Design of cinnamaldehyde amino acid Schiff base compounds based on the quantitative structure–activity relationship

Hui Wang¹ ², Mingyue Jiang¹, Shujun Li¹, Chung-Yun Hse², Chunde Jin³, Fangli Sun³ and Zhuo Li¹

¹Key Laboratory of Bio-Based Material Science and Technology of the Ministry of Education, Northeast Forestry University, Harbin 150040, People’s Republic of China
²Southern Research Station, USDA Forest Service, Pineville, LA 71360, USA
³Key Laboratory of Wood Science and Technology, Zhejiang Agriculture and Forestry University, Zhejiang 311300, People’s Republic of China

Cinnamaldehyde amino acid Schiff base (CAAS) is a new class of safe, bioactive compounds which could be developed as potential antifungal agents for fungal infections. To design new cinnamaldehyde amino acid Schiff base compounds with high bioactivity, the quantitative structure–activity relationships (QSARs) for CAAS compounds against Aspergillus niger (A. niger) and Penicillium citrinum (P. citrinum) were analysed. The QSAR models ($R^2 = 0.9346$ for $A. niger$, $R^2 = 0.9590$ for $P. citrinum$) were constructed and validated. The models indicated that the molecular polarity and the Max atomic orbital electronic population had a significant effect on antifungal activity. Based on the best QSAR models, two new compounds were designed and synthesized. Antifungal activity tests proved that both of them have great bioactivity against the selected fungi.

1. Introduction

Primary and opportunistic antifungal infections are a severe threat to human life and health [1]. As fungal resistance increases, many antifungal compounds have become ineffective [2]. It is therefore necessary to explore new, novel antifungal formulations to control fungal infections [3]. Natural products and modified natural-derived compounds have continued to play

© 2017 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, provided the original author and source are credited.
Figure 1. The synthesis route of cinnamaldehyde Schiff base.

a highly significant role in the discovery of antifungal agents [4]. Researchers have modified natural, antifungal compounds to meet key requirements for practical applications. Cinnamon oil is a natural, antifungal substance and its main component is cinnamaldehyde [5]. Numerous studies have reported that cinnamaldehyde could inhibit the growth of the pathogenic microorganisms *Aspergillus niger*, *Trametes versicolor* and *Staphylococcus aureus* [6]. Cinnamaldehyde also exhibited potential anti-tumour [7] and anti-diabetes [8] properties. Also, cinnamaldehyde is generally recognized as safe and is allowed as a food additive or antimicrobial agent by the US FDA (Food and Drug Administration) [9]. However, cinnamaldehyde as either an antimicrobial agent or food additive has many practical limitations largely due to its high volatility and strong odours [10].

Hence, many researchers have shifted their attention to cinnamaldehyde derivatives. Sharma et al. [11] synthesized cinnamaldehyde derivatives and cinnamaldehyde Schiff base. The results indicated that the presence of a methoxyl group on cinnamaldehyde benzene ring and cinnamaldehyde Schiff base led to a noticeable improvement in antifungal activity. Cinnamaldehyde Schiff base is an important class of cinnamaldehyde derivatives with excellent bioactivity and can be synthesized using a simple method [12]; the synthesis route is shown in figure 1. The bioactivity of cinnamaldehyde Schiff base compounds has been reported by many researchers. Zahan et al. [13] studied the dithiocarbazata cinnamaldehyde Schiff base compound and the metal complex compound. The bioactivity test showed that cinnamaldehyde Schiff base and metal complex exhibited comparative activity to cinnamaldehyde. Wei et al. [14] published a research on cinnamaldehyde amino acid Schiff base. Results indicated that Schiff base compounds were more active than the reference benzoic acid against *Bacillus subtilis*, *Escherichia coli* and *Saccharomyces cerevisiae*. Hence, it is meaningful to explore and design new cinnamaldehyde Schiff base compounds with favourable bioactivity. In a previous study, the antimicrobial activity of several cinnamaldehyde amino acid Schiff base compounds were studied [15]; the antimicrobial activity results implied that cinnamaldehyde amino acid Schiff base compounds possessed excellent antifungal activity, good water solubility and an odour. Cinnamaldehyde amino acid Schiff base has potential to be an antifungal agent. After an initial analysis regarding the structure and activity, its antifungal activity was found to be significantly influenced by its chemical structure. A comprehensive study on the relationship between activity and compounds should be conducted for designing the new cinnamaldehyde compounds. One approach is to design compounds using computer applications such as quantitative structure–activity relationship (QSAR) [16]. QSAR provides a mathematically quantified relationship between a molecule’s structural descriptors and a compound’s bioactivity at the molecular level, and can predict the activity of compounds including those not yet synthesized [17]. Using this approach, there is no need to synthesize each compound to discover those that possess the desired activity. Promising compounds can be further screened for synthesis in the laboratory.

This paper focuses on the use of QSAR for cinnamaldehyde amino acid Schiff base compounds to present a comprehensive analysis on the relationship between the bioactivity and structures of cinnamaldehyde amino acid Schiff base (CAAS) compounds. Under the guidance of QSAR models, two designed cinnamaldehyde compounds were synthesized and their antifungal activities were determined.

2. Material and methods

2.1. Materials

Cinnamon oil (95% cinnamaldehyde) was provided by Zhen xing Spices Oil Refinery of Ji’an City, China. All other chemicals were analytical grade reagents and were used as received without further purification. The structure of the CAAS compound is shown in figure 2. The functional groups of CAAS compounds are listed in tables 1 and 2. Test microorganisms were two mould species provided by the Chinese Center of Industrial Culture Collection (CICC), Beijing, China. They were *A. niger* (CICC2487) and *Penicillium citrinum* (CICC4010)).
The antifungal activity of all CAAS compounds was determined according to the Paper Disc Method against *A. niger* and *P. citrinum*. In brief, the procedure is as follows. Potato dextrose agar (PDA) medium with 2% agar was prepared and sterilized for use. The sterilized Petri dishes and PDA medium were sterilized under UV-irradiation for 20 min. The strain suspension was molten medium which was thoroughly mixed and then poured into the Petri dishes and allowed to solidify. The autoclaved discs (approx. 8 mm) were dipped into the test solution (concentration: 0.125 mol l\(^{-1}\)) for 10 min. After that, the discs were put onto the surface of the solid media strain suspension mixture. The test samples were cultured at 28°C for 2 days. All tests were carried out in triplicate and the diameter of the inhibition zone was the average of those of the three test zones.

### 2.2. Determination of antifungal activity

The antifungal activity of all CAAS compounds was determined according to the Paper Disc Method against *A. niger* and *P. citrinum*. In brief, the procedure is as follows. Potato dextrose agar (PDA) medium with 2% agar was prepared and sterilized for use. The sterilized Petri dishes and PDA medium were sterilized under UV-irradiation for 20 min. The strain suspension was molten medium which was thoroughly mixed and then poured into the Petri dishes and allowed to solidify. The autoclaved discs (approx. 8 mm) were dipped into the test solution (concentration: 0.125 mol l\(^{-1}\)) for 10 min. After that, the discs were put onto the surface of the solid media strain suspension mixture. The test samples were cultured at 28°C for 2 days. All tests were carried out in triplicate and the diameter of the inhibition zone was the average of those of the three test zones.
The antifungal activity rates and values of descriptors for CAAS compounds against *P. citrinum*.

| ID | R1          | R2          | AR   | lgAR  | max. atomic orbital electron population, $D_S$ | max. electroph. react. index for a C atom, $D_a$ | PNSA-2 total charge weighted PNSA | max. 1-electron react. index for a O atom, $D_{0}$ |
|----|-------------|-------------|------|-------|-----------------------------------------------|-----------------------------------------------|---------------------------------|-----------------------------------------------|
| 1  | (CH₂)₂COOK  | H           | 195.08 | 2.2902 | 1.9133 | 0.0229 | −198.7038 | 3.5194 × 10⁻⁴ |
| 2  | (CH₂)₂COOK  | -p-OCH₃     | 195.08 | 2.2902 | 1.9133 | 0.0267 | −196.8914 | 1.2945 × 10⁻⁴ |
| 3  | (CH₂)₂COOK  | -p-Cl       | 188.69 | 2.2758 | 1.9819 | 0.0177 | −197.1382 | −3.0860 × 10⁻⁷ |
| 4  | H           | H           | 79.58  | 1.9008 | 1.9089 | 0.0204 | −101.4046 | 1.0495 × 10⁻⁵ |
| 5  | H           | -p-OCH₃     | 89.85  | 1.9535 | 1.9089 | 0.0199 | −130.0808 | 1.0946 × 10⁻⁵ |
| 6  | H           | -p-Cl       | 164.31 | 2.2157 | 1.9919 | 0.0189 | −127.9426 | 1.1607 × 10⁻⁶ |
| 7  | CH₃         | H           | 80.87  | 1.9078 | 1.9095 | 0.0214 | −88.7516  | −3.2397 × 10⁻⁶ |
| 8  | CH₃         | -p-OCH₃     | 80.87  | 1.9078 | 1.9095 | 0.0206 | −119.6568 | −1.9399 × 10⁻⁶ |
| 9  | CH₃         | -p-Cl       | 207.92 | 2.3179 | 1.9819 | 0.0222 | −126.2349 | −1.0264 × 10⁻⁶ |
| 10 | CH₂CH(CH₃)₂ | H           | 71.88  | 1.8566 | 1.9088 | 0.0186 | −125.6587 | 1.4595 × 10⁻⁵ |
| 11 | CH₂CH(CH₃)₂ | -p-OCH₃     | 92.42  | 1.9658 | 1.9088 | 0.0199 | −156.3959 | −9.8043 × 10⁻⁵ |
| 12 | CH₂CH(CH₃)₂ | -p-Cl       | 197.65 | 2.2959 | 1.9819 | 0.0179 | −165.1279 | −4.5030 × 10⁻⁵ |
| 13 | CH(CH₂)₂    | H           | 74.44  | 1.8718 | 1.9095 | 0.0223 | −92.5453  | 1.5018 × 10⁻⁷ |
| 14 | CH(CH₂)₂    | -p-OCH₃     | 86.65  | 1.9378 | 1.9095 | 0.0227 | −124.8644 | 1.3298 × 10⁻⁶ |
| 15 | CH(CH₂)₂    | -p-Cl       | 238.77 | 2.3780 | 1.9819 | 0.0248 | −132.3244 | 1.9845 × 10⁻⁶ |
| 16 | CH₂Ar-OH    | H           | 88.57  | 1.9473 | 1.9087 | 0.0185 | −159.8114 | 4.1686 × 10⁻⁷ |
| 17 | CH₂Ar-OH    | -p-OCH₃     | 127.08 | 2.1041 | 1.9087 | 0.0161 | −172.6105 | 1.8280 × 10⁻⁵ |
| 18 | CH₂Ar-OH    | -p-Cl       | 181.00 | 2.2577 | 1.9819 | 0.0189 | −176.7350 | 1.8744 × 10⁻⁴ |
| 19 | CH₂Ar       | H           | 74.45  | 1.8718 | 1.9096 | 0.0205 | −131.3740 | 6.5028 × 10⁻⁵ |
| 20 | CH₂Ar       | -p-OCH₃     | 151.46 | 2.1803 | 1.9096 | 0.0194 | −179.6525 | 2.3434 × 10⁻³ |
| 21 | CH₂Ar       | -p-Cl       | 178.42 | 2.2514 | 1.9819 | 0.0172 | −168.2581 | 1.6235 × 10⁻⁶ |

In this experiment, a well-known commercial antifungal compound fluconazole served as control. The antifungal activity rate (AR) was calculated using the followed equation:

$$AR = \left( \frac{dT}{dC} \right) \times 100\%, \quad (2.1)$$

where $dT$ and $dC$ were the diameter of the inhibition zone for the test compounds and fluconazole, respectively. The antifungal activity rates (ARs) and the $\log_{10} AR$ (lgAR) of all CAAS compounds are listed in tables 1 and 2. The lgAR was used to compute the relationship between antifungal activity and structure of cinnamaldehyde compounds. The values of $dC$ for the control compound fluconazole were 18.7 mm and 13 mm against *A. niger* and *P. citrinum*, respectively.

2.3. The method used for quantitative structure–activity relationship calculations

The three-dimensional structures of the compounds were drawn using the Chembio 3D 12.0 software, and the chemical structures were imported for geometrical optimization using the AMPAC Agui 9.2.1 software [18]. Secondly, the output file of the 21 compounds’ structural information and the lgAR were imported in the CODESSA 2.7.16 software to compute molecular descriptors.

Then, a ‘best multilinear regression’ function was conducted in the CODESSA 2.7.16 software to calculate the regression relationship between chemical structures and antifungal activity. Then, the number of descriptors and the optimal QSAR models were determined by the ‘breaking point’ rule [19] (The squared correlation ($R^2$) of the model dramatically increased over the number of descriptors, but after a certain point, the increase is less significant [20]. That point is the ‘breaking point.’) To analyse the descriptor, charge distribution and density of cinnamaldehyde compounds were calculated by the
Gaussian 09 and Gauss View 5.0 software for further analysis of molecular descriptors. The calculation was performed in ground state/DFT/B3LYP/3-21G.

2.4. The validation of optimal models

Model validation was conducted using a process of internal validation and ‘leave one out’ cross-validation [21]. In short, 21 compounds were classified into three small groups termed a, b and c with seven compounds in each group. Each of two small groups were combined as the training set A(a + b), B(a + c), C(b + c), and the correspondingly remaining groups c, b and a were considered the test set. Using the same descriptors of the best model, ‘multilinear regression’ was conducted to obtain a regression model for the training set and this model was used to predict the $\text{lgAR}$ of the test set. The statistical results, correlation coefficient ($R^2$), Fisher value ($F$) and standard deviation ($s^2$) are listed in table 4. The ‘leave one out’ cross-validation is similar to the internal validation, which was conducted as follows. Every fourth compound was set as the external test set d(4, 8, 12, 16 and 20) and the other compounds were considered to be the training set $D$. Similarly, a training set model was obtained by computing the multilinear regression with the same descriptors of the best model. The obtained training set model was used to predict the corresponding external test set.

3. Results and discussion

3.1. Establishing the optimal quantitative structure–activity relationship models

A series of QSAR models were obtained after performing the ‘best multilinear regression’ procedure. The optimal model was determined by a rule called ‘breaking point’ shown in figure 3, which was implemented by analysing the plot of the descriptors of the obtained models versus squared correlation corresponding to those models. In figure 3, the squared coefficient increased rapidly until the point corresponding to four descriptors. After this point, the increase in value of the squared coefficient was not as great. Hence, this point is the ‘breaking point’, and the QSAR model corresponding to the breaking point is regarded as the optimal QSAR model. Additionally, the number of descriptors should meet the requirement of multilinear regression:

$$N \geq 3(D + 1),$$

where $N$ is the sample number (21) and the $D$ is descriptor number of the final QSAR models [22]. Therefore, the optimal QSAR models were selected using four descriptors. The value of each descriptor of the optimal models is listed in tables 1 and 2. These four descriptor parameters and statistical data corresponding to the optimal QSAR models were listed in table 3, and the definition and analysis for each descriptor parameter are presented in Results and discussion section.

According to the statistical data of the optimal models, the optimal QSAR models were described as fit multilinear regression equations (3.2) and (3.3). In the equations, descriptor parameters ($D$) are the independent variables, and the $\text{lgAR}$ is the calculated value of compounds. For CAAS compounds, the predicted value ($\text{lgAR}_{\text{calc}}$) was calculated according to the above equation and the relationships between experimental value ($\text{lgAR}_{\text{exp}}$) and predicted value ($\text{lgAR}_{\text{calc}}$) are presented for $A. niger$ and $P$. citrinum, respectively (figure 4). In figure 4, the $\text{lgAR}_{\text{exp}}$ and $\text{lgAR}_{\text{calc}}$ fit in a line $y = x$, with $R^2$ of 0.9572 and 0.9301 against $A. niger$ and $P. citrinum$, respectively, which implied that the best QSAR models possessed good predictability. In table 4, all the validation results are satisfactory. The average of the statistical results was very close to the best model.

\[
\text{lgAR}_{A.n} = -(3.2190 \pm 1.0892) \times 10^{-1} - (2.9745 \pm 2.1076 \times 10^{-1}) \times D_1 - (1.5306 \pm 1.9735 \times 10^{-1}) \times D_2 \\
- (3.2064 \times 10^{-2} \pm 7.8686 \times 10^{-3}) \times D_3 + (1.2940 \pm 4.0540 \times 10^{-1}) \times D_4 \tag{3.2}
\]

and

\[
\text{lgAR}_{P.c} = -(7.2473 \pm 6.1451 \times 10^{-1}) + (4.3411 \pm 3.1123 \times 10^{-1}) \times D_5 + (3.0016 \times 10^1 \pm 3.9781) \times D_6 \\
- (2.1543 \times 10^{-3} \pm 3.2208 \times 10^{-4}) \times D_7 + (8.7623 \pm 1.8233) \times 10^1 \times D_8. \tag{3.3}
\]
The optimal QSAR model obtained for CAAS compounds against Table 3.

Some important structure information significantly influenced the antifungal activity involved in the optimal QSAR model. According to the optimal model listed in Table 3 for the CAAS compounds and Penicillium citrinum,

### 3.2. Descriptor parameters analysis for the optimal quantitative structure–activity relationship model

Some important structure information significantly influenced the antifungal activity involved in the optimal QSAR model. According to the optimal model listed in Table 3 for the CAAS compounds

| no | descriptor | name of descriptor | t-test value | t-value | name of descriptor |
|----|------------|--------------------|--------------|---------|--------------------|

#### Table 3. The optimal QSAR model obtained for CAAS compounds against A. niger and P. citrinum.

| no | descriptor | name of descriptor | t-test value | t-value | name of descriptor |
|----|------------|--------------------|--------------|---------|--------------------|

#### A. niger model: $R^2 = 0.9346$, $F = 57.20$ and $s^2 = 0.0020$

0 $-3.2190 \times 10^1$ $1.0892 \times 10^1$ $-2.9555$ Intercept
1 $-2.9745$ $2.1076 \times 10^{-1}$ $-14.1132$ polarity parameter/square distance, $D_1$
2 $-1.5306$ $1.9735 \times 10^{-1}$ $-7.7556$ RNCG relative negative charge (QMNEG/QTMINUS), $D_2$
3 $-3.2064 \times 10^{-2}$ $7.8686 \times 10^{-3}$ $-4.0749$ ESP-HA-dependent HDCA-1[Quantum-Chemical], $D_3$
4 $-1.2940$ $4.0540 \times 10^{-1}$ $3.1918$ max. total interaction for a C–O bond, $D_4$

#### P. citrinum model: $R^2 = 0.9590$, $F = 93.47$ and $s^2 = 0.0018$

0 $-7.2473$ $6.1451 \times 10^{-1}$ $-11.7937$ Intercept
1 $4.3411$ $3.1123 \times 10^{-1}$ $13.9486$ max. atomic orbital electronic population, $D_5$
2 $3.0016 \times 10^1$ $3.9781$ $7.5454$ max. electroph. React. index for a C atom, $D_6$
3 $-2.1543 \times 10^{-1}$ $3.2208 \times 10^{-4}$ $-6.6889$ PNSA-2 total charge weighted PNSA, $D_7$
4 $8.7623 \times 10^1$ $1.8233 \times 10^1$ $4.8056$ max. 1-electron react. index for a O atom, $D_8$
Figure 5. The molecular electronic potential map of compounds 7 and 8 from Gaussian 09 W calculation.

Table 4. Validation of the QSAR models for A. niger and P. citrinum.

| training set | \( N \) | \( R^2 (\text{fit}) \) | \( F (\text{fit}) \) | \( s^2 (\text{fit}) \) | test set | \( N \) | \( R^2 (\text{pred}) \) | \( F (\text{pred}) \) | \( s^2 (\text{pred}) \) |
|--------------|-------|----------------|----------------|---------------|----------|-------|----------------|----------------|---------------|
| validation for the model of A. niger | | | | | | | | | |
| a + b        | 14    | 0.9410         | 35.88          | 0.0023        | c        | 7     | 0.8692         | 39.84          | 0.0207        |
| a + c        | 14    | 0.9678         | 67.68          | 0.0013        | b        | 7     | 0.8829         | 37.75          | 0.0281        |
| b + c        | 14    | 0.9367         | 33.31          | 0.0020        | a        | 7     | 0.9283         | 77.73          | 0.019         |
| average      | 14    | 0.9485         | 45.62          | 0.0019        | average  | 7     | 0.8868         | 51.77          | 0.0226        |
| \( D \)      | 16    | 0.9509         | 53.21          | 0.0016        | d        | 5     | 0.8770         | 28.53          | 0.0252        |
| validation for the model of P. citrinum | | | | | | | | | |
| a + b        | 14    | 0.9623         | 57.39          | 0.0021        | c        | 7     | 0.9485         | 110.52         | 0.0156        |
| a + c        | 14    | 0.9600         | 53.94          | 0.0021        | b        | 7     | 0.9334         | 84.05          | 0.0188        |
| b + c        | 14    | 0.9711         | 75.54          | 0.0014        | a        | 7     | 0.9235         | 72.43          | 0.0224        |
| average      | 14    | 0.9645         | 62.29          | 0.0019        | average  | 7     | 0.9351         | 89.00          | 0.0189        |
| \( D \)      | 16    | 0.9606         | 66.97          | 0.0020        | d        | 5     | 0.8934         | 33.53          | 0.0264        |

against A. niger, there were four structural descriptors that apparently affected the antifungal activity of the CAAS compounds. The most statistically significant descriptor was the polarity parameter/square distance, \( D_1 \). This is an electrostatic descriptor defined by the following equation [23,24]:

\[
P'' = \frac{Q_{\text{max}} - Q_{\text{min}}}{R_{\text{mm}}^2},
\]

where \( Q_{\text{min}} \) and \( Q_{\text{max}} \) are the most negative and the most positive atomic partial charges in the molecule, respectively, and \( R_{\text{mm}} \) is the distance between the most positive and the most negative atomic partial charges in the molecule. The polarity parameter reflects the polarity and characteristics of the charge distribution of the molecule. A compound with proper polarity can smoothly penetrate a fungal cell wall or cell membrane and interact with an active target. A change was observed on the value of \( P'' \) when the cinnamaldehyde compounds' structure changed, for example compound 7 had a \( P'' \) value of 0.1249. The charge distribution changed when –OCH\(_3\) was introduced on the benzene ring (figure 5). This charge distribution led to an increase in the value of \( Q_{\text{min}} \) and a decrease in \( P'' \) (2.5943 \times 10^{-3}).

The second descriptor is relative negative charge (RNCG) \( (D_2) \), and it is a quantum-chemical descriptor. RNCG is defined as the most negative charge divided by the total negative charge [25]:

\[
\text{RNCG} = \frac{Q_{\text{max}}}{Q^-},
\]

where \( Q_{\text{max}} \) is the most negative charge and \( Q^- \) is the total negative charge. As shown in table 3, RNCG negatively contributed to the AR against A. niger. Like compounds 2, 5, 8, 11, 14, 17 and 20 had a methoxy group (electron-donating group) on the benzene ring and resulted in an obvious increase in the descriptor RNCG compared with compounds that do not have a substituent group on the benzene ring (1, 4, 7, 10, 13, 16 and 19). Conversely, an electron-withdrawing group such as a halogen atom (chlorine) may decrease the value of \( D_2 \), because different substituent groups changed the charge distribution of...
cinnamaldehyde compounds; for instance in compounds 2 and 3, the only difference in structure is the benzene ring 4-substituent. Compound 2 has a methoxyl group and compound 3 has a chlorine atom in the p-position on the benzene ring. These differences in substituent groups lead to a lower RNCG value for compound 3 (0.1624) than for compound 2 (0.2546). The charge distribution of optimal structures of compounds 2 and 3 by Gaussian 09 could explain it (figure 6). In figure 6, the oxygen atom in the methoxyl group had the most negative charge ($-0.554$) in compound 2; the total negative charge and the most negative charge changed when the substituent group changed to a Cl atom. The most negative charge of compound 3 was $-0.495$.

The third most important descriptor is the ESP-HA-dependent HDCA-1 (H-acceptor dependent H-donor charged surface area, $D_3$), which is a quantum-chemical descriptor [26] that represents the hydrogen bonding donor ability of the CAAS compounds [27]. The formation of hydrogen bonding of CAAS compounds is easier as HDCA-1 increases [28]. In equation (3.2), a negative coefficient for HDCA-1 demonstrates the ability of the CAAS compounds to form hydrogen bonds which might be detrimental to antifungal activity.

The last parameter is maximum total interaction for a C–O bond ($D_4$), which is a semi-empirical descriptor that could be used to measure the bond strength between the two atoms involved [29]. A positive coefficient implied that the strength of the C–O bond had a positive contribution to antifungal activity of cinnamaldehyde compounds against $A. niger$.

According to the optimal QSAR models against $P. citrinum$, the most statistically significant descriptor was the maximum atomic orbital electronic population ($D_5$). It is an electrostatic descriptor and an index of nucleophilicity for cinnamaldehyde compounds [30]. The positive coefficient in the model indicated that the increase in $D_5$ denoted antifungal activity of cinnamaldehyde compounds against $P. citrinum$.

The second descriptor is maximum electrophilic reactivity index for a C atom ($D_6$), which is a quantum-chemical descriptor [31] that reflects the electrophilic reactivity of the C atom on cinnamaldehyde compounds. For a given atomic species A, the maximum electrophilic reactivity index for an A atom was defined as [32]

$$E_A = \sum_{i=1}^{n_A} \frac{C_{\text{LUMO},i}^2}{\varepsilon_{\text{LUMO}}} + 10$$

where $\varepsilon_{\text{LUMO}}$ is the energy of the lowest unoccupied molecular orbital (LUMO). Here, $C_{\text{LUMO},i}$ is the ith orbital coefficient of atom A on LUMO. Such summation is conducted over all valence atomic orbitals $i$ in atom A($i = 1 \ldots n_A$). In the best QSAR model against $P. citrinum$, a positive coefficient indicates that the antifungal activity increased as the magnitude of $D_6$ increased.

The third descriptor parameter is PNSA-2 total charge weighted PNSA ($D_7$). This descriptor is defined as the total negative charge multiplied by partial negative solvent-accessible surface area, which indicates the influence of negative charge distribution on the antifungal activity of cinnamaldehyde compounds [33].

The last descriptor is the maximum 1-electron reactivity index for an O atom ($D_8$), which is a quantum-chemical descriptor [34]. It is an important descriptor parameter selected by ‘best multilinear regression’ from about 400 descriptors. In equation (3.3), $D_8$ had a positive coefficient showing that increase in the magnitude of $D_8$ will increase the antifungal activity of the cinnamaldehyde compound against $P. citrinum$.

4. Design of new compounds

According to analysis results of the two best QSAR models, the most important factors for antifungal activity were the polarity parameter ($D_1$) and the maximum atomic orbital electronic population ($D_5$) against $A. niger$ and $P. citrinum$, respectively. For the most important structural parameter of the QSAR
model against *A. niger*, some special structural factors like the number of COO− and substituent groups on the benzene ring significantly decreased the value of the polarity parameter, and this decrease was very beneficial to increase antifungal activity. Hence, these special structural factors were chosen as the structural characteristics on the new design compounds. In addition, the number of COO− groups of cinnamaldehyde Schiff base compounds will change the polarity of the cinnamaldehyde compounds; that is the increase in COO− groups will result in an increase in the polarity of the compounds, which is beneficial to increase the water solubility of cinnamaldehyde compounds and enlarge their applied field.

With regard to the most important structural parameter of the QSAR model against *P. citrinum*, the value of D5 of the cinnamaldehyde compound was positive, contributing to the antifungal activity against *P. citrinum*. The structural characteristics like the number of COO− groups and halogen atoms increased the value of D5 obviously. However, it is generally believed that halogenated hydrocarbon possesses high toxicity [35]. This structural characteristic was not considered in design compounds. Above all, the structural characteristics of the number of COO− and −OCH3 groups were selected as the key character factors in the design of compounds. Hence two new compounds were designed and synthesized. Structured characterization results are shown as follows.

**Figure 7.** Structures of the new designed compounds.

**Table 5.** The antifungal activity rate and comparison between the predicted value and the experimental value of designed compounds Da and Db.

|                     | no | Cal lgAR | Exp lgAR | Exp lgAR | error |
|---------------------|----|----------|----------|----------|-------|
| *Aspergillus niger* | Da | 2.1851   | 128.55   | 2.1091   | 0.0760 |
|                     | Db | 2.0480   | 120.51   | 2.0810   | −0.0330 |
| *Penicillium citrinum* | Da | 2.0275 | 225.92 | 2.3540 | −0.3265 |
|                     | Db | 2.1125 | 182.28 | 2.2607 | −0.1482 |

The predicted lgAR of new compounds was obtained by the following steps. First, the structures of new compounds were drawn and inputted into the AMPAC 9.21 software to geometry-optimize them and the optimized structure files were saved. Then, the optimized structure files were inputted into the CODESSA 2.7.16 software to calculate the molecular descriptors. Finally, a predict function was conducted in the condition of the best model, and the calculated value (Cal lgAR) was obtained and listed in table 5. In table 5, the Cal lgAR value of two compounds for both fungi were greater than that of the control compound fluconazole. The Cal lgAR value of compound Da against *A. niger* was the greatest among those of all the compounds used for establishing the model.
The experimental results of antifungal activity have shown that new compounds exhibited better bioactivity than compounds listed in Table 1. From Table 5, the Exp.lgAR value was very close to that of Cal.lgAR for both new compounds against the two fungi. The average of absolute error and the relative error were 0.0545 and 2.55% against A. niger, and 0.2374 and 11.55% against P. citrinum. These small errors implied that two best QSAR models had good predictability and were satisfactory. From another perspective, two designed compounds could be treated as the external test set to validate the best QSAR models. Small errors indicated that these two QSAR models were reliable.

5. Conclusion

Two QSAR models of CAAS compounds against A. niger and P. citrinum with good statistical results were obtained and validated. The definition and analysis of the important descriptor parameters implied the chemical structural characteristics which influenced antifungal activity. The results indicated that molecular polarity and negative charge distribution of cinnamaldehyde compounds were important influences on antimicrobial activity. By analysis of the descriptor parameters of these two models, some guidance was obtained on chemical structure for the design of new cinnamaldehyde compounds. Two designed compounds exhibited excellent antifungal activity against both fungi and the experimental values were very close to the predicted values. All the results indicated that two best QSAR models possessed good predictability.

Ethics. Our research does not require any ethical approval from a local ethics committee because we carried out our work based on other sources.

Data accessibility. The datasets supporting this article have been uploaded as part of the electronic supplementary material.

Authors’ contributions. S.L. designed the research and revised the manuscript. H.W. carried out the major research work and statistical analyses. M.J. and Z.L. helped to test the antifungal activity of compounds. H.W. wrote the manuscript. H.C. coordinated the study and helped draft the manuscript. All authors gave their final approval for publication. F.S. and C.J. supported the research methods and revised the manuscript.

Competing interests. The authors declare no competing interests.

Funding. This research was financially supported by Fundamental Research Funds for the Central Universities (2572016A824 (H.W.)) and Zhejiang Key Level 1 (2014lygczy016 (S.L.)), China.

Acknowledgements. As a visiting student, H.W. would like to thank the support by the China Scholarship Council (CSC) and Southern Research Station, USDA Forest Service, United States.

References

1. Abranches P et al. 2015 Complexes of fluconazole with sodium p-sulfonatocalix[n]arenes: characterization, solubility and antifungal activity. RSC Adv. 5, 44 373–44 325. (doi:10.1039/CSRA05428K)

2. He D, Jian W, Liu X, Shen H, Song S. 2015 Synthesis, biological evaluation, and structure–activity relationship study of novel stillbene derivatives as potential fungicidal agents. J. Agric. Food Chem. 63, 1370–1377. (doi:10.1021/jf5052893)

3. Kathiravan MK, Salake AB, Chothe AS, Dudhe PB, Watode RP, Mukta MS, Gadhwe S. 2012 The biology and chemistry of antifungal agents: a review. Bioorg. Med. Chem. 20, 5678–5698. (doi:10.1016/j.bmcc.2012.04.045)

4. Qin S, Ying K, Jiang J-H, Xu L-H, Li W-J. 2011 Biodiversity, bioactive natural products and biotechnological potential of plant-associated endophytic actinobacteria. Appl. Microbiol. Biotechnol. 89, 457–473. (doi:10.1007/s00253-010-2923-6)

5. Nostro A, Scaffaro R, D’Amico M, Batta L, Filocamo A, Marino A, Bisignano G. 2012 Study on carvacrol and cinnamaldehyde polymeric films: mechanical properties, release kinetics and antibacterial and antifilm activities. Appl. Microbiol. Biotechnol. 96, 1029–1036. (doi:10.1007/s00253-012-4091-9)

6. Khan MSA, Ahamed I. 2011 Antifungal activity of essential oils and their synergy with fluconazole against drug-resistant strains of Aspergillus fumigatus and Trichophyton rubrum. Appl. Microbiol. Biotechnol. 90, 1083–1094. (doi:10.1007/s00253-011-3152-3)

7. Gan FF, Chua YS, Scarmagnani S, Palaniappan P, Franks M, Poolabalanam T, Bradshaw TD, Westwell AD, Hagen T. 2009 Structure–activity analysis of Z–modified cinnamaldehyde analogues as potential antitumor agents. Biochem. Biophys. Res. Commun. 387, 741–747. (doi:10.1016/j.bbrc.2009.07.104)

8. Babu PS, Prabuseenivasan S, Ignacimuthu S. 2007 Cinnamaldehyde—a potential antiobiotic agent. Phytomedicine 14, 15–22. (doi:10.1016/j.phymed.2006.11.005)

9. Feng K, Wen P, Yang H, Li N, Lou WY, Zong MH, Wu H. 2017 Enhancement of the antimicrobial activity of cinnamon essential oil-loaded electropositive nanofilm by the incorporation of lysozyme. RSC Adv. 7, 1572–1580. (doi:10.1039/C6RA25977D)

10. Rodriguez M et al. 2009 Synthesis and non-linear optical characterization of novel borinate derivatives of cinnamaldehyde. New J. Chem. 33, 1693–1702. (doi:10.1039/B804356G)

11. Sharma UK, Sood S, Sharma N, Rabi P, Kumar R, Sinha AK, Gulati A. 2013 Synthesis and SAR investigation of natural phenylpropene-derived methoxyolated cinnamaldehydes and their novel Schiff bases as potent antimicrobial and antioxidant agents. Med. Chem. Res. 22, 5129–5140. (doi:10.1007/s10004-013-0484-9)

12. Adabiardakani A, Hakimi M, Kargar H. 2012 Cinnamaldehyde Schiff base derivatives: a short review. World Appl. Program. 2, 472–476.

13. Kudrat-E-Zahan M, Islam M, Bashar MA. 2015 Synthesis, characteristics, and antimicrobial activity of some complexes of Mn (II), Fe (III) Cu (II), Ni (II), Cu (II), and Sb (III) containing bidentate Schiff base of SMDTC. Russ. J. Gen. Chem. 85, 667–672. (doi:10.1134/S1070363215080238)

14. Wei Q-Y, Xiong J-J, Jiang H, Zhang C, Ye W. 2011 The antimicrobial activities of the cinnamaldehyde adducts with amino acids. Int. J. Food Microbiol. 150, 164–170. (doi:10.1016/j.ijfoodmicro.2011.07.034)

15. Wang H, Yuan H, Li S, Li Z, Jiang M. 2016 Synthesis, antimicrobial activity of Schiff base compounds of cinnamaldehyde and amino acids. Bioorg. Med.
16. Deep A, Narasimhan B, Lim SM, Ramasamy K, Mishra RK, Mani V. 2016 4-Thiazolidinone derivatives: synthesis, antimicrobial, anticancer evaluation and QSAR studies. RSC Adv. 6, 109 485–109 494. (doi:10.1039/C6RA23006G)

17. Cramer RD, Bunce JD, Patterson, DE, Frank IE. 1988 Crossvalidation, bootstrapping, and partial least squares compared with multiple regression in conventional QSAR studies. Quant. Struct.-Act. Relatsh 7, 18–25. (doi:10.1002/qsar.19880070105)

18. Wang H, Nguyen TTH, Li S, Liang T, Zhang Y, Li J. 2015 Quantitative structure–activity relationship of antifungal activity of rosiderivatives. Bioorg. Med. Chem. Lett. 25, 347–354. (doi:10.1016/j.bmcl.2014.11.034)

19. Srivani P, Srinivas E, Raghu R, Sastry GN. 2007 Molecular modeling studies of pyridopurinone derivatives—potential phosphodiesterase 5 inhibitors. J. Mol. Graph. Model. 26, 378–390. (doi:10.1016/j.jmgm.2007.01.007)

20. Li J, Gao Y, Shang S, Rao X, Song J, Wang Z. 2014 Synthesis and quantitative structure–activity relationship (QSAR) studies of novel resin-based diamide insecticides. RSC Adv. 4, 58 190–58 199. (doi:10.1039/C4RA1025A)

21. Worachartcheewan A, Nantaenamrat C, Owasirikul W, Nonor T, Naruepantawat O, IyanapaIarn S, Prachayasittikul S, Prachayasittikul V. 2014 Insights into antioxidant activity of 1-adamantylthiopyridine analog using multiple linear regression. Eur. J. Med. Chem. 73, 258–264. (doi:10.1016/j.ejmech.2013.11.038)

22. Katarzynk AR, Stoyanova-Slavova IB, Dobchev DA, Karelson M. 2007 QSAR modeling of flash points: An update. J. Mol. Graph. Model. 26, 529–536. (doi:10.1016/j.jmgm.2007.03.006)

23. Fateni MH, Ghorbanzad’e M. 2019 In silico prediction of nematic transition temperature for liquid crystals using quantitative structure–property relationship approaches. Mol. Divers. 13, 483–491. (doi:10.1007/s11030-009-9115-y)

24. Marković V et al. 2011 Antiproliferative activity and QSAR studies of a series of new 4-aminomethylene derivatives of some pyrazol-5-ones. Bioorg. Med. Chem. Lett. 23, 4416–4421. (doi:10.1016/j.bmcl.2011.06.025)

25. Roy PP, Roy K. 2009 QSAR studies of CYP2D6 inhibitor aryloxypropanolamines using 2D and 3D descriptors. Chem. Biol. Drug Des. 73, 442–455. (doi:10.1111/j.1747-0285.2009.00791.x)

26. Massarelli I, Coi A, Pietra D, Nofal FA, Biagi G, Giorgi I, Leonardi M, Fiamingo F, Bianucci A. 2008 QSAR study on a novel series of 8-azaazadine analogues proposed as A1 adenosine receptor antagonists. Eur. J. Med. Chem. 43, 114–121. (doi:10.1016/j.ejmech.2007.02.009)

27. Xia B, Ma W, Zhang X, Fan B. 2007 Quantitative structure–retention relationship study of the constituents of saffron aroma in SPME-GC–MS based on the projection pursuit regression method. Talanta 77, 360–365. (doi:10.1016/j.talanta.2008.06.038)

30. Du H, Wang J, Hu Z, Yao X. 2008 Quantitative structure-retention relationship study of the constituents of saffron aroma in SPME–GC–MS based on the projection pursuit regression method. Talanta 77, 360–365. (doi:10.1016/j.talanta.2008.06.038)

31. Ivanescu O, Ivanescu T, Filip PA, Cabrol-Bass D. 1999 Estimation of the liquid viscosity of organic compounds with a quantitative structure–property model. J. Chem. Inform. Comput. Sci. 39, 515–524. (doi:10.1021/ci980117v)

32. Sun H, Zhao T, Zhang X, Song F. 2015 Series of 5-HT2C Agonists. J. Comput. Sci. Eng. 21, 671–676.

33. Couling DJ, Bernet RJ, Docherty KM, Dixon JK, Maginn EJ. 2006 Assessing the factors responsible for ionic liquid toxicity to aquatic organisms via quantitative structure–property relationship modeling. Green Chem. 8, 82–90. (doi:10.1039/B511333D)

34. Teixeira C, Barbault F, Rebehmed J, Liu K, Xie L, Lu H, Jiang S, Fan B, Maurel F. 2008 Molecular modeling studies of N-substituted pyrrole derivatives—potential HIV-1 gp41 inhibitors. Bioorg. Med. Chem. 16, 3029–3048. (doi:10.1016/j.bmc.2007.12.014)

35. Liu D, Thomson K, Kaiser K. 1982 Quantitative structure-toxicity relationship of halogenated phenols on bacteria. Bull. Environ. Contam. Toxicol. 29, 130–136. (doi:10.1007/BF01606940)