

# Use of Artificial Intelligence in the Design of Small Peptide Antibiotics Effective against a Broad Spectrum of Highly Antibiotic-Resistant Superbugs

Artem Cherkasov<sup>\*,†</sup>, Kai Hilpert<sup>†,‡</sup>, Håvard Jenssen<sup>†</sup>, Christopher D. Fjell<sup>‡</sup>, Matt Waldbrook<sup>†</sup>, Sarah C. Mullaly<sup>†</sup>, Rudolf Volkmer<sup>§</sup>, and Robert E.W. Hancock<sup>†,\*</sup>

<sup>†</sup>Centre for Microbial Diseases and Immunity Research, University of British Columbia, 2259 Lower Mall Research Station, Vancouver, British Columbia V6T 1Z3, Canada, <sup>‡</sup>Division of Infectious Diseases, Faculty of Medicine, University of British Columbia, 2733 Heather Street, Vancouver, British Columbia V5Z 3J5, Canada, and <sup>§</sup>Institut für Medizinische Immunologie, Universitätsklinikum Charité, Humboldt-Universität zu Berlin, Schumannstr. 20-21, 10117 Berlin, Germany. <sup>†</sup>These authors contributed equally to this work.

The spread of antibiotic resistance among bacterial pathogens especially in hospital environments but also now in the community has occurred at an alarming rate (1, 2). There is now a proliferation of so-called “Superbugs” that are resistant to multiple or all antibiotics and severely limit treatment options; these include methicillin-resistant *Staphylococcus aureus* (MRSA), vancomycin-resistant enterococci (VRE), and multidrug-resistant *Pseudomonas*, among others, and cause hundreds of thousands of infections annually. Despite these alarming trends and the obvious need for new interventions, pharmaceutical companies have largely withdrawn from the field of anti-infectives (3), and only two structurally novel antibiotics have entered the market in the past 40 years (4).

Virtually all species of life produce cationic antimicrobial (host defense) peptides. These peptides can be directly antimicrobial and/or can play an important role in the functioning and orchestration of innate immune and inflammatory responses of mammals, amphibians, and insects (5–7). These peptides are typically 12–50 amino acids in length with 2–9 excess basic residues (arginine or lysine) and up to 50% hydrophobic amino acids; they fall into four major structural categories based on their amphiphilic conformations that are preformed or occur after membrane interaction, namely,  $\beta$ -structures with 2–4  $\beta$ -strands, amphipathic  $\alpha$ -helices, loop structures, and extended structures (5).

**ABSTRACT** Increased multiple antibiotic resistance in the face of declining antibiotic discovery is one of society’s most pressing health issues. Antimicrobial peptides represent a promising new class of antibiotics. Here we ask whether it is possible to make small broad spectrum peptides employing minimal assumptions, by capitalizing on accumulating chemical biology information. Using peptide array technology, two large random 9-amino-acid peptide libraries were iteratively created using the amino acid composition of the most active peptides. The resultant data was used together with Artificial Neural Networks, a powerful machine learning technique, to create quantitative *in silico* models of antibiotic activity. On the basis of random testing, these models proved remarkably effective in predicting the activity of 100,000 virtual peptides. The best peptides, representing the top quartile of predicted activities, were effective against a broad array of multidrug-resistant “Superbugs” with activities that were equal to or better than four highly used conventional antibiotics, more effective than the most advanced clinical candidate antimicrobial peptide, and protective against *Staphylococcus aureus* infections in animal models.

\*Corresponding author,  
bob@cmdr.ubc.ca.

Received for review October 2, 2008  
and accepted November 16, 2008.

Published online December 4, 2008

10.1021/cb800240j CCC: \$40.75

© 2009 American Chemical Society

Such peptides have broad spectra of activity that can encompass bacteria, fungi, viruses, and parasites (8).

In recent years, such peptides have drawn significant attention as a possible source of novel antibiotics, given their rapid action on a broad range of bacterial strains, infrequent resistance development, and limited toxicity and immunogenicity; indeed, one peptide, MX-226, has demonstrated efficacy in limiting catheter colonization in phase IIIa clinical trials (9). However there are still major questions about this class of peptides including their substantial cost-of-goods (as amino acids are expensive building blocks), unknown toxicities, and rapid degradation by proteases. Furthermore we have only a rudimentary understanding of the structural basis of antimicrobial peptide activity. This relates in part to one of the most intractable problems in biology, namely, our inability to predict protein structure based on primary sequence, limiting rational peptide design. Cationic antimicrobial peptides always interact with membranes as part of their mechanism of action (10). For different peptides this can result in either disruption of membrane barriers or translocation across the membrane to attack cytosolic targets, and most peptides are considered “dirty drugs” that attack multiple targets (11). Cationic amphiphilic peptides insert into membranes by first binding to the negatively charged surfaces of bacterial cell membranes and then inserting through a combination of electrostatic and hydrophobic interactions; selectivity occurs in part because they tend to insert much more poorly into eukaryotic cell membranes that have zwitterionic lipids at their surfaces. Thus the presence in these peptides of cationic and hydrophobic amino acids appears to be essential.

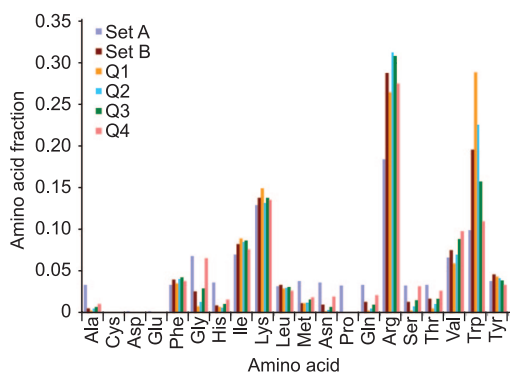
Some of these questions might be addressed if a sufficiently large pool of small effective peptides were available. The breakthrough of use of recent advancements in high-throughput peptide synthesis (peptide arrays on cellulose and rapid screening technologies) has allowed for the generation of peptide libraries of hundreds of peptides (12, 13), in a more efficient and cost-effective manner than previously employed, and has permitted a substantial reduction in the size of broad spectrum peptides. However, given the number of amino acid combinations (e.g.,  $5 \times 10^{11}$  possible 9-mers), these libraries cannot adequately explore sequence diversity.

The structural diversity of antimicrobial peptides makes it hard to relate their structure to activity. Thus de-

velopment of new more potent peptide candidates is a substantial challenge. Attempts have been made to overcome this by use of bioinformatics approaches such as the “linguistic model” for design of antimicrobial peptides (14) or rational design based on the known importance of charge and hydrophobicity. However these approaches have clear limitations due to the fluctuating structures of peptides in solution or when interacting with microbes and because they often utilize as starting points natural peptides that tend to be large and are often not very active.

The field of chemo-informatics involves computer-aided identification of new lead structures and their optimization into drug candidates (15). One of the most broadly used chemo-informatics approaches is called quantitative structure–activity relationship (QSAR) modeling, which seeks to relate the structural characteristics of a molecule (through a series of descriptors) to its measurable properties, such as biological activity. QSAR analysis has found a broad application in antimicrobial discovery. In a series of pilot studies we utilized a variety of chemical descriptors in combination with linear modeling methods such as principal component analysis and partial least-squares projections to successfully predict the antimicrobial activity of limited sets of sequence-specific cationic peptides (16). These models explicitly relate a series of input descriptors to an output prediction of activity, to permit an understanding of structure–activity relationships, but only for highly related peptides. However, when we applied these predictions methods separately to two individual libraries of peptides based on templates of the same size and composition but scrambled sequence, we were not able to extrapolate the derived relationships for one library to predict, with any significant accuracy, the activity of peptides in the other library (17).

To overcome the challenges in the design and optimization of antimicrobial peptides with superior activity toward drug-resistant bacterial pathogens, we utilized here experimental data from large random peptide libraries in combination with the descriptive power of atomic-resolution chemical descriptors and the predictive ability of an artificial neural net approach (a nonlinear modeling technique). By combining these disciplines we were able to design and correctly predict the biological activity of large pools of potent antimicrobial peptide candidates capable of combating life-threatening, multidrug-resistant pathogens.



**Figure 1. Occurrence of amino acids in the training and QSAR predicted data sets. The predicted activity quartiles from the 100,000 virtual peptide library are marked as Q1–Q4.**

## RESULTS AND DISCUSSION

**Rational Design of Antibacterial Peptide.** The bovine neutrophil cationic peptide bactenecin (RLCRIVIRVCR-NH<sub>2</sub>) is the smallest naturally occurring broad spectrum antimicrobial peptide (18). It is stabilized by an internal disulfide bridge, but the linear variant Bac2A (RLRIVIRVAR-NH<sub>2</sub>) shows a similar but still relatively modest activity against Gram-negative bacteria and somewhat improved activity against Gram-positive bacteria (19). Using the technique of SPOT synthesis hundreds of new peptides were synthesized on cellulose sheets and screened using a loss of energy-dependent luminescence activity in a *luxCDABE*-expressing *Pseudomonas aeruginosa* isolate. Combinations of single or multiple amino acid substitutions and sequence scrambling led to peptides with increased antibacterial activity (12, 13).

To verify that specific amino acid compositions and primary structures were necessary for antibacterial activity, a set of 200 random 9-mer peptides were synthesized on cellulose (using all natural coded amino acids except cysteine to avoid dimerization). No active peptides were found using the *lux* assay. Therefore using the amino acid preferences of the best peptides from our previous peptide libraries (20, 21), sets of 943 and 500 computer-selected random cellulose peptides (Sets A and B, respectively) were iteratively designed by moderately adjusting the amino acid composition of Set B to the preferences indicated by the best Set A peptides (Figure 1). The antibacterial activity of Set A measured by the *lux* assay indicated that 26% of peptides had

similar and 2.3% superior activity compared with that of Bac2A, while the adjustments made to the amino acid composition of Set B (increased content of the amino acids isoleucine, arginine, valine, and tryptophan and decreased amino acids alanine, aspartic acid, glutamic acid, glycine, histidine, methionine, asparagine, proline, glutamine, serine, and threonine; Figure 1) resulted in 48% and 5% of peptides having activity similar or superior to that of Bac2A, respectively. It is worth noting that although there was a potential for bias created by using Bac2A as the original starting point for library design, from the Bac2A sequence only the basic residue arginine appeared at high frequency in the final library and alanine, leucine, isoleucine, and valine that were present at frequencies of 17%, 8%, 17%, and 25%, respectively, in Bac2A were reduced to only 1%, 3%, 8%, and 8% of the amino acids in the second generation library (Figure 1). Thus the amino acid composition of the Set B peptides was used for random generation, by computer, of 100,000 virtual peptides (out of the approximately 70 billion possible 9-amino-acid peptide variants), with no position-specific restrictions and 16 different amino acids (omitting only four residues, three that were found to be unfavorable in previous libraries, the acidic residues aspartic acid and glutamic acid and the structure-breaking peptide proline, as well as cysteine to prevent the possibility of disulphide formation leading to dimerization).

**Computer-Aided Design of Antibacterial Peptides.** It has been demonstrated that the antimicrobial activity of cationic peptides can be influenced by their charge, hydrophobicity, and amphipathicity (22). Nonetheless, our current understanding of the exact factors determining peptide activity is incomplete. The small numbers of models developed to date have been based on closely related sequence variants and typically have utilized residue-based descriptors providing only moderately useful two-dimensional QSAR solutions (23–26). However, by abstracting from the conventional amino acid level used to approximate the character of peptides and taking into account all of their constituent atoms (including hydrogen), a number of “inductive” chemical descriptors (27–30) were calculated. These atomic-resolution characteristics have been demonstrated to be sensitive to the three-dimensional structure of a peptide and thus in preliminary investigations with closely related peptides (16, 17) appeared to describe peptide behavior more adequately than the isolated properties

of constituent amino acids. Thus we utilized this approach to train an atomic-based QSAR model based on the two tested 9-mer peptide libraries, Set A and Set B, described above.

The parameters of antimicrobial activity were utilized as dependent variables for modeling while a large number of conventional and inductive chemical descriptors (28) were used as independent variables. As this QSAR approach does not depend *per se* on the stereochemistry of the peptides, similar approaches could be used for retro peptides as well as peptidomimetics, providing sufficient were available to train the models. To relate chemical descriptors to antimicrobial activity, we employed the method of artificial neural networks, one of the most effective pattern recognition techniques utilizing artificial intelligence.

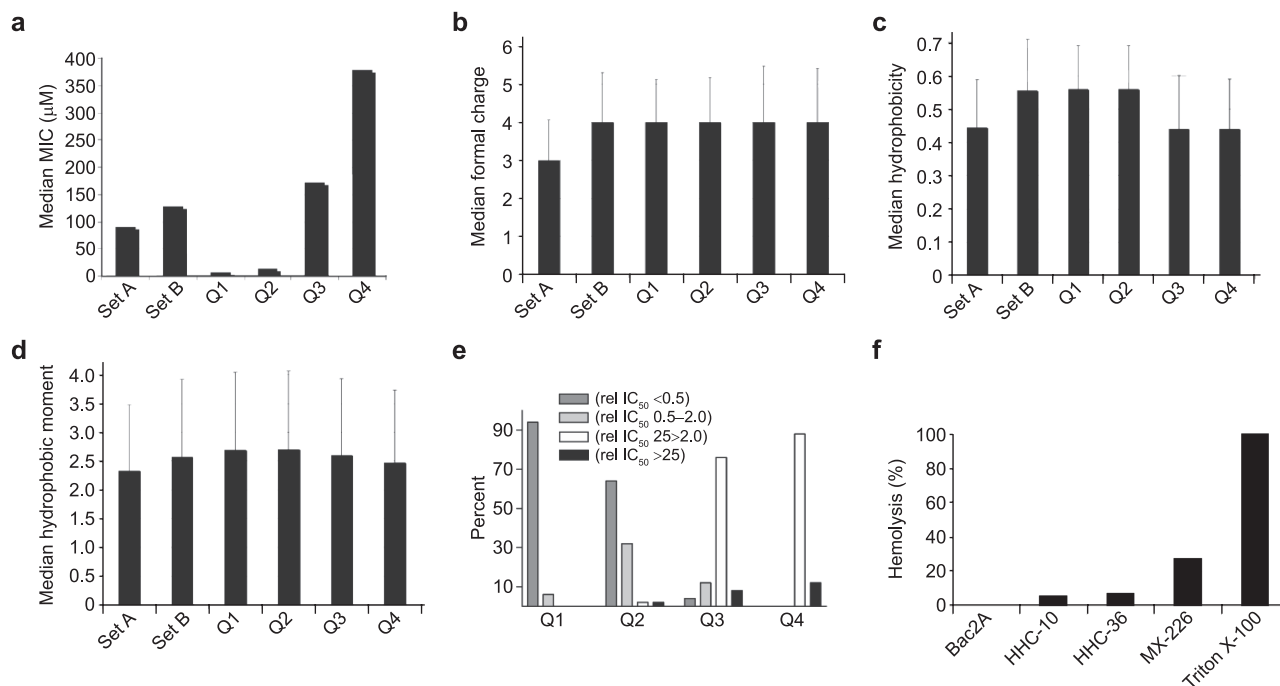
The developed QSAR solutions were then utilized to predict the activity of all peptides in our virtual library of 100,000 random 9-mer peptides. Overall these peptides were built from 16 amino acids, but the occurrence of tryptophan (19%), arginine (18%), and lysine (13%) was high and represented a total of 50% of the residues in the total 100,000 peptides, while the other 13 amino acids used represented the remaining 50%. The peptides were sorted by hypothetical antimicrobial potentials predicted by the neural network model and then grouped into four quartiles of 25,000 peptides, expected to contain high activity, medium activity, low activity, and completely inactive entries.

**In Vitro Testing of Best Virtual Candidates on *P. aeruginosa* PAO1.** To evaluate the practical applicability of the developed structure–activity models, 200 peptide candidates were selected from the entire range (50 from each quartile of the 100,000 entries), and their activities were tested against *P. aeruginosa* using the *lux* assay. This demonstrated a remarkable correlation of the measured relative IC<sub>50</sub> values of these peptides to the predictions of peptide activity by the QSAR model (luminescence readings for the peptides were subsequently compared to MIC values, giving a correlation coefficient by linear regression of  $r^2 = 0.986$ , confirming the accuracy of the luminescence screening results). Thus 98% of the peptides from quartile 1 were more active than the control peptide Bac2A (Figure 2, panel e). Quartile 2 peptides were less active overall but still demonstrated 88% highly active peptides. Quartile 3 contained substantially increased numbers of low activity peptides, and quartile 4 peptides were largely inactive.

It is important to note that peptides from each quartile had similar charge, hydrophobicity, and amphipathicity/hydrophobic moment (Figure 2, panels a–d), each of which have been considered to be critical determinants for the design of effective antimicrobial peptides (19, 25, 31). For example, two lead peptides selected from the top of first quartile, *i.e.*, HHC-10 (KRWWKWIRW) and HHC-36 (KRWWKWWRR), demonstrated high antibacterial activity, whereas peptides ranked in the lower portion of the third quartile were virtually inactive despite having very similar amino acid compositions (WHGVRW-WKW, WVKVWKYTW, WVRVYRYW, and AIRRWIRK). Thus the relative positioning of amino acids within the peptide sequence was clearly critical and likely reflected structural features that were discerned through the neural network predictions of effective combinations of chemical descriptors.

Assessment of the antibacterial activities of 20 peptides selected from the first three quartiles using standard MIC assays confirmed these observations (data not shown), demonstrating that the QSAR approach accurately forecasted the antimicrobial activity of *de novo* designed peptides. Thus, first quartile peptides demonstrated very significant activity against *P. aeruginosa* PAO1 with corresponding MIC values ranging from 1.4 to 6.8  $\mu\text{M}$ . In particular, two of the tested first quartile peptides, HHC-10 and HHC-36 (Table 1), outperformed the most active peptides in the training sets. This clearly stresses the importance of utilizing larger training sets and inductive descriptors that take into account interactions between adjacent amino acids, to permit the development of more effective QSAR models. Peptides selected from the second quartile also demonstrated substantial antimicrobial activity with MIC values from 4 to 12  $\mu\text{M}$ , and as predicted by the QSAR model, the third quartile selection did not return any generally active substances, and the fourth quartile peptides tested were all completely inactive (in agreement with the results presented in Figure 2, panel e).

**In Vitro Tests with Highly Antibiotic-Resistant Pathogens.** The two selected lead peptides, HHC-10 and HHC-36, were further evaluated for their *in vitro* activities, MIC, against several of the most multidrug-resistant and problematic pathogens, collectively termed Superbugs, including strains of multidrug-resistant *P. aeruginosa*, methicillin-resistant *Staphylococcus aureus* (MRSA), *Enterobacter cloacae* with derepressed chromosomal  $\beta$ -lactamase, extended spectrum



**Figure 2.** Antimicrobial activity and physical parameters for antimicrobial peptides from Training Sets A and B and peptides from the 100,000 peptide virtual library. **a)** Median MIC versus *P. aeruginosa* of antimicrobial peptides from the training Sets A (91  $\mu\text{M}$ ) and B (127  $\mu\text{M}$ ) and the corresponding median values for the 50 experimentally tested peptides separated into activity quartiles (7, 13, 172, and 379  $\mu\text{M}$ , corresponding to Q1, Q2, Q3, and Q4). **b)** Median values of formal net charge, **c)** hydrophobicity (expressed as hydrophobic fraction), and **d)** amphipathicity (expressed as hydrophobic moment) for peptides in training Sets A and B and for the 100,000 virtual peptides divided into activity quartiles Q1–Q4. Molecular weights of these peptides also varied very little (not shown). **e)** Activities of 200 peptides from the 100,000 peptide virtual library. Q1: first quartile peptides; Q2: second quartile; Q3: third quartile; Q4: fourth quartile; rel  $IC_{50}$  is the relative  $IC_{50}$ , the ratio of the  $IC_{50}$  for the experimental peptide to the  $IC_{50}$  of Bac2A. Peptides where the highest concentration failed to reduce the luminescence by at least 50% were identified as inactive. **f)** Toxic effect of control and lead peptides monitored by hemolytic activity at a peptide concentration of 188  $\mu\text{g mL}^{-1}$  compared to Bac2A and MX-226.

$\beta$ -lactamase producing *Escherichia coli* and *Klebsiella pneumoniae*, and vancomycin-resistant *Enterococcus faecalis* and *Enterococcus faecium* (VRE) (The results are expressed in micromolar concentration, but because these peptides have molecular weights of around 1000, MICs would have similar values in  $\mu\text{g mL}^{-1}$ .) The results demonstrated that both peptides had significant *in vitro* inhibitory activity against antibiotic-resistant bacteria. More specifically they exhibited MICs of 0.3–11  $\mu\text{M}$  against most of the tested Superbugs and were clearly superior to the only antimicrobial peptide to show efficacy to date in advanced clinical trials, MX-226 (9), which generally exhibited MICs of 10–76  $\mu\text{M}$ . For these exceptionally resistant bacteria, the peptides demonstrated far broader and in most cases more po-

tent activity compared to that of highly utilized clinical antibiotics such as the aminoglycoside tobramycin, fluoroquinolone ciprofloxacin, and  $\beta$ -lactams ceftazidime and imipenem, which represent the most potent, highly utilized variants of their respective chemical classes (Table 1). These results characterize the developed peptides as excellent antibiotic candidates for treating some of the most recalcitrant and dangerous human infections.

**Lack of Peptide Toxicity.** To assess possible host cell toxicity of the developed compounds, the two selected lead peptides were also tested for their toxicity. Hemolytic activity was monitored by hemin release over a broad range of peptide concentrations (4.4–251  $\mu\text{M}$ ). There was minimal red blood cell lysis at all concentra-

**TABLE 1. Activities against multidrug-resistant Superbugs<sup>a</sup>**

Bacterium	MIC (μM)							
	Bac2A	HHC-10	HHC-36	MX-226	Tobramycin	Ciprofloxacin	Imipenem	Ceftazidime
<i>P. aeruginosa</i>								
H103 (wild-type)	48	<b>0.8</b>	<b>0.7</b>	19	1.1	0.4	6.7	3.6
9 (Brazil MDR)	192	<b>5.9</b>	<b>5.7</b>	76	>274	386	428	>231
198 (Brazil MDR)	95	<b>1.5</b>	<b>1.4</b>	153	>274	97	107	231
213 (Brazil MDR)	192	<b>3</b>	<b>5.7</b>	38	>274	193	428	>231
LES400 (Liverpool MDR)	95	<b>5.9</b>	<b>11</b>	76	8.6	1.5	3.3	58
H1027 (Liverpool MDR)	12	<b>0.8</b>	<b>0.3</b>	38	69	3.0	27	231
H1030 (Liverpool MDR)	95	<b>3</b>	<b>1.4</b>	76	17	0.8	27	58
<i>P. maltophilia</i>								
ATCC 13637	3	<b>0.4</b>	<b>1.4</b>	19	8.6	0.8	428	0.5
<i>E. cloacae</i>								
218R Class C β-lactamase	24	<b>3</b>	<b>11</b>	9.6	1.1	0.4	0.4	58
<i>E. coli</i>								
63103 (ESBL)	24	<b>1.5</b>	<b>5.4</b>	38	137	>386	0.4	231
64771 (ESBL)	24	<b>1.5</b>	<b>2.7</b>	38	274	>386	0.4	>231
<i>K. pneumoniae</i>								
61962 (ESBL)	192	<b>25</b>	<b>174</b>	153	69	0.4	0.4	>231
63575 (ESBL)	192	<b>6.2</b>	<b>22</b>	76	34	193	0.4	>231
<i>S. aureus</i>								
ATCC 25923	24	<b>3</b>	<b>2.9</b>	9.6	0.5	0.4	0.4	29
C623 (MRSA)	24	<b>1.5</b>	<b>1.4</b>	19	>274	6.0	0.4	115
<i>E. faecalis</i>								
ATCC 29212	12	<b>12</b>	<b>43</b>	76	34	0.8	1.7	231
W61950 (VRE; VanA)	48	<b>99</b>	<b>&gt;174</b>	>153	>274	97	6.7	>231
f43559 (VRE; VanB)	12	<b>49</b>	<b>174</b>	153	69	97	6.7	>231
<i>E. faecium</i>								
mic80(VRE; VanA)	3	<b>3.1</b>	<b>11</b>	38	274	97	428	>231
t62764 (VRE; VanB)	48	<b>1.5</b>	<b>1.3</b>	76	>274	386	>428	>231

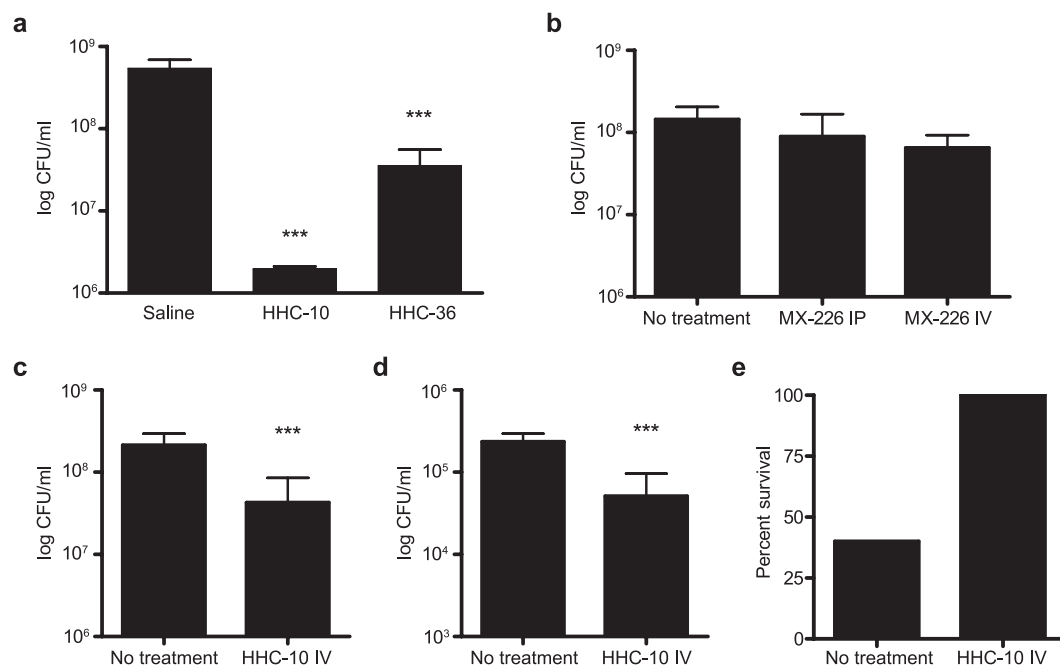
<sup>a</sup> Activities of selected peptides (HHC-10 and HHC-36) predicted through the QSAR analysis, compared to the original parent peptide Bac2A and the control peptide MX-226, as well as the some of the most highly utilized antibiotics in our society, aminoglycoside tobramycin, the fluoroquinolone ciprofloxacin, the carbapenem imipenem, and the cephalosporin ceftazidime. MIC values were measured in 3–5 replicates.

tions, and both lead peptides demonstrated substantially less lysis than observed with MX-226, indicating that these peptides were quite pathogen-specific (Figure 2, panel f). Similar very low toxicity was observed in metabolically active cells by using the WST-1 reagent to assay the reduction in mitochondrial activity of human peripheral blood mononuclear (PBMC) cells (data not shown).

**In Vivo Tests with *S. aureus*.** To examine the potential of these novel antimicrobial peptides *in vivo*, two of

the most potent broad spectrum candidates, HHC-10 and HHC-36, in addition to the control peptide MX-226, were selected, and their antimicrobial activities were assessed in a well-established mouse model of invasive Staphylococcal infection (32–35) that is commonly utilized to demonstrate antimicrobial activity (35). Female CD1 mice were infected intraperitoneally (IP) with high doses (~10<sup>9</sup> bacteria) of *S. aureus*, to avoid self-clearance through innate immune responses. When administered by the same route 4 h postinfection, both





**Figure 3.** Ability of the peptides HHC-10 and HHC-36 to protect mice against invasive *S. aureus* infections. Representative experiments are presented. Mice were infected with  $\sim 10^9$  CFU per mouse of *S. aureus* (ATCC 25923) in porcine gastric mucin and 4 h later were treated with the indicated peptides, as described. Twenty-four hours postinfection, mice were sacrificed and peritoneal bacterial CFU counts were assessed by overnight culture on LB plates. a) Peritoneal bacterial loads after IP treatment with saline, HHC-10, or HHC-36 administered 4 h postinfection. b) Peritoneal bacterial loads after no treatment or treatment with MX-226 *via* the IP route, administered 4 h postinfection. When the peptides were administered IP the survival frequency was not significantly different under any condition, using the Logrank test (Mantel-Cox test). c,d) Bacterial loads in the peritoneal lavage (c) and blood (d) after intravenous (IV) administration of as a single dose of HHC-10 (200  $\mu\text{g}$  per mouse or 8  $\text{mg kg}^{-1}$ ), 4 h after the initiation of infection. Dead animals were assigned the highest CFU count obtained in the experiment. The designation \*\*\* indicates  $p < 0.001$  by Chi-square analysis with a 95% confidence interval. e) Mortality after IV administration of peptide was 60% in the control group and 0% in the peptide treated animals, which is significantly difference with a  $P$  value of 0.0495 by Logrank test. Effects similar to those in panels c and d were observed when mice were treated with IV HHC-36 4 h postinfection (data not shown).

HHC-10 and HHC-36 ( $4 \text{ mg kg}^{-1}$  or  $100 \mu\text{g}$  per mouse) demonstrated substantial ability to reduce peritoneal counts 24 h later (Figure 3, panel a). Complete eradication of the bacteria was observed in some animals 24 h postinfection, though a decrease in bacterial counts and mortality was the most frequent result. In contrast to these peptides, the most advanced clinical peptide MX-226, when provided at the same dose and route, was unable to mediate protection (Figure 3, panel b). Significant protection was also observed when HHC-10 was inoculated through intravenous (IV) injection (Figure 3, panels c and d). None of the mice given HHC-10 *via* the IV route died, whereas 60% of the control animals succumbed to infection. Thus unlike most

other antimicrobial peptides investigated to date, these peptides may, if appropriately formulated, be effective against systemic infections.

**Conclusions.** These results reveal a new, effective approach for designing highly active broad spectrum peptides smaller than those found in Nature. Indeed tryptophan- and arginine-rich peptides do exist in Nature (*e.g.*, indolicidin, the parent peptide of MX-226 (36)) but are 12–13 amino acids long and have activities at least 10-fold poorer than peptides HHC-10 and HHC-36. The structure–activity models produced through neural network approaches substantially enriched the list of drug candidates. These results illustrate that adequate modeling of antimicrobial activity

of cationic peptides demands much more refined structure–activity approaches than simple considerations of polar and hydrophobic characteristics of constituent amino acids and indeed requires a detailed, atomic level of consideration of molecular structures. Since obtaining structural models of short peptides, *e.g.*, by utilizing two-dimensional NMR approaches, is very time-consuming, we suggest that only the use of powerful chemoinformatic approaches can bring about the rational design of antimicrobial peptides at a large-scale level. For example, if we were to consider a comprehensive library of 9-amino-acid peptides composed of only the five amino acids (tryptophan, arginine, lysine, isoleucine, leucine) that were predominant in active variants from our previous random peptide libraries, it would still be necessary to consider ~2 million sequence variants.

Recent advancements in high-throughput peptide synthesis and testing technologies have provided substantial support for the development of potent peptide-

based antibiotics. At the same time, there has been a lack of effective unbiased computational approaches enabling rational and large-scale design of novel candidates. We have demonstrated here that QSAR methodology combined with artificial intelligence/neural network approaches can significantly accelerate the discovery of antimicrobial peptides, even from semirandom starting points based on diverse peptide sequences. In particular, we have utilized experimental data, obtained from peptide array experiments that permit the assessment of antibacterial activity for large numbers of peptides, to train QSAR models that are capable of accurate *in silico* prediction of antimicrobials of diverse sequences. The developed computational solutions allowed effective virtual interrogation of a large number of peptide candidates and resulted in the development of new antibiotic leads that were effective against a broad spectrum of multidrug-resistant, life-threatening pathogenic bacteria.

## METHODS

**Peptide Synthesis.** Peptide syntheses on cellulose were performed using a pipetting robot (Abimed, Langenfeld, Germany) and Whatman 50 cellulose membranes (Whatman, Maidstone, U.K.) as described previously (21). The HPLC purified peptides (>95% pure) used for detailed characterization of antimicrobial activity *in vitro* and *in vivo* were purchased from Thermo electron cooperation (Ulm, Germany).

**Prediction of Peptide Activity.** The design and prediction of novel antibacterial peptides was performed using artificial neural networks, which represents one of the most broadly utilized machine-learning techniques that are based on the basic principles of brain organization and memory mechanisms. The theory behind our approach and the specifics of model training and optimization will be described elsewhere (37).

**Luminescence-Based IC<sub>50</sub> Assay of Peptides from Cellulose Support.** The peptides were cleaved from the dried membrane in an ammonia atmosphere overnight, resulting in free peptides with an amidated C-terminus. Subsequently the peptide spots were punched out, transferred into microtiter plates, diluted, and tested for their ability to reduce the energy (FMN)-dependent luminescence of *P. aeruginosa* strain H1001 *fliC::luxCDABE*, as described previously (20).

**Minimal Inhibitory Concentration (MIC) Determination.** The MIC of the peptides was measured using a modified broth microdilution method in Difco Mueller Hinton (MH) medium (Becton-Dickenson, Franklin Lakes, NJ) (19). Briefly, the peptides were dissolved and stored in glass vials. The assay was performed in sterile 96-well polypropylene microtiter plates (cat. 3790, Costar, Cambridge, MA). Serial dilutions of the peptides to be assayed were performed in 0.01% acetic acid containing 0.2% bovine serum albumin at 10-fold the desired final concentration. Ten microlitres of the 10-fold concentrated peptides were added to each well of a 96-well polypropylene plate containing 90  $\mu$ L of MH media per well. Bacteria were added to the

plate from an overnight culture at a final concentration of  $2-7 \times 10^5$  CFU mL<sup>-1</sup> and incubated overnight at 37 °C. The MIC was taken as the concentration at which no growth was observed.

MIC analyses were done on a panel of bacterial pathogens that were both susceptible and resistant to common antibiotics. *P. aeruginosa* PAO1 strain H103 (19), *P. maltophilia* ATCC 13637, *S. aureus* ATCC 25923 (19), *Enterococcus faecalis* ATCC 29212 (13), and *Enterobacter cloacae* 218R, constitutively expressing Class C chromosomal  $\beta$ -lactamase (38), were from our laboratory strains collection. A methicillin-resistant *S. aureus* (MRSA) clinical isolate was kindly provided by Anthony Chow (Vancouver General Hospital, Vancouver, Canada). Two *K. pneumoniae* and two *E. coli* clinical isolates expressing extended spectrum  $\beta$ -lactamases (ESBL) were kindly provided by George Zhanel (Health Sciences Centre, Winnipeg, Canada). Vancomycin-resistant clinical isolates of *E. faecalis* and *E. faecium* were obtained from Ana M. Paccagnella (BC Centre for Disease Control, Vancouver, Canada). Three clinical isolates (nos. 9, 198, and 213) of multidrug-resistant *P. aeruginosa* were kindly provided by Carlos Kiffer (University of São Paulo, Brazil). These isolates all have resistance to piperacillin/tazobactam, Meropenem, ceftazidime, ciprofloxacin, and cefepime, and no. 9 is also polymyxin B resistant. Three *P. aeruginosa* clinical isolates of the Liverpool epidemic strain (LES) (H1027, H1030, and LES400) (39) were all kindly provided by Craig Winstanley (University of Liverpool, UK). LES400 was resistant to gentamicin and tobramycin, and H1030 showed resistance to colistin, amikacin, gentamicin, and tobramycin. All tested bacterial strains were categorized as biohazard level 2 pathogens.

**Cytotoxicity Assessment.** Fresh human venous blood was collected from volunteers in Vacutainer collection tubes containing sodium heparin as an anticoagulant (BD Biosciences, Franklin Lakes, NJ), in accordance with UBC ethics approval and guidelines. The blood was diluted 1:1 in complete RPMI 1640 medium and separated over a Ficoll-Paque Plus gradient (Amer-



sham Biosciences) by centrifugation, separating the peripheral blood mononuclear cells (PBMC), white blood cells from the red blood cells. The toxic effect of the peptides was assessed by hemoglobin release from human red blood cells resulting from cell lysis. The red blood cells were washed three times in sterile 0.85% NaCl (saline) and centrifuged at 1500 rpm for 10 min. Concentrated red blood cells were diluted 3-fold in saline, and 150  $\mu$ L of this cell suspension was mixed with 50  $\mu$ L of peptide at concentrations ranging from 4.4 to 251  $\mu$ M, diluted in a mixture of 0.01% acetic acid and 0.2% BSA. Triton X-100 (1%) was used as a positive control, demonstrating 100% cell lysis, and sterile saline was considered as a negative control. The assay was carried out in 96-well polypropylene microtiter plates. The red blood cells and peptide dilutions were incubated on a rocking table to obtain circulation at 37 °C under 5% CO<sub>2</sub> pressure for 24 h. Hemoglobin release was monitored chromogenically at 414 and 546 nm using an ELISA plate reader.

To further test toxicity, PBMC ( $2 \times 10^5$ ) were seeded into 96-well plates (Sarstedt, Newton, NC) and incubated at 37 °C in 5% CO<sub>2</sub> overnight. The WST-1 assay measuring mitochondrial activity in PBMC was then performed after 24 h of incubation with the peptides (50  $\mu$ g mL<sup>-1</sup>) in a medium containing 0.01% acetic acid and 0.2% BSA (used for MIC testing). All experiments were done in triplicate.

**Animal Models.** The animal models were performed in accordance with UBC animal care ethics approval and guidelines, as per animal care certificate A04-0020. A dose of  $\sim 10^9$  CFU per mouse of *S. aureus* ATCC 25923 was suspended in MH broth containing 5% mucin and injected intraperitoneally into female CD1 mice. Peptides HHC-10, HHC-36, or MX-226, resuspended in PBS, pH 7.0, were administered IP or IV at 4 mg kg<sup>-1</sup>, 4 h postinfection. The animals were euthanized 24 h postinfection, and blood and peritoneal lavage were collected, diluted in PBS, and plated on MH agar plates that were incubated overnight at 37 °C and counted for viable bacteria. Animals that died during the study were assigned the highest colony forming unit count obtained in the experiment. The statistical differences between peptide-treated and nave mice were assessed using PRISM (GraphPad Software Inc., San Diego, CA) using contingency tables Chi-square test, with a confidence interval of 95%.

**Disclosure:** The peptides described here have been submitted as part of a U.S. patent application.

**Acknowledgment:** We gratefully acknowledge financial support from the Canadian Institutes for Health Research (CIHR) and the Foundation of the National Institutes of Health and CIHR through the Grand Challenges in Global Health Initiative. We thank Jessica Lee for technical support in creating the computer-based peptide libraries. R.H. is the recipient of a Canada Research Chair. K.H. holds a CIHR fellowship. C.D.F. received a Doctoral Research Award from CIHR.

## REFERENCES

- Hancock, R. E. (2007) The end of an era? *Nat. Rev. Drug Discovery* 6, 28.
- Theuretzbacher, U., and Toney, J. H. (2006) Nature's clarion call of antibacterial resistance: are we listening? *Curr. Opin. Invest. Drugs* 7, 158–166.
- Projan, S. J. (2003) Why is big Pharma getting out of antibacterial drug discovery? *Curr. Opin. Microbiol.* 6, 427–430.
- Spellberg, B., Powers, J. H., Brass, E. P., Miller, L. G., and Edwards, J. E., Jr. (2004) Trends in antimicrobial drug development: implications for the future. *Clin. Infect. Dis.* 38, 1279–1286.
- Hancock, R. E., and Sahl, H. G. (2006) Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies. *Nat. Biotechnol.* 24, 1551–1557.
- Oppenheim, J. J., and Yang, D. (2005) Alarmins: chemotactic activators of immune responses. *Curr. Opin. Immunol.* 17, 359–365.
- Zasloff, M. (2002) Antimicrobial peptides of multicellular organisms. *Nature* 415, 389–395.
- Jenssen, H., Hamill, P., and Hancock, R. E. (2006) Peptide antimicrobial agents. *Clin. Microbiol. Rev.* 19, 491–511.
- Hamill, P., Brown, K., Jenssen, H., and Hancock, R. E. (2008) Novel anti-infectives: is host defence the answer? *Curr. Opin. Biotechnol.* E-pub ahead of print. DOI: 10.1016/j.copbio.2008.10.006.
- Hancock, R. E., and Rozek, A. (2002) Role of membranes in the activities of antimicrobial cationic peptides. *FEMS Microbiol. Lett.* 206, 143–149.
- Peschel, A., and Sahl, H. G. (2006) The co-evolution of host cationic antimicrobial peptides and microbial resistance. *Nat. Rev. Microbiol.* 4, 529–536.
- Hilpert, K., Elliott, M. R., Volkmer-Engert, R., Henklein, P., Donini, O., Zhou, Q., Winkler, D. F., and Hancock, R. E. (2006) Sequence requirements and an optimization strategy for short antimicrobial peptides. *Chem. Biol.* 13, 1101–1107.
- Hilpert, K., Volkmer-Engert, R., Walter, T., and Hancock, R. E. (2005) High-throughput generation of small antibacterial peptides with improved activity. *Nat. Biotechnol.* 23, 1008–1012.
- Loose, C., Jensen, K., Rigoutsos, I., and Stephanopoulos, G. (2006) A linguistic model for the rational design of antimicrobial peptides. *Nature* 443, 867–869.
- Engel, T. (2006) Basic overview of chemoinformatics. *J. Chem. Inf. Model* 46, 2267–2277.
- Jenssen, H., Lejon, T., Hilpert, K., Fjell, C. D., Cherkasov, A., and Hancock, R. E. (2007) Evaluating different descriptors for model design of antimicrobial peptides with enhanced activity toward *P. aeruginosa*. *Chem. Biol. Drug Des.* 70, 134–142.
- Jenssen, H., Fjell, C. D., Cherkasov, A., and Hancock, R. E. (2008) QSAR modeling and computer-aided design of antimicrobial peptides. *J. Pept. Sci.* 14, 110–114.
- Romeo, D., Skerlavaj, B., Bolognesi, M., and Gennaro, R. (1988) Structure and bactericidal activity of an antibiotic dodecapeptide purified from bovine neutrophils. *J. Biol. Chem.* 263, 9573–9575.
- Wu, M., and Hancock, R. E. (1999) Improved derivatives of bactenecin, a cyclic dodecameric antimicrobial cationic peptide. *Antimicrob. Agents Chemother.* 43, 1274–1276.
- Hilpert, K., and Hancock, R. E. (2007) Use of luminescent bacteria for rapid screening and characterization of short cationic antimicrobial peptides synthesized on cellulose using peptide array technology. *Nat. Protoc.* 2, 1652–1660.
- Hilpert, K., Winkler, D. F., and Hancock, R. E. (2007) Peptide arrays on cellulose support: SPOT synthesis, a time and cost efficient method for synthesis of large numbers of peptides in a parallel and addressable fashion. *Nat. Protoc.* 2, 1333–1349.
- McPhee, J. B., Scott, M. G., and Hancock, R. E. (2005) Design of host defence peptides for antimicrobial and immunity enhancing activities. *Comb. Chem. High Throughput Screening* 8, 257–272.
- Ostberg, N., and Kaznessis, Y. (2005) Protegrin structure-activity relationships: using homology models of synthetic sequences to determine structural characteristics important for activity. *Peptides* 26, 197–206.
- Lejon, T., Strom, M. B., and Svendsen, J. S. (2001) Antibiotic activity of pentadecapeptides modelled from amino acid descriptors. *J. Pept. Sci.* 7, 74–81.
- Jenssen, H., Gutteberg, T. J., and Lejon, T. (2005) Modelling of anti-HSV activity of lactoferrin analogues using amino acid descriptors. *J. Pept. Sci.* 11, 97–103.
- Freder, V. (2006) QSAR analysis of antimicrobial and haemolytic effects of cyclic cationic antimicrobial peptides derived from protegrin-1. *Bioorg. Med. Chem.* 14, 6065–6074.
- Cherkasov, A. (2003) Inductive electronegativity scale. Iterative calculation of inductive partial charges. *J. Chem. Inf. Comput. Sci.* 43, 2039–2047.

28. Cherkasov, A. (2005) Inductive QSAR descriptors, distinguishing compounds with antibacterial activity by artificial neural network, *Int. J. Mol. Sci.* **6**, 63–86.
29. Karakoc, E., Sahinalp, S. C., and Cherkasov, A. (2006) Comparative QSAR and fragments distribution analysis of drugs, druglikes, metabolic substances, and antimicrobial compounds, *J. Chem. Inf. Model.* **46**, 2167–2182.
30. Cherkasov, A. (2006) Can “bacterial-metabolite-likeness” model improve odds of “in silico” antibiotic discovery? *J. Chem. Inf. Model.* **46**, 1214–1222.
31. Yeaman, M. R., and Yount, N. Y. (2003) Mechanisms of antimicrobial peptide action and resistance, *Pharmacol. Rev.* **55**, 27–55.
32. Suzuki, N., Suzuki, S., Duncan, G. S., Millar, D. G., Wada, T., Mirtos, C., Takada, H., Wakeham, A., Itie, A., Li, S., Penninger, J. M., Wesche, H., Ohashi, P. S., Mak, T. W., and Yeh, W. C. (2002) Severe impairment of interleukin-1 and Toll-like receptor signalling in mice lacking IRAK-4, *Nature* **416**, 750–756.
33. Mullaly, S. C., and Kubes, P. (2006) The role of TLR2 *in vivo* following challenge with *Staphylococcus aureus* and prototypic ligands, *J. Immunol.* **177**, 8154–8163.
34. Scott, M. G., Dullaghan, E., Mookherjee, N., Glavas, N., Waldbrook, M., Thompson, A., Wang, A., Lee, K., Doria, S., Hamill, P., Yu, J. J., Li, Y., Donini, O., Guama, M. M., Finlay, B. B., North, J. R., and Hancock, R. E. (2007) An anti-infective peptide that selectively modulates the innate immune response, *Nat. Biotechnol.* **25**, 465–472.
35. Tsuji, M., Takema, M., Miwa, H., Shimada, J., and Kuwahara, S. (2003) *In vivo* antibacterial activity of S-3578, a new broad-spectrum cephalosporin: methicillin-resistant *Staphylococcus aureus* and *Pseudomonas aeruginosa* experimental infection models, *Antimicrob. Agents Chemother.* **47**, 2507–2512.
36. Falla, T. J., Karunaratne, D. N., and Hancock, R. E. (1996) Mode of action of the antimicrobial peptide indolicidin, *J. Biol. Chem.* **271**, 19298–19303.
37. Fjell, C.D., Jenssen, H., Hilpert, K., Cheung, W.A., Panté, N., Hancock, R.E.W., and Cherkasov, A. (2009) Identification of novel antibacterial peptides by chemoinformatics and machine learning. Manuscript in preparation.
38. Bellido, F., Pechere, J. C., and Hancock, R. E. (1991) Reevaluation of the factors involved in the efficacy of new  $\beta$ -lactams against *Enterobacter cloacae*, *Antimicrob. Agents Chemother.* **35**, 73–78.
39. Salunkhe, P., Smart, C. H., Morgan, J. A., Panagea, S., Walshaw, M. J., Hart, C. A., Geffers, R., Tummier, B., and Winstanley, C. (2005) A cystic fibrosis epidemic strain of *Pseudomonas aeruginosa* displays enhanced virulence and antimicrobial resistance, *J. Bacteriol.* **187**, 4908–4920.