Supporting Information

Ab Initio Design of Potent Anti-MRSA Peptides based on Database Filtering Technology

Biswajit Mishra and Guangshun Wang*

This supporting information contains the following:

Part I: Experimental Section;

Part II: Supporting figures;

Part III: Supporting tables.

Part I: Experimental Section

The antimicrobial peptide database (http://aps.unmc.edu/AP/main.html**)**. To facilitate search, prediction, and design of AMPs, we established the APD in 2003.[2a] The first version of our database contained 525 entries. The second version reported 1280 AMPs.[2b] As of February 2012, the updated APD registered 1954 AMPs. Natural peptides (<100 amino acids) will be collected into our database if their amino acid sequences and biological activity have been determined. The APD enabled a systematic search of peptide information, including peptide name, amino acid sequence, length, charge, hydrophobic content, structure, and biological activity. Users can search antibacterial, antifungal, antiviral (anti-HIV), anti-parasital, anticancer, spermicidal, insecticidal, and chemotactic data. These structure and activity filters in the APD provide a solid basis for this *ab initio* design.

Peptides and hydrophobicity measurements. The peptides were chemically synthesized and purified (>95%) by GeneMed Synthesis, TX. One of the methods to estimate peptide hydrophobicity is to measure its retention time by HPLC.[8c] This was done on a Waters HPLC system equipped with an analytical reverse-phase Vydac C18 column (250 × 4.6 mm). The peptide was eluted by a linear gradient of acetonitrile (containing 0.1% TFA) from 5% to 95% at a flow rate of 1 ml/min. The peptide peak was detected by UV at 220 nm.

Organisms and antibacterial activity assays. The bacterial strains used in the study are *Bacillus subtilis* 168*,* methicillin-resistant *Staphylococcus aureus* USA300 LAC (Gram-positive), *Escherichia coli* K12, and *Pseudomonas aeruginosa* PAO1 (Gram negative). The antibacterial activities of the peptides were determined using the standard broth microdilution method.[9] In brief, the bacterial culture (90 μL) was treated with a series of peptide solutions (10 μL), each with two-fold dilution. After overnight incubation at 37^{o}C, the microplate was read on a ChroMate 4300 Microplate Reader at 630 nm (GMI, Ramsey, MN). The MIC was defined as the lowest peptide concentration that fully inhibited bacterial growth.

Measurements of the minimal hemolytic concentration. Peptide hemolysis was assayed using an established protocol.[3b] HL_{50} was the peptide concentration that caused 50% lysis of human erythrocytes. The cell selectivity (CS) indexes were calculated as the ratio of HL_{50} to the MIC of the peptide against *S. aureus* USA300.

Transmission electron microscopy. *S. aureus* USA300 were grown in LB medium at 37^{0}C to the mid-logarithmic phase. Five mL of the culture was treated with the peptide DFTamP1 at 2× or 8× MIC for 1 h. After washing with PBS, cells were fixed with a solution of 2% paraformaldehyde and 2% glutaraldehyde in 0.1 M PBS. The untreated control sample was processed in the same way. Ten μL of the bacterial suspension was adsorbed onto 200 mesh copper grids coated with formvar and silicon monoxide for 2

min, dried with the help of a blotting paper and then negatively stained by two drops of Nanovan® (Nanoprobes). Samples were observed in an FEI Technai G2 TEM operated at 80 kV accelerating voltage at the University of Nebraska Medical Center (UNMC).

Flow cytometry. Membrane permeation measurements were performed as described[10] using BacLight bacterial membrane potential kit (Molecular probes, Invitrogen) and a DNA-binding dye TO-PRO3 according to the manufacturer's guidelines on a FACSAria Flow Cytometer (BD Biosciences, San Jose, California).

NMR spectroscopy. The NMR sample (0.3 mL) contained 2 mM peptide, 80 mM deuterated sodium dodecylsulfate (SDS), and 10% D_2O as the field-locking signal. The pH of the NMR sample was adjusted to 5.4. Spectra were recorded on a 600-MHz Varian INOVA NMR spectrometer equipped with a triple-resonance cryogenic probe with a z-axis gradient capability. A set of 2D spectra, including NOESY, TOCSY, DQF-COSY, and natural abundance HSQC spectra, was recorded with the ^{1}H, ^{15}N, and ^{13}C carriers set at 4.77, 118.27, and 36.37 ppm, respectively. Thirty increments (128 scans) and 80 increments (256 scans) were collected for the ^{15}N (spectral width 2,200 Hz) and aliphatic ^{13}C (spectral width 12,000 Hz) dimensions, respectively. Chemical shifts were referenced as recommended.[15] Data were processed on a Silicon Graphics Octane workstation using NMRPipe[16] and analyzed by PIPP.[17]

Structure calculations. For structural calculations, major NMR restraints were derived from 2D NOESY spectra[11] and converted to distance restraints 1.8-2.8, 1.8-3.8, 1.8-5.0,

and 1.8-6.0 Å corresponding to strong, medium, weak, and very weak types of NOE peaks, respectively. Backbone anlges were predicted based on $^{1}H\alpha$, ^{15}N, $^{13}C\alpha$, and $^{13}C\beta$ chemical shifts as described.[9] An extended covalent structure was used as starting coordinates. In total, 100 structures were calculated using the simulated annealing protocol in the Xplor-NIH program.[18] Twenty structures were accepted and analyzed as described.[9]

Additional references

(15) Markley, J. L.; Bax, A.; Arata, Y.; Hilbers, C. W.; Kaptein, R.; Sykes, B. D.; Wright, P. E.; Wüthrich, K. *J. Biomol. NMR* **1998**, *12*, 1-23.

(16) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277-293.

(17) Garrett, D. S.; Gronenborn, A. M.; Clore, G. M. *J. Magn. Reson*. **1991**, *95*, 214-220.

(18) Schwieters, C. D.; Kuszewski, J. J.; Tjandra, N.; Clore, G. M. *J. Magn. Reson*. **2003**, *160*, 65–73.

Part II: Supporting Figures

Figure S1: **Residues G, K, and S are shared frequently occurring residues in natural AMPs.** Plotted are sequence profiles for 44 lantibiotics from bacteria (A), 131 cyclotides from plants (B), and 75 temporins from amphibians (C). These peptide groups were chosen because of structural uniformity. Lantibiotics tend to have multiple loops by forming thioethers, cyclotides usually adopt β-sheet, which is further stabilized by disulfide bonds, and temporins are normally helical.[1] Although these peptides adopt different structural scaffolds and kill bacteria in different mechanisms, *G, S, and K are abundant in these three peptide groups from different life kingdoms, indicative of their universal importance in natural AMPs* (Data obtained from the APD[2]).

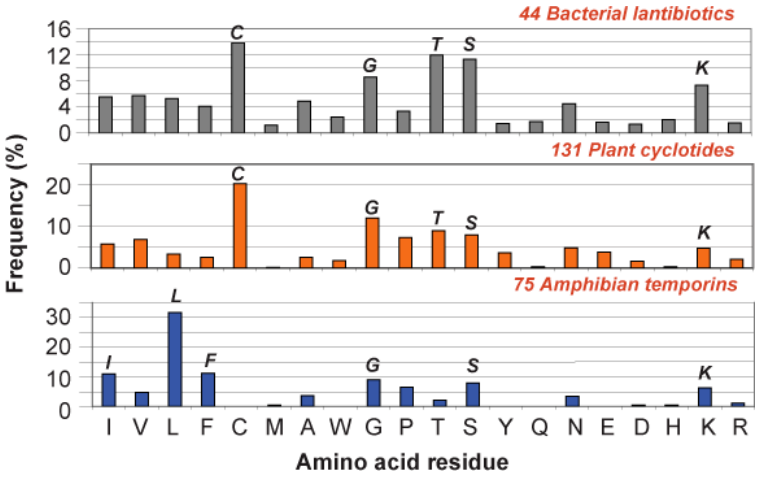


***Figure S2*: Rapid killing of *S. aureus* USA300 by DFTamP1 is also supported by traditional killing kinetics.** The experiments were set up in the same manner as the antibacterial assays above with the following additions. Bacteria (10^5 colony forming units of *S. aureus* USA300) were treated at a peptide level of 2 × MIC). Aliquots were taken at 15, 30, 50, 90, and 120 min, plated after 100-fold dilution, and the colonies were counted after overnight incubation at 37°C.[10] *This experiment indicates that DFTamP1 was able to rapidly kill S. aureus in 60 min.*

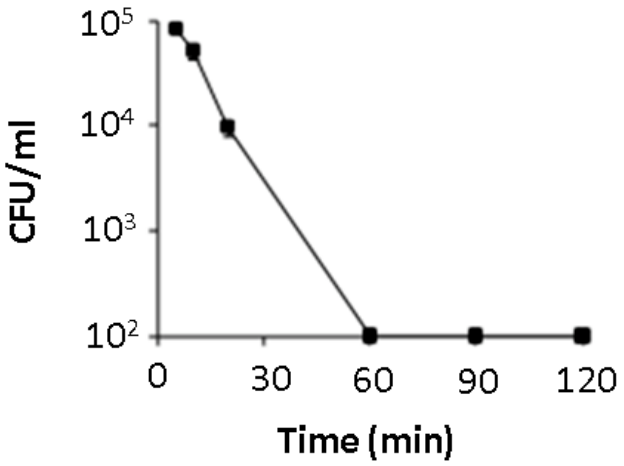


Figure S3: **Real time killing of *S. aureus* USA300 by DFTamP1 (A) and its mutant DFTamP1-pv (B)**. In the plots, bacterial growth or killing curves treated with 0, 1, 2, or 8× MICs of DFTamP1 were indicated in black, blue, violet, and red, respectively. For clarity, the curve at 4× MIC was not displayed since it is similar to that of 8× MIC. The microplate was incubated in a shaker at 100 rpm and 37°C. *This figure indicates that the designed peptide DFTamP1 is able to lyse and inhibit MRSA USA300 at 8× MICs, whereas DFTamP1-pv was unable to even at four-fold the corresponding concentrations of DFTamP1.*

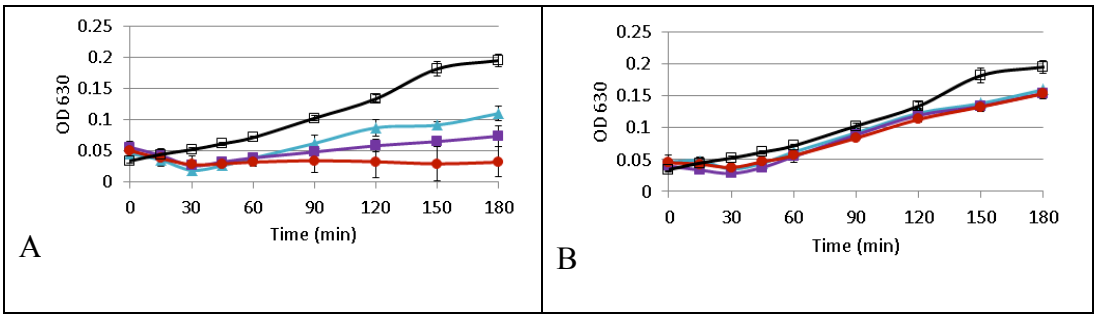
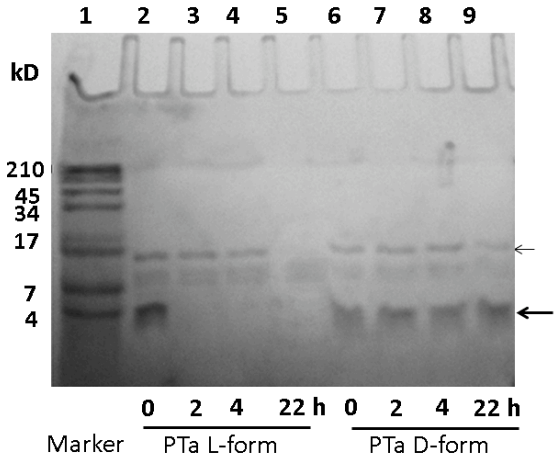


***Figure S4*: Stability comparison of the L- and D-forms of temporin-PTa in the presence of chymotrypsin**. A solution (100 μl) of the peptide and 12.5 μM chymotrypsin (peptide/protein molar ratio, 40:1) in 10 mM PBS buffer (pH 8) was incubated at 37°C. Aliquots (10 μl) of the reaction solutions were taken at 0, 2, 4, and 22 h and immediately mixed with 20 μl of 2× SDS loading buffer to stop the reaction. For the gel analysis, 10 μl of each sample was loaded to the well of a 5% stacking/18% resolving tricine gel. The samples in the lanes are: 1, protein marker; 2, temporin-PTa L-form at 0 h; 3, 2 h; 4, 4 h; 5, 22h; 6, temporin-PTa D-form at 0 h; 7, 2 h; 8, 4 h; 9, 22 h. In the gel, the bands for the peptide and chymotrypsin were indicated with large and small arrows, respectively. Note that the peptide bands were not detected in lanes 3-5 due to protease cleavage but remained essentially constant in lanes 7-9. *This figure indicates that in the presence of chymotrypsin* (small arrow)*, the L-form of temporin-PTa was degraded within 2 h (lanes 3-5), while the D-form was stable for at least 22 h* (lanes 7-9, big arrow).



Part III: Supporting Tables

Table S1. **Database motifs and their application in peptide design**[a]

Motifs	Maximum hits in the APD
I. The leucine motifs	
LL	449
LLL	11
LLLL	0
II. The 3-residue motifs containing "LL"	
GLL	147
LLG	75
SLL	58
LLS	90
KLL	69
LLK	58
III. The 4-residue motifs	
GKLL	41
KGLL	6
GSLL	18
SGLL	19
GGLL	25
KSLL	7
SKLL	1
SSLL	1
IV. The joining motifs	
LLG	75
LLS	90
LSL	37
LGL	31
LGK	89

[a]Data were obtained from the antimicrobial peptide database (*http://aps.unmc.edu/AP*) on February 2012 (total AMP entries: 1954). These hits changed little during our peptide design. Motifs incorporated into DFTamP1 were underlined.

Table S2. **Structural statistics of the designed peptide DFTamP1 in complex with sodium dodecylsulfate micelles**

Structural restraints	DFTamP1
no. of NOE restraints (total)	99
intra-residue	38
sequential	38
short-range	23
no. of backbone angles[a]	20
Structure and quality	
helical region	3-12
backbone rmsd (Å)[b]	0.36
NOE-derived distance violations (Å)	0.3
dihedral angle violations (deg)	5
Ramachandran plot[c]	
residues in the most favored region (%)	93.75
residues in the additionally allowed region (%)	3.75
residues in the generously allowed region (%)	2.50

[a]Predicted by the updated version of the TALOS program.[19]

[b]Calculated by MOLMOL[20] when residues 2-20 of the accepted ensemble of 20 structures were superimposed.

[c]Calculated by Procheck.[21]

Since ~94% of residues are in the most favored region, the structures have been determined to high quality.

Additional references

(19) Cornilescu, G.; Delaglio, F.; Bax, A. *J. Biomol. NMR* **1999**, *13*, 289-302.

(20) Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graphics* **1996**, *14*, 51-55.

(21) Laskowski, R. A.; Rullmannn, J. A.; MacArthur, M. W.; Kaptein, R.; Thornton, J. M. *J. Biomol. NMR* **1996**, *8***,** 477-486.

Table S3. **Sequence alignment, physical parameters and activity of the major types of Temporins**

APD ID	Peptide name	Sequence alignment[a]	Physical parameters[b]					Activity G+/G-
			AA	P	L	Pho	R+K	
1534	Temporin-SHf	----FF--FLSRIF	8	0	1	75	1	Y/Y
99	Temporin L	-FVQWFSKFLGRIL	13	0	2	61	2	Y/Y
101	Temporin G	-FFPVIGRILNGIL	13	1	2	61	1	Y/Y
110	Temporin-1Lb	NFLGTLINLAKKIM	14	0	3	57	2	Y/N
94	Temporin A	-FLPLIGRVLSGIL	13	1	4	61	1	Y/N
98	Temporin F	-FLPLIGKVLSGIL	13	1	4	61	1	Y/Y
97	Temporin E	-VLPIIGNLLNSLL	13	1	5	61	0	Y/N
859	Temporin H	-LSP --- NLLKSLL	10	1	5	50	1	Y/N
858	Temporin D	-LLPIVGNLLNSLL	13	1	6	61	0	n.f.[b]
95	Temporin B	-LLPIVGNLLKSLL	13	1	6	61	1	Y/N
100	Temporin K	-LLP --- NLLKSLL	10	1	6	60	1	Y/N
96	Temporin C	-LLPILGNLLNGLL	13	1	7	61	0	Y/N
N.A.[b]	DFTamP1	GLLSLL-SLLGKLL	13	0	8	61	1	Y/N

[a]Leucines (L) are shaded in yellow, prolines (P) in red, and positively charged residues in blue. Alignment gaps (i.e. missing residues) are represented with dashes (-).

[b]Peptide sequences, their physical parameters and activity data were obtained from the APD.[2] Additional information, including the original source, can be obtained by searching each entry in the database as well. AA: number of amino acid residues; P: number of prolines; L; number of leucines; Pho: percent hydrophobic residues; R+K, the sum of R and K residues; G+, activity against Gram-positive bacteria; G-, activity against Gram-negative bacteria; Y, Yes; N, No; n.f.: not found; N.A., not available.