Computational Design of Highly Selective Antimicrobial Peptides

Davor Juretić,*,† Damir Vukičević,† Nada Ilić,† Nikolinka Antcheva,‡ and Alessandro Tossi‡

Department of Physics, Faculty of Science, University of Split, 21000 Split, Croatia and Department of Life Sciences, University of Trieste, 34127 Trieste, Italy

Received August 31, 2009

We have created a structure-selectivity database (AMPad) of frog-derived, helical antimicrobial peptides (AMPs), in which the selectivity was determined as a therapeutic index (TI), and then used the novel concept of sequence moments to study the lengthwise asymmetry of physicochemical peptide properties. We found that the cosine of the angle between two sequence moments obtained with different hydrophobicity scales, defined as the D-descriptor, identifies highly selective peptide antibiotics. We could then use this descriptor to predict TI changes after point mutations in known AMPs, and to aid the prediction of TI for *de novo* designed AMPs. In combination with an amino acid selectivity index, a motif regularity index and other statistical rules extracted from AMPad, the D-descriptor enabled construction of the AMP-Designer algorithm. A 23 residue, glycine-rich, peptide suggested by the algorithm was synthesized and the activity and selectivity tested. This peptide, adepantin 1, is less than 50% identical to any other AMP, has a potent antibacterial activity against the reference organism, *E. coli*, and has a significantly greater selectivity (TI > 200) than the best AMP present in the AMPad database (TI = 125).

1. INTRODUCTION

New antibiotics are urgently needed due to the increasingly serious multidrug-resistance problem. Strong candidates are endogenous antimicrobial peptides (AMPs) with multimodal mechanisms of action.[1–4]

All higher organisms produce cationic antimicrobial peptides,[5] whose main targets of action are membranes and associated biological functions.[6] Antimicrobial peptides (AMPs) have remained effective against bacteria for millions of years.[7] Bacteria cannot easily reorganize, nor can they markedly alter, their membrane lipid composition or their requirement for a strong, inside-directed electric field (attracting cations). For these reasons, AMPs have a lower tendency to elicit resistance and often display a moderate selectivity for bacterial cells.

Known natural AMPs from different sources exceed 1000s in number (see the Web site: http://www.bbcm.univ.trieste.it/~tossi/antimic.html). A large group of these AMPs acquire an amphipathic α-helical secondary structure in an anisotropic and low dielectric constant environment, such as when they interact with the cytoplasmic membrane.[8] In most cases, their primary or secondary structure has not been related with quantitative parameters expressing their activity against either bacterial or host cells. Even when structural parameters affecting activity have been defined, it is a challenge to maintain or increase antibacterial activity while simultaneously decreasing hemolytic activity, since some optimal combination of specific attributes must be reached.[9,10] Rational design methods for improving the balance between antibacterial activity and cytotoxicity of AMPs are, therefore, highly desirable to harness this ancient host defense mechanism for therapeutic applications.

Increasing peptide selectivity requires a better definition of those structural features of AMPs which have evolved to ensure a selective action in fending off microbes without endangering host cells. This selectivity is usually measured and expressed as the therapeutic index TI = HC_{50}/MIC, where HC_{50} is peptide concentration needed to reach 50% lysis of red blood cells, while antimicrobial activity is expressed as the minimal inhibitory concentration (MIC) for bacterial growth. In this study, we have chosen as reference the Gram-negative bacterium *E. coli* for collecting the most widely available MIC data for AMPs isolated from different frog species, and HC_{50} data obtained using fresh human erythrocytes. Frog species are abundant sources of AMPs[11,12] and Gram-negative bacterial species have more complex mechanisms for developing multiple resistance to classical antibiotics.[13] Initially, we constructed the AMPad database by connecting peptide primary and secondary membrane-associated structure with experimental MIC, HC_{50}, and TI values and used data-mining procedures for statistical analysis of nonhomologous peptides from the database and identification of descriptors capable of distinguishing highly selective from mediocre peptide antibiotics. The rationale for this work and our second goal was to then prove the utility of selected descriptors and data mining procedures through computational "*ab initio*" design of highly selective peptide antibiotics with little identity to any known natural or synthetic AMP. We also present experimental tests that verify the design method.

2. METHODS

2.1. AMPad Database. AMP sequences and corresponding MIC and HC_{50} concentrations were obtained from a comprehensive literature survey and used to create the AMPad database (Supporting Information and corresponding references). Published results were selected only where

* Corresponding author e-mail: juretic@pmfst.hr.
† University of Split.
‡ University of Trieste.

10.1021/ci900327a CCC: $40.75 © XXXX American Chemical Society

antimicrobial activity was determined by the serial dilution method. Data from a total of 73 frog-derived AMPs were collected, such that each had predicted and/or experimentally verified propensity to form amphipathic helices in membrane-mimetic solvent. If MIC values were presented for several *E. coli* strains, the value for the most sensitive strain was selected. Fresh human blood from healthy donors was the normative for selecting the HC_{50} values. In a few cases, the HC_{50} values were extrapolated from given values below and above HC_{50}.

To extract rules as general as possible we divided the AMPad database in a training set of 36 nonhomologous peptides (less than 70% pairwise identity, SI Table S1) and a testing set of 37 remaining peptides (SI Table S2). All but two peptides (pexiganan and gaegurin 4 W16) in the training set are of natural origin.

Statistical analysis of best peptide antibiotics required an alternative designation of the set of best peptides (SBP) ($TI_{exp} \geq 20$, the first 11 and the first 15 peptides, respectively, from the training and testing sets from the AMPad database) and of the full set of peptides (SAP) (all AMPad peptides in SI Tables S1 and S2).

2.2. Calculation of Sequence Moments. Sequence profiles, such as the Kyte−Doolittle hydrophobicity profile,[14] are two-dimensional plots of sequence position-dependent amino acid attributes that are usually represented as smoothed values. In this work, the whole sequence profile for a chosen set of amino acid attributes is converted into one vector with two components (x and y). This is done by bending the sequence, for instance, in a right angle arc, and then by performing vector summation of attribute vectors associated with all amino acids in the sequence. The sequence bending angle is defined as the anticlockwise angle from the sequence C-terminal amino acid residue located at the positive x-axis. For the right angle bending arc the sequence N-terminal amino acid is located at the positive y-axis. The length and orientation of the attribute vector for a given amino acid residue in the sequence corresponds to its direct or smoothed attribute, while the residue sequence location determines the inclination of that vector with respect to the x-axis. The vector sum of attribute vectors, that replaces the sequence profile, is thus an indicator of peptide lengthwise asymmetry. It does not depend on any knowledge of the peptide's secondary structure and should thus not be confused with Eisenberg's hydrophobic moment,[15] which measures sidewise asymmetry of the peptide's regular structure (amphipathicity). The peptide lengthwise asymmetry and the observation that amino acids closer to the N-terminus are more relevant for AMP activity[8,16] is additionally modeled by weighting factors $(l + 1 - i)^q$, where l is the sequence length, i is residue sequence number, and q is a constant in the range from 0 to 1. The sequence moment $m = v(p)$ for position-dependent properties $p = (p_1,...,p_l)$, is then defined as the weighted sum of p_i-vectors associated with each amino acid i in a peptide sequence bent in an arc α. After performing vector summation of $\vec{p}_i(l + 1 - i)^q$ terms, the resulting sequence moment x and y coordinates are given by the following:

$$m = v(p) = v_{\alpha,q}(p) = \begin{pmatrix} \sum_{i=1}^{l} \cos\left(\frac{l-i}{l-1}\cdot\alpha\right)\cdot p_i\cdot(l+1-i)^q, \\ \sum_{i=1}^{l} \sin\left(\frac{l-i}{l-1}\cdot\alpha\right)\cdot p_i\cdot(l+1-i)^q \end{pmatrix}$$

Sequence profiles of amino acid attributes were converted into sequence moments by considering:

(a) 144 different scales of amino acid attributes used to study integral membrane proteins and protein folding.[17–19]

(b) three simple methods for finding sequence profiles p: (i) direct usage of amino acid attributes; (ii) smoothed amino acid attributes associated with a central amino acid in a sliding window (mean values); and (iii) smoothed amino acid attributes associated with a central amino acid in a sliding window without taking into account the attribute of the central amino acid itself (sequence environment values).

(c) 15 possible peptide bending angles $\alpha = \pi/15, \pi/14,...,\pi$.

(d) five possible weighting values $q = 0,1,1/2,1/3,1/4$.

By using the training data set of 36 nonhomologous peptides we obtained sequence moments (m', m'') for all pairs of sequence profiles (p', p'') corresponding to the same angle and same weighting value.

2.3. Converting Sequence Moments into Scalar Descriptors. One-parameter models were obtained by transforming each (m', m'') pair into 36 numbers $d_{i,j,k}(m', m'')$ defined by:

$$d_{i,j,k}(m', m'') = \begin{Bmatrix} 1, & i = 0 \\ |m'|, & i = 1 \\ 1/|m'|, & i = 2 \text{ and } |m'| \neq 0 \\ 0, & i = 2 \text{ and } |m'| = 0 \end{Bmatrix} \cdot \begin{Bmatrix} 1, & j = 0 \\ |m''|, & j = 1 \\ 1/|m''|, & j = 2 \text{ and } |m''| \neq 0 \\ 0, & j = 2 \text{ and } |m''| = 0 \end{Bmatrix} \cdot \begin{Bmatrix} 1, & k = 0 \\ \cos(\angle(m', m'')), & k = 1 \\ (\angle(m', m'')), & k = 2 \\ \sin(\angle(m', m'')), & k = 3 \end{Bmatrix}$$

Notice that some of these numbers correspond to one amino acid scale only. Such numbers are length or inverse length of sequence moment vectors m. Other scalar descriptors reflect the difference in sequence profiles calculated with two different amino acid scales, such as the angle between two sequence moments or some function involving the angle between two sequence moments.

2.4. Amino Acid Selectivity Index and Motif Regularity Index. An alternative data-mining procedure exploits amino acid motifs more frequently found in the best peptide antibiotics from the AMPad database.

The motif regularity index, which is introduced below, measures how well the designed peptide incorporates motifs that are the most common in the structure of antimicrobial peptides in the SBP data set (see Methods). It combines results of statistical analysis that led us to the following:

Table 1. Amino Acid Selectivity Index, AASI, for 20 Amino Acids

AA	*AASI*	*AA*	*AASI*
E	3.14	S	2.14
D	3.13	Y	2.01
Q	3.05	W	2.00
H	3.00	I	1.97
G	2.67	R	1.91
M	2.50	A	1.89
V	2.37	L	1.74
N	2.33	C	1.73
K	2.28	F	1.53
T	2.18	P	0.22

(a) definition of an amino acid selectivity index scale,

(b) the requirement that the residue succeeding each amino acid in the sequence should be one of the five most common successors of that amino acid in the SBP database, and

(c) the requirement that amino acid placed at the $i + 4$ position with respect to any amino acid (spatially close in an α-helical conformation) should be one of the five most common successors at distance $i + 4$ of that amino acid in the SBP data set.

The amino acid selectivity index *AASI* assigned to any given amino acid X is defined as follows:

$$\frac{\sum_{i=1}^{n} f_i(x) \cdot \ln(TI_i)}{\sum_{i=1}^{n} f_i(x)}$$

where n is the number of peptides in the SBP data set of best peptides, $f_i(x)$ is the number of occurrences of amino acid X in peptide i and TI_i is the therapeutic index of i-th peptide. The amino acid selectivity index thus selects those amino acid types whose relative frequency is highest among more effective frog-derived antibiotics (Table 1). Note that this index scale reflects the "selectivity" but not necessarily the antimicrobial "potency".

The requirements (a), (b), and (c) lead to the construction of 20×20 matrices with elements A_{ij} and B_{ij} (Tables 2 and 3).

The significance $s(XY)$ that amino acid Y is successor of amino acid X is given by the following equation:

$$s(XY) = \sum_{i=1}^{n} f_i(XY) + \sum_{i=1}^{n} \frac{f_i(Y)}{l_i \cdot n}$$

where $f_i(XY)$ is the number of occurrences of the XY motif in a peptide i, l_i is the length of the i-th peptide, and $f_i(Y)$ is the number of occurrences of amino acid Y in the i-th peptide.

The values 1,2,3 and 4 are assigned to four of the most significant successors among all amino acids and the value 5 is assigned to the fifth most significant successor only among amino acids E, D, Q, and H. In this way, four amino acids that are the most significant for selective frog-like peptide antibiotics are additionally favored (see Table 1).

The significance $s(X___Y)$ that amino acid Y is successor of amino acid X at the $i + 4$ position is given by the following equation:

$$s(X___Y) = \sum_{i=1}^{n} f_i(X___Y) + \sum_{i=1}^{n} \frac{f_i(Y)}{l_i \cdot n}$$

Table 2. The Most Significant Successor Amino Acids in the SBP Data Set[a]

	A	C	L	M	E	Q	H	K	V	I	F	Y	W	T	G	S	D	N	P	R
A	4	0	3	0	0	5	0	1	0	0	0	0	0	0	2	0	0	0	0	0
C	4	0	3	0	0	5	0	1	0	0	0	0	0	0	2	0	0	0	0	0
L	0	0	2	0	0	0	0	1	0	0	0	0	0	0	3	4	5	0	0	0
M	0	0	2	0	0	5	0	0	0	0	0	0	0	0	0	1	4	3	0	0
E	1	2	4	0	0	5	0	0	0	3	0	0	0	0	0	0	0	0	0	0
Q	1	2	0	0	0	5	0	0	0	0	0	0	0	0	3	0	0	0	4	0
H	0	0	0	0	5	0	0	4	2	1	0	0	0	0	0	3	0	0	0	0
K	4	0	0	0	0	0	5	3	0	2	0	0	0	0	1	0	0	0	0	0
V	3	0	1	0	0	5	0	4	0	0	0	0	0	0	2	0	0	0	0	0
I	0	0	0	0	0	0	0	1	0	0	3	0	0	4	2	0	5	0	0	0
F	0	0	3	0	0	5	0	1	2	0	0	0	0	0	4	0	0	0	0	0
Y	4	0	3	0	0	5	0	1	0	0	0	0	0	0	2	0	0	0	0	0
W	4	0	3	0	0	5	0	1	0	0	0	0	0	0	2	0	0	0	0	0
T	2	0	3	0	0	5	0	0	4	0	0	0	0	0	1	0	0	0	0	0
G	0	0	3	0	5	0	0	1	4	2	0	0	0	0	0	0	0	0	0	0
S	4	0	1	0	0	5	0	2	3	0	0	0	0	0	0	0	0	0	0	0
D	0	0	3	0	5	2	0	0	1	0	0	0	0	4	0	0	0	0	0	0
N	0	0	1	0	0	0	5	2	3	0	0	0	0	0	0	4	0	0	0	0
P	0	0	4	0	0	5	0	1	0	0	0	0	0	0	3	2	0	0	0	0
R	3	0	0	0	0	5	0	0	0	0	0	0	0	1	2	0	4	0	0	0

[a] In this table with A_{ij} elements we present the order of five most significant successors.

Table 3. The Most Significant Successor Amino Acids at the $i + 4$ Position in the SBP Data Set[a]

	A	C	L	M	E	Q	H	K	V	I	F	Y	W	T	G	S	D	N	P	R
A	1	0	3	0	5	0	0	0	4	2	0	0	0	0	0	0	0	0	0	0
C	0	0	4	0	0	5	0	2	0	0	0	0	0	0	1	0	0	3	0	0
L	4	0	2	0	0	0	5	1	0	0	0	0	0	0	3	0	0	0	0	0
M	4	0	3	0	0	5	0	2	0	0	0	0	0	0	0	0	0	0	0	1
E	0	0	4	0	0	5	0	3	0	0	0	0	0	0	0	1	2	0	0	0
Q	0	0	0	0	2	5	0	1	0	0	0	0	0	0	4	0	0	0	3	0
H	3	0	0	0	5	0	0	1	0	0	0	0	0	0	0	4	0	2	0	0
K	0	0	0	0	5	4	0	1	0	0	0	0	0	3	2	0	0	0	0	0
V	2	0	1	0	0	0	0	0	3	4	0	0	0	0	0	0	5	0	0	0
I	0	2	0	0	0	5	0	0	3	1	0	0	0	0	4	0	0	0	0	0
F	1	0	4	0	0	0	5	2	0	0	0	0	0	0	3	0	0	0	0	0
Y	4	0	3	0	0	5	0	1	0	0	0	0	0	0	2	0	0	0	0	0
W	3	0	1	0	0	5	0	2	0	4	0	0	0	0	0	0	0	0	0	0
T	0	0	3	0	5	0	1	2	4	0	0	0	0	0	0	0	0	0	0	0
G	0	0	1	0	0	0	5	2	0	0	0	0	0	0	3	0	4	0	0	0
S	0	0	0	0	5	2	0	3	4	0	0	0	0	0	1	0	0	0	0	0
D	0	1	0	0	0	5	0	3	0	0	0	0	0	2	4	0	0	0	0	0
N	0	0	0	0	0	4	5	1	0	0	0	0	0	0	2	3	0	0	0	0
P	4	0	3	0	0	5	0	1	0	0	0	0	0	0	2	0	0	0	0	0
R	0	0	4	0	0	5	0	2	0	0	0	0	0	0	3	0	1	0	0	0

[a] In this table with B_{ij} elements we present the order of the five most significant successors at the $i + 4$ position.

where $f_i(X___Y)$ is the number of occurrences of $X___Y$ motive in peptide i and l_i is the length of the i-th peptide. Analogously, as in Table 2, the value 5 is assigned only to E,D,Q and H.

The motif regularity index (MR) of a tested peptide sequence ps is defined by the following:

$$MR(ps) = \frac{\sum_{i=1}^{l-2} A_{ps(i)ps(i+1)} + \sum_{i=1}^{l-5} B_{ps(i)ps(i+4)}}{2l - 7}$$

where l is the length of the peptide sequence ps, A_{ij} is the (i,j) entry listed in Table 2 and B_{ij} is the (i,j) entry listed in Table 3. Lower MR values imply greater regularity.

2.5. Mean Hydrophobicity and Relative Amphipathicity. The global hydrophobicity and the amphipathicity for each synthesized peptide sequence were calculated using a hydrophobicity index (Hi) scale derived from the normalized and filtered consensus of 163 published scales, that arbitrarily ranges between maximum values of +10 for Phe, and −10 for Arg.[8,20] The hydrophobicity is given as the mean value ($H = (\Sigma \mathrm{Hi})/l$), where l = peptide length. The mean hydrophobic moment (μH_{max}) was calculated as described by Eisenberg et al. (1982).[15] The relative amphipathicity ($\mu H/\mu H_{max}$) for each peptide was then determined with respect to the value of the maximum hydrophobic moment for a perfectly amphipathic, 18-residue peptide composed only of the most hydrophobic and hydrophilic residues, Phe and Arg ($\mu H_{max} = 6.4$ with our scale), respectively. This relative measure of amphipathicity is less likely to vary according to the scale used than an absolute value.

The online server http://www.bbcm.univ.trieste.it/~tossi/HydroCalc/HydroMCalc.html can be used for H and μH calculations as described above.

2.6. Peptide Synthesis, Purification and Characterization. Adepantin 1 amide, ascaphin-1 amide (ASC 1) and pseudin-2 amide (PSEU 2) were synthesized by Fmoc-solid phase peptide synthesis on a Microwave-enhanced CEM Liberty synthesizer using Fmoc-PAL-PEG-PS resin (substitution 0.34 mmol/g). The cleaved/deprotection cocktail was trifluoroacetic acid (TFA), water and triisopropylsilane (95:2.5:2.5).

Peptides were synthesized chemically in good yield and the crude purity was over 90%. The peptides were then purified by reversed phase HPLC on a semipreparative Waters column (XTerra C18, 7 μm, 300 Å, 19 × 300 mm) using a 25−45% CH_3CN in 60 min gradient with a 5 mL/min flow. Several suspension/lyophilization cycles in 10 mM HCl were used to remove traces of TFA remaining from HPLC. Peptides quality and purity were verified by mass spectrometry (Esquire 4000 Ion Trap, Bruker Daltonics Inc., Germany), while peptide concentrations were determined by mass and verified by the Waddell method.[21] On the basis of analytical HPLC and mass spectrometry, purity was >99%.

The propensity of each synthesized peptide to assume a helical conformation was probed by CD spectroscopy on a Jasco J-715 spectropolarimeter (Jasco, Japan), using 2 mm path length quartz cells and peptide concentrations of 20 μM, in 5 mM sodium phosphate buffer pH7, in the absence and presence of 50% (v/v) trifluoroethanol (TFE). All spectra are the mean of at least two trials, each with the accumulation of three scans. Helical content (% Helix) was estimated from the molar ellipticity at 222 nm, according to the method of Chen et al.[22]

2.7. Hemolytic Activity (HC_{50}). The hemolytic activity of peptides was assessed using freshly isolated human erythrocytes from healthy donors, by monitoring the release of hemoglobin at 405 nm. Aliquots of cell suspension (0.5% erythrocytes) were incubated in triplicate with different peptide concentrations (from 10 μM to 800 μM) in PBS at 37 °C for 30 min. Total lysis (100% hemolysis) was determined by addition of 0.2% Triton X-100. The HC_{50} value was taken as the mean concentration of the peptide producing 50% hemolysis.

2.8. Antimicrobial Activity (MIC). The bacteriostatic activity of the peptides was determined against *Escherichia coli* ATCC 25922, *Pseudomonas aeruginosa* ATCC 27853, and *Staphylococcus aureus* ATCC (10^5 cfu/mL bacteria) as minimum inhibitory concentrations (MIC), using the microdilution susceptibility test in MH broth as described previously[16] or as the effect on bacterial growth kinetics (10^6 cfu/mL bacteria in PBS) in the presence of different peptide concentrations (2, 4, and 8 μM), monitoring the optical density at 600 nm for 4 h.

3. RESULTS

3.1. Sequence Moments and the D-Descriptor. Instead of just using peptide mean properties as given descriptors to correlate with antimicrobial activity and selectivity, we have decided to exploit the observation[9,23] that position-dependent physicochemical properties are also quite important. The definition of sequence moments from residue hydrophobicity scales (see Methods) enabled us to convert sequence profiles into vectors. After the subsequent conversion of these sequence moments into many possible scalar descriptors, we then used the correlation between a given scalar descriptor and measured TI in the training data set of 36 nonhomologous peptides as the selection criterion for optimal descriptors. With a choice of scalar descriptors corresponding to a single scale of amino acid attributes, the highest correlation with measured TI reached only the $r^2 = 0.52$. Much better results were obtained with a triplet defined by choosing two different amino acid scales and one smoothing technique (see below). In this way, we assigned 2700 descriptors to each triplet. The best correlation between descriptor and observed TI in the training data set was assigned to the chosen triplet. Out of thousands of triplets, only three corresponded to correlation coefficients $r^2 \geq 0.81$ in the training data set. The sequence moments for these best descriptors were obtained by summing all sequence environment vectors, i.e., by using the smoothing process that excludes the central amino acid in the sliding window.

The simplest descriptor (among 3 best descriptors corresponding to these triplets) was the cosine of the angle between sequence moments associated with Janin's[24] and Guy's hydrophobicity scale,[25] a peptide sequence bending arc of $\alpha = \pi/2$ and weighting factor $q = 1/3$. This was named the D-descriptor and was associated with a correlation $r^2 = 0.83$ when the training data set was used (Figure 1). The remaining two descriptors were significantly more complex and only one of them was associated with a slightly higher correlation coefficient: $r^2 = 0.838$. Hence, our choice fell on the D-descriptor.

After calculating the D-descriptor values for the training set peptides the predicted TI was found as the best linear fit: $TI = 50.126 - 44.803D$ (Figure 1). Predicted TI values, by using this linear one-descriptor fit, are provided for all peptides in the AMPad database (Supporting Information). The correlation between measured and predicted TI so obtained for the test set of peptides (SI Figure S1) was then expressed as a determination coefficient: $R^2 = 0.64$, an acceptable indication of correlation.

It should be pointed out that statistical methods were used here in a nonstandard manner. Namely, the data in both the training and testing data set are not normally distributed. Hence, standard statistical practice would be not to use Pearson r correlation, but to use a nonparametric method

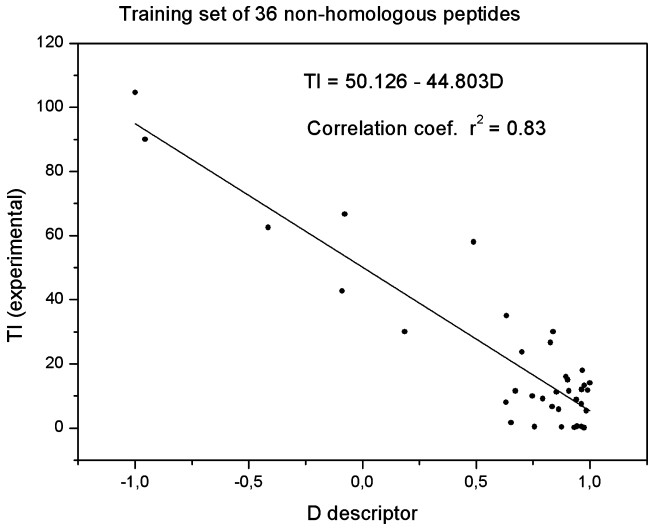

Figure 1. One parameter linear model for correlation between measured therapeutic index (*y*-axis) and predicted D-descriptor (*x*-axis) in the case of the training set of nonhomologous peptides. The D-descriptor is the cosine of the angle between two sequence moments for a chosen peptide.

instead. However, we are not interested in an index that correctly orders peptides according to their TI. Instead, our goal is to find an index that well predicts which peptides have a very high TI value. It can be easily shown that a given percent error in predicting peptides with high values of TI much more influences the Pearson coefficient than the same percent error in predicting peptides with low values of TI. Therefore, r extracts information which is significant to our study.

We also addressed the concern that only 11 frog-derived peptide antibiotics with high TI determined the selection of the D-descriptor among many other possible descriptors. In order to verify that our predictions were not purely accidental, we performed the same computation, but for each of amino acid scales we randomly permuted the data. In three such experiments the best values of r^2 were: 0.71, 0.63, and 0.65, and only very few descriptors had comparable values. These correlation values are significantly lower than in our calculations with the correct distribution of amino acid attribute values for each amino acid scale, where we furthermore had over a hundred triplets and more than 200 000 descriptors with $r^2 \geq 0.71$.

Figure 2 illustrates that angular separation of sequence moments for the D-descriptor choice is quite different for a highly selective (PGLa) and a mediocre peptide antibiotic (pseudin 2). Predicted TI values for PGLa and pseudin 2 are 95 and 8, respectively, which is in good agreement with measured TI values (105 for PGLa, while in our hands C-terminal amidated pseudin 2 had measured TI = 6). Details for the D-descriptor calculation in the case of the PGLa and measured TI values with associated sequences and references can be found in the Supporting Information.

Using our method to optimize the selectivity of known frog-type helical AMPs by introducing point mutations is also feasible, due to the sensitivity of the D-descriptor, so that the corresponding predicted TI can exhibit a large increase for some point mutations. For example, the point mutation F5W in magainin 2[23] increases its measured TI from 20 to 125, and as it causes an increased separation of sequence moments that leads to a more negative value for the D-descriptor, this results in a considerable increase in the predicted TI (from 19 to 93), so that the experimental and predicted values are quite close in both cases (Figure 3).

As another example, Pál et al. have charted the variation of TI when the charge in pseudin 2 analogues is varied from +2 to +9.[26] Our algorithm resulted in a clear correlation between measured and predicted TI within the charge range +2 to +6 (Figure 4). Lower and higher net positive charge decreases both peptide activity and selectivity for peptides with about 20 amino acids.[9] At the *x*-axis of the Figure 4, higher predicted TI is associated with increased net positive charge of the peptides. The E14K substitution is an exception, since it increased pseudin charge from +2 to +4 but did not increase predicted and measured TI. Low predicted TI values are all associated with increased hydrophobicity due to different phenylalanine substitutions without net charge change from the pseudin 2 value of +2.

We have also tested this capacity in our laboratory by synthesizing and measuring the antimicrobial and hemolytic activity of point-mutated analogs of pseudin 2: PSEU 2(F9A), and ascaphin 1: ASC 1(F2I). The rationale for choosing these point mutations was the observed large TI increase in magainin 2 (Figure 3) after replacing the phenylalanine residue closest to the N-terminus with a less hydrophobic residue according to the CCS scale.[20] Our algorithm predicted a TI increase from 7 to 89 for the PSEU 2 mutation, while the measured increase was from 6 to >30. It predicted an increase from 50 to 79 for the ASC 1 mutation against a measured increase from 40 to >60 (Table 4). In both cases, the trend is correct, further confirming the usefulness of our method.

3.2. The Implementation of the AMP-Designer Algorithm and Associated Restrictions. Although useful in predicting increases in TI, the D-descriptor has its limitations, and could not be the sole foundation of our second goal for rational computational design of new peptide antibiotics with high TI. It predicts TI in the range 5 <TI< 95, while experimental TI values range from less than 0.5 for hemolytics and toxins to more than 100 for the better peptide antibiotics. It is however not intended to provide an accurate TI prediction but to distinguish frog-type peptides as belonging to high or mediocre TI range.

AMP design must take into account that the designed peptides should have an appropriate net positive charge, hydrophobic moment, and hydrophobicity, in addition to a high predicted TI. Even so, the number of possible peptides with about 20 amino acid residues (enough to span a membrane in helical conformation) is potentially extremely large. Therefore, we used a recursive algorithm and imposed various rational statistical and physicochemical restrictions based on the structural and chemical characteristics of the best peptide antibiotics from our AMPad database (SBP peptides, see Methods and Supporting Information).

The set of applied restrictions for our lead peptides containing 23 amino acid residues were thus:

(1) a predicted TI > 85 (close to the maximal possible TI = 95 in the D-descriptor model);

(2) a net positive charge = 4 or 5 (as frequently observed in natural helical AMPs[8]);

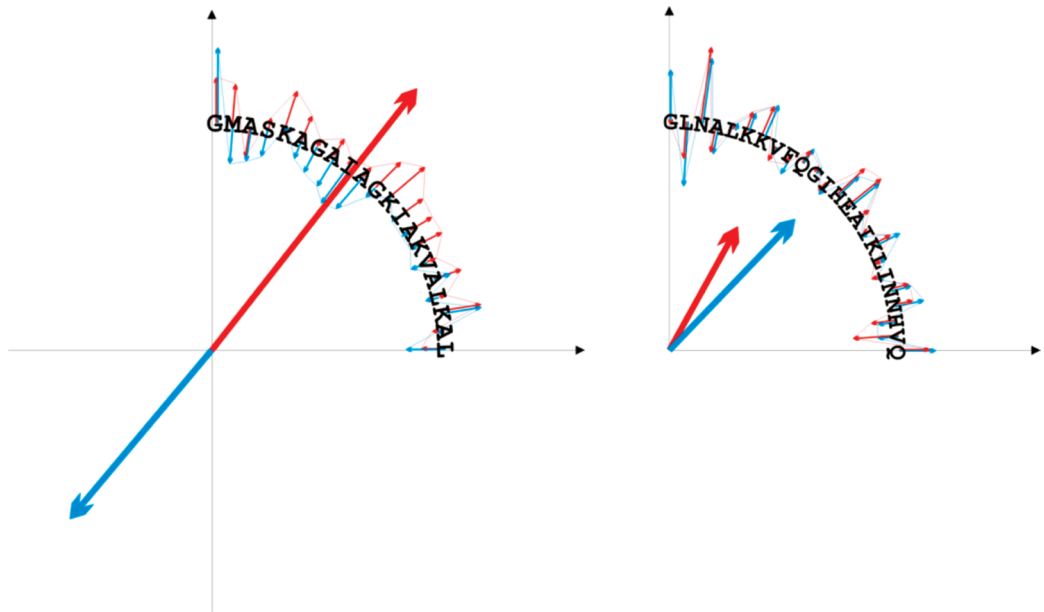

Figure 2. Sequence moments for PGLa (left panel) and pseudin 2 (right panel). For each residue, small blue arrows are calculated by using Guy's amino acid index scale,[25] while small red arrows are determined by using Janin's amino acid index scale.[24] The D-descriptor is the cosine of the angle between vector sums (large arrows) of sequence moments for individual residues from a chosen peptide.

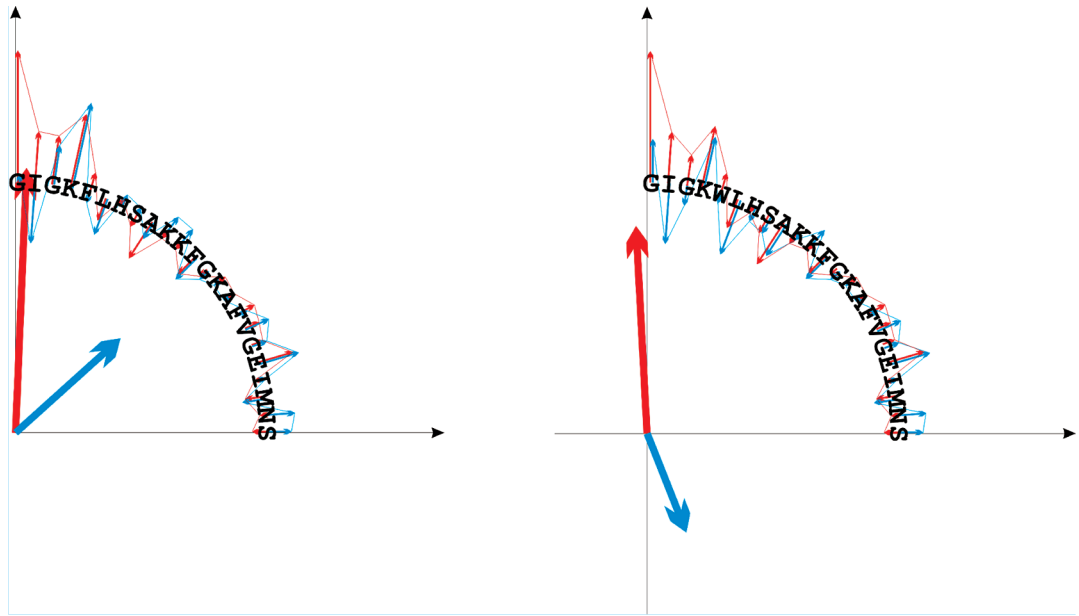

Figure 3. A (left panel). Sequence moments for the peptide antibiotic magainin 2. B (right panel). Point mutation (F5W) in magainin 2 greatly increases its measured therapeutic index.[23] The increased separation of sequence moments after that point mutation leads to a more negative value for the D-descriptor and to a considerable increase in predicted TI (from TI = 18.7 to TI = 92.6).

(3) a mean hydrophobicity (using the CCS scale[20]) = 0.0 to −1.2 (as most often found for helical AMPs[8]). In fact, out of 26 SBP peptides (see Methods) only four are outside that range;

(4) nonpolar residues (A, L, M, V, I, F, W) are separated from polar residues (E, D, Q, N, G, K, R) in a helical wheel projection (limiting design to helix-forming amphipathic peptides that have glycine residues in the polar helix face);

(5) no more than two identical neighboring amino acid residues are allowed (to avoid trivial cases with long segments of identical residues that do not occur in natural AMPs);

(6) the three C-terminal amino acid residues should be the C-terminal residues in at least one good natural peptide antibiotic (together with conditions 7 and 8 this condition limits the design freedom at both peptide termini);

(7) the first amino acid residue should be *G*, because most peptides in SBP data set start with glycine;

(8) the second amino acid residue should be either *L* or *I*, since most of peptides starting with *G* in the SBP data set have *L* or *I* at the second position;

(9) residues with the highest amino acid selectivity index associated with SBP peptides are *E*, *D*, *Q*, *H* and *G* (Table 1), hence at least nine such residues should be present in a peptide (more than nine residues becomes incompatible with other restrictions and produces a null set of suggested peptides);

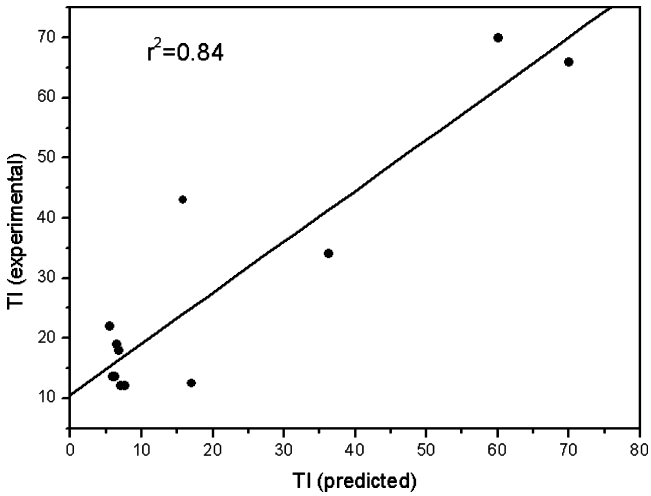

Figure 4. Predicted and measured[26] TI for pseudin 2 analogues with 2 to 6 positive charges. Predicted TI values were derived from the Figure 1 linear one parameter model.

(10) the *motif regularity index* (see Methods for its definition) should be less than 2.5 (ensuring that frequent amino acid motifs from the SBP data set are taken into account).

We do not claim that these 10 conditions/restrictions are the only or optimal choices. They are simply implemented to produce a workable load of suggested peptide antibiotics expected to have high TI and good antibacterial activity against Gram-negative bacterium. On the basis of these restrictions, the algorithm selected the following peptides out of 8×10^{29} possible ones:

GLKGLLGKALKGIGKHIGKAQGC;
GLKGLLGKALGEAKGLLGKHKGC;
GITQGVLKGIGKHVGKALKGIGC;
GIGKHVGKALKGLKGLLKGLGEC;
GIGKHVGKALKGVKGLLKGLGEC;
GIGKHVGKALGELKGLLKGLKGC;
GIGKHVGKALKGLKGLLKGLGES; adepantin 1

Six of these sequences end with a cysteine residue, which may be a vestige from the so-called "Rana-box", a cysteine-bridged stretch present at the C-terminus of many frog-derived helical antimicrobial peptides,[8,27] which is not correlated with increased activity or selectivity.[27] This likely means that restriction 6 has to be modified. Hence, we considered the last peptide, the adepantin 1, as the best candidate for synthesis and characterization. The peptide has a predicted TI = 86, mean hydrophobicity[20] H = −0.97, and a relative amphipathicity[28] of 0.66 (66% that of a perfectly amphipathic peptide). The underlined "small" motifs [A,G,S]XXX[A,G,S] in the primary structure of the adepantin 1: **GI<u>GKHVG</u>KALKGLK<u>GLLKGLGES</u>**, promote helix−helix interactions in the membrane.[29]

Adepantin 1 is unlike any other helical peptide. Its unusual primary structure with seven glycine residues is ≤50% identical to any known helical AMP, being closest to bombinin (BMN_BOMVA, accession P01505, 50% identity). Its name derives from an abbreviation for **A**utomatically **De**signed **P**eptide **Anti**biotic **N**o. **1**.

3.3. Synthesis, Characterization and Testing of Peptide Antibiotics. The suggested peptide was synthesized in good purity (SI Figure S2 and Table S3), together with two known frog peptides from the training database. Ascaphin 1 and pseudin 2 were selected to span the range of corresponding peptide attributes for good and mediocre peptide antibiotics. The observed structural and functional properties of adepantin 1 were favorable for antimicrobial activity (SI Figure S3 and Table S3). For instance, despite the presence of numerous glycine residues, it had an estimated helical content of about 70% in the presence of the helix promoting solvent TFE.

Antimicrobial activity assays confirmed that the adepantin 1 was selectively active against Gram-negative bacteria (Table 5). Its C-terminally amidated form had a MIC value of 2−4 μM against *E. coli*, against which it was designed, and a moderate activity also toward *P. aeruginosa*, whereas its activity was poor against the Gram-positive bacterium *S. aureus*. The designed peptide did not quite reach 50% lysis of red blood cells even at the highest concentration tested, 800 μM (Table 5 and Figure 5). By comparison, under the same conditions pseudin showed a HC_{50} of 50 μM and ascaphin of 80 μM. Using these data to determine experimental TI values, and comparing them with predicted TI, we found values for control peptides of 6 against the predicted 7 for pseudin, 40 against the predicted 54 for ascaphin (Table 5 and SI Table S1). These data, together with those for adepantin 1 with a measured TI value of over 200 against the predicted 86, verify the design method.

4. DISCUSSION

Our first goal was to find a rational computational method for guiding the increase in therapeutic index (TI) of antimicrobial peptides. We choose to search for a one-parameter linear model which would correlate measured and predicted TI values for the compiled data set of frog-derived nonhomologous AMPs. This was achieved with the selection of the D-descriptor by using Pearson's correlation coefficient as the selection criterion which gives the greatest weight to peptides with high TI. To the best of our knowledge, nonhomologous antimicrobial peptides have never been used before in quantitative structure−activity studies of TI. Furthermore, most structure−activity studies of AMPs use multilinear models for predicting peptide activity.[32] Models with only one peptide descriptor were not expected to work, due to a well-known necessity to use several interrelated structural and physicochemical parameters that modulate activity and specificity, such as charge, secondary structure, amphipathicity, and hydrophobicity.[8]

While we have shown that the introduced D-descriptor is a useful tool for estimating peptide selectivity in the high TI range or for estimating the direction of the TI change after point mutations, the D-descriptor has its limitations: (i) similar procedures for finding best descriptors based on peptide lengthwise asymmetry, instead of using peptide mean values for some attributes, can be used outside the domain of frog-derived helical peptides, but the D-descriptor described here is tightly connected with the training database used for its extraction; (ii) selection depends both on the chosen computational procedure and on the quality of available experimental MIC and HC_{50} data (note that for particular bacterial species and peptides, the reported experimental TI can vary from 2-fold to 4-fold);[33] (iii) the D-descriptor predicts maximal TI = 95 also in the case of all-Gly, all-His, and all-Ala peptides, which are not likely to be peptide antibiotics, so that false positives may occur

Table 4. Therapeutic Index Changes after Single Point Mutations As Measured in Our Laboratory

[a]peptide	sequence	[a]MIC (μM) *E. coli*	[a]hemolysis (%) (10 μM)	(100 μM)	(200 μM)	TI[b] (exp)	TI[b] (pred)
PSEU 2	GLNALKKVFQGIHEAIKLINNHVQ-NH_2	8	5	73	83	6	7
PSEU 2 A9	GLNALKKV**A**QGIHEAIKLINNHVQ-NH_2	8−16	3	8	30	>30	90
ASC 1	GFRDVLKGAAKAFVKTVAGHIAN-NH_2	2	6	64	80	40	50
ASC 1 I2	G**I**RDVLKGAAKAFVKTVAGHIAN-NH_2	4	1	25	44	>60	80

[a] See Methods for the description of peptide synthesis, purification, characterization, and activity testing. For PSEU 2 A9 and ASC 1 I2 the predicted/measured MW were 2608.1/2608.0 and 2335.8/2335.6, respectively. [b] TI (exp) determinations in the case of pseudin-2 amide (PSEU 2) and ascaphin-1 amide (ASC 1) were extrapolated from data presented in Figure 5 and Table 5, respectively. Predicted TI values were calculated from the Figure 1 linear one parameter model.

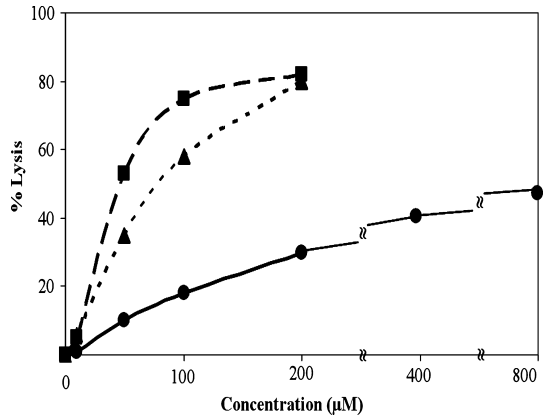

Figure 5. Hemolytic activity of adepantin 1 (-●-), ASC 1 (-▲-) and PSEU 2 (-■-). The hemolytic activity of the peptides was calculated as percent lysis of human erythrocytes (0.5% suspension), with increasing peptide concentrations (from 10 − 800 μM) in PBS, after incubation at 37°*C* for 30 min, and are the mean of two experiments performed in triplicate. Total lysis (100% hemolysis) was determined by addition of 0.2% Triton X-100.

for unusual sequences. A maximal TI is predicted for such peptides because for the chosen pair of amino acid scales all sequence environments are then of opposite sign (see Supporting Information Methods, Table A, normalized values). This follows from different classification of Gly, Ala, and His by the Janin and Guy scales. For instance, Gly is considered as a buried residue by the Janin's scale, while Gly partitions as if it is polar according to Guy's scale.

Our main research goal of designing and testing novel peptide antibiotics thus required a more complex and articulated combined use of the D-descriptor together with other computed, statistical, and experimental data pertaining to the peptides in the database. While we maintained the generality of our approach, in the sense that it does not depend on homology to known structures as a main bottleneck for designing new structures, we had to use additional concepts and data mining procedures such as the definition and calculation of an amino acid selectivity index and motif regularity index. An interesting result of this simple statistical analysis was the importance of negatively charged residues E and D, in addition to glycine for enhancing selectivity (Table 1). The presence of negatively charged Asp/Glu residues is in fact observed in more than 70% of native amphipathic cationic peptides and has been suggested to influence peptide structuring and dimerization,[10] possibly through electrostatic bonding of parallel or antiparallel peptide dimers in the membrane environment. Glutamic acid can also inhibit hemolysis perhaps by hydrogen bonding to cholesterol in eukaryotic cell membranes.[34]

The strength of our Designer algorithm depends on multiple flexible screening criteria. For instance, the choice of polar and nonpolar residues, which must be separated in a helical wheel projection (condition 4) used by the AMP-Designer, leaves the possibility for some polar residues to appear at the hydrophobic helix side and for some nonpolar residues to appear at the hydrophilic helix side. This is often the case with natural peptide antibiotics and would tend to decrease their hydrophobic moment (mean relative hydrophobic moment, see Methods), but TI can still be high, as is the case for the PGLa. The number of suggested peptides can be additionally reduced either by including additional polar residues (such as His) or by imposing the range of allowed mean hydrophobic moments.

The range of allowed mean hydrophobicities,[20] H, is also a flexible choice. For instance, the H range from −0.35 to −1.35 would leave only two out of the 26 best peptide antibiotics (our SBP data set) outside that range and the AMP-Designer output would then be 10 peptides including adepantin 1 and its point mutation adepantin 1 L13V. Another flexible option is the number of amino acids with high amino acid selectivity index: *E*, *D*, *Q*, *H,* and *G* to use (condition 9), which depends on the peptide length. Together with a required high net positive charge (condition 2) and low ceiling for the motif regularity index (condition 10), this ensures frequent appearance of glycine in designed peptides.

The experimentally confirmed, high selectivity of the designed peptide adepantin 1 may in part derive from its unusual primary structure, with 7 glycine residues out of 23. On the basis of a statistical analysis of natural helical AMPs,[8] glycine residues have an increased frequency at position 7 or 14. It was shown that placing a Gly residue at position 7 in an artificial 19-residue helical AMP,[28] somewhat increased selectivity, while placing Gly in position 14, or both 7 and 14 considerably reduced antimicrobial potency; so there seem to be positional requirements for its use in increasing selectivity. Using the D-descriptor in connection with the AMP-Designer algorithm, it was possible to simultaneously introduce 6 internal glycine residues without markedly reducing activity toward Gram-negative microorganisms, while dramatically increasing selectivity. The distribution of glycines at the polar helix face ensures that glycine-containing "small" motifs appear within most peptides suggested by the algorithm. From our results, it follows that both the D-predictor and performed statistical analysis of amino acid motifs have singled out a glycine distribution pattern as an important factor for increasing peptide selectivity. However, additional experiments are needed to explore the hypothesis that selectivity and target specificity acquisi-

Table 5. Biological Activities and Therapeutic Index Determination

peptide[a]	MIC[b] (μM)			HC^c_{50}	TI[d]	% bacterial growth inhibition[e] (*E. coli*)		
	E. coli ATCC	*P. aeruginosa*	*S. aureus*	(μM)		2 μM	4 μM	8 μM
Adepantin-1	2−4	16	>128	>800	200	10	75	100
ASC 1	2	16	32	80	40	65	100	100
PSEU 2	8	>64	32	50	6	0	25	100

[a] adepantin 1 = GIGKHVGKALKGLKGLLKGLGES-NH_2, ASC 1 = ascaphin 1,[30] PSEU 2 = pseudin 2.[31] [b] Carried out in 100% (v/v) MH broth with 10^5 cfu/mL bacteria in the logarithmic phase - mean of at least three experiments performed in duplicate. [c] Calculated from % lysis vs concentration plots, see Figure 5- the mean of two experiments performed in triplicate. [d] TI, therapeutic index [e] Calculated for the indicated concentration of peptide, at the 4th hour, from bacterial growth kinetics curves, monitoring OD at 600 nm (100% v/v MH broth, 10^6 cfu/mL logarithmic phase bacteria) - mean of two experiments performed in triplicate.

tion is in part due to promotion of helix−helix interactions in membrane milieu by "small motifs".[29]

The flexibility of the adepantin-1 peptide and its Gly distribution pattern may have the dual role of facilitating unstructured monomers to pass easily through cell wall components[10] and of increasing the probability of transient monomer aggregation and pore formation at the level of the cytoplasmic membrane. Being different from any known AMP, the adepantin 1 peptide can be considered as the first member of a new glycine-rich class of helical AMP lead compounds for potential medical applications against Gram-negative bacteria.

Extensive characterization of model helical peptides carried out in the past[8,16,28] has shown that in general, Gram-negative bacteria are more susceptible than Gram-positive ones, and their membranes are also permeabilized more slowly. Our algorithm has selected for features which maintain high selectivity and a potent activity against Gram-negatives (Table 5), as it was trained on activity data for this type of microorganism, but is lacking necessary features for Gram-positive ones. Ongoing work is now aimed at setting up a similar database for activity against the latter type of bacteria, so it may be possible in the future to design peptides with a wider spectrum, while maintaining a low cytotoxicity.

The importance of using a recursive algorithm with in-built simple restrictions is seen in the enormous reduction of all possible 23 amino acid long peptides (from 8×10^{29} to 106 designed peptides) when conditions 2−9 were imposed. Adding condition 10 (the motif regularity index requirement MR < 2.5) led to a total of 37 peptides. Finally, a condition 1 led to a final manageable choice of only seven lead peptides. Therefore, the D-descriptor was useful in eliminating 81% of designed peptides produced by conditions 2−10. We expect the AMP-Designer to be helpful in exploring a broad range of helical frog-type peptide antibiotics, which biological evolution has left unexplored or which may not be explored due to the increased rarity or extinction of amphibian species.[12]

5. SOFTWARE TOOLS AVAILABILITY

The TI estimator based on the D-descriptor is available online at http://split.pmfst.hr/split/dserv1. The AMP-Designer and other developed C++ software tools are available at http://sites.google.com/site/adepantin1.

ACKNOWLEDGMENT

The authors are grateful to Viktor Bojović for the on-line availability of dserv1 and SPLIT software, to Bono Lučić who provided us with some amino acid scales, and to Dražen Petrov and Bojan Žagrović for fruitful discussions. *Funding*: The work was supported in part by Croatian Ministry of Science, Education and Sport (Grant Nos. 177-1770495-0476 (D.J.), 177-0000000-0884 (D.V.), and 037-0000000-2779 (D.V.)) and by a Friuli Venezia Giulia LR26 grant for the R^3A^2 network project.

Supporting Information Available: Tables S1 and S2 with additional references (AMPad database), Figure S1 (test set determination coefficient), Methods (calculating the sequence moment for PGLa), Figure S2 (Analytical HPLC and ESI−MS of adepantin 1), Figure S3 (CD spectra of peptides) and Table S3 (physicochemical properties of peptides).

This information is available free of charge via the Internet at http://pubs.acs.org/.

REFERENCES AND NOTES

(1) Brogden, K. A. Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria. *Nat. Rev. Microbiol.* **2005**, *3*, 238–250.
(2) Marr, A. K.; Gooderham, W. J.; Hancock, R. E. W. Antibacterial peptides for therapeutic use: obstacles and realistic outlook. *Curr. Opin. Pharmacol.* **2006**, *6*, 468–472.
(3) Zhang, L.; Falla, T. J. Host defense peptides for use as potential therapeutics. *Curr. Opin. Investig. Drugs* **2009**, *10*, 164–171.
(4) Bowie, J. H.; Jackway, R. J.; Separovic, F.; Carver, J. A.; Tyler, M. J. Host defence peptides from the secretion of the skin glands of frogs: membrane active peptides from the genera Litoria, Uperoleia and Crinia. In *Bioactive Peptides*, Howl, J., Jones, S. , Eds.; Taylor and Francis Inc.: Abingdon, United Kingdom, 2009; pp 333−355.
(5) Zasloff, M. Antimicrobial peptides of multicellular origin. *Nature* **2002**, *415*, 389–395.
(6) Westerhoff, H. V.; Juretić, D.; Hendler, R. W.; Zasloff, M. Magainins and the disruption of membrane-linked free-energy transduction. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 6597–6601.
(7) Yeaman, M. R.; Yount, N. Y. Unifying themes in host defence effector polypeptides. *Nat. Rev. Microbiol.* **2007**, *5*, 727–740.
(8) Tossi, A.; Sandri, L.; Giangaspero, A. Amphipathic, α-helical antimicrobial peptides. *Biopolymers* **2000**, *55*, 4–30.
(9) Dathe, M.; Nikolenko, H.; Meyer, J.; Beyermann, M.; Bienert, M. Optimization of the antimicrobial activity of magainin peptides by modification of charge. *FEBS Lett.* **2001**, *501*, 146–150.
(10) Jiang, Z.; Vasil, A. I.; Hale, J.; Hancock, R. E. W.; Vasil, M. L.; Hodges, R. S. Effects of net charge and the number of positively charged residues on the biological activity of amphipathic α-helical cationic antimicrobial peptides. *Biopolymers (Pept. Sci.)* **2008**, *90*, 369–383.
(11) Vanhoye, D.; Bruston, F.; Nicolas, P.; Amiche, M. Antimicrobial peptides from hylid and ranin frogs originated from a 150-million-year-old ancestral precursor with a conserved signal peptide but a hypermutable antimicrobial domain. *Eur. J. Biochem.* **2003**, *270*, 2068–2081.
(12) Apponyi, M. A.; Pukala, T. L.; Brinkworth, C. S.; Maselli, V. M.; Bowie, J. H.; Tyler, M. J.; Booker, G. W.; Wallace, J. C.; Carver, J. A.; Separovic, F.; Doyle, J.; Llewellyn, L. E. Host-defence peptides of Australian anurans: structure, mechanism of action and evolutionary significance. *Peptides* **2004**, *25*, 1035–1054.
(13) Mahamoud, A.; Chevalier, J.; Alibert-Franco, S.; Kern, W. V.; Pagès, J.-M. Antibiotic efflux pumps in Gram-negative bacteria: the inhibitor response strategy. *J. Antimicrob. Chemother.* **2007**, *59*, 1223–1229.

(14) Kyte, J.; Doolittle, R. F. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* **1982**, *157*, 105–132.
(15) Eisenberg, D.; Weiss, R. M.; Terwilliger, C. T.; Wilcox, W. Hydrophobic moments and protein structure. *Faraday Symp. Chem. Soc.* **1982**, *17*, 109–120.
(16) Tossi, A.; Tarantino, C.; Romeo, D. Design of synthetic antimicrobial peptides based on sequence analogy and amphipathicity. *Eur. J. Biochem.* **1997**, *250*, 549–558.
(17) Juretić, D.; Lučić, B. ; Zucić, D. ; Trinajstić, N. Protein transmembrane structure: recognition and prediction by using hydrophobicity scales through preference functions. In *Theoretical and Computational Chemistry*; Párkányi, C. , Ed.; Elsevier: Amsterdam, The Netherlands, 1998; Vol. 5, pp 405−445.
(18) Gromiha, M. M. A statistical model for predicting protein folding rates from amino acid sequence with structural class information. *J. Chem. Inf. Model.* **2005**, *45*, 494–501.
(19) Yuan, Z.; Zhang, F.; Davis, M. J.; Bodén, M.; Teasdale, R. D. Predicting the solvent accessibility of transmembrane residues from protein sequence. *J. Proteome Res.* **2006**, *5*, 1063–1070.
(20) Tossi, A.; Sandri, L.; Giangaspero, A. New consensus hydrophobicity scale extended to non-proteinogenic amino acids. In *Peptides 2002, Proceedings of the 27th European Peptide Symposium, Sorrento, August 31st - September 6th, 2002*; Benedetti, E., Pedone, C. Eds., Edizioni Ziino: Napoli, Italy, 2002; pp 416−417.
(21) Waddell, W. J. A simple ultraviolet spectrophotometric method for the determination protein concentration. *J. Lab. Clin. Med.* **1956**, *48*, 311–314.
(22) Chen, Y. H.; Yang, J. T.; Chau, K. H. Determination of the helix and beta form of proteins in aqueous solution by circular dichroism. *Biochemistry* **1974**, *13*, 3350–3359.
(23) Tachi, T.; Epand, R. F.; Epand, R. M.; Matsuzaki, K. Position-dependent hydrophobicity of the antimicrobial magainin peptide affects the mode of peptide−lipid interactions and selective toxicity. *Biochemistry* **2002**, *41*, 10723–10731.
(24) Janin, J. Delta*G*-transfer from buried interior to solvent accessible surface. *Nature* **1979**, *277*, 491–492.
(25) Guy, H. R. Amino acid side-chain partition energies and distribution of residues in soluble proteins. *Biophys. J.* **1985**, *47*, 61–70.
(26) Pál, T.; Sonnevend, Á.; Galadari, S.; Conlon, J. M. Design of potent, non-toxic antimicrobial agents based upon the structure of the frog skin peptide, pseudin-2. *Regul. Pept.* **2005**, *129*, 85–91.
(27) Simmaco, M.; Mignogna, G.; Barra, D. Antimicrobial peptides from amphibian skin: what do they tell us. *Biopolymers.* **1998**, *47*, 435–450.
(28) Zelezetsky, I.; Pacor, S.; Pag, U.; Papo, N.; Shai, Y.; Sahl, H.-G.; Tossi, A. Controlled alteration of the shape and conformational stability of α-helical cell-lytic peptides: effect on mode of action and cell specificity. *Biochem. J.* **2005**, *390*, 177–188.
(29) Senes, A.; Gerstein, M.; Engelman, D. M. Statistical analysis of amino acid patterns in transmembrane helices: The GxxxG motif occurs frequently and in association with β-branched residues at neighboring positions. *J. Mol. Biol.* **2000**, *296*, 921–936.
(30) Conlon, J. M.; Sonnevend, A.; Davidson, C.; Smith, D. D.; Nielsen, P. F. The ascaphins: a family of antimicrobial peptides from the skin secretions of the most primitive extant frog, *Ascaphus truei*. *Biochem. Biophys. Res. Commun.* **2004**, *320*, 170–175.
(31) Olson, L.; Soto, A.; Knoop, F. C.; Conlon, J. M. Pseudin-2: an antimicrobial peptide with low hemolytic activity from the skin of the paradoxical frog. *Biochem. Biophys. Res. Commun.* **2001**, *288*, 1001–1005.
(32) Langham, A. A.; Khandelia, H.; Schuster, B.; Waring, A. J.; Lehrer, R. I.; Kaznessis, Y. N. Correlation between simulated physicochemical properties and hemolycity of protegrin-like antimicrobial peptides: Predicting experimental toxicity. *Peptides* **2008**, *29*, 1085–1093.
(33) Chen, Y.; Vasil, A. I.; Rehaume, L.; Mant, C. T.; Burns, J. L.; Vasil, M. L.; Hancock, R. E. W.; Hodges, R. S. Comparison of biophysical and biologic properties of α-helical enantiomeric antimicrobial peptides. *Chem. Biol. Drug. Des.* **2006**, *67*, 162–173.
(34) Tytler, E. M.; Anantharamaiah, G. M.; Walker, D. E.; Mishra, V. K.; Palgunachari, M. N.; Segrest, J. P. Molecular basis for prokaryotic specificity of magainin-induced lysis. *Biochemistry* **1995**, *34*, 4393–4401.

CI900327A