Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology)
ISSN 1673-1581 (Print); ISSN 1862-1783 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Quantitative structure-activity relationships of antimicrobial fatty acids and derivatives against *Staphylococcus aureus**

Hui ZHANG[§1], Lu ZHANG[§1], Li-juan PENG[1], Xiao-wu DONG[2], Di WU[1], Vivian Chi-Hua WU[3], Feng-qin FENG[†‡1]

([1]*Department of Food Science and Nutrition, Zhejiang University, Hangzhou 310058, China*)

([2]*ZJU-ENS Joint Laboratory of Medicinal Chemistry, College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310058, China*)

([3]*Department of Food Science and Human Nutrition, University of Maine, Orono 04469, USA*)

[†]E-mail: fengfq@zju.edu.cn

Received Feb. 20, 2011; Revision accepted July 27, 2011; Crosschecked Dec. 7, 2011

Abstract: Fatty acids and derivatives (FADs) are resources for natural antimicrobials. In order to screen for additional potent antimicrobial agents, the antimicrobial activities of FADs against *Staphylococcus aureus* were examined using a microplate assay. Monoglycerides of fatty acids were the most potent class of fatty acids, among which monotridecanoin possessed the most potent antimicrobial activity. The conventional quantitative structure-activity relationship (QSAR) and comparative molecular field analysis (CoMFA) were performed to establish two statistically reliable models (conventional QSAR: R^2=0.942, Q^2_{LOO}=0.910; CoMFA: R^2=0.979, Q^2=0.588, respectively). Improved forecasting can be achieved by the combination of these two models that provide a good insight into the structure-activity relationships of the FADs and that may be useful to design new FADs as antimicrobial agents.

Key words: Fatty acid derivatives, Quantitative structure-activity relationship, Comparative molecular field analysis, Antimicrobial activity

doi:10.1631/jzus.B1100049 **Document code:** A **CLC number:** TS221

1 Introduction

Research into alternative antimicrobial agents has become important because of consumer's demand for fresh, natural, and safe foods. Fatty acids and derivatives (FADs) are one group of chemicals found in nature and are considered to have little or no toxicity in low concentrations. They have been proven to exert antimicrobial activity and have attracted much attention during the past 40 years (Bergsson *et al.*, 2001; Kelsey *et al.*, 2006; Habulin *et al.*, 2008).

To date, there have been a limited number of reports on the relationship between structure and antimicrobial activity of FADs. From 1972 to 1984, the group of Kabara examined the antimicrobial activities of a series of FADs, and found that the activities might be influenced by structural properties, such as carbon chain length, unsaturation, location of unsaturated bonds, hydroxide radical, esterification, and functional groups (Kabara *et al.*, 1972; 1977; Kabara, 1984). Hou and Forman (2000) studied the growth inhibition of plant pathogenic fungi by several hydroxy fatty acids, and assumed that the positions of hydroxy groups on the fatty acids played an important role in their antifungal activities. Nobmann *et al.* (2009) investigated the antimicrobial efficacies of monosubstituted carbohydrate fatty acid esters and ethers against *Listeria* spp. and some other foodborne microorganisms, and suggested that α and β configurations of the carbohydrate moiety and the nature of the bond that connects to fatty acids have a significant effect on efficacy.

The group of Siebert was the first to try the

‡ Corresponding author
§ The two authors contributed equally to this work
* Project (No. 31071501) supported by the National Natural Science Foundation of China

quantitative structure-activity relationship (QSAR) method to study the structure-activity relationship of fatty acids (Hsiao and Siebert, 1999; Nakai and Siebert, 2003; 2004). They applied principal component analysis (PCA) to 11 properties of 17 organic acids (including 10 fatty acids) commonly used in food systems, and they successfully constructed models for 10 bacteria. This initial success suggested the possibility of the existence of linear relationships between the structure and the antimicrobial activity of organic acids. In the present work, 57 different FADs were collected and the structure-activity relationship against *Staphylococcus aureus* CMCC (B) 26003, a typical Gram-positive methicillin-resistant strain, was examined using two QSAR methods: conventional QSAR analysis (using three-dimensional (3D) structure descriptors) and comparative molecular field analysis (CoMFA) in order to establish predictive models, as an important guideline for the molecular design of new FAD antimicrobials against methicillin-resistant *S. aureus* strains.

2 Materials and methods

2.1 Chemicals

All FADs used were of the highest purity grade (above 99%), were commercially obtained from Sigma, and used without further purification. Double-distilled water was used in all experiments.

2.2 Bacteria

Staphylococcus aureus CMCC (B) 26003 was provided by the Institute of Microbiology, Chinese Academy of Sciences, Beijing, China. The strain was maintained on nutrient agar (NA; Hangzhou Microbiological Agents Co., Ltd., China) slants and transferred each month to maintain viability at the Department of Food Science and Nutrition, Zhejiang University, Hangzhou, China. Working cultures were obtained by inoculating a loop of pure culture into tryptic soy broth (TSB, pH 7.2) medium and incubating at 37 °C for 24 h.

2.3 Antimicrobial activity assay

The experiment was designed according to the methods of Branen and Davidson (2004) with modifications. FADs were dissolved in 100% ethanol, diluted by TSB to obtain various concentrations (2, 4, 8, 16, 32, 63, 125, 250, 500, 1000, 2000, and 4000 μg/ml), and dispensed into the wells of a sealed flat-bottom 96-well microtiter plate. Overnight cultures of the test organisms were diluted to approximately 10^4 colony-forming units/ml (CFU/ml) in TSB broth (pH 7.2) and 100 μl volumes were added to each well of this sealed microtiter plate. The experiments were conducted in duplicate and controls were included as appropriate. The 96-well plates were incubated for 24 h at 37 °C, and the optical density at 595 nm (OD_{595}) for 0 and 24 h of the culture was read on a Multiskan MK3 microplate reader (Multiskan MK3, Thermo Labsystems, Finland). Minimum inhibitory concentrations (MICs) at 24 h were defined as the lowest concentration at which the bacterial growth was completely inhibited (OD_{595} difference <0.05). MIC$\geq$4000 μg/ml was defined as no antimicrobial activity.

2.4 Data sets for analysis

To cover the potential range as widely as possible, 57 compounds including fatty acids and their monoglycerides, diglycerides, alcohols, methyl/ethyl esters, and acetates were collected. The FADs without antimicrobial activity were not used in QSAR analysis. According to the different molecular structures, those with definite MIC values (37 compounds, listed in Table 1) were classified as fatty acids (Type I), monoglycerides of fatty acids (Type II), and fatty alcohols (Type III), and were randomly divided into a training set (28 compounds) and a test set (9 compounds). Their antimicrobial activities were expressed as −logMIC.

2.5 Conventional QSAR

The 3D structures of the compounds were refined by the semi-empirical quantum chemical method AM1 implemented in MOPAC 10.0 computer software (James Stewart, Springs, CO, USA).

The DRAGON Professional software (Talete srl, Milano, Italy) was employed to calculate 3224 structure descriptors which were classified as 22 different kinds of topological descriptors including: edge adjacency indices, Burden eigenvalue descriptors, topological charge indices, eigenvalue-based indices, geometrical descriptors, radial distribution function (RDF) descriptors, 3D molecule representation of

structure based on electron diffraction (3D-MoRSE) descriptors, weighted holistic invariant molecular (WHIM) descriptors, GETAWAY descriptors, 2D binary fingerprints, etc. (Todeschini and Consonni, 2000). Descriptors with constant values were discarded, the remaining number of descriptors in this study was 1315.

The correlation between antimicrobial activity and structure properties was obtained by using the variable selection enhanced replacement method (ERM) (Mercader *et al.*, 2008a; 2008b) and multiple linear regression analysis (MLRA) methods. Compared with genetic algorithm (GA) and forward stepwise methods, the QSAR models obtained by replacement method (RM) showed the best statistical parameters (Gonzalez *et al.*, 2008; Mercader *et al.*, 2010). ERM is an enhanced method based on RM, and it is better for a larger (more than a thousand) pool of descriptors. The selected descriptors were then used to build a QSAR model using MLRA. The predictive stability and robustness of the model were first verified by internal cross-validation (leave-one-out (LOO) method), and then were further checked by external cross-validation. These methods were all implemented in MATLAB 7.6 (the Math Works, Natick, MA, USA).

The performance of the model was evaluated using the following statistical parameters: squared correlation coefficient (R^2), standard error of the estimate (s), significance of the model (P), Fisher ratio value (F), LOO cross-validation coefficient for conventional QSAR model (Q^2_{LOO}), and prediction coefficient (R^2_{EXT}).

2.6 CoMFA

The 3D structures of all compounds were minimized using SYBYL 6.9 (TRIPOS, St. Louis, MO, USA). The geometries were optimized by Powell's method using the Tripos force field. Conformation analysis was also carried out using the SYBYL/GRID search module. The lowest-energy conformations were considered as the bioactive conformations. Compound **20** (monotridecanoin) with the most potent activity (Table 1) was selected as the template of the alignment. Furthermore, the common –COO structure (Fig. 1) was selected to superimpose all the compounds using the SYBYL/fit-atom module. The result of the alignment of the molecules is shown in Fig. 2 (in page 90). To calculate the steric and electrostatic fields in CoMFA, a probe sp3 carbon atom with +1.0 net charge was employed, and the CoMFA fields generated were scaled by the CoMFA-standard (STD) method with an energy level of 30 kcal/mol. The partial least-square (PLS) method was used to set up a correlation between the molecular fields and the −logMIC denoting antimicrobial activities of tested samples. The optimal number of the components was calculated using the LOO cross-validation method. The highest cross-validated Q^2 (an LOO cross-validation coefficient for CoMFA model) that resulted in an optimum number of components was applied in non-cross-validated analysis, in which the final model was built.

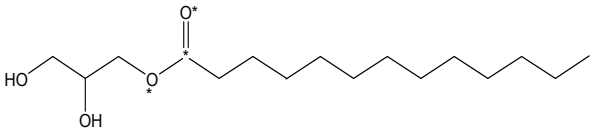


Fig. 1 Template molecule and common atoms (*) used for atom-fit alignment

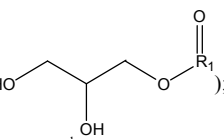
3 Results

3.1 Antimicrobial activities of FADs

The MIC values of all FADs against *S. aureus* were tested in this experiment, and the values of −logMIC are listed in Table 1. Among the FADs, the monoglycerides of fatty acids (Type II) showed the most potent activities, with an average activity of 3.968. In comparison, the average activities of fatty acids (Type I) and fatty alcohols (Type III) were 2.644 and 2.602, respectively. The compounds with −logMIC≥3.500 (with potent antimicrobial activity) were either monoglycerides of fatty acids with a length of 10–14 carbons, or unsaturated fatty acids with a chain length longer than 18 carbons and their monoglyceride counterparts. Linolenic acid was the most potent fatty acid, and monotridecanoin was the most potent monoglyceride. Those FADs without antimicrobial activity (MIC≥4000 μg/ml) were the tested diglycerides, methyl/ethyl esters, acetates, and *trans* fatty acids (monobutyrin, ethyl caprylate, methyl caprate, ethyl caprate, 1,2-dilaurin, methyl laurate, ethyl laurate, lauryl acetate, palmitic acid,

Table 1 Structures and biological activities of fatty acids and their derivatives

Compd.	R_1	Type[a]	Mol. ID	−logMIC				
				Obs.	Pre.[b]	Res.[b]	Pre.[c]	Res.[c]
1	C1:0	I	Formic acid	1.663	2.002	−0.339	1.753	−0.090
2	C2:0	I	Acetic acid	2.080	1.757	0.323	1.983	0.097
3[d]	C3:0	I	Propionic acid	2.171	1.822	0.349	1.992	0.179
4	C4:0	I	Butyric acid	1.945	1.749	0.196	2.015	−0.070
5[d]	C6:0	I	Hexanoic acid	2.366	1.867	0.499	2.256	0.110
6	C6:2 (2, 4)	I	Sorbic acid	2.050	2.290	−0.240	2.038	0.012
7	C8:0	I	Octanoic acid	1.858	2.009	−0.151	2.134	−0.276
8[d]	C8:0	II	Monocaprylin	3.242	3.889	−0.647	3.500	−0.258
9[d]	C10:0	I	Capric acid	2.838	2.162	0.676	2.332	0.506
10	C10:0	II	Monocaprin	3.886	4.082	−0.196	3.899	−0.013
11[d]	C11:0	I	Undecanoic acid	2.571	2.204	0.367	2.402	0.169
12	C11:0	II	Monoundecanoin	4.212	3.969	0.243	4.185	0.027
13	C11:1 (10)	I	10-Undecenoic acid	2.266	2.074	0.192	2.406	−0.140
14[d]	C11:1 (10)	II	Monoundecenoin	3.613	3.354	0.259	3.931	−0.318
15	C12:0	III	Lauryl alcohol	2.270	2.260	0.010	2.302	−0.032
16	C12:0	II	Monolaurin	4.535	4.489	0.046	4.548	−0.013
17	C12:1 (11)	I	11-Dodecenoic acid	2.598	2.473	0.125	2.466	0.132
18	C12:1 (11)	II	Monododecenoin	4.231	4.460	−0.229	4.369	−0.138
19	C13:0	I	Tridecanoic acid	2.933	2.759	0.174	2.629	0.304
20	C13:0	II	Monotridecanoin	4.858	4.722	0.136	4.670	0.188
21[d]	C13:1 (12)	I	12-Tridecenoic acid	2.929	2.315	0.614	2.711	0.218
22	C14:0	I	Myristic acid	2.359	2.706	−0.347	2.543	−0.184
23[d]	C14:0	III	Tetradecanol	2.933	2.433	0.500	2.355	0.578
24[d]	C14:0	II	Monomyristin	4.277	3.832	0.445	4.591	−0.314
25	C14:1 (9)	I	Myristoleic acid	2.957	2.733	0.224	2.861	0.096
26	C16:1 (9)	I	Palmitoleic acid	3.008	3.316	−0.308	2.909	0.099
27	C16:1 (9)	II	Monopalmitolein	3.420	3.378	0.042	3.769	−0.349
28	C18:0	I	Stearic acid	1.852	2.119	−0.267	1.777	0.075
29	C18:1 (9)	I	Oleic acid	3.946	3.430	0.516	3.973	−0.027
30	C18:2 (9, 12)	II	Monolinolein	2.851	2.826	0.025	2.764	0.087
31	C18:3 (9, 12, 15)	I	Linolenic acid	4.241	4.273	−0.032	4.223	0.018
32	C18:3 (9, 12, 15)	II	Monolinolenin	3.748	4.080	−0.332	3.635	0.113
33	C18:3 (6, 9, 12)	I	γ-linolenic acid	3.645	3.659	−0.014	3.762	−0.117
34	C18:3 (6, 9, 12)	II	Mono-γ-linolenin	4.644	4.584	0.060	4.627	0.017
35	C20:4 (5, 8, 11, 14)	II	Monoarachidonin	4.073	3.970	0.103	3.984	0.089
36	C20:5 (5, 8, 11, 14, 17)	I	Eicosapentaenoic acid	3.384	3.141	0.243	3.452	−0.068
37	C22:6 (4, 7, 10, 13, 16, 19)	I	Docosahexaenoic acid	2.517	2.719	−0.202	2.358	0.159

[a] I: fatty acids (); II: monoglyceride of fatty acids (); III: fatty alcohols (). [b] Calculated using the conventional QSAR model. [c] Calculated using the CoMFA model. [d] These compounds were used as a test set and not included in model derivation. Calculation for the conventional QSAR and CoMFA models is based on the −logMIC (MIC: mol/L) value. Compd.: compound; Mol.: molecular; Obs.: observed; Pre.: predictive; Res.: residual

monopalmitin, monostearin, monoolein, elaidic acid, linoleic acid, *trans*-10,*cis*-12 conjugated linoleic acid, *cis*-9,*trans*-11 conjugated linoleic acid, linolelaidic acid, eicosanoic acid, erucic acid, and monoerucin).

3.2 Conventional QSAR model

Based on the five most representative structure descriptors (Table 2) selected by ERM, the following MLRA model was obtained:

$$-\log\mathrm{MIC}=2.918+0.345a-0.326b-3.896c+0.697d+1.239e, \quad (1)$$

where a is RDF030p, b is RDF080p, c is E2p, d is Inflammat-80, and e is nHBonds (N=28, R^2=0.942, s=0.254, F=75.309, P=0.000, Q^2_{LOO} =0.910, R^2_{EXT} = 0.746).

The R^2 value of 0.942 indicates that the inner predictive power of the model was high. The Q^2_{LOO} value, which was also above 0.900, confirms the significant stability of the model. The test set was then applied to verify the external predictive power of the model. The theoretical results of compounds from the test set were in good agreement with the experimental values with R^2_{EXT} at 0.746. The experimental −logMIC values versus the calculated values and the residual versus experimental values obtained from the conventional model are plotted in Fig. 3. It is seen that the residual values distributed uniformly and were lower than three times the standard deviation.

The meanings of the five descriptors are shown in Table 2 and the values of each molecule are listed in Table 3. Table 4 shows the standardized coefficients of the variables, which indicate that the two RDF descriptors contribute mostly to the antimicrobial activity. RDF descriptors are based on the distance distribution of the atoms. The RDF descriptors of a molecule of n atoms could be interpreted as the probability distribution of finding an atom in a spherical volume of radius R. These 3D descriptors suggested the occurrence of the linear dependence between the antimicrobial activity and the 3D molecular distribution of atoms calculated at radius of 3.0 and 8.0 Å from the specific geometrical centers of each molecule (Hemmer *et al.*, 1999). Therefore, this

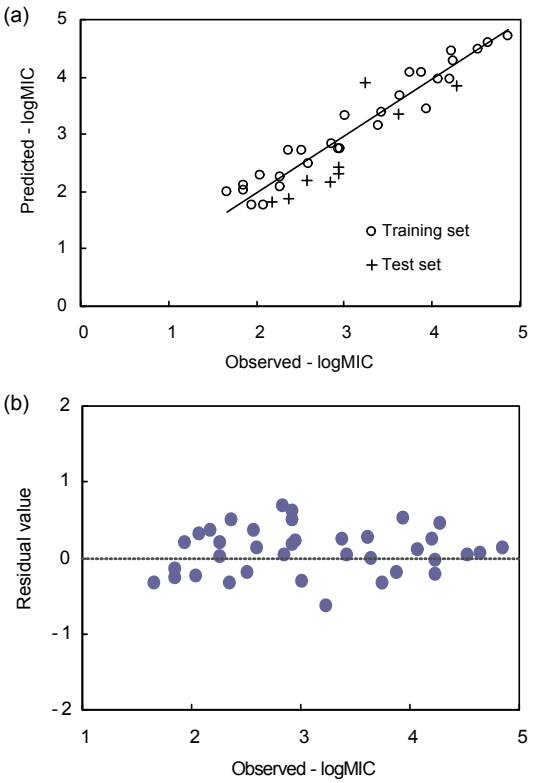


Fig. 3 Analysis of conventional QSAR model

(a) Observed −logMIC values plotted against predicted −logMIC values. R^2=0.942. The standard deviation of the test set was 0.816. The unit of MIC is mol/L. (b) Residual values plotted against experimental values

Table 2 Definition of selected structure descriptors[a]

Structure descriptor	Type	Meaning
RDF030p[b]	RDF descriptors	Radial distribution function-3.0/weighted by atomic polarizabilities
RDF080p[b]	RDF descriptors	Radial distribution function-8.0/weighted by atomic polarizabilities
E2p[c]	WHIM descriptors	2nd component accessibility directional weighted holistic invariant molecular (WHIM) index/weighted by atomic polarizabilities
Inflammat-80[d]	Molecular properties	Ghose-Viswanadhan-Wendoloski anti-inflammatory-like index at 80%
nHBonds	Functional group counts	Number of intramolecular H-bonds (with N, O, F)

[a] Todeschini and Consonni, 2000; [b] Hemmer *et al*., 1999; [c] Todeschini and Gramatica, 1997; [d] Ghose *et al*., 1999

Table 3 Values of selected structure descriptors

Compound	Molecular ID	RDF030p	RDF080p	E2p	Inflammat-80	nHBonds
1	Formic acid	0.170	0.000	0.250	0	0
2	Acetic acid	0.023	0.000	0.300	0	0
3[d]	Propionic acid	0.258	0.000	0.304	0	0
4	Butyric acid	0.396	0.000	0.335	0	0
5[d]	Hexanoic acid	0.901	0.018	0.348	0	0
6	Sorbic acid	0.710	0.107	0.215	0	0
7	Octanoic acid	1.415	0.017	0.357	0	0
8[d]	Monocaprylin	2.235	0.221	0.248	0	1
9[d]	Capric acid	1.926	0.017	0.363	0	0
10	Monocaprin	3.314	2.428	0.288	1	1
11[d]	Undecanoic acid	2.786	1.922	0.269	0	0
12	Monoundecanoin	3.362	2.455	0.319	1	1
13	10-Undecenoic acid	2.742	2.071	0.286	0	0
14[d]	Monoundecenoin	3.189	2.559	0.453	1	1
15	Lauryl alcohol	2.527	0.029	0.390	0	0
16	Monolaurin	4.140	2.150	0.280	1	1
17	11-Dodecenoic acid	2.191	0.001	0.308	0	0
18	Monododecenoin	4.240	1.820	0.324	1	1
19	Tridecanoic acid	3.081	1.248	0.209	0	0
20	Monotridecanoin	5.072	2.277	0.292	1	1
21[d]	12-Tridecenoic acid	3.221	2.197	0.256	0	0
22	Myristic acid	3.678	1.564	0.249	0	0
23[d]	Tetradecanol	3.039	0.029	0.391	0	0
24[d]	Monomyristin	4.341	2.456	0.262	0	1
25	Myristoleic acid	5.802	3.416	0.275	0	0
26	Palmitoleic acid	5.123	0.134	0.340	0	0
27	Monopalmitolein	5.955	2.854	0.488	0	1
28	Stearic acid	4.479	2.491	0.393	0	0
29	Oleic acid	5.711	0.141	0.362	0	0
30	Monolinolein	7.027	6.601	0.411	0	1
31	Linolenic acid	9.321	1.217	0.375	0	0
32	Monolinolenin	10.474	5.814	0.460	0	1
33	γ-linolenic acid	9.067	1.613	0.477	0	0
34	Mono-γ-linolenin	9.243	3.336	0.429	0	1
35	Monoarachidonin	12.037	11.558	0.464	0	2
36	Eicosapentaenoic acid	13.172	7.617	0.471	0	0
37	Docosahexaenoic acid	13.800	9.012	0.518	0	0

[d] These compounds were used as a test set and not included in model derivation

Table 4 Standardized coefficients of variables in conventional QSAR model

Structure descriptor	Standardized coefficient
RDF030p	1.380
RDF080p	−0.995
E2p	−0.346
Inflammat-80	0.278
nHBonds	0.719

model suggests that the more atom pairs within distance of 3.0 Å and the less within 8.0 Å, and the higher polarizabilities these atoms possessed, for example of linolenic acid, the higher antimicrobial activity of the FADs can achieve.

The number of intramolecular hydrogen bonds was another important structure property. According to this model, it can be speculated that the difference

of antimicrobial activity between fatty acids and their monoglycerides can be attributed to the formation of intramolecular hydrogen bonds.

WHIM descriptors were obtained as statistical indices of the atoms projected onto the three principal components from weighted covariance matrices of the atomic coordinates. They can be viewed as a useful template in searching for the principal axes with respect to a defined atomic property (Todeschini and Gramatica, 1997). E2p denotes atom polarizability distribution and density around origin and along the 2nd component axis. In this model, lower distribution and density of atom polarizabilities around origin and along the 2nd component axis were favored for high antimicrobial activity.

Inflammat-80 (Ghose *et al.*, 1999) denotes whether the physicochemical properties were in anti-inflammatory drug-qualifying range (covering approximately 80% of anti-inflammatory drugs). The range was: Ghose-Crippen octanol-water partition coefficient (ALOGP) between 1.4 and 4.5, molar refractivity (AMR) between 59 and 119, molecular weight (M_W) between 212 and 447, and number of atoms between 24 and 59. It could be speculated that compounds with physicochemical properties in this range would have a higher possibility to be antimicrobial agents.

3.3 CoMFA model

CoMFA (Cramer *et al.*, 1988) is widely used in drug design because of its rapid generation of QSAR models (Wei *et al.*, 2005; Caballero *et al.*, 2007; Dong *et al.*, 2009). A CoMFA model (Table 5) was developed with R^2=0.979 and Q^2_{LOO}=0.588. The contributions of steric and electrostatic fields were 72.9% and 27.1%, respectively. The test set was also used to verify the model's external predictivity. The experimental and predicted activity values for training set and test set compounds are shown in Table 1 and Fig. 4. It is shown in Fig. 4b that the residual values distributed uniformly and were lower than three times the standard deviation.

The steric contour plot is shown in Fig. 5a. It was found that a medium-sized green-colored contour was surrounding the C12 (taking the C atom of –COOH in fatty acids, –COO– in monoglycerides of fatty acids, and –COH in fatty alcohols as the C1, similarly hereinafter) and C13 atoms. This means that the bulky substituents at C12 or C13 would increase the potency of antimicrobial activity. For example, compounds **18** (monododecenoin) and **20** (monotridecanoin) exhibited more potent activity than **8** (monocaprylin) and **10** (monocaprin). Two big-sized yellow-colored contours were found surrounding some carbon chains after C14. This indicates that bulky substituents at

Table 5 Summary of CoMFA analysis

Parameter	Value
Q^2_{LOO}	0.588
Number of components	5
Non cross-validated R^2	0.979
Standard error of estimate s	0.156
F	206.300
Steric contribution	72.9%
Electrostatic contribution	27.1%
Prediction R^2_{EXT}	0.852

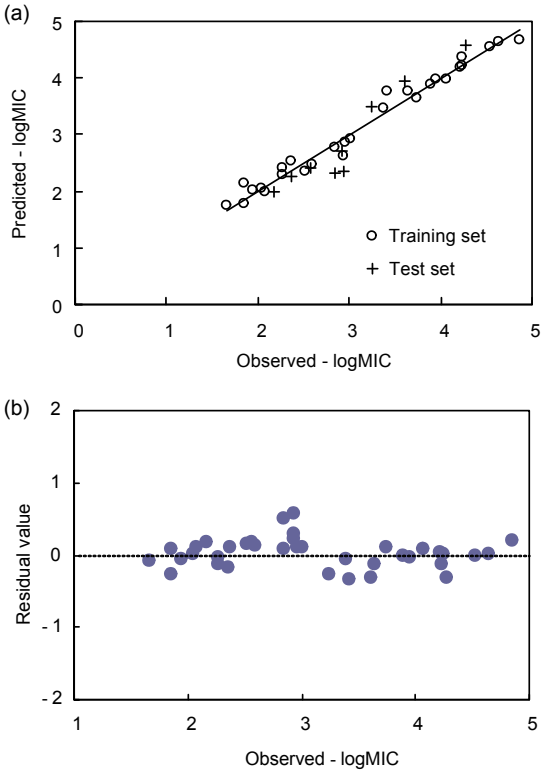


Fig. 4 Analysis of CoMFA model

(a) Observed −logMIC values plotted against predicted −logMIC values. R^2=0.979. The standard deviation of the test set was 0.896. The unit of MIC is mol/L. (b) Residual values plotted against experimental values

these regions would decrease the potency of antimicrobial activity. For instance, compounds **28** (stearic acid) and **30** (monolinolein) displayed less antimicrobial activity, which could be the result of these two compounds being too bulky.

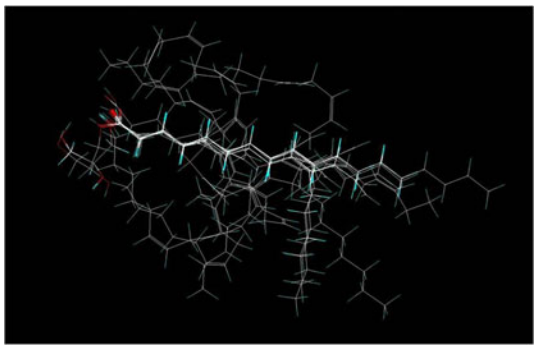


Fig. 2 Alignment of the training and test set compounds

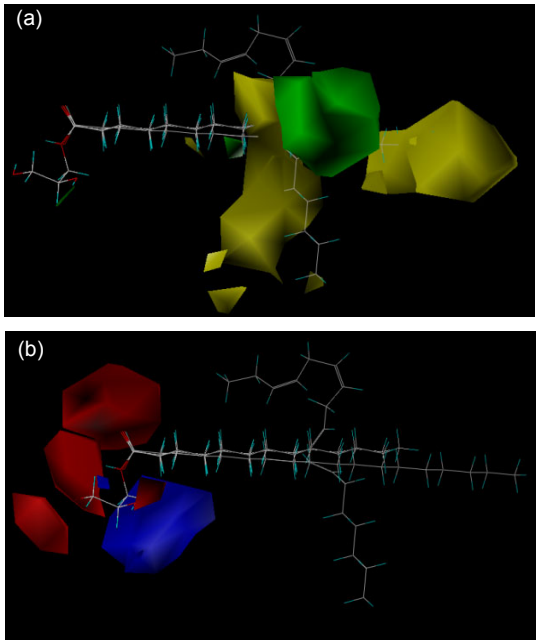


Fig. 5 Electrostatic map and steric map from the CoMFA model

Compounds **15** (lauryl alcohol), **20** (monotridecanoin), **26** (palmitoleic acid), **28** (stearic acid), and **32** (monolinolenin) are shown inside the field. (a) The favorable steric areas with more bulks are indicated by green isopleths (80% contribution), whereas the disfavorable steric areas are shown by yellow isopleths (20% contribution); (b) The favorable electrostatic areas with positive charges are indicated by blue isopleths (78% contribution), whereas the favorable electrostatic areas with negative charges are shown by red isopleths (22% contribution)

The electrostatic contour plot is shown in Fig. 5b. Four medium-sized red-colored contours were found surrounding these oxygen atoms. This means that the increasing of electron density caused by oxygen atoms contributes significantly to antimicrobial activity. It was found that a small-sized blue-colored contour was in the region between the glyceride ester groups and fatty acid carbon chains, which means that negatively charged substituents were not favored.

4 Discussion

It was shown that *S. aureus*, a Gram-positive bacterium, was most effectively inactivated by exposure to monoglycerides of fatty acids with a length of 10–14 carbons, or unsaturated fatty acids and their monoglycerides with a length longer than 18 carbons. This is in good agreement with previous studies in which Kabara *et al.* (1977) and Kelsey *et al.* (2006) showed that medium-chain saturated fatty acids and long-chain unsaturated fatty acids were potent antimicrobials against bacteria, especially Gram-positive bacteria. Monoglycerides were the most active derivatives and were usually more potent than their corresponding fatty acids. However, monotridecanoin (13:0) displayed the most potent activity in this study. It was previously reported (Kabara *et al.*, 1977) that monotridecanoin showed no activity against *S. aureus*. *Cis*-9,*trans*-11 and *trans*-10,*cis*-12 conjugated linoleic acids showed no activity in this study, which was not consistent with Kelsey *et al.* (2006). This may be due to the different sensitivity of *S. aureus* strains to certain lipids. On the other hand, it is the first observation that the compounds such as γ-linolenic acid (18:3), monotridecanoin (13:0), mono-γ-linolenin (18:3), and monoarachidonin (20:4) possess potent anti-*S. aureus* activity.

If the number of descriptors is too small, it is not easy to obtain a good model because some important factors may be neglected. If the number of descriptors is too large, overfitting may occur leading to the wrong prediction. To ensure the quality of the established model, the number of descriptors was determined as five in this study.

From the conventional QSAR model, it can be seen that the molecular space volume and distribution

of atom polarizabilities have a significant influence on antimicrobial activity of FADs, because RDF030p, RDF080p, and E2p are related to the distribution of atoms and their polarizabilities. Atomic polarizability is the relative tendency of a charge distribution, like the electron cloud of an atom to be distorted from its normal shape by an external electric field, which represents atomic electric susceptibility to an external electric field (Nagle, 1990). It is well known that the formation of intramolecular hydrogen bonds changes the electron cloud shape; thus, it is also related to molecular electric properties. Therefore, the molecular space volume and electric properties were the most prevailing structure factors to affect antimicrobial activity.

From the CoMFA model, it can be seen that steric effect was the main effect. The contour plot demonstrated that bulky substituents at C12 or C13, for example of monotridecanoin, but not at C≥14, and negatively charged substituents at the –COOH side are favored for high activity.

Conventional QSAR analysis was used to examine the whole molecule, while CoMFA was prone to focus on structure moieties. The analysis of combining these two models can lead to more reliable results. When predicted values of one compound by the QSAR and CoMFA models were similar (difference ≤0.35), the prediction seemed to be more accurate, as seen in compounds **1**, **2**, **4**, **6**, **7**, **10**, **12**, **13**, **15**, **16**, **17**, **18**, **19**, and **20**. When predicted values by the two models were considerably different (difference >0.35), the predictive accuracy tended to be lower, as seen in compounds **5**, **8**, **14**, **21**, **24**, and **29**.

Conventional QSAR and CoMFA analyses were performed to explore comprehensive structure-activity relationships and two statistically reliable models (conventional QSAR: R^2=0.942, Q^2_{LOO} = 0.910; CoMFA: R^2=0.979, Q^2=0.588) were established. Combining these two models, it can be concluded that in order to display higher antimicrobial activity, FAD molecular structures should possess certain properties. In terms of steric properties, the structure should have bulky substituents at C12 or C13, but not at C≥14, and have more atom pairs with a distance close to 3.0 Å, but fewer near 8.0 Å, which means a denser atom distribution and smaller space volume. In terms of electronic properties, the structure should contain more negative atoms in specific regions, which increases molecular negativity, generates hydrogen bonds, and forms an appropriate distribution of atom polarizabilities. In general, there should be special electron clouds with shape and susceptibility to external electric fields. In terms of physical properties, the structure should have inflammatory drug-like properties (ALOGP 1.4–4.5, AMR 59–119, M_{W} 212–447, and number of atoms 24–59).

Hsiao and Siebert (1999) were the first ones who applied the QSAR method to antimicrobial activity of organic acids. They applied PCA to 11 properties such as M_{W}, number of carbon atoms, and melting point, and obtained four principal components corresponding to polar groups (PC1), the number of double bonds (PC2), molecular size (PC3), and solubility in nonpolar solvents (PC4) to establish models with high R^2 values (Hsiao and Siebert, 1999; Nakai and Siebert, 2003; 2004).

It is well known that molecular solubility is closely related with its polarity: the higher polarity of the molecule, the better solubility in polar solvents, which is called the theory of "similarity and intermiscibility" (Schmid, 2001). Thus, both PC1 and PC4 represented polarity of molecule, and suggested that polarity had great impact on MICs, which coincided with the results in our study. Both the favor of negative charges suggested by the CoMFA model, and the importance of atomic polarizabilities by conventional QSAR model, indicated that electronic properties, including molecular electron cloud shape and its susceptibility to an external electric field, played a significant role. PC2 mainly represented the number of carbon-carbon double bonds. Kabara (1984) analyzed the influence of carbon-carbon double bonds and speculated that the addition of a *cis* double bond increased the antimicrobial activity of a straight-chain fatty acid. A second double bond further increased the activity, while a third double bond was not as effective. This conclusion was only based on serial 18 carbon fatty acids, but differed from another study (Kodicek, 1949), which indicated that the formation of double bonds would change the electron cloud shape as well as the molecular polarity. In addition, double bonds would distort the conformation of the carbon chain. The conformations of linolenic acid (18:3), monoarachidonin (20:4), eicosapentaenoic acid (20:5), and docosahexaenoic acid (22:6) were

distorted by their double bonds and became more spherical, resulting in a decrease of molecular space volume. These compounds displayed potent activity, while some other compounds with the same long-carbon chain, such as monostearin (18:0), eicosanoic acid (20:0), erucic acid (22:1), and monoerucin (22:1), did not have enough double bonds to make the distortion, leading to the results that show their MICs were all above 4000 μg/ml. Therefore, the changes of molecular space volume and electronic properties may be the intrinsic factors of double bonds. PC3 corresponded to molecular size, which indicates that the steric effect was another important factor to antimicrobial activity, as suggested in this study.

Good accordance was found between this study and Siebert's research (Hsiao and Siebert, 1999; Nakai and Siebert, 2003; 2004). Their work was based on different training sets, structure descriptor types, and analysis methods, and the result was reasonable since many physicochemical properties, such as melting point and solubility, had good linear relationships with the structures, as demonstrated by ample quantitative structure-property relationship (QSPR) studies. The results of our work support their conclusions, and may be more reliable because of the more extensive range of FAD samples and the combined application of conventional QSAR and CoMFA methods. The novel 3D structure descriptors were first used in the conventional QSAR analysis to consider the 3D structure features of FADs, and combined with CoMFA analysis to predict more reliable results and provide a better understanding on the nature of antimicrobial FADs.

5 Conclusions

The antimicrobial activities of 57 FADs were examined and 37 were antimicrobials. Compared to fatty acids and fatty alcohols, monoglycerides of fatty acids displayed higher antimicrobial activities, and monotridecanoin was the most potent among the monoglycerides. Conventional QSAR and CoMFA analyses were performed to explore comprehensive structure-activity relationships and two statistically reliable models (conventional QSAR: R^2=0.942, Q^2_{LOO} =0.910; CoMFA: R^2=0.979, Q^2=0.588) were established. Combining these two models is very useful in predicting new potent FADs as anti-*S. aureus* agents prior to synthesis.

References

Bergsson, G., Arnfinnsson, J., Steingrimsson, O., Thormar, H., 2001. In vitro killing of *Candida albicans* by fatty acids and monoglycerides. *Antimicrob. Agents Chemother.*, **45**(11):3209-3212. [doi:10.1128/AAC.45.11.3209-3212.2001]

Branen, J.K., Davidson, P.M., 2004. Enhancement of nisin, lysozyme, and monolaurin antimicrobial activities by ethylenediaminetetraacetic acid and lactoferrin. *Int. J. Food Microbiol.*, **90**(1):63-74. [doi:10.1016/S0168-1605(03)00172-7]

Caballero, J., Saavedera, M., Fernandez, M., Gonzalez-Nilo, F.D., 2007. Quantitative structure-activity relationship of rubiscolin analogues as delta opioid peptides using comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA). *J. Agric. Food Chem.*, **55**(20):8101-8104. [doi:10.1021/jf071031h]

Cramer, R.D., Patterson, D.E., Bunce, J.D., 1988. Comparative molecular-field analysis (CoMFA). I. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.*, **110**(18):5959-5967. [doi:10.1021/ja00226a005]

Dong, X.W., Liu, T., Yan, J.Y., Wu, P., Chen, J., Hu, Y.Z., 2009. Synthesis, biological evaluation and quantitative structure-activities relationship of flavonoids as vasorelaxant agents. *Bioorg. Med. Chem.*, **17**(2):716-726. [doi:10.1016/j.bmc.2008.11.052]

Ghose, A.K., Viswanadhan, V.N., Wendoloski, J.J., 1999. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. *J. Comb. Chem.*, **1**(1):55-68. [doi:10.1021/cc9800071]

Gonzalez, M.P., Teran, C., Saiz-Urra, L., Teijeira, M., 2008. Variable selection methods in QSAR: an overview. *Curr. Top. Med. Chem.*, **8**(18):1606-1627. [doi:10.2174/156802608786786552]

Habulin, M., Sabeder, S., Knez, Z., 2008. Enzymatic synthesis of sugar fatty acid esters in organic solvent and in supercritical carbon dioxide and their antimicrobial activity. *J. Supercrit. Fluids*, **45**(3):338-345. [doi:10.1016/j.supflu.2008.01.002]

Hemmer, M.C., Steinhauer, V., Gasteiger, J., 1999. Deriving the 3D structure of organic molecules from their infrared spectra. *Vibrat. Spectrosc.*, **19**(1):151-164. [doi:10.1016/S0924-2031(99)00014-4]

Hou, C.T., Forman, R., 2000. Growth inhibition of plant pathogenic fungi by hydroxy fatty acids. *J. Ind. Microbiol. Biotechnol.*, **24**(4):275-276. [doi:10.1038/sj.jim.2900816]

Hsiao, C.P., Siebert, K.J., 1999. Modeling the inhibitory effects of organic acids on bacteria. *Int. J. Food Microbiol.*, **47**(3):189-201. [doi:10.1016/S0168-1605(99)00012-4]

Kabara, J.J., 1984. Antimicrobial agents derived from

fatty-acids. *J. Am. Oil Chem. Soc.*, **61**(2):397-403. [doi:10.1007/BF02678802]

Kabara, J.J., Swieczkowski, D.M., Truant, J.P., Conley, A.J., Truant, J.P., 1972. Fatty-acids and derivatives as antimicrobial agents. *Antimicrob. Agents Chemother.*, **2**(1):23-28. [doi:10.1128/AAC.2.1.23]

Kabara, J.J., Vrable, R., Liekenjie, M.S.F., 1977. Antimicrbial lipids-natural and synthetic fatty-acids and monoglycerides. *Lipids*, **12**(9):753-759. [doi:10.1007/BF02570908]

Kelsey, J.A., Bayles, K.W., Shafii, B., McGuire, M.A., 2006. Fatty acids and monoacylglycerols inhibit growth of *Staphylococcus aureus*. *Lipids*, **41**(10):951-961. [doi:10.1007/s11745-006-5048-z]

Kodicek, E., 1949. The effect of unsaturated fatty acids on Gram-positive bacteria. *Symp. Soc. Exp. Biol.*, **3**:217-232.

Mercader, A.G., Duchowicz, P.R., Fernandez, F.M., Castro, E.A., 2008a. Modified and enhanced replacement method for the selection of molecular descriptors in QSAR and QSPR theories. *Chemom. Intell. Lab. Syst.*, **92**(2):138-144. [doi:10.1016/j.chemolab.2008.02.005]

Mercader, A.G., Duchowicz, P.R., Fernandez, F.M., Castro, E.A., 2008b. QSAR Search Tool Tutorial-Replacement method. Available from http://www.mathworks.com/matlabcentral/fileexchange/19578

Mercader, A.G., Duchowicz, P.R., Fernandez, F.M., Castro, E.A., 2010. Replacement method and enhanced replacement method versus the genetic algorithm approach for the selection of molecular descriptors in QSPR/QSAR theories. *J. Chem. Inf. Model.*, **50**(9):1542-1548. [doi:10.1021/ci100103r]

Nagle, J.K., 1990. Atomic polarizability and electronegativity. *J. Am. Chem. Soc.*, **112**(12):4741-4747. [doi:10.1021/ja00168a019]

Nakai, S.A., Siebert, K.J., 2003. Validation of bacterial growth inhibition models based on molecular properties of organic acids. *Int. J. Food Microbiol.*, **86**(3):249-255. [doi:10.1016/S0168-1605(02)00551-2]

Nakai, S.A., Siebert, K.J., 2004. Organic acid inhibition models for *Listeria innocua*, *Listeria ivanovii*, *Pseudomonas aeruginosa* and *Oenococcus oeni*. *Food Microbiol.*, **21**(1):67-72. [doi:10.1016/S0740-0020(03)00043-1]

Nobmann, P., Smith, A., Dunne, J., Henehan, G., Bourke, P., 2009. The antimicrobial efficacy and structure activity relationship of novel carbohydrate fatty acid derivatives against *Listeria* spp. and food spoilage microorganisms. *Int. J. Food Microbiol.*, **128**(3):440-445. [doi:10.1016/j.ijfoodmicro.2008.10.008]

Schmid, R., 2001. Recent advances in the description of the structure of water, the hydrophobic effect, and the like-dissolves-like rule. *Monatshefte für Chemie*, **132**(11):1295-1326. [doi:10.1007/s007060170019]

Todeschini, R., Gramatica, P., 1997. 3D-modelling and prediction by WHIM descriptors. Part 5. Theory development and chemical meaning of the WHIM descriptors. *Quant. Struct.-Act. Relat.*, **16**(2):113-119. [doi:10.1002/qsar.19970160203]

Todeschini, R., Consonni, V., 2000. Handbook of Molecular Descriptors. Wiley-VCH, Weinheim and New York. [doi:10.1002/9783527613106]

Wei, D.G., Yang, G.F., Wan, J., Zhan, C.G., 2005. Binding model construction of antifungal 2-aryl-4-chromanones using CoMFA, CoMSIA, and QSAR analyses. *J. Agric. Food Chem.*, **53**(5):1604-1611. [doi:10.1021/jf048313r]