

## Rational development of antimicrobial peptides for therapeutic use: design and production of highly active compounds

G. Maccari<sup>1</sup>, R. Nifosi<sup>2</sup> and M. Di Luca<sup>2\*</sup>

<sup>1</sup> Center for Nanotechnology Innovation @NEST, Istituto Italiano di Tecnologia, Piazza San Silvestro 12-56127 Pisa, Italy

<sup>2</sup> NEST, Scuola Normale Superiore and Istituto Nanoscienze-CNR, Piazza San Silvestro 12-56127 Pisa, Italy

\* corresponding author: mariagrazia.diluca@sns.it

The emergence of drug-resistant pathogenic microbial strains has created an urgent need for new anti-infective molecule development. In order to avoid the spread of bacterial resistance, there is pressing demand to design a novel class of antibiotics having different mechanism of action in comparison to existing drugs. Natural antimicrobial peptides (AMPs) represent a novel class of molecules with a broad spectrum of activity and a low rate in inducing bacterial resistance. However, until now, many AMPs have failed in clinical trials because of several drawbacks that strongly limit their applicability such as degradation, cytotoxicity and high production cost. Thus, to overcome the limitations of native peptides, a rational *in-silico* approach to AMPs design becomes a promising strategy that drastically reduce production costs and the time required for evaluation of activity and toxicity. This chapter will focus on the strategies and methods for *de-novo* design of potentially active AMPs.

**Keywords** AMPs; *de-novo* peptide design; antimicrobial activity; QSAR; Machine Learning; Molecular Dynamics

### 1. Introduction

In recent years, the increasing and rapid spread of pathogenic microorganisms resistant to conventional antibiotics represents a major global health problem associated with bacterial infections both in hospital settings and in the community. Despite successes in antimicrobial drug development, resistance to every new antibiotic molecule has appeared in bacterial populations within a few years of its introduction determining a reduction in the therapeutic efficacy [1]. The decline in effectiveness of current therapies spurs research for the identification of novel molecules endowed with antimicrobial activities and new mechanisms of action. Currently, antimicrobial peptides (AMPs) received an increasing attention as potential therapeutic agents, since they represent a novel class of molecules with a wide spectrum of activity and a low rate in inducing bacterial resistance. Over the past decades a large number of naturally occurring AMPs have been identified or predicted from various organisms as effector molecules of the innate immune system playing a crucial role in the first line of defence [2]. Their mechanism of action mainly consists in membrane permeabilization either by forming discrete pores or by disrupting the membrane bilayer resulting in cell death. However the inhibition of nucleic acid, protein, and cell-wall component synthesis and some of the essential microbial enzymatic activities, were also described [2]. In addition, recent studies have shown the ability of some AMPs to act against microbial biofilms, in particular during early phases of biofilm development [3].

In this regard, AMPs are considered new candidate molecules particularly promising for a use in antimicrobial therapy. Although some AMPs are already in clinical and commercial use [4], the future design and optimization of novel AMPs will need to improve their antimicrobial activity, minimize the cytotoxicity and reduce the proteolytic degradation or the biological fluid inhibition [5].

Traditional design and optimization studies of peptides are known to be expensive and time-consuming. A rational *in-silico* approach to AMPs design can drastically reduce production costs and the time required for evaluation of activity and toxicity. Different types of computer-assisted design strategies have been developed to address the difficult problem of relating primary sequence to peptide structure and activity. Alignment-based approaches are intuitive and give good results with long sequences, but lack in understanding of the mechanism of action. Machine learning algorithms, combined with stochastic optimization methods are adopted to relate antimicrobial activity to some structural characteristic in order to screen large virtual libraries [4]. Finally, biophysical studies like molecular dynamic (MD) simulations are the most accurate approach to a rational design of antimicrobial molecules, providing a working hypothesis of the mechanism of action [4].

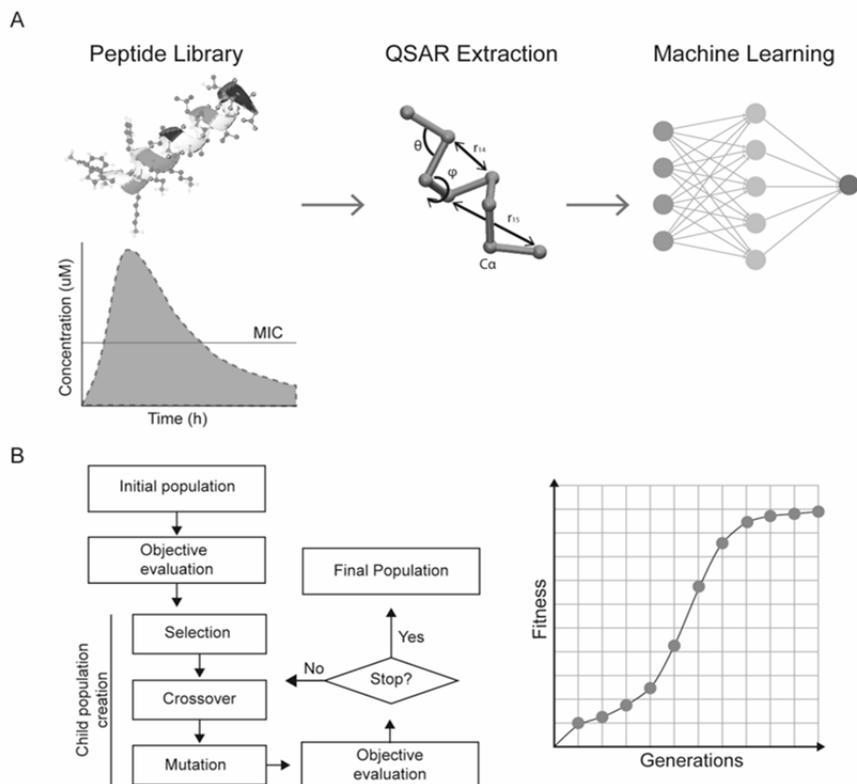
This chapter describes the strategies and methods for *de-novo* design, production and purification of potentially active AMPs and discusses advantages and disadvantages of AMPs for therapeutic application.

### 2. Virtual screening and optimization of antimicrobial peptides by statistical analysis

Recent research on AMPs has focused on methods to search through the constellation of known or predicted peptide sequences – either empirically or computationally – for peptides with desired properties and these approaches are continually evolving. *In silico* peptide design and prediction is based basically on two different types of approaches. In the first one, the primary sequence information is associated with a measure of its activity – quantitative or qualitative –

through a series of sample sequences. A statistical model is then constructed by machine learning and/or lexical methods. In contrast, in the second branch of approaches, biophysical studies are applied in order to evaluate peptide's interaction and folding within the membrane. While the first method is optimal for virtually screening a library of peptides, the second allows to study more extensively the mechanism of action of AMPs. Statistical methods for AMPs screening can be distinguished from linguistic-based models, in which natural peptide sequences are treated like words and grammar rules are applied to identify text patterns, and quantitative structure activity relationships (QSAR) models, where chemical structure is correlated to measurement of biological activity.

Common statistical-based design strategies involve a series of fundamental points (**Figure 1A**): 1) Preparation of a dataset representing active and inactive molecules; 2) Building of a mathematical model, able to distinguish peptide activity in a qualitative or quantitative fashion; 3) Application of the model to design novel active molecules (**Figure 1B**). In this paragraph, the process of dataset construction, model preparation and validation will be outlined. Some practical applications of statistical method to AMP design will be illustrated.



**Fig. 1** Schematic process representation. A) Statistical-based design strategy. A dataset library is created and screened for antimicrobial activity; a series of QSAR descriptor is selected to create a regression model with a machine learning method. B) Design of new bioactive molecules. A schematization of the genetic algorithm for selecting new active molecules. As the number of generation increases, the fitness function tents to a plateau.

## 2.1. Dataset preparation

In statistical analysis, dataset preparation is probably the most delicate part in the process of model construction. During this phase, a list of peptides is collected in an ordered database and a specific activity is associated with primary and/or secondary sequence information. AMPs feature a remarkable variety of structural motifs, resulting in a huge variety of the primary sequence, from the amino acidic composition to the total length. Because of this variety, a rich and complete dataset of active and inactive peptides is difficult to obtain without introducing biases. For these reasons, during the years different bioinformatics methods were applied in order to collect as much as possible information of natural and synthetic AMPs from literature. Although the process of information gathering can be automated (for example by iterative scanning of public sources [6]), because of the difficulty and sensitivity of the information crawling process, manually attended datasets are more appreciated.

**Table 1** Chronological list of AMPs databases.

| Year | Database  | Web site  | Content               |
|------|-----------|---|-----------------------|
| 2002 | AMSDb     | <a href="http://www.bbcm.univ.trieste.it/~tossi/pag1.htm">http://www.bbcm.univ.trieste.it/~tossi/pag1.htm</a> | Plant and Animal AMPs |
| 2007 | AMPer     | <a href="http://marray.cmdr.ubc.ca/cgi-bin/amp.pl">http://marray.cmdr.ubc.ca/cgi-bin/amp.pl</a>               | Plant and Animal AMPs |
| 2007 | BACTIBASE | <a href="http://bactibase.pfba-lab-tun.org/main.php">http://bactibase.pfba-lab-tun.org/main.php</a>           | Bacteriocins          |
| 2008 | RAPD      | <a href="http://faculty.ist.unomaha.edu/chen/rapd/">http://faculty.ist.unomaha.edu/chen/rapd/</a>             | Recombinant AMPs      |
| 2009 | PhytAMP   | <a href="http://phytamp.pfba-lab-tun.org/main.php">http://phytamp.pfba-lab-tun.org/main.php</a>               | Plant AMPs            |
| 2009 | APD2      | <a href="http://aps.unmc.edu/AP/main.php">http://aps.unmc.edu/AP/main.php</a>                                 | Natural AMPs          |
| 2010 | CAMP      | <a href="http://www.bicnirrh.res.in/antimicrobial/">http://www.bicnirrh.res.in/antimicrobial/</a>             | All AMPs              |
| 2012 | DAMPD     | <a href="http://apps.sanbi.ac.za/dampd/">http://apps.sanbi.ac.za/dampd/</a>                                   | All AMPs              |
| 2012 | YADAMP    | <a href="http://yadamp.unisa.it/">http://yadamp.unisa.it/</a>   | All AMPs              |

The advent of antimicrobial peptide databases has made it easier for scientists to understand and rationally design AMPs. Table 1 summarizes the major databases of AMPs, classified into three categories: 1) general databases from AMPs, 2) specialized databases for natural AMPs, and 3) specialized databases for non-natural AMPs. The AMSDb is one of the first attempts to manually organize AMP data in an exhaustive manner, covering the sequences of gene-encoded antimicrobial peptides and proteins from animal and plant sequences. Other public databases were published over the years, providing a set of tools for the automatic analysis and prediction of active AMP sequences. In 2009 an updated version of antimicrobial peptide database (APD2) [7] was released. The server, in addition to presenting an extensive and updated list of peptides, provides information on amino acidic frequency, the presence of conserved motifs and additional chemophysical information. Another useful source of AMP sequences and information is the manually curated collection of antimicrobial peptides (CAMP) [8] database, that includes additional tools like a local BLAST query system and a prediction tool based on amino acidic frequencies and Random Forests (RF) algorithm. More recently, two manually curated datasets were developed with more information about biological activity against difference strains [9] and chemophysical and taxonomical characteristics [10]. Another approach for the preparation of AMPs dataset is the screening of random peptide libraries. This method has the advantage to reduce the bias of active sequence selection and to give precise and uniform quantitative information of peptide activity [11]. Furthermore, complex prediction models often require a large number of measured values of AMP activity to fit the correspondingly large set of parameters. To this end, solid-phase synthesis and high-throughput screening of large peptide arrays has become a common practice in drug discovery. However, because of the huge number of amino acidic combinations an exhaustive screening of random libraries is unfeasible. For example, a full combinatorial assay of peptides with length up to 10 residues would result in  $20^{10}$  different sequences, an unfeasible number of combinations. For this reason, systematic studies tend to limit the number of peptides by analyzing a fixed number of amino acids positions with a precise combination of substitution [12]. On the basis of the analysis of natural AMPs, the amino acidic space is limited to charged residues and moderately hydrophobic sequences; to avoid technical problems during the synthesis phase, cysteine and methionine residues are excluded, owing to potential cross linking or oxidation. In this way the number of combinations is extremely reduced, at the cost of some bias introduction, since a large number of substitution is excluded a priori.

## 2.2. Peptide representation

Possibly the most intuitive attempt to model AMPs based on natural peptide sequence is a linguistic model in which sequences in one letter amino acid code are considered as 'text' and formal grammar rules are applied to identify text patterns in naturally occurring peptides. Different studies were performed on the basis of this method. Generally speaking, a dataset of known antimicrobial peptides is used as a template for the extrapolation of some grammar rules [13], like amino acid frequency or particular motives occurrence [4]. These studies are generally based on the systematic substitution of a series of amino acids to identify residues and positions that are important for activity. In most cases their activity was investigated by introducing variations on the basis of general concepts, like charge and amphiphilicity, crucial to antimicrobial peptide activity. CM, a chimerical alpha-helical antimicrobial peptide obtained from the fusion of cecropin A and mellitin - two natural AMPs firstly isolated from insects [14] - is an example of AMP derived from template-based studies. Even though these studies have shed light on the importance of specific amino acids and residue positions to peptide activity, such local approaches fail to account for amino acidic interactions. Various experiments have demonstrated the importance of secondary structure in antimicrobial activity; therefore, in order to introduce secondary structure information, different strategies have been adopted. Sequence alignments or position-specific scoring matrices (PSSM) are among the most successful attempts to integrate sequence order information to linguistic analysis of AMPs [6,8,15]. However, these approaches are limited to natural amino acids, since there is not enough sequence information of non-natural amino acidic substitutions to build an exhaustive statistical model. A rational analysis of AMP based on template structures has been successfully applied to design AMPs with a specific structure and elevated activity against gram+ and gram-. This design concept involves the use of positively

charged or hydrophobic amino acidic residues accurately placed in specific positions of a template sequence [16]. In this study, a hybrid approach was used, by the rational insertion of non-natural residues on the basis of chemophysical characteristics. In contrast to template based methods and computational simulations (like molecular dynamics), QSAR models apply numerical analyses to describe the relationship between chemophysical characteristics and biological activity. These chemophysical characteristics, named *descriptors*, are usually derived from peptide's primary sequence and allow an insight into the physical characteristics regulating the biological activity. QSAR descriptors can be empirically derived from experimental measures such as molecular weight, partition coefficient or HPLC retention time, but also theoretically calculated. Calculated descriptors can be related to peptide's primary structure or chemical composition, as well as 2D or 3D structure. Moreover, single descriptors can be combined to describe different – but related – chemophysical characteristics, like polarity and hydrophobicity, to reduce variable hyperspace in a coarse fashion. In a pioneering study, Hellberg et al. developed highly condensed variables derived from a principal component analysis (PCA) of several experimental or theoretical physicochemical properties for the 20 naturally occurring amino acids [11]. QSAR analysis of peptides using these types of descriptors has proven effective in predicting AMP activity as well as different physiological functions [17][18]. Although QSAR has become an integral part of screening programs in pharmaceutical drug-discovery pipelines of small compounds, its application to the discovery of AMPs is relatively recent. In AMP design and prediction, it is of relevance to understand how modifications in peptide sequences may lead to better performing drugs.

In order to account for sequence information in QSAR analysis, different strategies have been adopted. From one side, different classes of QSAR descriptors have been developed, generally named 3D QSAR or inductive QSAR [19], where intra and inter molecular interactions are considered. Another approach is the representation of structural information by theoretical calculations, molecular dynamics or experimental studies. A measure of sequence information can be taken into account by analyzing the correlation between QSAR variables along the primary sequence. Auto and Cross-covariance (ACC) analysis, a measure originally introduced by Wold [20] is a set of discrete numbers that take into account some sequence order effects by measuring auto covariance (AC) within the same descriptor and cross covariance (CC) between two different descriptors. For a given protein sequence, ACC variables describe the average interactions between residues distributed a certain *lag* apart throughout the whole sequence. Higher lag values result in describing distant interactions along the peptide sequence.

$$AC_{(i),d} = \sum_{k=1}^{L-d} (Z_i^k * Z_i^{k+d}) / L; CC_{(i,i+1),d} = \sum_{k=1}^{L-d} (Z_i^k * Z_{i+1}^{k+d}) / L;$$

**Equation 1** Classical Auto and Cross-covariance equations

Besides describing the sequence order, ACC has the ability to transform each amino acid sequence of variable length into uniform equal-length vectors. This feature is very important in data mining methods, where a fixed length vector describing each instance is required. Furthermore, this particular feature is often used to face the 'curse of dimensionality' problem. In Life Science, datasets typically have many more variables than samples, representing a problem for peptide representation as well as model construction. There are two different approaches to deal with this problem. The first is to shrink the number of variables into a smaller one with PCA, ACC or other techniques. The second one is to choose a classification algorithm designed for this scope, as we will discuss below. During the years, different versions of ACC were introduced, but the concept remains that different chemophysical descriptors are correlated between each other in a given order along the primary sequence, and this correlation describes the antimicrobial activity.

In a recent paper, a combination of the linguistic and the QSAR approach has been proposed [21] by analyzing in a two-steps process the primary sequence of candidate AMPs. A template-based model was used for the first step, analyzing sequences on the basis of BLAST alignment algorithm. As the sequence alignment method could not deal with all the AMP sequences, a feature selection method, based on QSAR descriptors, was applied. The combination of these two methods gave some improvement in the overall prediction results.

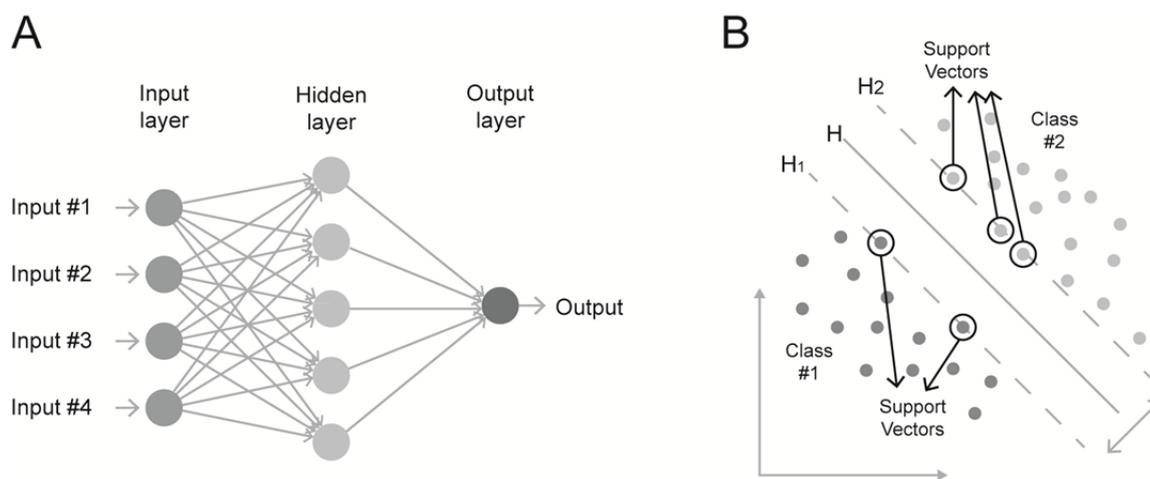
**Table 2** Brief selection of computer-based peptide design approaches

| Year | Description  | Method                  | Reference |
|------|--|-------------------------|-----------|
| 2006 | Grammatical rules are derived from natural AMPs in order to design new peptides. | Template-based methods  | [13]      |
| 2008 | Random AMP generation with template-based methods                                | Hidden Markov Models    | [15]      |
| 2010 | Dataset of AMP with different tools and prediction methods                       | SVM, RF, AA frequency   | [8]       |
| 2011 | GA Optimization of AMP based on a library of biased peptides                     | ANN, genetic algorithms | [22]      |

|      |   |   |      |
|------|---|---|------|
| 2011 | Prediction of AMP based on a combination of sequence alignment and feature selection.                             | Sequence alignment, k-mean clustering.      | [21] |
| 2013 | Prediction and optimization of AMP activity by multi-object optimization by feature selection of QSAR descriptors | Random Forest, Multi-objective optimization | [23] |

### 2.3. Model construction

In order to develop mathematical models to correlate sequence and structure to biological activity, many statistical learning methods and multivariate approach can be applied. There are two main categories of prediction to answer two different questions: regression models (for predicting the activity of a peptide as a continuous variable such as MIC or a surrogate such as in the luminescence assay) or classification where the model is trained to classify an unknown peptide (instance) as simply active or inactive, giving as result a confidence value. Different regression models can be applied depending on the type of data and the choice of descriptors. Linear methods primarily have been used in determination of QSAR for AMPs because of the simple calculation and interpretation. The most used linear methods are PCA, partial least squares projections to latent structures (PLS) and multivariate design. PCA is a mathematical procedure which is able to transform a number of possibly correlated variables into a smaller number of uncorrelated variables called principal components. It is a way to explore the relationships among AMP analogues and among variables. Then the relation between the biological activity and these principal components associated for each AMP can be handled using a PLS regression method. PLS model tries to find the multidimensional direction in the X space that explains the maximum multidimensional variance direction in the Y space. The main advantage of these techniques is they can be used where multiple regression fails: when there are fewer peptides than variables in the model. Choices of a prediction technique also involve tradeoffs between model accuracy and meaningfulness. Nonlinear techniques, like support vector machines (SVM) and artificial neural networks (ANN) are considered to give better results when the correlation between QSAR descriptors and biological activity is not completely clear. ANN is a mathematical model based on the simulation of some properties of biological neural networks. A network of descriptors is defined as input nodes or neurons. These nodes are connected together, forming a network that interacts in a hidden layer and sums up into an output node (Figure 2A). SVM is a classification method, where two classes are separated by a hyperplane (H) (Figure 2B). The optimized plane is defined as the one that maximizes the margin between the two classes which is measured by the distances between plane H and the planes cutting the nearest sample points on both sides of H, namely, H<sub>1</sub> and H<sub>2</sub>. In particular, the sample points located exactly on planes H<sub>1</sub> and H<sub>2</sub> are defined as support vectors. For the purpose of classification, the nonlinear techniques of support vector machines and ANNs are considered to give superior results, but at the cost of introducing rather opaque models that cannot easily be used to shed light on the underlying mechanisms involved.



**Fig. 2** Machine learning methods. A) In Artificial Neural networks, a series of input descriptors are encoded into a hidden layer that converges to an output result. B) Illustration of the principles of SVM: SVM separates classes along their maximal margin of separation. Support vectors are selected to set the maximal margin between class clusters

Decision trees are another method to classify an unknown instance in different classes. Each node in the tree represents a particular attribute to test. In order to classify an unknown instance, it is routed down the tree according to the values of the attributes tested in successive nodes, and when a leaf is reached, the instance is classified according to the class assigned to the leaf. Different decision trees algorithms were developed; in particular Random Forest (RF) has become very popular for pattern recognition and classification, mainly because RF provides two aspects that are very important for data mining: high prediction accuracy and information on variable importance for classification [24]. RF

is an ensemble recursive partitioning method where many decision trees are trained using subsets of samples and descriptors with replacement. RFs have been widely used in AMP prediction and optimization, with performances that compare well to other classification algorithms such as SVM and ANN [8]. The advantage is a better control on variable interactions and correlations, which are ubiquitous in data sets generated in the Life Sciences.

#### 2.4. Design and optimization of active AMPs

In common computer-based design, a sophisticated activity estimator is associated with a technique that enables stochastic optimization to iteratively optimize peptide activity. In fact, due to the huge number of possible combinations, stochastic optimization methods such as Genetic algorithms (GA) may be a preferred tool to perform directed random searches in large problem spaces [25]. GAs are evolutionary algorithms that mimics Nature's adaptive approach to the environment, in which the process of evolution is performed through successive generation of mutation, deletion and so on. Although some implementations of evolutionary algorithms suffer from certain well-known computational inefficiencies (for example, a dependence of parameter initialization, partially insufficient sampling and premature convergence), they have proven high applicability in many practical studies [26]. In evolutionary algorithms, each potential AMP candidate is treated like an entity belonging to a population. A fitness function is used as a measure of its biological activity. This value can reflect its predicted activity from a regression or a classification model. In the second case, it will express the confidence assigned to the prediction result. To find the best solution in the fitness landscape, the population of AMP candidates must be able to adapt to the ruggedness of the landscape, avoiding a premature convergence in areas of low fitness. Generally speaking, at the beginning of the evolutionary process, the population represents a set of random candidates. During each generation, peptides are selected according to their fitness and new peptides are generated from the combination and mutation of the selected ones. As the simulation goes on, the population tends to present an increasing mean fitness value, until convergence. An example of genetic algorithms applied to AMP selection and optimization is [22], where a training library of 1400 peptides was used with a regression model in order to discriminate highly active AMPs. Initially, a training library biased on amino acidic composition (rich in lysine, arginine, tryptophan and hydrophobic residues) was synthesized and tested for antimicrobial activity against *P.aeruginosa*. A neural network was trained on the basis of screening results and used to screen an *in silico* library of 100000 peptides. A set of candidates was screened, synthesized and experimentally validated for their antimicrobial activity. These peptides demonstrate high activity against reference strains as well as clinical pathogens. Interestingly, similar peptides with overall chemophysical characteristics showed deeply different activity, demonstrating the importance of amino acidic arrangement along peptide's primary sequence.

When simultaneous optimization of two or more conflicting objectives is required, a class of GA called multi-objective evolutionary algorithms (MOEA) can be used to provide an optimal solution. In fact, MOEAs have the peculiarity to optimize different objectives separately, maintaining a set of good trade-off solutions ranked by their dominance at a given instant of the evolution process [27]. The advantage of this technique is that final peptides can be screened on the basis of their biological activity as well as other chemophysical characteristics, without favouring one single objective in particular. In our paper [23] we dealt with the AMP discovery task by using a combination of MOEA, machine learning algorithm and QSAR analysis. The resulting tool was used to screen two *ab-initio* alpha helical AMPs and one optimized version of CM18. Furthermore, the type of QSAR descriptor chosen in this work allowed to account for non-natural amino acidic insertion. An *ab-initio* AMP with two norleucine residues was designed and tested against *S. aureus* and *P. aeruginosa* strains together with all the designed peptides. In addition, MD simulations were performed to confirm the peptide's secondary structure.

While different valid tools exist for the design of peptidomimetic endowed with bactericidal activity, it has been proven to be tricky to improve other features such as *in vivo* stability and toxicity in view of their clinical applications. In an attempt to improve antimicrobial activity of a selected peptide, general optimization strategies may be followed including cyclization [28], increase of positive charge or hydrophobicity [29] or improving segregation of polar and hydrophobic residue starting from a sequence template [16]. Recent studies have investigated the effect of different basic residues on peptide antibacterial activity and cell selectivity [30]. The primary amine of Lys interacts less electrostatically with zwitterionic phospholipids than the guanidinium group of Arg, suggesting that Arg-to-Lys substitution would increase selectivity against negatively charged phospholipids of bacterial membranes and cause the peptide to be less lytic to erythrocytes.

All the findings suggest that the activity of AMPs is determined by a subtle combination of factors such as sequence, hydrophobicity and position of cationic residue. At the same time, a large number of studies have attempted to establish more precise quantitative structure-activity relationship (QSAR) between the primary sequence of AMPs and their activity toward selected microorganisms. Unfortunately, a general statement is almost impossible to obtain due to the complexity of the target and the mechanism of action lacking a defined target.

### 3. AMP design by MD simulations

A rational design of AMPs relies on detailed understanding of their mechanisms of action. The ability to interact with the pathogen membrane is pivotal for the efficacy of AMPs, either because the antipathogenic activity entails destabilization of the cell membrane or because, in the case of intracellular killing mechanisms, the AMP molecules need to be internalized in order to perform their activity. Experimental studies aimed at unravelling peptide-membrane interactions rely on fluorescence and vibrational spectroscopy, circular dichroism, and NMR [31]. These techniques provide essential pieces of information such as peptide position and orientation in the bilayer, peptide secondary structure and peptide aggregation state. However, a complete picture may fail to emerge due to the lack of crucial details, in particular because the peptide-membrane interactions often involve the formation of transient structural arrangements. In this regard, molecular modelling techniques, mainly MD simulations, may be able to provide the necessary bridges to achieve a complete and high-resolution understanding of membrane-peptide interaction processes. Due to the inherent difficulties in these computational methods, regarding sampling and force field accuracy, such methods must be complemented by the above-mentioned experimental observations, in order to be validated. The synergy between these different approaches successfully provides the researchers with accurate models of AMP membrane destabilization mechanisms.

MD simulations are based on the numerical solution of Newton's equation to describe the dynamics of atoms interacting via molecular mechanics force fields. These force fields describe the inter- and intra-molecular interactions as a sum of different contributions. They are typically separated in a covalent part describing bond, angle, and dihedral interactions, and a non-covalent part taking into account electrostatic and Van der Waals forces. For peptide-membrane systems the number of atoms can be quite large (typically larger than 50000) so the simplified biomolecular force fields such as CHARMM, GROMOS and AMBER are employed, rather than more sophisticated molecular force fields containing, for example, electrostatic polarization terms, which are computationally more expensive. These methods are routinely applied to systems of 50000-100000 atoms, including the peptide in one or more copies, a patch of lipid bilayer (with no more than few hundreds of lipids) and the surrounding water molecules, using periodic boundary conditions to avoid border effects and because the electrostatic part is treated in a more computationally convenient fashion. Simulated times are of hundreds of nanoseconds. In these time scales some interesting peptide-membrane association events are observed, though there is no guarantee of an exhaustive exploration of the configurational space, and artefacts may arise because of limited sampling. The starting configuration should be carefully chosen, because a completely random arrangement of the various molecular components is unlikely to lead to the wanted event in the limited simulated time scales. So, for example, peptide aggregation inside the membrane may be simulated starting from the needed number of peptide copies already embedded in the membrane.

An additional step in the simplification of the system description is provided by coarse-grained methods [32]. While the above-mentioned force fields explicitly describe all atoms (in certain cases, like GROMOS, neglecting some of the hydrogen atoms in the so-called unified atom approach) coarse-grained (CG) force fields consider appropriately chosen groups of atoms as single interacting centers. The critical point is how to describe the interactions between these groups. The CG approach enables systems to be simulated for longer times (two/three orders of magnitude), the downside being a loss in accuracy. The all-atoms and CG approach can be combined to afford the advantages of both classes of methods (see below). Typical CG simulations target membrane patches with thousands of lipids and several peptide molecules, running from the micro to the millisecond timescale.

Critical aspects of MD simulations applied to lipid-membrane interactions, beyond the already mentioned force field accuracy and configurational sampling, regard the choice of the membrane model. With respect to the cellular membrane, the simulated systems contain only few lipid components, so they are closer to experiments involving artificial bilayers, with controlled lipid composition. However, since generally the membrane-disrupting peptides are also shown to destabilize the bilayer in artificial vesicles, simple membrane models may for the most part be appropriate. The size of the simulated bilayer patch may also be a critical parameter. Some peptide may act by selectively modifying the surface tension of the outer leaflet of the bilayer, and cause curvature-induced destabilization [33]. Such effects may be masked by the use of periodic boundary conditions if the bilayer patch is too small.

#### 3.1. MD studies of mechanism of action

In the following, we shall briefly summarize some of the computational studies with no intent of being exhaustive (for other reviews on the subject the interested reader is referred to [34–39]), but focusing on those providing insight in peptide/membrane interaction mechanisms and in the design of novel AMPs.

**Table 3** Model AMP mode of action. Brief summary of some MD studies of the mechanism of action model AMPs.

| Peptide          | Structure  | Mechanism of action                              | References |
|------------------|--|--|------------|
| Magainin-2 (MAG) | 23 AA peptide, alpha helical amphipathic                           | Membrane perturbation (toroidal pore)            | [40–44]    |
| Protegrin        | 16-18 AA peptide, beta-hairpin configuration with disulfide bonds. | Membrane perturbation (toroidal pore)            | [45,46]    |
| Alamethicin      | 20 AA peptide, alpha helical configuration                         | Membrane perturbation (barrel-stave)             | [39,47–49] |
| Melittin         | 26 AA peptide, alpha helical configuration                         | Membrane perturbation (disordered toroidal pore) | [50,51]    |
| Cateslytin       | 15 AA peptide with beta-sheet configuration                        | Membrane perturbation (disordered toroidal pore) | [52]       |
| Indolicidin      | 13 AA peptide rich in Trp  | Membrane perturbation (presumably barrel-stave)  | [53]       |
| Buforin II       | 21 AA peptide with alpha-helical configuration                     | Non-lytic mechanism (disordered toroidal pore)   | [54]       |
| Ovispirin        | 18 AA peptide, Alpha-helix configuration                           | Membrane perturbation (carpet-mode)              | [55][56]   |

Magainin-2 (MAG) is one of the most well characterized AMP, also from the molecular modelling point of view. MAG is 23-amino acid long, cationic, and amphipathic peptide with antibacterial, antifungal, antiviral, and anticancer activity. Its mechanism of action involves bilayer perturbation, as confirmed by how magainin induces leakage of lipid vesicles and how it reduces the bending rigidity of GUVs even at very low peptide concentration [40]. The commonly accepted mechanism of magainin activity involves the formation of toroidal pores. In this pore model, the peptides interact with the lipid head groups forming the walls of the pore.

MD simulations showed how a single magainin interacts with a POPC bilayer patch, causing a local thinning effect [57], mediated by the contacts between lysine residues and lipid polar head. Marrink and co-workers [58] observed the spontaneous formation of a pore-like defect in DPPC membrane in MD simulation featuring multiple magainin-H2 (an analogue of magainin-2, with a higher hydrophobic moment) peptides at the water/bilayer interface. They showed that the formation of the “pore” is concentration dependent, as observed in the experiments, and requires peptide aggregation. As confirmed by similar studies on other AMPs, these simulations point to a revise of the “toroidal pore” model. In particular the pore is rather disordered, both in the arrangement of the various peptides (only one peptide spans the center of the pore, the others being diffusely distributed on the rim of the ~2-nm wide pore with a parallel orientation) and in the peptide internal structure. The latter is found to be rather flexible, so that helical structure in the membrane does not seem to be a strict requirement.

Similar disordered toroidal pores, with only one or two peptides perpendicularly spanning the bilayer, were observed in MD simulations of Melittin [59], and of the  $\beta$ -sheet AMPs Cateslytin with a zwitterionic membrane [60]. Disordered transient toroidal pores were also formed in MD simulations of Buforin II [61], a non-lytic AMP derived from histone H2A. Santo et al. [51] MD study aimed at understanding the different behaviour of Melittin and Magainin-2 detected in experiments with GUV [40], reporting that, while mellitin induces graded dye release, Magainin-2 induces all-or-none release. From coarse-grained MD simulations this different behaviour is traced back to different aggregation propensities of the two peptides in the membrane, with magainin 2 aggregating in larger numbers in the pore region, thus creating larger holes. By contrast, Melittin does not aggregate in large numbers, resulting in smaller pores in the bilayer. Using the MARTINI coarse-grain force field, Woo et al [62] simulated a larger system comprising up to 1600 Magainin-2 peptides and a mixed DPPC/POPG patch of lateral dimension up to 0.1  $\mu\text{m}$ . Their findings confirm the formation of disordered toroidal pores, as those revealed in previous all-atom studies. However, when the peptides are placed only on one side of the membrane, the peptide-bound surface acquires a positive curvature, resulting in spontaneous buckling of the bilayer at the  $\mu\text{s}$  timescale. This buckling is presumably associated with the nucleation of giant transient pores, which in turn would explain recent estimates of an effective pore cross section of about 20 nm in the initial phase of the leakage from GUV [63].

Several computational studies were devoted to Protegrin AMPs (for a review see [64]), a 16-18 amino acid long peptides rich in Arg residues with a beta-hairpin configuration stabilized by two Cys-Cys disulfide bonds. In particular all-atom simulations of pre-arranged barrel-stave and toroidal pores of Protegrin I peptides highlighted a higher stability of the latter organization [65].

Alamethicin, a 20-amino acid peptide, is the only AMP for which the barrel-stave channel model has been conclusively demonstrated [39]. In this model, the peptide aggregates forming a cylinder positioned in a transmembrane orientation. The hydrophobic regions of the peptide interact with the core of the bilayer, while the polar faces form the walls of the hydrophilic channel. The lipid head-groups remain well separated into outer and inner leaflets, whereas they merge in the toroidal pore model. Tieleman et al [66] performed all-atom simulations in the tens of ns timescale of different putative arrangements of peptides in channels and found out that the most stable configuration is a regular pore with six peptides. In a mixed coarse-grained/all-atom study Thogersen et al [67] studied peptide aggregation and pore formation, starting from a configuration in which 25 peptide copies, each in the alpha-helix configuration, are placed in a strict transmembrane orientation, forming a regular lattice on a 330 DMCP bilayer patch. During the CG MD simulation in the  $\mu$ s timescale the peptides spontaneously aggregate in a growing cluster. There is, however, a certain degree of disorder in the orientation with respect to the membrane, with several peptides flipping to a membrane-parallel orientation. Subsequent all-atom simulations starting from the CG MD snapshots reveal also deviations from the initial alpha-helical structure. This heterogeneity in both the peptide arrangement in the aggregate and in the peptide internal structure agrees with solid-state NMR measurements reported in the same work and in a further MD/liquid-state NMR study [68]. Other all-atom MD studies were aimed at discriminating between the possible peptide-membrane interaction mechanisms by evaluating the stability of the Alamethicin hexamer inside either a barrel-stave or a toroidal pore [69], pointing to a higher stability of Alamethicin secondary structure in the barrel-stave case.

A mixed experimental/MD simulation study was performed on Indolicidin [70], a 13 AA AMP rich in Trp, Arg/Lys and Pro residues, with the aim of designing mutants retaining antimicrobial activity but with decreased hemolytic activity. The mutations were suggested by simulating the process of Indolicidin absorption and insertion into models of erythrocyte and bacterial membrane (pure POPC and POPC:POPG mixture respectively). The perturbation of the various membrane models was evaluated in terms of bilayer thinning and reduction in the order parameter of the acyl chains during the molecular dynamics. The erythrocyte membrane model was more perturbed after insertion into the bilayer, whereas peptide adsorption was already sufficient to reduce the order parameter of the bacterial-membrane bilayer. Mutations were consequently designed to reduce the insertion into the bilayer, which is assisted by the hydrophobic Trp residues, while maintaining the positively charged Lys residues, which are instead important for adsorption. By mutating some of the Trp residues into Phe, and increasing the number of Lys, it was possible to achieve reduced hemolytic activity and enhanced antimicrobial activity.

Ovispirin is a cathelicidin-like model peptide, 18 AA long with potent, broad spectrum antimicrobial activity. It has been experimentally demonstrated to disrupt the bacterial membrane by the carpet-mode mechanism of action [55]. In this model, peptides associate with the surface of the external layer creating a difference in surface tension, then released by creating membrane defects. MD simulations were performed in order to unravel more details of the mechanism of action of this peptide [56]. Long-time simulations of Ovispirin in a zwitterionic lipid micelle were carried out, starting from an unbiased peptide configuration. The simulation has shown to converge to the correct experimentally observed binding conformation of the peptide. However, due to the intrinsically slow mode of action, a complete internalization by the carpet-mode mechanism of action was unable to be shown.

Lazaridis and co-workers [71] conducted a survey among several helical AMPs, with the aim of correlating the biological activity to chemico-physical quantities such as peptide-membrane binding free energies, either taken from experimental data or calculated using implicit membrane models. Though the correlation are rather weak, deeper insertion into the bilayer was found to correlate with hemolytic activity, whereas surface area occupation correlates with antimicrobial activity. Interestingly, structural flexibility correlates with both activities. This result agrees with the previously mentioned MD studies, all pointing to a certain degree of flexibility in the internal peptide structure.

## 4. AMP synthesis and purification

In AMP design and discovery, there is often need to massively synthesize and analyze potentially active compounds, either for dataset building or experimental validation. Combinatorial chemistry embraces a diversity of techniques that, rapidly and efficiently, enabled the production of large collection of compounds rapidly tested for desirable properties. Moreover, the possibility to automate the process of synthesis and screening by robot and computer-controlled systems takes the combinatorial chemistry to a high-throughput level. However, is not uncommon the need to synthesize peptides with non-natural modifications, like non canonical amino acids incorporation or backbone modifications. Depending on the experimental requirement, different approaches can be followed.

### 4.1. Combinatorial Chemistry

The process involved in the development of lead candidates is time consuming and limited by the number of individual compounds that can be synthesized. Combinatorial chemistry procedures have emerged as a useful tool in drug discovery, specially for lead finding and lead optimization. Different techniques are adopted in Combinatorial chemistry, allowing a rapid and efficient production of large collection of compounds, as well as a convenient and rapid

screening phase. Two main branches of libraries are mainly defined: indexed libraries, where each peptide has a known sequence and non-indexed libraries, where the exact sequence of each peptide is unknown.

Indexed libraries are obtained using parallel synthetic methods that provide final pure samples of single peptides, which facilitates the interpretation of the biological results. Parallel strategies are not convenient for the synthesis of large peptide libraries; however this approach is useful for the lead optimization of a previously identified active compound. For instance, a recent study reported a high-throughput method for the selection of broad-spectrum peptide antibiotics [72] from a biased combinatorial library with structure {RRG}WOLOLOLOY{GRR}-amide where N- and C- terminal cassettes are randomly present or absent and the O residue represents one of the following amino acids: NDTRGAVY. Peptide activities were tested against three different strands together with the activity on synthetic biological membranes. Another example of AMPs activity study with high-throughput methods is an article about peptides derived from PAF26, active against phyto-pathogens [73].

Non-Indexed libraries are prepared using iterative or positional scanning strategies where multiple sequences are produced fixing the amino acids in selected positions and randomizing the amino acids of the other sites. These strategies allow the production of large libraries and peptides are obtained as mixtures. Therefore deconvolution procedures are required to identify the most important amino acid at each position in terms of activity.

A common combinatorial technique is the Solid-phase peptide synthesis (SPPS), which uses organic synthesis on solid support, allowing the synthesis of natural peptides which are difficult to express in bacteria and the incorporation of non-natural amino acids. In SPPS, cycles of coupling-wash-deprotection-wash are repeated for each residue added to the neo-formed peptide chain. Small solid porous beads are treated with functional units ('linkers') on which peptide chains can be built. The peptide remains covalently attached to the bead until the end of the synthesis process, when it will be cleaved by anhydrous hydrogen fluoride or trifluoroacetic acid. The peptide is thus 'immobilized' on the solid-phase and can be retained during a filtration process, whereas liquid-phase excess reagents are flushed away. Unlike ribosome protein synthesis, solid-phase peptide synthesis proceeds in a C-terminal to N-terminal fashion. The free N-terminal amine of a solid-phase attached peptide is coupled to a single amino acid unit. Each amino acid monomers is protected by a *protecting group* at the N-termini, in order to control the reaction process. Currently, two protecting groups are used in SPPS, *t-Boc* and *Fmoc*. After the coupling process, the newly attached unit is de-protected, revealing a new N-terminal amine to which a further amino acid may be attached. Wash cycles are performed after each reaction, removing excess reagent with all the growing peptide of interest remaining covalently attached to the insoluble resin.

SPPS is limited by yields, and typically peptides and proteins in the range of 70 amino acids are pushing the limits of synthetic accessibility. Synthetic difficulty is also sequence dependent; typically amyloid peptides and proteins are difficult to make. Longer lengths can be accessed by using native chemical ligation to couple two peptides together with quantitative yields. During the years, SPPS has been extensively applied to AMPs both for candidate synthesis or naturally-occurring peptides studies [74].

#### 4.2. The SPOT method

A valuable tool in AMP design and discovery is the possibility to screen large number of potentially active compounds. The use of peptide microarrays allows to screen a high number of peptides on a small chip. However, due to the miniscule amounts of peptides synthesized directly on chips, this approach can be difficult and often solution phase assays and analysis are unreliable. These disadvantages can be overcome using peptide macroarrays. The mostly widely used macroarray technique is the SPOT method, where peptides are synthesized directly onto cellulose membranes [75]. Using the SPOT technique, it is possible to synthesize and to screen up to 8,000 peptides, peptide mixtures, or other organic compounds on a letter-size membrane (19 cm × 28 cm). Peptides produced by the SPOT technology can be cleaved from the membrane and directly used in screening for antimicrobial activity. Furthermore, the peptides can also be synthesized on modified membranes and their ability to kill bacteria can be tested even when tethered to a surface.

Because of its convenience, the SPOT method is often used to prepare peptide libraries in order to train machine learning models and infer chemophysical characteristics for a wanted biological activity. In AMP design and discovery, a recent study [76] analyzed different libraries of candidates and tested AMP activity via a luminescence-based assay. Initially, a set of randomized peptide sequences was synthesized and tested against *P.aeruginosa*. Later, a second library was generated on the basis of the information derived from the first, biasing the amino acidic composition according to the most active peptides.

## 5. Conclusions

During the past two decades it has become evident that increasing bacterial drug resistance has created an urgent need for new classes of antibiotics. Even if the panel of traditional antibiotics can still manage drug resistant pathogens, at the moment, AMPs seem to represent one of the most promising future strategies for defeating this threat. At industrial level, several companies worldwide are focused on the development of AMPs with several molecules both at the preclinical and clinical stage. This demonstrates that despite the first clinical trials failing (Pexiganan and Isegran), there is still a lot of general optimism for their use in future clinical practice. Some of the challenges facing the de-

velopment of peptidic drugs have already been overcome, starting from the industrial production of T-20 peptide (Fuzeon) at the multi-ton scale production, incredibly boosting the production of new peptides on a large scale with beneficial effects on the cost of all starting materials. Other challenges are common to other class of molecules that may be defined as innovative. Several strategies have been devised to optimize AMPs with promising activity, ranging from inclusion of non-natural amino acids and a new method applicable to high-throughput screening to multimerization of linear sequences. Even if we cannot exclude the fact that resistance may evolve whenever bacterial populations are consistently exposed to elevated levels of AMPs, this concern should not discourage their further study and development; instead, they may help us to rationalize their use in future, for example preventing the big mistake made in the past of distribution of large amounts of antibiotics, and thereby minimizing the emergence of resistant organisms.

## References

- [1] Fernández L, Hancock REW. Adaptive and mutational resistance: role of porins and efflux pumps in drug resistance. *Clinical microbiology reviews* 25. 2012: 661–681.
- [2] Peters BM, Shirliff ME, Jabra-Rizk MA. Antimicrobial peptides: primeval molecules or future drugs? *PLoS pathogens* 6. 2010: e1001067.
- [3] Batoni G, Maisetta G, Brancatisano FL, Esin S, Campa M. Use of antimicrobial peptides against microbial biofilms: advantages and limits. *Current medicinal chemistry* 18. 2011: 256–279.
- [4] Fjell CD, Hiss JA, Hancock REW, Schneider G. Designing antimicrobial peptides: form follows function. *Nature reviews Drug discovery* 11. 2012: 37–51.
- [5] Maisetta G, Di Luca M, Esin S, Florio W, Brancatisano FL, et al. Evaluation of the inhibitory effects of human serum components on bactericidal activity of human beta defensin 3. *Peptides* 29. 2008: 1–6.
- [6] Fjell CD, Hancock REW, Cherkasov A. AMPer: a database and an automated discovery tool for antimicrobial peptides. *Bioinformatics (Oxford, England)* 23. 2007: 1148–1155.
- [7] Wang G, Li X, Wang Z. APD2: the updated antimicrobial peptide database and its application in peptide design. *Nucleic acids research* 37. 2009: D933–7.
- [8] Thomas S, Karnik S, Barai RS, Jayaraman VK, Idicula-Thomas S. CAMP: a useful resource for research on antimicrobial peptides. *Nucleic acids research* 38. 2010: D774–80.
- [9] Piotto SP, Sessa L, Concilio S, Iannelli P. YADAMP: yet another database of antimicrobial peptides. *International journal of antimicrobial agents* 39. 2012: 346–351.
- [10] Seshadri Sundararajan V, Gabere MN, Pretorius A, Adam S, Christoffels A, et al. DAMPD: a manually curated antimicrobial peptide database. *Nucleic acids research* 40. 2012: D1108–12.
- [11] Rathinakumar R, Wimley WC. Biomolecular engineering by combinatorial design and high-throughput screening: small, soluble peptides that permeabilize membranes. *Journal of the American Chemical Society* 130. 2008: 9849–9858.
- [12] Marks JR, Placone J, Hristova K, Wimley WC. Spontaneous membrane-translocating peptides by orthogonal high-throughput screening. *Journal of the American Chemical Society* 133. 2011: 8995–9004.
- [13] Loose C, Jensen K, Rigoutsos I, Stephanopoulos G. A linguistic model for the rational design of antimicrobial peptides. *Nature* 443. 2006: 867–869.
- [14] Andreu D, Ubach J, Boman A, Wählin B, Wade D, et al. Shortened cecropin A-melittin hybrids. Significant size reduction retains potent antibiotic activity. *FEBS letters* 296. 1992: 190–194.
- [15] Fjell CD, Jenssen H, Fries P, Aich P, Griebel P, et al. Identification of novel host defense peptides and the absence of alpha-defensins in the bovine genome. *Proteins* 73. 2008: 420–430.
- [16] Giangaspero A, Sandri L, Tossi A. Amphipathic  $\alpha$  helical antimicrobial peptides. *European Journal of Biochemistry* 268. 2001: 5589–5600.
- [17] Flower DR, Macdonald IK, Ramakrishnan K, Davies MN, Doytchinova IA. Computer aided selection of candidate vaccine antigens. *Immunome research* 6 Suppl 2. 2010: S1.
- [18] Lapins M, Wikberg JE. Kinome-wide interaction modelling using alignment-based and alignment-independent approaches for kinase description and linear and non-linear data analysis techniques. *BMC bioinformatics* 11. 2010: 339.
- [19] Pissurlenkar RRS, Malde AK, Khedkar SA, Coutinho EC. Encoding Type and Position in Peptide QSAR: Application to Peptides Binding to Class I MHC Molecule HLA-A\*0201. *QSAR & Combinatorial Science* 26. 2007: 189–203.
- [20] Wold S, Jonsson J, Sjöström M, Sandberg M, Rännar S. DNA and peptide sequences and chemical processes multivariately modelled by principal component analysis and partial least-squares projections to latent structures. *Analytica Chimica Acta* 277. 1993: 239–253.
- [21] Wang P, Hu L, Liu G, Jiang N, Chen X, et al. Prediction of antimicrobial peptides based on sequence alignment and feature selection methods. *PloS one* 6. 2011: e18476.
- [22] Fjell CD, Jenssen H, Cheung WA, Hancock REW, Cherkasov A. Optimization of antibacterial peptides by genetic algorithms and cheminformatics. *Chemical biology & drug design* 77. 2011: 48–56.
- [23] Maccari G, Di Luca M, Nifosí R, Cardarelli F, Signore G, et al. Antimicrobial peptides design by evolutionary multiobjective optimization. *PLoS computational biology* 9. 2013: e1003212.
- [24] Michaelson JJ, Sebat J. forestSV: structural variant discovery through statistical learning. *Nature methods* 9. 2012: 819–821.
- [25] Weaver DC. Applying data mining techniques to library design, lead generation and lead optimization. *Current opinion in chemical biology* 8. 2004: 264–270.
- [26] Solmajer T, Zupan J. Optimization algorithms and natural computing in drug discovery. *Drug Discovery Today: Technologies* 1. 2004: 247–252.

- [27] Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 6. 2002: 182–197.
- [28] Wessolowski A, Bienert M, Dathe M. Antimicrobial activity of arginine- and tryptophan-rich hexapeptides: the effects of aromatic clusters, D-amino acid substitution and cyclization. *The journal of peptide research : official journal of the American Peptide Society* 64. 2004: 159–169.
- [29] Dathe M, Meyer J, Beyermann M, Maul B, Hoischen C, et al. General aspects of peptide selectivity towards lipid bilayers and cell membranes studied by variation of the structural parameters of amphipathic helical model peptides. *Biochimica et biophysica acta* 1558. 2002: 171–186.
- [30] Yang S-T, Shin SY, Lee CW, Kim Y-C, Hahn K-S, et al. Selective cytotoxicity following Arg-to-Lys substitution in tritripticin adopting a unique amphipathic turn structure. *FEBS letters* 540. 2003: 229–233.
- [31] Bocchinfuso G, Bobone S, Mazzuca C, Palleschi A, Stella L. Fluorescence spectroscopy and molecular dynamics simulations in studies on the mechanism of membrane destabilization by antimicrobial peptides. *Cellular and molecular life sciences : CMLS* 68. 2011: 2281–2301.
- [32] Tozzini V. Coarse-grained models for proteins. *Current opinion in structural biology* 15. 2005: 144–150.
- [33] Yesylevskyy S, Marrink S-J, Mark AE. Alternative mechanisms for the interaction of the cell-penetrating peptides penetratin and the TAT peptide with lipid bilayers. *Biophysical journal* 97. 2009: 40–49.
- [34] Marrink SJ, de Vries AH, Tieleman DP. Lipids on the move: simulations of membrane pores, domains, stalks and curves. *Biochimica et biophysica acta* 1788. 2009: 149–168.
- [35] Matyus E, Kandt C, Tieleman D. Computer Simulation of Antimicrobial Peptides. *Current Medicinal Chemistry* 14. 2007: 2789–2798.
- [36] Gurtovenko AA, Anwar J, Vattulainen I. Defect-mediated trafficking across cell membranes: insights from in silico modeling. *Chemical reviews* 110. 2010: 6077–6103.
- [37] J. Bond P, Khalid S. Antimicrobial and Cell-Penetrating Peptides: Structure, Assembly and Mechanisms of Membrane Lysis via Atomistic and Coarse-Grained Molecular Dynamic Simulations. *Protein & Peptide Letters* 17. 2010: 1313–1327.
- [38] Chen L, Gao L. How the Antimicrobial Peptides Kill Bacteria: Computational Physics Insights. *Communications in Computational Physics*. 2012.
- [39] Bocchinfuso G, Bobone S, Mazzuca C, Palleschi A, Stella L. Fluorescence spectroscopy and molecular dynamics simulations in studies on the mechanism of membrane destabilization by antimicrobial peptides. *Cellular and molecular life sciences : CMLS* 68. 2011: 2281–2301.
- [40] Almeida PF, Pokorny A. Mechanisms of antimicrobial, cytolytic, and cell-penetrating peptides: from kinetics to thermodynamics. *Biochemistry* 48. 2009: 8083–8093.
- [41] Kandasamy SK, Larson RG. Binding and insertion of alpha-helical anti-microbial peptides in POPC bilayers studied by molecular dynamics simulations. *Chemistry and physics of lipids* 132. 2004: 113–132.
- [42] Leontiadou H, Mark AE, Marrink SJ. Antimicrobial peptides in action. *Journal of the American Chemical Society* 128. 2006: 12156–12161.
- [43] Santo KP, Berkowitz ML. Difference between magainin-2 and melittin assemblies in phosphatidylcholine bilayers: results from coarse-grained simulations. *The journal of physical chemistry B* 116. 2012: 3021–3030.
- [44] Woo H-J, Wallqvist A. Spontaneous buckling of lipid bilayer and vesicle budding induced by antimicrobial peptide magainin 2: a coarse-grained simulation study. *The journal of physical chemistry B* 115. 2011: 8122–8129.
- [45] Bolintineanu DS, Kaznessis YN. Computational studies of protegrin antimicrobial peptides: a review. *Peptides* 32. 2011: 188–201.
- [46] Jang H, Ma B, Lal R, Nussinov R. Models of toxic beta-sheet channels of protegrin-1 suggest a common subunit organization motif shared with toxic alzheimer beta-amyloid ion channels. *Biophysical journal* 95. 2008: 4631–4642.
- [47] Peter Tieleman D, Hess B, Sansom MSP. Analysis and Evaluation of Channel Models: Simulations of Alamethicin. *Biophysical Journal* 83. 2002: 2393–2407.
- [48] Thøgersen L, Schiøtt B, Vosegaard T, Nielsen NC, Tajkhorshid E. Peptide aggregation and pore formation in a lipid bilayer: a combined coarse-grained and all atom molecular dynamics study. *Biophysical journal* 95. 2008: 4337–4347.
- [49] Dittmer J, Thøgersen L, Underhaug J, Bertelsen K, Vosegaard T, et al. Incorporation of antimicrobial peptides into membranes: a combined liquid-state NMR and molecular dynamics study of alamethicin in DMPC/DHPC bicelles. *The journal of physical chemistry B* 113. 2009: 6928–6937.
- [50] Sengupta D, Leontiadou H, Mark AE, Marrink S-J. Toroidal pores formed by antimicrobial peptides show significant disorder. *Biochimica et biophysica acta* 1778. 2008: 2308–2317.
- [51] Santo KP, Berkowitz ML. Difference between magainin-2 and melittin assemblies in phosphatidylcholine bilayers: results from coarse-grained simulations. *The journal of physical chemistry B* 116. 2012: 3021–3030.
- [52] Jean-François F, Elezgaray J, Berson P, Vacher P, Dufourc EJ. Pore formation induced by an antimicrobial peptide: electrostatic effects. *Biophysical journal* 95. 2008: 5748–5756.
- [53] Tsai C-W, Hsu N-Y, Wang C-H, Lu C-Y, Chang Y, et al. Coupling molecular dynamics simulations with experiments for the rational design of indolicidin-analogous antimicrobial peptides. *Journal of molecular biology* 392. 2009: 837–854.
- [54] Elmore DE. Insights into buforin II membrane translocation from molecular dynamics simulations. *Peptides* 38. 2012: 357–362.
- [55] Yamaguchi S, Huster D, Waring A, Lehrer RI, Kearney W, et al. Orientation and dynamics of an antimicrobial peptide in the lipid bilayer by solid-state NMR spectroscopy. *Biophysical journal* 81. 2001: 2203–2214.
- [56] Khandelia H, Kaznessis YN. Molecular dynamics simulations of the helical antimicrobial peptide ovipirin-1 in a zwitterionic dodecylphosphocholine micelle: insights into host-cell toxicity. *The journal of physical chemistry B* 109. 2005: 12990–12996.
- [57] Kandasamy SK, Larson RG. Binding and insertion of alpha-helical anti-microbial peptides in POPC bilayers studied by molecular dynamics simulations. *Chemistry and physics of lipids* 132. 2004: 113–132.
- [58] Leontiadou H, Mark AE, Marrink SJ. Antimicrobial peptides in action. *Journal of the American Chemical Society* 128. 2006: 12156–12161.

- [59] Sengupta D, Leontiadou H, Mark AE, Marrink S-J. Toroidal pores formed by antimicrobial peptides show significant disorder. *Biochimica et biophysica acta* 1778. 2008: 2308–2317.
- [60] Jean-François F, Elezgaray J, Berson P, Vacher P, Dufoure EJ. Pore formation induced by an antimicrobial peptide: electrostatic effects. *Biophysical journal* 95. 2008: 5748–5756.
- [61] Elmore DE. Insights into buforin II membrane translocation from molecular dynamics simulations. *Peptides* 38. 2012: 357–362.
- [62] Woo H-J, Wallqvist A. Spontaneous buckling of lipid bilayer and vesicle budding induced by antimicrobial peptide magainin 2: a coarse-grained simulation study. *The journal of physical chemistry B* 115. 2011: 8122–8129.
- [63] Tamba Y, Ariyama H, Levadny V, Yamazaki M. Kinetic pathway of antimicrobial peptide magainin 2-induced pore formation in lipid membranes. *The journal of physical chemistry B* 114. 2010: 12018–12026.
- [64] Bolintineanu DS, Kaznessis YN. Computational studies of protegrin antimicrobial peptides: a review. *Peptides* 32. 2011: 188–201.
- [65] Jang H, Ma B, Lal R, Nussinov R. Models of toxic beta-sheet channels of protegrin-1 suggest a common subunit organization motif shared with toxic alzheimer beta-amyloid ion channels. *Biophysical journal* 95. 2008: 4631–4642.
- [66] Peter Tieleman D, Hess B, Sansom MSP. Analysis and Evaluation of Channel Models: Simulations of Alamethicin. *Biophysical Journal* 83. 2002: 2393–2407.
- [67] Thøgersen L, Schiøtt B, Vosegaard T, Nielsen NC, Tajkhorshid E. Peptide aggregation and pore formation in a lipid bilayer: a combined coarse-grained and all atom molecular dynamics study. *Biophysical journal* 95. 2008: 4337–4347.
- [68] Dittmer J, Thøgersen L, Underhaug J, Bertelsen K, Vosegaard T, et al. Incorporation of antimicrobial peptides into membranes: a combined liquid-state NMR and molecular dynamics study of alamethicin in DMPC/DHPC bicelles. *The journal of physical chemistry B* 113. 2009: 6928–6937.
- [69] Mihajlovic M, Lazaridis T. Charge distribution and imperfect amphipathicity affect pore formation by antimicrobial peptides. *Biochimica et biophysica acta* 1818. 2012: 1274–1283.
- [70] Tsai C-W, Hsu N-Y, Wang C-H, Lu C-Y, Chang Y, et al. Coupling molecular dynamics simulations with experiments for the rational design of indolicidin-analogous antimicrobial peptides. *Journal of molecular biology* 392. 2009: 837–854.
- [71] He Y, Lazaridis T. Activity determinants of helical antimicrobial peptides: a large-scale computational study. *PloS one* 8. 2013: e66440.
- [72] Rathinakumar R, Wimley WC. High-throughput discovery of broad-spectrum peptide antibiotics. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 24. 2010: 3232–3238.
- [73] Muñoz A, López-García B, Pérez-Payá E, Marcos JF. Antimicrobial properties of derivatives of the cationic tryptophan-rich hexapeptide PAF26. *Biochemical and Biophysical Research Communications* 354. 2007: 172–177.
- [74] Andreu D, Merrifield RB, Steiner H, Boman HG. Solid-phase synthesis of cecropin A and related peptides. *Proceedings of the National Academy of Sciences of the United States of America* 80. 1983: 6475–6479.
- [75] Hilpert K, Winkler DFH, Hancock REW. Peptide arrays on cellulose support: SPOT synthesis, a time and cost efficient method for synthesis of large numbers of peptides in a parallel and addressable fashion. *Nature protocols* 2. 2007: 1333–1349.
- [76] Cherkasov A, Hilpert K, Jenssen H, Fjell CD, Waldbrook M, et al. Use of artificial intelligence in the design of small peptide antibiotics effective against a broad spectrum of highly antibiotic-resistant superbugs. *ACS chemical biology* 4. 2009: 65–74.