Bioinformatics of Antimicrobial Peptides: From Data Collecting to Web servers for Automatic Design of Peptide Antibiotics

Davor Juretić[1]*, Alessandro Tossi[2], Nédia Kamech[3], Nada Ilić[1], Viktor Bojović[4], Mario Novković[1], Juraj Simunić[1], Dražen Petrov[5], Bono Lučić[4], Marija Miljak[1], Josip Ivica[1], Mara Kozić[1] and Damir Vukičević[1]

[1]Faculty of Science, University of Split, Nikole Tesle 12, 21000 Split, Croatia

[2]Department of Life Sciences, University of Trieste, Via Giorgieri 5, 34127 Trieste, Italy

[3]Université Pierre et Marie Curie, Paris 06, Equipe Biogenèse des signaux peptidiques, ER3, 7 Quai Saint-Bernard, Paris cedex 05, France.

[4]Ruđer Bošković Institute, P.O. Box 180, HR-10002 Zagreb, Croatia.

[5]Max F. Perutz Laboratories, University of Vienna, Campus Vienna Biocenter, 1030 Vienna, Austria

*****Author to whom correspondence should be addressed (E-mail: juretic@pmfst.hr)

Abstract. New peptide antibiotics may be able to fill the discovery void of novel antibacterial agents during last decades. University of Split initiated search for new classes of antimicrobial peptides with low toxicity to human cells and good antibacterial activity, which resulted during past several years in finding dozen peptide antibiotics with high therapeutic index. To facilitate additional discoveries we provided free for use and friendly on-line web servers for design of peptide antibiotics.

Keywords: Antimicrobial Peptides, Antibiotics, Web Servers, Drug Design

INTRODUCTION

This paper describes essence of the 21 November, 2012 presentation by the first author at Croatian Academy of Sciences and Arts in Zagreb, Croatia. During short introductory speech, Academician Vlado Paar mentioned previously unknown connection in 1991 between recently born Internet and Prof. D. Juretić role in breaching information blockage around surrounded and bombarded Dubrovnik citizens. A growth of Internet services enabled also most of scientific achievements during the last decade in the field of biological physics and bioinformatics based on ideas and concepts originating from the University of Split, Faculty of Science. These publications deal with three research goals: a) the construction of accurate algorithm for secondary structure prediction of those parts of membrane proteins that are tightly associated with membrane lipids[1], b) modelling bioenergetics and enzyme kinetics by using maximum entropy production principle and maximum information (Shannon) entropy principle[2-4], and c) finding and designing new classes of peptide antibiotics with high therapeutic index (TI)[5-10]. A common thread for these research topics is a desire to get a better insight into structure-function relationships of membrane-active peptides and integral membrane proteins. The topic b) also connects advances in irreversible thermodynamics with applications in different scientific fields. Topics a) and c) are described here with main emphasis on latest successes in producing novel peptide antibiotics expected to be effective against multiple drug resistant strains of bacteria.

Bioinformatics and SPLIT: Historical background

Faculty of Science, Split, has long tradition in producing scientific papers from the field of bioinformatics, going back to 1991, mainly through efforts of Prof. D. Juretić and his associates[1,11-25]. This group also established a custom of constructing dedicated web servers for scientific calculations previously described in publications dealing with advancements in structural bioinformatics. The intention was to provide user friendly services both for students and researchers in that field. Now sadly deceased, Dr. Damir Zucić from the University of Osijek, gave an initial impetus into this direction during his collaboration with the Split group and he helped the construction of mirror servers at the University of Split. Two such servers, in continuous operation since 1998 and 2001 are respectively: http://split.pmfst.hr/split/ and http://split4.pmfst.hr/split/4/ . According to the statistical analysis of server usage in 2007, SPLIT 4.0 server for predicting topology of integral membrane proteins was used by 309 universities from 52 countries, including 20 universities ranked as the best in the world. Two years before, in 2005, server SPLIT 4.0 was ranked among three best servers for predicting sequence position of membrane-spanning helices[26]. It is perhaps worthwhile to mention that the SPLIT server was the first on-line server for scientific calculations in Croatia and at the same time the first server for bioinformatic calculations constructed in Croatia close to the end of previous century. Its user-friendly interface enabled early export of intellectual services from Croatia so that numerous foreign students and researchers expressed their appreciation at being able to use our bioinformatic tools free of charge.

Motivation for antimicrobial peptides research

First antibiotics have been found as natural compounds used by one class of microorganisms (fungi) to eliminate bacterial attacks or competition[27]. They had very specific molecular targets, usually inhibiting enzymes that are unique and essential for bacteria[28,29]. This was valuable advantage for their medical usage, because, apart from rare allergic reactions, such antibiotics are almost completely non-toxic for human cells. At the same time the specificity of molecular targets made it easier for bacteria to develop resistance, to share it with other bacterial strains, and even to develop multiple resistance against several conventional antibiotics[30]. For instance, multiply resistant *Pseudomonas aeruginosa* (gram-negative bacteria) and methicillin resistant *Staphylococcus aureus* (MRSA, gram-positive) and their unrelenting global spread are now major problems in Croatia too[31]. As additional strains of bacteria develop resistance there is an urgent need for next generation antibiotics, but frequency of finding novel antibiotic classes decreased with each passing decade[29], while those few approved for medical practice, the antibiotics of “last resort”, are more toxic and also prone to induce resistance. A double peril of increased bacterial resistance and decreased availability of new antibiotics may lead to decreased longevity and worsening quality of life.

Antimicrobial peptides (AMPs) also act as antibiotics and are present as host defense contrivance in all unicellular and multicellular species examined so far[32]. Some of them are active against different bacterial types, fungi, protozoa and even cancer cells[33]. We have antibiotics with such chemical composition in our bodies as well, usually specialized for defense of particular organs in collaboration with the immune cells. Insects and frogs are abundant sources of AMPs, with some frog species known to have more than 200 different

AMPs[34]. Altogether, around five thousand of known frog and toad species are estimated to contain more than 100 thousand different AMPs[35]. Amphibians experience different microbe laden environments during their life-cycle, while insects and other invertebrates lack an acquired immune system. Rapid de novo synthesis of a battery of AMPs is common defense response of insects to infections.

Once when researchers realized that a gold mine of potential new antibiotic classes exists in insects and amphibians (about 25 years ago), a vigorous work started in isolating, characterizing and testing natural AMPs in parallel with exploration of chemical modifications needed to make them more active against bacteria and less toxic for human cells. This research activity occasionally resulted in press releases with attention-grabbing titles such as: “Modified bee peptide slays deadly bacteria” or “Frog skin may provide antimicrobial peptides effective against multidrug-resistant infections”. When examined in more details, these claims were based on published research papers which described new antimicrobial peptides as lead candidates for developing peptide antibiotics for medical practice. Unfortunately, none of so described numerous peptide-candidates succeeded to break through the regulatory barrier to get a permission for medical usage. To get better insight why so promising antibacterial agents, used millions of years by nature without being made useless by rise of bacterial resistance, did not appear yet in our pharmacies, we should strive to understand their mechanism of action and toxicity problem for human cells, which is connected with AMPs being membrane-active compounds.

Why are AMPs membrane-active bactericidal compounds?

The question what is the mechanism of AMPs bacteriostatic and/or bactericidal activity appeared already in 1987, when first AMPs from African clawed frog *Xenopus laevis* were isolated from skin tissue by Micheal Zasloff. He named them magainins using the Hebrew word “magain”, meaning shield, reflecting their function as an antimicrobial shield[36]. Magainin-2 is the best known member of magainin AMPs family which turned out to have negligible toxicity for human cells and simple linear structure containing 23 amino acids: GIG**K**FLHSA**KK**FG**K**AFVGEIMNS. Net positive charge of this peptide, due to four lysines (red bold letters K), is common property (cationicity) of more than two thousand different anuran defense peptides discovered so far, recently collected in the DADP database http://split4.pmfst.hr/dadp/ by Novković et al.[8]. Most of them do not have any regular secondary structure in water solution, but also most of them acquire α-helical structure when they come in contact with external surface of bacterial cytoplasmic membrane. All-D-enantiomers of peptide antibiotics are equally active[37]. Hence, membrane lipid matrix is primary target of AMPs action rather than chiral proteins.

Another common property of helix-forming AMPs is amphipathicity. One helix side is hydrophilic and charged, while another helix side is hydrophobic and neutral. Helix length of about 20 amino acid residues is enough to span the cytoplasmic membrane. Taken together, these observations suggested the first concept for possible mechanism of AMP action. Oligomers of AMPs can arrange themselves in a barrel form so that all monomers are perpendicular to membrane surface, all hydrophilic helix sides form the membrane aqueous channel (pore) as inner side of the barrel, and all hydrophobic helix sides form the external

barrel surface in contact with membrane lipids[38]. Transient pore formation, due to AMP-membrane interaction, can be responsible for observed increase of membrane permeability to water, cations and anions. Regardless of pore design[33], high permeability increase would certainly reduce bacterial protonmotive force and electric field, making it impossible for bacteria to synthesize ATP by ATP-synthase activity. Blocking bacterial bioenergetics in this manner stops bacterial growth and eventually leads to its death. D. Juretić and collaborators proved that magainins cause drastic decrease of electrochemical proton gradient in bacteria, mitochondria and artificial vesicles with incorporated cytochrome-oxidase proton pumps[39-44].

How to distinguish toxins from antibiotics?

The analysis with the SPLIT algorithm revealed very low preference for peptide antibiotics to enter deeper into membrane interior. Experiments also indicated that peptide antibiotic monomers prefer location at membrane surface to deeper entrance into membrane interior[45]. Cases when SPLIT predicts the formation of membrane-spanning helix are limited to toxic AMPs, such as melittin from bee venom and brevinins-1 from frogs (Figure 1).

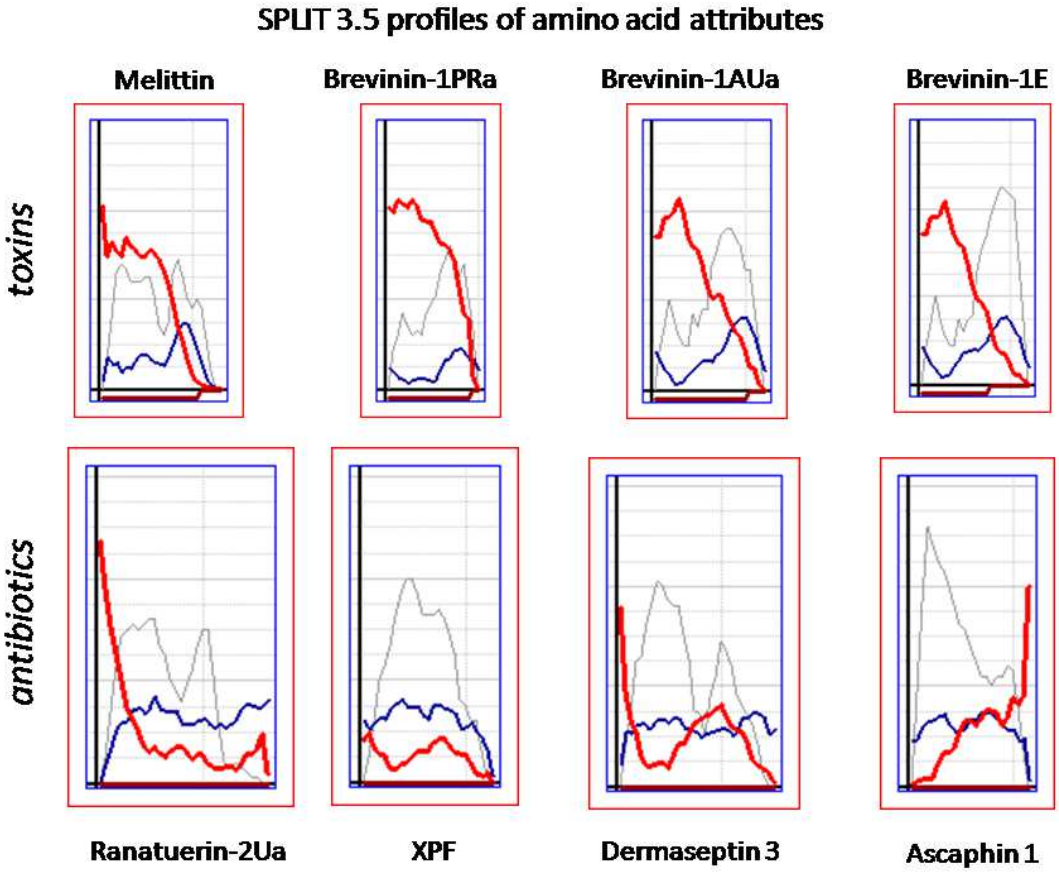

Figure 1. Preference profiles for peptide toxins (upper part) and peptide antibiotics (lower part). Sequence number at the x-axis starts from peptide N-terminal at the origin. First 20 amino acids at x-axis reach up to first thin vertical line. The bold straight line just under the x-axis is SPLIT 3.5 prediction for sequence location of transmembrane helix (http://split.pmfst.hr/split/). Preference profiles are: *red profile* for membrane-buried helix preference, *blue profile* for beta-strand preference, *gray profile* for modified hydrophobic moment index in the case of α-helix[21].

It is of obvious importance to have a quick and easy way of distinguishing AMP toxins, which kill both human and bacterial cells, from potential AMP antibiotics with selective activity against microbial cells. SPLIT algorithm may be useful according to Figure 1, although it was designed to predict topology of integral membrane proteins as its main goal. The TI predictor using SPLIT 3.5 output was constructed by D.J., V.B. and B.L. (available at SPLIT 3.5 home page but unpublished). We found it necessary, however, to develop dedicated computational tools in order to find, design or identify nontoxic AMPs.

Previous attempts to design non-toxic peptide antibiotics with high therapeutic index

A short history of still unsuccessful attempts to introduce magainin analogues into medical practice, after almost 20 years of trying to achieve that goal, may be useful in the case of pexiganan example. Pexiganan, or MSI-78, with highly charged sequence containing nine lysines: GIG**K**FL**KK**A**KK**FG**K**AFV**K**IL**KK**, has excellent broad spectrum activity against gram-negatives and gram-positives[46,47], but it toxicity against human red blood cells increased 20 times with respect to its parent compound[48]. It appears that rational means of reducing pexiganan toxicity to human cells[6] were not explored by Magainin Pharmaceuticals, a company established by M. Zasloff after he left National Institutes of Health. This limited pexiganan usage to topical applications and contributed to the failure of its approval by Food and Drugs Aministration, USA[49].

We can surmise that toxicity to human cells is one of reasons keeping back AMPs from being introduced into medical practice. How one can reduce AMP toxicity without losing its antimicrobial activity? Is there some parameter related to AMP selectivity for bacterial cells that can be easily defined, measured and possibly predicted as well? Such a parameter is well known in pharmaceutics and medicine. While denoted as the therapeutic index (TI), in the case of AMPs, it is usually defined as the ratio of peptide concentration which lyses 50% of mammalian erythrocytes (HC_{50}) to minimal peptide concentration (MIC) which stops overnight growth of bacteria: TI = HC_{50}/MIC. The TI can be increased either by toxicity decrease expressed quantitatively as HC_{50} increase, or by stronger bacteriostatic activity expressed as the MIC decrease. Using some natural AMP as a template to introduce amino acid substitutions at particular sequence positions it is possible to achieve a significant increase in the TI, which is usually due to toxicity decrease[50]. It is difficult, however, to achieve both goals at the same time: toxicity decrease and increase in the antibacterial activity. Most toxicity decreasing substitutions also decrease the peptide antibacterial activity. Also, even best experts cannot avoid subjectivity in a trial and error method for designing better AMP analogues. Just increasing peptide charge, for instance, may fare well for increasing TI and decreasing MIC, but only up to the point when each additional charge leads to increased toxicity and increased MIC[51,52]. In any case, a subjective method of playing with many different amino acid substitutions is expensive practice with low probability of achieving a desired goal of significant TI increase without loss of antimicrobial activity.

RESULTS AND DISCUSSION

Data-mining procedure

In order to predict the therapeutic index, data must be collected and analyzed using a data-mining procedure to propose quantitative structure-selectivity models. While measuring peptide concentrations needed for TI determination is simple in principle, in practice many different methods have been used. It was therefore crucial to select only those published results for anuran peptide sequences and corresponding MIC and HC_{50} concentrations that originated from laboratories with many-years experience in working with such peptides and standardized measurement procedures. As outlined in the Introduction, amphibians have well developed ability to produce AMPs due to the life style of these animals. The construction of the DADP database[8] was a crucial step, which took place (in part) much earlier (from years 2007 to 2009) and made possible the selection procedure described here. The decision to construct the training set of less than 70% identical sequences, with many peptide pairs having very low if any similarity, ensured general nature of our search for appropriate TI-predictor model. A restriction to amphibian helical AMPs tested on *Escherichia coli* strains and mammalian red blood cells (mostly human) enabled us to extract some common general rules from well specified training set of peptides and corresponding activity data.

The sequence moment concept

One can notice already from Figure 1 that N to C-terminal asymmetry exists in hydrophobicity, buried helix preference and amphipaticity profiles of amino acid attributes. This was confirmed through analysis of training set peptides. Also greater biological activity and selectivity was noticed for N-terminal than C-terminal parts of most α-helical AMPs[53] opening the possibility to correlate lengthwise asymmetry in structural profiles and biological activity by using an innovative data-mining procedure. This was done by introducing the sequence moment concept for converting sequence profiles of amino acid attributes into vectors. Bending the line of amino acid codes into an arc allows for associating vector with each amino acid letter (Figure 2). Its length depends on chosen amino acid attribute, such as hydrophobicity, and on chosen smoothing procedure used to calculate sequence profile, so that it can be named, for instance, hydrophobicity index vector. When all such hydrophobicity index vectors are summed, the vector sum is denoted as the sequence moment. Although the sequence moment represents a huge reduction of information contained in corresponding sequence profile, its direction and length preserve the information about lenghtwise asymmetry of sequence profile. Interestingly, when two different hydrophobicity scales are used, corresponding sequence moments can point toward similar or quite different direction even when these scales are highly correlated. We therefore explored if the angle between sequence moments and its cosine (named the D-descriptor) can be related to therapeutic index. In other words, we were looking for the best pair of amino acid attributes (amino acid scales) which would produce the best correlation with the therapeutic index[5]. The data-mining procedure pointed toward Janin's[54] and Guy's scale[55] as the best pair of amino acid scales.

Magainin 2 and its F5W analogue (Figure 2) will serve to illustrate the point that sequence moments, constructed by using Janin's (red arrows) and Guy's hydrophobicity scale (blue arrows), are sensitive indicators of significant selectivity change (the TI increase) caused by only one amino acid substitution[56].

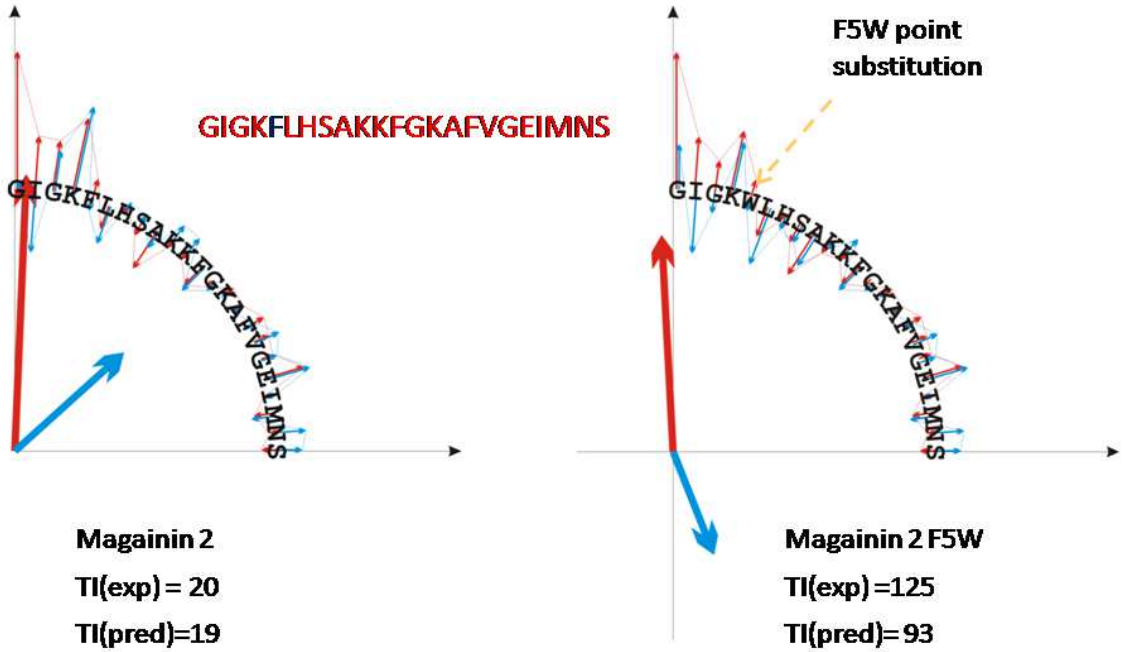

Figure 2. Sequence moments (large arrows). D-descriptor, the cosine of the angle between sequence moments, leads to predicted therapeutic index in accord with measured TI(exp). It is sensitive to a single amino acid change (Phe into Trp in the case of magainin-2)

Predicting the therapeutic index and testing prediction accuracy

The correlation coefficient among measured TI values TI(exp) and calculated D-descriptor values is $r^2 = 0.83$ for 36 non-homologous anuran AMPs from the training set peptides, while the determination coefficient for the test set peptides (a total of 37 peptides) is $r^2 = 0.64$[5]. These are satisfactory indications of correlation for a linear one-descriptor fit: TI(pred)= 50.1 – 44.8D. The TI predictor was tested with 11 pseudin analogues within the charge range +2 to +6, and the correlation achieved among measured and predicted TI values was also very good ($r^2 = 0.84$)[5]. Therefore, our TI predictor works equally well with similar and non-similar peptides with their antimicrobial and hemolytic activity tested in many different laboratories. We followed a good practice of constructing on-line web servers for scientific calculations and provided the TI-predictor server for everybody interested in bioinformatics of AMPs to use it freely: http://split4.pmfst.hr/split/dserv1/ . The TI predictor was tested with analogues of natural peptides synthesized in laboratories of our foreign collaborators Prof. Dr. Alessandro Tossi, University of Trieste, Trieste, Italy and Dr. Nédia Kamech, University Pierre and Marie Curie, Paris, France. Predictions were confirmed in each case[5,9,10].

Predicting amino acid substitutions expected to increase the therapeutic index

The computational tool for predicting the change in the TI upon single or double amino acid substitutions was constructed jointly by Davor Juretić and Damir Vukičević, both professors at the Faculty of Science, University of Split. It was named the Mutator by one of us (D.V.). The Mutator algorithm did not explore DNA single point mutations. It explored only those point mutations in AMPs that substituted one amino acid with another and retained only modified peptides predicted to have higher TI than their parent compounds. The TI-predictor is the core of the Mutator algorithm, which additionally contains filters and subroutines designed to eliminate from consideration all substitutions with predicted effect of decreasing peptide amphipathicity and of decreasing frequency of amino acid motifs commonly found among AMPs with high antimicrobial activity. In the end, the Mutator accepted only substitutions with predicted TI being equal or greater from TI(parent peptide)+ 10. For *E. coli* ATCC 25922 strain, predicted trend for TI increase without loss of antibacterial activity was confirmed in synthesized and tested ascaphin-8 and XT-7 analogues[9]. For four of these analogues the HC_{50} was greater than 800 μM, which is about 10 times lower toxicity with respect to parent AMPs. In practice, it is a negligible toxicity. Therefore, the Mutator is proposed as an objective tool for creating analogues of natural AMPs with better therapeutic potential due to their decreased toxicity for human cells. It is also provided as a free-of-charge on-line scientific server that everybody can use: http://split.pmfst.hr/mutator/. A combination of DADP database perusing and the Mutator server usage by copy-paste procedure for peptides selected from DADP is a powerful and simple method for in silico design of new peptide antibiotics that even high school and university students can use within their research projects with a teacher or a mentor that has some knowledge of bioinformatics.

"In silico" experiments and experimental verifications

A serendipitous benefit of the Mutator working with peptides that have similar activity against gram-negative (*E. coli*) and gram-positive (*S. aureus*) bacteria is that significant increase in measured therapeutic index can be achieved for both classes of bacteria. Examples of improved TI for ascaphin-8 and XT-7 analogues (as output of the Mutator algorithm) are [I^2, K^{19}]-ascaphin-8 with predicted and measured TI(*E. coli*) being 80 and > 480 respectively, and [K^2, K^{16}]-XT-7 with predicted and measured TI(*E. coli*) being 89 and > 128 respectively[9]. Almost equally good experimental result was achieved for *S. aureus*, a gram-positive bacteria that was not used in training the TI-predictor and Mutator algorithms. Measured TI(*S. aureus*) was > 170 and > 267 for ascaphin and XT analogue respectively[9].

The Mutator algorithm and corresponding server are not restricted to predicting just two amino acid substitutions in natural anuran peptides. If predicted TI after two best substitutions is still modest in comparison with maximal possible prediction of TI = 95, another round of Mutator application can be started by using the output sequence after the first round as a query sequence for the second round. In practice, one round of Mutator algorithm application is often enough to obtain predicted sequence with estimated TI higher than 85. An important cavaet is that the Mutator can produce null result too. Initial query sequence may have very low predicted TI, a situation that makes it difficult for the Mutator to predict any substitution expected to increase TI. A summary of such in-silico experiments makes it clear that the Mutator is not able to predict new classes of peptide antibiotics, but

predicting analogues of natural AMPs with expected high increase in selectivity with respect to parent peptides is still a valuable service to research community interested in peptide antibiotics. For example, Mutator's predicted analogues of ascaphin-8 and XT-7 maintain their strong broad-spectrum activity against gram-negatives and gram-positives and can be considered as novel peptide antibiotics due to their very low toxicity to human cells[9].

Transformation of cell-penetrating peptide into antimicrobial peptide

Cell penetrating peptides (CPPs) are able to cross the plasma membrane of eukaryotic cells[57]. Some CPPs are also antimicrobial peptides and vice versa[58,59]. While CPPs are very interesting and useful due to their unique ability to act as cargo delivery peptides, thus facilitating the delivery of drugs to the cytoplasm of targeted cells, their antimicrobial activity and computational methods how to enhance it has not attracted much attention so far[60]. The MAP peptide KLALKLALKALKAALKLA-NH_2 is an example of a cationic and amphiphatic CPP with moderate antimicrobial activity against gram-negatives and undesirably high toxicity for red blood cells[61]. A combination of manually introduced substitutions and three rounds of the Mutator server usage produced the MAP analogue with predicted high selectivity TI(pred) = 95, increased charge, increased hydrophobic moment and decreased hydrophobicity. Synthesis of this analogue and biophysical experiments revealed that it assumes an α-helical structure in membrane-mimicking solvents and confirmed its strong bacteriostatic activity against gram-negatives (MIC around one μM), but toxicity was not completely eliminated[62]. Measured therapeutic index was around TI(*E. coli*) = 30. Perhaps this is not surprising, because the parent peptide (MAP) does not belong to anuran AMPs used for training the TI-predictor.

The Designer algorithm and adepantins

New classes of peptide antibiotics can be predicted by the Designer algorithm, also constructed jointly by Professors Damir Vukičević and Davor Juretić from the University of Split. One of us (DJ) named them adepantins, an abbreviation for **A**utomatically **De**signed **P**eptide **Anti**biotics. Before going into brief description how the Designer algorithm works, let us show first how it was verified in practice (Figure 3) by synthesis and activity/selectivity testing in the laboratory of Prof. Alessandro Tossi from the University of Trieste[5,6,10]. It turned out that the bacteriostatic activity of adepantins (the MIC value) is very good for *E. coli* (in the range from 1.0 to 4.0 μM) and essentially absent against *S. aureus*. The toxicity of adepantins is negligible, while their therapeutic index is very high (Table 1).

The Designer algorithm uses the output of the PredictorSelector algorithm, which selects several descriptors associated with correlation higher than 0.90 (either positive or negative) among descriptor values and experimentally determined TI(exp) values when 36 non-homologous frog-derived antimicrobial peptides are used (the training data set). The D-descriptor (see *Sequence moment concept* paragraph), so extracted, is however not enough to achieve a goal of rational computational design of novel AMPs. The AMP design must take into account that net positive charge, amphipathicity and hydrophobicity of designed peptides should be in the proper range for helical peptide antibiotics.

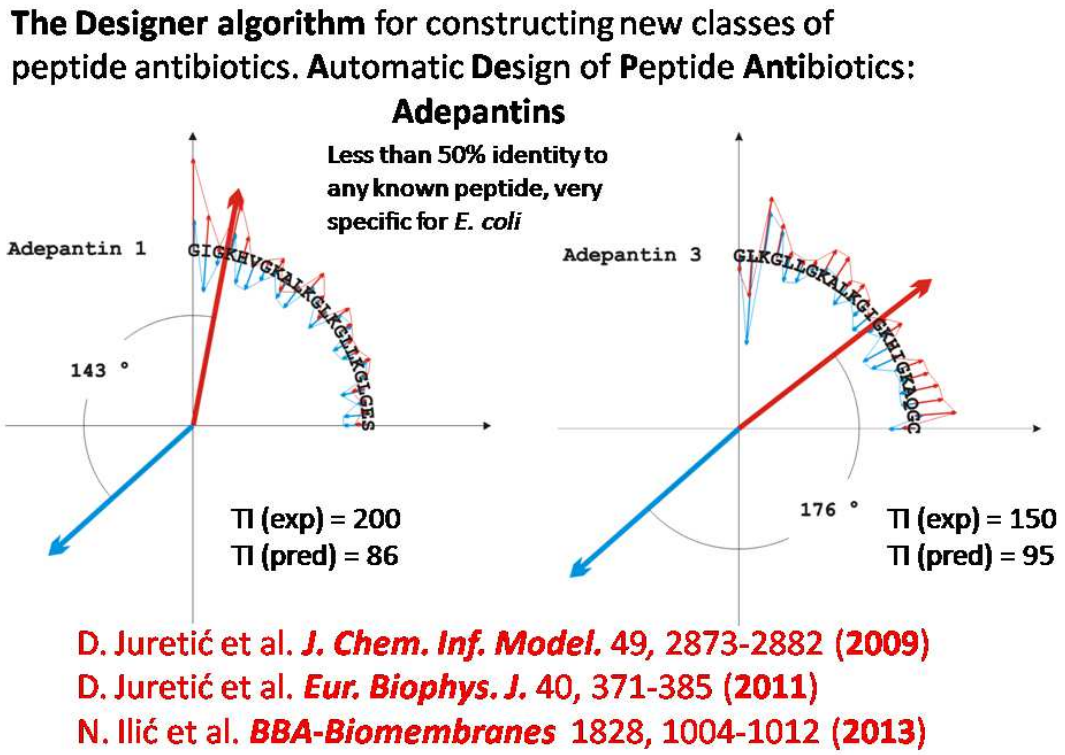

Figure 3. Adepantins 1 and 3 compared with respect to their predicted and measured therapeutic index. For details see the caption of Figure 2 and corresponding text.

By using a recursive algorithm, with flexible statistical and physicochemical restrictions, the Designer software constructs and selects a small number of antimicrobial peptides expected to form helical structure in membrane environment and also to have high selectivity against gram-negative bacteria such as the *E. coli*. The Designer (designer.cpp) and PredictorSelector (PredictorSelector.cpp) source codes have been created and compiled with the Microsoft Visual C++ software 6.0. For local usage these source codes and instructions how to use them are freely available at the site: http://sites.google.com/site/adepantin1

Table 1 summarizes the results obtained with adepantins by MSc and PhD students from the University of Split and University of Trieste. The Designer software selected only seven adepantins out of $8x10^{29}$ possible peptides having 23 amino acids in their sequence (not quite arbitrary choice of sequence length, because it is equal to magainin-2 sequence length and close to the mean length of anuran AMPs). We synthesized and tested only three out of these seven adepantins[5,10]. Adepantins 1 and 2 are different only in the C-terminal amino acid, while adepantin 3 is the most different (among seven adepantins) peptide in comparison to adepantin 1 (Table 2). Peptides with amidated C-terminal have one additional positive charge, each lysine also contributes one positive charge, while histidine charge ranges from 0 to +1 depending on the pH and its local environment.

Adepantin dimers

Adepantin dimers (Table 2) were also synthesized and tested[10]. Dimers have remarkably good antibacterial activity against all gram-negative bacterial strains we tested

(*Escherichia coli, Pseudomonas aeruginosa, Salmonella typhimurium, Klebsiella pneumoniae*), but probably due to doubling of their net positive charge, their toxicity to human red blood cells is about 20 times higher than for corresponding monomers. Like monomers, adepantin dimers are inactive against gram-positive *Staphylococcus aureus* (MIC equal to 128 μM). Glycine rich pattern, associated with all adepantins (Table 2), may contribute to adepantins selectivity for gram-negatives.

Table 1. Antimicrobial and cytotoxic activities of adepantin peptides against *E. coli*[a]

Antibiotic	Adepantin-1	Adepantin-2[b]	Adepantin-3	ADP-2 dimer	ADP-3 dimer
TI(pred)	86	91	95	93	95
TI(exp)	160	400	>125	32	40
MIC(μM)	3	1	4	0.5	0.5
HC_{50}(μM)	480	400	>500	16	20

[a]Bactericidal concentrations are two-fold higher than bacteriostatic (MIC) concentrations
[b]Table cells with yellow background highlight the best peptide (adepantin-2) with measured lowest MIC(*E. coli*) and highest TI.

Table 2. Sequences of designed adepantin peptides[a]

Peptide	Amino acid sequence
Adepantin-1	GIGKHVGKALKGLKGLLKGLGES–NH_2
Adepantin-2	GIGKHVGKALKGLKGLLKGLGEC–NH_2
Adepantin-3	GLKGLLGKALKGIGKHIGKAQGC–NH_2
ADP-2 dimer	GIGKHVGKALKGLKGLLKGLGECCEGLGKLLGKLGKLAKGVHKGIG-NH_2
ADP-3 dimer	GLKGLLGKALKGIGKHIGKAQGCCGQAKGIHKGIGKLAKGLLGKLG-NH_2

[a]All sequences are amidated at their C-terminal. Colour code: red for lysines and histidines (K and H), blue for hydrophobic residues isoleucine, leucine and valine (I, L, V) and yellow for glycine (G).

Mechanism of action of adepantins and future prospects

Taking into account all activity data[10] it appears that adepantin 2 monomer is the best lead compound for specific eradication of *E. coli* strains. Additional experiments[10] confirmed the validity of our computational approachs to design AMPs. Adepantins adopt a helical structure in contact with negatively charged membranes (typical for bacterial cells), which favours insertion into and permeabilisation of these, whilst they only appear to interact with surface of neutral membranes (such as those of human cells). Our results also indicate that the outer membrane and lipopolysaccharide layer of gram-negative bacteria is not a significant barrier for adepantins, while cell wall of gram-positive bacteria and/or their different phospholipid composition keep them resistant to adepantins, an observation that makes adepantins potentially useful tool for exploring selectivity mechanism at the level of different bacterial types too. For practical goal of achieving cheaper adepantins synthesis, their excellent selectivity against gram-negatives and very low toxicity with respect to eukaryotic cells, may be also valuable if gram-positive bacterial strains and/or yeast cells are used as possible microbial factories for the production of these antibiotics.

Discovery of new natural AMPs classes as an intrinsically difficult problem

New classes of natural peptide antibiotics certainly exist in nature and will be discovered in the future, providing that progress of civilisation and deforestation does not cause extermination of endangered species before scientists get a chance to examine their rich supply of antimicrobial peptides. At present, such studies are underfunded, so that researchers are confronted with a painful choice which species deserve to be subjected to genome analysis which may reveal (among many other insights) sequences of their host defense peptides. Still, so many new species are added each year to the list of species with decoded genome and proteome, that the bottleneck in discovering new natural classes of peptide antibiotics is more the abundance of available data for analysis and shortage of bioinformatic methods how to do it, than the absence of data for extinct/exterminated species. The main problem with existing bioinformatic tools such as BLAST, or recently created HMMER, is their focus on finding similar sequences. It is possible of course to disregard warnings these tools have when entering into a twilight zone with non-significant similarity to query AMP. The problems then multiply, because a new question arises, namely, how one can have any degree of certainty that novel natural AMPs is found, if it is not annotated as an AMP and is not similar to known AMPs. However, an insight into conservation of precursors for AMPs can teach us what would be the best alternative approach.

Using conserved signal peptides as bait to catch novel AMPs

We can again use the DADP database[8] as a starting point for this research goal. One can easily notice unusually high conservation of signal peptides when compared with the conservation of mature AMPs (Table 3). Such conservation is unusual because for exported polypeptides functionally important part is only a mature protein, not the signal sequence, which is removed and disregarded after it has been recognized by a signal recognition particle and used to help with protein export. On average, signal peptides of secretory proteins evolve two to five times faster than mature secretory proteins[63], which is just the opposite of what is observed for AMP precursors[64].

As a rule, precursors of anuran AMPs have tripartite structure consisting of signal peptide at the N-terminal, acidic propeptide in middle part, and mature AMP at the C-terminal. Anuran AMP precursors from Table 3 are one of many examples of precursors with identical signal peptides (yellow background). Acidic propeptide parts are also identical for two AMP precursors from *Odorrana* family, but corresponding AMPs (nigrosin-2Sb and odorranain-H1) have only 57% pairwise identity. Temporin-1P from *Lithobates* family has only 15% pairwise identity with nigrosin and 28% identity in acidic propeptides.

Teleost fishes more often have mature AMP immediately after the signal peptide, while acidic propeptide is at the C-terminal of an AMP precursor (Table 3). There is still a case of identical signal peptides (yellow background) from AMP precursors belonging to the same fish familly (*Moronidae*) but pairwise identity between mature AMPs (moronecidin and dicentracin alignment) and between acidic propeptides at the C-terminal is similarly high (91% and 88% respectively). For a different family of fishes (*Sciaenidae*) all elements of tripartite structure are different from corresponding parts of the moronicidin precursor with the least difference found for signal peptides (77 % identity), greater difference for mature

AMPs moronecidin and piscidin-like peptide (41% identity), and with only 5% identity in acidic propeptides. For another family of teleost fishes (*Gasterosteidae*) a pairwise comparison of moronecidin and hfp1 AMP precursor[7] results in 73%, 14%, and 21% identity respectively for signal peptides, mature AMPs and acidic propeptides. The only consistent result of this analysis is a much better conservation of signal peptides both for anurans and fishes when compared with remaining parts of AMP precursor structure

Table 3. Examples for conserved signal sequences and variable mature AMPs[a].

Anuran (frog) precursor for AMP			
Signal peptide	Acidic propeptide	Mature AMP	AMP / *genus*
MFTLKKSLLLLFFLGTINLSLC	QDETNAEEE-RRDEEVAKMEEI-KR	GILSGVLGMGKKIVCGLSGLC	Nigrosin-2Sb/ *Odorrana*
MFTLKKSLLLLFFLGTINLSLC	QDETNAEEE-RRDEEVAKMEEI-KR	GLFGKILGVGKKVLCGLSGMC	Odorranain-H1/ *Odorrana*
MFTLKKSLLLLFFLGTINLSLC	EEERDADEEERRDDSDESNVEVEKR	FLPIVGKLLSGLL	Temporin-1P/ *Lithobates*
Teleost (fish) precursor for AMP			
Signal pepide	Mature AMP	Acidic propeptide	AMP / *family*
MKCATLFLVLSMVVLMAEPGDA	FFHHIFRGIVHVGKTIHRLVTG	AEQDQQDQQYQQEQQEQQA QQYQRFNRERAAFD	Moronecidin/ *Moronidae*
MKCATLFLVLSMVVLMAEPGDA	FFHHIFRGIVHVGKSIHKLVTG	AQQDQQDQQYQQDQQDQQA EQYQRFNRERAAFD	Dicentracin/ *Moronidae*
MKCTALFLVLSLVVLMAEPGEC	IWGLIAHGVGHVGRLIHGLIRG	AEEQHVQLDKRSLSYDPPKK LQW-----RE----D	Piscidin-like peptide/ Sciaenidae
MKYVTIFLVLSLVVLMADPGDC	SFKKFWGGVKAIFKGARKGWK	EHRAIARSHRGQEQQGQQVV NYEGQPYWQD	sticklefish hfp1 *Gasterosteidae*

[a] bold letters in red colour are used to highlight differences

Covalent connection of well conserved signal peptides with mature AMPs in the same precursor sequence can be an inspiration to perform indirect search for novel natural AMPs by searching biological databases for those classes of signal sequences that are known to be associated with AMPs. Such a search in the UNIPROT database with anuran signal peptide query from the first column of Table 3 reveals much higher abundance of AMPs than direct search for mature AMPs, but novel AMP classes are unlikely to be found in the UNIPROT or GenBank collection of sequences. However, there is one huge database, the database of EST sequences, containing parts or whole cDNA sequences, which is largely unexplored for the presence of novel AMP classes.

Hfp1 peptide. A case of wrong (human) annotation in the EST database?

The third method used to find new classes of peptide antibiotics led to the last result in the Table 3 (the hfp1 precursor)[7]. It required: 1) a good choice of signal peptide which is associated with many different AMPs, 2) limited variations in signal peptide sequence still leading to many different AMPs and easily recognized tripartite structure, 3) the TBLASTN tool available at http://blast.ncbi.nlm.nih.gov/, and 4) a choice of expressed sequence tags

database (EST) with signal peptide as a query. For example, with the signal peptide query for moronecidin (Table 3) two interesting hits are found: a) *Trichoplax adhaerens* cDNA, E-value 0.036 and maximal identity of 76%, and *Homo sapiens* cDNA, E-value 0.097 and maximal identity of 76%. Other TBLASTN hits in EST with E-values less than 0.1 lead to mostly known AMP precursors from different families of teleost fishes, but *Trichoplax adhaerens* hit and *Homo sapiens* hit span the whole range between the simplest free-living animals (placozoans)[65] and humans, provided of course that these are not annotation errors.

In the case of "Homo sapiens" hit we proved that it contains AMP precursor (the last row of Table 3), which most likely originated from some species of sticklefish *Gastereosteus aculeatus*[7]. In other words, it is indeed an annotation error which happened during decoding of human genome. Nevertheless, so found mature peptide with provisional name hfp1 and sequence SFKKFWGGVKAIFKGARKGWK was synthesized, tested and shown to have low toxicity and very good broad spectrum antibacterial activity (Figure 4). It is also unique, the first member of a new class of AMPs, being less than 14% identical to all other peptides in the UNIPROT database. We concluded that conserved signal peptides (SPs) can be used as bioinformatic bait for finding fish and anuran AMPs, because the search with SPs is more effective than use of mature AMP alone.

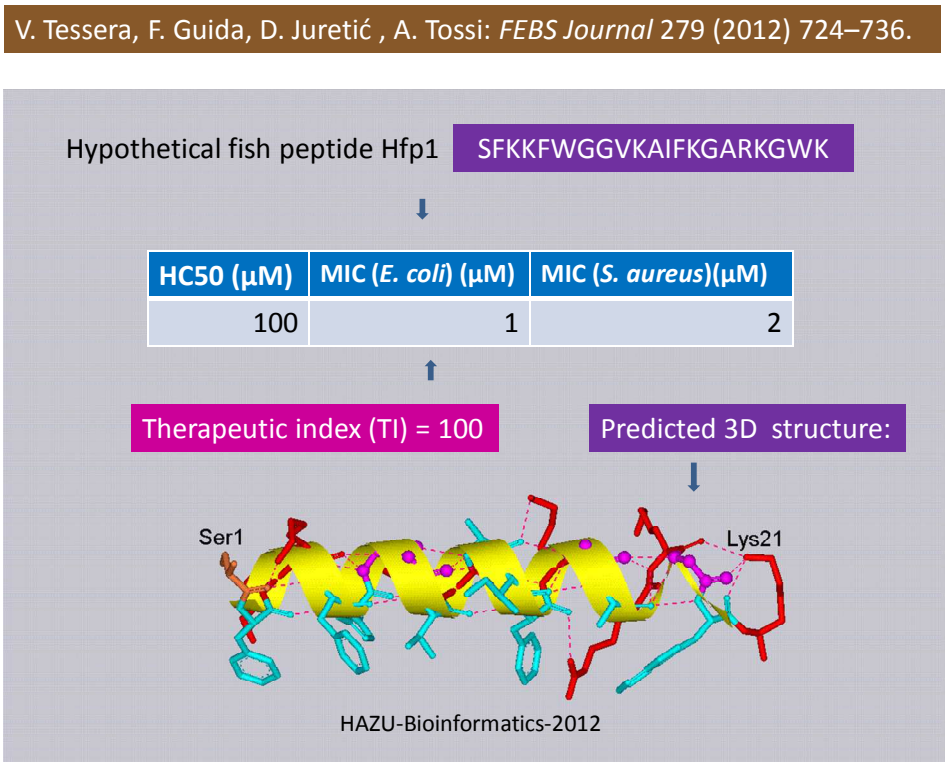

Figure 4. AMP found in EST databases and named Hfp1, due to its annotation as human peptide, most likely originates from some sticklefish species.

Sequential dimer preparation

Some weakly active, very short AMPs, can be extended by repeating twice the same sequence pattern. In the case of helix forming dimer, it is expected that sequential dimer made from two monomers, each about 10 amino acids long, will gain the ability to bridge the

membrane in favorable circumstances and so stimulate membrane permeabilization. We decided to use the PGLa-H from the DADP database reported to have modest antibacterial activity[66]. When synthesized and tested at the University of Trieste it did not show antibacterial activity. In contrast, its sequential dimer had MIC (*E. coli*) in the range from one to two μM and high therapeutic index[67]. We concluded that sequential dimerization of well chosen short AMPs can be a powerful method for achieving substantial increase in antibacterial activity, while maintaining low toxicity against human cells.

Remaining tasks

A computational method for MIC prediction is more difficult to develop and may require using 3D structure-activity descriptors[68]. We still do not know which structural aspects of AMPs are crucial for selective destabilization of gram-negative and gram-positive cytoplasmic membrane. For potential therapeutic applications, novel AMPs described in this paper should be tested with greater variety of multiple resistant bacterial strains. Applications are not limited to AMPs antibacterial activity. For instance, antitumor activity is possible, but has not been examined so far for those AMPs. The potential of the Designer and Mutator algorithms is much broader then reported here. With few little changes in the algorithm the Designer algorithm can produce “in silico” hundreds of additional potential peptide antibiotics which would need experimental verifications. A combination of Mutator algorithm application and other bioinformatic approaches for finding or constructing AMPs is also capable of producing hundreds of AMP analogues. Food conservation and topical applications, either alone or in synergy with conventional antimicrobials, are some of possible knowledge-transfer avenues for novel peptide antibiotics from university laboratories into a commercial sector.

ACKNOWLEDGEMENT

This work was partly supported by the Croatian Ministry of Science, Education and Sports grants (177-1770495-0476 to D.J., 098-1770495-2919 to B.L. and V.B., 177-0000000-0884 and 037-0000000-2779 to D.V.) and the Friuli Venezia Giulia L26 project R3A2. N.I. acknowledges an Alpe Adria 2009 grant. J.I. and M.M. acknowledge the support of the Erasmus Student Exchange Programme between University of Split, Croatia, and University of Trieste, Italy.

SAŽETAK

Novi peptidni antibiotici bi mogli popuniti zabrinjavajući nedostatak u otkrićima novih klasa antibiotika u ovom stoljeću. Sveučilište u Splitu potaklo je potragu za novim klasama peptidnih antibiotika s niskom toksičnošću za ljudske stanice i dobrom antibakterijskom aktivnošću. Kroz nekoliko proteklih godina ta potraga dovela je do otkrića desetak peptidnih antibiotika s visokim terapeutskim indeksom. Da bi olakšali dodatna otkrića konstruirali smo i on-line web poslužitelje za konstrukciju peptidnih antibiotika koje svi zainteresirani mogu slobodno i lako koristiti.

REFERENCES

1. D. Juretic, L. Zoranic, and D. Zucic, *J. Chem. Inf. Comput. Sci.* **42** (2002) 620-632.
2. D. Juretic and P. Zupanovic, *Comp. Biol. Chem.* **27** (2003) 541-553.

3. R. C. Dewar, D. Juretić, and Paško Županović, *Chem. Phys. Lett.* **430** (2006) 177-182.
4. A. Dobovišek, P. Županović, M. Brumen, Ž. Bonačić-Lošić, D. Kuić, and D. Juretić, *Biophys. Chem.* **154** (2011) 49-55.
5. D. Juretić, D. Vukičević, N. Ilić, N. Antcheva, and A. Tossi, *J. Chem. Inf. Model.* **49** (2009) 2873-2882.
6. D. Juretić, D. Vukičević, D. Petrov, M. Novković, V. Bojović, B. Lučić, N. Ilić, and A. Tossi, *Eur. Biophys. J.* **40** (2011) 371-385.
7. V. Tessera, F. Guida, D. Juretić, and A. Tossi, *FEBS J.* **279** (2012) 724–736.
8. M. Novković, J. Simunić, V. Bojović, A. Tossi, and D. Juretić, *Bioinformatics* **28** (2012) 1406-1407.
9. N. Kamech, D. Vukičević, A. Ladram, C. Piesse, J. Vasseur, V. Bojović, J. Simunić, and D. Juretić, *J. Chem. Inf. Model.* **52** (2012) 3341–3351.
10. N. Ilić, M. Novković, F. Guida, D. Xhindoli, M. Benincasa, A. Tossi, and D. Juretić, *BBA-Biomembranes* **1828** (2013) 1004-1012.
11. D. Juretic and R.W. Williams, *J. Math. Chem.* **8** (1991) 229-242.
12. D. Juretic, *Croat. Chem. Acta* **65** (1993) 921-932.
13. D. Juretic, B.K. Lee, N. Trinajstic, and R.W. Williams, *Biopolymers* **33** (1993) 255-273.
14. D. Juretic, N. Trinajstic, and B. Lucic, *J. Math. Chem.* **14** (1993) 35-45.
15. D. Juretic, B. Lucic, and N. Trinajstic, *Croat. Chem. Acta* **66** (1993) 201-208.
16. D. Juretic, B. Lucic, and N. Trinajstic, *J. Mol. Struc-Theochem.* **338** (1995) 43-50.
17. D. Juretic and R. Pesic, *Croat. Chem. Acta* **68** (1995) 215-232.
18. B. Lucic, S. Nikolic, N. Trinajstic, A. Juric, and D. Juretic, *Croat. Chem. Acta* **68** (1995) 435-450.
19. B. Lucic, S. Nikolic, N. Trinajstic, and D. Juretic, *J. Chem. Inf. Comp. Sci.* **35** (1995) 532-538.
20. D. Juretic, D. Zucic, B. Lucic, and N. Trinajstic, *Comput. Chem.* **22** (1998), 279-294.
21. D. Juretic and A. Lucin, *J. Chem. Inf. Comput. Sci.* **38** (1998) 575-585.
22. D. Juretic, A. Jeroncic, and D. Zucic, *Croat. Chem. Acta* **72** (1999) 975-997.
23. D. Juretic, A. Jeroncic, and D. Zucic, *Periodicum Biologorum* **101** (1999) 339-347.
24. D. Zucic and D. Juretic, *Croat. Chem. Acta* **77** (2004) 397-401.
25. D. Juretic, B. Lucic, and N. Trinajstic, *Periodicum Biologorum* **107** (2005) 379-383.
26. J.M. Cuthbertson, D.A. Doyle, and M.S.P. Sansom, *Prot. Eng. Des. Sel.* **18** (2005) 295-308.
27. K.M. Overbye and J.F. Barrett, *Drug Discov. Today* **10** (2005) 45-52.
28. H. Yoneyama and R. Katsumata, *Biosci. Biotechnol. Biochem.* **70** (2006) 1060-1075.
29. L.L. Silver, *Clin. Microbiol. Rev.* **24** (2011) 71-109.
30. A.N. Neely and I.A. Holder, *Burns* **25** (1999) 17-24.
31. A.T. Andrašević and T. Tambić (Eds), *Antibiotic resistance in Croatia, 2009*, The Croatian Academy of Medical Sciences, Zagreb, 2010.
32. M. Zasloff, *Nature* **415** (2002) 389-395.
33. Y. Li, Q. Xiang, Q. Zhang, Y. Huang, and Z. Su, *Peptides* **37** (2012) 207-2015.
34. X. Yang, W-H. Li, and Y. Zhang, *J. Proteome Res.* **11** (2012) 306-319.
35. D. Vanhoye, F. Bruston, P. Nicolas, and M. Amiche, *Eur. J. Biochem.* **270** (2003) 2068-2081.
36. M. Zasloff, *Proc. Natl. Acad. Sci. USA* **84** (1987) 5449-5453.
37. D. Wade, A. Boman, B. Wahlin, C.M. Drain, D. Andreu, H.G. Boman, and R.B. Merrifield, *Proc. Natl. Acad. Sci. USA* **87** (1990) 4761-4765.
38. D.M. Ojcius and J.D.E. Young, *Trends Biochem. Sci.* **16** (1991) 225-229.

39. D. Juretic, H.-C. Chen, J.H. Brown, J.L. Morell, R.W. Hendler, and H.V. Westerhoff, *FEBS Lett.* **249** (1989) 219-223.
40. H.V. Westerhoff, D. Juretic, R.W. Hendler, and M. Zasloff, *Proc. Natl. Acad. Sci. USA* **86** (1989) 6597-6601.
41. H.V. Westerhoff, R.W. Hendler, M. Zasloff, and D. Juretic, *Biochim. Biophys. Acta* **975** (1989) 361-369.
42. D. Juretic, *Stud. Biophys.* **138** (1990) 79-86.
43. D. Juretic, R. W. Hendler, F. Kamp, W. S. Caughey, M. Zasloff, and H. V.Westerhoff, *Biochemistry* **33** (1994) 4562-4570.
44. H.V. Westerhoff, M. Zasloff, J.L. Rosner, R.W. Hendler, A.De Wall, A Vaz Gomez, A.P.M. Jongsma, A. Riethorst, and D. Juretic, *Eur. J. Biochem*. **228** (1995) 257-264.
45. B. Bechinger, M. Zasloff, and S.J. Opella, *Biophys. J*. **74** (1998) 981-987.
46. W.L. Maloy and U.P. Kari, *Biopolymers* **37** (1995) 105-122.
47. Y. Ge, D.L. MacDonald, K.J. Holroyd, C. Thornsberry, H, Wexler, and M. Zasloff, *Antimicrob. Agents Ch.* **43** (1999) 782-788.
48. S. Rotem, I. Radzishevsky, and A. Mor, *Antimicrob. Agents Ch.* **50** (2006) 2666-2672.
49. L.M. Gottler and A. Ramamoorthy, *Biochim. Biophys. Acta* **1788** (2009) 1680-1686.
50. A. Hawrani, R.A. Howe, T.R. Walsh, and C.E. Dempsey, *J. Biol. Chem.* **283** (2008) 18636-18645.
51. M. Dathe, H. Nikolenko, J. Meyer, M. Beyermann, and M. Bienert, *FEBS Lett*. **501** (2001) 146-150.
52. Z. Jiang, A.I. Vasil, J.D. Hale, R.E.W. Hancock, M.L. Vasil, and R.S. Hodges, *Biopolymers* **90** (2008) 369-383.
53. A. Tossi, L. Sandri and A. Giangaspero, *Biopolymers* **55** (2000) 4−30.
54. J. Janin, *Nature* **277** (1979) 491-492.
55. H. R. Guy, *Biophys. J.* **47** (1985) 61-70.
56. T. Tachi, R.F. Epand, R.M. Epand, and K. Matsuzaki, *Biochemistry* **41** (2002) 10723-10731**.**
57. F. Milletti, *Drug Discov. Today*. **17** (2012) 850-860.
58. S.T. Henriques, M.N. Melo, and M.A.R.B. Cantanho, *Biochem. J.* **399** (2006) 1-7.
59. K. Splith and I. Neundorf, *Eur. Biophys. J.* **40** (2011) 387-397.
60. W.L. Zhu, H. Lan, I.S. Park, J.I. Kim, H.Z. Jin, K.S. Hahm, and S.Y. Shin, *Biochem. Biophys. Res. Commun.* **349** (2006) 769-774.
61. E. Strandberg, D. Tiltak, M. Ieronimo, N. Kanithasen, P. Wadhwani, and A.S. Ulrich, *Pure Appl. Chem.* **79** (2007) 717-728.
62. J. Ivica, *Algorithmic and knowledge-based modifications of cell-pepetrating peptide leading to its improved physicochemical properties and antimicrobial activity*. MSc thesis, Faculty of Science, University of Split, Split, Croatia, January 2013.
63. E.J.B. Williams, C. Pal, and L.D. Hurst, *Gene* **253** (2000) 313-322.
64. P. Nicolas and D. Vanhoye, and M. Amiche, *Peptides* **24** (2003) 1669-1680.
65. M. Srivastava et al. *Nature* **454** (2008) 955-960.
66. F. Hou, J. Li, P. Pan, J. Xu, L. Liu, W. Liu, B. Song, N. Li, J. Wan, and H. Gao, *Int. J. Antimicrob. Agents* **38** (2011), 510-515.
67. M. Miljak, *Dimeric analogue of a natural antimicrobial peptide with improved activity and selectivity: A study connecting bioinformatics, biophysics and biochemistry.* MSc thesis, Faculty of Science, University of Split, Split, Croatia, January 2013.
68. M. Kozić, Asymmetric antimicrobial peptides and their measured properties, MSc thesis, Faculty of Science, University of Split, Split, Croatia, December 2012.