We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



17

Prediction and Rational Design of Antimicrobial Peptides

William F. Porto, Osmar N. Silva and Octávio L. Franco
Universidade Católica de Brasília,
Centro de Análises Proteômicas e Bioquímicas,
Brazil

1. Introduction

In recent decades the activity of conventional antibiotics against pathogenic bacteria has decreased due to the development of resistance. This phenomenon has generated the so-called 'superbugs', which are multi-resistant bacteria. In this context, antimicrobial peptides (AMP) appear as an alternative to control them. AMPs have been found in several sources, including animals, plants and fungi, constituting the first line of host defence against pathogens. However, the use of AMPs as therapeutic agents has some limitations, such as stability, cytotoxicity and mainly their amino acid length, since amino acids are expensive building blocks. Despite these limitations they have compensatory properties, including secondary activities such as immunomodulation or antitumor ones. Several methods have been applied since the 1990s for rational AMPs design, in order to generate analogues with improved activity, looking to reduce limitations and increase advantages. Computer-aided identification and design of AMPs play a crucial role in this area. The discovery of AMP properties, through the first rational design studies, will allow the development of methods for prediction of AMPs, which in turn, should lead to identification prior to synthesis of novel analogues. Thus, this chapter will be dedicated to describing important techniques in prediction and rational design of AMPs and their applications for drug development.

1.1 Multi-resistant bacteria: The 'superbugs'

A number of lethal infections became tractable and curable after the discovery and subsequent use of antimicrobial agents in clinical therapy, as the case of syphilis, rheumatic fever and cellulitis. However, this success has dimmed over the course of time due to the uncontrolled and inappropriate use of antibiotics, including the administration of under or overestimated doses, the insufficient duration of treatment and mistakes in the choice of drugs. Currently various microorganisms are resistant to antimicrobials, leading to the emergence and spread of so-called 'superbugs' resistant to virtually all available antibiotics on the market (Breidenstein *et al.*, 2011).

Among a variety of mechanisms of bacterial resistance, the production of β-lactamases is the main resistance factor of Gram-negative bacteria to β-lactam antibiotics.

Indeed, resistance to β-lactam antibiotics has increased in recent years, being mediated by a variety of mechanisms, most commonly the cleavage of β-lactam ring, antibiotics efflux and/or reduced drug uptake due to loss of outer membrane porin proteins (Pfeifer *et al*., 2010). The large number of bacteria resistant to multiple antibiotics represents a challenge in the treatment of infections, since the rate of obtaining new antibiotics today cannot match the increasingly large number of resistant strains. Our next step must include the careful use of antibiotics in clinical and agricultural fields as well as the search for novel drugs.

1.2 Antimicrobial peptides

AMPs have emerged as an alternative strategy for the treatment of infections caused by resistant bacteria. These peptides are evolutionarily ancient molecules that have been isolated from microorganisms, plants, invertebrates, fish, amphibians, birds and mammals, including humans. They play an important role in the innate immune system and are the first line of defence to protect internal and external surfaces of the host (reviewed in Silva *et al*., 2011). The AMPs may have a broad spectrum of antibacterial and antifungal activities. Moreover, in some cases, antiviral, antiparasitic and antitumor activities have also been observed (Nijnik & Hancock, 2009). Despite the enormous diversity in their sequences and structures, the majority of AMPs show a positive charge (+2 to +9), 12-100 amino acid residues and variable three-dimensional structures. Among them are included α-helices (*e.g.*, magainin, cecropin and cathelicidin), β-sheets (*e.g.*, hepcidin and human α-defensin 1), a combination of α-helices and β-sheets (*e.g.*, human β-defensin 1 and plant defensins), head-to-tail cyclized fold (*e.g.,* cyclotides), as well as extended and flexible loops (*e.g.,* indolicidins) (reviewed in Silva *et al*., 2011). In addition to their action against microorganisms, AMPs have activities related to innate and adaptive immunity (immunomodulatory activity) that include the induction or modulation of proinflammatory cytokines and chemokines production, chemotaxis, apoptosis, inhibition of inflammatory response, recruitment and stimulation of proliferation of macrophages, neutrophils, eosinophils and T lymphocytes (Nijnik & Hancock, 2009).

The AMPs have a wide variety of mechanisms, showing that they clearly act bound to the lipid bilayer, using it as a primary target and leading to a membrane disruption (reviewed in Silva *et al*., 2011). It was at first believed that the initial AMP mechanism of action was solely on the cell membrane. However, AMPs can also perform their functions through interactions with intracellular targets or by disturbing cellular processes, as well as causing synthesis inhibition of the cell wall, nucleic acids or proteins (Brogden, 2005).

The AMPs are molecules of great relevance to the pharmaceutical, biotechnology and food industries. The structural diversity and chemical nature displayed by these molecules is a condition that has led researchers to consider them as natural antibiotics, an innovative alternative to conventional antibiotics as a new class of drugs to prevent and treat systemic and topical infections (Gordon *et al*., 2005). Due to these facts, some AMPs are already utilized with clinical and commercial purposes, including ambicin (nisin), polymixin B and gramicidin S (Bradshaw, 2003). However a restriction on the use of AMPs for therapeutic use is their limited stability (especially when composed of L-amino acids), toxicity against eukaryotic cells, susceptibility to proteolytic degradation and development of allergies. Thus, the rational design of AMPs emerges as an important tool that aims to develop AMPs with maximum performance against resistant bacteria.

1.3 Computer-aided identification and design of AMPs

Rational design of AMPs is a modern approach to antibiotic development, nevertheless, a more detailed target characterization is needed. Indeed, a target with sufficient differences between the host and the pathogen is necessary, in order to reduce or abolish adverse effects, according to the principle of selective toxicity. The principal barrier to the use of AMPs as antibiotics lies in their cytotoxicity for mammalian cells. This is perhaps not surprising since AMP activity is mostly dependent on membrane-peptide interaction. However, for AMPs become useful as broad-spectrum antibiotics it would be necessary to dissociate toxicity to mammalian cells from antimicrobial activity, which could be reached by increasing antimicrobial activity or reducing haemolytic activity, or both (Chen *et al*., 2005). Another obstacle to the use of AMPs as antibiotics is their susceptibility to proteolysis, since peptides formed by L-amino acid are sensitive to degradation and clearance of serum components. These problems can be solved through amino acid substitutions, including replacement of L-amino acids to D-amino acids. These substitutions may promote alterations in amphipathicity/hydrophobicity, leading to a reduction in the cytotoxicity of the peptides to mammalian cells, without changing the antimicrobial activity, besides leaving the AMPs less susceptible to proteolytic degradation (Chen *et al*., 2005; Pag *et al*., 2004).

The first studies of rational design of AMPs generated several analogues of known AMPs (*e.g.,* cathelicidins, defensins, magainins and cecropins). Nevertheless, many of them were less active than the original prototype. In fact, these studies played a critical role in identifying the AMP properties involved in antimicrobial activity. These properties served as the basis for developing approaches for antimicrobial activity prediction, through several methods, such as support vector machine (SVM, Lata *et al*., 2007; Porto *et al*., 2010; Thomas *et al*., 2010), artificial neural network (ANN, Fjell *et al*., 2009; Torrent *et al*., 2011) and quantitative structure-activity relationship (QSAR, Jenssen *et al*., 2007) as will be further detailed. By using machine learning methods, this field became more scientific than descriptive. Nevertheless, the AMP mode of action is still an open subject since there are no definite models of prediction or rational design, and so novel methods tend to appear. Certainly, the AMPs emerge as a promising class of therapeutics, despite their limitations. Methods of prediction and rational design play a crucial role in improving AMP performance against resistant bacteria. Therefore, designing novel AMPs requires progress in methods for identifying the best candidate peptides prior to synthesis and then testing them against bacteria. Methods of rational design and prediction of AMPs emerged from early 90s and 2000s, respectively, and they will be reviewed in the next sections.

2. Methods of rational design

Rational design methods aim to create novel peptides with improved antimicrobial activity, lower toxicity to human cells and reduced size. In other words, it is much more specific in creating a pharmaceutical with higher specificity to microorganisms, avoiding side effects. This review classifies the rational design methods into three major classes: physicochemical, template-based and *de novo* methods. The first two methods use a previously known AMP as the basis for designing studies. While physicochemical approaches generate several analogues with different physicochemical properties, the template-based methods search for

size reduction, adding selectivity and/or killer activity to known sequences. Furthermore, *de novo* methods that generate AMP without a template sequence, using only frequencies or patterns. Essentially, these three classes define the rational design methods, but there are also hybrid methods.

2.1 Physicochemical methods of rational design of AMPs

The first rational design methods were based on the most commonly proposed AMP mechanism of action, which is membrane disruption. This process is first mediated by electrostatic interactions among positive charged residues and negatively charged lipid heads, and then by insertions of hydrophobic residues into the membrane. The majority of physicochemical methods use α-helical peptides as the basis for study. Since α-helical peptides present wide distribution and the broadest activities spectrum, their physicochemical properties can be easily measured. In addition to charge and hydrophobicity, another property that can be easily measured is the hydrophobic moment, given by Eisenberg's equation (Eisenberg *et al*., 1982):

$$\mu_H = \sqrt[2]{[\sum_i H_i \cos(\delta i)]^2 + [\sum_i H_i \sin(\delta i)]^2} \quad (1)$$

Where δ is the angle separating side chains along the backbone (100° for α-helix); *i* is the number of residues and H_i is the hydrophobicity of amino acid *i* in a determined hydrophobicity scale, such as Eisenberg's (Eisenberg *et al*., 1982) or Kite-Doolittle's (Kite & Doolittle, 1982). In fact, it is more common to use a normalized hydrophobic moment, dividing it by the total amino acid residues. These physicochemical properties are, apparently, directly involved in interactions between α-helical AMPs and bacterial membranes, by some "rules". First, increasing hydrophobicity boosts the lipid's affinity. Second, enhancing the hydrophobic moment may favour the α-helix peptide fold, and third, increasing the net charge could lead to a higher interaction with anionic membranes (Drin and Antonny, 2010).

Using this approach, Dathe *et al*. (1997) developed several magainin 2 analogues and an 18-residue model peptide with KLA repetitions, modulating their activity by changing only hydrophobicity, hydrophobic moment and the angle of positively charged face helix (Figure 1). Moreover other features were conserved, such as helix propensity and total charge. This showed that when the hydrophobicity and hydrophobic moment increase, the antimicrobial and haemolytic activities from those peptides also increase (Dathe *et al*., 1997). It also showed that the angle of positively charged face has little influence on antimicrobial activity. Haemolytic activity increases if the angle is more obtuse than the original. Nonetheless, it varies according to peptide. For example, an angle of 120° applied to KLA model peptide increases haemolytic activity, but the same angle applied to magainin 2 does not affect its activity (Dathe *et al*., 1997). On the other hand, very low hydrophobicity abolishes the antimicrobial activity of those peptides, which can be compensated by increasing the hydrophobic moment. Therefore, while increasing those parameters the peptide becomes unspecific, a selective peptide may be reached with moderated hydrophobicity, increasing the hydrophobic moment and keeping the angle of charged face small. Further, the same group would show that changes in net charge of magainin 2 also

modulate its activity (Dathe *et al.*, 2001). This study designed six magainin 2 analogues, keeping helix propensity, hydrophobicity, hydrophobic moment and the angle of charged residues, and changing only the net charge. So for each charge modification, one or more amino acid substitutions were required to keep the other parameters (*i.e.*, the MK6 analogue has a charge of +7, while magainin 2 has +4; the identity between them is only 39%, but the other properties were very similar). A charge threshold was observed to develop an analogue with specificity to bacteria, and increasing the charge from +3 to +5 made the peptide more active against bacteria and less toxic to erythrocytes. Nevertheless increasing its charges to +6 or +7 could generate a very haemolytic analogue. The relation between the angle of charged face and haemolytic activity has a bias: the net charge must be great, but not too great. In the case of magainin 2 this threshold is +5. Therefore, increasing the peptide charge makes the peptide lose its specificity to bacterial membranes. Perhaps, when the charge is too positive, the neutral membrane ends up interacting with the peptide as an acidic membrane.

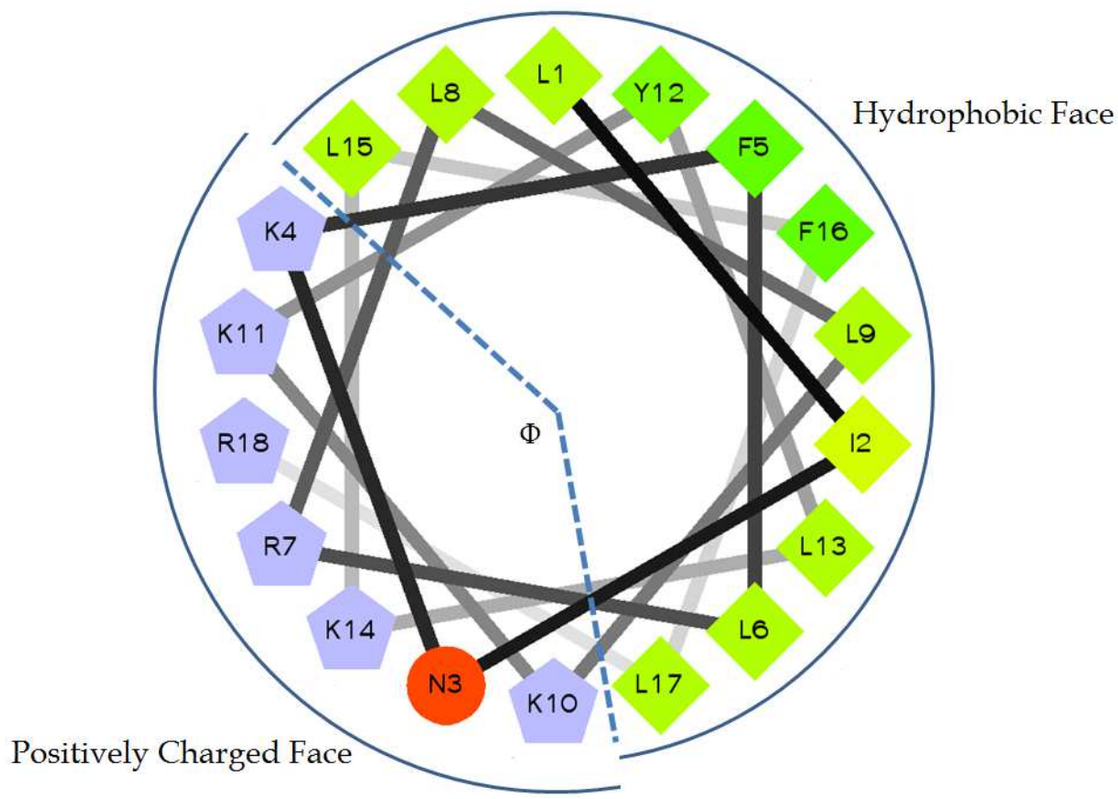


Fig. 1. Schematic representation of an α-helical amphipathic peptide. The angle of polar face is indicated by Φ. Positive charged residues are represented in pentagons, polar ones as circles and nonpolar as diamonds.

Giangaspero *et al.* (2001) also observed similar results in their study about α-helical peptides with non-proteinogenic amino acids. In fact, this work employs a hybrid method, first using a *de novo* technique and, subsequently, a physicochemical one. *De novo* design uses a model developed through amino acid frequencies by type of amino acid (structure determining,

hydrophobic, hydrophilic, positively charged, negatively charged and polar uncharged). Then the model was filled up with a restricted set of amino acids: norleucine to hydrophobic positions and ornithine, glutamine or glutamic acid to hydrophilic ones. These residues were chosen since they have the same side-chain length, ensuring a homogeneous cross-section to the helix. This step generates two peptides, P19 (5) and P19 (6). In the next step, 18 novel sequences were derived from the initial model in order to verify the effects on activity by charge, helicity, amphipathicity, hydrophobicity and size reduction. Four sequences were developed with different charges: +1, +3, +8 and +9. As observed by Dathe *et al.* (2001), charge reduction leads to a decrease in antimicrobial activity. However, Giangaspero *et al.* (2001) propose that the activity is independent of positioning of charged residues within the helical domain. The addition of two ornithine or glutamic acid residues to the N-terminal of those peptides can increase or decrease the activity, respectively: adding two ornithine residues to the analogue with charge +1, its charge became +3 and it became active, with a similar spectrum of analogue with charge +3.

The amphipathicity and hydrophobicity were tested by developing a shuffled peptide version of P19 (6). This peptide has a moderate, and restricted to Gram-negative, antimicrobial activity when compared to P19 (6), even with the same amino acid composition, charge and hydrophobicity, showing that amphipathic arrangement is important to activity (Giangaspero *et al.*, 2001). Size reduction also was tested, by deleting either N- or C-terminal from the most active peptide. This reduces or abolishes the activity, however, switching polar to nonpolar residues, resulting in activity recovery, being similar to or better than the original peptide. These data indicate that in small peptides, there must be equilibrium among charges, helix formers and hydrophobic residues. Helicity modifications were also measured, showing that an increase in the helix propensity also increases the antimicrobial potency. However, it has little additional effect on peptides that have a high helix propensity. On the other hand, decreasing the helix propensity, by proline or D-amino acid insertions could clearly decrease the antimicrobial activity.

Nonetheless, Chen *et al.* (2005) observed a different relationship between helicity and antimicrobial activity in their study about analogues of V_{681}, a designed amphipathic α-helix antimicrobial peptide. Its nonpolar face comprises 12 amino acid residues, while the polar face shows 14 of these (Figure 1). The central residue from each face was chosen for substitutions, Ser^{11} and Val^{13} for polar and nonpolar faces, respectively. Amino acid substitutions were made by increasing or decreasing the peptide's hydrophobicity and/or amphipathicity. Each analogue were generated by only one amino acid substitution, being divided in two groups, the ones with alterations in the polar face named S11X, where 'S' is replaced by 'X'; and the second group with alterations in the nonpolar face named V13X, with the same logic as S11X analogues. Five L-amino acids (Leu, Val, Ala, Ser and Lys) plus glycine were selected to replace the central residues, representing a wide range of hydrophobicity, on a decreasing scale in the following order: Leu > Val > Ala > Gly > Ser > Lys. Moreover, D-enantiomers of each selected L-amino acid were also incorporated in the same positions in order to disrupt helical structures generating a total of 20 analogues. It was observed that some D-amino acid analogues were stronger than their L-amino acid equivalents. Probably, D-amino acids analogues overcome the helix disruption through other properties, such as hydrophobicity or amphipathicity. Moreover, they also observed that changes in the hydrophilic face of V_{681} does not reduce peptide activity against Gram-

positive or -negative bacteria or human erythrocytes, in contrast to changes in hydrophobic face. A similar result was observed by Blondelle *et al.* (1996).

The D-enantiomer of analogue V13K was used by Jiang *et al.* (2011) as the basis of another physicochemical study. The all-D-enantiomer analogues were developed in order to create peptides with specificity to Gram-negative bacteria. Five analogues (D11, D14, D15, D16 and D22) were designed to investigate the influence of charge (analogue D11), hydrophobicity (analogue D22), insertions of charged residues into nonpolar face (analogue D14) and composition of the nonpolar face (analogues D15 and D16). They observed that when charge and hydrophobicity increase (comparing V13K and D11), antimicrobial activity also increases, as observed by Dathe *et al.* (2001) and Giangaspero *et al.* (2001). By increasing hydrophobicity (comparing D11 to D22), haemolytic activity increases, confirming the data proposed by Dathe *et al.* (1997). However, by introducing a second lysine in the nonpolar face (comparing D22 to D14), hydrophobicity can be kept higher and haemolytic activity can decrease. Finally, the composition of the nonpolar face (comparing D11 and D14 to D15 and D16, respectively), D15 and D16 were generated by switching all large side-chain hydrophobic residues for leucine residues. Those changes increase hydrophobicity and antimicrobial activity, but they have different effects on haemolytic activity, while D15 becomes more haemolytic and D16 becomes less haemolytic, probably due to the presence of second lysine residue in the polar face of D16. These data show that the same physicochemical rules can be applied to D- or L-enantiomers.

Few studies with another kind of folding have been reported. In 1999, Wu & Hancock carried out a study based on bactenecin, a 12-amino acid residue peptide that adopts a β-turn structure cyclized via disulphide bond. Linear analogues of bactenecin show its activity depleted. However, C-terminal amidation partially restores the activities. In this study, several changes in both forms (linear and cyclic) of bactenecins were evaluated. Several analogues were designed to test the importance of ring size (numbers of amino acids between the cysteine residues), charge, and amphipathicity. The results are similar to α-helical peptides, in which it was observed that increasing the charge leads to an improvement in antimicrobial activity. Moreover, the same study showed that the positions of charged residues are more important than the number of positive charged residues. Increasing the ring size by insertion of a tryptophan in the middle of the ring increases the activities, while a proline residue insertion was able to abolish the activity. Additionally, the cyclic analogues also have agglutination activities, in contrast to linear versions. The linear analogue Bac2A-NH2 was the most desirable candidate generated in this study, due to its broad spectrum of activity and absence of agglutination activity. Further, several analogues of Bac2A were developed through point substitutions, scrambling, and deletions in sequence; IDR1018 is the most promising of all Bac2A analogues. Besides bactericidal activity, IDR1018 also displays chemokine induction activity and suppresses pro-inflammatory responses to Gram-negative bacteria (Wieczorek *et al.*, 2010).

Conversely, this kind of analysis, considering the minimum inhibitory concentration (MIC) as a consequence of structural and physicochemical properties, leads us to false conclusions. MIC values can be very similar for peptides with different properties. This is easily observed when peptides KLA12 and KLA7 (Dathe *et al.*, 1997) are compared. They have similar MICs; nevertheless, their hydrophobicity and hydrophobic moment are different. The hydrophobicity of KLA7 is a half of KLA12, while its hydrophobic moment is 1.15 times

higher than KLA12. Moreover, this kind of study is almost completely restricted to α-helical peptides. The lack of study of other varieties of folding might bring novel information about the relationship between physicochemical properties and antimicrobial activities.

2.2 Sequence template methods of rational design of AMP

Sequence template methods involve generating novel AMPs based on a known sequence, whether of an active or an inactive peptide. These approaches can seek to reduce size, add selectivity and/or increase the activity. In several cases, the information generated by physicochemical methods can be used to reach these objectives, by switching residues, changing the net charge or pursuing minor peptides with the same properties, without performing a physicochemical study itself.

In 1996, Thennarasu & Nagaraj developed three analogues of pardaxin by switching some amino acid residues for others with different properties. Pardaxin is a toxic peptide secreted by the sole fish from the genus *Pardachirus*. At low concentrations pardaxin is able to form ion channel-like structures and at high concentrations that causes cell membrane disruption. This toxin can also induce neurotransmitter release from neurons. Firstly, the authors identified the probable region responsible for membrane permeation activity. Preliminary studies have shown that the C-terminal region did not have this activity, since the positive charges were concentrated at N-terminal. Then, the first designed analogue was the N-terminal 18 residue segment, named 18P. The second analogue, 18A, was designed by switching the residue Pro^7 to an alanine residue, since proline residues cause structural distortions to helix backbone. The last analogue, 18Q, was developed switching the two lysines (Lys^8 and Lys^{16}) for glutamines in the 18A sequence, since glutamine residues play an important role in channel formation by peptides with neutral charges. Having designed these analogues, their activities were examined against *Escherichia coli*, *Staphylococcus aureus* and human erythrocytes. 18P analogue showed activity only against *E. coli*, while 18A showed haemolytic activity in addition to antimicrobial activity against *E. coli*. On the other hand, 18Q showed only haemolytic activity. No activities against *S. aureus* were observed. Although the minimum identity among the sequences was 83.3% (18P and 18Q) the activities were different. While 18P showed simply antimicrobial activity, 18Q had only haemolytic one. These differences can be explained by their intrinsic structures. Circular dichroism (CD) analysis showed that 18P had a low propensity to occur in helical conformation when compared to 18A, even though both have a typical helical CD spectrum, while 18Q adopted a clear β-structure, probably forming an amphipathic β-sheet, even in ~65% of 2,2,2-trifluoroethanol, indicating the importance of structure-activity relationship.

Ueno *et al.* (2011) developed a strategy that does not generate great conformational changes in relation to original sequences. This strategy is based on acid-amide substitutions by switching aspartic acids and glutamic acids to asparagine and glutamine, respectively. Since these substitutions are conservative, the structure has few changes and if the original peptide has basic residues, there will be an increased charge in the novel peptides. This strategy was successfully applied to three pro-regions of nematode cecropins. The pro-regions are inactive against bacteria and human erythrocytes. After modifications, the sequences became antimicrobial peptides with slight haemolytic activity. CD spectra reveal that the structure of original peptides and its analogues are similar.

Likewise, Ahn *et al*. (2006) developed an AMP based on the 11-residue α-helical domain from tenecin 1, an insect defensin isolated from *Tenebrio molitor*. This α-helical domain, named L1, shows no antimicrobial activity. In defensins, the activity is related to the γ-core motif, comprised of a β-hairpin (Yount & Yeaman, 2004). However, L1 shows physicochemical properties similar to well-known AMPs, except the net charge. L1 has a net charge of +2, while AMPs have a charge of +4 or +5. Three analogues of L1 were developed (L2, L3 and L4) by switching some residues. L4 was the most active analogue, showing even greater activity than tenecin 1. L4 was developed by switching an aspartic acid and a histidine for lysine residues, increasing the charge to +3. Thereafter, L4 showed activity against bacteria and fungi, including *E. coli*, *Pseudomonas aeruginosa* and *Candida albicans,* besides activity against *S. aureus* and *Micrococcus luteus,* while tenecin 1 had no activity toward these pathogens.

Another work involving defensins was developed in 2008 by Landon *et al*. In this study, 70 chimeric defensins were designed by combining conserved regions of *Anopheles gambiae* defensin and variable regions of other insect defensins. From these, 45 were expressed in yeast *Saccharomyces cerevisiae*. Five of them were selected for study. These five hybrid defensins originated from combinations of *A. gambiae* defensin (DEF-AAA) with defensins from *Belostoma gigas*, *T. molitor*, *Acrocinus longimanus* and/or *Drosophila melanogaster*. All hybrid defensins have the same structural scaffold of a cysteine-stabilized αβ motif. On the other hand, their activities against *S. aureus* multi-resistant strains were different. Two analogues were more effective *in vitro* against *S. aureus* (DEF-AcAA and DEF-DAA). DEF-DAA was toxic to mice models, with a lethal dose of 30 $mg{\cdot}kg^{-1}$, while DEF-AcAA showed a lethal dose higher than 100 $mg{\cdot}kg^{-1}$. Indeed, since the active site of defensins consists of the γ-core, these two hybrid defensins have an identical γ-core sequence, indicating that the N-terminal loop of defensins may also contribute to the activity. So the *in vivo* activity of DEF-AcAA and DEF-AAA was evaluated against *S. aureus* peritonitis model on mice. The results showed that both defensins have the same efficacy for *S. aureus* multi-sensitive strain, with a dose of 3 $mg{\cdot}kg^{-1}$. Nevertheless, on the same model with a multi-resistant strain DEF-AcAA was shown to be the most effective with a dose of 3 $mg{\cdot}kg^{-1}$, while DEF-AAA needed a dose of 10 $mg{\cdot}kg^{-1}$. These results demonstrate that these AMPs are more efficient than vancomycin, which requires a dose ranging from 10 to 30 $mg.kg^{-1}$ for treatment on the same model.

Also focusing on development of peptides with potential for systemic use, Sigurdardottir *et al.* (2006) identified a 21-amino acid fragment of human cathelicidin antimicrobial peptide LL-37 with similar or stronger activities than the complete peptide. LL-37 is an attractive candidate for treatment of sepsis, due to its broad spectrum of antimicrobial activity, the immune system's cell chemotactic abilities and also abilities to bind and neutralize bacterial lipopolysaccharides. However, LL-37 also has cytotoxic activity against eukaryotic cells. Therefore, using the helical propensity prediction of AGADIR (Lacroix, 1998) and the amino acid preference for α-helix terminals, they identified a fragment starting from Gly^{14} going up to Arg^{34}, named GKE. For comparisons, two other 21-amino acid fragments were derived, one from the N-terminal (LLG) and other from C-terminal (FKR). GKE was more active than LL-37 against bacteria and fungi. Moreover, GKE and LL-37 showed similar chemotaxis and inhibition of nitric oxide production activities. The same patterns were not observed for LLG and FKR fragments. Interestingly, all fragments showed 100% of identity to LL-37. However, they differed in helix propensity, although GKE and FKR do not present much

difference in helical content and have 85.7% of identity. The amino acid preference for α-helix terminals might explain those differences in activity.

Another successful strategy is the use of synthetic combinatory libraries (Blondelle *et al*., 1996). A synthetic combinatory library is composed of a set of mixtures of peptides generated from a template. In this approach, several positions of template sequence were chosen for amino acid substitutions. One of them was chosen for individual defined substitutions, where each modification is controlled, generating one subset for each modification. Thus it can generate 20 subsets, by using the proteinogenic amino acids. The other selected positions are randomly filled up in each subset. Following this strategy, Blondelle *et al*. (1996) used the sequence 'YKLLKKLLKKLKKLLKKL-NH_2' as a template for design of two synthetic combinatory libraries, the first one with changes in hydrophobic face and the second one in hydrophilic face. In hydrophobic face, Leu^4 was chosen for individual defined substitutions (represented by "O"), while Leu^7, Leu^{11} and Leu^{14} were chosen for random filling (represented by "X"); this library was represented by the sequence $YKLO^4KKX^7LKKX^{11}KKX^{14}LKKL\text{-}NH_2$. The same logic was applied to hydrophilic face, being represented by the sequence $YKLLKO^6LLX^9KLKX^{13}LLX^{16}KL\text{-}NH_2$. The assays showed that substitutions in hydrophobic face depleted the activity of the original peptide. During the second stage of design, the library with changes in polar face was redefined by changing the residue used for individual defined substitutions (positions 6, 9, 13 and 16). In this way, the best residues for each position could be selected. Thus leucine was selected at position 6 as a representative hydrophobic residue, proline at position 9, proline and glycine at position 13, and phenylalanine, isoleucine and proline at position 16, six peptides being generated by combining those selected residues (1 x 1 x 2 x 3 = 6). Five of the six designed peptides showed a 10-fold improvement in activity.

Indeed, sequence template methods can generate novel AMPs with enhanced activity against bacteria and lower effectiveness toward mammalian cells than original sequences. However, there are no directives for using this kind of method. Which residues are important for activities? What must be substituted to obtain a higher deleterious activity and lower side effects? Are the physicochemical properties useful for this kind of study? Can the methodology applied to peptide A be applied to peptide B? For these methods, it is necessary to identify the governing principle that allows the enhancement of antimicrobial activity. In fact, there must be something beyond switching residues or development of chimeric proteins. Identifying this principle will be helpful for developing novel methods of rational design.

2.3 *De novo* methods of rational design of AMP

De novo methods are very interesting in terms of achieving a yield from multiple AMPs with little amino acid conservation. Instead of using one pivotal sequence to develop analogues, *de novo* methods can use amino acid patterns or amino acid frequencies and positioning preferences, generating several sequences with no clear relation.

Tossi *et al.* (1997) developed a *de novo* method that considers the length of the peptide, its cationicity, its amphipathicity and its helicity, in addition to sequence patterns. First, 20 residues from the N-terminal of 85 natural α-helical AMPs were aligned without gap insertions or any attempt to improve the alignment. Based on this alignment, frequencies

and kinds of amino acids were extracted, in order to create a novel pattern. Despite the simplicity of this method, a well-defined pattern of residue distribution was developed. Next, the pattern was filled up with the most frequent amino acids in each position, generating an AMP with 20 amino acid residues. The same method was applied to mammalian cathelecidin with some modifications, using a reduced number of sequences and adding gaps in the alignment, creating three novel patterns ranging from 18 to 22 amino acids (Tossi *et al*., 1997). Thus, the patterns were filled up with the most frequent amino acids. The helicity was evaluated by using secondary structure predictions and helical wheel diagrams. The four designed peptides showed a potent and broad-spectrum activity against Gram-positive and Gram-negative bacteria.

Some years later, Loose *et al*. (2006) developed a similar but more sophisticated *de novo* method, the linguistic model. According to this model, AMPs seem to be a formal language with grammar composed of several rules (patterns) and a vocabulary (amino acids). Instead of using alignments to define the patterns, the TEIRESIAS algorithm was used for pattern discovery (Rigoutsos & Floratos, 1998). Thus, ~700 grammars sequences were established, and then all possible grammar sequences with 20 amino acid residues were written out. Sequences with at least 60% of identity with natural AMPs were removed, resulting in 12 million remaining sequences. Next, by removing sequences with at least 70% of identity, 41 candidates were obtained. From these, one peptide was insoluble, but 18 had MIC at maximum of 256 $\mu g \cdot ml^{-1}$ and the remaining peptides showed no activity against *E. coli* and *Bacillus cereus*. Through this method, novel antimicrobial peptides were designed without any information about their structures (Loose *et al*., 2006). In fact, as well as generating novel AMPs, this work was of great importance in that it explained some results of physicochemical methods of rational design. For each grammatical peptide, a non-grammatical peptide was designed with the same amino acid composition, by shuffling the sequence. Giangaspero *et al.* (2001) had already used this strategy and the shuffled peptide showed a reduced activity. However, Loose *et al*. (2006) generated shuffled peptides with no grammars, expecting that they had no activity because they were non-grammatical peptides. As result, only two shuffled peptides were active. From this, it could be seen that there are no direct relations between scalar physicochemical properties and antimicrobial activity, because shuffled and grammatical peptides have the same charge, hydrophobicity, size and molecular mass (Loose *et al*., 2006). This explains why the conclusions of physicochemical methods were not completely correct, and in some cases, controversial (*i.e.,* equal MICs but different physicochemical properties and *vice-versa*). The scalar physicochemical properties had led researchers to false conclusions because they have a secondary role in activity.

On the other hand, a property widely used in physicochemical methods does change when the sequence is shuffled, and that is the hydrophobic moment, a vector property. Loose *et al*. (2006) also used the hydrophobic moment. In a second step of rational design, the best designed sequence was submitted to a redesign process to increase its activity using a heuristic approach, and one of proposals of this redesign process was to "improve the segregation of positive and hydrophobic residues based on a helical projection" or, in other words, to improve the hydrophobic moment. The fact that no structural information is needed is certainly an advantage, but this method has some limitations, such as the difficulty in designing larger proteins with complex structures. So this method is restricted

to generating AMPs similar to those that are deposited in the main data set, *i.e.*, the two most active peptides obtained through this method have 50 and 60% of identity to natural AMPs.

All these methods have been helpful, in their time, in reaching a better understanding of relationships between sequence, structure, physicochemical properties and antimicrobial activity. Overall, these methods have been effective in designing potent AMPs able to kill bacteria at low concentrations. Furthermore, they have also been helpful in the development of antimicrobial prediction tools, as will be seen in the next section.

3. Methods to predict antimicrobial activity

The understanding of antimicrobial peptides' behaviour led some groups to propose different approaches to predict antimicrobial activity, and this field saw much progress in last years. Several methods of antimicrobial activity prediction emerged from studies of rational design, mainly the physicochemical and *de novo* methods. The rules extracted from rational design methods can be extrapolated to other sequences with good reliability by computer-aided predictions. As a result, several tools have been developed, such as prediction tools from Collection of Antimicrobial Peptides (CAMP, Thomas *et al*., 2010) and AntiBP Server (Lata *et al*., 2010). Overall, there are two main strategies for predicting AMPs, the empirical methods and the supervised machine learning ones.

3.1 Empirical methods of AMP prediction

The empirical methods are qualitative, being based only on characters of AMPs without taking into account peptides without antimicrobial activity. In fact, these models are based on rules or patterns correlated to antimicrobial activity. However, the methods cannot be extrapolated to other classes of AMP, being restricted to the class that generated the model. Moreover, they have no standard accuracy measurement, since there is no larger set of non-antimicrobial sequences to test them. In fact, there is no accuracy value, making it complicated to compare the methods, which are summarized in Table 1.

The most simple prediction method is that employed by the Antimicrobial Peptides Database prediction tool (Wang & Wang, 2004). In this case no artificial intelligence was used. It is based only on logical questions about the sequence. It returns a positive prediction whenever a sequence is less than 50 amino acid residues in length, has hydrophobicity below 75% and a cationic net charge. Moreover, the prediction is also going to be positive if the sequences present an even number of cysteines. This method seems to be merely based on rules extracted from physicochemical studies. However, this method neglects some AMPs (*e.g.*, anionic and hydrophilic antimicrobial peptides).

Despite these clear limitations the APD prediction tool was used by Nagarajan *et al.* (2006) for validating their prediction method. This method was developed in order to mine protein data sets, being based on Fourier transformations and Euclidian distances. Comparisons were made with a power spectrum generated by the Fourier transformation of five indices. The indices were based on hydrophobicity, charge, polarity, cysteine content and amino acid distribution. For analysis, the method was applied to six antimicrobial peptides with 16 amino acid residues. In all cases, the power spectrum shows a peak at period 5, and the

major contribution to power spectrum is given by hydrophobicity index. Those power spectra were used to generate a reference power spectrum used in further comparisons. A set of 10,000 random peptides with 16 amino acid residues were generated by PERL scripts. The power spectrum of each random sequence was obtained and compared to the reference through Euclidian distance. From 10,000 random sequences, only three hits were obtained. Two of three hits had positive charge and were predicted as AMPs by the APD prediction tool. However, the three hits showed at least 30% of identity to a known AMP.

Similarly, Fernandes *et al.* (2009) developed a classification method based on fuzzy modelling, also focusing on data set mining. This approach is based on the linguistic model developed by Loose *et al.* (2006) and also on the peptide's amphipathicity. It made each screening into a data set, searching for sequences with a defined pattern. The found sequences were then classified by fuzzy modelling. This consists of a surface plot generated by two membership functions: a triangular function relating the ratio of polar to charged residues and a Gaussian hydrophobicity membership. The best candidates fall into a region between 2:1 and 1:1 polar to charged residues and the regions of moderated hydrophobicity in Gaussian membership, identifying the amphipathic sequences. Assuming hydrophobicity to be low or the ratio to be lower than 1, the sequence is a weak AMP; if hydrophobicity is medium and the ratio is adequate the peptide is a specific AMP; and if hydrophobicity or the ratio is high, the peptide is non-specific. The system was tested in NCBI's non-redundant protein data set (NR) and the seed sequence was Cn-AMP1 (Mandal *et al*., 2009). Through this, three sequences were obtained from a total of 7,153,872 sequences in NR.

Another method that involves patterns is the multidimensional signatures developed by Yount & Yeaman (2004). This method is based on recognition of sequence patterns and motifs in three dimensional structures to correlate them to antimicrobial activity. It was successfully applied to cysteine-stabilized peptides. In this work, a γ-core motif was recognized by the patterns "X[1,3]GXCX[3,9]C", "CX[3,9]CXGX[1,3]" and "CX[3,9]GXCX[1,3]", where X corresponds to any natural amino acid and the numbers between brackets represent sequence variations (*i.e.,* X[3,9] represents an extension of three to nine residues, being composed of any natural amino acid residue). Based on a data set of 500 antimicrobial peptides with length of up to 75 amino acids residues and cysteine content, prototypic sequences were chosen as representative of their classes. The conserved motif GXC was identified by visual inspection of multiple alignments. However, in some sequences this motif was inverted, so the three patterns of γ-core motif were proposed. Structurally the three patterns are absolutely conserved, corresponding to an antiparallel β-sheet composed of two strands. In order to validate the model, two peptides without previously reported antimicrobial activity were selected, the sweet-tasting protein brazzein and the toxin charybdotoxin, both containing the γ-core motif in their 3D structures. These two peptides exerted direct antimicrobial activity against bacteria and fungi. The method was also validated by identification of γ-core sequence into well-known antimicrobial peptides without known 3D structure. The γ-core motif was identified in tachyplesins before its structure became available. Its three-dimensional structure really exhibits the motif of two antiparallel β-strands.

Jenssen *et al.* (2007) also developed and tested their model *in vitro*. They constructed a mathematical model for prediction based on the statistical methods, principal component analysis (PCA) and partial least squares (PLS). This model was filled up with three major

classes of descriptors, (i) amino acid (charge, hydrophobicity and size); (ii) a series of contact energy for each pair of amino acids and (iii) 78 biophysical inductive and conventional quantitative QSAR descriptors. These data were extracted from a single-substitution Bac2a-library containing 228 peptides. This model was capable of predicting 84% of tested peptides.

Method	How is the evaluation done?	References
APD Prediction Tool	Well-known AMP properties	Wang & Wang, 2004
Fourier Transformation	Physicochemical properties	Nagarajan *et al*., 2006
Fuzzy Modelling	Sequence motifs and physicochemical properties	Fernandes *et al*., 2009
PCA/PLS	Amino acid descriptors, pairs of amino acid descriptors and QSAR descriptors	Jenssen *et al*., 2007
Multidimensional Signatures	Sequence and structure motifs	Yount & Yeaman, 2004

Table 1. Empirical antimicrobial prediction methods.

In fact, the last two methods are the most important among the methods discussed so far, mainly due to the *in vitro* validation of predictions. Without this kind of validation, these methods become only good hypotheses, without contributing much knowledge. However, they can achieve a more accurate prediction when they are more restricted to some class of AMP, without a generalization model.

3.2 Supervised machine learning methods of AMP prediction

Supervised machine learning methods for predicting antimicrobial activities have a well-established validation procedure, allowing these methods to be compared. The reliability of these methods is evaluated by several parameters, the main three being calculated as follows:

$$Sensitivity = \frac{TP}{TP+FN} \times 100 \quad (2)$$

$$Specificity = \frac{TN}{TN+FP} \times 100 \quad (3)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP} \times 100 \quad (4)$$

TP corresponds to the number of true positives; FN, the false negatives; TN, the true negative; and FP, the false positives. However, the evaluation of precision on positive predictions can be done by calculating the positive predictive value (PPV), given by the following equation:

$$PPV = \frac{Sensitivity}{Sensitivity+(100-Specificity)} \times 100 \ \therefore \ \frac{TP}{TP+FP} \times 100 \quad (5)$$

Here, comparisons among the methods are going to be made based on PPV and accuracy values (Table 2), since for discovering novel AMPs it is more important that the probability of a positive prediction be true. Overall, these methods require two data sets, the training set and the blind set. The training set is composed of two subsets, the positive data set (the AMPs) and the negative data set (the non-AMPs). Through these sets the algorithm is trained and then, tested against the blind set. From the results against the blind set, the true positives, negative and false positives and negatives are estimated and then the parameters (*e.g.*, accuracy) are calculated.

There are two major challenges in the usage of supervised machine learning to predict antimicrobial activity: the AMPs' size variation, and the absence of a dataset for non-antimicrobial peptides (Lata *et al.*, 2007). There are at least two choices of positive set, APD (Wang & Wang, 2004) and CAMP (Thomas *et al.*, 2010). Nevertheless, there are no non-antimicrobial data bases to use as a negative set. Another difficulty is the variation in size of AMPs, since the machine learning techniques need fixed length input vectors. Several strategies have been developed in order to overcome these problems.

Lata *et al.* (2007) developed the first supervised machine learning methods for prediction of antimicrobial activity. In this pioneer work, three algorithms were tested: SVM, ANN and quantitative matrices (QM). The positive data set was composed of 436 AMPs from APD and the negative set was composed of an equal number of non-secretory proteins randomly selected from SwissProt. Initially, an SVM model using amino acid composition of whole sequence was built with 20 inputs, one for each amino acid. This model achieves the highest accuracy of all generated models in 5-fold cross validation (89.04%). However, the authors proposed that is impossible to utilize this approach to search for AMPs in genomes or proteins due to the enormous size variation. Thus, it was decided that a fixed length would be used, using binary patterns, where each amino acid is represented by one binary pattern. SVM models with the 5, 10, 15 or 20 first N-terminal residues were constructed. The best accuracy observed in 5-fold cross validation was in SVM model with 15 residues (87.85%). Therefore, another two approaches using SVM were developed, the C-terminal approach (with 15 C-terminal residues) and the N+C-terminal approach (with 30 residues, 15 from N-terminal and 15 from C-terminal). The C-terminal approach achieves an accuracy of 85.16 %, while the N+C-terminal one achieves 92.11% in 5-fold cross validation. The three approaches were applied to QM and ANN. In both cases, the N+C-Terminal approach achieved the best accuracies in 5-fold cross validation, 90.37% and 88.17%, respectively. In a blind data set composed of 24 mature sequences extracted from SwissProt, the N+C-terminal approach had the higher performance in all algorithms, achieving a PPV of 91.66% for all algorithms.

In 2010, this system was improved, but only the SVM was used (Lata *et al.*, 2010). In this new version, the positive data set was composed of 999 AMPs and the negative data set was constructed with an equal number of non-secretory proteins extracted from SwissProt. The blind set was composed of 466 AMPs from SwissProt, none of which were present in the positive set. The N+C-terminal approach continued to show higher accuracy (91.64%). Despite the drop in precision (92.11 to 91.64%), the improved version was more reliable because the number of sequences used in training and testing were higher than the previous version.

Thomas *et al.* (2009) used two other methods in addition to SVM. Random forest (RF) and discriminant analysis (DA) were implemented. RF showed the finest accuracy (93.2%), followed by SVM (91.5%) and DA (87.5%). The positive data set was composed of 2578 AMPs and 4011 sequences derived from SwissProt or randomly generated sequences. 70% of each set was used for training the machines and the other 30% composed the blind set. The algorithms were trained with 275 features, including composition, physicochemical properties and structural characteristics of each amino acid. In contrast to that of Lata *et al.* 2007 and 2010, the method developed by Thomas *et al.* (2009) was able to predict antimicrobial activity for sequences with variable size.

Method	Positive Set	Negative Set	Accuracy (%)	PPV (%)	References
SVM	APD	SwissProt	92.11	92.11	Lata *et al.*, 2007
QM	APD	SwissProt	90.37	90.65	Lata *et al.*, 2007
ANN	APD	SwissProt	88.17	88.17	Lata *et al.*, 2007
SVM	APD	SwissProt	91.64	90.43	Lata *et al.*, 2010
RF	CAMP	SwissProt and Random Sequences	93.20	95.13	Thomas *et al.*, 2010
SVM	CAMP	SwissProt and Random Sequences	91.50	93.42	Thomas *et al.*, 2010
DA	CAMP	SwissProt and Random Sequences	87.50	87.45	Thomas *et al.*, 2010
SVM	APD	Random Transmembrane Sequences	83.02	96.54	Porto *et al.*, 2010
ANN	CAMP	SwissProt	89.20	88.60	Torrent *et al.*, 2011
ANN	AMP Library	Non-AMP Library	86.50	92.94	Fjell *et al.*, 2009

Table 2. Supervised machine learning methods of antimicrobial prediction.

Our group has developed an SVM model based on physicochemical properties for prediction of peptides with cysteine knot motifs (Porto *et al.*, 2010). Despite the absence of direct correlation between antimicrobial activity and physicochemical properties, their use solved the problem of size variation. However, it generated another problem, which is that shuffled sequences have the same scalar properties, since they are simple averages and the order of residues does not imply average modifications. That problem is avoided by including the hydrophobic moment, since the modification of sequence clearly modifies the hydrophobic moment. For the second challenge, a set of predicted transmembrane proteins was used as the set of non-antimicrobial peptides, since the transmembranes are non-secretory proteins. Through this approach, an overall accuracy of 83% was observed in a blind dataset. This model can be helpful to predict antimicrobial activity of a wide number of cysteine-stabilized peptides, such as conotoxins, proteinase inhibitors, metallothioneins, defensins and cyclotides. The only requirement is the presence of disulphide bonds in the peptide structure.

Also using physicochemical properties, Torrent *et al.* (2011) developed an ANN with eight properties: isoelectric point (pI), peptide length, α-helix, β-sheet and turn structure propensity, *in vivo* and *in vitro* aggregation propensity and hydrophobicity. The main data set was composed of 1157 AMPs from CAMP and 991 non-AMPs from SwissProt. The training set was composed of 1074 peptides, while the testing and validation sets contained 537 peptides. The system achieved an accuracy of 90%. Indeed, the aggregation propensity was seen to be crucial for this method in much the same way as the hydrophobic moment in the method developed by Porto *et al.* (2010). The aggregation propensity changes if the sequence is shuffled, but the other six properties do not. When the aggregation parameter is removed, the system's reliability decreases.

The methods discussed so far show that AMP size variation problem is easy to solve, by using fixed sizes or physicochemical properties. Both strategies achieve similar accuracies. On the other hand, the non-AMP data set only seems to be easier to solve by using random proteins or proteins from SwissProt. However, comparing AMPs to randomly selected proteins from SwissProt is almost the same as comparing oranges to strawberries; it is relatively easy to distinguish each, generating high accuracies. Moreover, as shown in section 2, two peptides with high identities and subsequently similar properties can have different activities, as is the case of peptides derived from pardaxin (Thennarasu & Nagaraj, 1996) and LL-37 (Sigurdardottir *et al.*, 2006).

Lately, a combined approach of QSAR and machine learning techniques has been developed (Fjell *et al.*, 2009). Through 44 QSAR descriptors, an ANN was built based on 1433 random nine-mer peptides. The ANN was trained to predict sequence activity in relation to the control peptide Bac2A. For model evaluation, a library of nine-mer peptides composed of approximately 100,000 sequences was screened *in silico*. These sequences were divided into four classes: (I) most likely to be more active than the control; (II) likely to be more active than the control; (III) likely to be less active than the control; and (IV) most likely to be less active than the control. The topmost 50 positions of each class were synthesized and tested. For class I, an accuracy of 94% was observed, although the overall accuracy was around 85%.

The methods discussed here show that the great difficulties in antimicrobial activity prediction are the absence of a non-antimicrobial database and the enormous variation in sequence size. The greatest challenge for prediction methods is perhaps the heterogeneity of AMPs, which are part of a group with different sequences, structures and mechanisms of action.

4. Conclusions and prospects

In the future, novel treatments against resistant bacteria should be developed, including strategies that use unnatural AMPs as their basis. The development of unnatural AMPs can be carried out by various methods, including those discussed here. This kind of study brings new knowledge and also generates novel AMPs, in turn boosting the development of prediction methods that can help evaluate rational designed AMPs. A more accurate prediction model may be developed when the patterns of the linguistic model can be used to train machine learning techniques. In addition, a more efficient approach to pattern recognition is needed, since a single sequence is insufficient for patterns identification. In this view TEIRESIAS could be used once two or more sequences were need by this approach. This methodology will be helpful not only for novel AMPs development, but also for other

protein chemistry fields. The development of this new methodology is a real challenge and could reduce current limitations, leading us to develop novel and more potent antimicrobial peptide agents.

5. References

Ahn, H.S.; Cho, W.; Kang, S.H.; Ko, S.S.; Park, M.S.; Cho, H. & Lee, K.H. (2006). Design and synthesis of novel antimicrobial peptides on the basis of a helical domain of Tenecin 1, an insect defensin protein, and structure–activity relationship study. *Peptides*. Vol.27 ,No.1, pp. 640-648

Blondelle, S.E.; Takahashi, E.; Houghten, R.A. & Pérez-Payá, E. (1996). Rapid identification of compounds with enhanced antimicrobial activity by using conformationally defined combinatorial libraries. *Biochem J.* Vol.313, pp. 141-148

Bradshaw, J. (2003). Cationic antimicrobial peptides: issues for potential clinical use. *BioDrugs*. Vol.17, No.4, pp.233-240

Breidenstein, E.B.; de la Fuente-Núñez, C. & Hancock, R.E. (2011). *Pseudomonas aeruginosa*: all roads lead to resistance. *Trends Microbiol*. Vol.19, No.8, pp. 419-426

Brogden, K.A. (2005). Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? *Nat Rev Microbiol*. Vol.3, No.3, pp. 238-250

Chen, Y.; Mant, C.T.; Farmer, S.W.; Hancock, R.E.W.; Vasil, M.L. & Hodges, R.S. (2005). Rational design of α-helical antimicrobial peptides with enhanced activities and specificity/therapeutic index. *J Biol Chem.* Vol.280, No.13, pp. 12316-12329

Dathe, M.; Nikolenko, H.; Meyer, J.; Beyermann, M. & Bienert, M. (2001). Optimization of the antimicrobial activity of magainin peptides by modification of charge. *FEBS Lett.* Vol.541, pp. 146-150

Dathe, M.; Wieprecht, T.; Nikolenko, H.; Handel, L.; Maloy, W.L.; MacDonald, D.L.; Beyermann, M. & Bienert M. (1997). Hydrophobicity, hydrophobic moment and angle subtended by charged residues modulate antibacterial and haemolytic activity of amphipathic helical peptides. *FEBS Lett*. Vol.403, No.2, pp. 208-212

Drin, G. & Antonny, B. (2010). Amphipathic helices and membrane curvature. *FEBS Lett.* Vol.584, pp. 1840-1847.

Eisenberg, D.; Weiss, R.M.; Terwilliger, T.C. & Wilcox, W. (1982). Hydrophobic moments and protein structure. *Faraday Symp Chem Soc*. Vol.17, pp. 109-120

Fernandes, F.C.; Porto, W.F. & Franco, O.L. (2009). A wide antimicrobial peptides search method using fuzzy modeling. *LNCS*. Vol.5676, pp. 147-150

Fjell, C.D.; Jenssen, H.; Hilpert, K.; Cheung, W.A.; Panté, N.; Hancock, R.E.W. & Cherkasov, A. (2009). Identification of novel antibacterial peptides by chemoinformatics and machine learning. *Journal of Medicinal Chemistry*. Vol.52, No.1, pp. 2006-2015

Giangaspero, A.; Sandri, L. & Tossi, A. (2001). Amphipathic α helical antimicrobial peptides, a systemic study of the effects of structural and physical properties on biological activity. *Eur J Biochem*. Vol.268, pp. 5589-5600

Gordon, Y.J., Romanowski, E.G. & McDermott, A.M. (2005). A review of antimicrobial peptides and their therapeutic potential as anti-infective drugs. *Curr Eye Res*. Vol.30, No.7, pp. 505-515

Jenssen, H.; Lejon, T.; Hilpert, K.; Fjell, C.D.; Cherkasov, A. & Hancock, R.E.W. (2007). Evaluating Different Descriptors for Model Design of Antimicrobial Peptides with Enhanced Activity Toward *P. aeruginosa*. *Chem Biol Drug Des.* Vol.70, pp. 134-142

Jiang, Z.; Vasil, A.I.; Gera, L.; Vasil, M.L. & Hodges, R.S. (2011). Rational design of α-helical antimicrobial peptides to target gram-negative pathogens, *Acinetobacter baumannii* and *Pseudomonas aeruginosa*: utilization of charge, 'specificity determinants,' total hydrophobicity, hydrophobe type and location as design parameters to improve the therapeutic ratio. *Chem Biol Drug Des.* Vol.77, pp. 225-240

Kite, J. & Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. *J Mol Biol.* Vol.157, pp. 105-132

Lacroix, E.; Viguera A.R. & Serrano, L. (1998). Elucidating the folding problem of α-helices: Local motifs, long-range electrostatics, ionic strength dependence and prediction of NMR parameters. *J Mol Biol.* Vol.284, pp. 173-191

Landon, C.; Barbault, F.; Legrain, M.; Guenneugues, M. & Vovelle, F. (2008). Rational design of peptides active against the gram positive bacteria *Staphylococcus aureus*. *Proteins*. Vol.72, pp. 229-239

Lata, S.; Mishra, N.K. & Raghava, G.P.S. (2010). AntiBP2: improved version of antibacterial petides prediction. *BMC Bioinformatics.* Vol.11

Lata, S.; Sharma, B.K. & Raghava, G.P.S. (2007). Analysis and prediction of antibacterial peptides. *BMC Bioinformatics.* Vol.8, No.236

Loose, C.; Jensen, K.; Rigoutsos, I. & Stephanopoulos, G. (2006). A linguistic model for the rational design of antimicrobial peptides. *Nature.* Vol.443, No.1, pp. 867-869

Mandal, S.M.; Dey, S.; Mandal, M.; Maria-Neto, S. & Franco, O.L. (2009). Identification and structural insights of three novel antimicrobial peptides isolated from green coconut water. *Peptides*. Vol.30, No.4, pp. 633-637

Nagarajan, V.; Kaushik, N.; Murali, B.; Zhang, C.; Lakhera, S.; Elasri, M.O. & Deng, Y. (2006). A Fourier Transformation based Method to Mine Peptide Space for Antimicrobial Activity. *BMC Bioinformatics*. 7(Suppl 2):S2.

Nijnik, A. & Hancock, R.E.W. (2009). Host defence peptides: antimicrobial and immunomodulatory activity and potential applications for tackling antibiotic-resistant infections. *Emerging Health Threats Journal*. Vol.2, No.e1, pp. 1-7

Pfeifer, Y.; Cullik, A. & Witte, W. (2010). Resistance to cephalosporins and carbapenems in Gram-negative bacterial pathogens. *Int J Med Microbiol*. Vol.300, No.6, pp. 371-379

Poirel, L.; Naas, T. & Nordmann, P. (2010). Diversity, epidemiology, and genetics of class D beta-lactamases. *Antimicrob Agents Chemother.* Vol.54, No. 1, pp. 24-38

Porto, W.F.; Fernandes, F.C. & Franco, O.L. (2010). An SVM model based on physicochemical properties to predict antimicrobial activity from protein sequences with cysteine knot motifs. *LNCS*. Vol.6268, pp. 59-62

Rigoutsos, L. & Floratos, A. (1998). Combinatorial pattern discovery in biological sequences: The TEIRESIAS algorithm. *Bioinformatics.* Vol.14, pp. 55-67

Sigurdardottir, T.; Andersson, P.; Davoudi, M.; Malmsten, M.; Schmidtchen, A. & Bodelsson, M. (2006). *In silico* identification and biological evaluation of antimicrobial peptides based on human cathelicidin LL-37. *Antimicrob Agents Chemother.* Vol.50, No.9, pp. 2983-2989

Silva, O.N.; Mulder, K.C.; Barbosa, A.A.; Otero-Gonzalez, A.J.; López-Abarrategui, C.; Dias, S.C.; Rezende, T.M. & Franco, O.L. (2011). Exploring the pharmacological potential

of promiscuous host-defense peptides: from natural screenings to biotechnological applications. *Front Microbiol*. Vol.2

Thennarasu, S. & Nagaraj, R. (1996). Specific antimicrobial and hemolytic activities of 18-residue peptides derived from the amino terminal region of the toxin pardaxin. *Protein Eng*. Vol.9, No.12, pp. 1219-1224

Thomas, S.; Karnik, S.; Barai, R.S.; Jayaraman, V.K. & Idicula-Thomas, S. (2010). CAMP: a useful resource for research on antimicrobial peptides. *Nucl Acid Res.* Vol.38, Database issue.

Torrent, M.; Andreu, D.; Nogués, V.M. & Boix, E. (2011). Connecting peptide physicochemical and antimicrobial properties by a rational prediction model. *PLoS ONE*. Vol.6, No.2

Tossi, A.; Tarantino, C. & Romeo, D. (1997). Design of synthetic antimicrobial peptides based on sequence analogy and amphipathicity. *Eur J Biochem*. Vol.250, pp. 549-558

Ueno, S.; Minaba, M.; Nishiuchi, Y.; Taichi, M.; Tamada, Y.; Yamazaki, T. & Kato, Y. (2011). Generation of novel cationic antimicrobial peptides from natural non-antimicrobial sequences by acid-amide substitution. *Ann Clin Microbiol Antimicrob.* Vol.10, No.11

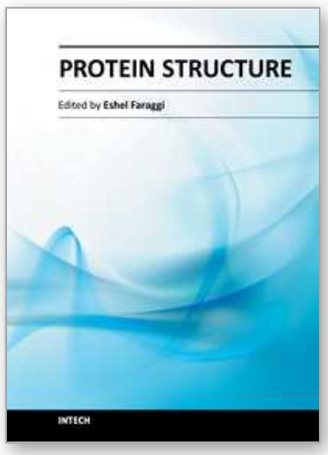
Wang, Z. & Wang, G. (2004). APD: the Antimicrobial Peptide Database. *Nucl Acid Res*. Vol.32, No.1, pp. D590-D592

Wieczorek, M.; Jenssen, H.; Kindrachuk, J.; Scott, W.R.P.; Elliott, M.; Hilpert, K.; Cheng, J.T.J.; Hancock, R.E.W. & Straus, S.K. (2010). Structural studies of a peptide with immune modulating and direct antimicrobial activity. *Chemistry & Biology*. Vol.17, No.1, pp. 970-980

Wu, M. & Hancock, R.E.W. (1999). Improved Derivatives of Bactenecin, a Cyclic Dodecameric Antimicrobial Cationic Peptide. *Antimicrob Agents Chemother.* Vol.43, No.5, pp. 1274-1276

Yount, N.Y. & Yeaman, M.R. (2004). Multidimensional signatures in antimicrobial peptides. *PNAS*. Vol.101, No.1, pp. 7363-7368

Zhang, X. (2010). Human in check: new threat from superbugs equipped with NDM-1. *Protein Cell*. Vol.1, No.12, pp.1051-1052.



Protein Structure
Edited by Dr. Eshel Faraggi

ISBN 978-953-51-0555-8
Hard cover, 396 pages
Publisher InTech
Published online 20, April, 2012
Published in print edition April, 2012

Since the dawn of recorded history, and probably even before, men and women have been grasping at the mechanisms by which they themselves exist. Only relatively recently, did this grasp yield anything of substance, and only within the last several decades did the proteins play a pivotal role in this existence. In this expose on the topic of protein structure some of the current issues in this scientific field are discussed. The aim is that a non-expert can gain some appreciation for the intricacies involved, and in the current state of affairs. The expert meanwhile, we hope, can gain a deeper understanding of the topic.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:

William F. Porto, Osmar N. Silva and Octávio L. Franco (2012). Prediction and Rational Design of Antimicrobial Peptides, Protein Structure, Dr. Eshel Faraggi (Ed.), ISBN: 978-953-51-0555-8, InTech, Available from: http://www.intechopen.com/books/protein-structure/prediction-and-rational-design-of-antimicrobial-peptides

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the Creative Commons Attribution 3.0 License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen