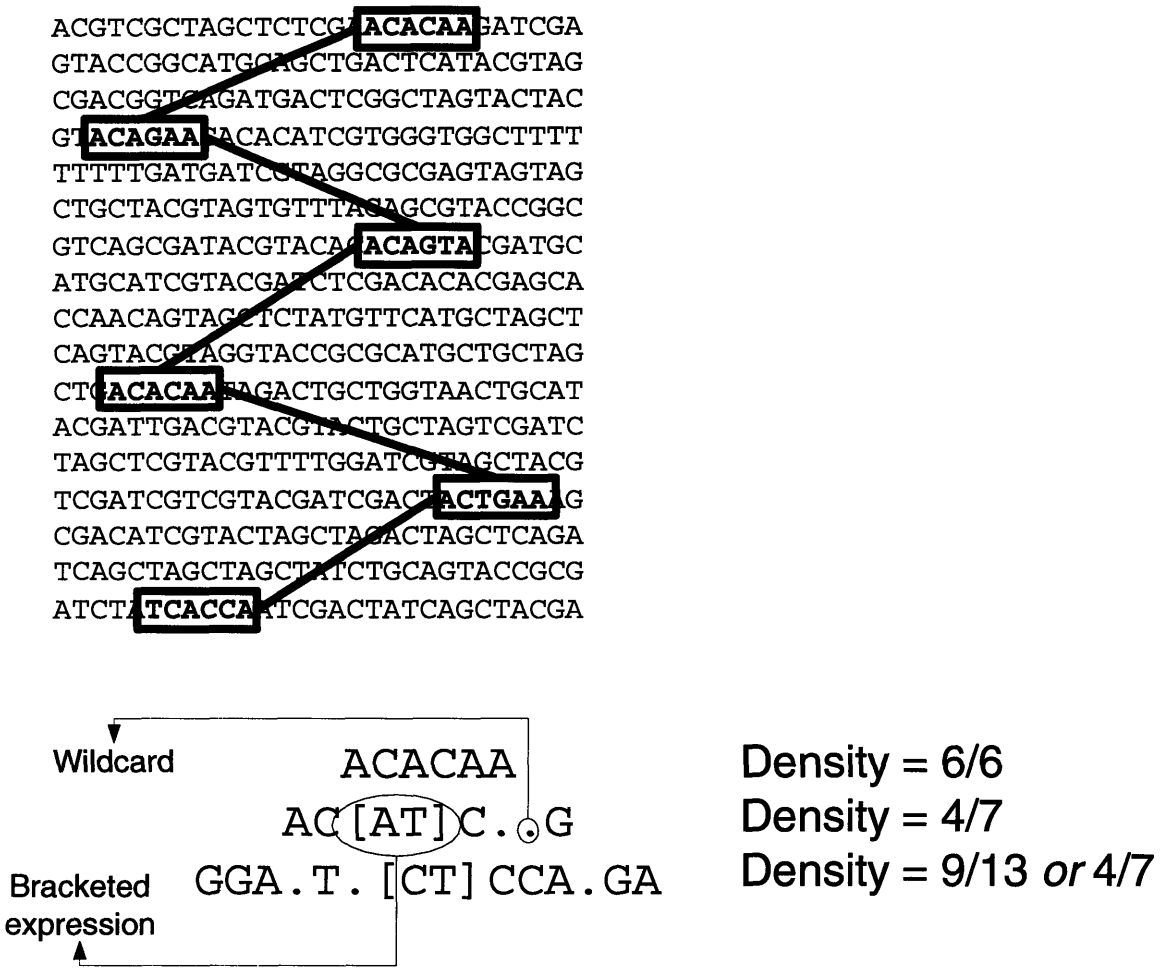



**Figure 4.3: Sample results of a TEIRESIAS search of a sample text**. Sample patterns which meet an L/W of 6/6, 4/7, and 9/13 are shown. An L/W of 4/7 means that in a sliding window of 7 characters, at least 4 must be uniquely specified (not bracket or wild card). Additional parameters in TEIRESIAS include L, the minimum length of a pattern and K, the minimum number of times a pattern must appear in the training set to be considered significant. Figure adapted from KL Jensen.

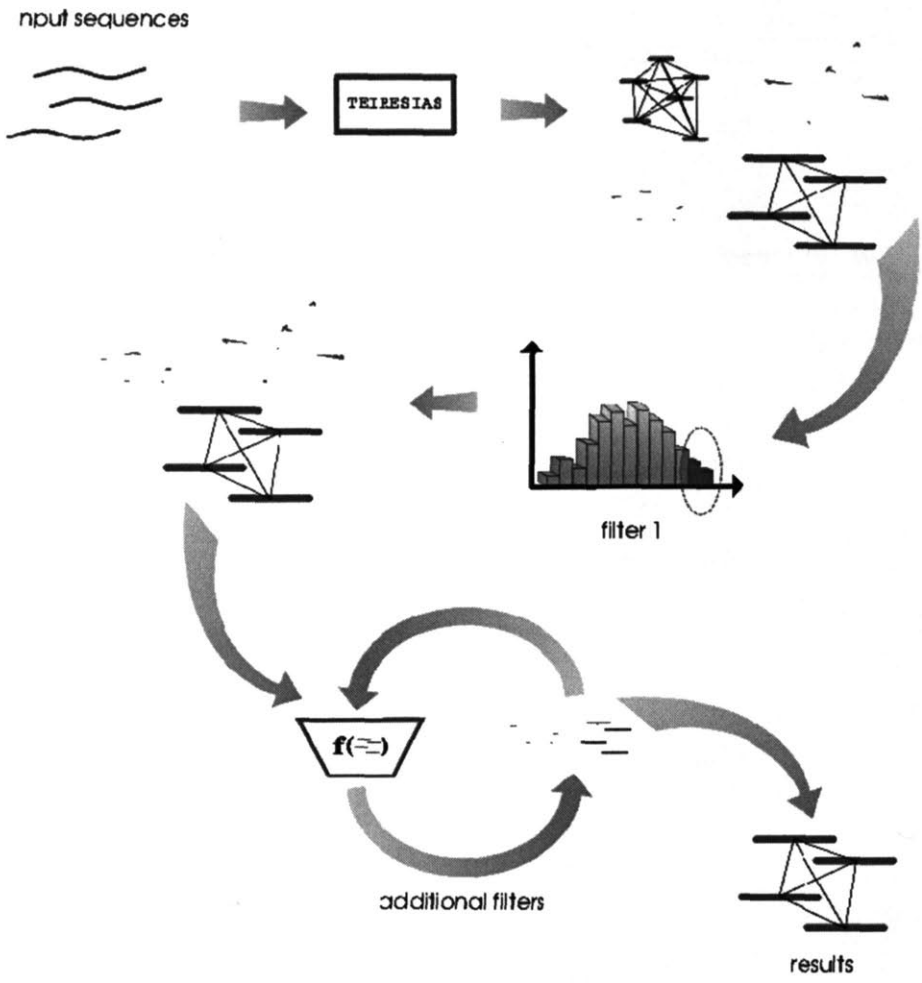



**Figure 4.4: A generalized flowsheet for pattern discovery using TEIRESIAS.** Input sequences are exhaustively searched by TEIRESIAS for patterns meeting certain parameter specifications. A filter for statistical overrepresentation is applied next, such as a Markov model (filter 1), limiting the size of the candidate set. Additional filters including a chi-squared test across defined subgroups or clique formation may also applied until the candidate set. Figure reproduced from KL Jensen.

#### 4.1.2.3 Application of Teiresias to AmPs

Using the set of 526 AmP sequences from APD, we ran the Teiresias pattern discovery tool with the following settings: L = 6, W = 6, and K = 2. The resulting grammar set was masked from the input sequences and the process was repeated using L = 7, W = 15, K = 5 with the following amino acid equivalency groups [[AG], [DE], [FYW], [KR], [ILMV], [QN], [ST]]. (See Rigoutsos and Floratos (1998) for a background on pattern discovery, masking, and terminology used in these methods.) With Teiresias, we derived a set of 684 regular grammars that occur commonly in a set of 526 well-characterized eukaryotic AmP sequences from the Antimicrobial Peptide Database (APD) (Wang and Wang 2004).

Together, these ~700 grammars describe the "language" of the AmP sequences. In this linguistic metaphor, the peptide sequences are analogous to sentences and the individual amino acids are analogous to the words in a sentence. Each grammar describes a common arrangement of amino acids, similar to popular phrases in English. For example, the frog AmP brevinin-1E contains the amino acid sequence fragment PKIFCKITRK, which matches the grammar P[KAYS][ILN][FGI]C[KPSA][IV][TS][RKC][KR] from our database. (The bracketed expression [KAYS] indicates that, at the second position in the grammar, either lysine, alanine, tyrosine, or serine is equally acceptable.) Based on this match, we would say that the brevinin-1E fragment is "grammatical".

The grammars were further processed for convenience in AmP design. Teiresias outputs its grammars in regular expression format, using wildcards. To make the grammars more selective, we de–referenced each wildcard in the grammars to a

bracketed expression. That is, we replaced each wildcard with the set of amino acids implied by the grammar's offset list. Finally, to allow partial matches as short as 10 amino acids, we divided each grammar into sub–grammars using a sliding–window of size 10, resulting in 1551 grammars of length ten.

#### 4.1.2.4 Filtering for selectivity

Using the set of 526 AmP sequences from APD, we ran the Teiresias pattern discovery tool with the following settings: L = 6, W = 6, and K = 2 (a detailed description of the Teiresias input parameters and associated tools is available elsewhere (Rigoutsos and Floratos 1998)). The resulting grammar set was masked from the input sequences and the process was repeated using L = 7, W = 15, K = 5 with the following amino acid equivalency groups [[AG], [DE], [FYW], [KR], [ILMV], [QN], [ST]]. (See Rigoutsos and Floratos (1998) for a background on pattern discovery, masking, and terminology used in these methods.)

Teiresias outputs its grammars in regular expression format, using wildcards. To make the grammars more selective, we de–referenced each wildcard in the grammars to a bracketed expression. That is, we replaced each wildcard with the set of amino acids implied by the grammar's offset list. Finally, to allow partial matches as short as 10 amino acids, we divided each grammar into sub–grammars using a sliding–window of size 10, resulting in 1551 grammars of length ten.

By design, each grammar in this set of ~700 grammars is ten amino acids in length and is specific to AmPs --- at least 80% of the matches for each grammar in

Swiss–Prot/TrEMBL are found in peptides annotated as AmPs (Bairoch and Apweiler 2000). APD is a subset of Swiss–Prot/TrEMBL.

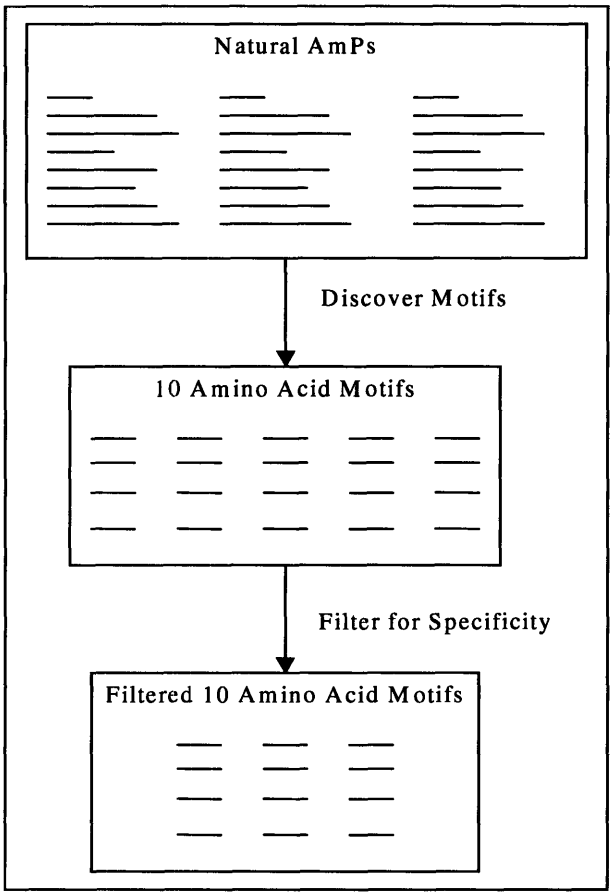



**Figure 4.5: Schematic of the discovery and filtering process for antimicrobial motifs.** TEIRESIAS was utilized to discover motifs from the approximately 500 natural AmPs. These motifs were broken to 10 amino acid segments, which is long enough to have relevant motifs, but short enough they could be manipulated and combined to produce the desired design length of 20 amino acids. Additionally, the motifs were filtered to be specific to AmPs by ensuring that 80% of the sequences matched in Swiss–Prot/TrEMBL were annotated as antimicrobial.

### 4.1.3 Designing novel AmPs from grammars

The grammars discovered using pattern discovery tools are rules to link function with the primary sequence of a peptide. We would hypothesize that any sequence which has significant homology to these grammars would be likely to be antimicrobial. Therefore, a wide variety of approaches could be used to design novel peptide sequences based on these grammars. An appealing strategy for reassembling the grammars into new designs would have the following properties:

1. **Sequences highly covered by grammars.** Given the hypothesis that the grammars are associated with antimicrobial activity, a sequence which has a greater proportion of its length covered by antimicrobial grammars will more likely exhibit antimicrobial activity. Further, the likelihood for antimicrobial activity is expected to increase if at least portions of the designed sequences are homologous to multiple grammars.
2. **Sequences diverse relative to each other.** Because the set of sequences that have homology to antimicrobial grammars is far too large to synthesize exhaustively, it would be most efficient to sample a very diverse subset of sequences from the grammar space. It may be the case that sequences from similar regions of grammar space have similar activity profiles or toxicity characteristics. In order to maximize the probability of finding a sequence with a desirable activity spectrum early in the screen, diverse sequences should be selected for validation.

3. **Sequences diverse relative to natural AmPs.** A primary reason to use a grammar-based approach to design is that it enables the design of sequences which are only partially homologous to natural AmPs. Well established techniques have been used to optimize known AmPs through truncations or single mutations. The uniqueness and power of a grammar based approach is best exploited through creating diverse sequences. Furthermore, if therapeutics are developed that are highly homologous to natural AmPs, they will be more likely to lead to cross resistance in natural peptides.

#### 4.1.3.1 *In silico* evolution

For initial studies in AmP design, an *in silico* evolution process was used to generate designed sequences. This is a three step approach which involved the following:

1. **Random sequence generation.** On the order of one million random amino acid sequences were generated on a computer in order to sample a broad section of sequence space. While this is only a small portion of all available sequences, it may be large enough to generate a subset of sequences with moderate homology to antimicrobial grammars.
2. **Scoring of sequences.** The randomly generated sequences were then scored based on their degree of homology with antimicrobial grammars. Given that the random sequences are likely to have only limited homology to the grammars, a scoring metric suitable for weak

homology was devised. In this metric, the fraction of the random sequence that are homologous to any grammar is determined through an exhaustive homology search between all designs and grammars. This scoring metric has a maximum score of 1.0 and will not differentiate between sequences entirely covered by grammars based on the degree of overlap of the grammars.

3. **Mutate high-scoring sequences**. A variety of methods can then be used to create additional diversity from high-scoring sequences. One suitable strategy is to perform single or multiple mutations, which may be conservative or non-conservative. Alternatively, high scoring sequences may be allowed to breed, meaning that portions of one sequence is truncated and appended to a portion of another sequence to further enhance diversity. Steps 2 and 3 of this process may be used in an iterative process in which the minimum score for a sequence retained in the mutation step is increased between each round of evolution. This will create a rich set of highly scoring sequences.

A diagram of this *in silico* is shown in Figure 4.6.

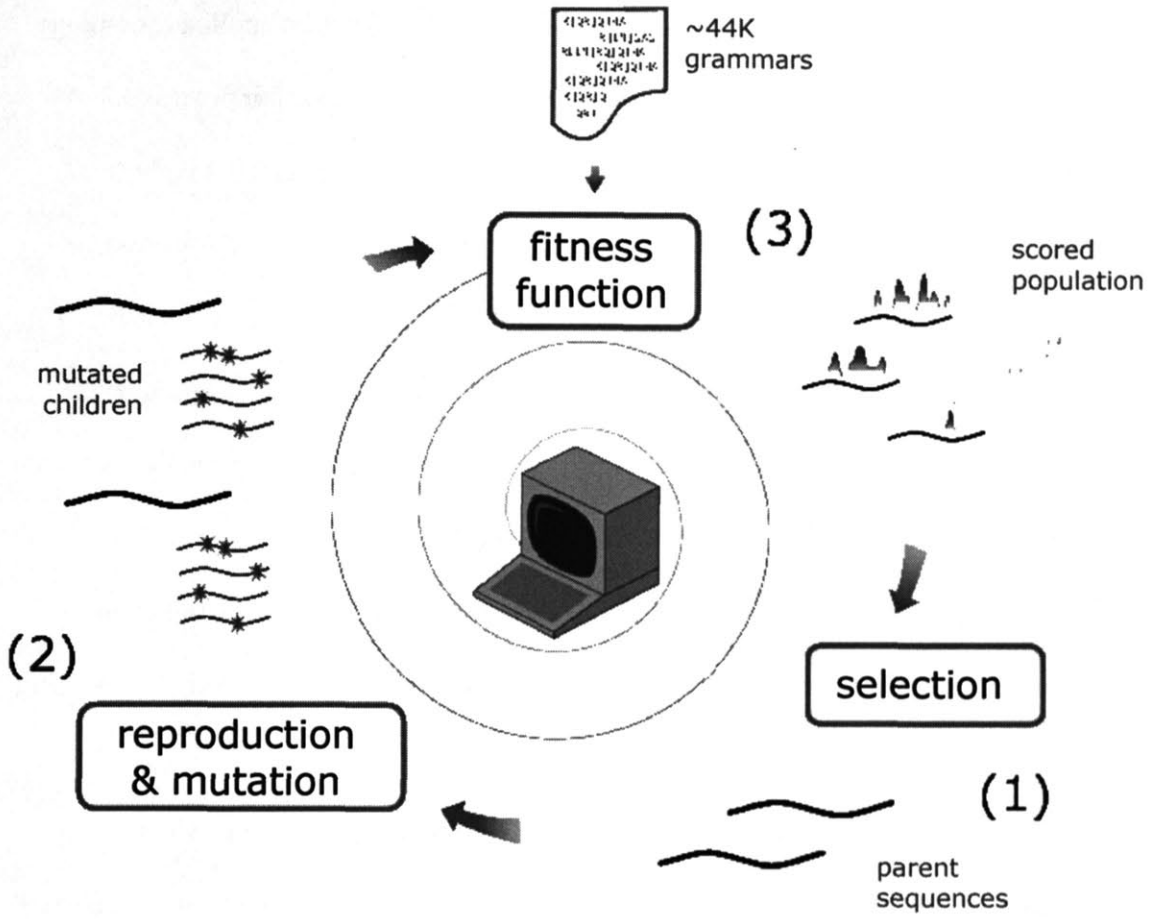



**Figure 4.6: *In silico* evolution procedure for designing AmPs.** A random set of sequences are generated *in silico*. Sequences are allowed to recombine with each other and are mutated. Sequences are evaluated using a fitness function, which scores sequences based on the fraction of the peptide that is covered by antimicrobial motifs. High-scoring sequences are retained and iteratively allowed to recombine and mutate to produce sequences highly covered by motifs. Figure reproduced from KL Jensen.

An *in silico* evolution approach is attractive because it mimics the natural process by which AmPs are designed. This assumes that the fitness function applied (homology to antimicrobial grammars) is a reasonable proxy for a natural selection process for antimicrobial defense. Additionally, the starting set of sequence is created without any user bias, increasing the diversity of sequences that go through the scoring process. Furthermore, because the process starts with a completely random set of sequences, this process may be rerun indefinitely to continue generating novel leads for experimental characterization. This could be used in a design scheme in which sequences that are experimentally validated are added to a modified training set, and grammars are rediscovered. The quality of the grammars, and thus the scoring metric, would continue to improve after each iteration, improving the efficiency of the *in silico* design at creating active sequences.

Nevertheless, there are drawbacks with an evolutionary approach. Most importantly, the designed set is highly dependent on an initial random set. If only a small portion of the random set has any homology to the grammars, a small set of sequences may come to dominate the evolution. This may lead to a final set of sequences that are not highly diverse relative to each other. The breadth of the initial training set is limited by the computational expense of the exhaustive scoring and mutation process. Additionally, the process is an inefficient way to develop high scoring sequences. If parameters are set to retain diversity, the degree of antimicrobial coverage through each round of evolution increases slowly. Further, a great deal of computational effort is spent scoring sequences having little homology to the grammars. A more efficient approach to

sequence design would incorporate knowledge of the grammars into the starting set of candidates.

#### 4.1.3.2 Combinatorial enumeration

As described above, an *in silico* evolution approach has a number of shortcomings, including low efficiency at generating high-scoring sequences, and limited diversity in the set of peptides designed. We sought a more rigorous and efficient approach to avoid these problems. In order to accomplish this, the information contained in the grammar set was used as a starting place for designing new sequences. To this end, we combinatorially enumerated all sequences of a desired length that were entirely covered by grammars. This addresses the 3 design criteria stated in Section 4.1.3 by accomplishing the following:

1. **Sequences highly covered by grammars**. Inherently, all sequences we created will be homologous over their entire length to at least one grammar. This means that all designs would score a 1.0 using the fractional coverage scoring metric developed for *in silico* evolution. A modified scoring metric will be developed for highly covered sequences and described below.
2. **Sequences diverse relative to each other**. While there will be less diversity in the starting set than when using *in silico* evolution, enumerating all heavily grammatical sequences will give much larger set of high-scoring sequences. Sequence diversity may then be

enforced by using clustering to select unique members from this high-scoring set.

3. **Sequences diverse relative to natural AmPs**. As in the second criteria, using combinatorial enumeration will create a rich set of high-scoring sequences. Thereafter, a strict filter may be applied to remove sequences that are too similar to natural peptides, and a large, highly grammatical set will be left from which to select candidates for characterization.

To design unnatural AmPs, we combinatorially enumerated all grammatical sequences of length 20. A length of 20 amino acids was chosen because peptides of this length may easily be chemically synthesized for validation, and it is near the median length of natural linear AmPs. While it is advantageous to have shorter peptides for therapeutic applications to reduce the cost of synthesis, peptides much below 12 amino acids have reduced selectivity, and 20 amino acids gives a reasonable scale on which to reassemble 10-amino acid motifs. An alternative approach using motifs shorter than 10-amino acids would likely have less substantially justified motifs with some of the short motifs based on chance.

First, for each grammar, we wrote out all possible grammatical amino acid sequences. More than one 10-amino acid sequence is generated from each motif given that wild-cards and bracketed expression allow multiple choices for amino acids at each position. As an example, the grammar [IVL]K[TEGDK]V[GA]K[AELNH][VA][GA]K has multiple acceptable amino acids at 6 positions, leading to 3*5*2*5*2*2 = 600 unique sequences. There are roughly 3 million such 10-mers that correspond to antimicrobial

patterns. Then we wrote out all possible 20 amino acid sequences for which each window of 10 amino acids is found in the set of 3 million 10-mers. From this set, we removed any 20-mers that had six or more amino acids in a row in common with a naturally occurring AmP. There are roughly 12 million such 20-mers, each of which is a "tiling" of ten 10-mers. An example of a 20-mer entirely covered by 10 amino acid grammars is shown in Figure 4.7.

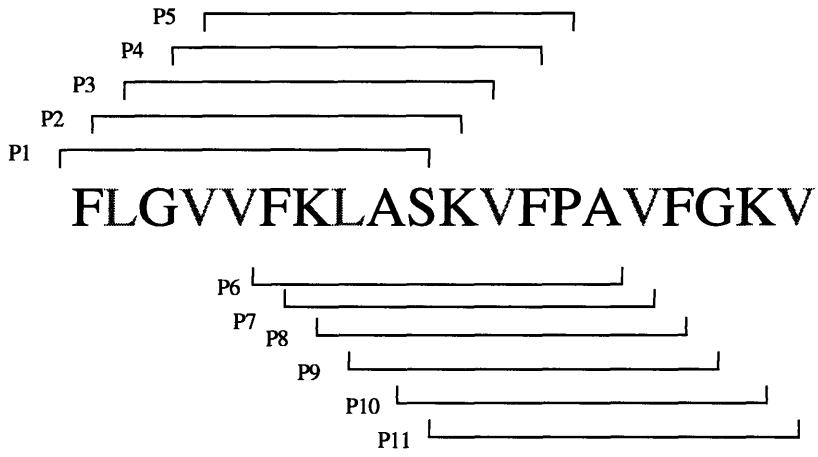



**Figure 4.7: Example designed sequence covered by antimicrobial motifs.** A 20-amino acid designed sequence (D28) that is covered over each 10-amino acid window by an antimicrobial pattern, P1-P11 .

#### 4.1.3.3 Clustering and scoring of designs

In order to select designed AmPs to synthesize, a scoring metric is required. As discussed in the text, if one of the AmP grammars matches a particular 10 residue stretch of a 20-mer, we would call that 10 residue stretch "grammatical." Here, the 20-mers were designed such that each window of 10 amino acids is grammatical. Obviously though, some grammars are more common in the set of known AmPs than others. So, rather than using a binary metric --- grammatical or not --- we developed a score *S* which is the degree to which a given 20-mer is grammatical. This score is computed by making a sequence dot plot matrix (Maizel and Lenk 1981). In the dot plot, the columns represent the positions, 1-20, of the 20-mer and the rows represent the concatenated sequences of the ~1000 naturally occurring AmPs. A dot is placed in the matrix wherever a grammar matches both a naturally occurring AmP and the 20-mer. Then score *S* is then just the number of dots in the matrix.

First, we desired to create sequences that were diverse relative to natural AmPs. To accomplish this, we removed any 20-mers that had six or more amino acids in a row in common with a naturally occurring AmP. Next, we desired to ensure that our designs were diverse relative to each other. To this end, we clustered these 12 million sequences using the Mcd-hit software at 70% identity (Li, Jaroszewski et al. 2001). From these clusters, we chose 42 high scoring sequences to test experimentally. This process is illustrated in Figure 4.8.

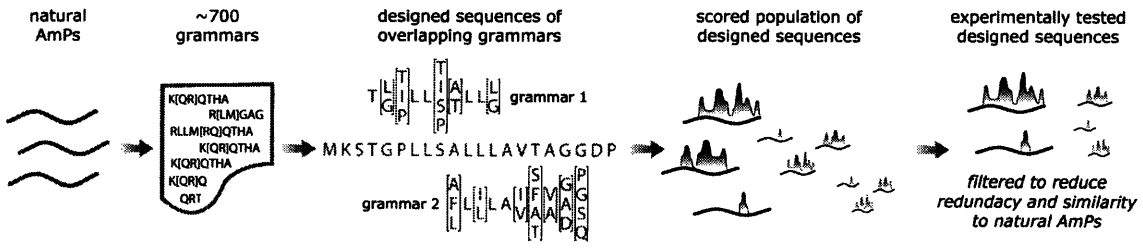



**Figure 4.8: A schematic of the *in silico* peptide design strategy.** Grammars are induced from the set of natural AmP sequences using Teiresias. Overlapping grammars are stitched together to create novel 20 amino acid sequences that correspond to the antimicrobial syntax. Designed sequences are scored based on the prevalence of the underlying grammar in natural peptides. After removing designs with 6 or more consecutive amino acids in common with a natural peptide, a representative set of sequences is selected for experimental validation.

### 4.1.3.4 Designing negative controls

We sought to demonstrate that antimicrobial activity of designed sequences would result from the grammars contained, rather than bulk physiochemical properties of the peptide. While a reasonable negative control would be to use randomly generated peptide sequences, a more rigorous approach would match the properties of the control to the designed sequences with the exception of the presence of grammars.

To this end, for each of the 42 synthetic peptides, we also designed a shuffled sequence, in which the order of amino acids was rearranged randomly such that the sequence did not match any grammars. These shuffled peptides had the same amino acid composition as their synthetic counterparts and thus, the same molecular weight, charge, and pI: bulk physiochemical factors often correlated with antimicrobial activity. We hypothesized that because the shuffled sequences were "ungrammatical" they would have no antimicrobial activity, despite having the same bulk physiochemical characteristics. In addition, we selected 8 peptides from the APD as positive controls and 6 20-mers from non-antimicrobial proteins as negative controls.

It should be noted that one would expect that a portion of the shuffled controls would have antimicrobial activity. Given the appropriate bulk physical properties, it is reasonable to assume a random shuffling has the potential to produce new active sequences even though it does not match grammars derived from AmPs found in nature, which is a small set compared to the set of all sequences which have antimicrobial activity. In fact, shuffling a strong natural AmP, cecropin-melittin hybrid, produced a series of peptide with weaker, but measurable activity. Nevertheless, we hypothesize that

non-grammatical sequences should be active less frequently than similar grammatical sequences.

## 4.2 Experimental validation

After computational tools were used to generate 100 peptide sequences, including positive and negative controls, a cost-effective and reliable method to synthesize and assay the sequences was required. The methods must be appropriate for 20 amino acid peptides and ideally should be automatable so that a broader set of predictions could be validated in future work. First, methods of peptide synthesis will be discussed, followed by methods for evaluating antimicrobial activity.

### *4.2.1 Peptide synthesis*

A variety of methods are possible for the synthesis of peptides. These methods vary in cost, the amount and purity of peptide produced, and the ease to which they can be adapted to a diversity of sequences. In the following sections the advantages and disadvantages of recombinant production, *in vitro* translation, and chemical synthesis will be discussed.

#### 4.2.1.1 *Recombinant production*

One method used for the creation of peptides is through recombinant synthesis. Nevertheless, this approach has a series of undesirable properties for validating AmP activity. First, the produced AmPs may be toxic to the host cells. Additionally, the grammar based designs are greatly diverse, so there is not a cost-effective approach to produce plasmids encoding each of these custom designed peptides. Further, in order to create sufficient amounts of peptide for characterization against a series of bacteria, fermentations and purifications would be required. For these reasons, recombinant production for validating our predictions was infeasible. Recombinantly produced peptides for well-studied AmPs were purchased and used as standards in this work.

#### 4.2.1.2 *In vitro* translation

As discussed in Chapter 3, *in vitro* translation avoids some of the shortcomings of recombinant production. There is no concern over toxicity to host cells and cells do not have do be lysed and membranes removed before assaying the peptides. Through the series of improvements in fusion partner optimization, the cost per peptide to synthesize sufficient material to test against 2 bacteria at up to 16 ug/ml was reduced from around $800 to $40, making this approach economically feasible. Additionally the synthesis and purification was demonstrated entirely in 96-well plates, indicating the process could be automated.

There are, however, drawbacks associated with an *in vitro* translation platform. First, the fact that the *in vitro* translation mixture impacts MIC measurements necessitates a series of purification steps. While these are, in principle, automatable, the robotic systems required are not available for our use, so they steps must be performed manually. Additionally, while a combinatorial approach for producing oligos coding for grammatical antimicrobial peptides was described, using this approach limits the diversity of sequences that may be explored because it requires reuse of grammatical subunits. There is not a cost-effective methods for producing oligos for completely unique sequences that could emerge from some design approaches. Finally, the concentrations produced are only sufficient to assay up to 16 ug/ml, so only highly active peptides would be found. It would be advantageous to have information on peptides which have moderate activity at 256 ug/ml or below, because one could have a fuller picture of the landscape of activity. There are classes of peptide for which *in vitro* translation would be the superior method. For instance sequences which are prohibitively long to be chemically synthesized and are toxic to host cells would be ideal for *in vitro* production. Nevertheless, for linear 20-mers, advances in automated small-scale chemical synthesis were made during the time of this study so that it became a feasible approach to validate our predictions.

#### 4.2.1.3 Chemical synthesis

Chemical synthesis is carried out by adding amino acids one at a time through a series of chemical reactions to produce a desired sequence on polymer resin beads. In

order to add exactly one amino acid to each peptide being elongated at each synthesis steps, blocked peptides are used (Fields and Noble 1990). For example, these peptides may have an Fmoc group on the N-terminal of each amino acid being added. After a single amino acid is added, the free amino acids are washed away, and the end of the recently added amino acid is deprotected through the addition of 20% piperidine in DMF. After a series of washes, the next amino acid is added to the reaction, which attaches to the free end of the peptide chain which has been deprotected. Additionally, the side chains of amine or acid containing amino acids are protected with a separate blocking agent that remains attached throughout the entire peptide synthesis. After the entire peptide has been synthesized stepwise, all deprotecting groups are removed by trifluoroacetic acid and the peptide is cleaved from the resin beads. A diagram of the procedure is given in Figure 4.9.

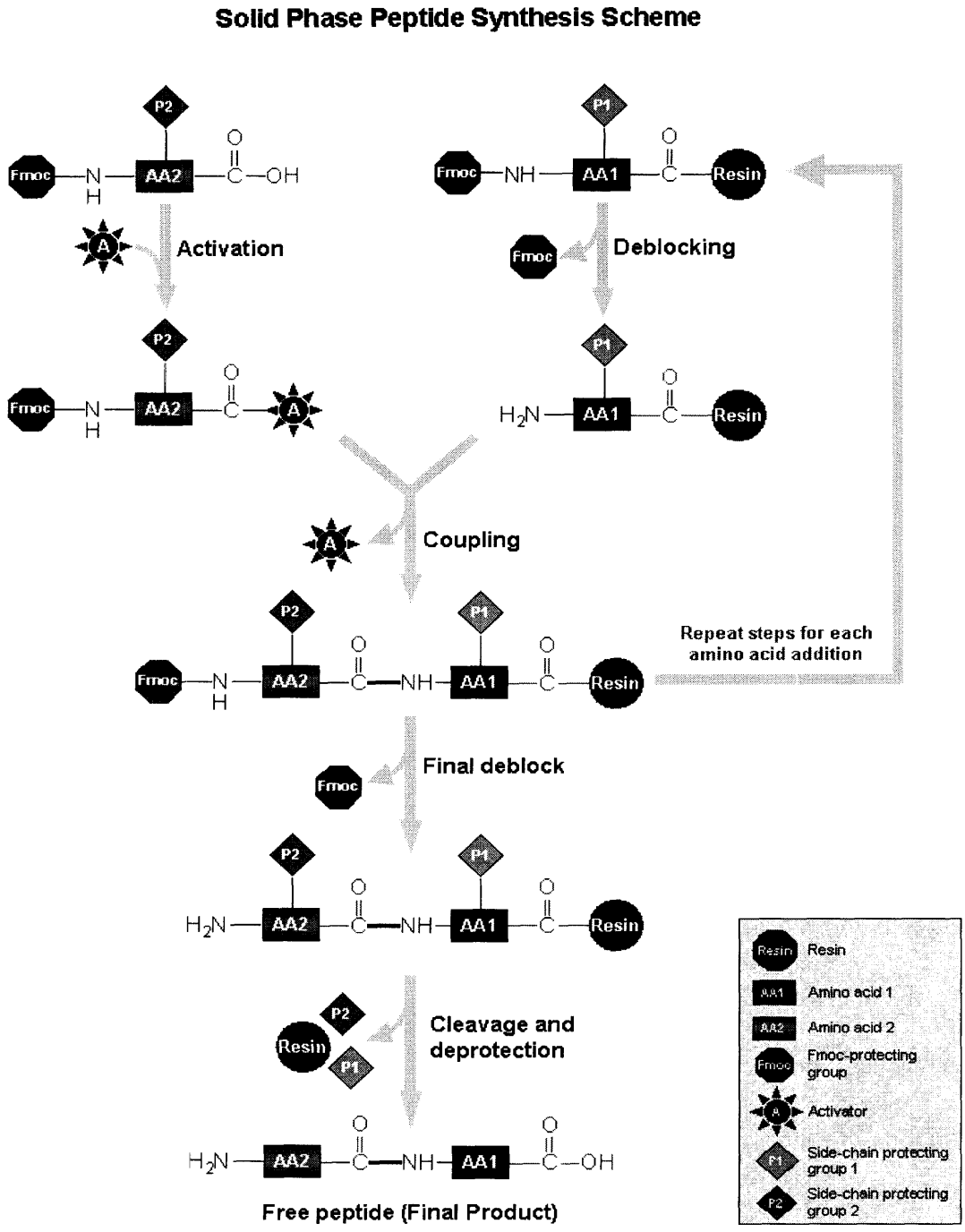



**Figure 4.9: Schematic for Fmoc synthesis of peptides**. Peptides are synthesized one amino acid at with alternating coupling and deprotection steps. Figure reproduced from http://www.sigmaaldrich.com/Brands/Sigma_Genosys/Custom_Peptides/Key_Resources/Solid_Phase_Synthesis.html.

While, at the outset of this work, custom synthesis of 20-mers cost in excess of $500, costs fell to tens of dollars per peptide later in my work. This reduction was enabled by advances in producing robotics to synthesize peptides at a small scale in sufficient purity that they could be used without HPLC purification. Chemical synthesis possesses a number of important attributes that made it ideal for validating our predictions. First, the system is automated and can produce completely independent sequences, allowing the evaluation of a diverse set of peptides. Additionally, the yield of small-scale chemical synthesis is dramatically higher than *in vitro* translation, which enables the testing of each peptide against multiple bacteria at up to 256 ug/ml, rather than just 16 ug/ml. Finally, chemical synthesis allows the incorporation of unnatural amino acids and chemical modification of the termini of the peptides.

For these reasons, Fmoc chemistry was used to synthesize peptides for validating our grammar based designs by using the Intavis Multipep Synthesizer (Intavis LLC, San Marcos, CA) at the MIT Biopolymers Lab. Mass spectrometry was used to confirm the accuracy of the synthesis and typical purities obtained with the synthesizer were >85%. Mass spectroscopy for a representative set of peptides is shown in Table 4.1. Recombinantly produced standards for Cecropin P1, Cecropin Melittin Hybrid, Melittin, Magainin 2, and Parasin were purchased from the American Peptide Company (Sunnyvale, CA). In antimicrobial assays, four of the five recombinant peptides had identical MICs to the chemically synthesized versions from MIT biopolymers, with the last being one dilution different (Cecropin P1). Also, one designed peptide was synthesized by MIT biopolymers 4 separate times and the resulting peptides had consistent activities against both *E. coli* and *B. cereus*.

| Peptide | Sequence | Molecular Weight (theoretical) | Molecular Weight (MALDI-TOF) |
|---|---|---|---|
| D28 | FLGVVFKLASKVFPAVFGKV | 2153.9 | 2154.0 |
| S28 | GVSVAGAKKVKVLFVFPFLF | 2153.9 | 2153.6 |
| D23 | ALGKLASKVFPAVYCTISRK | 2152.9 | 2153.1 |
| S23 | KYGPALVIAVKKSCSLTFRA | 2152.9 | 2151.8 |
| Magainin II | GIGKFLHSAKKFGKAFVGEIMNS | 2467.4 | 2467.1 |
| Parasin | KGRGKQGGKVRAKAKTRSS | 2000.7 | 1999.4 |

**Table 4.1: Molecular weight of selected peptides.** Theoretical molecular weights of peptides, and measured molecular weights using MALDI-TOF.

### *4.2.2 Antimicrobial assays*

Researchers have used a host of different assays to evaluate antimicrobial activity. While some assays measure the concentration at which AmPs kill bacteria (the minimum bactericidal concentration, or MBC), other measure the minimum concentration at which AmPs inhibit growth (the minimal inhibitory concentration, or MIC). The assays vary greatly in the ease with which they may adapted to high-throughput evaluation and the amount of labor required to test each peptide. While some assays are quantitative, others are qualitative. Additionally, only some of the methods provide information on the kinetics of the killing or inhibitory activity. Finally, the frequency with which each assay is used and the degree to which clinicians believe the results of different assays varies. Commonly used assays include plate counting, fluorescence, and microdilution methods described below.

#### 4.2.2.1 Plate counting

As described in the National Center for Clinical Laboratory Standards (NCCLS) guideline M-26A, the gold standard for evaluating the antibiotic activity of an agent is through a plate counting technique. In this assay, the agent tested is incubated at desired concentrations with a target bacteria at a given concentration (usually 5 x 10^5 cfu/ml) in Cation-adjusted Mueller Hinton Broth (CMHB) for 16-20 hours. The optimal temperatures for incubation varies by microbe (e.g. 37°C for *E. coli* and 30°C for *B. cereus*). After that time, a sample of the mixture is diluted in Phosphate Buffered Saline (PBS) to create a series of 10-fold dilutions. A 50 ul sample of each of these dilutions is spread on a standard 150 mm agar plate and grown overnight at the optimal temperature, and then enumerated. In order to obtain accurate counts, the dilution should be selected which gives between 50 and 500 colonies per plate, and at least replicate plating should be performed. An extension of this assay is a time-kill curve, in which dilution and plating is carried out at specified times after the start of incubation of the AmPs and bacteria in the 96-well plate. Time-kill curves demonstrate the rate at which the count of viable bacteria decreases.

The advantages of this assay are that it is quantitative, and gives a measurement of MBC as well as MIC. The major disadvantage of this method is that is ill-suited for a screening assay because of the enormous effort that is required for each agent tested. If it is desired to test 8 concentrations of an agent, with 5 dilutions required for each concentration (since it is unknown what the order of magnitude of the resulting counts will be), with triplicate plating, 120 plates must be created and counted for a single

peptide. While automated methods are available for creating dilutions and even imaging and plate counting, the spreading of plates is a significant bottleneck. Additionally, if a number of agents are being tested together, the length of time required to make dilutions and plate would likely introduce inconsistencies between the first sample evaluated and the last sample evaluated. For these reasons, the plate counting method will only be used as a validation tool to study leads that emerge from other screening methods.

#### 4.2.2.2 Fluorescent reporting

The plate counting technique to find MICs would be prohibitively slow if hundreds of peptides are to be tested. An alternative method is to use bioluminescence, a technique developed in recent years to study microbial viability. A gene coding for luciferase may be cloned into each target pathogen. Luciferase genes may be obtained from insect and bacterial sources (Simon, Fremaux et al. 2001; Vesterlund, Paltta et al. 2004). Bacterial luciferases tend to have water-insoluble or volatile substrates and are difficult to express in many hosts. Insect luciferases, on the other hand, catalyze the breakdown of D-luciferin, which is available as a soluble sodium salt (Lehtinen, Virta et al. 2003; Sakakibara, Murakami et al. 2003). Further, D-luciferin passes through cell membranes, allowing fluorescence within intact cells. Finally, insect luciferases are easily expressed in prokaryotes and mammalian cells. In the presence of ATP, insect luciferase catalyzes the breakdown of D-luciferin in the presence of ATP and releases a fluorescent signal, as shown in the following reaction:

$O_2$ + D-luciferin + ATP => ADP + PP + oxyluciferin + H2O + light

Fluorescence is only observed in cells that are producing ATP, so the fluorescence of a population of bacteria is proportional to the fraction of the cells that are viable.

In order to use this assay, a luciferase encoding plasmid must be incorporated either as a plasmid, as in (Lehtinen, Virta et al. 2003), or more stably into the genome of the target bacteria (Hilpert, Volkmer-Engert et al. 2005). We obtained *E. coli* and *B. subtilis* containing an existing plasmid (pCSS962) coding for luciferase, and appropriate plasmids for other microbes may be made. Antibiotics are used to select only those microbes that express this plasmid. For this plasmid, both chloramphenicol and kanamycin resistance genes are used for selection. An antimicrobial assay may be carried out by incubating the AmP and microbe for 16-20 hours at the desired incubation temperature. At the end of incubation 500 uM D-luciferin is added at a 1:1 dilution and a fluorescence plate reader is used to measure relative fluorescence by exciting at 485 nm and reading emission at 530 nm in order to quantify the bacteria. Known dilutions of viable bacteria may be used as a calibration curve for fluorescence versus viable bacterial count. Plate counting may confirm the accuracy of the counts.

One concern with the use of a luciferase plasmid within target bacteria is that an antibiotic may be necessary to ensure plasmid expression. If an antibiotic is required, it may complicate the analysis of an AmP in a screen because the additive effect of the AmP and antibiotic is measured. This may be particularly misleading for AmPs, which are known to permeabilize membranes and may enhance the potency of the other

antibiotic. An AmP which strongly encourages penetration of the other antibiotic without any killing effect on its own would be mistakenly labeled a potent candidate. While some error is acceptable for a screen since leads will be validated, clinicians may be suspicious of this approach.

Fortunately, for this assay, an antibiotic may not be needed if the copy number of the plasmid is sufficient that it is largely retained during the course of the assay. Cells containing the plasmid could be grown to a high concentration in a selective medium with the antibiotic, and then diluted to a fresh medium without the antibiotic for the AmP assay. In order to test if the plasmid was retained, the specific fluorescence among a population of viable cells was measured. First, the plasmid containing bacteria was grown overnight in the presence of kanamycin (10 ug/ml) and chloramphenicol (68 ug/ml). The culture was then diluted 500-fold in the presence of chloramphenicol, kanamycin, both, or neither and grown to an OD between 0.2 and 0.5. Dilutions of 3 and 10 fold were made, the sample was mixed 1:1 with 500 uM D-Luciferin, and relative fluorescence read on a plate reader. The resulting fluorescence versus OD is shown in Figure 4.10.

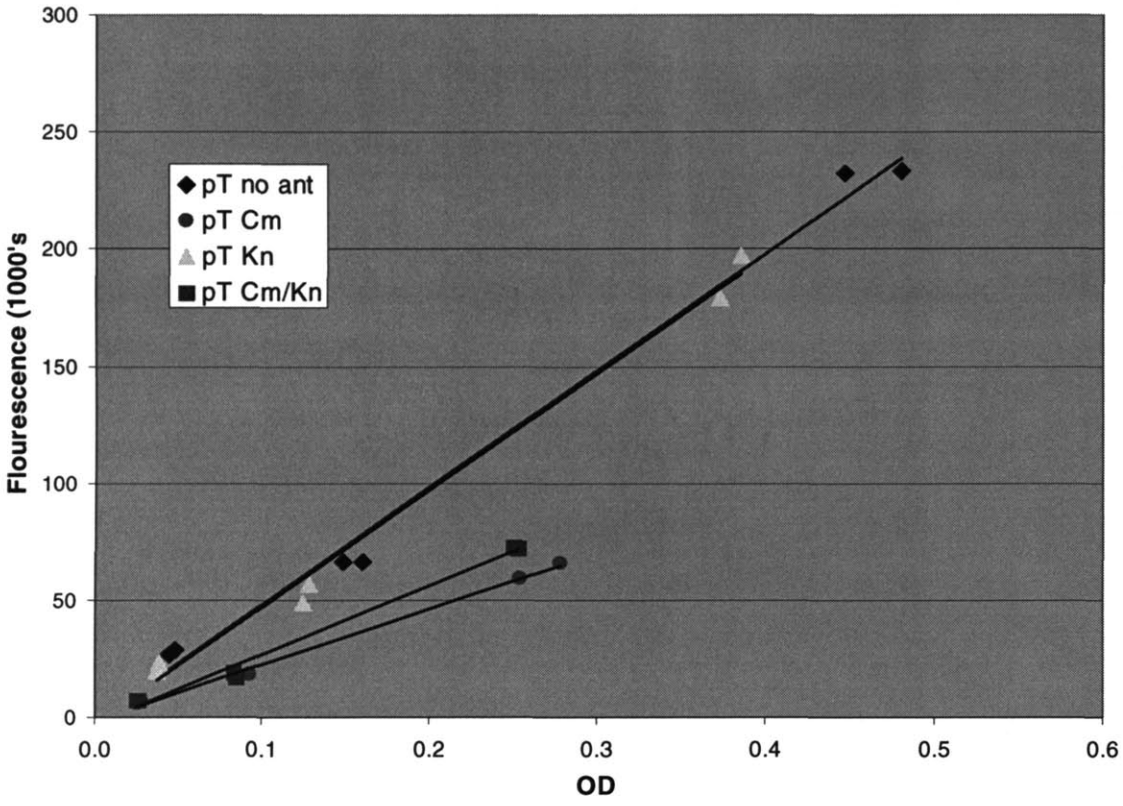



**Figure 4.10: Fluorescence versus cell density for luciferase containing *E. coli* with various antibiotics**. Viable *E. coli* was grown in the presence of Chloramphenicol (Cm), Kanamycin (Kn), both, or neither. Both cultures with Cm showed reduced fluorescent responses. For the duration of the assay, plasmid loss did not occur in the absence of antibiotic, with equivalently strong fluorescence for samples with and without Kn. Carrying out an antimicrobial assay without a second antibiotic would give clearer measurement of AmP activity.

As shown in Figure 4.10, no plasmid loss occurs over the course of the assay since the specific fluorescence is identical with and without kanamycin present. This would allow the AmP screen to be run without the presence of an extra antibiotic in the mixture, simplifying the interpretation of the screening results. It was also noted that chloramphenicol reduced the specific fluorescence of the plasmid containing bacteria.

A fluorescence assay is convenient because it may be run against multiple dilutions of an AmP in a 96 well plate. Furthermore, the fluorescence reading reflects the viable fraction of the cells over time, so both growth inhibition and killing kinetics may be monitored. Nevertheless, these fluorescence methods are not widely accepted as accurate measurements of antimicrobial activity within the clinical community. Additionally, the burden of the luciferase production has the potential to affect MIC measurements. It could also be possible that an AmP changes the specific fluorescence of the target bacteria, as chloramphenicol did. Finally, D-Luciferin is an expensive reagent required for all measurements.

An alternative method for quantifying a viable population is through the use of LIVE/DEAD staining (Ramji, Baig et al. 2005). These commercially available stains contain two dyes with selectively stain viable or non-viable cells. SYTO9 is a green dye that stains all cells, including those with intact membranes. Propidium Iodide (PI) is a red dye that may not penetrate intact cells, and thus stains the DNA only with disrupted cell membranes. The fraction of red cells to green cells gives the percentage viability and may be quantified using microscopy, or fluorescence plate readers. Additionally, flow cytometry may be applied to quantify live and dead populations (Veal, Deere et al. 2000; Budde and Rasch 2001). As with luciferase techniques, LIVE/DEAD quantification, is

not broadly accepted in the clinical community. Further, LIVE/DEAD staining really is an assay for membrane integrity, which is not equivalent to viability, especially for membrane permeabilizing agents such as AmPs.

#### 4.2.2.3 Microdilution assay

The most commonly accepted clinical method for screening antimicrobial agents is a microdilution growth inhibition assay. This assay is describe in NCCLS M-26A, and optimized for AmPs in the Hancock assay for cationic peptides (Wu and Hancock 1999). The growth medium can have significant impact on the resulting MICs, particularly for cationic peptides, and should be chosen to reflect the clinical environment (Schwab, Gilligan et al. 1999; Dorschner, Lopez-Garcia et al. 2006). Serial dilutions of peptides in 0.2% Bovine Serum Albumin and 0.01% Acetic acid were made at 10x the desired testing concentration. Target bacteria were grown in Mueller Hinton Broth (MHB, or cation-adjusted CMHB) to an OD600 of 0.1 to 0.3 and diluted down to 2-7 x10^5 cfu/ml in fresh MHB, as confirmed by plating serial dilutions. Five ul of the peptide dilutions was incubated with 45 ul of the target in sterile, capped, polypropylene strip tubes for 16-20 hrs. The MIC was defined as the minimum concentration that prevented growth based on visual inspection of OD. If verification of visual OD quantification is desired, dose response curves may be created by taking OD600 measurements on a 96 well plate reader after diluting 30 ul of the sample 1:3 in MHB. The 16-20 hour incubation may also be carried out in a 96-well plate.

A microdilution assay may be automated and does not require significant manual manipulation, making it ideal for a screen. Its results are widely accepted by clinicians, and any bacteria may be used. Finally, no additional antibiotics or dyes must be added which could affect bacterial metabolism or viability. When desired, samples from an overnight incubation may be streaked on MHB agar plates to assess the viability of the bacteria, and thus determine an MBC. The shortcoming of the microdilution method is that it does not measure time-kill kinetics. Nevertheless, microdilution is ideally suited for a screen, and leads from the screen may be further characterized for killing characterization. For these reasons, the Hancock assay for cationic peptides was used to evaluate the AmPs designed with pattern-based methods. For more in depth studies in complex environments such as biofilms, specialized assays may be applied to leads (Zelver, Hamilton et al. 1999). A diagram of the microdilution assay is given in Figure 4.11 and a sample result is shown in Figure 4.12.

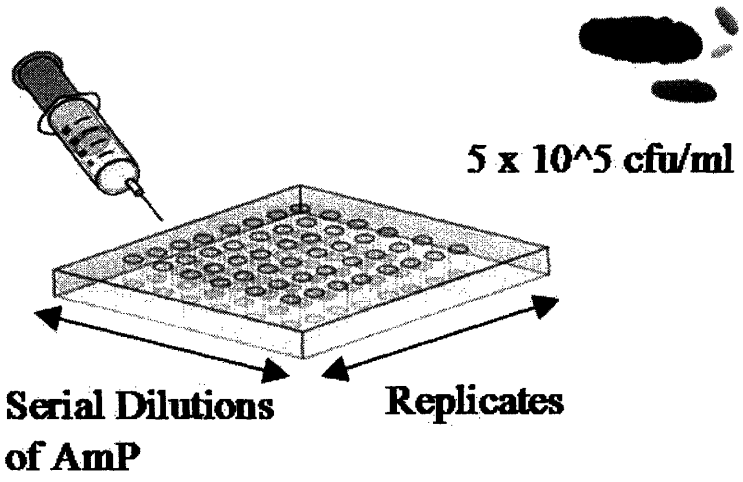



**Figure 4.11: Schematic of a microdilution assay**. Within a 96-well plate, all wells are inoculated with 5 x 10^5 cfu/ml target bacteria in Cation Adjusted Mueller Hinton Broth. A series of 2-fold dilutions of an AmP is added along one axis at 10x the desired testing concentration. Along the second axis, replicates of the AmP may be assayed, or different AmPs. Figure reproduced from JF Moxley.

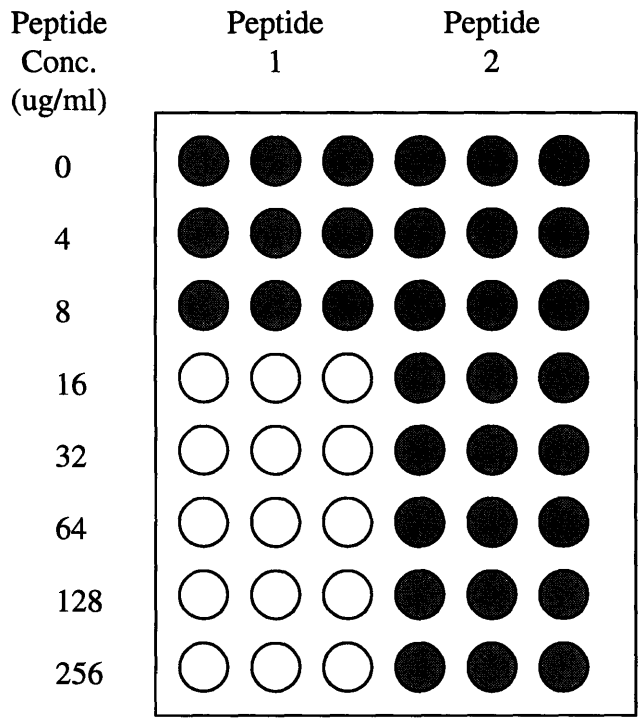



**Figure 4.12: Schematic of MIC results using a microdilution assay**. Following 16-20 hour incubation of peptide with target bacteria, growth is observed (filled) for all dilutions of peptide 2, whereas growth was inhibited at and above 16 ug/ml for peptide 1. The MIC is 16 ug/ml for peptide 1 and is greater than 256 ug/ml for peptide 2.

### 4.2.3 MICs of natural AmPs

We characterized the activity of each synthetic AmP using a broth microdilution assay described in Section 4.2.2.3. This assay measures the Minimum Inhibitory Concentration (MIC) at which the peptide inhibits growth of the target organism. It was desirable to determine the spectrum of activity for each peptide, so *Bacillus cereus* and *Escherichia coli* were used as representative gram-positive and gram-negative bacteria. First, a selection of natural AmPs were evaluated to ensure that our results were consistent with previous studies. A few strong, well-studied AmPs that are commercially available through recombinant production were studied, including cecropin P1, cecropin melittin, cecropin A magainin, magainin 2, and melittin. Additionally, a random selection of peptides from the AmP database were selected including parasin, ranelexin, and pyrrhocoricin. The MICs for each of these natural AmPs are shown in Table 4.2.

| Natural AmPs | *E. coli* MIC (ug/ml) | *B. cereus* MIC (ug/ml) |
|---|---|---|
| Cecropin P1 | 2 | >256 |
| Cecropin Melittin | 2 | 8 |
| Magainin 2 | 64 | 256 |
| Melittin | 16-32 | 8 |
| Parasin | >256 | >256 |
| Ranelexin | 64 | 32 |
| Cecropin A Magainin | 32 | >256 |
| Pyrrhocoricin | >128 | >128 |

**Table 4.2: Representative MICs for natural AmPs.** Powerful, well-studied, AmPs including cecropin, melittin, magainin, and combinations thereof were tested against *E. coli* and *B. cereus*. Some peptides, such as cecropin P1, are highly specific to Gram negative bacteria. Others, such as cecropin-melittin hybrid, are broad spectrum. A random selection of other natural peptides used in our training set including Parasin, Ranelexin, and Pyrrhocoricin revealed that some AmPs were not active against our target bacteria in the concentration range tested. This implies that designs drawn from these peptides may not necessarily have activity.

The activity of the natural peptides varies from narrow to broad spectrum. For instance, cecropin P1 is highly active against *E. coli*, with no measurable activity against *B. cereus*. Intriguingly, a number of the natural peptides such as parasin and pyrrhocoricin are not active against either bacteria in the concentration range tested (note that pyrrhocoricin was only tested to 128 ug/ml because of solubility limits). They may have antimicrobial activity against other bacteria not tested, or may have been documented as antimicrobial because they are evolutionarily related to other AmPs based on an upstream or pro-region that is cleaved. This indicates that even if we create a design tool that perfectly transmits antimicrobial information from the training set to a designed set, not all of the designs would be expected to have activity. These MICs were shared with Professor Robert Hancock who pioneered this antimicrobial assay and he confirmed that the results matched his expectations. Finally, the MICs matched those measured for recombinant standards for 4 of 5 natural AmPs for which recombinant versions are available, with the MIC for last peptide being only one dilution different. This indicates that the chemical synthesis is a reliable method for creating AmPs for evaluation in this length range.

### *4.2.4 MICs of control sequences*

We characterized the activity of each synthetic AmP using a broth microdilution assay described in Section 4.2.2.3 against *B. cereus* and *E. coli*. Of the negative controls - 6 peptides randomly selected from the middle of non-antimicrobial proteins from Swiss-Prot/TrEMBL - none had activity. As shown in Table 4.3, six of the eight naturally-

occurring AmPs in the positive control group show activity in the concentration range tested against at least one of the bacteria. Even if our design tool were to perfectly translate activity information from the training set to the designed set, activity in all sequences would not be achieved.

NATURAL AMPs

| Peptide | Sequence | E. coli | B. cereus |
|---|---|---|---|
| Cecropin P1 | SWLSKTAKKLENSAKKRISEGIAIAIQGGPR | 2 | + |
| Cec Mel Hyb | KWKLFKKIGAVLKVL-NH2 | 2 | 8 |
| Magainin 2 | GIGKFLHSAKKFGKAFVGEIMNS | 64 | 256 |
| Melittin | GIGAVLKVLTTGLPALISWIKRKRQQ-NH2 | 16 | 8 |
| Parasin | KGRGKQGGKVRAKAKTRSS | + | + |
| Ranalexin | FLGGLIKIVPAMICAVTKKC | 64 | 32 |
| Cec A Mag 2 | KWKLFKKIGIGKFLHSAKKF | 32 | + |
| Pyrrhocoricin | VDKGSYLPRPTPPRPIYNRN | ++ | ++ |

NON-AMP SEQUENCES

| Peptide number | Sequence | E. coli | B. cereus |
|---|---|---|---|
| 10 | THIHMNARLLIRSPFTDPQL | + | + |
| 19 | RKRKSDVDFEAEFELFEDDD | + | + |
| 38 | AATGTGKTAAFALPVLERLI | + | + |
| 50 | PYSLKNGENWLLSEEIIRYP | + | + |
| 74 | KFDPLEGAPMARGIVLEKVG | + | + |
| 86 | HTGRSGPATGHSGHSSTHGS | + | + |

+ = MIC greater than 256 ug/mL
++ = MIC greater than 128 ug/mL, not sufficiently soluble to test at 256 ug/mL

**Table 4.3: Minimum inhibitory concentrations for natural peptides.** Strong, natural AmPs including cecropin, melittin, magainin, and combinations thereof have antimicrobial activity that is either broad spectrum or specific. Randomly selected AmPs including parasin, ranelexin, and pyrrhocoricin had weaker or no activity against these target bacteria. This indicates that even if the design tool perfectly translates the activity distribution from the natural set of peptides, not all designs will be active. Randomly selected sequences from non-antimicrobial proteins did not display antimicrobial activity.

### 4.2.5 MICs of designed sequences

Of 42 designed and 42 shuffled peptides, 2 of the designed were insoluble. Of the 40 soluble designed peptides, 18 had an MIC of 256 ug/ml or less against at least one of the bacterial targets. This represents a design success rate of nearly 50 %. These results were repeatable with variation of one dilution or less, which is regarded as the standard error for MIC measurements. While more sequences are active against gram-negative bacteria, many of the leads are broad spectrum.

Not only was the fraction of design peptides that were active significant, highly active leads emerged from this screen of just 42 peptides. Two of the designed peptides, D28 (FLGVVFKLASKVFPAVFGKV) and D51 (FLFRVASKVFPALIGKFKKK), inhibited *B. cereus* growth at 16 ug/ml, which is close to the MICs of the strong positive controls melittin and cecropin-melittin hybrid (8 ug/ml).

DESIGNED PEPTIDES

| Peptide number | Sequence | E. coli | B. cereus |
|---|---|---|---|
| 1 | ALFSLASKVVPSVFSMVTKK | + | + |
| 2 | VVFRVASKVFPAVYCTVSKK | 128 | + |
| 5 | FLFGLASKVFPAVYCKVTRK | 64 | 256 |
| 6 | LSAVGKIASKVVPSVIGAFK | + | + |
| 7 | PVIGKLASKVVPSVFSMIKR | + | + |
| 9 | GLMSLVKDIAKLAAKQGAKQ | 256 | + |
| 15 | SALGRVASKVFPAVYCSITK | + | + |
| 22 | LGALFRVASKVFPAVISMVK | 256 | 64 |
| 23 | ALGKLASKVFPAVYCTISRK | 128 | + |
| 24 | GFIGKLASKVVPSVYCKVTG | 128 | + |
| 25 | PVVFSVASKVVPSLISALKR | + | + |
| 28 | FLGVVFKLASKVFPAVFGKV | 64 | 16 |
| 29 | PAVFKIASKVVPSVYCKVSR | 128 | + |
| 30 | GALFGLASKVFPAVFGAFKK | 256 | + |
| 31 | SAVGKLASKVFPAVFSMVTK | + | + |
| 33 | VKDLAKFIAKTVAKQGGCYL | ++ | ++ |
| 34 | GVVGKLASKVVPSVFGSFTK | + | + |
| 35 | LPVVFRVASKVFPALISKLT | + | 256 |
| 36 | SAVGSVASKVVPSLISKVTK | + | + |
| 39 | MKSIAKFIAKTVAKQGAKQG | + | + |
| 42 | LPAVFKLASKVVPSVFGLVK | + | + |
| 43 | SFVFKLASKVVPSVFSALTR | 256 | 256 |
| 44 | SVIGKIASKVVPSVYCAISK | + | + |
| 45 | PVVGRVASKVFPAVIGLVKK | + | + |
| 51 | FLFRVASKVFPALIGKFKKK | 64 | 16 |
| 55 | LSFVGRVASKVVPSLISMIK | 256 | + |
| 56 | SALGRLASKVVPAVIGKVTT | + | + |
| 57 | LGVVGSLASKVVPAVISKVK | + | + |
| 62 | LPAVFKLASKVFPAVYCKAS | 128 | + |
| 63 | LPVLFKLASKVFPAVFSSLK | 256 | 64 |
| 65 | VVGRVASKVVPSLIGLFTTK | + | + |
| 69 | SVVFGVASKVVPSVIGKVKT | + | + |
| 75 | FLPFVGRIASKVVPSVIGKV | + | + |
| 77 | GKKLAKTIAKEVAKQGAKFA | 64 | + |
| 81 | PFVGRVASKVVPSVYCAITR | Not soluble | |
| 82 | FVGSLASKVVPSVFGAIKTK | + | + |
| 83 | LPVVFKIASKVVPSVISKIT | + | + |
| 84 | GAVFGVASKVVPSVFSAIKK | + | + |
| 85 | FVGGVASKVVPSVYCKVSKK | + | + |
| 88 | VVFKLASKVVPSVYCTITKK | 256 | + |
| 96 | GALFSLASKVVPAVIGLIKK | + | 256 |

+ = MIC greater than 256 ug/mL
++ = MIC greater than 128 ug/mL, not sufficiently soluble to test at 256 ug/mL

**Table 4.4: Minimum inhibitory concentrations for designed peptides.** Of 42 designed peptides, 40 were soluble, and 18 had activity against either *B. cereus* and *E. coli* at 256 ug/ml or below.

### 4.2.6 MICs of shuffled sequences

As described in Section 4.1.3.4, a strict set of control peptides which had the same overall charge, pI, and molecular weight were designed. Table 4.5 shows that only 2 of the 38 soluble shuffled peptides displayed activity against *B. cereus* and *E. coli*. Therefore, grammatical sequences are 10-fold more likely to have activity than similar, non-grammatical sequences. This indicates that activity is not an artifact of bulk physiochemical properties. Table 4.6 shows a summary of activity for designed, shuffled, and control peptides against *B. cereus* and *E. coli*.

SHUFFLED PEPTIDES

| Peptide number | Sequence | E. coli | B. cereus |
|---|---|---|---|
| 1 | MVVFSVPKFKSTVAKLLSSA | + | + |
| 2 | TAKVVVFVSFSYVVPKKRAC | + | + |
| 5 | FLPVLVKVFRYSKKTAAGCF | ++ | 64 |
| 6 | GVSSPIVAVKFKGAVASLIK | + | + |
| 7 | SRVPLKSPVKIVGSKVMIFA | + | + |
| 9 | GLKKDALQSIVKKAQLAAMG | + | + |
| 15 | LYSPTCVKAAVSRFIGKVSA | + | + |
| 22 | SVPSVGAVLFFKRAAVMKLI | + | + |
| 23 | KYGPALVIAVKKSCSLTFRA | + | + |
| 24 | GGSTLGVFVKKSKACVIVPY | Not soluble | |
| 25 | KSPFVLVVSSRVAAVIKSLP | + | + |
| 28 | GVSVAGAKKVKVLFVFPFLF | + | + |
| 29 | KVYVVKIAVPCFPKSARSVS | + | + |
| 30 | KVVLFGAAGAKLFKASFFGP | Not enough material | |
| 31 | FMKVLAVFGSVVTSAPKASK | + | + |
| 33 | ALVYAGIKKTAFLKVQKCDG | + | + |
| 34 | SVKPVGSSVVKGTALVKFFG | + | + |
| 35 | KVFIATLVVSSFLLAKPPRV | + | + |
| 36 | STVKVASKLAVVVSPISKGS | + | + |
| 39 | AKKAQKSGAQTIVKIFAKGM | + | + |
| 42 | VVAKKFFVLVKGLAPVLSPS | + | + |
| 43 | ASPTVFRSSVFLSLFVVAKK | + | + |
| 44 | IASAVPVCVKGKISKSYISV | + | + |
| 45 | VKRAGKGVAVVPSPLFKIVV | + | + |
| 51 | RKVAPALIKSFVFLFKFKKG | + | + |
| 55 | SSSIPIKMVLVRALVFVKSG | + | + |
| 56 | TLVGVVAKLVATKIGSSPRA | + | + |
| 57 | PKVVGLSIVVVKAKVSSALG | + | + |
| 62 | PSLLYKAKAVFCKPSAVAVF | ++ | ++ |
| 63 | VSVKKVLPFAPLKSLLSFAF | 256 | 256 |
| 65 | FKVVISKPGLSVRVGTALVT | ++ | ++ |
| 69 | VFSVKGGKPSVVIKVVVAST | + | + |
| 75 | SKFPLAGIFSVPGVKRVVVI | + | + |
| 77 | VIAFAKTKEAKAKLKGQAKG | + | + |
| 81 | PAVYKSIVGFSPVARVTVCR | | |
| 82 | KTVPVVLKASIKVSSAGFGF | + | + |
| 83 | KIVKVITVKSISPASLVPVF | ++ | ++ |
| 84 | SVKVAKSVIPSAVFAGGKVF | + | + |
| 85 | KVGKGSYPCSFVKVVAKVSV | + | + |
| 88 | VKTKCSVPAVVYILVKTFKS | + | + |
| 96 | LPVLFSSAIAKVGIKLGAKV | + | + |

+ = MIC greater than 256 ug/mL
++ = MIC greater than 128 ug/mL, not sufficiently soluble to test at 256 ug/mL

**Table 4.5: Minimum inhibitory concentrations for shuffled peptides.** Of 42 designed peptides, 38 were soluble, and only 2 had activity against either *B. cereus* and *E. coli* at 256 ug/ml or below. The fraction of sequences that had activity is almost 10-fold lower than for the designed sequences.

| Class | *E.coli* (gram-negative) | | *B. cereus* (gram-positive) | | Either *E. coli* or *B. cereus* |
|---|---|---|---|---|---|
| | MIC ≤ 256 ug/ml | MIC ≤ 64 ug/ml | MIC ≤ 256 ug/ml | MIC ≤ 64 ug/ml | MIC ≤ 256 ug/ml |
| Designed | 16 / 40 | 4 / 40 | 8 / 40 | 4 / 40 | 18 / 40 |
| Shuffled | 1 / 38 | 0 / 38 | 2 / 38 | 1 / 38 | 2 / 38 |
| Natural AmPs | 6 / 8 | 6 / 8 | 4 / 8 | 3 / 8 | 6 / 8 |

**Table 4.6: Summary of activity of designed, shuffled, and natural AmPs.** Of peptides designed using antimicrobial motifs, nearly half were active against one of the two target bacteria at 256 ug/ml or lower. Shuffled peptides, with the same physiochemical properties, but no homology with antimicrobial motifs showed only 2 active sequences. Thus, the designed population is enriched by approximately 10-fold. Further, only 6 of 8 natural peptides were active in the range tested, indicating that even if our design tool perfectly translated activity from the training set, activity rates would not approach 100%.

### 4.2.7 Killing curves and bactericidal measurements

To validate MIC determinations, we measured optical density at varying concentrations of a representative set of designed, shuffled, and natural peptides. The microdilution assay described in Section 4.2.2.3 was carried out, and the degree of growth was quantified by measuring the OD600 in a 96-well plate reader, rather than by visual inspection. Figure 4.13 shows that peptides inhibit growth at the MICs determined by visual inspection. The MBC was measured by dipping sterile toothpicks in wells from an overnight incubation of the standard antimicrobial assay and streaking on MHB agar plates. Incubating these samples overnight indicated that there were generally no viable bacteria above the MIC, with the MBC within 2x of the MIC.

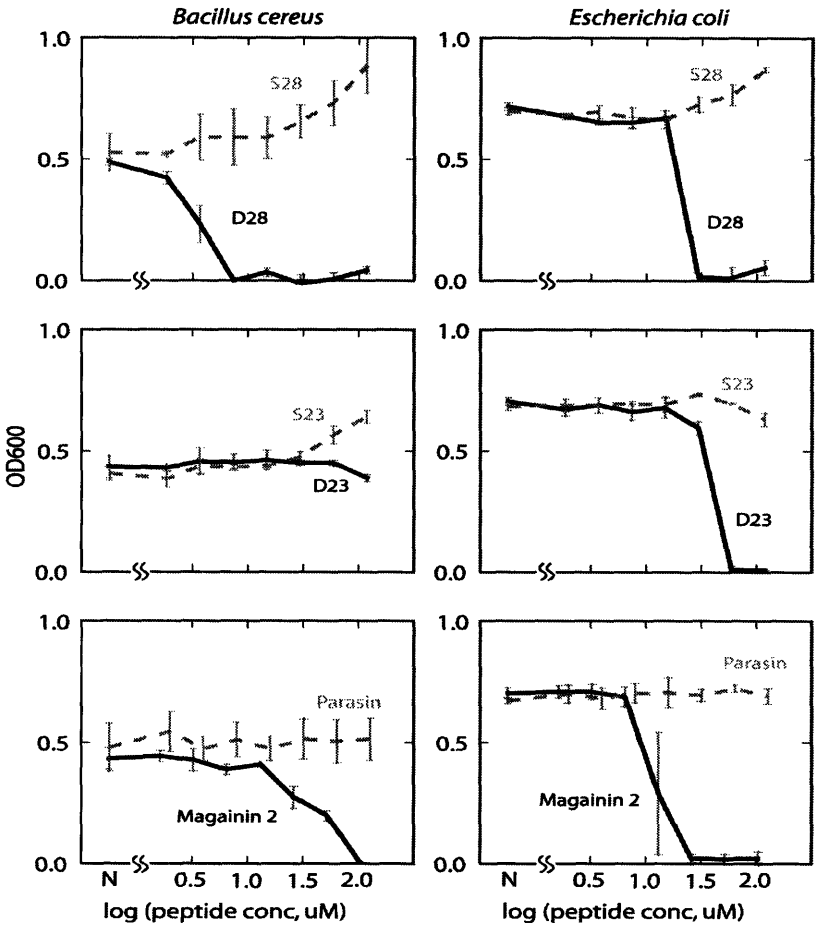



**Figure 4.13: Dose response curves for peptides against *B. cereus* and *E. coli*.**

Designed peptides D28 and D23 are shown with their shuffled controls, S28 and S23. Natural AmPs Magainin 2 (active against these bacteria) and Parasin (inactive) are also shown. The visual determination of MIC corresponded with an OD600 below 0.1 for all samples. No peptide is present at concentration N. Data is for 3 replicates, 4 OD measurements each. Error bars = sd.

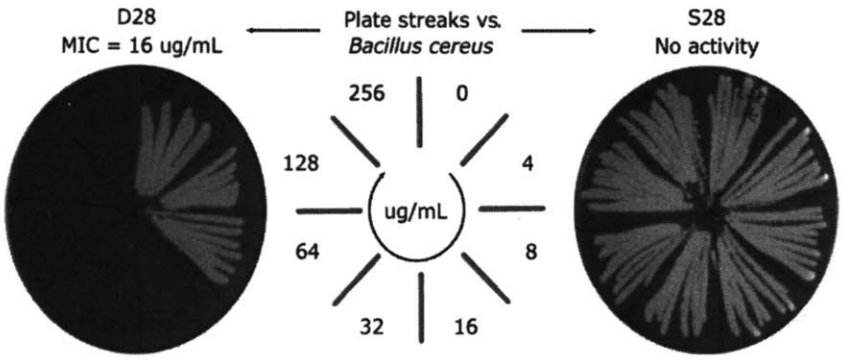



**Figure 4.14: Plating for viability following incubation with a lead design.** Samples of *B. cereus* incubated with varying concentrations of D28 and S28 overnight are streaked on agar plates, which confirms that the bacteria are not viable at or above the MIC (16 ug/ml).

## 4.3 High value targets

The prior section describes a proof of concept study for the use of pattern based AmP design. Relative to shuffled controls, a 10-fold enrichment in the fraction of active sequences was demonstrated along with a nearly 50% success rate in design. Both gram-positive and gram-negative bacteria were evaluated and strong broad spectrum activity was found for 2 leads. While *E. coli*, is an important hospital pathogen, the value of the design work would be increased if efficacy was found against pathogens that are critical in hospital infections or for national security.

### *4.3.1 Gram positive targets*

Peptides with gram-positive activity are particularly exciting because of the prevalence of drug-resistant, nosocomial *Staphylococcal aureus* and the threat of bioterror agents such as *B. anthracis*, or anthrax. *S. aureus* is an aggressive pathogen and the prevalence of drug-resistant *S. aureus* increased from 32% to 44% over just three years ending in 1999 (Lowy 2003). This bacteria frequently colonizes a patient's skin or nares and is often involved in catheter or bloodstream infections. The risk posed by bioterror agents such as anthrax was brought to the forefront in 2001. These weapons threaten both civilians and the military, and government has made the stockpiling of existing antibiotics such as Cipro™ a priority, along with supporting the development of novel agents to fight anthrax. Therefore, we assayed seven designed peptides that had gram-positive activity, including the highly active D28 and D51 peptides, against the

Smith Diffuse strain of *S. aureus* and the Sterne strain of *B. anthracis*. The results are shown in Table 4.7.

| MIC (ug/ml) | *S. aureus* | *B. anthracis* |
|---|---|---|
| 8 | D28 | |
| 16 | D51 | D28, D51 |
| 32 | | |
| 64 | D22 | D22 |
| 128 | D63, S51 | D5, D35, D43, D63 |
| 256 | D5, D35, D43 | S51 |
| >256 | S5, S22, S28 S35, S43, S63 | S5, S22, S28, S35, S43, S63 |
| D = Designed based on motifs; S = Shuffled control | | |

**Table 4.7: Minimal inhibitory concentration of designed and shuffled sequences against *S. aureus* and *B. anthracis*.** All seven designed peptides that were active against *B. cereus*, were active against both *S. aureus* and *B. anthracis* at 256 ug/ml or below. Only one of the seven shuffled bacteria displayed antimicrobial activity.

### 4.3.2 Antimicrobial activity against S. aureus and B. anthracis

As shown in Table 4.7, all seven peptides had activity against both bacteria at 256 ug/ml or less, whereas only one of the seven shuffled controls had activity (the only active shuffled peptide had MICs of 128 ug/ml against *S. aureus* and 256 ug/ml against *B. anthracis).* Moreover, two designed peptides, D28 and D51, had MICs of 16 ug/ml against *Bacillus anthracis*, which is equivalent to the activity of Cecropin-melittin hybrid, a strong natural peptide. D28 also had an MIC of 8 ug/ml against *S. aureus.* Therefore, our diversity generation tool produces leads against critical target bacteria for hospital infections and national defense.

## 4.4 Diversity of designed sequences

One goal of our design approach was to create sequences which are diverse relative to natural AmPs. This was enforced through the application of a filter that removed any sequence that had more than 5 amino acids in a row in common with a natural AmP. Nevertheless, our designed AmPs will, by design, show some degree of homology with natural AmPs because the grammars are based on native sequences. Peptide D28, for example, was matched by grammars derived from 11 natural AmPs including brevinin, temporin, and ponericin. Smith-Waterman alignments of our designed peptides against all natural AmPs in the Swiss-Prot/TrEMBL database reveal that the degree of homology is limited. In particular, our two most active peptides, D51 and D28, have only 50 and 60% sequence identity with the nearest natural AmP,

respectively. Our design tool maintains a high success rate for predicting activity while accomplishing the goal of increasing diversity.

## 4.5 Summary

We have demonstrated a pattern-based design approach for AmPs. This tool produced a high fraction of active sequences, which is significantly enriched relative to shuffled controls with the same bulk physiochemical properties. We also demonstrated the utility of the designed sequences against important bacterial targets including *S. aureus* and *B. anthracis*.

Our linguistic approach to designing synthetic AmPs may be successful due to the pronounced modular nature of naturally–occurring AmP amino acid sequences. As we have shown, this approach can be used to rationally expand the AmP sequence space without using structure–activity information or complex simulations of the interactions of a peptide with a membrane. The peptides designed in this work are different from previously designed synthetic AmPs in that they bear limited homology to any known protein, which may be desirable for AmPs used in clinical settings (Tossi, Sandri et al. 2000). Some researchers argue that widespread clinical use of AmPs that are too similar to human AmPs will inevitably elicit bacterial resistance, compromising our own natural defenses and posing a threat to public health (Bell and Gouyon 2003). Additionally, using our approach to develop an arsenal of highly diverse antimicrobials would further reduce concerns about the development of antimicrobial resistance. We hope that this approach

will help to expand the diversity of known AmPs well beyond those found in nature, possibly leading to new candidates for AmP–based antibiotic therapeutics.

Our linguistic design approach may be most valuable as a method for rationally constraining a sequence-based search for novel AmPs. Diverse leads generated by our algorithms may be optimized using approaches described in the literature (Hilpert, Volkmer-Engert et al. 2005). But, the linguistic approach described here has a number of limitations. First, sequence families that are poorly conserved on an amino acid level would not benefit from this approach. Second, we suspect that the small size of AmPs is helpful. Due to the simple nature of regular grammars, they would be less useful for designing larger proteins and, in particular, proteins with complex tertiary or quaternary structures. Nevertheless, AmPs are a class of therapeutic candidates whose size and mechanism of action make them ideal for a pattern-based design approach. This AmP design approached was published as (Loose, Jensen et al. 2006).

# Chapter 5

# Refining AmP Design

In Chapter 4, a motif-based approach to designing novel AmPs was described that predicted antimicrobial sequences with a high probability of success. Furthermore, two leads emerged from just 42 designs that had broad spectrum activity at clinically relevant concentrations. The motif-based tool may have great utility as the first step of a two-stage design approach. In the first stage, diverse leads could be created using the strategy described in Chapter 4. In the second stage, these leads are optimized using a variety of heuristics. In the following sections, a proof-of-concept optimization is carried out for one of the leads discovered. Additionally, strategies are presented to improve designs by developing information on the toxicity and activity of natural AmPs used in our training set.

## 5.1 Optimization of activity

Researchers have developed a variety of methods to improve the activity of natural AmPs. Through structure-function studies, they have found that increasing positive charge, hydrophobicity, and amphipathicity may increase activity (Hilpert,

Volkmer-Engert et al. 2005). A number of these approaches involve the use of a helical projection of a peptide, since many linear AmPs form alpha helices that segregate hydrophobic from positively charged sidechains when they interact with a bacterial membrane. A helical projection is created by placing the amino acids in order around an alpha-helical template and viewing down the length of the helix to observe the manner in which the amino acids are spatially separated.

### *5.1.1 Helical projection based optimization*

In order to demonstrate that the leads we develop may be further improved using these techniques, we optimized our best candidate, peptide D28. A helical projection of this peptide is shown in Figure 5.1.

We created 44 variants of D28 by introducing mutations that were selected to increase positive charge, increase hydrophobicity, and improve segregation of positive and hydrophobic residues. Additionally, an interior proline residue was removed because prolines typical create a break in helical structure. While a proline may be useful in segregating the activity of two distinct domains of an AmP, they may limit the cooperativity of two portions of the peptide. Each of the 44 variants was chemically synthesized and tested using methods in Section 4.2.2.3. The resulting MICs are given in Table 5.1.

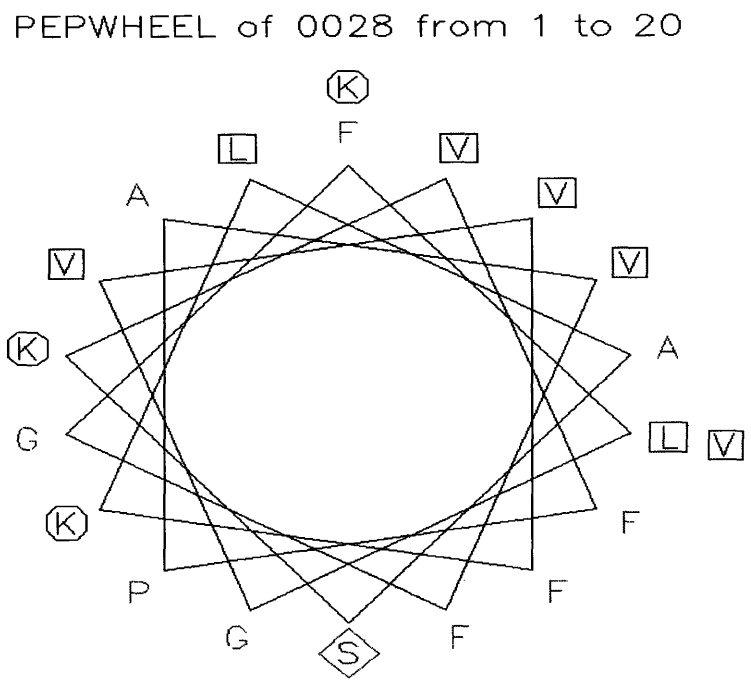



**Figure 5.1: Helical projection of lead peptide used for optimization.** Because linear peptides often form alpha helices when interacting with bacteria, heuristic methods were used based on this projection to increase positive charge, increase hydrophobicity, improve amphipathicity, and remove a helix-breaking proline to create an optimized sequence.

Revised D28

| Peptide number | Sequence | E. coli | B. cereus | S. aureus |
|---|---|---|---|---|
| D28 | FLGVVFKLASKVFPAVFGKV | 64 | 16 | 8 |
| D28 | FLGVVFKLASKVFPAVFGKV | 64 | 16 | 4 |
| D28 | FLGVVFKLASKVFPAVFGKV | 64 | 16 | 8 |
| R1 | FLKVVFKLASKVFPAVFGKV | 32 | 16 | 4 |
| R2 | FLGKVFKLASKVFPAVFGKV | 16 | 16 | 32 |
| R3 | FLGVVFKKASKVFPAVFGKV | 256 | 256 | * |
| R4 | FLGVVFKLASKVFKAVFGKV | 128 | 8 | 8 |
| R5 | FLGVVFKLASKVFPAVFKKV | 64 | 16 | 4 |
| R6 | FLKKVFKLASKVFPAVFGKV | 4 | 32 | 64 |
| R7 | FLGKVFKKASKVFPAVFGKV | 32 | * | * |
| R8 | FLGKVFKLASKVFKAVFGKV | 16 | 8 | 4 |
| R9 | FLGKVFKLASKVFPKVFGKV | 16 | 32 | 64 |
| R10 | FLGKVFKLASKVFPAVFKKV | 16 | 128 | 128 |
| R11 | FLKVVFKKASKVFPAVFGKV | 128 | * | * |
| R12 | FLGVVFKKASKVFKAVFGKV | 32 | 32 | 256 |
| R13 | FLGVVFKKASKVFPKVFGKV | 128 | * | * |
| R14 | FLGVVFKKASKVFPAVFKKV | 256 | * | * |
| R15 | FLGGVFKLASKVFPAVFGKV | 64 | 32 | 64 |
| R16 | FLGSVFKLASKVFPAVFGKV | 64 | 32 | 64 |
| R17 | FLGVVFKGASKVFPAVFGKV | * | 256 | * |
| R18 | FLGVVFKSASKVFPAVFGKV | * | * | * |
| R19 | FLGGVFKGASKVFPAVFGKV | * | * | * |
| R20 | FLGSVFKSASKVFPAVFGKV | * | * | * |
| R21 | FLGVVVKLASKVFPAVFGKV | * | 32 | 16 |
| R22 | FLGVVFKLVSKVFPAVFGKV | 256 | 16 | 8 |
| R23 | FLGVVFKLAVKVFPAVFGKV | * | * | * |
| R24 | FLGVVFKLASKVVPAVFGKV | 256 | 64 | 16 |
| R25 | FLGVVFKLASKVFPAVVGKV | 256 | 64 | 32 |
| R26 | FLGKVVKGASKVFPAVFGKV | * | * | * |
| R27 | FLGKVFKGVSKVFPAVFGKV | 128 | * | * |
| R28 | FLGKVFKGAVKVFPAVFGKV | 128 | * | * |
| R29 | FLGKVFKGASKVVPAVFGKV | * | * | * |
| R30 | FLGVVFKGASKVFPAVVGKV | * | * | * |
| R31 | FLGKVVKKASKVFPAVFGKV | 128 | * | * |
| R32 | FLGKVFKKVSKVFPAVFGKV | 32 | * | * |
| R33 | FLGKVFKKAVKVFPAVFGKV | 32 | * | 256 |
| R34 | FLGKVFKKASKVVPAVFGKV | 256 | * | * |
| R35 | FLGKVFKKASKVFPAVVGKV | 128 | * | * |
| R36 | FLGVVFKLASKVFPAVFGKV-NH3 | 64 | 16 | 4 |
| R37 | FLGVVFKLASKVFGAVFGKV | * | 16 | 8 |
| R38 | FLKVVFKLASKVFGAVFGKV | 128 | 16 | 16 |
| R39 | FLGKVFKLASKVFGAVFGKV | 16 | 8 | 8 |
| R40 | FLGVVFKKASKVFGAVFGKV | 64 | 128 | * |
| R41 | FLGVVFKLASKVFGKVFGKV | 64 | 8 | 8 |
| R42 | FLGVVFKLASKVFGAVFKKV | 256 | 8 | 8 |
| R43 | FLGKVFKGASKVFGAVFGKV | 32 | 128 | * |
| R44 | FLGGVFKKASKVFGAVFGKV | 32 | 128 | * |

+ = MIC greater than 256 ug/mL
++ = MIC greater than 128 ug/mL, not sufficiently soluble to test at 256 ug/mL

**Table 5.1: Minimum inhibitory concentrations for variants of a lead peptide**. By creating just 44 variants of a lead peptide, an optimized sequence was found with 2-4-fold improved MICs (R8). This study demonstrates that diverse leads created with a pattern-based approach may be optimized in a second stage of design.

### 5.1.2 Antimicrobial activity of optimized sequences

As shown in Table 5.1, 18 of the 44 D28 variants showed improved activity against *E. coli*, *B. cereus*, or *S. aureus*. Many of the D28 variants with improved activity against *B. cereus* included a mutation at an internal proline, either to lysine or glycine. These mutations removed the helix breaking proline and may allow the entire structure to form an alpha helix. One variant (R8) had MICs of 16 ug/ml against *E. coli* and 8 ug/ml against *B. cereus* and 4 ug/ml against *S. aureus* (relative to 64, 16, and 8 ug/ml, respectively, for D28). This demonstrates that a lead may be quickly optimized to improve its broad spectrum activity. If this process were automated, one would have the opportunity to produce a much larger number of diverse leads, which could be more thoroughly optimized.

Additionally, all of the work above was carried out using standard assays for inhibitory activity. Because AmPs are known to work through a membrane destabilization attack that is usually lethal, the minimum bactericidal concentrations (MBCs) are frequently similar to the minimum inhibitory concentrations (MICs). In order to measure MBCs, the standard MIC assay was carried out, and at the end of the 16-20 hour incubation, a sterile toothpick was dipped in each dilution of the 96 well plate and streak on CMHB agar plates. These plates were grown overnight and all samples were checked viable cells. As expected, D28 and all six of its variants that were assayed for bactericidal activity had an MBC within a 2-fold dilution of their MIC.

## 5.2 Evaluating mammalian toxicity

Antimicrobial peptides are known to discriminate between bacterial and eukaryotic membranes based on charge, membrane potential, and presence of cholesterol, among other characteristics. Nevertheless, systemic antibiotics require large doses to maintain therapeutic concentrations throughout the body. In part, because AmPs have not necessarily been designed by nature to be non-toxic to a human host in a systemic setting, many sequences have significant toxicity at a therapeutic concentration.

A diversity generating design tool will help enable these toxicity barriers to be overcome. First, a great variety of peptides may be assayed for toxicity *in vitro*, allowing researchers to explore new regions of sequence space that may be populated by highly selective AmPs. Additionally, this design tool will allow the generation of diverse AmPs which may be tested for toxicity. This information could be used to create a filter that could be used to eliminate future designs that are likely to be toxic.

### *5.2.1 Methods for assessing hemolytic activity*

For the development of AmPs as systemic antibiotics, the first toxicity screen typically carried out is for hemolytic activity. This screen is preferred because it does not require cell culture and gives information on a critical toxicity that many AmPs display. The hemolytic activity of the peptides may be quantified by measuring the concentration at which 50% hemolysis of Human Red Blood Cells occurs (HC50) using standard high-throughput methods (Hamuro et al., 1999). A stock of 10% washed pooled red blood

cells is diluted to 0.25% with a buffer of 150 mM NaCl and 10 mM Tris at pH 7.0. Twenty ul of peptide is diluted into 80 ul of the 0.25% red blood cell solution and incubated for 1 hour at 37°C in a Costar round bottom polypropylene microtiter plate. Cells are spun down at 6000g, the supernatant removed, and the OD414 measured on a Fusion Microplate Reader. Total hemolysis will be defined by Melittin at 128 ug/ml or sterile DI water. Using these methods, the hemolytic activity of lead AmPs D28 and D51 was assessed. While D28 had an HC50 of 64 ug/ml, D51 had an HC50 greater than 512 ug/ml.

## 5.3 Characterizing natural AmPs

The AmP design carried out in Section 4.1 was based on a list of sequences in a database of natural AmPs. It was unknown which natural AmPs were active against the targeted bacteria. Further, the strength and spectrum of activities of natural AmPs was unknown. Also, no standardized toxicity data was available for the set of natural AmPs. While publications exist that describe the activity and toxicity of many AmPs, the methods for synthesizing or assaying are not standardized, and different strains of target bacteria are used.

It is likely designs could be improved given information on the activity and toxicity of sequences in the training set. This information could be used during the pattern discovery phase by removing sequences that are not active against bacteria of interest. Alternatively, the information could be incorporated during the reassembly of

patterns or scoring phase. Significantly toxic members could be removed from the training set or the score of any motifs derived from them could be penalized.

### 5.3.1 Screening for antimicrobial activity

In order to develop this information, we desired to characterize a diverse, representative set of natural AmPs. The sequences were chosen from the University of Nebraska Medical Center AmP database, the same set used for previous design. Only linear sequences between 10 and 30 amino acids with no disulfide bonds were used because they could be created using a chemical synthesizer with a high probability of successful synthesis. The members of the database were clustered using k-means clustering to select 100 natural AmPs for evaluation. Of these 100 peptides, 82 were successfully synthesized and were soluble at sufficient concentration to measure their MICs up to 256 ug/ml. Antimicrobial activities against *E. coli* ATCC 25922, *S. aureus* ATCC 25923, and *S. epidermidis* ATCC 14990 were measured using standard techniques given in 4.2.2.3. A distribution of the resulting activities are shown in Figure 5.2.

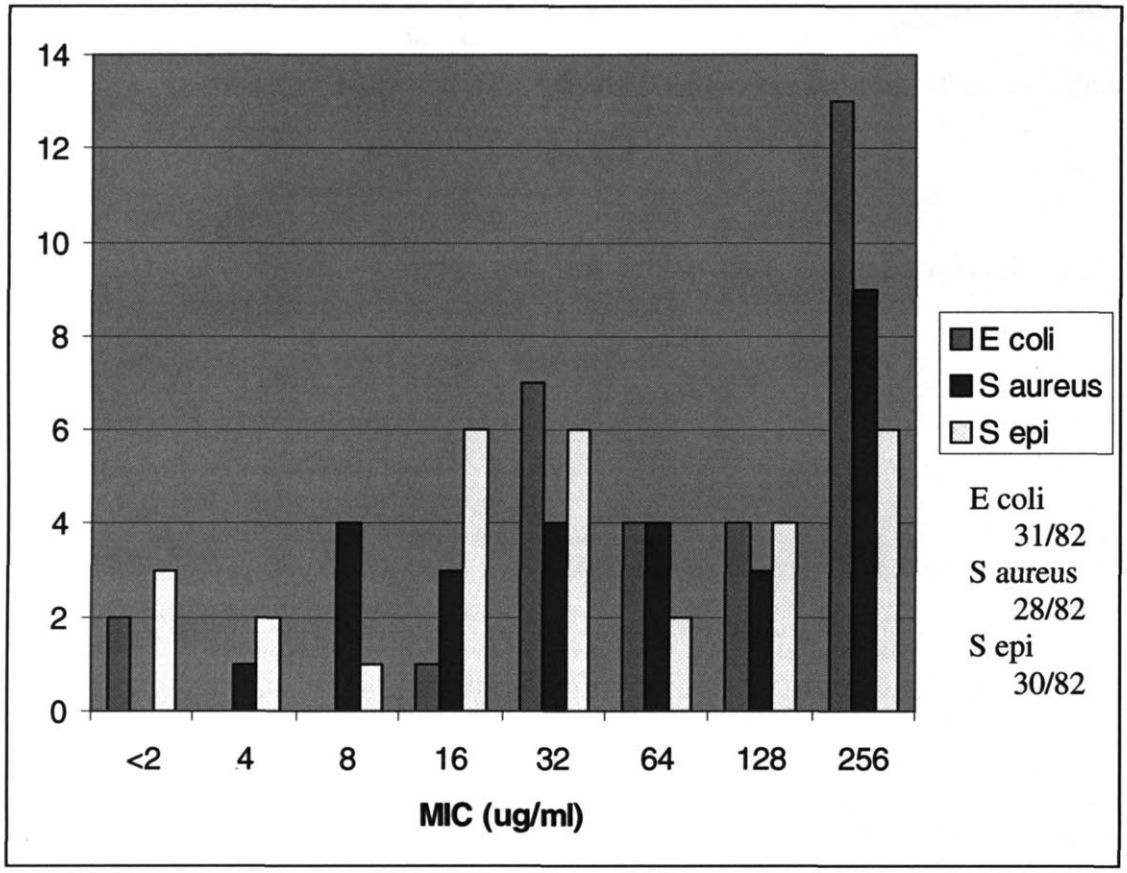



**Figure 5.2: Distribution of antimicrobial activity for natural AmPs**. The MICs for 82 representative natural AmPs was evaluated *E. coli*, *S. aureus*, or *S. epidermidis*. Only approximately 40% of the peptides had an MIC at or below 256 ug/ml against each bacteria. Eliminating natural AmPs from the training set that are inactive against target bacteria may improve the success rate for new designs.

### 5.3.2 Antimicrobial activity of natural AmPs

In seen in Figure 5.2, for each target bacteria, approximately 40% of natural AmPs are active at 256 ug/ml or below. This likely limited the fraction of AmPs that were designed that had activity in this range and may explain why roughly half of our designed sequences were not active against similar bacteria. The natural AmPs that were not active in this screen may have been annotated as AmPs because they are active against bacteria not tested or at higher concentrations than were tested. Alternatively, they could have an evolutionary similarity to or share a conserved pro-region with other AmPs. Narrowing the training set to AmPs active against target bacteria or developing a scoring system that reflects natural AmPs activity would likely be more successful than our initial design and is discussed in Future Work.

## 5.4 Correlation of antimicrobial activity and toxicity

As described in Section 5.3, the antimicrobial activities of natural AmPs vary widely. While incorporating MIC information for natural AmPs in our design may improve activity, it would be even more advantageous if toxicity information were simultaneously utilized. Ideally, the hemolytic activity of individual AmP sequences would not be highly correlated with antimicrobial activity so that design tools could be tuned to produce highly active and non-toxic sequences.

### *5.4.1 Screening for hemolytic activity*

In order to examine this correlation, the hemolytic activities of the 82 natural peptides screened in Section 5.3 were characterized using methods given in Section 5.2.1. The resulting distribution of toxicity versus MIC is shown in Figure 5.3.

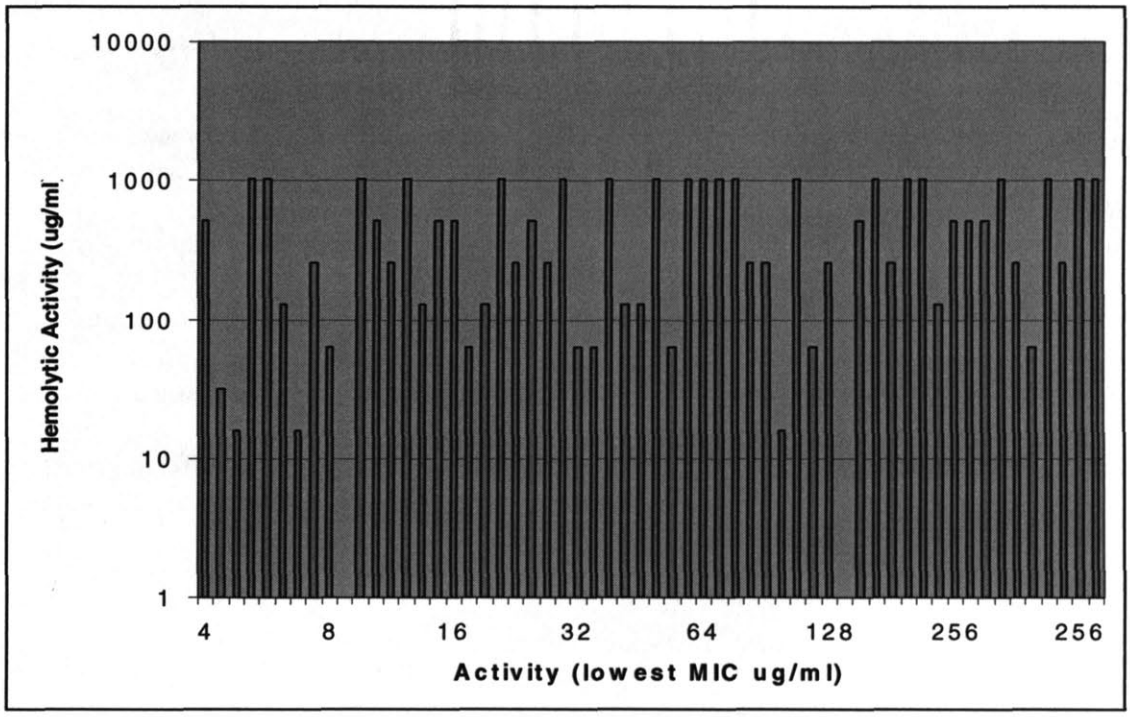



**Figure 5.3: Hemolytic activity versus antimicrobial activity for natural AmPs.** The hemolytic concentrations for 82 natural AmPs were plotted against the lowest MIC for each peptide against *E. coli*, *S. aureus*, or *S. epidermidis*. Strong hemolytic activity is not significantly correlated with strong antimicrobial activity, with some low MIC peptides demonstrating little hemolysis. This indicates that a scoring approach could be designed to enrich a population of highly active and non-toxic AmPs.

### 5.4.2 Correlation of hemolytic and antimicrobial activity

As shown in Figure 5.3, hemolytic activity is not highly correlated with antimicrobial activity. There are many AmPs with strong antimicrobial activity (low MIC) that have low toxicity (a high concentration required for hemolytic activity). Ideally, activity and toxicity and activity could be incorporated into a scoring algorithm to enrich the desirable set of highly active and non-toxic AmPs, as discussed in Future Work.

### 5.4.3 Distribution of therapeutic indices for natural AmPs

Figure 5.3 demonstrated that antimicrobial and hemolytic activity were not strongly correlated. An important metric for the selectivity of AmPs for bacterial cells is the therapeutic index. This is defined as:

Therapeutic Index = Hemolytic Concentration / Antimicrobial Concentration

For the purposes of our designs, the HC50, at which the peptide lyses 50% of red blood cells is used for the hemolytic concentration and the minimal MIC for *E. coli*, *S. aureus*, and *S. epidermidis* is used as the antimicrobial concentration. Figure 5.4 shows the distribution of therapeutic indices for the set of 82 natural peptides described in Section 5.3. The therapeutic indices vary widely, and AmPs with MICs of 32 ug/ml or below tend to have a higher therapeutic index.

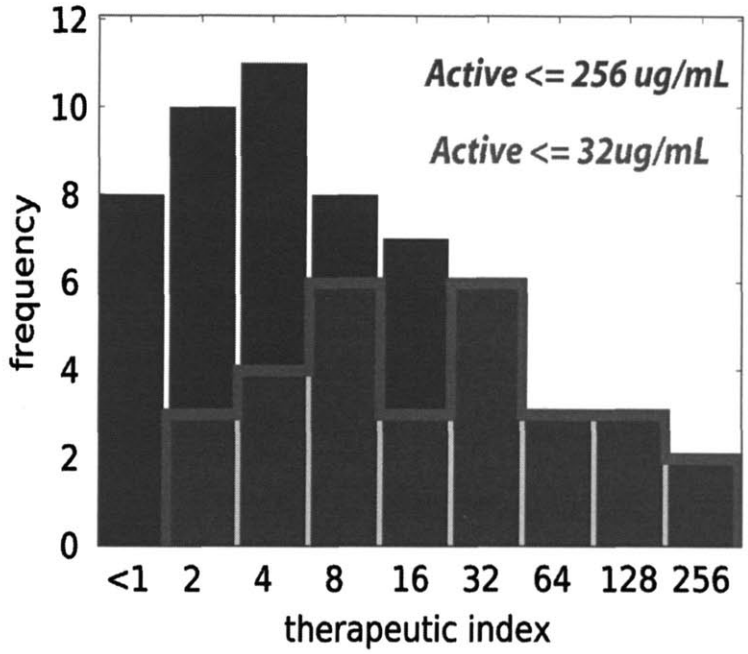



**Figure 5.4: Therapeutic index distribution for weak and strong natural AmPs.** A histogram of therapeutic indices for natural AmPs in solution (hemolytic concentration / MIC). The MIC was defined as the smallest of MICs against E. coli, S. aureus, or, S. epidermidis. The smaller outlined distribution is for strong AmPs which have MICs of 32 ug/ml or below. Because there is a wide distribution of MICs, the activity and toxicity of peptides are not highly coupled, allowing highly active and non-toxic variants to be designed. Figure reproduced from JF Moxley.

## 5.5 Effect of c-terminal amidation

One of the advantages to a high-throughput design tool and assay is that it allows users to deduce rules or test hypotheses for AmP design. One attractive question we chose to explore is whether covering the negatively charged C-terminus of an AmP would systematically improve activity for a diversity of AmPs. It has been reported for individual AmPs that amidating the C-terminus, which eliminates the negative charge, may increase activity (Machado, Sforca et al. 2007). Nevertheless, a systematic study of the effect of amidation on activity has not been reported.

### *5.5.1 Effect of amidation of representative AmPs*

In order to test the general effect of amidating AmPs, a representative set of designed, shuffled, and natural peptides were chemically synthesized with and without amidation and assayed against *E. coli* and *B. cereus*. The MICS are shown in Table 5.2.

| Peptide | E coli MIC (ug/ml) | | B cereus (ug/ml) | |
|---|---|---|---|---|
| | Free Pep | Amidated | Free Pep | Amidated |
| 0 | <4 | <4 | > 256 | > 256 |
| 5 | 128 | 32 | 128 | 64 |
| 20 | 32 | 16 | <4 | <4 |
| 28 | 64 | 32 | 8 | 8 |
| 40 | 32 | 32 | 128 | 128 |
| 46 | > 256 | 64 | > 256 | 64 |
| 51 | 64 | 16 | 16 | 8 |
| 62 | 256 | 64 | > 256 | 256 |
| 72 | <4 | <4 | 16 | <4 |

**Table 5.2: Effect of C-terminal amidation on antimicrobial activity for representative peptides.** The MIC of free acid peptides was found to improve 2-4 fold if the peptide was amidated. The improvement may result from the shielding of the negatively charged C-terminus, which may inhibit interaction with a bacterial membrane.

As shown in Figure 5.2, amidated AmPs tend to have stronger activity than their free acid counterparts. The improvement in activity in this sampling was encouraging, but it was desired to see if this effect was consistent across a broader selection of natural AmPs. If it were, this would enable us to improve the activity of all of our designs an average of 2 to 4-fold.

### 5.5.2 Effect of amidation on a broader set of natural AmPs

In order to test the generality of the effect of amidation on activity, all of the natural peptides characterized in Section 5.3 were synthesized with amidated c-termini. MICs were measured against *E. coli* ATCC 25922, *S. aureus* ATCC 25923, and *S. epidermidis* ATCC 14990. A histogram of MICs against each bacteria is shown in Figures 5.5 through 5.7.

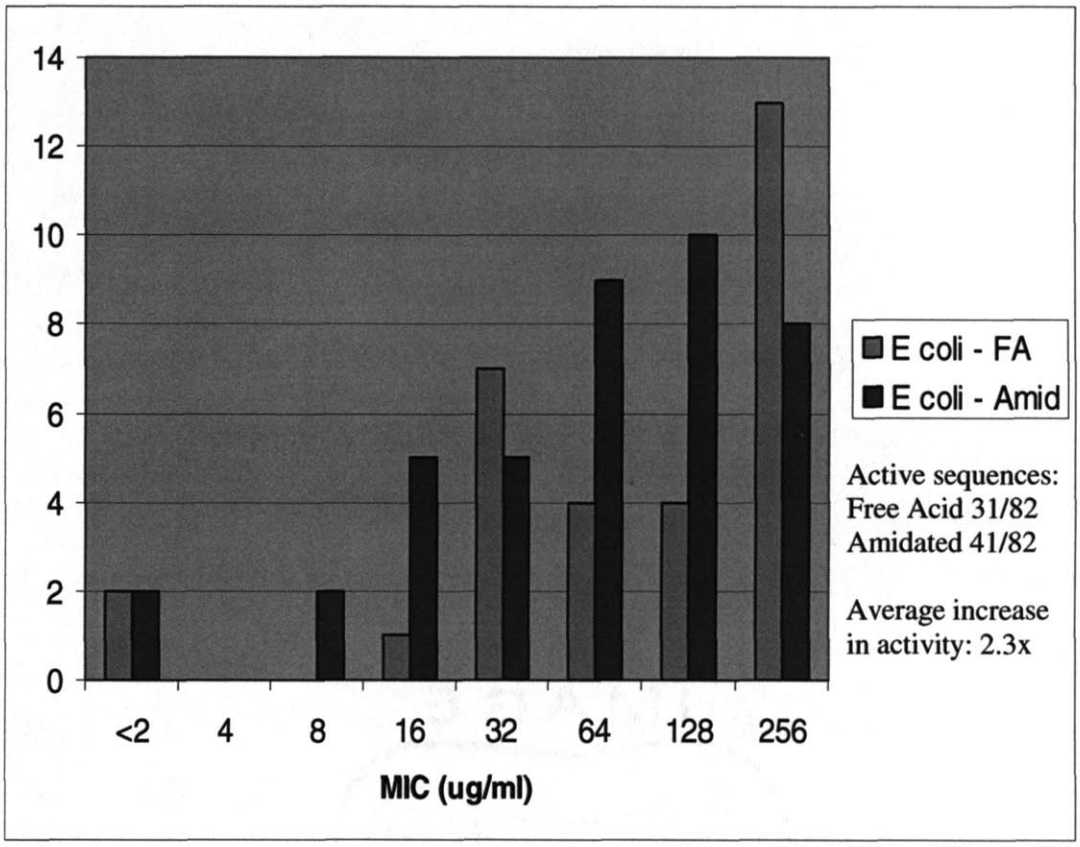



**Figure 5.5: Effect of amidation on the MIC against *E. coli*.** The fraction of peptides with MICs of 256 ug/ml or below increases from 31/82 for free acid peptides (FA) to 41/82 for amidated peptides (Amid), and results in a greater number of highly active (<16 ug/ml) peptides.

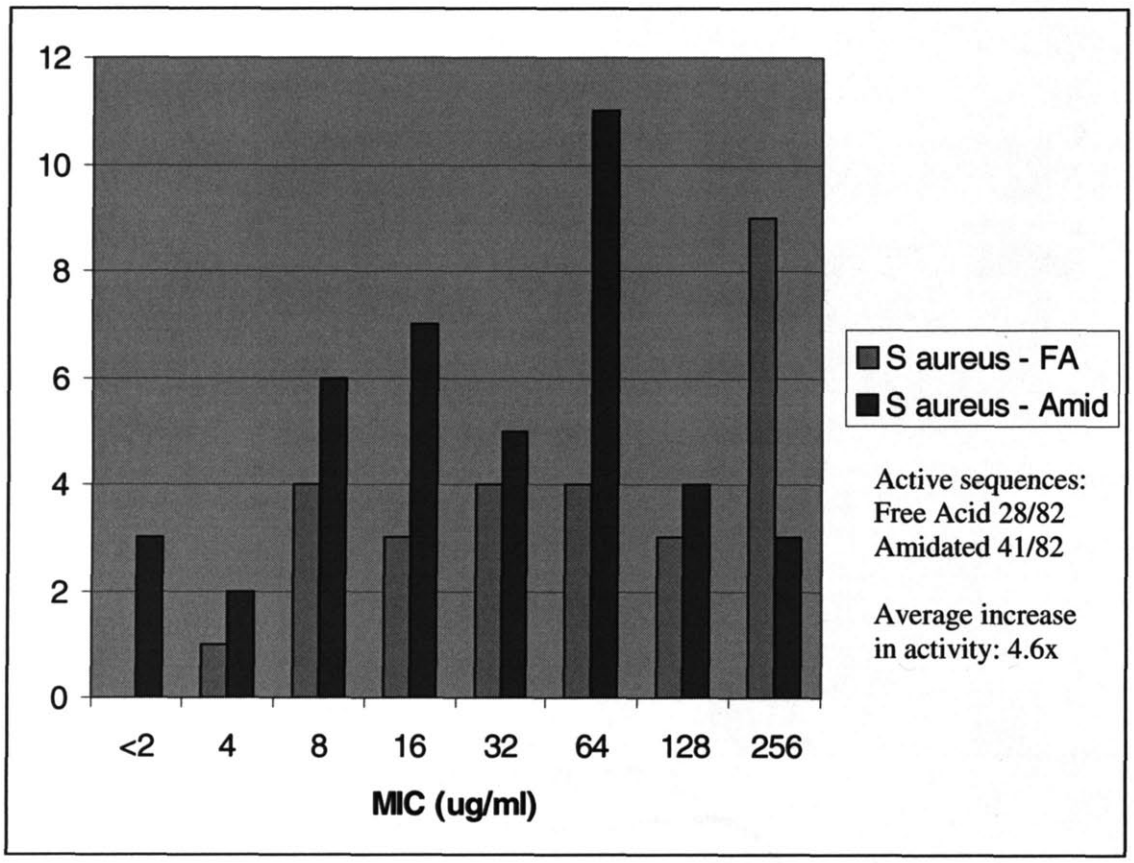



**Figure 5.6: Effect of amidation on the MIC against *S. aureus*.** The fraction of peptides with MICs of 256 ug/ml or below increases from 28/82 for free acid peptides (FA) to 41/82 for amidated peptides (Amid), and results in a greater number of highly active peptides.

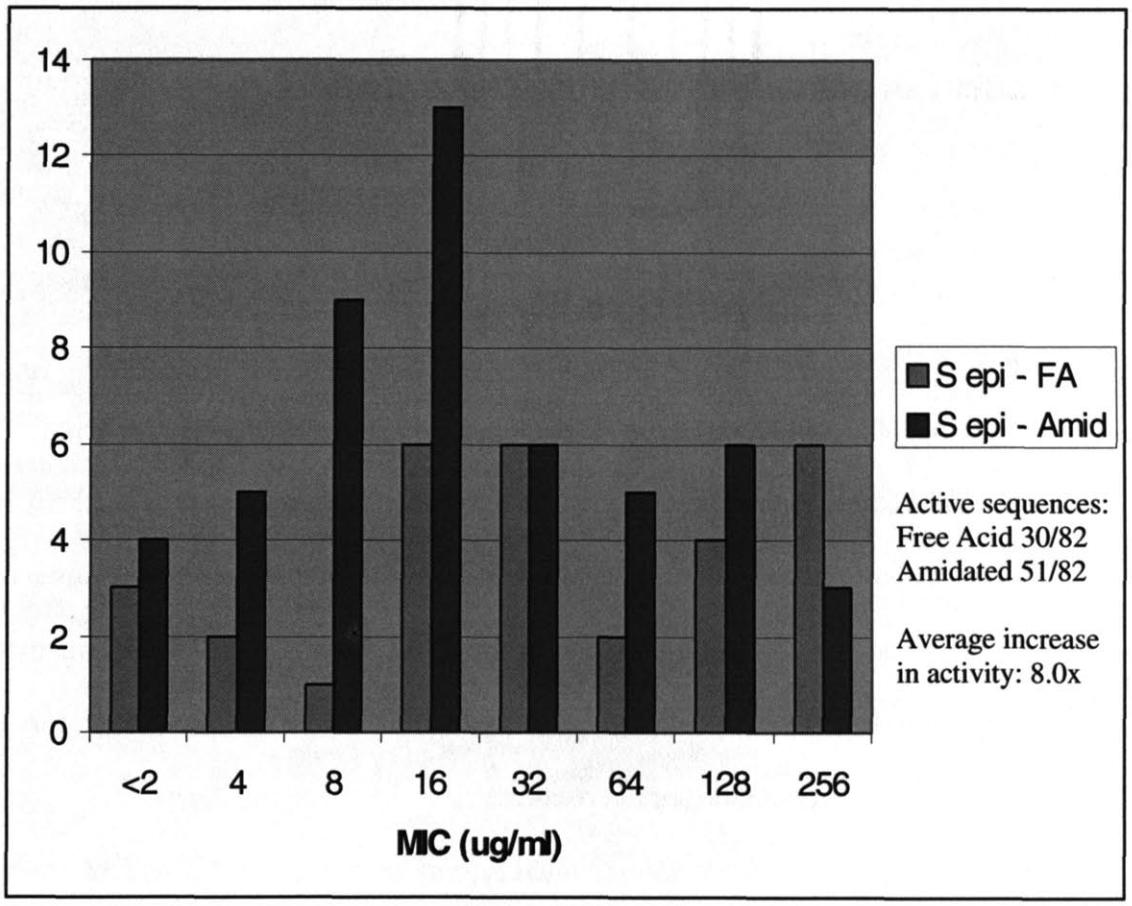



**Figure 5.7: Effect of amidation on the MIC against *S. epidermidis*.** The fraction of peptides with MICs of 256 ug/ml or below increases from 30/82 for free acid peptides (FA) to 51/82 for amidated peptides (Amid), and results in a greater number of highly active peptides.

### 5.5.3 Effect of amidation on hemolytic activity

In the previous section, it was demonstrated that amidating the C-terminus of natural AmPs gives a broad spectrum improvement in antimicrobial activity, with the most significant improvement against gram-positive bacteria. This result would be particularly useful if there was not a concomitant increase in hemolytic activity. Even without an improvement in therapeutic index, an equivalent increase in activity and toxicity is desirable because it reduces the concentration of drug that would be required to achieve a therapeutic effect, lowering the cost of the drug for each dose. To this end, the hemolytic activities of the 82 natural AmPs screened in Section 5.3 were quantified for the free acid and amidated versions of the peptides.

As seen in Table 5.3, while antimicrobial activity increases 2.3, 4.6, and 8.0 -fold for *E. coli* ATCC 25922, *S. aureus* ATCC 25923, and *S. epidermidis* ATCC 14990, respectively, the average hemolytic activity increases only 1.9 fold. Therefore the average therapeutic index increases against all bacteria, with an 2.4-fold improvement against *S. aureus*, and an 4.0-fold improvement against *S. epidermidis*. In future designs, amidated versions of all leads should be tested in hopes of achieving this same therapeutic index improvement.

| | *E. coli* | *S. aureus* | *S. epidermidis* |
|---|---|---|---|
| Avg improvement in MIC from amidating | 2.3x | 4.6x | 8.0x |
| Avg improvement in therapeutic index from amidating | 1.2x | 2.4x | 4.0x |

**Table 5.3: Improvement in antimicrobial activity and therapeutic index from amidation**. Amidating the C-terminus of peptides increased hemolytic activity an average of only 1.9x, whereas antimicrobial activities (MICs) improved 2.3-8.0x. This resulted in improvement in the therapeutic index, defined as the hemolytic concentration divided by the MIC.

## 5.6 Summary

Ideally, a pattern-based AmP design approach would be used as the first step of a two-stage design tool, with leads being optimized in the second step. Additionally, our designs would likely improve if the members of the training set were better characterized. For this reason, the antimicrobial activity and toxicity of a representative set of natural, linear AmPs was evaluated. We found that antimicrobial activity was not highly correlated with toxicity, indicating that highly active and non-toxic sequences maybe designed. Further, we systematically showed that amidating the c-termini of AmPs improves activity and therapeutic index. Suggestions for further improving a pattern-based AmP design methodology are given in Section 7.2.1.

# Chapter 6

# Medical Device Applications

As described in Section 2.4.2, the development of AmPs as systemic antibiotics requires half-life, stability, and toxicity challenges to be overcome. Creating a localized application for AmPs sidesteps many of these problems. Furthermore, hospital infections are frequently associated with medical devices, so a dramatic clinical impact could be achieved by creating AmPs localized on medical devices. In this Chapter, the role of medical devices in infections is outlined. Existing antimicrobial technologies, along with their shortcomings, are discussed. Finally, the potential of immobilized AmPs as antimicrobial coatings is described, along with a series of formulations that have been developed.

## 6.1 Infections associated with medical devices

The CDC estimates that hospital infections afflict 1.7 million patients and kill 99,000 people in the US each year (http://www.hhs.gov/asl/testify/t060329.html). Medical devices such as catheters and implants are frequently involved in these infections. In fact, the majority of hospital infections in critically ill patients are associated with

medical devices (Darouiche 2001). In the following sections, the pathways for device colonization, an analysis of devices susceptible to infection, and the problems posed by biofilms are described.

### 6.1.1 Pathways for medical device infection

Medical devices provide a surface for bacterial growth and, when penetrating the skin, a route for external bacteria to enter the bloodstream. Patients in a hospital setting are particularly likely to be colonized by drug-resistant bacteria, which are particularly difficult to treat. Figure 6.1 depicts the risks posed by these two classes of devices. If these infections reach the bloodstream, they lengthen average hospital stays from 5.4 to 20 days and cost up to $50,000 to treat (US Antimicrobial Coatings Market, Frost and Sullivan, 13 Feb 2006). Further, mortality from bloodstream infections have been reported from 5 to 25% based on the health of the patient group (Pittet, Tarara et al. 1994).

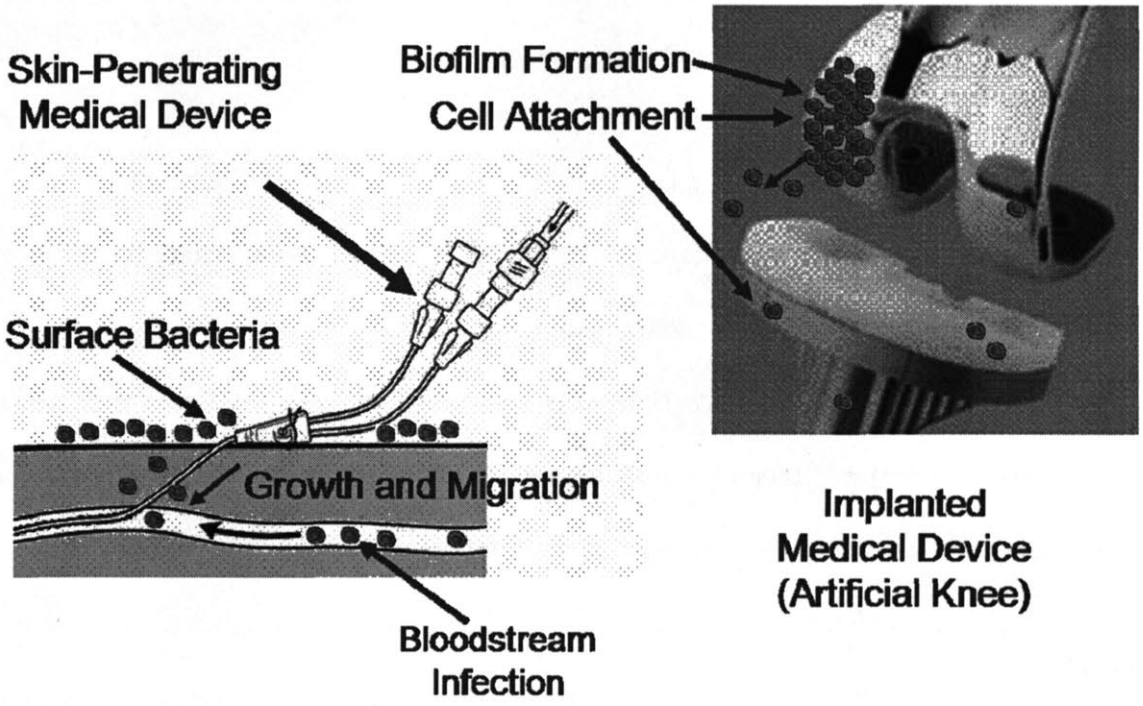



**Figure 6.1: Paths for infection of medical devices**. Skin-penetrating medical devices, such as venous catheters, allow a path for skin bacteria to enter the bloodstream. Fully implanted devices are suitable surfaces for bacterial growth and biofilm formation. Figure reproduced from M Hencke.

### 6.1.2 Products associated with infection

As described in Section 6.1.1, both fully implanted and skin-penetrating devices are susceptible to infection. The large infection costs and mortality associated with these products provide drivers for the development of antimicrobial technology to prevent these infections. The sizes of a number of medical device markets, along with the total amount spent on treating infections associated with each of these devices, are given in Table 6.1. Infection costs per device are also calculated.

For many of these devices, the cost of treating infections exceeds the cost of the device itself. This is most dramatically evident for central venous catheters (CVCs), whose infection cost per device used is ten-fold higher than device cost. Because central venous catheters directly access the bloodstream, infections frequently become bloodborne, leading to these high costs. If an effective technology were developed to reduce the likelihood of infection by 50%, there would be a clear incentive for adoption by hospital purchasers even with premium pricing. All of these markets offer either a large potential premium or a large total market, and even the lowest achievable premium ($70 for urinary catheters) is high enough to justify market entry. Smaller markets such as ventricular assist devices may be particularly rewarding. These devices have an average associated infection cost of $20,000 per device, so a coating that reduces this rate by 50% provides a savings of $10,000 per device. Ideally, a versatile technology could be developed that addresses devices of all shapes, sizes, and materials.

| Device | Market Size (SM) | Total Infection Cost (SM) | Price per Device ($) | Infection Cost per Device ($) |
|---|---|---|---|---|
| Fracture fixation (nails, rods) | 1,498 | 1,500 | 750 | 750 |
| Urinary catheters | 283 | 1,900 | 10 | 70 |
| Artificial heart valves | 339 | 48 | 14,000 | 2,000 |
| Vascular grafts | 115 | 275 | 650 | 1,600 |
| Ventricular assist devices | 130 | 44 | 59,000 | 20,000 |
| Central venous catheters | 344 | 3,440 | 45 | 450 |

**Table 6.1: Market sizes and infection costs for medical devices.** A sampling of medical devices are shown, many of which have average infection costs per devices greater than the cost of each device. In addition to improving patient health, an economic justification exists for an effective antimicrobial coating is added, even if a premium is charged. (Veenstra, Saint et al. 1999; Darouiche 2004).

### 6.1.3 Biofilm formation

Once bacteria colonize a device surface, they frequently form biofilms. These biofilms are comprised of highly differentiated and complex regions in which interior bacteria are sheltered from agents in the bloodstream (Costerton, Stewart et al. 1999). Once a biofilm is formed, the bacteria have reduced susceptibility to both the immune system and systemic antibiotics (Caraher, Reynolds et al. 2007). For many device infections, the lack of systemic treatment options requires that physicians remove the medical device, treat with antibiotics, and then insert a new device. For orthopedics such as artificial hips and knees, which have been cemented in place, this operation is very traumatic, especially given that many of the patients are elderly.

The details of biofilm formation and the biofilm environment have been a focus of infectious disease researchers in recent years. The mechanism of lifestyle switching have been elucidated for some bacteria, but the complex process of niche-formation within biofilms is not fully understood (Kolter and Greenberg 2006). The factors which induce biofilm formation vary, but it is often triggered by a quorum sensing process that is activated once a sufficient local concentration of bacteria are present, and it may be accelerated through stresses to the bacteria (Balaban, Gov et al. 2003). For instance, systemic antibiotics have been shown to induce biofilm formation (Rachid, Ohlsen et al. 2000). Following biofilm formation, clumps of bacteria may detach and create infections in a variety of tissues distant from the original infection site (Fux, Wilson et al. 2004). Because treatment of fully formed biofilms is rarely successful, preventing colonization

and biofilm formation in the first place may be a more effective strategy to block implant infection.

## 6.2 Existing antimicrobial technology

Given the enormous impact of device-related infection, a variety of antimicrobial products have been designed in an attempt to reduce colonization. These devices use the slow-release of existing antibiotics such as silver other antibiotics to attack bacteria in the neighborhood of the device. Technologies on the market and under development are described in the following sections, followed by the shortcomings of these approaches.

### *6.2.1 Coatings on the market and in development*

Great attention has been paid to central venous catheters given the significant infection cost. There are currently three antimicrobial CVCs with significant clinical use. ARROWg+ard® Blue catheters (Arrow International) are impregnated with a combination of chlorhexidine (Kuyyakanond and Quesnel 1992) and silver sulfadiazine, whose antimicrobial activity is primarily due to silver's disruption of the electron transport chain and DNA replication (Silver, Phung le et al. 2006). These catheters have been shown in clinical studies to reduce catheter colonization by 44% (Veenstra, Saint et al. 1999). Cook Critical Care's Spectrum® line of catheters utilizes the slow-release of minocycline, which disrupts protein synthesis (Speer, Shoemaker et al. 1992) and

rifampin, which inhibits RNA polymerase (Kim, Kim et al. 2005). These catheters have been shown in clinical studies to reduce catheter colonization by 69% (Raad, Darouiche et al. 1997). Edwards Lifesciences' Vantex® catheters release silver, carbon, and platinum ions, with most of the antimicrobial activity attributed to the silver ions. These catheters have a demonstrated reduction in catheter colonization of approximately 35%, which may be limited in part by the in vivo sequestration of silver ions by albumin in the blood stream (Corral, Nolla-Salas et al. 2003). An alternative approach, using long, cationic hydrophobic polymers has shown antimicrobial activity, but has yet to demonstrate sufficient selectivity to enable use on an implanted device (Tiller, Liao et al. 2001; Lewis and Klibanov 2005).

### *6.2.2 Shortcomings of slow-release coatings*

Inherently, slow-release formulations lose efficacy over time. Much of the research to improve existing technologies has focused on extending lifespan by increasing the loading of antibiotics. One strategy has been to increase the polymer charge on the surface of the device to ionically bind a higher concentration of antibiotic (Bassetti, Hu et al. 2001). For many applications, the lifespan remains insufficient. Additionally, increasing drug load or directly impregnating a polymer with antibiotic may impact the physical performance of the device. In the following sections, additional concerns with existing slow-release technology, including the induction of drug resistance and hypersensitivity, are described.

#### 6.2.2.1 Drug resistance

Drug resistance to antimicrobials used in existing slow-release formulations is well documented. Furthermore, selection experiments using these antimicrobials or the antimicrobial impregnated catheters creates bacterial strains with increased resistance *in vitro*, particularly for rifampin (Sampath, Tambe et al. 2001; Tambe, Sampath et al. 2001). Both chlorhexidine, silver sulfadiazine, and minocycline may also induce bacterial resistance (Speer, Shoemaker et al. 1992; Silver, Phung le et al. 2006). Additionally, the existing drug-resistant bacteria are less susceptible to these technologies. For instance, rifampin-resistant staphylococcal strains, which are common, show reduced susceptibility to rifampin impregnated catheters (Sampath, Tambe et al. 2001). Clinicians are hesitant to preventively use antibiotic-impregnated catheters because of the risk of encouraging the development of drug-resistance, and the knowledge that many resistant strains already exist.

#### 6.2.2.2 Systemic Toxicity and Hypersensitivity

Slow-release coating also pose a risk because the active agents may lead to toxicity in the bloodstream or at a distant target in the body. For example, chlorhexidine is known to result in hypersensitivity reactions (Terazawa, Shimonaka et al. 1998). Additionally, toxicity has been reported for silver ions that enter circulation (Trop, Novak et al. 2006). Inherently, a functional coating which does not release its active agent will have reduced concerns over toxicity since the agent is permanently attached at the site of

interest. Clinicians and the FDA may be more receptive to an immobilized approach given these safety advantages.

## 6.3 AmPs as device coatings

AmPs have a number of advantageous properties for developing antimicrobial coatings from them. First, they work through a mechanism that is unlikely to induce drug resistance. This alleviates concerns from clinician about furthering this problem. Also, they are membrane targeting, and may retain their function when permanently attached to the surface of a device. By permanently attaching the AmPs, lifespan may be longer than slow-release coatings and systemic toxicity risks reduced. In the following sections, data that demonstrates the potential of an immobilized AmP coating for medical devices is presented.

### *6.3.1 Immobilized AmPs retain activity*

Some antimicrobial peptides maintain activity when permanently tethered to resin beads. As shown in Haynie et al., some antimicrobial peptides retain activity when synthesized on and permanently tethered to polymer surfaces (Haynie, Crum et al. 1995). We sought to verify this activity for a variety of AmPs. To this end, we synthesized peptides on polystyrene resin beads with permanent linkers.

### 6.3.1.1 Methods to create permanently tethered AmPs on resin beads

Peptide synthesis described in Section 4.2.1.3 is frequently carried out on polymer resin beads. These beads present amine groups on their surface upon which to synthesize an amino acid stepwise from the C-terminus to the N-terminus. Typically, these beads have a cleavable groups that allows one to free the peptide after synthesis for purification. However, amine resins are available without cleavable groups, such as Anaspec tentagel S-NH2. After synthesis, these 90 micron beads were washed and the peptide left covalently attached to the beads by its C-terminus. For quality assurance, the same synthesis procedures was used so the peptides could be cleaved from the resin and analyzed using MALDI-TOF. A diagram of AmPs synthesized on and permanently attached to resin beads is shown in Figure 6.2.

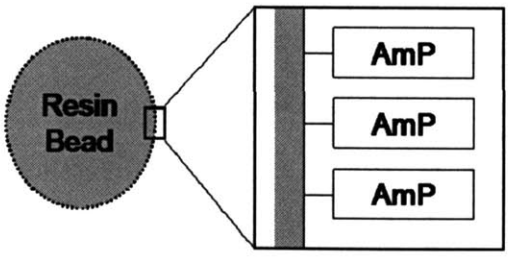



**Figure 6.2: Diagram of polymer resin beads with immobilized AmPs.** A diagram of a resin bead with peptide synthesized on the surface. These beads have bactericidal activity when suspended with target bacteria.

### 6.3.1.2 Antimicrobial activity of AmPs synthesized on resin beads

In order to evaluate the activity of tethered AmPs, an antimicrobial assay was applied in which AmP-beads were suspended with target bacteria and the viability of the bacteria was monitored over time. This resin was incubated at 4, 16, and 64 mg/ml with a suspension of bacteria at $10^5$ cfu/ml in CMHB for 1 hour and the concentration of surviving bacteria quantified by plating serial dilutions. A broad spectrum of bacteria most often associated with medical device infection was tested, including *E. coli* K12 and *S. aureus* ATCC 25923. A polymer resin with a shuffled AmP was used to ensure the resin itself does not affect cell viability.

As shown in Figure 6.3, cecropin melittin hybrid functionalized beads killed both gram-positive and gram-negative bacteria. A shuffled peptide did not display activity. Further, the supernatant from these samples was removed and tested using the same incubation procedures described above. The supernatant did not display antimicrobial activity, indicating that the killing does not result from a leached agent. Nevertheless, the harsh conditions of peptide synthesis are not suitable for medical devices. A superior approach would allow fully synthesized peptides to be covalently linked to designed functional groups on the surface of a medical device. Ideally, only mild conditions would be used.

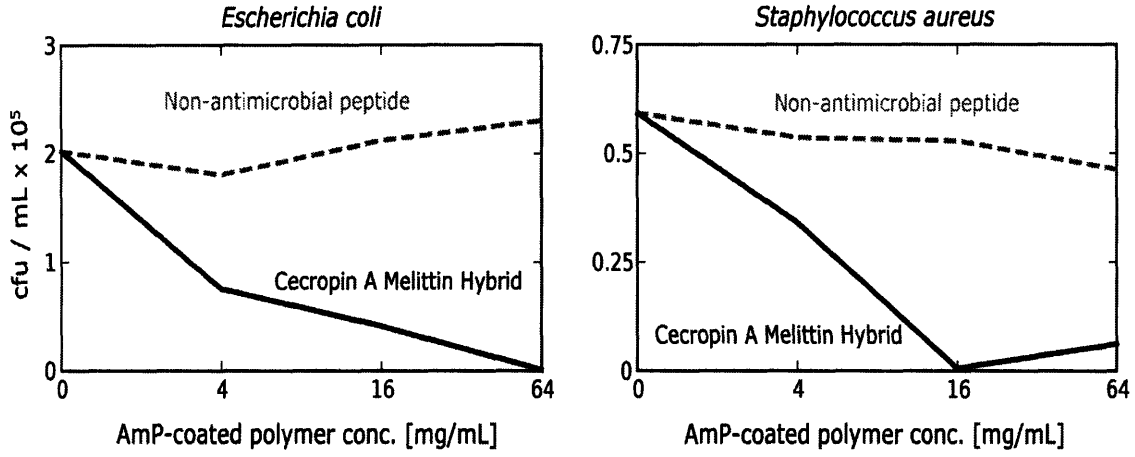



**Figure 6.3: Reduction in bacterial viability due to incubation with AmP covered resin beads.** We created a polymer resin that was covered with covalently tethered AmP. This resin was incubated with a suspension of target bacteria, K12 E. coli or ATCC 25923 S. aureus, at ~$10^5$ cfu/ml in CMHB for 1 hour, and the concentration of surviving bacteria was quantified with a standardized plating method. A non-antimicrobial control was used to ensure the polymer resin itself did not affect cell viability. The supernatant of the active beads did not show antimicrobial activity, indicating the AmP is not leached.

## *6.3.2 Chemistries for attaching AmPs to device surfaces*

As discussed in Section 6.3.1.2, immobilized AmPs retain activity, but a mild synthetic approach is required to attach fully synthesized AmPs to the surface of medical devices. Amines are a convenient functional group for attachment because they can be created on a variety of device surfaces. These amines can be converted to maleimide using sulfo-GMBS. The maleimide created may be used to couple specifically to a cysteine in the peptide. The remaining synthetic methods for creating immobilized AmPs were developed and documented by L Ferreira and T Squier.

### 6.3.2.1 Coupling AmPs via sulfo-GMBS

First, we sought to couple AmPs to a surface similar to that used in Section 6.3.1.2. For this reason, we used the same NH2-microparticles (TentaGel S-NH2 resin). Thus, reattached peptides could be compared to peptides synthesized on beads so that similar bead sizes and surface density could be achieved. A single cysteine was incorporated at the C-terminal of the cecropin-melittin peptide to produce a similar orientation to the peptide synthesized on beads.

First, the number of free amino groups was quantified using the ninhydrin assay. Approximately 6.7 mg of microparticles were suspended in 1 mL of 1 M acetate buffer pH 5.0 containing 12.5 mg of ninhydrin (Sigma). The suspension was kept in boiling water for 15 min. After 15 min the sample was removed and 15 mL of an ethanol/water

mixture (1/1, v/v) was added. The reaction mixture was allowed to cool to room temperature for 1 hour, away from light. Ninhydrin reacts with free amino groups and creates a blue water-soluble compound. The amount of free amino groups in the beads was spectrophotometrically determined by measuring the absorbance of the supernatant at 570 nm, after the 1 hour cooling time. Glycine was used as a reference material.

Peptides were coupled using sulfo-GMBS chemistry. 3.4 mg of sulfo-GMBS was reacted with 15 mg of NH2-microparticles suspended in 0.5 mL PBS buffer having a pH of 7.4 at room temperature for two hours with mild agitation (vortex, 100 rpm). After two hours, the beads were centrifuged for two minutes at 2500 rpm and washed five times with 1 mL of PBS buffer. In the last wash, the beads were re-suspended in 0.5 mL PBS buffer and reacted with 5 mg of a cysteine-incorporating Cecropin-Melittin hybrid peptide, overnight, at room temperature with mild agitation (100 rpm). The beads were again washed 5 times with 0.5 mL PBS buffer and then re-suspended in 1 mL of PBS and kept at 4 °C overnight. The following morning the supernatant was removed, and the beads were washed 5 times with 1mL of PBS buffer. The beads were re-suspended in 1mL and stored at 4°C. The peptide immobilized in the beads was determined by BCA assay (Sigma), using cecropin melittin as the standard. The amount of peptide bound to beads was determined indirectly from the difference between the initial total peptide exposed to the beads and the amount of peptide recovered in the several washes. The concentration of peptide bound to the beads was approximately 0.91 mg per 15 mg of beads, which corresponds to 0.060 mg of peptide per 72.7 $mm^2$ of bead surface area, assuming the bead is non-porous.

#### 6.3.2.2 Antimicrobial activity of AmPs attached via sulfo-GMBS

The peptide conjugated beads were tested against *E. coli* by incubating with 1 x $10^7$ cfu/ml K12 *E. coli* in CMHB which had been stained with 30 μM propidium iodide and 6 μM SYTO9 stains from a standard Molecular Probes LIVE/DEAD kit. As determined with a fluorescence microscope, 50% of the bacteria in solution were killed after one hour, whereas no killing occurred with control beads that had shuffled peptides. To assess whether the killing effect was due to the immobilized peptide, the medium that was incubated with the beads was centrifuged at 3000g for 2 minutes, the supernatant was removed, and the supernatant was inoculated with 1 x $10^7$ cfu/ml stained *E. coli* in CMHB for 1 hour. No killing was observed. This indicates that the immobilized peptide is the effective component against bacteria.

### *6.3.3 Reattaching AmPs to membranes*

While the experiments described in Section 6.3.2.2 demonstrated the antimicrobial activity of an AmP attached via sulfo-GMBS, the assay is not well suited to quantitative measurements. Different samples had various levels of bead agglomeration, so the exposed surface area may not be consistent. A more appropriate model for the interaction of a device surface with bacteria would be a flat functionalized membrane. Additionally, standardized assays exist for non-leaching antimicrobial surfaces which may be applied for a flat functionalized surface.

### 6.3.3.1 Methods to attach AmPs to a cellulose membrane

Since the sulfo-GMBS chemistry reacts with an amine on the device surface, we chose to immobilize on a membrane with terminal amine groups that is used for the solid state synthesis of peptides (Intavis Product number 30.100). This surface presents 0.340 μmoles of NH2 per cm2, as determined by a picric acid assay. The terminal amine groups of the membrane were reacted with the succinimide groups of sulfo-GMBS and in a subsequent step the maleimide groups of sulfo-GMBS was reacted with the thiol groups of the cysteine containing cecropin melittin peptide. The amount of peptide bound to the membrane was determined indirectly from the difference between the initial total peptide exposed to the beads and the amount of peptide recovered in the several washes. The quantity of immobilized peptide was approximately 2.0 mg per $cm^2$ of membrane. This peptide-conjugated membrane was tested for immobilized bactericidal activity against *E. coli* ATCC 2592.

### 6.3.3.2 Antimicrobial activity of AmP functionalized membranes

The standardized method for evaluating a non-leaching antimicrobial surfaces is described in ASTM 2149. Given materials constraints, we scaled this assay down to smaller volumes. The modified assay used was as follows: an overnight culture of a target bacteria in a growth medium such as Cation Adjusted Mueller Hinton Broth, is diluted to approximately $1x10^5$ cfu/ ml in pH 7.4 Phosphate Buffered Saline using a predetermined calibration between OD600 and cell density. A 0.5 $cm^2$ sample of

immobilized antimicrobial surface is added to 0.75 ml of the bacterial suspension. The sample should be covered by the liquid and incubated at 37°C with a sufficient amount of mixing that the solid surface is seen to rotate through the liquid. After 1 hour of incubation, serial dilutions of the bacterial suspension are plated on agar plates and allowed to grow overnight for quantifying the viable cell concentration. Using this procedure, the cecropin melittin conjugated membrane produced a 4.2-log reduction of *E. coli* in solution over 1 h. Testing the amine-functionalized membrane without an antimicrobial peptide conjugated to it for immobilized bactericidal activity did not show a significant reduction in viable bacteria (<0.1 log reduction).

### 6.3.4 Killing through multiple bacterial challenges

Samples identical to those generated in Section 6.3.3.1 were synthesized and stored at 4° C in pH 7.4 PBS for more than three weeks. When this peptide-conjugated membrane was tested against for immobilized bactericidal activity against *E. coli* using methods given in Section 6.3.3.2, an average of a 1.8-log reduction of bacteria in solution occurred over 1 h. The samples were then removed from the testing solution, and placed in fresh PBS. Samples underwent 10 minutes of ultrasonication, were switched to fresh PBS, and underwent an additional 30 minutes of sonication. They were then rinsed and retested for bactericidal activity against *E. coli* ATCC 25922. The washed samples produced an average of a 3.3-log reduction in viable bacteria over 1 hour. Therefore, the active agent is not consumed during the one hour assay, and repeated killing is possible.

### 6.3.5 Confirmation that activity results from non-leached material

In order to convince clinicians and the FDA that activity results from the immobilized agent, rather than a leached material, a variety of studies were undertaken. As described in Section 6.3.1.2, direct observation of the interaction between the antimicrobial surface and bacteria stained with LIVE/DEAD viability stain demonstrated killing within minutes of bacteria contacting the surface. This could, however, result from locally released AmPs. One reason this may be unlikely is that many of the samples that demonstrate killing had been washed in PBS in excess of 10 times, well beyond the point that leachable AmP could be detected using a BCA assay. Additionally, some samples were stored for more than 3 weeks before use in an assay which lasted just one hour. While bacteria were not present during this storage period, much of the non-covalently attached material would have been released before the assay. Nevertheless, a standardized test for leached agent from a permanently immobilized surface was carried out, which is described below.

After performing the antimicrobial assay described in Section 6.3.3.2, the supernatant was removed for characterization during both rounds of killing before and after washing. At the end of the 1 hour incubation between the sample and a solution of bacteria, 0.4 ml of bacterial solution was removed. The 0.4 ml was centrifuged at 3000g for 5 minutes to remove remaining bacteria. A sample of 0.2 ml of supernatant was

removed and added to 0.05 ml of *E. coli* ATCC 25922 at 5 x $10^5$ cfu/ml, giving a final concentration of 1 x $10^5$ cfu/ml, as in the standard antibacterial assay. This mixture was incubated at 37°C with the same degree of mixing as in the immobilized bactericidal activity assay, and serial dilutions were plated at the end of 1 hour.

The supernatant from both the 1st and 2nd rounds of killing in Sections 6.3.4 did not show a measurable amount of killing (< 0.1-log reduction in viable bacteria). Because the surface demonstrated killing, but the supernatant above the surface does not demonstrate any killing, the immobilized antimicrobial surface is substantially non-leaching.

## 6.4 Improved device coatings

In the previous section, permanently immobilized AmPs as antimicrobial surfaces were described. In addition to demonstrating broad spectrum activity, it would be desirable to achieve the maximal activity per amount of peptide immobilized (specific activity). This would reduce manufacturing costs and may lessen concerns over toxicity. In the following sections, a strategy to optimize activity by controlling the orientation of peptide during immobilization is described. Further, an alternative hydrogel formulation is presented which may have superior properties in certain clinical settings. Finally, the procedure for evaluating the hemolytic activity of the immobilized antimicrobial surface is given, with the surface proving to be largely non-hemolytic.

### 6.4.1 Effect of orientation on specific activity

In the two patent applications that describe reattachment of an AmP to a polymer surface, the peptide is attached via any amine group (US Patent Applications 20040126409, 20050065072). AmPs tend to have multiple positively charged amines as side chains, which contribute to their activity. Using a chemistry that links to any amine, such as 1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (EDC) chemistry, will tether an AmP with a random distribution of orientations, depending on which amine is linked. Additionally, a given AmP may be tethered through multiple amines. Some orientations of peptide, or peptides tethered by multiple sites may have reduced bactericidal activity.

A superior strategy would be to immobilize the peptide on the surface of the substrate in such a manner that the portion of the peptide presented to interact with bacteria upon exposure is uniform for all immobilized molecules of a given peptide. Ideally, the location of the single tethering site within the peptide should be controlled to maximize activity. Alternatively, multiple attachment residues within the same region of the peptide could yield a single orientation. Typically, the N-terminus of the peptide should be presented to target cells for highest activity, although this may vary depending on the peptide. The sulfo-GMBS described has the advantage that the orientation can be controlled by placing only a single cysteine within the peptide being tethered. Since most linear AmPs do not contain cysteines, allowing only one cysteine is not significantly restrictive for peptide design.

### 6.4.1.1 Methods to generated AmP surfaces without orientation

In order to assess the effect of orientation on activity, samples of membranes with oriented and non-oriented peptide at similar surface densities would be desired. A sample with non-oriented peptide tethered through any amine groups was created on the same cellulose surface used in Section 6.3.3.1. The following protocol was followed: A cellulose membrane (1×1 $cm^2$) containing terminal amine groups was incubated with a solution of Methyl N-succinimidyl adipate (MSA, Pierce) (1.54 mg in 0.1 mL of a solution of DMSO in PBS pH 7.4 (1:9, v/v)) for 2 h, at room temperature. The membrane was then washed several times (5 × 1 mL) with PBS and incubated in phosphate buffer pH 9.5 (2 mL) overnight. After that time, the membrane was washed with PBS pH 7.4 (5 × 1 mL) and 0.1M citrate buffer pH 7.5 (5 × 1 mL). The majority of the terminal amine groups of the cellulose membrane reacted with MSA. The content of amine groups was 0.340 μmol/cm2 and 0.039 μmol/cm2, before and after MSA reaction, respectively.

The terminal COOH groups of MSA were coupled with the terminal NH2 groups of the antimicrobial peptide using EDC chemistry. The membrane was reacted with 0.5 mL of N-(3-dimethylaminopropyl)-N´-ethyl-carbodiimide hydrochloride (EDC) solution (4.8 mg/mL in 0.1 M sodium citrate buffer pH 5.0) for 30 minutes and afterwards washed with PBS (3 × 1 mL). The activated membrane was subsequently reacted with a cysteine-incorporating Cecropin-Melittin hybrid peptide (5 mg in 1 mL of PBS), overnight, at room temperature, with mild agitation (100 rpm), and finally washed with PBS (10 times, 1 mL washes) and kept in PBS, 4°C, until use. The peptide immobilized on the membrane was determined by the BCA assay (Sigma), using free peptide as a standard.

The amount of peptide bound to the membrane was calculated from the difference between the initial total peptide exposed to the membrane and the amount of peptide recovered in the several washes.

### 6.4.1.2 Methods to generate AmP surfaces without orientation

We sought to assess the specific activity for oriented versus non-oriented AmPs immobilized on surfaces. To this end, samples were created with methods given in Section in 6.3.3.1 and 6.4.1.1. For non-oriented AmP, the peptide content was 1.84 ± 0.27 mg/cm$^2$ (n=2). The content of peptide per surface area was similar to the one immobilized using an oriented peptide (sulfo-GMBS chemistry) (1.75 ± 0.36 mg per cm2, n=3). Both oriented and non-oriented peptides with similar surface densities were evaluated for immobilized bactericidal activity against *E. coli* ATCC 25922 using methods given in Section 6.3.3.2. The oriented peptide produced a 3.0-log reduction in viable bacteria, whereas the non-oriented peptide produced only a 1.6-log reduction. This shows that the immobilization of an antimicrobial peptide in an oriented way creates a higher specific biological activity. A higher specific activity lowers costs and reduces concern over toxicity.

#### 6.4.1.3 Activity of oriented AmP at lower concentrations

A cysteine-incorporating Cecropin-Melittin hybrid peptide was immobilized to the amine presenting cellulose membrane with the sulfo-GMBS chemistry as described in Section 6.3.3.1, with the exception that the concentration of peptide in solution during immobilization was varied from 0.125 mg/ml to 5.0 mg/ml. Samples were assayed for immobilized bactericidal activity as described in Section 6.3.3.2. When a concentration of 5 mg/ml was used during immobilization, the resulting surface produced a 2.0-log reduction of *E. coli* ATCC 25922 in 1 hour. However, when the concentration of peptide during immobilization was reduced to 0.125 mg/ml, a 1.8-log reduction still occurred, which is at a significantly lower density than the non-oriented peptide in Example 6. Thus, a greater immobilized bactericidal activity is achieved per mass of peptide used when the peptide is oriented (higher specific activity)

### *6.4.2 Hydrogel formula*

Hydrogels have a number of appropriate properties for coating medical devices. They can reduce biofouling and improve biocompatibility (Norton, Tegnell et al. 2005). In addition, a wide variety of chemistries may be used to create hydrogels, which may be designed to present attachment sites for AmPs. The design and evaluation of an AmP functionalized dextran hydrogel is described below.

### 6.4.2.1 Production of hydrogels

Dextran gels were prepared by the UV crosslinking of dextran acrylate macromonomer. Dextran-acrylate with a degree of substitution of 23.3% (400 mg) (see Ferreira et al., Biomaterials 2002, 23, 3957-3967 for details, in preparation) was dissolved in PBS (1.8 ml) and Irgacure (5 mg/ml, 250 μL) was gently mixed into the solution. Cross-linking of the solution was initiated by exposure UV-light over a 10 minute period. The resulting gel was cut into several disks (8 mm diameter) using a biopsy punch, and washed overnight in water. Prior to the functionalization reaction, each dextran disk was soaked in 95% ethanol for 20 minutes, shrinking the gel. Then, the shrunken gel was soaked in a solution of sodium periodate (5.3 mg/ml, 1 ml) in PBS, for 1 hour with mild agitation (vortex, 100 rpm). After this time the disk was washed (5 times) in PBS to remove any un-reacted sodium periodate. The disk was then placed in a solution of ethylene diamine dihydrochloride (66 mg/ml, 1 ml) and the reaction was allowed to continue for 1 ½ hours with mild agitation (vortex, 100 rpm). After this step the disk was rinsed thoroughly (5 times) in PBS. A solution of sodium cyanoborohydride (15 mg/mL, 1 ml) was prepared in PBS and allowed to cool to room temperature for 10 minutes after mixing. The disk was allowed to react, without agitation, in the sodium cyanoborohydride solution for 30 minutes followed by thorough rinsing and overnight soaking in PBS. The functionalized gel was soaked in 95% ethanol for 20 minutes followed by soaking in a sulfo-GMBS (10 mg/ml, 0.4 ml) solution for 2 hours at room temperature, with mild agitation (vortex, 100 rpm). Excess sulfo-GMBS was removed by rinsing with PBS (5 times). The disk was then soaked again in 95% ethanol for 2 minutes

followed by soaking in a cysteine-incorporating Cecropin-Melittin hybrid peptide solution (5 mg/ml) overnight, at room temperature, with mild agitation (vortex, 100 rpm). The disk was washed 10 times (0.5 ml, PBS), over a 2 day period, and the washings were kept for the determination of peptide released. BCA assay showed 3.22 mg of peptide was immobilized on the dextran disk.

#### 6.4.2.2 Activity of hydrogels

The solid dextran gels were assayed for antimicrobial activity in Section 6.3.3.2. Approximately 0.5 cm2 samples of gel were used at the same agitation as for the cellulose samples, which was sufficient to suspend the gel throughout the assay. A gel functionalized with a cysteine-incorporating Cecropin-Melittin hybrid peptide demonstrated a 2.9-log reduction in E. coli ATCC 25922, whereas a gel without Cecropin-Melittin hybrid peptide did not display a significant reduction in viable bacteria (< 0.1-log).

### *6.5.3 Hemolytic activity*

For blood contacting devices, a coating must be non-hemolytic. Neither cecropin melittin nor the tethering approach have been optimized for hemocompatibility. Nevertheless, a cysteine-incorporating Cecropin-Melittin hybrid peptide was immobilized to the amine presenting cellulose membrane with the sulfo-GMBS chemistry as described

in Section 6.3.3.1 was tested for hemolytic activity. A stock of 10% washed pooled red blood cells (Rockland Immunochemicals Inc, Gilbertsville, PA) was diluted to 0.25% with a hemolysis buffer of 150 mM NaCl and 10 mM Tris at pH 7.0. A 0.5 $cm^2$ antimicrobial sample was incubated with 0.75 ml of 0.25% red blood cell suspension for 1 hour at 37°C. The solid sample was removed and cells spun down at 6000 g, the supernatant removed, and the OD414 measured on a spectrophotometer. Total hemolysis was defined by diluting 10% of washed pooled red blood cells to 0.25% in sterile DI water and incubating for 1 hour at 37°C, and 0% hemolysis was defined by a suspension of 0.25% red blood cells in hemolysis buffer without a solid sample. The peptide immobilized sample produced only 4.95% hemolysis using this assay, demonstrating that the sample is a substantially non-hemolytic surface. With optimization of the AmP sequence or coating parameters, it may be possible to further reduce hemolysis.

## 6.6 Commercial potential

A team was formed to evaluate the potential of immobilized AmPs as a commercial coating for medical devices. Through extensive conversations with users and industry, a business plan was developed. The societal impact and commercial opportunities of our technology in addressing medical device infections has generated excitement in the entrepreneurship and venture capital communities. In the previous year, the following competitions recognized the potential of our technology:

- First Place, MIT 100K Venture Capital Competition 2006 (164 entries)

- First Place, Harvard GSAS Biotechnology Business Plan Competition
- First Place, Cambridge-MIT Pitch Competition
- First Place, Oxford University Business Plan Competition (117 entries from 10 countries)

In light of this success, we incorporated a company, SteriCoat, in July 2006 to exploit this technology. Development work moving towards *in vivo* studies is described in Future Work in Section 7.2.2.

# Chapter 7

# Conclusions and recommendations

## 7.1 Summary and conclusions

A successful approach to create diverse AmPs was demonstrated. First, a high-throughput method to produce AmPs was developed using *in vitro* translation. This approach is automatable and may produce agents that would be toxic to recombinant hosts. The cost of the system was reduced and yield optimized through a series of improvements. An oligonucleotide synthesis scheme that mimicked pattern-based design allowed the reuse of primers and the study of pattern synergy. Additionally, a robust batch system was developed and yield maximized by optimizing the fusion partners co-translated with the AmPs (Loose, Langer et al. 2007). Finally, an effective purification scheme was developed so that components of the *in vitro* mixture would not impact MIC measurements.

Methods to design AmPs based on semi-conserved patterns were evaluated. Nearly 50% of the designed sequences were active against target bacteria tested at 256 ug/ml or below. This represents a 10-fold enrichment over shuffled control peptides which had the same bulk physiochemical properties but were not homologous to antimicrobial patterns. Further, two lead sequences were found which had MICs of 64

ug/ml against *E. coli* and 16 ug/ml against *B. cereus.* This activity was extended to critical targets including *S. aureus* and *B. anthracis* (Loose, Jensen et al. 2006).

We also demonstrated that AmP leads that emerged from pattern-based design could be optimized through a variety of heuristic approaches. Broad spectrum activity was improved, with an optimized sequence having MICs of 16 ug/ml against *E. coli* and 8 ug/ml against *B. cereus* and 4 ug/ml against *S. aureus.* Additionally, the groundwork was laid for improved design by characterizing the antimicrobial activity and toxicity of a representative set of natural, linear AmPs. This information may enable an improved scoring metric, leading to more active designs, as described in Section 7.2.1. We also systematically showed that amidated AmPs had stronger activities and higher therapeutic indices than free acid peptides.

Finally, a series of immobilized AmP coatings were created for medical devices. Through interactions with clinicians, it was found that infections were frequently associated with catheters and implants. AmPs have suitable characteristics for medical device coatings because they remain active when tethered to the device surface, potential increasing active lifespan and lowering toxicity. Proof-of-concept surfaces we created demonstrated broad spectrum activity through repeated bacterial challenges without significant hemolytic activity.

Collectively, this work leads to the following conclusions:

- *In vitro* translation is a suitable platform for the creation of a diverse set of peptides that may be toxic to a recombinant host.

- Translation yield may be improved and functionality enhanced by the choice of an appropriate fusion partner.
- Further, designing the fusion partner and peptide to be the optimal length, approximately 100 amino acids, maximizes translation yield while providing stability from proteases. This was demonstrated for multiple sequences and translation systems.
- A high-throughput platform for the *in vitro* translation and evaluation of AmPs may be enabled given an appropriate purification scheme.
- Natural AmPs contain hundreds of semi-conserved motifs in their sequences that are specific to AmPs relative to other peptides and proteins in SWISS-PROT.
- Further, these semi-conserved motifs may be assembled in novel combinations to produce active AmPs with a high probability of success.
- Diverse leads emerging from a pattern based design approach may be optimized to give broad spectrum agents.
- Amidation of natural peptides improves their antimicrobial activity and therapeutic index.
- Immobilized AmPs retain broad-spectrum activity and are effective against repeated bacterial challenges.
- While the clinical use of AmPs has been limited due to peptide cost, toxicity, and instability, an improved design methodology may yield sequences that have superior properties for development. Additionally, innovative strategies

for the delivery or formulation of AmPs may also bypass obstacles to clinical development.

## 7.2 Recommendations and future studies

The pattern-based design presented here was executed without any knowledge of the antimicrobial activity and toxicity of members of the training set. Improved design would likely be possible if this information was incorporated into the pattern discovery or scoring portion of the AmP design. A portion of the data for natural peptides was created as described in Section 5.3. This data should be supplement, and improved scoring algorithms hypothesized and tested, as described below.

### *7.2.1 Iterative optimization of AmPs*

In the work presented in this thesis, 50% of designs were active against *E. coli* or *B. cereus*. Nevertheless, 70% of natural AmPs were found not to be active against each target bacteria at 256 ug/ml or below. Given activity and toxicity data, a weighting system could be developed for each motif based on the characteristics of the peptides from which the motif is derived. These motif weights should provide a useful metric for scoring designed sequences. An iterative approach could be used in which each set of designed peptides is characterized and active peptides are added to the training set before discovery and design are executed again. After many rounds of design and analysis, one

would hope to enrich the space of highly active and non-toxic sequences. An overview of this procedure is given in Figure 7.1.

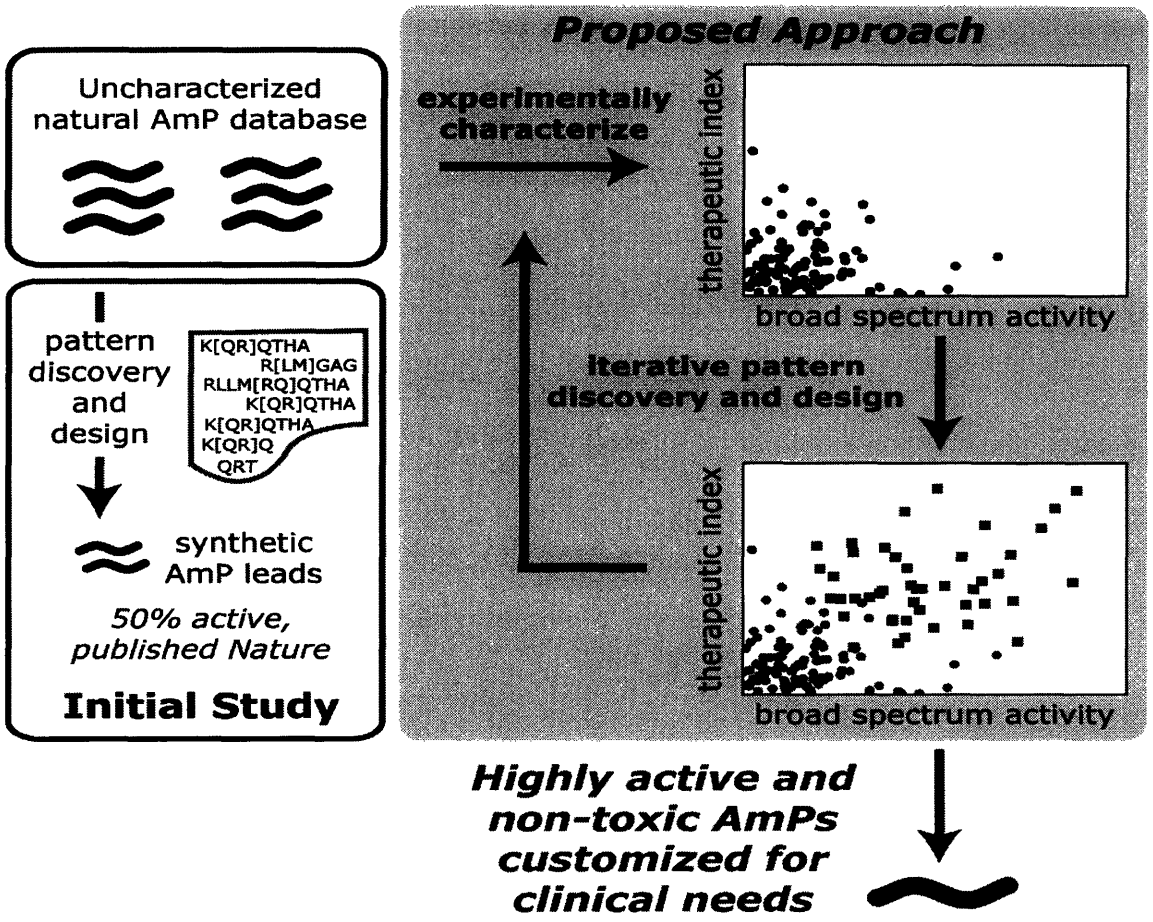



**Figure 7.1: Proposed method for discovery and evaluation of antimicrobial peptides.** In our initial study, we applied pattern discovery to a sequence database of natural AmPs, designed synthetic AmP leads, and demonstrated activity for 50% of these designed sequences. However, preliminary experimental characterization of this sequence database has shown only a small fraction of these AmPs as non-toxic and active against clinical bacterial targets. To improve activity an upgraded database of experimentally characterized peptides should be created. Furthermore, using this improved training set, an iterative pattern discovery and design approach that leverages our accumulating knowledge of peptide activity could produce highly active and non-toxic AmPs customized for clinical needs. Figure designed with JF Moxley.

### 7.2.2 *In vivo* activity and toxicity

In Chapter 6, a number of active and non-hemolytic medical device coatings comprised of immobilized AmPs were presented. In order to further develop this technology, animal biocompatibility of one of these active surfaces should be characterized as soon as possible. Additionally, the activity should be improved by optimizing the AmP sequence used and altering coating parameters including tethering length and density. Finally, the *in vivo* antimicrobial activity should be quantified in accepted models of medical device infection.

# Chapter 8

# Appendix

## 8.1 Sequences for *in vitro* translation studies

### Truncated GFP constructs

**35 amino acid truncation of GFP (Bbs1 linker underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTTCA
CAGAA*TAAT*GG<u>GTCTTC</u>A**TAATAG**ATTCCAGCACACTGGCGGCCGTTACTAGTGGAT
CCGGCTGCTAACAAAGCCCGAAAGGAAGCGAGTTGGCTGCTGCCACCGCTGAGCAA
TAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAA
GGAGGAACTATATCCGGATATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTC
GATAGTGGCTCCAAGTAGCGAA

**43 amino acid truncation of GFP (Bbs1 linker underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACCTCACAGAA*TAAT*GG<u>GTCTTC</u>A**TAATAG**ATTCCAGCA
CACTGGCGGCCGTTACTAGTGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCGAG
TTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGG
GTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATATCCACAGGACGGG
TGTGGTCGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAGCGAA

**57 amino acid truncation of GFP (Bbs1 linker underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGGTCACAGAA*TAAT*GG<u>GTCTTC</u>A**TAATAG**ATTCCAGCACACTGGCGGCCGTTAC

TAGTGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCGAGTTGGCTGCTGCCACCG
CTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTT
TGCTGAAAGGAGGAACTATATCCGGATATCCACAGGACGGGTGTGGTCGCCATGATC
GCGTAGTCGATAGTGGCTCCAAGTAGCGAA

**70 amino acid truncation of GFP (Bbs1 linker underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCAATGCTCACAGAA*TAAT*GG
GTCTTCA**TAATAG**ATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGGCTGCTAA
CAAAGCCCGAAAGGAAGCGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATA
ACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTAT
ATCCGGATATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTCGATAGTGGCTC
CAAGTAGCGAA

**82 amino acid truncation of GFP (Bbs1 linker underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCAATGCTTTTCAAGATACCC
GGATCATATGAAACGGCATGACTCACAGAA*TAAT*GG**GTCTTCA**TAATAGATTCCAGC
ACACTGGCGGCCGTTACTAGTGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCGA
GTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACG
GGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATATCCACAGGACGG
GTGTGGTCGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAGCGAA

**95 amino acid truncation of GFP (Bbs1 linker underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCAATGCTTTTCAAGATACCC
GGATCATATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACA
GGAATCACAGAA*TAAT*GG**GTCTTCA**TAATAGATTCCAGCACACTGGCGGCCGTTACT
AGTGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCGAGTTGGCTGCTGCCACCGC
TGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTT
GCTGAAAGGAGGAACTATATCCGGATATCCACAGGACGGGTGTGGTCGCCATGATC
GCGTAGTCGATAGTGGCTCCAAGTAGCGAA

**139 amino acid truncation of GFP (Bbs1 linker underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT

GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCAATGCTTTTCAAGATACCC
GGATCATATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACA
GGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAAGTCA
AGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAG
AAGATGGAAACATTCTTGGACACTCACAGAA*TAAT*GGGTCTTCA**TAATAG**ATTCCAG
CACACTGGCGGCCGTTACTAGTGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCG
AGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAAC
GGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATATCCACAGGACG
GGTGTGGTCGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAGCGAA

**203 amino acid truncation of GFP (Bbs1 linker underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCAATGCTTTTCAAGATACCC
GGATCATATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACA
GGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAAGTCA
AGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAG
AAGATGGAAACATTCTTGGACACAAATTGGAATACAACTATAACTCACACAATGTAT
ACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGACAC
AACATTGAAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATT
GGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCCACATCACAGAA*TAAT*G
G<u>GTCTTC</u>A**TAATAG**ATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGGCTGCTA
ACAAAGCCCGAAAGGAAGCGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCAT
AACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTA
TATCCGGATATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTCGATAGTGGCT
CCAAGTAGCGAA

**250 amino acid truncation of GFP (Bbs1 linker underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCAATGCTTTTCAAGATACCC
GGATCATATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACA
GGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAAGTCA
AGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAG
AAGATGGAAACATTCTTGGACACAAATTGGAATACAACTATAACTCACACAATGTAT
ACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGACAC
AACATTGAAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATT
GGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTT
CGAAAGATCCCAACGAAAAGAGAGACCACATGGTCCTTCTTGAGTTTGTAACAGCTG
CTGGGATTACACATGGCATGGATGAACTGTACCAACCCTCACAGAA*TAAT*GG<u>GTCTTC</u>
A**TAATAG**ATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGGCTGCTAACAAAGC
CCGAAAGGAAGCGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTT
GGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGA

TATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAG
CGAA

## Truncated GFP constructs with S1

**35_S1 (Bbs1 linker underlined, S1 bold and underlined)**

ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTTCA
CAGAA*TAAT***GATGATGACGATAAGAATAAAGTTAAGAAACCACTTACAGGAGCAC**
**ATAGACTATTGTTCACTTTTTTATTTGTATAATAGAA***CCCC*GGGTCTTC**ATAATAG**
ATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGGCTGCTAACAAAGCCCGAAAG
GAAGCGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCT
CTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATATCCAC
AGGACGGGTGTGGTCGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAGCGAA

**43_S1 (Bbs1 linker underlined, S1 bold and underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACCTCACAGAA*TAAT***GATGATGACGATAAGAATAAAGT**
**TAAGAAACCACTTACAGGAGCACATAGACTATTGTTCACTTTTTTATTTGTATAA**
**TAGAA***CCCC*GGGTCTTC**ATAATAG**ATTCCAGCACACTGGCGGCCGTTACTAGTGGAT
CCGGCTGCTAACAAAGCCCGAAAGGAAGCGAGTTGGCTGCTGCCACCGCTGAGCAA
TAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAA
GGAGGAACTATATCCGGATATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTC
GATAGTGGCTCCAAGTAGCGAA

**57_S1 (Bbs1 linker underlined, S1 bold and underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGGTCACAGAA*TAAT***GATGATGACGATAAGAATAAAGTTAAGAAACCACTTAC**
**AGGAGCACATAGACTATTGTTCACTTTTTTATTTGTATAATAGAA***CCCC*GGGTCTT
C**ATAATAG**ATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGGCTGCTAACAAAG
CCCGAAAGGAAGCGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCT
TGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGG
ATATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTA
GCGAA

**70_S1 (Bbs1 linker underlined, S1 bold and underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT

GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCAATGCTCACAGAA*TAAT***GA**
**TGATGACGATAAGAATAAAGTTAAGAAACCACTTACAGGAGCACATAGACTATT**
**GTTCACTTTTTTATTTGTATAATAGAA***CCCC*GGGTCTTC**ATAATAG**ATTCCAGCAC
ACTGGCGGCCGTTACTAGTGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCGAGT
TGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGG
TCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATATCCACAGGACGGGT
GTGGTCGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAGCGAA

**82_S1 (Bbs1 linker underlined, S1 bold and underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCAATGCTTTTCAAGATACCC
GGATCATATGAAACGGCATGACTCACAGAA*TAAT***GATGATGACGATAAGAATAAAG**
**TTAAGAAACCACTTACAGGAGCACATAGACTATTGTTCACTTTTTTATTTGTATA**
**ATAGAA***CCCC*GGGTCTTC**ATAATAG**ATTCCAGCACACTGGCGGCCGTTACTAGTGGA
TCCGGCTGCTAACAAAGCCCGAAAGGAAGCGAGTTGGCTGCTGCCACCGCTGAGCA
ATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAA
AGGAGGAACTATATCCGGATATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGT
CGATAGTGGCTCCAAGTAGCGAA

**95_S1 (Bbs1 linker underlined, S1 bold and underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCAATGCTTTTCAAGATACCC
GGATCATATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACA
GGAATCACAGAA*TAAT***GATGATGACGATAAGAATAAAGTTAAGAAACCACTTACA**
**GGAGCACATAGACTATTGTTCACTTTTTTATTTGTATAATAGAA***CCCC*GGGTCTTC
**ATAATAG**ATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGGCTGCTAACAAAGC
CCGAAAGGAAGCGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTT
GGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGA
TATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAG
CGAA

**139_S1 (Bbs1 linker underlined, S1 bold and underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCAATGCTTTTCAAGATACCC
GGATCATATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACA
GGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAAGTCA
AGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAG

AAGATGGAAACATTCTTGGACACTCACAGAA*TAAT***GATGATGACGATAAGAATAAA**
**GTTAAGAAACCACTTACAGGAGCACATAGACTATTGTTCACTTTTTTATTTGTAT**
**AATAGAA***CCCC*GGGTCTTC**ATAATAG**ATTCCAGCACACTGGCGGCCGTTACTAGTGG
ATCCGGCTGCTAACAAAGCCCGAAAGGAAGCGAGTTGGCTGCTGCCACCGCTGAGC
AATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGA
AAGGAGGAACTATATCCGGATATCCACAGGACGGGTGTGGTCGCCATGATCGCGTA
GTCGATAGTGGCTCCAAGTAGCGAA

**203_S1 (Bbs1 linker underlined, S1 bold and underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCAATGCTTTTCAAGATACCC
GGATCATATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACA
GGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAAGTCA
AGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAG
AAGATGGAAACATTCTTGGACACAAATTGGAATACAACTATAACTCACACAATGTAT
ACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGACAC
AACATTGAAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATT
GGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCCACATCACAGAA*TAATG*
**ATGATGACGATAAGAATAAAGTTAAGAAACCACTTACAGGAGCACATAGACTAT**
**TGTTCACTTTTTTATTTGTATAATAGAA***CCCC*GG**GTCTTC**A**TAATAG**ATTCCAGCA
CACTGGCGGCCGTTACTAGTGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCGAG
TTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGG
GTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATATCCACAGGACGGG
TGTGGTCGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAGCGAA

**250_S1 (Bbs1 linker underlined, S1 bold and underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCAATGCTTTTCAAGATACCC
GGATCATATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACA
GGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAAGTCA
AGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAG
AAGATGGAAACATTCTTGGACACAAATTGGAATACAACTATAACTCACACAATGTAT
ACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGACAC
AACATTGAAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATT
GGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTT
CGAAAGATCCCAACGAAAAGAGAGACCACATGGTCCTTCTTGAGTTTGTAACAGCTG
CTGGGATTACACATGGCATGGATGAACTGTACCAACCCTCACAGAA*TAATG***ATGATG**
**ACGATAAGAATAAAGTTAAGAAACCACTTACAGGAGCACATAGACTATTGTTCA**
**CTTTTTTATTTGTATAATAGAA***CCCC*GG**GTCTTC**A**TAATAG**ATTCCAGCACACTGGC
GGCCGTTACTAGTGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCGAGTTGGCTG
CTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGA

GGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATATCCACAGGACGGGTGTGGTC
GCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAGCGAA

## Truncated GFP constructs with stop codons inserted before S1

**35_S1_stop_ins (Bbs1 linker underlined, S1 bold and underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGT**TAA**
**TAG**AA*TAAT***GATGATGACGATAAGAATAAAGTTAAGAAACCACTTACAGGAGCA**
**CATAGACTATTGTTCACTTTTTTATTTGTATAATAGAA***CCCC*GGGTCTTC**ATAATA**
**G**ATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGGCTGCTAACAAAGCCCGAAA
GGAAGCGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCC
TCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATATCCA
CAGGACGGGTGTGGTCGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAGCGAA

**43_S1_stop_ins (Bbs1 linker underlined, S1 bold and underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACC**TAATAG**AA*TAAT***GATGATGACGATAAGAATAAAGT**
**TAAGAAACCACTTACAGGAGCACATAGACTATTGTTCACTTTTTTATTTGTATAA**
**TAGAA***CCCC*GGGTCTTC**ATAATAG**ATTCCAGCACACTGGCGGCCGTTACTAGTGGAT
CCGGCTGCTAACAAAGCCCGAAAGGAAGCGAGTTGGCTGCTGCCACCGCTGAGCAA
TAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAA
GGAGGAACTATATCCGGATATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTC
GATAGTGGCTCCAAGTAGCGAA

**57_S1_stop_ins (Bbs1 linker underlined, S1 bold and underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGG**TAATAG**AA*TAAT***GATGATGACGATAAGAATAAAGTTAAGAAACCACTTAC**
**AGGAGCACATAGACTATTGTTCACTTTTTTATTTGTATAATAGAA***CCCC*GGGTCTT
C**ATAATAG**ATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGGCTGCTAACAAAG
CCCGAAAGGAAGCGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCT
TGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGG
ATATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTA
GCGAA

**70_S1_stop_ins (Bbs1 linker underlined, S1 bold and underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT

GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCAATGC**TAATAG**AA*TAAT***GA**
**TGATGACGATAAGAATAAAGTTAAGAAACCACTTACAGGAGCACATAGACTATT**
**GTTCACTTTTTTATTTGTATAATAGAA***CCCC*GGGTCTTCA**TAATAG**ATTCCAGCAC
ACTGGCGGCCGTTACTAGTGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCGAGT
TGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGG
TCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATATCCACAGGACGGGT
GTGGTCGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAGCGAA

**82_S1_stop_ins (Bbs1 linker underlined, S1 bold and underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCAATGCTTTTCAAGATACCC
GGATCATATGAAACGGCATGAC**TAATAG**AA*TAAT***GATGATGACGATAAGAATAAA**
**GTTAAGAAACCACTTACAGGAGCACATAGACTATTGTTCACTTTTTTATTTGTAT**
**AATAGAA***CCCC*GGGTCTTCA**TAATAG**ATTCCAGCACACTGGCGGCCGTTACTAGTGG
ATCCGGCTGCTAACAAAGCCCGAAAGGAAGCGAGTTGGCTGCTGCCACCGCTGAGC
AATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGA
AAGGAGGAACTATATCCGGATATCCACAGGACGGGTGTGGTCGCCATGATCGCGTA
GTCGATAGTGGCTCCAAGTAGCGAA

**95_S1_stop_ins (Bbs1 linker underlined, S1 bold and underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCAATGCTTTTCAAGATACCC
GGATCATATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACA
GGAA**TAATAG**AA*TAAT***GATGATGACGATAAGAATAAAGTTAAGAAACCACTTACA**
**GGAGCACATAGACTATTGTTCACTTTTTTATTTGTATAATAGAA***CCCC*GGGTCTTC
A**TAATAG**ATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGGCTGCTAACAAAGC
CCGAAAGGAAGCGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTT
GGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGA
TATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAG
CGAA

**139_S1_stop_ins (Bbs1 linker underlined, S1 bold and underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCAATGCTTTTCAAGATACCC
GGATCATATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACA
GGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAAGTCA
AGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAG

AAGATGGAAACATTCTTGGACAC**TAATAG**AA*TAAT***GATGATGACGATAAGAATAAA**
**GTTAAGAAACCACTTACAGGAGCACATAGACTATTGTTCACTTTTTTATTTGTAT**
**AATAGAA***CCCC*GG**GTCTTC**A**TAATAG**ATTCCAGCACACTGGCGGCCGTTACTAGTGG
ATCCGGCTGCTAACAAAGCCCGAAAGGAAGCGAGTTGGCTGCTGCCACCGCTGAGC
AATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGA
AAGGAGGAACTATATCCGGATATCCACAGGACGGGTGTGGTCGCCATGATCGCGTA
GTCGATAGTGGCTCCAAGTAGCGAA

**203_S1_stop_ins (Bbs1 linker underlined, S1 bold and underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCAATGCTTTTCAAGATACCC
GGATCATATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACA
GGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAAGTCA
AGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAG
AAGATGGAAACATTCTTGGACACAAATTGGAATACAACTATAACTCACACAATGTAT
ACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGACAC
AACATTGAAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATT
GGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCCACA**TAATAG**AA*TAAT*G
**ATGATGACGATAAGAATAAAGTTAAGAAACCACTTACAGGAGCACATAGACTAT**
**TGTTCACTTTTTTATTTGTATAATAGAA***CCCC*GG**GTCTTC**A**TAATAG**ATTCCAGCA
CACTGGCGGCCGTTACTAGTGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCGAG
TTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGG
GTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATATCCACAGGACGGG
TGTGGTCGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAGCGAA

**250_S1_stop_ins (Bbs1 linker underlined, S1 bold and underlined)**
ATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTA
TAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGACTAAAGGTGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAA
TTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCAATGCTTTTCAAGATACCC
GGATCATATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACA
GGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAAGTCA
AGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAG
AAGATGGAAACATTCTTGGACACAAATTGGAATACAACTATAACTCACACAATGTAT
ACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGACAC
AACATTGAAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATT
GGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTT
CGAAAGATCCCAACGAAAAGAGAGACCACATGGTCCTTCTTGAGTTTGTAACAGCTG
CTGGGATTACACATGGCATGGATGAACTGTACCAACCC**TAATAG**AA*TAAT***GATGATG**
**ACGATAAGAATAAAGTTAAGAAACCACTTACAGGAGCACATAGACTATTGTTCA**
**CTTTTTTATTTGTATAATAGAA***CCCC*GG**GTCTTC**A**TAATAG**ATTCCAGCACACTGGC
GGCCGTTACTAGTGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCGAGTTGGCTG
CTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGA

GGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATATCCACAGGACGGGTGTGGTC
GCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAGCGAA

## Truncated Luciferase constructs

**36 amino acid truncation of Luciferase (Bbs1 linker underlined)**
TGGTCCGGTAACGTGCTGAGCCCGGCCAAGCTTACTCCCCATCCCCCTGTTGACAATT
AATCATCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAA
ACAGGATCCAAATGGAAGACGCCAAAAACATAAAGAAAGGCCCGGCGCCATTCTAT
CCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGC
CCTGGTTTCACAGAA*TAAT*GGGTCTTC**ATAATAG**ATTCCAGCACACTGGCGGCCGTTA
CTAGTGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCGAGTTGGCTGCTGCCACC
GCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTT
TTGCTGAAAGGAGGAACTATATCCGGATATCCACAGGACGGGTGTGGTCGCCATGAT
CGCGTAGTCGATAGTGGCTCCAAGTAGCGAA

**51 amino acid truncation of Luciferase (Bbs1 linker underlined)**
TGGTCCGGTAACGTGCTGAGCCCGGCCAAGCTTACTCCCCATCCCCCTGTTGACAATT
AATCATCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAA
ACAGGATCCAAATGGAAGACGCCAAAAACATAAAGAAAGGCCCGGCGCCATTCTAT
CCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGC
CCTGGTTCCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGAACATCTCACA
GAA*TAAT*GGGTCTTC**ATAATAG**ATTCCAGCACACTGGCGGCCGTTACTAGTGGATCC
GGCTGCTAACAAAGCCCGAAAGGAAGCGAGTTGGCTGCTGCCACCGCTGAGCAATA
ACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGG
AGGAACTATATCCGGATATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTCGA
TAGTGGCTCCAAGTAGCGAA

**69 amino acid truncation of Luciferase (Bbs1 linker underlined)**
TGGTCCGGTAACGTGCTGAGCCCGGCCAAGCTTACTCCCCATCCCCCTGTTGACAATT
AATCATCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAA
ACAGGATCCAAATGGAAGACGCCAAAAACATAAAGAAAGGCCCGGCGCCATTCTAT
CCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGC
CCTGGTTCCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGAACATCACGTA
CGCGGAATACTTCGAAATGTCCGTTCGGTTGGCAGAAGCTATGAAACGATCACAGAA
*TAAT*GGGTCTTC**ATAATAG**ATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGGCT
GCTAACAAAGCCCGAAAGGAAGCGAGTTGGCTGCTGCCACCGCTGAGCAATAACTA
GCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGA
ACTATATCCGGATATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTCGATAGT
GGCTCCAAGTAGCGAA

**83 amino acid truncation of Luciferase (Bbs1 linker underlined)**
TGGTCCGGTAACGTGCTGAGCCCGGCCAAGCTTACTCCCCATCCCCCTGTTGACAATT
AATCATCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAA
ACAGGATCCAAATGGAAGACGCCAAAAACATAAAGAAAGGCCCGGCGCCATTCTAT
CCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGC
CCTGGTTCCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGAACATCACGTA
CGCGGAATACTTCGAAATGTCCGTTCGGTTGGCAGAAGCTATGAAACGATATGGGCT

GAATACAAATCACAGAATCGTCGTATGCAGTGAATCACAGAA*TAAT*GGGTCTTCA**TA**
**ATAG**ATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGGCTGCTAACAAAGCCCG
AAAGGAAGCGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGG
GCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATAT
CCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAGCGA
A

**101 amino acid truncation of Luciferase (Bbs1 linker underlined)**
TGGTCCGGTAACGTGCTGAGCCCGGCCAAGCTTACTCCCCATCCCCCTGTTGACAATT
AATCATCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAA
ACAGGATCCAAATGGAAGACGCCAAAAACATAAAGAAAGGCCCGGCGCCATTCTAT
CCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGC
CCTGGTTCCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGAACATCACGTA
CGCGGAATACTTCGAAATGTCCGTTCGGTTGGCAGAAGCTATGAAACGATATGGGCT
GAATACAAATCACAGAATCGTCGTATGCAGTGAAAACTCTCTTCAATTCTTTATGCC
GGTGTTGGGCGCGTTATTTATCGGAGTTGCATCACAGAA*TAAT*GG*GTCTTC*A**TAATAG**
ATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGGCTGCTAACAAAGCCCGAAAG
GAAGCGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCT
CTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATATCCAC
AGGACGGGTGTGGTCGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAGCGAA

**129 amino acid truncation of Luciferase (Bbs1 linker underlined)**
TGGTCCGGTAACGTGCTGAGCCCGGCCAAGCTTACTCCCCATCCCCCTGTTGACAATT
AATCATCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAA
ACAGGATCCAAATGGAAGACGCCAAAAACATAAAGAAAGGCCCGGCGCCATTCTAT
CCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGC
CCTGGTTCCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGAACATCACGTA
CGCGGAATACTTCGAAATGTCCGTTCGGTTGGCAGAAGCTATGAAACGATATGGGCT
GAATACAAATCACAGAATCGTCGTATGCAGTGAAAACTCTCTTCAATTCTTTATGCC
GGTGTTGGGCGCGTTATTTATCGGAGTTGCAGTTGCGCCCGCGAACGACATTTATAA
TGAACGTGAATTGCTCAACAGTATGAACATTTCGCAGCCTACCGTAGTGTTTGTTTCC
TCACAGAA*TAAT*GG*GTCTTC*A**TAATAG**ATTCCAGCACACTGGCGGCCGTTACTAGTG
GATCCGGCTGCTAACAAAGCCCGAAAGGAAGCGAGTTGGCTGCTGCCACCGCTGAG
CAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTG
AAAGGAGGAACTATATCCGGATATCCACAGGACGGGTGTGGTCGCCATGATCGCGT
AGTCGATAGTGGCTCCAAGTAGCGAA

**169 amino acid truncation of Luciferase (Bbs1 linker underlined)**
TGGTCCGGTAACGTGCTGAGCCCGGCCAAGCTTACTCCCCATCCCCCTGTTGACAATT
AATCATCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAA
ACAGGATCCAAATGGAAGACGCCAAAAACATAAAGAAAGGCCCGGCGCCATTCTAT
CCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGC
CCTGGTTCCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGAACATCACGTA
CGCGGAATACTTCGAAATGTCCGTTCGGTTGGCAGAAGCTATGAAACGATATGGGCT
GAATACAAATCACAGAATCGTCGTATGCAGTGAAAACTCTCTTCAATTCTTTATGCC
GGTGTTGGGCGCGTTATTTATCGGAGTTGCAGTTGCGCCCGCGAACGACATTTATAA
TGAACGTGAATTGCTCAACAGTATGAACATTTCGCAGCCTACCGTAGTGTTTGTTTCC
AAAAAGGGGTTGCAAAAAATTTTGAACGTGCAAAAAAAATTACCAATAATCCAGAA
AATTATTATCATGGATTCTAAAACGGATTACCAGGGATTTCAGTCGATGTACACGTTC

GTCACATCACAGAA*TAAT*GGGTCTTCA**TAATAG**ATTCCAGCACACTGGCGGCCGTTA
CTAGTGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCGAGTTGGCTGCTGCCACC
GCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTT
TTGCTGAAAGGAGGAACTATATCCGGATATCCACAGGACGGGTGTGGTCGCCATGAT
CGCGTAGTCGATAGTGGCTCCAAGTAGCGAA

**211 amino acid truncation of Luciferase (Bbs1 linker underlined)**
TGGTCCGGTAACGTGCTGAGCCCGGCCAAGCTTACTCCCCATCCCCCTGTTGACAATT
AATCATCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAA
ACAGGATCCAAATGGAAGACGCCAAAAACATAAAGAAAGGCCCGGCGCCATTCTAT
CCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGC
CCTGGTTCCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGAACATCACGTA
CGCGGAATACTTCGAAATGTCCGTTCGGTTGGCAGAAGCTATGAAACGATATGGGCT
GAATACAAATCACAGAATCGTCGTATGCAGTGAAAACTCTCTTCAATTCTTTATGCC
GGTGTTGGGCGCGTTATTTATCGGAGTTGCAGTTGCGCCCGCGAACGACATTTATAA
TGAACGTGAATTGCTCAACAGTATGAACATTTCGCAGCCTACCGTAGTGTTTGTTTCC
AAAAAGGGGTTGCAAAAAATTTTGAACGTGCAAAAAAAATTACCAATAATCCAGAA
AATTATTATCATGGATTCTAAAACGGATTACCAGGGATTTCAGTCGATGTACACGTTC
GTCACATCTCATCTACCTCCCGGTTTTAATGAATACGATTTTGTACCAGAGTCCTTTG
ATCGTGACAAAACAATTGCACTGATAATGAATTCCTCTGGATCTACTGGGTTACCTA
AGGGTGTGGCCCTTCCGTCACAGAA*TAAT*GGGTCTTCA**TAATAG**ATTCCAGCACACT
GGCGGCCGTTACTAGTGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCGAGTTGG
CTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCT
TGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATATCCACAGGACGGGTGTG
GTCGCCATGATCGCGTAGTCGATAGTGGCTCCAAGTAGCGAA

**339 amino acid truncation of Luciferase (Bbs1 linker underlined)**
TGGTCCGGTAACGTGCTGAGCCCGGCCAAGCTTACTCCCCATCCCCCTGTTGACAATT
AATCATCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAA
ACAGGATCCAAATGGAAGACGCCAAAAACATAAAGAAAGGCCCGGCGCCATTCTAT
CCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGC
CCTGGTTCCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGAACATCACGTA
CGCGGAATACTTCGAAATGTCCGTTCGGTTGGCAGAAGCTATGAAACGATATGGGCT
GAATACAAATCACAGAATCGTCGTATGCAGTGAAAACTCTCTTCAATTCTTTATGCC
GGTGTTGGGCGCGTTATTTATCGGAGTTGCAGTTGCGCCCGCGAACGACATTTATAA
TGAACGTGAATTGCTCAACAGTATGAACATTTCGCAGCCTACCGTAGTGTTTGTTTCC
AAAAAGGGGTTGCAAAAAATTTTGAACGTGCAAAAAAAATTACCAATAATCCAGAA
AATTATTATCATGGATTCTAAAACGGATTACCAGGGATTTCAGTCGATGTACACGTTC
GTCACATCTCATCTACCTCCCGGTTTTAATGAATACGATTTTGTACCAGAGTCCTTTG
ATCGTGACAAAACAATTGCACTGATAATGAATTCCTCTGGATCTACTGGGTTACCTA
AGGGTGTGGCCCTTCCGCATAGAACTGCCTGCGTCAGATTCTCGCATGCCAGAGATC
CTATTTTTGGCAATCAAATCATTCCGGATACTGCGATTTTAAGTGTTGTTCCATTCCA
TCACGGTTTTGGAATGTTTACTACACTCGGATATTTGATATGTGGATTTCGAGTCGTC
TTAATGTATAGATTTGAAGAAGAGCTGTTTTTACGATCCCTTCAGGATTACAAAATTC
AAAGTGCGTTGCTAGTACCAACCCTATTTTCATTCTTCGCCAAAAGCACTCTGATTGA
CAAATACGATTTATCTAATTTACACGAAATTGCTTCTGGGGGCGCACCTCTTTCGAAA
GAAGTCGGGGAAGCGGTTGCAAAACGCTTCCATCTTCCAGGGATACGACAAGGATC
ACAGAA*TAAT*GGGTCTTCA**TAATAG**ATTCCAGCACACTGGCGGCCGTTACTAGTGGA
TCCGGCTGCTAACAAAGCCCGAAAGGAAGCGAGTTGGCTGCTGCCACCGCTGAGCA
ATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAA

AGGAGGAACTATATCCGGATATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGT
CGATAGTGGCTCCAAGTAGCGAA

**437 amino acid truncation of Luciferase (Bbs1 linker underlined)**

TGGTCCGGTAACGTGCTGAGCCCGGCCAAGCTTACTCCCCATCCCCCTGTTGACAATT
AATCATCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAA
ACAGGATCCAAATGGAAGACGCCAAAAACATAAAGAAAGGCCCGGCGCCATTCTAT
CCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGC
CCTGGTTCCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGAACATCACGTA
CGCGGAATACTTCGAAATGTCCGTTCGGTTGGCAGAAGCTATGAAACGATATGGGCT
GAATACAAATCACAGAATCGTCGTATGCAGTGAAAACTCTCTTCAATTCTTTATGCC
GGTGTTGGGCGCGTTATTTATCGGAGTTGCAGTTGCGCCCGCGAACGACATTTATAA
TGAACGTGAATTGCTCAACAGTATGAACATTTCGCAGCCTACCGTAGTGTTTGTTTCC
AAAAAGGGGTTGCAAAAAATTTTGAACGTGCAAAAAAAATTACCAATAATCCAGAA
AATTATTATCATGGATTCTAAAACGGATTACCAGGGATTTCAGTCGATGTACACGTTC
GTCACATCTCATCTACCTCCCGGTTTTAATGAATACGATTTTGTACCAGAGTCCTTTG
ATCGTGACAAAACAATTGCACTGATAATGAATTCCTCTGGATCTACTGGGTTACCTA
AGGGTGTGGCCCTTCCGCATAGAACTGCCTGCGTCAGATTCTCGCATGCCAGAGATC
CTATTTTTGGCAATCAAATCATTCCGGATACTGCGATTTTAAGTGTTGTTCCATTCCA
TCACGGTTTTGGAATGTTTACTACACTCGGATATTTGATATGTGGATTTCGAGTCGTC
TTAATGTATAGATTTGAAGAAGAGCTGTTTTTACGATCCCTTCAGGATTACAAAATTC
AAAGTGCGTTGCTAGTACCAACCCTATTTTCATTCTTCGCCAAAAGCACTCTGATTGA
CAAATACGATTTATCTAATTTACACGAAATTGCTTCTGGGGGCGCACCTCTTTCGAAA
GAAGTCGGGGAAGCGGTTGCAAAACGCTTCCATCTTCCAGGGATACGACAAGGATA
TGGGCTCACTGAGACTACATCAGCTATTCTGATTACACCCGAGGGGGATGATAAACC
GGGCGCGGTCGGTAAAGTTGTTCCATTTTTTGAAGCGAAGGTTGTGGATCTGGATAC
CGGGAAAACGCTGGGCGTTAATCAGAGAGGCGAATTATGTGTCAGAGGACCTATGA
TTATGTCCGGTTATGTAAACAATCCGGAAGCGACCAACGCCTTGATTGACAAGGATG
GATGGCTACATTCTGGAGACATAGCTTACTGGGACGAAGACGAACACTTCTTCATAG
TTGACCGCTCACAGAA*TAAT*GG<u>GTCTTC</u>A**TAATAG**ATTCCAGCACACTGGCGGCCGTT
ACTAGTGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCGAGTTGGCTGCTGCCAC
CGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTT
TTTGCTGAAAGGAGGAACTATATCCGGATATCCACAGGACGGGTGTGGTCGCCATGA
TCGCGTAGTCGATAGTGGCTCCAAGTAGCGAA

**550 amino acid truncation of Luciferase (Bbs1 linker underlined)**

TGGTCCGGTAACGTGCTGAGCCCGGCCAAGCTTACTCCCCATCCCCCTGTTGACAATT
AATCATCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAA
ACAGGATCCAAATGGAAGACGCCAAAAACATAAAGAAAGGCCCGGCGCCATTCTAT
CCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGC
CCTGGTTCCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGAACATCACGTA
CGCGGAATACTTCGAAATGTCCGTTCGGTTGGCAGAAGCTATGAAACGATATGGGCT
GAATACAAATCACAGAATCGTCGTATGCAGTGAAAACTCTCTTCAATTCTTTATGCC
GGTGTTGGGCGCGTTATTTATCGGAGTTGCAGTTGCGCCCGCGAACGACATTTATAA
TGAACGTGAATTGCTCAACAGTATGAACATTTCGCAGCCTACCGTAGTGTTTGTTTCC
AAAAAGGGGTTGCAAAAAATTTTGAACGTGCAAAAAAAATTACCAATAATCCAGAA
AATTATTATCATGGATTCTAAAACGGATTACCAGGGATTTCAGTCGATGTACACGTTC
GTCACATCTCATCTACCTCCCGGTTTTAATGAATACGATTTTGTACCAGAGTCCTTTG
ATCGTGACAAAACAATTGCACTGATAATGAATTCCTCTGGATCTACTGGGTTACCTA
AGGGTGTGGCCCTTCCGCATAGAACTGCCTGCGTCAGATTCTCGCATGCCAGAGATC

CTATTTTTGGCAATCAAATCATTCCGGATACTGCGATTTTAAGTGTTGTTCCATTCCA
TCACGGTTTTGGAATGTTTACTACACTCGGATATTTGATATGTGGATTTCGAGTCGTC
TTAATGTATAGATTTGAAGAAGAGCTGTTTTTACGATCCCTTCAGGATTACAAAATTC
AAAGTGCGTTGCTAGTACCAACCCTATTTTCATTCTTCGCCAAAAGCACTCTGATTGA
CAAATACGATTTATCTAATTTACACGAAATTGCTTCTGGGGGCGCACCTCTTTCGAAA
GAAGTCGGGGAAGCGGTTGCAAAACGCTTCCATCTTCCAGGGATACGACAAGGATA
TGGGCTCACTGAGACTACATCAGCTATTCTGATTACACCCGAGGGGGATGATAAACC
GGGCGCGGTCGGTAAAGTTGTTCCATTTTTTGAAGCGAAGGTTGTGGATCTGGATAC
CGGGAAAACGCTGGGCGTTAATCAGAGAGGCGAATTATGTGTCAGAGGACCTATGA
TTATGTCCGGTTATGTAAACAATCCGGAAGCGACCAACGCCTTGATTGACAAGGATG
GATGGCTACATTCTGGAGACATAGCTTACTGGGACGAAGACGAACACTTCTTCATAG
TTGACCGCTTGAAGTCTTTAATTAAATACAAAGGATATCAGGTGGCCCCCGCTGAAT
TGGAATCGATATTGTTACAACACCCCAACATCTTCGACGCGGGCGTGGCAGGTCTTC
CCGACGATGACGCCGGTGAACTTCCCGCCGCCGTTGTTGTTTTGGAGCACGGAAAGA
CGATGACGGAAAAAGAGATCGTGGATTACGTGGCCAGTCAAGTAACAACCGCGAAA
AAGTTGCGCGGAGGAGTTGTGTTTGTGGACGAAGTACCGAAAGGTCTTACCGGAAA
ACTCGACGCAAGAAAAATCAGAGAGATCCTCATAAAGGCCAAGAAGGGCGGAAAGT
CCAAATTGTCACAGAA*TAAT*GGGTCTTCA**TAATAG**ATTCCAGCACACTGGCGGCCGT
TACTAGTGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCGAGTTGGCTGCTGCCAC
CGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTT
TTTGCTGAAAGGAGGAACTATATCCGGATATCCACAGGACGGGTGTGGTCGCCATGA
TCGCGTAGTCGATAGTGGCTCCAAGTAGCGAA

# Chapter 9

# References

Amblard, M., J. A. Fehrentz, et al. (2005). "Fundamentals of modern peptide synthesis." Methods Mol Biol **298**: 3-24.

Andreu, D. and L. Rivas (1998). "Animal antimicrobial peptides: an overview." Biopolymers **47**(6): 415-33.

Andreu, D., J. Ubach, et al. (1992). "Shortened cecropin A-melittin hybrids. Significant size reduction retains potent antibiotic activity." FEBS Lett **296**(2): 190-4.

Asthana, N., S. P. Yadav, et al. (2004). "Dissection of antibacterial and toxic activity of melittin: a leucine zipper motif plays a crucial role in determining its hemolytic activity but not antibacterial activity." J Biol Chem **279**(53): 55042-50.

Bairoch, A. and R. Apweiler (2000). "The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000." Nucleic Acids Res **28**(1): 45-8.

Balaban, N., Y. Gov, et al. (2003). "Prevention of Staphylococcus aureus biofilm on dialysis catheters and adherence to human cells." Kidney Int **63**(1): 340-5.

Baneyx, F. and G. Georgiou (1991). "Construction and characterization of Escherichia coli strains deficient in multiple secreted proteases: protease III degrades high-molecular-weight substrates in vivo." J Bacteriol **173**(8): 2696-703.

Bassetti, S., J. Hu, et al. (2001). "Prolonged antimicrobial activity of a catheter containing chlorhexidine-silver sulfadiazine extends protection against catheter infections in vivo." Antimicrob Agents Chemother **45**(5): 1535-8.

Befus, A. D., C. Mowat, et al. (1999). "Neutrophil defensins induce histamine secretion from mast cells: mechanisms of action." J Immunol **163**(2): 947-53.

Bell, G. and P. H. Gouyon (2003). "Arming the enemy: the evolution of resistance to self-proteins." Microbiology **149**(Pt 6): 1367-75.

Blondelle, S. E., E. Takahashi, et al. (1996). "Rapid identification of compounds with enhanced antimicrobial activity by using conformationally defined combinatorial libraries." Biochem J **313 (Pt 1)**: 141-7.

Boggiano, C., N. Reixach, et al. (2003). "Successful identification of novel agents to control infectious diseases from screening mixture-based peptide combinatorial libraries in complex cell-based bioassays." Biopolymers **71**(2): 103-16.

Boman, H. G. (2003). "Antibacterial peptides: basic facts and emerging concepts." J Intern Med **254**(3): 197-215.

Boman, H. G., D. Wade, et al. (1989). "Antibacterial and antimalarial properties of peptides that are cecropin-melittin hybrids." FEBS Lett **259**(1): 103-6.

Bot, A. I., D. J. Smith, et al. (2001). "Receptor-mediated targeting of spray-dried lipid particles coformulated with immunoglobulin and loaded with a prototype vaccine." Pharm Res **18**(7): 971-9.

Brogden, K. A. (2005). "Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria?" Nat Rev Microbiol **3**(3): 238-50.

Brogden, K. A., J. M. Guthmiller, et al. (2005). "The nervous system and innate immunity: the neuropeptide connection." Nat Immunol **6**(6): 558-64.

Brotz, H., G. Bierbaum, et al. (1998). "The lantibiotic mersacidin inhibits peptidoglycan synthesis by targeting lipid II." Antimicrob Agents Chemother **42**(1): 154-60.

Brown, K. L. and R. E. Hancock (2006). "Cationic host defense (antimicrobial) peptides." Curr Opin Immunol **18**(1): 24-30.

Budde, B. B. and M. Rasch (2001). "A comparative study on the use of flow cytometry and colony forming units for assessment of the antibacterial effect of bacteriocins." Int J Food Microbiol **63**(1-2): 65-72.

Busso, D., R. Kim, et al. (2004). "Using an Escherichia coli cell-free extract to screen for soluble expression of recombinant proteins." J Struct Funct Genomics **5**(1-2): 69-74.

Calhoun, K. A. and J. R. Swartz (2005). "An economical method for cell-free protein synthesis using glucose and nucleoside monophosphates." Biotechnol Prog **21**(4): 1146-53.

Caraher, E., G. Reynolds, et al. (2007). "Comparison of antibiotic susceptibility of Burkholderia cepacia complex organisms when grown planktonically or as biofilm in vitro." Eur J Clin Microbiol Infect Dis.

Carreno, C., M. E. Mendez, et al. (2000). "Nsc and Fmoc Nalpha-amino protection for solid-phase peptide synthesis: a parallel study." J Pept Res **56**(2): 63-9.

Chen, Y., C. T. Mant, et al. (2005). "Rational design of alpha-helical antimicrobial peptides with enhanced activities and specificity/therapeutic index." J Biol Chem **280**(13): 12316-29.

Ciornei, C. D., T. Sigurdardottir, et al. (2005). "Antimicrobial and chemoattractant activity, lipopolysaccharide neutralization, cytotoxicity, and inhibition by serum of analogs of human cathelicidin LL-37." Antimicrob Agents Chemother **49**(7): 2845-50.

Cirioni, O., C. Silvestri, et al. (2006). "Experimental study on the efficacy of combination of alpha-helical antimicrobial peptides and vancomycin against Staphylococcus aureus with intermediate resistance to glycopeptides." Peptides **27**(11): 2600-6.

Coates, A., Y. Hu, et al. (2002). "The future challenges facing the development of new antimicrobial drugs." Nat Rev Drug Discov **1**(11): 895-910.

Conlon, J. M., A. Sonnevend, et al. (2005). "A family of acyclic brevinin-1 peptides from the skin of the Ryukyu brown frog Rana okinavana." Peptides **26**(2): 185-90.

Corral, L., M. Nolla-Salas, et al. (2003). "A prospective, randomized study in critically ill patients using the Oligon Vantex catheter." J Hosp Infect **55**(3): 212-9.

Costerton, J. W., P. S. Stewart, et al. (1999). "Bacterial biofilms: a common cause of persistent infections." Science **284**(5418): 1318-22.

Darouiche, R. O. (2001). "Device-associated infections: a macroproblem that starts with microadherence." Clin Infect Dis **33**(9): 1567-72.

Darouiche, R. O. (2004). "Treatment of infections associated with surgical implants." N Engl J Med **350**(14): 1422-9.

Davis, G. D., C. Elisee, et al. (1999). "New fusion protein systems designed to give soluble expression in Escherichia coli." Biotechnol Bioeng **65**(4): 382-8.

Deslouches, B., S. M. Phadke, et al. (2005). "De novo generation of cationic antimicrobial peptides: influence of length and tryptophan substitution on antimicrobial activity." Antimicrob Agents Chemother **49**(1): 316-22.

Dorschner, R. A., B. Lopez-Garcia, et al. (2006). "The mammalian ionic environment dictates microbial susceptibility to antimicrobial defense peptides." Faseb J **20**(1): 35-42.

Drin, G., S. Cottin, et al. (2003). "Studies on the internalization mechanism of cationic cell-penetrating peptides." J Biol Chem **278**(33): 31192-201.

Ellerby, H. M., W. Arap, et al. (1999). "Anti-cancer activity of targeted pro-apoptotic peptides." Nat Med **5**(9): 1032-8.

Falla, T. J. and R. E. Hancock (1997). "Improved activity of a synthetic indolicidin analog." Antimicrob Agents Chemother **41**(4): 771-5.

Falquet, L., M. Pagni, et al. (2002). "The PROSITE database, its status in 2002." Nucleic Acids Res **30**(1): 235-8.

Ferre, R., E. Badosa, et al. (2006). "Inhibition of plant-pathogenic bacteria by short synthetic cecropin A-melittin hybrid peptides." Appl Environ Microbiol **72**(5): 3302-8.

Fields, G. B. and R. L. Noble (1990). "Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids." Int J Pept Protein Res **35**(3): 161-214.

Finlay, B. B. and R. E. Hancock (2004). "Can innate immunity be enhanced to treat microbial infections?" Nat Rev Microbiol **2**(6): 497-504.

Frecer, V., B. Ho, et al. (2004). "De novo design of potent antimicrobial peptides." Antimicrob Agents Chemother **48**(9): 3349-57.

Fux, C. A., S. Wilson, et al. (2004). "Detachment characteristics and oxacillin resistance of Staphyloccocus aureus biofilm emboli in an in vitro catheter infection model." J Bacteriol **186**(14): 4486-91.

Giacometti, A., O. Cirioni, et al. (2000). "In-vitro activity and killing effect of polycationic peptides on methicillin-resistant Staphylococcus aureus and interactions with clinically used antibiotics." Diagn Microbiol Infect Dis **38**(2): 115-8.

Giangaspero, A., L. Sandri, et al. (2001). "Amphipathic alpha helical antimicrobial peptides." Eur J Biochem **268**(21): 5589-600.

Gimenez, D., C. Andreu, et al. (2006). "The introduction of fluorine atoms or trifluoromethyl groups in short cationic peptides enhances their antimicrobial activity." Bioorg Med Chem **14**(20): 6971-8.

Goldman, M. J., G. M. Anderson, et al. (1997). "Human beta-defensin-1 is a salt-sensitive antibiotic in lung that is inactivated in cystic fibrosis." Cell **88**(4): 553-60.

Gottesman, S. (1996). "Proteases and their targets in Escherichia coli." Annu Rev Genet **30**: 465-506.

Hamamoto, K., Y. Kida, et al. (2002). "Antimicrobial activity and stability to proteolysis of small linear cationic peptides with D-amino acid substitutions." Microbiol Immunol **46**(11): 741-9.

Hancock, R. E. and A. Patrzykat (2002). "Clinical development of cationic antimicrobial peptides: from natural to novel antibiotics." Curr Drug Targets Infect Disord **2**(1): 79-83.

Haynie, S. L., G. A. Crum, et al. (1995). "Antimicrobial activities of amphiphilic peptides covalently bonded to a water-insoluble resin." Antimicrob. Agents Chemother. **39**(2): 301-307.

Henry, S., D. V. McAllister, et al. (1999). "Microfabricated microneedles: A novel approach to transdermal drug delivery." J Pharm Sci **88**(9): 948.

Hilpert, K., R. Volkmer-Engert, et al. (2005). "High-throughput generation of small antibacterial peptides with improved activity." Nat Biotechnol **23**(8): 1008-12.

Hisata, K., K. Kuwahara-Arai, et al. (2005). "Dissemination of methicillin-resistant staphylococci among healthy Japanese children." J Clin Microbiol **43**(7): 3364-72.

Hoess, R. H. (2001). "Protein design and phage display." Chem Rev **101**(10): 3205-18.

Holguin, A., J. L. Faudon, et al. (2007). "Susceptibility of HIV-1 non-B subtypes and recombinant variants to Enfuvirtide." J Clin Virol **38**(2): 176-80.

Hongbiao, W., N. Baolong, et al. (2005). "Biological activities of cecropin B-thanatin hybrid peptides." J Pept Res **66**(6): 382-6.

Houghten, R. A., C. Pinilla, et al. (1991). "Generation and use of synthetic peptide combinatorial libraries for basic research and drug discovery." Nature **354**(6348): 84-6.

Jacob, L. and M. Zasloff (1994). "Potential therapeutic applications of magainins and other antimicrobial agents of animal origin." Ciba Found Symp **186**: 197-216; discussion 216-23.

Jenssen, H., J. H. Andersen, et al. (2004). "A wide range of medium-sized, highly cationic, alpha-helical peptides show antiviral activity against herpes simplex virus." Antiviral Res **64**(2): 119-26.

Jenssen, H., P. Hamill, et al. (2006). "Peptide antimicrobial agents." Clin Microbiol Rev **19**(3): 491-511.

Jewett, M. C. and J. R. Swartz (2004). "Mimicking the Escherichia coli cytoplasmic environment activates long-lived and efficient cell-free protein synthesis." Biotechnol Bioeng **86**(1): 19-26.

Jiang, X., K. Oohira, et al. (2002). "Reduction of protein degradation by use of protease-deficient mutants in cell-free protein synthesis system of Escherichia coli." J Biosci Bioeng **93**(2): 151-6.

Josephson, K., M. C. Hartman, et al. (2005). "Ribosomal synthesis of unnatural peptides." J Am Chem Soc **127**(33): 11727-35.

Kamysz, W., M. Okroj, et al. (2003). "Novel properties of antimicrobial peptides." Acta Biochim Pol **50**(2): 461-9.

Kanikkannan, N., J. Singh, et al. (1999). "Transdermal iontophoretic delivery of bovine insulin and monomeric human insulin analogue." J Control Release **59**(1): 99-105.

Kim, D. M., T. Kigawa, et al. (1996). "A highly efficient cell-free protein synthesis system from Escherichia coli." Eur J Biochem **239**(3): 881-6.

Kim, D. M. and J. R. Swartz (2001). "Regeneration of adenosine triphosphate from glycolytic intermediates for cell-free protein synthesis." Biotechnol Bioeng **74**(4): 309-16.

Kim, H., S. H. Kim, et al. (2005). "Mechanism of natural rifampin resistance of Streptomyces spp." Syst Appl Microbiol **28**(5): 398-404.

Kim, S., S. S. Kim, et al. (2003). "In vitro activities of native and designed peptide antibiotics against drug sensitive and resistant tumor cell lines." Peptides **24**(7): 945-53.

Kirikae, T., M. Hirata, et al. (1998). "Protective effects of a human 18-kilodalton cationic antimicrobial protein (CAP18)-derived peptide against murine endotoxemia." Infect Immun **66**(5): 1861-8.

Koczulla, A. R. and R. Bals (2003). "Antimicrobial peptides: current status and therapeutic potential." Drugs **63**(4): 389-406.

Kolter, R. and E. P. Greenberg (2006). "Microbial sciences: the superficial life of microbes." Nature **441**(7091): 300-2.

Kuyyakanond, T. and L. B. Quesnel (1992). "The mechanism of action of chlorhexidine." FEMS Microbiol Lett **79**(1-3): 211-5.

Lamb, H. M. and L. R. Wiseman (1998). "Pexiganan acetate." Drugs **56**(6): 1047-52; discussion 1053-4.

Latham, P. W. (1999). "Therapeutic peptides revisited." Nat Biotechnol **17**(8): 755-7.

Lee, D. G., K. S. Hahm, et al. (2004). "Structure and fungicidal activity of a synthetic antimicrobial peptide, P18, and its truncated peptides." Biotechnol Lett **26**(4): 337-41.

Lee, J. H., J. H. Kim, et al. (2000). "High-level expression of antimicrobial peptide mediated by a fusion partner reinforcing formation of inclusion bodies." Biochem Biophys Res Commun **277**(3): 575-80.

Lee, J. H., M. S. Kim, et al. (2002). "Enhanced expression of tandem multimers of the antimicrobial peptide buforin II in Escherichia coli by the DEAD-box protein and trxB mutant." Appl Microbiol Biotechnol **58**(6): 790-6.

Lee, J. H., I. Minn, et al. (1998). "Acidic peptide-mediated expression of the antimicrobial peptide buforin II as tandem repeats in Escherichia coli." Protein Expr Purif **12**(1): 53-60.

Lehrer, R. I. (2004). "Paradise lost and paradigm found." Nat Immunol **5**(8): 775-6.

Lehtinen, J., M. Virta, et al. (2003). "Fluoro-luminometric real-time measurement of bacterial viability and killing." J Microbiol Methods **55**(1): 173-86.

Leone-Bay, A., M. Sato, et al. (2001). "Oral delivery of biologically active parathyroid hormone." Pharm Res **18**(7): 964-70.

Leontiadou, H., A. E. Mark, et al. (2006). "Antimicrobial peptides in action." J Am Chem Soc **128**(37): 12156-61.

Lewis, K. and A. M. Klibanov (2005). "Surpassing nature: rational design of sterile-surface materials." Trends Biotechnol **23**(7): 343-8.

Li, W., L. Jaroszewski, et al. (2001). "Clustering of highly homologous sequences to reduce the size of large protein databases." Bioinformatics **17**(3): 282-3.

Lien, S. and H. B. Lowman (2003). "Therapeutic peptides." Trends Biotechnol **21**(12): 556-62.

Loose, C., K. Jensen, et al. (2006). "A linguistic model for the rational design of antimicrobial peptides." Nature **443**(7113): 867-9.

Loose, C. R., R. S. Langer, et al. (2007). "Optimization of Protein Fusion Partner Length for Maximizing in Vitro Translation of Peptides." Biotechnol Prog.

Loury, D., J. R. Embree, et al. (1999). "Effect of local application of the antimicrobial peptide IB-367 on the incidence and severity of oral mucositis in hamsters." Oral Surg Oral Med Oral Pathol Oral Radiol Endod **87**(5): 544-51.

Lowy, F. D. (2003). "Antimicrobial resistance: the example of Staphylococcus aureus." J Clin Invest **111**(9): 1265-73.

Lupetti, A., A. Paulusma-Annema, et al. (2000). "Candidacidal activities of human lactoferrin peptides derived from the N terminus." Antimicrob Agents Chemother **44**(12): 3257-63.

Machado, A., M. L. Sforca, et al. (2007). "Truncation of amidated fragment 33-61 of bovine alpha-hemoglobin: Effects on the structure and anticandidal activity." Biopolymers.

Maizel, J. V., Jr. and R. P. Lenk (1981). "Enhanced graphic matrix analysis of nucleic acid and protein sequences." Proc Natl Acad Sci U S A **78**(12): 7665-9.

Makovitzki, A., D. Avrahami, et al. (2006). "Ultrashort antibacterial and antifungal lipopeptides." Proc Natl Acad Sci U S A **103**(43): 15997-6002.

Marr, A. K., W. J. Gooderham, et al. (2006). "Antibacterial peptides for therapeutic use: obstacles and realistic outlook." Curr Opin Pharmacol **6**(5): 468-72.

Matsuura, T., H. Yanagida, et al. (2007). "Nascent chain, mRNA, and ribosome complexes generated by a pure translation system." Biochem Biophys Res Commun **352**(2): 372-7.

McPhee, J. B., M. G. Scott, et al. (2005). "Design of host defence peptides for antimicrobial and immunity enhancing activities." Comb Chem High Throughput Screen **8**(3): 257-72.

Metz-Boutigue, M. H., A. E. Kieffer, et al. (2003). "Innate immunity: involvement of new neuropeptides." Trends Microbiol **11**(12): 585-92.

Michel-Reydellet, N., K. Calhoun, et al. (2004). "Amino acid stabilization for cell-free protein synthesis by modification of the Escherichia coli genome." Metab Eng **6**(3): 197-203.

Mignone, F., C. Gissi, et al. (2002). "Untranslated regions of mRNAs." Genome Biol **3**(3): REVIEWS0004.

Monkarsh, S. P., Y. Ma, et al. (1997). "Positional isomers of monopegylated interferon alpha-2a: isolation, characterization, and biological activity." Anal Biochem **247**(2): 434-40.

Mygind, P. H., R. L. Fischer, et al. (2005). "Plectasin is a peptide antibiotic with therapeutic potential from a saprophytic fungus." Nature **437**(7061): 975-80.

Nakano, H., Y. Kawarasaki, et al. (2004). "Cell-free protein synthesis systems: increasing their performance and applications." Adv Biochem Eng Biotechnol **90**: 135-49.

Nizet, V., T. Ohtake, et al. (2001). "Innate antimicrobial peptide protects the skin from invasive bacterial infection." Nature **414**(6862): 454-7.

Owen, S. M., D. Rudolph, et al. (2004). "A theta-defensin composed exclusively of D-amino acids is active against HIV-1." J Pept Res **63**(6): 469-76.

Owens, L. D. and T. M. Heutte (1997). "A single amino acid substitution in the antimicrobial defense protein cecropin B is associated with diminished degradation by leaf intercellular fluid." Mol Plant Microbe Interact **10**(4): 525-8.

Ozawa, K., M. J. Headlam, et al. (2004). "Optimization of an Escherichia coli system for cell-free synthesis of selectively N-labelled proteins for rapid analysis by NMR spectroscopy." Eur J Biochem **271**(20): 4084-93.

Papagianni, M. (2003). "Ribosomally synthesized peptides with antimicrobial properties: biosynthesis, structure, function, and applications." Biotechnol Adv **21**(6): 465-99.

Park, C. B., H. S. Kim, et al. (1998). "Mechanism of action of the antimicrobial peptide buforin II: buforin II kills microorganisms by penetrating the cell membrane and inhibiting cellular functions." Biochem Biophys Res Commun **244**(1): 253-7.

Park, J. M., J. E. Jung, et al. (1995). "Antimicrobial peptides from the skin of a Korean frog, Rana rugosa." Biochem Biophys Res Commun **209**(2): 775.

Park, Y., H. N. Kim, et al. (2004). "Design of novel analogues with potent antibiotic activity based on the antimicrobial peptide, HP(2-9)-ME(1-12)." Biotechnol Lett **26**(6): 493-8.

Perron, G. G., M. Zasloff, et al. (2006). "Experimental evolution of resistance to an antimicrobial peptide." Proc Biol Sci **273**(1583): 251-6.

Peschel, A. and H. G. Sahl (2006). "The co-evolution of host cationic antimicrobial peptides and microbial resistance." Nat Rev Microbiol **4**(7): 529-36.

Pilon, A., P. Yost, et al. (1997). "Ubiquitin fusion technology: bioprocessing of peptides." Biotechnol Prog **13**(4): 374-9.

Pittet, D., D. Tarara, et al. (1994). "Nosocomial bloodstream infection in critically ill patients. Excess length of stay, extra costs, and attributable mortality." Jama **271**(20): 1598-601.

Putsep, K., G. Carlsson, et al. (2002). "Deficiency of antibacterial peptides in patients with morbus Kostmann: an observation study." Lancet **360**(9340): 1144-9.

Raad, I., R. Darouiche, et al. (1997). "Central venous catheters coated with minocycline and rifampin for the prevention of catheter-related colonization and bloodstream infections. A randomized, double-blind trial. The Texas Medical Center Catheter Study Group." Ann Intern Med **127**(4): 267-74.

Rachid, S., K. Ohlsen, et al. (2000). "Effect of subinhibitory antibiotic concentrations on polysaccharide intercellular adhesin expression in biofilm-forming Staphylococcus epidermidis." Antimicrob Agents Chemother **44**(12): 3357-63.

Ramji, N., A. Baig, et al. (2005). "Sustained antibacterial actions of a new stabilized stannous fluoride dentifrice containing sodium hexametaphosphate." Compend Contin Educ Dent **26**(9 Suppl 1): 19-28.

Raventos, D., O. Taboureau, et al. (2005). "Improving on nature's defenses: optimization & high throughput screening of antimicrobial peptides." Comb Chem High Throughput Screen **8**(3): 219-33.

Rigoutsos, I. and A. Floratos (1998). "Combinatorial pattern discovery in biological sequences: The TEIRESIAS algorithm." Bioinformatics **14**(1): 55-67.

Robinson, W. E., Jr., B. McDougall, et al. (1998). "Anti-HIV-1 activity of indolicidin, an antimicrobial peptide from neutrophils." J Leukoc Biol **63**(1): 94-100.

Rotem, S., I. Radzishevsky, et al. (2006). "Physicochemical properties that enhance discriminative antibacterial activity of short dermaseptin derivatives." Antimicrob Agents Chemother **50**(8): 2666-72.

Rungpragayphan, S., Y. Kawarasaki, et al. (2002). "High-throughput, cloning-independent protein library construction by combining single-molecule DNA amplification with in vitro expression." J Mol Biol **318**(2): 395-405.

Rungpragayphan, S., H. Nakano, et al. (2003). "PCR-linked in vitro expression: a novel system for high-throughput construction and screening of protein libraries." FEBS Lett **540**(1-3): 147-50.

Ryge, T. S., X. Doisy, et al. (2004). "New indolicidin analogues with potent antibacterial activity." J Pept Res **64**(5): 171-85.

Sakakibara, T., S. Murakami, et al. (2003). "Enumeration of bacterial cell numbers by amplified firefly bioluminescence without cultivation." Anal Biochem **312**(1): 48-56.

Sampath, L. A., S. M. Tambe, et al. (2001). "In vitro and in vivo efficacy of catheters impregnated with antiseptics or antibiotics: evaluation of the risk of bacterial resistance to the antimicrobials in the catheters." Infect Control Hosp Epidemiol **22**(10): 640-6.

Sawasaki, T., T. Ogasawara, et al. (2002). "A cell-free protein synthesis system for high-throughput proteomics." Proc Natl Acad Sci U S A **99**(23): 14652-7.

Schmid, M. B. (2006). "Do targets limit antibiotic discovery?" Nat Biotechnol **24**(4): 419-20.

Schwab, U., P. Gilligan, et al. (1999). "In vitro activities of designed antimicrobial peptides against multidrug-resistant cystic fibrosis pathogens." Antimicrob Agents Chemother **43**(6): 1435-40.

Searls, D. B. (2002). "The language of genes." Nature **420**(6912): 211-7.

Shai, Y. (2002). "Mode of action of membrane active antimicrobial peptides." Biopolymers **66**(4): 236-48.

Shike, H., X. Lauth, et al. (2002). "Bass hepcidin is a novel antimicrobial peptide induced by bacterial challenge." Eur J Biochem **269**(8): 2232-7.

Shu, Q., Z. Shi, et al. (2006). "Protection against Pseudomonas aeruginosa pneumonia and sepsis-induced lung injury by overexpression of beta-defensin-2 in rats." Shock **26**(4): 365-71.

Sigurdardottir, T., P. Andersson, et al. (2006). "In silico identification and biological evaluation of antimicrobial peptides based on human cathelicidin LL-37." Antimicrob Agents Chemother **50**(9): 2983-9.

Silphaduang, U. and E. J. Noga (2001). "Peptide antibiotics in mast cells of fish." Nature **414**(6861): 268-9.

Silver, S., T. Phung le, et al. (2006). "Silver as biocides in burn and wound dressings and bacterial resistance to silver compounds." J Ind Microbiol Biotechnol **33**(7): 627-34.

Simon, L., C. Fremaux, et al. (2001). "Luminescent method for the detection of antibacterial activities." Appl Microbiol Biotechnol **57**(5-6): 757-63.

Speer, B. S., N. B. Shoemaker, et al. (1992). "Bacterial resistance to tetracycline: mechanisms, transfer, and clinical significance." Clin Microbiol Rev **5**(4): 387-99.

Spirin, A. S. (2004). "High-throughput cell-free systems for synthesis of functionally active proteins." Trends Biotechnol **22**(10): 538-45.

Spirin, A. S., V. I. Baranov, et al. (1988). "A continuous cell-free translation system capable of producing polypeptides in high yield." Science **242**(4882): 1162-4.

Steinberg, D. A., M. A. Hurst, et al. (1997). "Protegrin-1: a broad-spectrum, rapidly microbicidal peptide with in vivo activity." Antimicrob Agents Chemother **41**(8): 1738-42.

Su, W. W. (2005). "Fluorescent proteins as tools to aid protein production." Microb Cell Fact **4**(1): 12.

Swartz, J. R., M. C. Jewett, et al. (2004). "Cell-free protein synthesis with prokaryotic combined transcription-translation." Methods Mol Biol **267**: 169-82.

Taboureau, O., O. H. Olsen, et al. (2006). "Design of novispirin antimicrobial peptides by quantitative structure-activity relationship." Chem Biol Drug Des **68**(1): 48-57.

Tai, C. C., A. A. Nirvani, et al. (2004). "Methicillin-resistant Staphylococcus aureus in orthopaedic surgery." Int Orthop **28**(1): 32-5.

Tamamura, H., T. Ishihara, et al. (1996). "Analysis of the interaction of an anti-HIV peptide, T22 ([Tyr5, 12, Lys7]-polyphemusin II), with gp120 and CD4 by surface plasmon resonance." Biochim Biophys Acta **1298**(1): 37-44.

Tambe, S. M., L. Sampath, et al. (2001). "In vitro evaluation of the risk of developing bacterial resistance to antiseptics and antibiotics used in medical devices." J Antimicrob Chemother **47**(5): 589-98.

Tenno, T., N. Goda, et al. (2004). "[High-throughput construction of expression vector for protein domain analysis]." Tanpakushitsu Kakusan Koso **49**(16): 2587-94.

Terazawa, E., H. Shimonaka, et al. (1998). "Severe anaphylactic reaction due to a chlorhexidine-impregnated central venous catheter." Anesthesiology **89**(5): 1296-8.

Terpe, K. (2003). "Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems." Appl Microbiol Biotechnol **60**(5): 523-33.

Tian, Z. G., D. Teng, et al. (2007). "Multimerization and fusion expression of bovine lactoferricin derivative LfcinB15-W4,10 in Escherichia coli." Appl Microbiol Biotechnol.

Tiller, J. C., C. J. Liao, et al. (2001). "Designing surfaces that kill bacteria on contact." Proc Natl Acad Sci U S A **98**(11): 5981-5.

Tiozzo, E., G. Rocco, et al. (1998). "Wide-spectrum antibiotic activity of synthetic, amphipathic peptides." Biochem Biophys Res Commun **249**(1): 202-6.

Tossi, A., L. Sandri, et al. (2000). "Amphipathic, alpha-helical antimicrobial peptides." Biopolymers **55**(1): 4-30.

Tossi, A., M. Scocchi, et al. (1994). "Identification and characterization of a primary antibacterial domain in CAP18, a lipopolysaccharide binding protein from rabbit leukocytes." FEBS Lett **339**(1-2): 108-12.

Trop, M., M. Novak, et al. (2006). "Silver-coated dressing acticoat caused raised liver enzymes and argyria-like symptoms in burn patient." J Trauma **60**(3): 648-52.

Tytler, E. M., G. M. Anantharamaiah, et al. (1995). "Molecular basis for prokaryotic specificity of magainin-induced lysis." Biochemistry **34**(13): 4393-401.

Utsugi, T., A. J. Schroit, et al. (1991). "Elevated expression of phosphatidylserine in the outer membrane leaflet of human tumor cells and recognition by activated human blood monocytes." Cancer Res **51**(11): 3062-6.

Veal, D. A., D. Deere, et al. (2000). "Fluorescence staining and flow cytometry for monitoring microbial cells." J Immunol Methods **243**(1-2): 191-210.

Veenstra, D. L., S. Saint, et al. (1999). "Efficacy of antiseptic-impregnated central venous catheters in preventing catheter-related bloodstream infection: a meta-analysis." Jama **281**(3): 261-7.

Veenstra, D. L., S. Saint, et al. (1999). "Cost-effectiveness of antiseptic-impregnated central venous catheters for the prevention of catheter-related bloodstream infection." Jama **282**(6): 554-60.

Vesterlund, S., J. Paltta, et al. (2004). "Rapid screening method for the detection of antimicrobial substances." J Microbiol Methods **57**(1): 23-31.

Wakabayashi, H., H. Matsumoto, et al. (1999). "N-Acylated and D enantiomer derivatives of a nonamer core peptide of lactoferricin B showing improved antimicrobial activity." Antimicrob Agents Chemother **43**(5): 1267-9.

Walker, J. R., J. R. Roth, et al. (2001). "An in vivo study of novel bioactive peptides that inhibit the growth of Escherichia coli." J Pept Res **58**(5): 380-8.

Wang, Z. and G. Wang (2004). "APD: the Antimicrobial Peptide Database." Nucleic Acids Res **32**(Database issue): D590-2.

Won, H. S., S. J. Jung, et al. (2004). "Systematic peptide engineering and structural characterization to search for the shortest antimicrobial peptide analogue of gaegurin 5." J Biol Chem **279**(15): 14784-91.

Wu, M. and R. E. Hancock (1999). "Improved derivatives of bactenecin, a cyclic dodecameric antimicrobial cationic peptide." Antimicrob Agents Chemother **43**(5): 1274-6.

Wu, M. and R. E. Hancock (1999). "Interaction of the cyclic antimicrobial cationic peptide bactenecin with the outer and cytoplasmic membrane." J Biol Chem **274**(1): 29-35.

Yan, H., S. Li, et al. (2003). "Individual substitution analogs of Mel(12-26), melittin's C-terminal 15-residue peptide: their antimicrobial and hemolytic actions." FEBS Lett **554**(1-2): 100-4.

Yang, F., L. G. Moss, et al. (1996). "The molecular structure of green fluorescent protein." Nat Biotechnol **14**(10): 1246-51.

Yeaman, M. R. and N. Y. Yount (2003). "Mechanisms of antimicrobial peptide action and resistance." Pharmacol Rev **55**(1): 27-55.

Yount, N. Y. and M. R. Yeaman (2004). "Multidimensional signatures in antimicrobial peptides." Proc Natl Acad Sci U S A **101**(19): 7363-8.

Zaiou, M. (2007). "Multifunctional antimicrobial peptides: therapeutic targets in several human diseases." J Mol Med.

Zasloff, M. (2002). "Antimicrobial peptides in health and disease." N Engl J Med **347**(15): 1199-200.

Zasloff, M. (2002). "Antimicrobial peptides of multicellular organisms." Nature **415**(6870): 389-95.

Zasloff, M., B. Martin, et al. (1988). "Antimicrobial activity of synthetic magainin peptides and several analogues." Proc Natl Acad Sci U S A **85**(3): 910-3.

Zelezetsky, I., U. Pag, et al. (2005). "Identification and optimization of an antimicrobial peptide from the ant venom toxin pilosulin." Arch Biochem Biophys **434**(2): 358-64.

Zelver, N., M. Hamilton, et al. (1999). "Measuring antimicrobial effects on biofilm bacteria: from laboratory to field." Methods Enzymol **310**: 608-28.

Zhang, L. and T. J. Falla (2006). "Antimicrobial peptides: therapeutic potential." Expert Opin Pharmacother **7**(6): 653-63.

Zhang, L., W. Yu, et al. (2002). "Contribution of human alpha-defensin 1, 2, and 3 to the anti-HIV-1 activity of CD8 antiviral factor." Science **298**(5595): 995-1000.