



In silico design of antimicrobial oligopeptides based on 3D-QSAR modeling and bioassay evaluation

Guangping Li¹ · Yuxuan Wang¹ · Yan Shen^{1,2,3} · Haiqiong Guo⁴ · Qingxiu He¹ · Yong Hu^{1,2} · Haibin Liu⁵ · Zhihua Lin^{1,2,3} · Yuanqiang Wang^{1,2,3}

Received: 13 May 2021 / Accepted: 20 August 2021 / Published online: 2 September 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

The emergence of multidrug resistance bacteria poses a great health threat. Therefore, it is a crucial demand to obtain new antibacterial drugs. Antimicrobial peptides (AMPs) have the characteristics of wide antimicrobial spectrum and lower drug resistance, hence, it is hopeful to substitute for classical antibiotics. In this study, two classic methods, comparative molecular field analysis (CoMFA) and comparative molecular similarity index analysis (CoMSIA), were used to analyze the structural feature of AMPs against *Staphylococcus aureus* or *Escherichia coli*, respectively. Subsequently, the models established by three-dimensional quantitative structure–activity relationships (3D-QSAR) (for *S. aureus*, CoMFA: $Q^2 = 0.512$, $R^2 = 0.943$, $F = 59.916$; CoMSIA: $Q^2 = 0.645$, $R^2 = 0.993$, $F = 339.242$; for *E. coli*, CoMFA: $Q^2 = 0.507$, $R^2 = 0.913$, $F = 66.862$; CoMSIA: $Q^2 = 0.573$, $R^2 = 0.966$, $F = 96.84$) have good predictability and stability was constructed. Seven novel small AMPs were designed and synthesized based on the theoretical model. The novel AMPs showed potent antibacterial activity against *S. aureus* and *E. coli* while causing low host toxicity. This study provides a potential therapeutic option using 3D-QSAR models guiding the design and modification of novel AMPs, to address the prevalent infections caused by MDR bacterial.

These authors contributed equally: Guangping Li, Yuxuan Wang, Yan Shen

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1007/s00044-021-02789-4>.

✉ Haibin Liu
liuhaibin@dongeejiao.com

✉ Zhihua Lin
zhlin@cqut.edu.cn

✉ Yuanqiang Wang
wangyqnn@cqut.edu.cn

¹ School of Pharmacy and Bioengineering, Chongqing University of Technology, Chongqing 400054, China

² Chongqing Key Laboratory of Medicinal Chemistry & Molecular

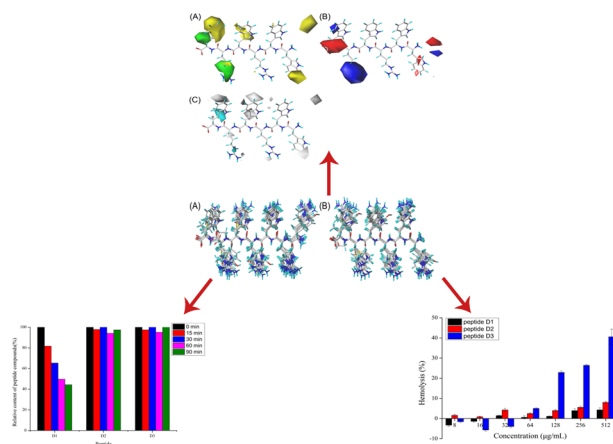
Pharmacology, Chongqing University of Technology, Chongqing 400054, China

³ Key Laboratory of Target Based Drug Screening and Activity Evaluation, Chongqing University of Technology, Chongqing 400054, China

⁴ State Key Laboratory of Silkworm Genome Biology, Southwest University, Chongqing 400716, China

⁵ National Engineering Research Center for Gelatin-based Traditional Chinese Medicine, Dong-E-E-Jiao Co, Shenzhen 252201, China

Graphical Abstract



Keywords 3D-QSAR · CoMFA · CoMSIA · Antimicrobial peptides · *S. aureus* · *E. coli*

Introduction

Antibiotics have been widely used for their efficiency in the prevention and treatment of bacterial infections [1] for more than 70 years. However, the abuse of antibiotics may cause the bacterial resistance, which can eliminate the effectiveness of these medications [2]. Cases for this are abundant, for example, the vancomycin-resistant enterococci bacteria [3], the methicillin-resistant *Staphylococcus aureus* [4], and the ampicillin-resistant *Escherichia coli* [5]. The continuous emergence of various resistant bacteria and even super bacteria has made bacterial infectious diseases an urgent problem to be solved [6, 7]. Hence, the development of new, effective, and safe antibacterial drugs is quite necessary.

Antimicrobial peptides (AMPs) are effector molecules of the innate immune system in multicellular organisms that can resist the invasion of exogenous pathogens [8]. These substances were first discovered and studied in the natural immunity of insects around the 1980s [9–11]. AMPs can be found in various organisms such as fruit flies [12, 13], bees [14, 15], marine animals [16], and mammals [17]. So far, hundreds of AMPs have been found and they can be structurally divided into four major classes: α -helical, β -sheet, loop, and extended peptides [18]. Compared with traditional antibiotics, the advantages of AMPs include good thermal stability, high sterilization rate, broad antimicrobial spectrum, inhibition of fungi, parasites, and viruses growth. Therefore, AMPs are potentially an ideal solution to antibiotic resistance.

Traditional experimental methods for identifying antibacterial peptides from polypeptides are expensive

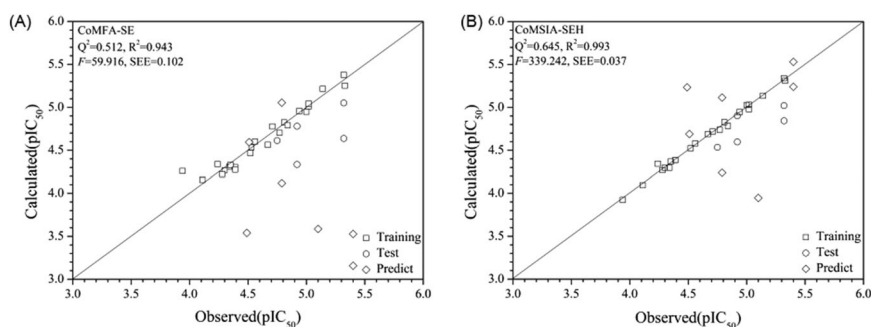
and time-consuming [19]. An alternative solution is to use computational methods to screen and predict the AMPs [20, 21]. Compared with traditional experimental methods, computational methods are more efficient and economical [22]. Heretofore, fruitful achievements have been made by the computational methods in the investigation of AMPs [23].

Among the computational methods, QSAR models are frequently used to predict the activity and properties of an unknown set of molecules. They are essentially regression or classification models, linking biological activity and physical and chemical properties with molecular structures [24, 25]. In recent years, researchers have successfully applied the three-dimensional quantitative structure–activity relationships (3D-QSAR) models, which are the extensions of traditional QSAR models by considering the 3D structures of the molecules [26], to evaluate the antibacterial activity of AMPs and provide theoretical guidance for the synthesis of novel AMPs [27]. Therefore, the 3D-QSAR models provided the impetus for designing a new generation of synthetic AMPs.

In this study, we collected the molecular structures of 29 AMPs against *S. aureus* and 30 AMPs against *E. coli* from the literature. Two frequently used 3D-QSAR models (comparative molecular field analysis (CoMFA) and comparative molecular similarity index analysis (CoMSIA)) [28] were then constructed and evaluated. Based on these validated models, seven peptides were designed, and tested for their cytotoxicity, plasma stability, and activity against *S. aureus* and *E. coli*. The present study can serve as the theoretical support for the future design of novel AMPs.

Table 1 Related data of the CoMFA and CoMSIA models

Bacteria	Method	LOO		PLS			Normalized coefficients					External validation	
		np	Q^2	R^2	SEE	F	S	E	H	D	A	r^2_{pred}	SDEP _{ext}
<i>S. aureus</i>	CoMFA	5	0.512	0.943	0.102	59.916	0.731	0.269				0.746	0.471
	CoMSIA	7	0.645	0.993	0.037	339.242	0.164	0.323	0.513			0.888	0.350
<i>E. coli</i>	CoMFA	3	0.507	0.913	0.099	66.862	0.577	0.423				0.578	0.435
	CoMSIA	5	0.573	0.966	0.065	96.841		0.316		0.401	0.284	0.623	0.411

Fig. 1 The plots of observed and calculated value of the peptides, together with the predicted value of the seven novel peptides of the antimicrobial peptides against *S. aureus*

Results and discussion

The results of 3D-QSAR

Analysis of CoMFA models

The CoMFA model contains two force fields: the steric field (S) and the electrostatic field (E). For both AMPs, through the computational data comparison (Supplementary Tables S1 and S3), the $S + E$ field showed the best result (AMPs against *S. aureus*: $Q^2 = 0.51$, $np = 5$, $R^2 = 0.943$, $SEE = 0.102$. AMPs against *E. coli*: $Q^2 = 0.577$, $np = 3$, $R^2 = 0.913$, $SEE = 0.099$, the related data are shown in Table 1); for the external validation peptides, AMPs against *S. aureus*: $r^2_{\text{pred}} = 0.746$, $\text{SDEP}_{\text{ext}} = 0.471$; AMPs against *E. coli*: $r^2_{\text{pred}} = 0.578$, $\text{SDEP}_{\text{ext}} = 0.435$). The Q^2 for both models are above 0.5, indicating the constructed models could be useful in predicting the antibacterial activity of the peptides. The R^2 in both models are above 0.8, indicating that there is a strong correlation between the experimental and predicted values in the training dataset. The contribution values for the steric field were higher than the electrostatic field in each model, showing that the steric field contributed greatly. Figures 1A and 2A are plots of the training (square dots) and test (circle dots) activity values of the peptides along with the predicted values of the seven novel peptides (rhombus dots) of the AMPs against *S. aureus* and *E. coli*. For both CoMFA models, except for a few outliers, most of the points were uniformly scattered around the curve fitting.

Analysis of CoMSIA models

For the CoMSIA models, there are five force fields, namely electrostatic (E), steric (S), hydrophobic (H), hydrogen bond acceptor field (A), and hydrogen bond donor field (D). In total, there are 31 CoMSIA force field combinations. Through data comparison (Supplementary Tables S2 and S4), the $S + E + H$ fields were selected as the best field of the AMPs against *S. aureus* ($Q^2 = 0.645$, $np = 7$, $R^2 = 0.993$, $SEE = 0.037$, $F = 339.242$). The contribution rates of the steric, electrostatic, and hydrophobic field are 16.4, 32.3, and 51.3%. For the AMPs against *E. coli*, the CoMSIA-EDA model performs the best ($Q^2 = 0.573$, $np = 5$, $R^2 = 0.966$, $SEE = 0.065$, $F = 96.841$). The contribution rate of the electrostatic, hydrogen bond donor, and hydrogen bond acceptor field are 31.6%, 40.1%, and 28.4%, respectively, meaning that the hydrogen bond donor field contributed most in this model. The scatter plots of the CoMSIA model are shown in Figs. 1B and 2B, the alignment of the dots indicates the robustness of both models. The predictive ability of the peptides in the test set for the CoMSIA models were evaluated by the external validation method (AMPs against *S. aureus*: $r^2_{\text{pred}} = 0.888$, $\text{SDEP}_{\text{ext}} = 0.350$. AMPs against *E. coli*: $r^2_{\text{pred}} = 0.623$, $\text{SDEP}_{\text{ext}} = 0.411$).

Analysis of the contour maps

AMPs against *S. aureus* Figure 3A, B shows the contour maps of the steric and electrostatic fields in the CoMFA

Fig. 2 The plots of observed and calculated value of the peptides, together with the predicted value of the seven novel peptides of the antimicrobial peptides against *E. coli*

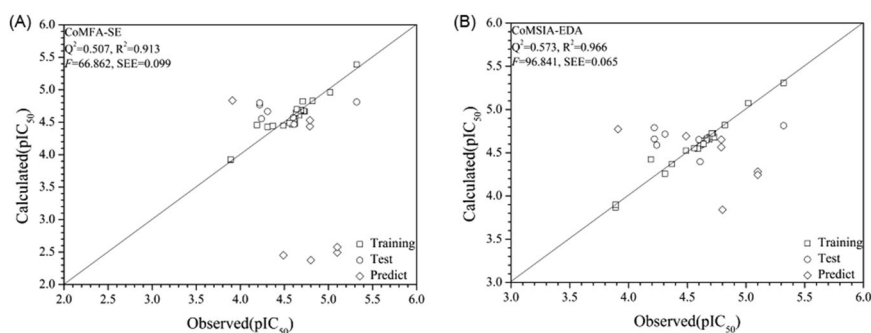
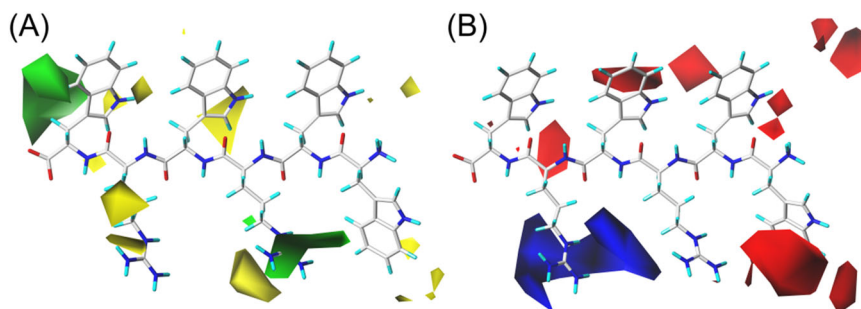


Fig. 3 Contour maps of CoMFA analysis in combination with peptide 1. **A** Steric field. Green contours indicate sterically bulky groups are favorable, yellow contours indicate the opposite. **B** Electrostatic field. Red contours indicate negative charge is favorable, blue contours indicate the opposite



model of AMPs against *S. aureus*. We used peptide 1 as a template to illustrate the three-dimensional contour maps of the CoMFA model. In Fig. 3A, there are yellow contours on the 1-, 3-, 4-, and 5-places of the peptide, suggesting the amino acid residues with small molecule in these regions are favorable to increase the molecule activity. Green contours can be observed on the 6-place of the peptide, indicating regions where increased steric bulk are predicted to enhance the peptide activity. This can be used to explain why peptide 1 (WWRWRW-NH₂, consisting of a tryptophan (W) at the 6-place), has greater activity than peptide 19 (RRWWCN-NH₂), which has an asparagine (N) at its 6-place (the relative molecular mass (MS) of tryptophan (W) is larger than asparagine (N)). In the CoMFA electrostatic contour map (Fig. 3B), red contours occur in the 1-, 2-, and 4-places of the peptide and the 5-place of the peptide chain, showing the interaction can be enhanced by negatively charged substituents at these regions. The blue contours on the 5-place of the peptide indicates that positively charged substituents in that position are favorable.

The steric, electrostatic, hydrophobic fields are used to construct the CoMSIA contours maps (Fig. 4). In Fig. 4A, the yellow contours on the 1-, 4-, and 6-places of the peptide indicate that the amino acid residues with a small molecular weight at these regions are favorable for the peptide activity. The increased peptide activity can also be contributed by large molecular weight on the 5-place of the peptide and the 6-place of the peptide near the peptide chain as shown by the green contours. In Fig. 4B, there are red areas in the N-end and C-end of the peptide chain, and in the 1-place of the peptide,

indicating that negatively charged groups can be added here for increased peptide activity. The blue contours on the 5- and 6-places of the peptide indicate positively charged groups here were favorable. In Fig. 4C, gray areas can be mainly observed on the 1- and 4-places of the peptide, indicating that hydrophilic groups in these areas are favorable for the peptide activity, while light blue region on the 6-place of the peptide means that hydrophobic groups at these regions can contribute to the enhanced activity. The difference in the antimicrobial activity of peptides 1 and 19 can also be explained by the hydrophobic tryptophan group and the hydrophilic asparagine group at the 6-place on these two peptides.

Antimicrobial peptides against *E. coli* Figure 5 shows the contour maps of CoMFA model (using peptide 20 as a template molecule). Figure 5A shows the steric counter map, green areas on the 4- and 6-places of the peptide and the yellow regions observed far from the peptide chain indicate the regions where increased or decreased steric bulk are predicted to enhance the peptide activity. Figure 5B highlights the areas where increased (red region) and decreased (blue region) electron density, respectively, can lead to the increased activity of the peptides. Take training test peptide 18 (RRWWCD-NH₂) and test set peptide 24 (RRRWWW-NH₂) in the *E. coli* set as an example. Aspartic (D) is a negatively charged amino acid and has a relatively smaller molecular weight locates at the 6-place of peptide 18, while tryptophan (W) is positively charged and locates at the same region of peptide 24. In consistent with the observed value, the contour map suggests that peptide 24 has a relatively

Fig. 4 Contour maps of CoMSIA analysis in combination with peptide 1. **A** Steric field. Green contours mean sterically bulky groups is favorable, yellow contours mean the opposite. **B** Electrostatic field. Red contours mean negative charge is favorable, blue contours mean the opposite. **C** Hydrophobic field. Gray contours mean hydrophilic groups is favorable, color light blue means that hydrophobic group was favorable

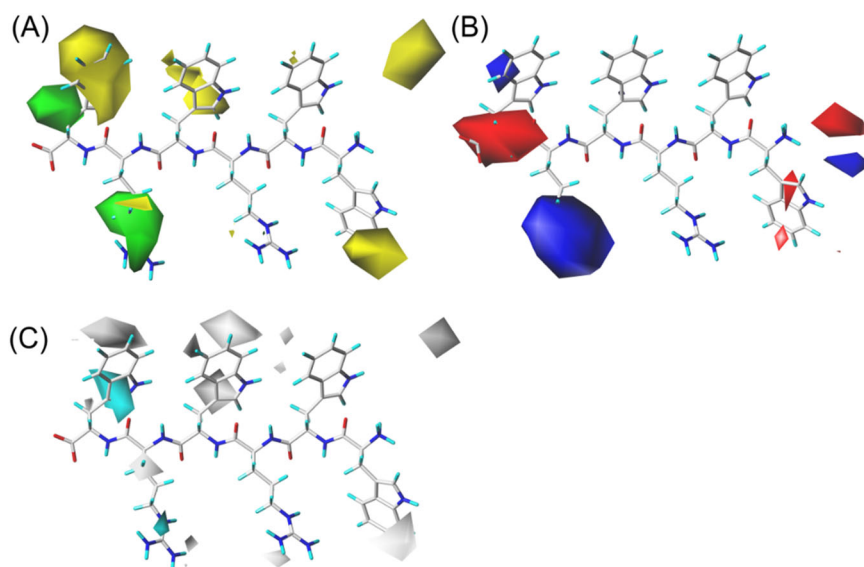


Fig. 5 Contour maps of CoMFA analysis in combination with peptide 20. **A** Steric field. Green contours mean sterically bulky groups is favorable, yellow contours mean the opposite. **B** Electrostatic field. Red contours mean negative charge is favorable, blue contours mean the opposite

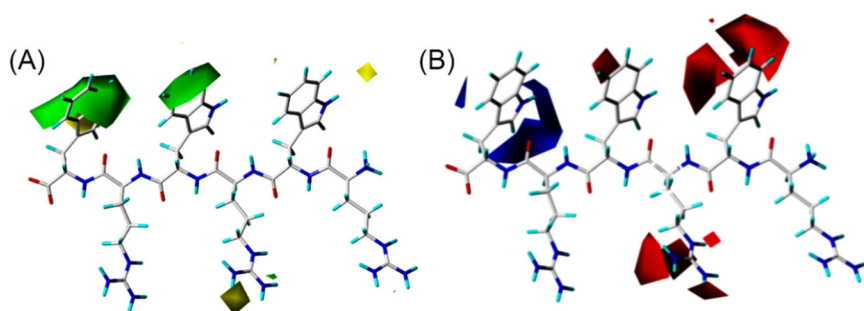
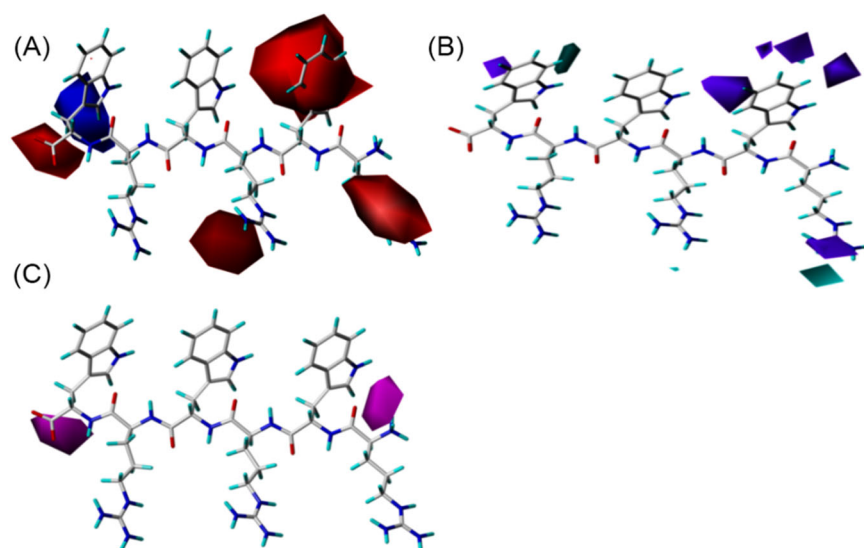


Fig. 6 Contour maps of CoMSIA analysis in combination with compound 20. **A** Electrostatic field. Red contours mean negative charge is favorable, blue contours mean the opposite. **B** Donor field. Purple contours mean hydrogen bond donors are unfavorable, light blue mean the opposite. **C** Acceptor field. Light purple contours mean the hydrogen bond donor is unfavorable



greater antimicrobial activity than peptide 18 because positively charged and large molecular weight groups at the 6-place can increase the activity of the peptide.

Figure 6 was the contour maps of CoMSIA model. In Fig. 6A, there are red areas on the C-end of the peptide chain and blue areas on the 6-place of the peptide,

indicating negatively and positively charged groups, respectively, added here are favorable. Figure 6B shows that in the 1-, 2- and 6-places of the peptide, there were purple areas, describing that hydrogen bond donors were unfavorable at this location. The light blue areas on the 1- and 6-places away from the peptide chain mean that

Table 2 Observed and predicted activities of CoMFA and CoMSIA models of antimicrobial peptides against *S. aureus* and *E. coli*

Antimicrobial peptides against <i>S. aureus</i>					Antimicrobial peptides against <i>E. coli</i>				
No.	Peptide	Obs	Pred		No	Peptide	Obs	Pred.	
			CoMFA	CoMSIA				CoMFA	CoMSIA
Training dataset									
1	WWRWRW-NH ₂	5.33	5.252	5.311	1	WRWRWR-NH ₂	5.02	4.96	5.074
2	RWWWR-NH ₂	5.32	5.378	5.340	2	RRWWCR-NH ₂	4.82	4.828	4.820
3	WRWRWR-NH ₂	5.14	5.215	5.135	3	RRWWCV-NH ₂	4.73	4.665	4.678
4	RWRYRW-NH ₂	5.00	4.948	5.027	4	YWRWRW-NH ₂	4.72	4.67	4.727
5	RRWWCR-NH ₂	4.94	4.957	4.945	5	RRWWCF-NH ₂	4.69	4.679	4.653
6	WWRRRW-NH ₂	5.02	5.011	5.030	6	RRWWCY-NH ₂	4.69	4.69	4.683
7	RRWWCI-NH ₂	4.81	4.823	4.827	7	RRWWCL-NH ₂	4.67	4.607	4.665
8	RRWWCM-NH ₂	4.77	4.703	4.739	8	RRWWCS-NH ₂	4.66	4.668	4.644
9	RRWWCS-NH ₂	4.34	4.309	4.297	9	FRWWHR-NH ₂	4.63	4.644	4.606
10	RRWWCY-NH ₂	4.71	4.777	4.718	10	RRWWCA-NH ₂	4.62	4.475	4.594
11	RRWWCH-NH ₂	4.56	4.601	4.580	11	WWRRRW-NH ₂	4.62	4.551	4.636
12	RRRWWW-NH ₂	4.53	4.532	4.530	12	RRWWCH-NH ₂	4.61	4.515	4.585
13	RRWQWR-NH ₂	4.52	4.467	4.523	13	RRWWCM-NH ₂	4.59	4.463	4.546
14	RRWWCA-NH ₂	4.39	4.302	4.385	14	RRWWCT-NH ₂	4.56	4.49	4.554
15	RRWWCT-NH ₂	4.35	4.330	4.368	15	RRWWCI-NH ₂	4.49	4.449	4.524
16	RRWWCP-NH ₂	4.30	4.266	4.297	16	RRWWCN-NH ₂	4.31	4.428	4.258
17	RRWWCG-NH ₂	4.28	4.222	4.268	17	RRWWCC-NH ₂	4.19	4.455	4.424
18	RRWWCQ-NH ₂	4.11	4.157	4.094	18	RRWWCD-NH ₂	3.89	3.921	3.867
19	RRWWCN-NH ₂	3.94	4.259	3.923	19	RRWWCE-NH ₂	3.89	3.928	3.898
20	RRWWCW-NH ₂	4.39	4.275	4.387	20	RWRWRW-NH ₂	5.32	5.388	5.306
21	YWRWRW-NH ₂	5.02	5.044	4.979	21	RRWWCK-NH ₂	4.71	4.822	4.726
22	RRWWCL-NH ₂	4.84	4.790	4.784	22	RKKWFW-NH ₂	4.64	4.698	4.602
23	RRWWCV-NH ₂	4.67	4.563	4.689	23	RRWWCQ-NH ₂	4.37	4.444	4.37
24	RRWWCC-NH ₂	4.24	4.338	4.343					
Test dataset									
25	RWRWRW-NH ₂	5.32	5.051	5.021	24	RRRWWW-NH ₂	5.32	4.812	4.814
26	FRWLLF-NH ₂	4.92	4.779	4.901	25	RRWWCG-NH ₂	4.60	4.571	4.653
27	RRWWCF-NH ₂	4.92	4.334	4.597	26	RWRYRW-NH ₂	4.31	4.666	4.716
28	RRWWCK-NH ₂	4.75	4.612	4.534	27	RRWQWR-NH ₂	4.22	4.767	4.66
29	RRWWRF-NH ₂	5.32	4.637	4.843	28	RRWWCP-NH ₂	4.61	4.467	4.396
					29	RRWWRF-NH ₂	4.22	4.798	4.791
					30	RRWWCW-NH ₂	4.24	4.554	4.589

hydrogen bond donors were favorable at these regions. In Fig. 6C, light purple regions can be observed at the C-end and N-end of the peptide chain, showing that hydrogen bond acceptors here are beneficial for increasing the peptide activities. Again, for peptides 18 and 24, at the 3- and 5-position of these peptides are hydrogen bond acceptors cysteine (C), and hydrogen bond donor tryptophan (W), respectively. Based on the contour map, we can deduce the fact that peptide 24 has a greater activity than peptide 18.

Synthesis of novel peptides

According to the established QSAR models, we designed and synthesized seven AMPs. After the synthesis and purification of these peptides, we found the molecular weights detected by the relative MS were fully consistent with the theoretical calculations. Tested by high performance liquid chromatography (HPLC), the purities of the synthesized peptides were showed to reach more than 95%. The predicted values, the purities of the seven AMPs, and the

Table 3 Sequence and physicochemical properties of the designed peptides

No.	Sequence	MS		HPLC Purity	Spectrogram
		Calculated	Observed		
D1	RRWWRW-NH ₂	1045.2	1044.4	96.76%	Figure S1
D2	RRWWR-PEA	962.2	962.4	96.28%	Figure S2
D3	RRWWRW-PEA	1148.4	1148.4	95.70%	Figure S3
D4	RRFFRF-PEA	1031.3	1030.8	98.79%	Figure S4
D5	RR(BIP)(BIP)R-PEA	1034.3	1035.0	98.44%	Figure S5
D6	FRWWQR-NH ₂	978.1	977.6	99.5%	Figure S6
D7	RRQFW-NH ₂	978.1	977.4	96.32%	Figure S7

Table 4 Antibacterial activity of antimicrobial peptides

Bacterial	Minimal inhibitory concentrations MIC ($\mu\text{g/mL}$)							
	D1	D2	D3	D4	D5	D6	D7	IPM
<i>S. aureus</i>	64	64	16	8	8	16	16	<0.5
<i>E. coli</i>	64	64	16	32	16	16	16	4

HPLC and MS spectra are shown in Tables 2–3 and Supplementary Figs. S1–S7.

Antimicrobial activity of the novel peptides

The in vitro antimicrobial properties of seven novel peptides were evaluated against *S. aureus* and *E. coli*. As shown in Table 4, the minimal inhibitory concentration (MIC) values of the newly designed peptides range between 8.0 and 64 $\mu\text{g/mL}$. All the designed peptides showed potential antimicrobial activities. Observing Table 5, it can be found that the MIC values of the newly designed AMPs are somewhat greater than those of the natural peptides collected in the literature and some are not. Part of the reason for this phenomenon is that when designing peptides, this article not only considers the antibacterial activity, but also considers to a certain extent the shortcomings of reducing the hemolytic toxicity of the natural AMPs in the original literature and improving the structural instability of the original literature.

Furthermore, we analyzed the relationship between structure and antibacterial activity of these peptides. The structures as well as the observed and predicted pIC_{50} values of the novel peptides are shown in Table 5. From this table, we found that there was a little difference between the observed and calculated value of these seven peptides for both models. The observed and predicted CoMFA values of D2–D5 are quite different. The reason may be that we introduced the phenethylamine at the C-terminus when designing the peptide, and the peptides

used to construct the model were all natural peptides, and there was no peptide containing the phenethylamine in the training set. It may be based on this that the result of poor forecasting effect is caused. However, the natural peptides we designed still fit the model very well. Moreover, although the prediction result of CoMFA is not very well, the result of CoMSIA is quite reasonable. Forasmuch, these results indicate that the two models were well constructed, providing theoretical basis and support for the design of new AMPs.

Plasma stability assay of the synthesized AMPs

A serious disadvantage of AMPs, which hinders their practical application, is that they are susceptible to be degraded. Based on the results of plasma stability test, three novel peptides with better effects and representativeness (D1–D3) were selected for experimental research. The relative content of residual AMPs in plasma at different time periods was shown in Fig. 7. The residual percentage of AMP RRWWRW-NH₂ in plasma at 15, 30, 60, and 90 min are 81.6%, 65.3%, 49.7%, and 44.3% respectively, while the residuals of peptides D2 and D3 in plasma at 90 min are 97.4 and 97.6%. The results indicate that peptide D1 lacks stability in plasma. However, more than 97% of peptides D2 and D3 remained stable after 90 min of plasma incubation, indicating that these two peptides do not easily degrade in plasma and can maintain a long-term stability to exert their antimicrobial effect. Analyzing the structure of these three AMPs, we found that the two peptides (D2 and D3) with higher stability are C-terminal phenylethylamine derivatives of AMP D1. Therefore, this study speculates that for antibacterial peptides like peptide D1, C-terminal phenylethylamine can be considered in the later structural design and modification to improve its stability in plasma.

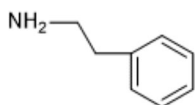
Hemolytic toxicity test of the synthesized AMPs

The cytotoxicity of AMPs in normal mammalian cells is another important factor that limits their clinical application. The cationic AMP works by electrostatically attracting the negatively charged phosphatidylglycerol and cardiolipin in the bacterial cell membrane to destroy the bacterial cell membrane structure [29]. However, the cytomembrane of eukaryotic cell also contains a small amount of negatively charged phospholipids (such as phosphatidylserine and phosphatidylinositol), which may cause the combination of AMPs and eukaryotic cell membranes and thus causing cytotoxicity [30]. To determine the cytotoxicity of these novel peptides, the hemolytic activity of these peptides against sheep red blood cells was measured. As shown in Fig. 8, peptides D1 and D2 were not toxic to red blood cells. The hemolytic activity of red

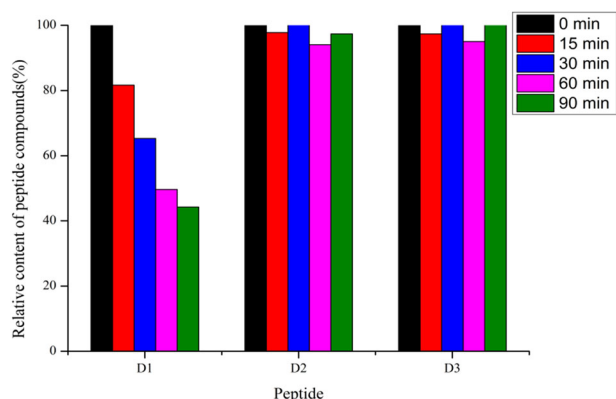
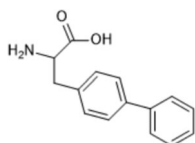
Table 5 The structure and observed/predicted pIC₅₀ of the novel peptides

Antimicrobial peptides against <i>S. aureus</i>					Antimicrobial peptides against <i>E. coli</i>		
No.	Peptide	Obs	Pred.		Obs.	Pred.	
			CoMFA	CoMSIA		CoMFA	CoMSIA
D1	RRWWRW-NH ₂	4.51	4.594	4.690	3.91	4.835	4.774
D2	RRWWR-PEA ^a	4.49	3.539	5.234	4.49	2.451	4.692
D3	RRWWRW-PEA ^a	5.10	3.586	3.947	5.10	2.493	4.283
D4	RRFFRF-PEA ^a	5.40	3.528	5.240	4.80	2.376	3.841
D5	RR(BIP) ^b (BIP)R-PEA ^a	5.40	3.157	5.530	5.10	2.576	4.244
D6	FRWWQR-NH ₂	4.79	5.054	5.115	4.79	4.437	4.565
D7	RRQFW-NH ₂	4.79	4.118	4.238	4.79	4.53	4.653

^aPEA (β -phenylethylamine)



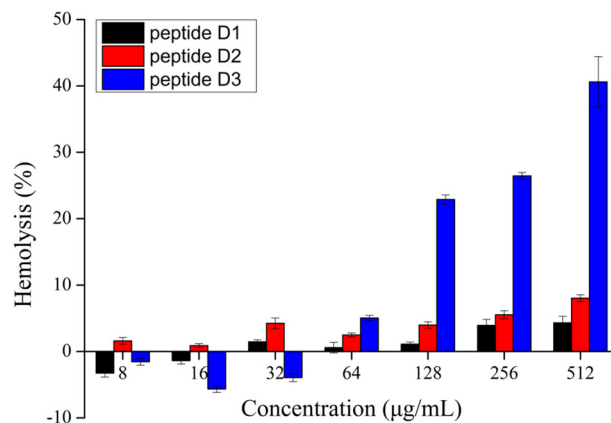
^bBIP

**Fig. 7** Stability of antimicrobial peptides in sheep plasma

blood cells at the concentration of 512 $\mu\text{g}/\text{mL}$ are 2.50 and 4.33% ($<10\%$). For peptide D3, hemolytic activity of red blood cells at the concentration of 64 $\mu\text{g}/\text{mL}$ is 5.05%. However, when the concentration reaches 128 $\mu\text{g}/\text{mL}$, the hemolytic activity is 22.88%. Further modifications should be made on this peptide to decrease host toxicity.

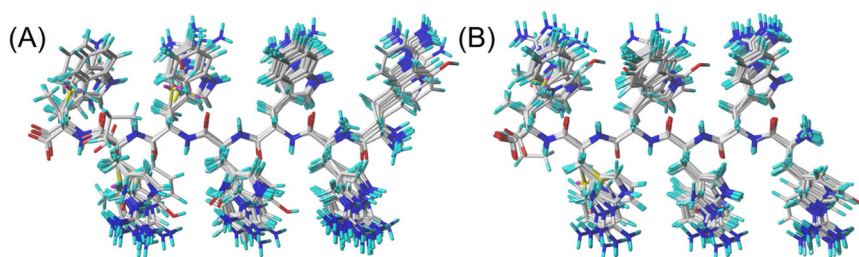
Conclusion

In this study, SYBYL 2.1 was used to establish the CoMFA and CoMSIA models to study the 3D-QSAR of

**Fig. 8** Hemolytic toxicity of antimicrobial peptides in vitro

29 AMPs against *S. aureus* and 30 AMPs against *E. coli* extracted from the literature. The Q^2 and R^2 values indicate that these models are well established for both two AMPs. The contour maps of the model visually reflect the structure–activity relationship of peptides. Based on these models, we designed and synthesized seven AMPs and predicted their activity values. We then experimentally validated that these seven peptides have certain antibacterial activity against gram-positive bacterium (*S. aureus*) and gram-negative bacterium (*E. coli*) with good purity, stability, and low toxicity. Our results indicate that the well-constructed 3D-QSAR models can

Fig. 9 The alignments of antimicrobial peptide.
A Antimicrobial peptides against *S. aureus*.
B Antimicrobial peptides against *E. coli*



provide important theoretical basis for the design, modification and synthesis of the new AMPs.

Materials and methods

Data sources

Molecular structures of 29 AMPs against *S. aureus* and 30 AMPs against *E. coli* and their MIC were extracted from literatures [29, 31–42]. The inhibitory effect of these peptides on *S. aureus* and *E. coli* were expressed as pIC_{50} ($\text{pIC}_{50} = -\log\text{IC}_{50}$). The dataset was randomly divided into training sets (24 AMPs against *S. aureus* and 23 AMPs against *E. coli*) and test sets (5 AMPs against *S. aureus* and seven AMPs against *E. coli*).

Molecular construction and optimization

The 3D structures of all peptides in the training and test sets were constructed using SYBYL 2.1 The Gasteiger–Hückel charge [43] was used to calculate the peptides' charges. The energy minimizations were conducted using the Tripos force field [44] with the max iterations of 1000 and the gradient was 0.005 kcal/ (mol Å). Conformation with the lowest energy was selected as the active conformation. The “align database” command in SYBYL 2.1 was used for superimposing the collected AMPs. The optimized peptides with the maximum activity (lowest energy) were selected as the template for superimposition. The alignments of AMPs are shown in Fig. 9.

CoMFA and CoMSIA modeling

As classic methods, CoMFA and CoMSIA models are widely used in 3D-QSAR studies. CoMFA and CoMSIA models can reflect the activity of the compounds through two fields (electrostatic and steric field) [45] and five fields (electrostatic, steric, hydrophobic, hydrogen bond acceptor field, and donor field) [46], respectively. The partial least square (PLS) models [47] were derived to analyze the extension of the multiple regression. Cross-validation was performed by the leave-one-out method (LOO) [48] to

calculate Q^2 and get the optimum number of components (np). The non-cross-validated correlation coefficient (R^2), F values, and error of estimate (SEE) of the model were calculated to evaluate the reliability and predictivity ability of the models [49]. The external prediction ability of the model was evaluated by the predicted r^2 ($r^2_{\text{pred}} > 0.5$) and external standard deviation error of prediction (SDEP_{ext}) using the following equations [50]:

$$r^2_{\text{pred}} = 1 - \frac{\sum_1^{n_{\text{ext}}} (y_i - \hat{y}_i)^2}{\sum_1^{n_{\text{ext}}} (y_i - \bar{y}_i)^2} \quad (1)$$

$$\text{SDEP} = \sqrt{\frac{\sum_1^{n_{\text{ext}}} (y_i - \hat{y}_i)^2}{n_{\text{ext}}}} \quad (2)$$

where y_i and \hat{y}_i represent the observed and calculated values, \bar{y}_i is the average of the observed value in the training set, and n_{ext} represents the number of the test set.

Synthesis of novel AMPs

After the sequence of bovine lactoferrin is optimized by design, LfcinBG_{4,9} (lactoferrin B_{4,9}, RRWQWR) contains only six natural amino acid residues, which not only has higher antibacterial activity but also has low hemolysis. It is an ideal template for the molecular design of AMPs. This research is based on the AMP design theory obtained by the previous research group, which is “short peptide chain, high charge, strong amphipathic, and more dominant amino acids,” using LfcinBG_{4,9} as a template, and collecting cationic AMPs from the professional AMP database APD2 sequence and antibacterial activity. Combined with some parameters of AMPs, including polar angle, specific secondary structure, amphiphilicity, hydrophobic moment, net charge number, hydrophobicity [51], etc., using bioinformatics and three-dimensional structure–activity relationship analysis methods, a series of rational designs are obtained, antibacterial hexapeptide. In addition, there are reports in the literature that the introduction of phenethylamine can enhance the binding stability of peptides on the one hand [52–54]. Therefore, when designing the structure in this article, we try to combine phenethylamine with peptides. Based on the above reasons and according to the models, seven new peptides were designed.

The novel peptides were synthesized under the solid-phase synthesis method as described [55]. Briefly, the dichloro resin was taken as the carrier in general, wherein halogen chlorine stays at the active site. The resin needs to be dissolved first. Then, the C-end carboxyl of the first amino acid reacts with the active site chlorine on resin. After the first amino acid was connected to resin, it is connected to the second amino acid after dehydration condensation. After that, Fmoc protection was conducted. Operations were repeated according to the designed amino acid sequence, the rest amino acids were connected in sequence, and acetylation of N-end was completed. Finally, the polypeptide was cut from resin with a cutting reagent [56, 57] and the naked carboxyl was formed.

Antimicrobial activity assay

MIC of each peptide against gram-positive bacterium (*S. aureus*) and gram-negative bacterium (*E. coli*) was determined using the broth microdilution assay as described with slight modification [58]. Briefly, mid-logarithmic phase cells were diluted to 2.0×10^5 CFU/mL in Mueller-Hinton Broth growth medium. Fifty microliters of the diluted cell suspension were mixed in 96-well plates with 50 μ L peptide in PBS solution at different stock concentrations (2–512 μ g/mL). The suspensions were then incubated at 37 °C for 12 h. The growth of bacteria was determined by measuring the absorbance at 600 nm using a microplate reader. MIC was defined as the lowest concentration of investigated peptide that completely inhibited bacteria growth.

Hemolysis assay

The hemolytic activity of each peptide was determined as described with slight modification [59]. Briefly, fresh sheep RBCs were washed three times with normal saline, then resuspended into the 3% red cell suspension. One hundred microliters of sheep red blood cell suspension was incubated with 100 μ L peptide solutions at different concentrations ranging from 2 to 512 μ g/mL. Sheep red blood cells suspended in normal saline alone were used as negative control, while cells lysed with 0.1% Triton-X100 were taken as positive control. After incubation for 0.5 h at 37 °C, the suspension was centrifuged at 3000 rpm for 10 min. One hundred microliters of supernatant was added to 96-well plates and absorbance was recorded at 570 nm. The experiment was repeated three times and the hemolysis ratio was an average value based on the result of three repeats. Hemolysis ratio [60] = $[(OD_{\text{test hole}} - OD_{\text{negative hole}})/(OD_{\text{positive hole}} - OD_{\text{negative hole}})] \times 100\%$.

Plasma stability assay

Twenty-five percent of sheep plasma was incubated at 37 °C for 30 min. Two hundred and fifty microliters of sheep plasma was mixed with 50 μ L peptide solution at a concentration of 1 mg/mL. The mixture was incubated in a biochemical incubator, shaking at 100 rpm at 37 °C. After incubation for 0, 10, 30, 60, and 90 min, 200 μ L TFA was added to stop the reaction of peptide in plasma. The mixture was cooled for 30 min at 4 °C and then centrifuged at 1200 rpm for 30 min. Two hundred microliters of supernatant was extracted and analyzed using HPLC-QqQ-MS [61]. The 135 V electrospray ionization source was used for scanning.

Acknowledgements This work was supported by the National Natural Science Foundation of China (31400667), Scientific and Technological Research Program of Chongqing Municipal Education Commission (KJ1400932, KJ1500902, KJ1600908), Chongqing Research Program of Basic Research and Frontier Technology (cstc2014jcyjA10044, cstc2013jcyjA10019), and Program for Innovation Team Building at Institutions of Higher Education in Chongqing (CXTDX201601031). The authors would like to acknowledge the funding support to the Xie Laboratory from the NIH NIDA (P30 DA035778A1). The authors would like to thank Li Yang for helping to revise the manuscript in English.

Compliance with ethical standards

Conflict of interest The authors declare no competing interests.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Kümmerer K. Antibiotics in the aquatic environment—a review—part I. *Chemosphere*. 2009;75:417–34. <https://doi.org/10.1016/j.chemosphere.2008.11.086>.
- Levy SB, Marshall B. Antibacterial resistance worldwide: causes, challenges and responses. *Nat Med Suppl*. 2007;10:S122.
- Murray BE. Vancomycin-resistant enterococcal infections. *N Engl J Med*. 2000;342:710–21. <https://doi.org/10.1056/nejm200003093421007>.
- Panlilio AL, Culver DH, Gaynes RP, Banerjee S, Henderson TS, Tolson JS, et al. Methicillin-resistant *Staphylococcus aureus* in U.S. hospitals, 1975–1991. *Infect Control Hosp Epidemiol*. 1992;13:582–6. <https://doi.org/10.2307/30148460>.
- Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*. 2004;60:2126–32. <https://doi.org/10.1107/S0907444904019158>.
- Yong DG, Toleman MA, Giske CG, Cho HS, Sundman K, Kyungwon L, et al. Characterization of a new metallo- β -lactamase gene, blaNDM-1, and a novel erythromycin esterase gene carried on a unique genetic structure in *Klebsiella pneumoniae* sequence type 14 from India. *Antimicrob Agents Chemother*. 2009;53:5046–54.
- Sabrina R, Dagmar Z. Membrane-active host defense peptides—challenges and perspectives for the development of novel anticancer drugs. *Chem Phys Lipids*. 2011;164:766–81.
- Koczulla AR, Bals DR. Antimicrobial peptides. *Drugs*. 2003;63:389–406.

9. Bakula M. Antibacterial compounds in the cell-free haemolymph of *Drosophila melanogaster*. *J Insect Physiol.* 1970;16:185–97.
10. Hink WF, Briggs JD. Bactericidal factors in haemolymph from normal and immune wax moth larvae, *Galleria mellonella*. *J Insect Physiol.* 1968;14:1025–34.
11. Steiner H, Hultmark D, Engström A, Bennich H, Boman HG. Sequence and specificity of two antibacterial proteins involved in insect immunity. *Nature* 292: 246–248. 1981. *J Immunol.* 2009;182:6635.
12. Samakovlis C, Kylsten P, Kimbrell DA, Engström A, Hultmark D. The andropin gene and its product, a male-specific antibacterial peptide in *Drosophila melanogaster*. *EMBO J.* 1991;10:163.
13. Bulet P, Dimarcq JL, Hetru C, Lagueux M, Charlet M, Hegy G, et al. A novel inducible antibacterial peptide of *Drosophila* carries an O-glycosylated substitution. *J Biol Chem.* 1993;268:14893–7.
14. Casteels P, Ampe C, Riviere L, Damme JV, Elicone C, Fleming M, et al. Isolation and characterization of abaecin, a major antibacterial response peptide in the honeybee (*Apis mellifera*). *FEBS J.* 1990;187:381.
15. Casteels P, Ampe C, Jacobs F, Vaeck M, Tempst P. Apidaecins: antibacterial peptides from honeybees. *EMBO J.* 1989;8:2387.
16. Rameshkumar G, Ravichandran S, Kaliyavarathan G, Ajithkumar T. Antimicrobial peptide from the crab, *Thalamita crenata* (Latreille, 1829). *World J Fish Mar Sci.* 2009;1:74–9.
17. Lehrer RI, Ganz T. Antimicrobial peptides in mammalian and insect host defence. *Curr Opin Immunol.* 1999;11:23–7. [https://doi.org/10.1016/S0952-7915\(99\)80005-3](https://doi.org/10.1016/S0952-7915(99)80005-3).
18. Y P, KS H. Antimicrobial peptides (AMPs): peptide structure and mode of action. *J Biochem Mol Biol.* 2005;38:507.
19. Da CN, Cobacho NB, Viana JF, Lima LA, Sampaio KB, Dohms SS, et al. The next generation of antimicrobial peptides (AMPs) as molecular therapeutic tools for the treatment of diseases with social and economic impacts. *Drug Discov Today.* 2017;22:234–48.
20. Gautam A, Sharma A, Jaiswal S, Fatma S, Arora V, Iquebal MA, et al. Development of antimicrobial peptide prediction tool for aquaculture industries. *Probiotics Antimicrob Proteins.* 2016;8:1–9.
21. Fjell C, Jenssen H, Hilpert K, Cheung W, Pante N, Hancock R, et al. Identification of novel antibacterial peptides by chemoinformatics and machine learning. *J Med Chem.* 2009;52:2006–15.
22. Liu S, Fan L, Sun J, Lao X, Zheng H. Computational resources and tools for antimicrobial peptides. *J Pept Sci.* 2017;23:4–12.
23. Bolintineanu DS, Kaznessis YN. Computational studies of protegrin antimicrobial peptides: a review. *Peptides.* 2011;32:188–201. <https://doi.org/10.1016/j.peptides.2010.10.006>.
24. Tong JB, Jiang GY, Yuan-Yuan LI, Kang-Nan LI. 3D-QSAR study on antimicrobial peptides by CoMFA method. *J At Mol Phys.* 2017;34(06),990–996.
25. Chen B, Zhang T, Bond T, Gan Y. Development of quantitative structure activity relationship (QSAR) model for disinfection byproduct (DBP) research: a review of methods and resources. *J Hazard Mater.* 2015;299:260–79. <https://doi.org/10.1016/j.jhazmat.2015.06.054>.
26. Jitender V, Vijay MK, Evans CC. 3D-QSAR in drug design—a review. *Curr Topics Med Chem.* 2010;10:95–115. <https://doi.org/10.2174/156802610790232260>.
27. Liu S, Bao J, Lao X, Zheng H. Novel 3D structure based model for activity prediction and design of antimicrobial peptides. *Sci Rep.* 2018;8:11189.
28. Bordás B, Kómvics T, Lopata A. Ligand-based computer-aided pesticide design. A review of applications of the CoMFA and CoMSIA methodologies. *Pest Manag Sci.* 2003;59:393–400. <https://doi.org/10.1002/ps.614>.
29. Hancock RE. Cationic peptides: effectors in innate immunity and novel antimicrobials. *Lancet Infect Dis.* 2001;1:156.
30. Huang Y, Huang J, Chen Y. Alpha-helical cationic antimicrobial peptides: relationships of structure and function. *Protein Cell.* 2010;1:143–52. <https://doi.org/10.1007/s13238-010-0004-3>.
31. Sundriyal S, Sharma RK, Jain R, Bharatam PV. Minimum requirements of hydrophobic and hydrophilic features in cationic peptide antibiotics (CPAs): pharmacophore generation and validation with cationic steroid antibiotics (CSAs). *J Mol Model.* 2008;14:265–78.
32. Strøm MB, Rekdal O, Svendsen JS. Antimicrobial activity of short arginine- and tryptophan-rich peptides. *J Pept Sci.* 2002;8:431–7.
33. Strøm MB, Haug BE, Skar ML, Stensen W, Stiberg T, Svendsen JS. The pharmacophore of short cationic antibacterial peptides. *J Med Chem.* 2003;46:1567–70.
34. Murugan RN, Jacob B, Kim EH, Ahn M, Sohn H, Seo JH, et al. Non hemolytic short peptidomimetics as a new class of potent and broad-spectrum antimicrobial agents. *Bioorg Med Chem Lett.* 2013;23:4633–6.
35. Albada HB, Prochnow P, Bobersky S, Langklotz S, Schriek P, Bandow JE, et al. Tuning the activity of a short Arg-Trp antimicrobial peptide by lipidation of a C- or N-terminal lysine side-chain. *ACS Med Chem Lett.* 2012;3:980.
36. Liu Z, Brady A, Young A, Rasimick B, Chen K, Zhou C, et al. Length effects in antimicrobial peptides of the (RW)_n series. *Antimicro Agents Chemother.* 2007;51:597–603.
37. Chou S, Shao C, Wang J, Shan A, Xu L, Dong N, et al. Short, multiple-stranded β -hairpin peptides have antimicrobial potency with high selectivity and salt resistance. *Acta Biomater.* 2015;30:78–93.
38. Lee SH, Kim SJ, Lee YS, Song MD, Kim IH, Won HS. De novo generation of short antimicrobial peptides with simple amino acid composition. *Regul Pept.* 2011;166:36–41.
39. Muñoz A, López-García B, Marcos JF. Comparative study of antimicrobial peptides to control citrus postharvest decay caused by *Penicillium digitatum*. *J Agric Food Chem.* 2007;55:8170.
40. Blondelle SE, Lohner K. Combinatorial libraries: a tool to design antimicrobial and antifungal peptide analogues having lytic specificities for structure-activity relationship studies. *Pept Sci.* 2015;55:74–87.
41. Giménez D, Andreu C, MI DO, Varea T, Diaz D, Asensio G. The introduction of fluorine atoms or trifluoromethyl groups in short cationic peptides enhances their antimicrobial activity. *Bioorg Med Chem.* 2006;14:6971–8.
42. Remuzgo C, Oewel TS, Daffre S, Lopes TRS, Dyszy FH, Schreier S, et al. Chemical synthesis, structure–activity relationship, and properties of shepherin I: a fungicidal peptide enriched in glycine-glycine-histidine motifs. *Amino Acids.* 2014;46:2573–86.
43. Tsai K-C, Chen Y-C, Hsiao N-W, Wang C-L, Lin C-L, Lee Y-C, et al. A comparison of different electrostatic potentials on prediction accuracy in CoMFA and CoMSIA studies. *Eur J Med Chem.* 2010;45:1544–51. <https://doi.org/10.1016/j.ejmech.2009.12.063>.
44. Clark M, Cramer RD III, Van Opdenbosch N. Validation of the general purpose Tripos 5.2 force field. *J Comput Chem.* 1989;10:982–1012.
45. Kim KH, Greco G, Novellino E. A critical review of recent CoMFA applications. *Perspectives in Drug Discovery & Design.* 1998;12–14:257–315.
46. Ghaleb A, Aouidate A, Ghamali M, Sbair A, Bouachrine M, Lakhlifi T. 3D-QSAR modeling and molecular docking studies on a series of 2,5 disubstituted 1,3,4-oxadiazoles. *J Mol Struct.* 2017;1145:278–84. <https://doi.org/10.1016/j.molstruc.2017.05.065>.
47. Sawatsky ML, Clyde M, Meek F. Partial least squares regression in the social sciences. *Mathematics.* Vol. 6. 2003. p. 792–5.
48. Burman P, Chow E, Nolan D. A cross-validated method for dependent data. *Biometrika.* 1994;81:351–8. <https://doi.org/10.1093/biomet/81.2.351>.
49. Depriest SA, Mayer D, Naylor CB, Marshall GR. 3D-QSAR of angiotensin-converting enzyme and thermolysin inhibitors: a comparison of CoMFA models based on deduced and

- experimentally determined active site geometries. *J Am Chem Soc.* 1993;115:5372–84.
50. Tosco P, Balle T. A 3D-QSAR-driven approach to binding mode and affinity prediction. *J Chem Inf Model.* 2012;52:302–7. <https://doi.org/10.1021/ci200411s>.
51. Gudmundsson GH, Lidholm DA, Asling B, Gan R, B, Boman HG. The cecropin locus. Cloning and expression of a gene cluster encoding three antibacterial peptides in *Hyalophora cecropia*. *JBC.* 1991;266:11510–7.
52. Anderson MA, Ogbay B, Kisselev OG, Cistola DP, Marshall GR. Alternate binding mode of C-terminal phenethylamine analogs of G α (340–350) to photoactivated rhodopsin. *Chem Biol Drug Res.* 2006;68:295–307.
53. Nilsson AC, Janson H, Wold H, Fugelli A, Andersson K, Håkangård C. et al. LTX-109 is a novel agent for nasal decolonization of methicillin-resistant and -sensitive *Staphylococcus aureus*. *Antimicrob Agents Chemother.* 2014;59:145–51.
54. Torsten, W, Margitta, D, Michael, B, Krause E, Maloy WL, MacDonald DL, et al. Peptide hydrophobicity controls the activity and selectivity of magainin 2 amide in interaction with membranes. *Biochemistry.* 1997;36:6124–32.
55. Chan WC, White PD. Fmoc solid phase peptide synthesis: a practical approach. Oxford University Press. vol. 222. (ResearchGate, 2000).
56. Montanya E. A comparison of currently available GLP-1 receptor agonists for the treatment of type 2 diabetes. *Expert Opin Pharmacother.* 2012;13:1451–67.
57. Bechinger B, Gorr SU. Antimicrobial peptides: mechanisms of action and resistance. *J Dent Res.* 2017;96:0022034516679973.
58. Weseler A, Geiss HK, Saller R, Reichling. A novel colorimetric broth microdilution method to determine the minimum inhibitory concentration (MIC) of antibiotics and essential oils against *Helicobacter pylori*. *Pharmazie.* 2005;60(7):498–502.
59. Gauthier C, Legault J, Girard-Lalancette K, Mshvildadze V, Pichette AJB. Haemolytic activity, cytotoxicity and membrane cell permeabilization of semi-synthetic and natural lupane- and oleanane-type saponins. *Bioorg Med Chem.* 2009;17:2002–8.
60. Hou F, Li J, Pan P, Xu J, Liu L, Liu W, et al. Isolation and characterisation of a new antimicrobial peptide from the skin of *Xenopus laevis*. *Int J Antimicrob Agents.* 2011;38:510–5.
61. Yost RA, Enke CG. Triple quadrupole mass spectrometry for direct mixture analysis and structure elucidation. *Anal Chem.* 1979;51:1251–64. <https://doi.org/10.1021/ac50048a002>.