JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

Synthesis and Quantitative Structure−Activity Relationship (QSAR) Study of Novel *N*-Arylsulfonyl-3-acylindole Arylcarbonyl Hydrazone Derivatives as Nematicidal Agents

Zhiping Che,[†] Shaoyong Zhang,[‡] Yonghua Shao,[#] Lingling Fan,[†] Hui Xu,*[,†] Xiang Yu,[†] Xiaoyan Zhi,[†] Xiaojun Yao,[#] and Rui Zhang[†]

[†]Laboratory of Pharmaceutical Design and Synthesis, College of Sciences, Northwest A&F University, Yangling 712100, Shaanxi Province, People's Republic of China

[‡]Provincial Engineering Laboratory of Biopesticide Preparation, Zhejiang A&F University, Lin'an 311300, Zhejiang Province, People's Republic of China

[#]Department of Chemistry, Lanzhou University, Lanzhou 730000, Gansu Province, People's Republic of China

Supporting Information

ABSTRACT: In continuation of our program aimed at the discovery and development of natural-product-based pesticidal agents, 54 novel *N*-arylsulfonyl-3-acylindole arylcarbonyl hydrazone derivatives were prepared, and their structures were well characterized by ^{1}H NMR, ^{13}C NMR, HRMS, ESI-MS, and mp. Their nematicidal activity was evaluated against that of the pine wood nematode, *Bursaphelenchus xylophilus* in vivo. Among all of the derivatives, especially **V-12** and **V-39** displayed the best promising nematicidal activity with LC_{50} values of 1.0969 and 1.2632 mg/L, respectively. This suggested that introduction of R^1 and R^2 together as the electron-withdrawing substituents, R^3 as the methyl group, and R^4 as the phenyl with the electron-donating substituents could be taken into account for further preparation of these kinds of compounds as nematicidal agents. Six selected descriptors are a WHIM descriptor (E1m), two GETAWAY descriptors (R1m+ and R3m+), a Burden eigenvalues descriptor (BEHm8), and two edge-adjacency index descriptors (EEig05x and EEig13d). Quantitative structure−activity relationship (QSAR) studies demonstrated that the structural factors, such as molecular mass (a negative correlation with the bioactivity) and molecular polarity (a positive correlation with bioactivity), are likely to govern the nematicidal activities of these compounds. For this model, the correlation coefficient ($R^2_{\text{training set}}$), the leave-one-out cross-validation correlation coefficient (Q^2_{LOO}), and the 7-fold cross-validation correlation coefficient ($Q^2_{\text{7-fold}}$) were 0.791, 0.701, and 0.715, respectively. The external cross-validation correlation coefficient (Q^2_{ext}) and the root-mean-square error for the test set ($\text{RMSE}_{\text{test set}}$) were 0.774 and 3.412, respectively. This study will pave the way for future design, structural modification, and development of indole derivatives as nematicidal agents.

KEYWORDS: *indole, hydrazone, structural modification, botanical pesticide, nematicidal activity, QSAR, Bursaphelenchus xylophilus*

INTRODUCTION

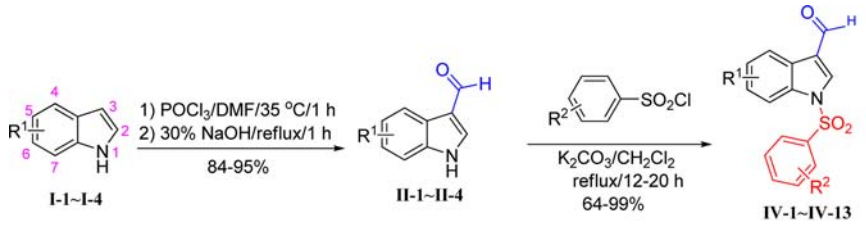
Pine wood nematode, *Bursaphelenchus xylophilus*, is the cause of pine wilt disease, which has been devastating forests worldwide.[1] Moreover, at present there are only a few commercial nematicides left in use, and their repeated applications over the years have led to the enhancement of biodegradation mechanisms in soil and the development of pest resistance.[2−4]

To prevent pine wilt disease and overcome the problems of resistance development and environmental pollution, therefore, the research and development of efficacious nematicidal agents has received much attention internationally in recent years.[5−11] In the meantime, during the long period of evolution, plants must resist attackers over their lifetime by producing and exuding secondary metabolites, and pesticides produced from plant secondary metabolites may result in less or slower resistance development and lower pollution.[12,13] Hence, the discovery of new pesticidal compounds directly from plant secondary metabolites, or by using them as lead compounds for further structural modifications, has recently been one of the important procedures in the research and development of new pesticides.[14−17] Some botanical pesticides such as nicotine, pyrethrum, and neem extracts are characteristic examples made from plants as defenses against pests.[18]

Indole (**I-1**, Scheme 1), an aromatic heterocyclic compound, is a constituent of many natural plants, such as *Robinia pseudacacia*, jasmines, certain citrus plants, and orange blossoms. Due to its crucial heterocyclic skeleton, extensive efforts using **I-1** as a lead compound have been made for the preparation of potent anti-human immunodeficiency virus type 1 (HIV-1) inhibitors (e.g., delavirdine),[19,20] hepatitis C virus (HCV) inhibitors,[21] antimicrobial agents,[22] glutamate carboxypeptidase II (GCPII) inhibitors,[23] antifungal agents,[24] and so on. In contrast, to the best of our knowledge, little work has been conducted on the structural modifications of indoles as nematicidal agents against the pine wood nematode, *B. xylophilus*. Recently, we have found

Received: February 3, 2013
Revised: May 19, 2013
Accepted: May 20, 2013
Published: May 20, 2013

Scheme 1. Synthetic Route for the Preparation of IV-1−IV-13

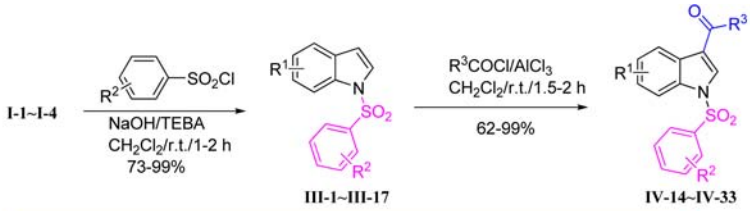


I-1: R^1 = H
I-2: R^1 = 6-Me
I-3: R^1 = 5-CN
I-4: R^1 = 5-NO_2

II-1: R^1 = H
II-2: R^1 = 6-Me
II-3: R^1 = 5-CN
II-4: R^1 = 5-NO_2

IV-1: R^1 = H, R^2 = 4-Me; **IV-2**: R^1 = H, R^2 = 4-OMe; **IV-3**: R^1 = H, R^2 = 4-NHAc; **IV-4**: R^1 = H, R^2 = 4-Cl; **IV-5**: R^1 = H, R^2 = 3-NO_2; **IV-6**: R^1 = H, R^2 = 3-NO_2, 4-Cl; **IV-7**: R^1 = 6-Me, R^2 = 4-Me; **IV-8**: R^1 = 6-Me, R^2 = 4-OMe; **IV-9**: R^1 = 6-Me, R^2 = 4-Cl; **IV-10**: R^1 = 6-Me, R^2 = 3-NO_2; **IV-11**: R^1 = 5-CN, R^2 = 4-Me; **IV-12**: R^1 = 5-CN, R^2 = 3-NO_2; **IV-13**: R^1 = 5-NO_2, R^2 = 4-Me.

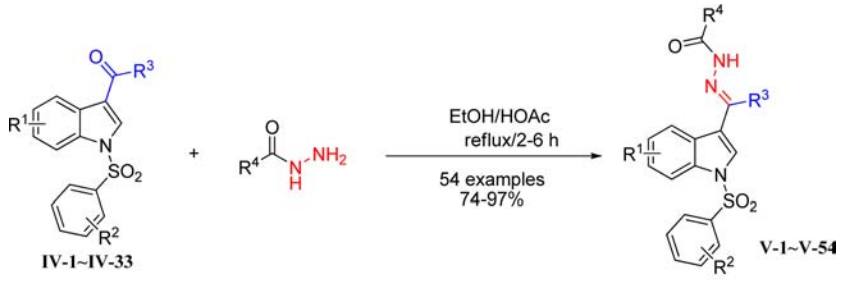
Scheme 2. Synthetic Route for the Preparation of IV-14−IV-33



III-1: R^1 = R^2 = H; **III-2**: R^1 = H, R^2 = 4-Me; **III-3**: R^1 = H, R^2 = 4-OMe; **III-4**: R^1 = H, R^2 = 4-Cl; **III-5**: R^1 = H, R^2 = 3-NO_2; **III-6**: R^1 = 6-Me, R^2 = H; **III-7**: R^1 = 6-Me, R^2 = 4-Me; **III-8**: R^1 = 6-Me, R^2 = 4-Et; **III-9**: R^1 = 6-Me, R^2 = 4-OMe; **III-10**: R^1 = 6-Me, R^2 = 4-Cl; **III-11**: R^1 = 6-Me, R^2 = 3-NO_2; **III-12**: R^1 = 5-CN, R^2 = H; **III-13**: R^1 = 5-CN, R^2 = 4-Me; **III-14**: R^1 = 5-CN, R^2 = 4-Cl; **III-15**: R^1 = 5-NO_2, R^2 = H; **III-16**: R^1 = 5-NO_2, R^2 = 4-Me; **III-17**: R^1 = 5-NO_2, R^2 = 4-Cl.

IV-14: R^1 = R^2 = H, R^3 = Me; **IV-15**: R^1 = H, R^2 = 4-Me, R^3 = Me; **IV-16**: R^1 = H, R^2 = 4-OMe, R^3 = Me; **IV-17**: R^1 = H, R^2 = 4-Cl, R^3 = Me; **IV-18**: R^1 = H, R^2 = 3-NO_2, R^3 = Me; **IV-19**: R^1 = 6-Me, R^2 = H, R^3 = Me; **IV-20**: R^1 = 6-Me, R^2 = 4-Me, R^3 = Me; **IV-21**: R^1 = 6-Me, R^2 = 4-Et, R^3 = Me; **IV-22**: R^1 = 6-Me, R^2 = 4-OMe, R^3 = Me; **IV-23**: R^1 = 6-Me, R^2 = 4-Cl, R^3 = Me; **IV-24**: R^1 = 6-Me, R^2 = 3-NO_2, R^3 = Me; **IV-25**: R^1 = 5-CN, R^2 = H, R^3 = Me; **IV-26**: R^1 = 5-CN, R^2 = 4-Me, R^3 = Me; **IV-27**: R^1 = 5-CN, R^2 = 4-Cl, R^3 = Me; **IV-28**: R^1 = 5-NO_2, R^2 = H, R^3 = Me; **IV-29**: R^1 = 5-NO_2, R^2 = 4-Me, R^3 = Me; **IV-30**: R^1 = 5-NO_2, R^2 = 4-Cl, R^3 = Me; **IV-31**: R^1 = H, R^2 = H, R^3 = Et; **IV-32**: R^1 = H, R^2 = H, R^3 = *n*-pentyl; **IV-33**: R^1 = H, R^2 = 4-Me, R^3 = *n*-pentyl.

Scheme 3. Synthetic Route for the Preparation of V-1−V-54



R^1 = H, Me, CN, NO_2; R^2 = H, Me, Et, OMe, NHAc, Cl, NO_2; R^3 = H, Me, Et, *n*-pentyl; R^4 = Ph, Ph(*m*-Me), Ph(*p*-OMe), Ph(*m*-Cl), Ph(*p*-NO_2), Ph(*p*-OH), 2-thienyl, 3-pyridyl.

that some fraxinellone-based hydrazone derivatives exhibited pronounced insecticidal activity,[25] and some *N*-arylsulfonyl-3-acetylindoles showed potent anti-HIV-1 activity.[26] In continuation of our program aimed at the discovery and development of novel natural-product-based pesticidal agents,[24−28] consequently, we herein synthesized 54 novel *N*-arylsulfonyl-3-acylindole arylcarbonyl hydrazone derivatives (**V-1−V-54**; Scheme 3) by introduction of the hydrazone fragments on the *N*-arylsulfonyl-3-acylindolyl skeleton. Their nematicidal activity was evaluated against *B. xylophilus*. In addition, the quantitative structure−activity relationship (QSAR) studies of **V-1−V-54** were also investigated.

dx.doi.org/10.1021/jf400536q | *J. Agric. Food Chem.* 2013, 61, 5696−5705

MATERIALS AND METHODS

General. All reagents and solvents were of reagent grade or purified according to standard methods before use. Analytical thin-layer chromatography (TLC) and preparative thin-layer chromatography (PTLC) were performed with silica gel plates using silica gel 60 GF_{254} (Qingdao Haiyang Chemical Co., Ltd.). Melting points were determined on an XT-4 digital melting-point apparatus (Beijing Tech Instrument Co., Ltd.) and were uncorrected. Infrared spectra (IR) were recorded on a Bruker TENSOR 27 spectrometer. Nuclear magnetic resonance spectra (NMR) were recorded on a Bruker Avance DMX 300, 400, or 500 MHz instrument in $CDCl_3$ or DMSO-d_6 (^{1}H at 300, 400, or 500 MHz and ^{13}C at 125 MHz) using tetramethylsilane (TMS) as the internal standard. Electrospray ion trap mass spectrometry (ESI-TRAP-MS) and electron ionization mass spectra (EI-MS) were carried out with a Bruker ESI-TRAP Esquire 6000 plus mass spectrometry instrument and an HP 5988 instrument, respectively. High-resolution mass spectra (HR-MS) were carried out with an IonSpec 4.7 T FTMS instrument.

*Synthesis of 3-Formylindoles (**II-1**–**II-4**).* A mixture of *N*,*N*-dimethylformamide (DMF, 5 mL) and phosphorus oxychloride ($POCl_3$, 0.5 mL) was stirred at 0 °C for 20 min. Then a solution of indoles (**I-1**–**I-4**, 5 mmol) in DMF (2 mL) was added dropwise to the above mixture. After the addition, the mixture was stirred at 35 °C for 1 h, and water was added, followed by the addition of 30% aqueous sodium hydroxide (NaOH) to adjust the pH value to 8–9. The mixture was then refluxed for 1 h. On cooling, the mixture was poured into ice water, and the precipitated product was collected, washed by water, and recrystallized from absolute methanol to afford **II-1**–**II-4** in 84–95% yields.

*Data for **II-1**:* yield = 91%; pink solid; mp = 190–192 °C [lit., 195–198 °C];[29] ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.93 (s, 1H), 8.29 (s, 1H), 8.08 (d, J = 8.0 Hz, 1H), 7.50 (d, J = 8.0 Hz, 1H), 7.20–7.28 (m, 2H); EI-MS, m/z (%) 145 (M^+, 96).

*Data for **II-2**:* yield = 95%; brown solid; mp = 187–189 °C [lit., 186–188 °C];[29] ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.89 (d, J = 2.0 Hz, 1H), 8.21 (s, 1H), 7.95 (d, J = 8.4 Hz, 1H), 7.30 (s, 1H), 7.04 (d, J = 8.0 Hz, 1H), 2.41 (s, 3H); EI-MS, m/z (%) 159 (M^+, 68).

*Data for **II-3**:* yield = 88%; yellow solid; mp = 241–243 °C [lit., 244–245 °C];[29] ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.01 (s, 1H), 8.52 (s, 1H), 8.47 (s, 1H), 7.71 (d, J = 8.4 Hz, 1H), 7.65 (d, J = 8.0 Hz, 1H); EI-MS, m/z (%) 170 (M^+, 70).

*Data for **II-4**:* yield = 84%; yellow solid; mp >300 °C [lit., 312–313.3 °C];[30] ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.03 (s, 1H), 8.94 (d, J = 2.0 Hz, 1H), 8.58 (s, 1H), 8.15 (dd, J = 2.4, 9.2 Hz, 1H), 7.71 (d, J = 9.2 Hz, 1H); EI-MS, m/z (%) 190 (M^+, 100).

*Synthesis of N-Arylsulfonylindoles (**III-1**–**III-17**).* A mixture of **I-1**–**I-4** (1 mmol), benzyltriethylammonium chloride (TEBA, 0.1 mmol), NaOH (1.8 mmol), and arylsulfonyl chlorides (1.2 mmol) in dry CH_2Cl_2 (5 mL) was stirred at room temperature. After 1–2 h of stirring, the reaction was complete according to TLC analysis, and water (10 mL) was added to the mixture, which was extracted with CH_2Cl_2 (30 mL × 3). Subsequently, the combined organic phase was washed by brine (30 mL), dried over anhydrous Na_2SO_4, and concentrated under reduced pressure to give **III-1**–**III-17** in 73–99% yields. The example data of **III-1** and **III-2** are shown as follows, whereas data of **III-3**–**III-17** can be found in the Supporting Information.

*Data for **III-1**:* yield = 97%; white solid; mp = 78–79 °C [lit., 78–79 °C];[31] ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.99 (dd, J = 8.4, 4.0 Hz, 1H), 7.87 (d, J = 8.0 Hz, 2H), 7.56 (d, J = 3.6 Hz, 1H), 7.50 (d, J = 8.0 Hz, 2H), 7.40 (d, J = 7.6 Hz, 2H), 7.29 (d, J = 8.4 Hz, 1H), 7.21 (d, J = 7.6 Hz, 1H), 6.66 (d, J = 2.8 Hz, 1H); EI-MS, m/z (%) 257 (M^+, 85).

*Data for **III-2**:* yield = 96%; white solid; mp = 83–84 °C [lit., 87–88 °C];[32] ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.97 (d, J = 10.8 Hz, 1H), 7.75 (d, J = 10.8 Hz, 2H), 7.56 (d, J = 4.4 Hz, 1H), 7.97 (d, J = 10.0 Hz, 1H), 7.20 (m, 4H), 6.65 (d, J = 4.8 Hz, 1H), 2.33 (s, 3H); EI-MS, m/z (%) 271 (M^+, 100).

*Synthesis of 3-Formyl-N-arylsulfonylindoles (**IV-1**–**IV-13**).* A mixture of **II-1**–**II-4** (2 mmol), arylsulfonyl chlorides (4 mmol), and K_2CO_3 (6 mmol) in dry CH_2Cl_2 (10 mL) was refluxed for 12–20 h. Then the reaction mixture was filtered. The corresponding filtrate was collected, concentrated under reduced pressure, and purified by PTLC to produce **IV-1**–**IV-13** in 64–99% yields. The example data of **IV-1** and **IV-2** are shown as follows, whereas data of **IV-3**–**IV-13** can be found in the Supporting Information.

*Data for **IV-1**:* yield = 71%; white solid; mp = 148–150 °C [lit., 149 °C];[33] ^{1}H NMR (500 MHz, $CDCl_3$) δ 10.09 (s, 1H), 8.25 (d, J = 7.5 Hz, 1H), 8.22 (s, 1H), 7.95 (d, J = 8.0 Hz, 1H), 7.85 (d, J = 8.5 Hz, 2H), 7.34–7.42 (m, 2H), 7.29 (d, J = 8.0 Hz, 2H), 2.36 (s, 3H); ESI-MS, m/z (%) 300 ([M + H]$^+$, 100).

*Data for **IV-2**:* yield = 75%; white solid; mp = 137–139 °C [lit., 137 °C];[33] ^{1}H NMR (500 MHz, $CDCl_3$) δ 10.08 (s, 1H), 8.26 (d, J = 7.5 Hz, 1H), 8.22 (s, 1H), 7.95 (d, J = 8.5 Hz, 1H), 7.89–7.91 (m, 2H), 7.34–7.42 (m, 2H), 6.94 (d, J = 9.0 Hz, 2H), 3.81 (s, 3H); ESI-MS, m/z (%) 316 ([M + H]$^+$, 25).

*Synthesis of 3-Acyl-N-Arylsulfonylindoles (**IV-14**–**IV-33**).* To a stirred mixture of $AlCl_3$ (3 mmol) and R^3COCl (acetyl chloride, propionyl chloride, or *n*-hexanoyl chloride, 1.5 mmol) in dry CH_2Cl_2 (5 mL) at room temperature was added dropwise a solution of **III-1**–**III-17** (1 mmol) in dry CH_2Cl_2 (2 mL). After the addition, the mixture was stirred at room temperature for 1.5–2 h, and the reaction process was checked by TLC analysis. Then water (10 mL) was added to the mixture, which was extracted with CH_2Cl_2 (30 mL × 3). Subsequently, the combined organic phase was washed with saturated aqueous $NaHCO_3$ (30 mL × 2) and brine (30 mL), dried over anhydrous Na_2SO_4, concentrated in vacuo, and purified by PTLC to afford **IV-14**–**IV-33** in 62–99% yields. The example data of **IV-14** and **IV-15** are shown as follows, whereas data of **IV-16**–**IV-33** can be found in the Supporting Information.

*Data for **IV-14**:* yield = 72%; white solid; mp = 158–159 °C [lit., 159–160 °C];[34] ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.32–8.34 (m, 1H), 8.21 (s, 1H), 7.92–7.97 (m, 3H), 7.48–7.61 (m, 3H), 7.34–7.39 (m, 2H), 2.58 (s, 3H); EI-MS, m/z (%) 299 (M^+, 74).

*Data for **IV-15**:* yield = 91%; white solid; mp = 142–144 °C [lit., 145–146 °C];[35] ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.31–8.34 (m, 1H), 8.21 (s, 1H), 7.90–7.94 (m, 1H), 7.82 (d, J = 8.4 Hz, 2H), 7.26–7.37 (m, 4H), 2.57 (s, 3H), 2.36 (s, 3H); EI-MS, m/z (%) 313 (M^+, 35).

*Synthesis of N-Arylsulfonyl-3-acylindole Arylcarbonyl Hydrazone Derivatives (**V-1**–**V-54**).* A mixture of **IV-1**–**IV-33** (0.5 mmol), the corresponding hydrazides (0.5 mmol), and HOAc (2 drops) in ethanol (5 mL) was refluxed for 2–6 h. When the reaction was complete according to TLC analysis, the mixture was allowed to cool and filtered to give the solid, which was further recrystallized from absolute ethanol to produce target compounds **V-1**–**V-54** in 74–97% yields. The example data of **V-1**–**V-20** are shown as follows, whereas data of **V-21**–**V-54** can be found in the Supporting Information.

*Data for **V-1**:* yield = 90%; white solid; mp = 215–216 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.63 (s, 1H), 8.45 (d, J = 7.5 Hz, 1H), 8.36 (s, 1H), 7.92–7.98 (m, 5H), 7.61 (d, J = 6.5 Hz, 1H), 7.56 (d, J = 7.0 Hz, 2H), 7.40–7.44 (m, 4H), 2.31 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 163.4, 146.3, 142.6, 135.2, 134.2, 133.9, 132.1, 130.8, 130.3, 128.9, 128.0, 127.4, 127.3, 126.3, 124.7, 123.8, 118.7, 113.5, 21.5; ESI-MS, m/z (%) 418 ([M + H]$^+$, 100). HRMS (ESI): calcd for $C_{23}H_{20}N_3O_3S$ ([M + H])$^+$, 418.1219; found, 418.1226.

*Data for **V-2**:* yield = 83%; white solid; mp = 182–183 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.64 (s, 1H), 8.45 (d, J = 7.5 Hz, 1H), 8.36 (s, 1H), 7.97–8.00 (m, 3H), 7.95 (d, J = 7.5 Hz, 2H), 7.62 (t, J = 7.5 Hz, 1H), 7.56 (t, J = 7.5 Hz, 2H), 7.46 (t, J = 7.5 Hz, 1H), 7.41 (t, J = 7.5 Hz, 1H), 7.11 (d, J = 9.0 Hz, 2H), 3.79 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 164.4, 142.7, 135.2, 133.9, 132.1, 130.3, 129.8, 128.9, 128.4, 128.0, 127.4, 126.2, 124.6, 123.7, 118.5, 115.6, 113.5, 56.3; ESI-MS, m/z (%) 434 ([M + H]$^+$, 100). HRMS (ESI): calcd for $C_{23}H_{20}N_3O_4S$ ([M + H])$^+$, 434.1169; found, 434.1161.

*Data for **V-3**:* yield = 79%; white solid; mp = 267–268 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.64 (s, 1H), 8.45 (d, J = 7.5 Hz, 1H), 8.33 (s, 1H), 7.93–8.00 (m, 5H), 7.79 (d, J = 8.0 Hz, 2H), 7.55–7.61 (m, 3H), 7.39–7.45 (m, 2H), 2.05 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 169.6, 163.4, 145.2 142.6, 135.2, 133.9, 132.1, 130.3, 130.1, 128.9, 128.8, 128.0, 127.4, 126.2, 124.6, 123.7, 119.3, 118.5,

113.5, 24.5; ESI-MS, m/z (%) 483 ([M + Na]$^+$, 100). HRMS (ESI): calcd for $C_{24}H_{21}N_4O_4S$ ([M + H])$^+$, 461.1278; found, 461.1284.

Data for ***V-4****:* yield = 97%; white solid; mp = 208−209 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.63 (s, 1H), 8.47 (d, J = 7.5 Hz, 1H), 8.39 (s, 1H), 8.08 (d, J = 8.5 Hz, 2H), 7.94−7.99 (m, 3H), 7.70 (d, J = 7.0 Hz, 2H), 7.62 (t, J = 7.0 Hz, 1H), 7.56 (t, J = 7.5 Hz, 2H), 7.48 (t, J = 7.5 Hz, 1H), 7.43 (t, J = 7.5 Hz, 1H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 163.4, 142.4, 140.5, 135.8, 135.2, 133.8, 132.2, 130.6, 130.2, 129.2, 128.9, 128.1, 127.5, 126.5, 124.9, 123.9, 119.1, 113.5; ESI-MS, m/z (%) 460 ([M + Na]$^+$, 100). HRMS (ESI): calcd for $C_{22}H_{17}N_3O_3SCl$ ([M + H])$^+$, 438.0673; found, 438.0681.

Data for ***V-5****:* yield = 76%; yellow solid; mp = 155−156 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.70 (s, 1H), 8.62 (s, 1H), 8.50−8.53 (m, 3H), 8.46 (d, J = 8.0 Hz, 1H), 8.05 (d, J = 8.5 Hz, 1H), 7.90−7.93 (m, 3H), 7.62 (t, J = 7.0 Hz, 1H), 7.56 (t, J = 7.5 Hz, 2H), 7.51 (t, J = 8.0 Hz, 1H), 7.44 (t, J = 7.5 Hz, 1H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 163.4, 148.6, 142.3, 138.3, 135.2, 133.8, 133.0, 132.5, 132.2, 130.2, 129.9, 128.9, 128.0, 127.6, 126.7, 125.2, 124.0, 122.0, 119.6, 113.5; ESI-MS, m/z (%) 471 ([M + Na]$^+$, 70). HRMS (ESI): calcd for $C_{22}H_{17}N_4O_5S$ ([M + H])$^+$, 449.0914; found, 449.0917.

Data for ***V-6****:* yield = 78%; yellow solid; mp = 217−218 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.81 (s, 1H), 8.63 (s, 1H), 8.48 (d, J = 7.5 Hz, 1H), 8.41 (s, 1H), 8.34−8.35 (m, 1H), 8.02−8.04 (m, 2H), 7.95 (d, J = 7.5 Hz, 2H), 7.63 (t, J = 7.0 Hz, 1H), 7.56 (t, J = 7.5 Hz, 2H), 7.51 (t, J = 7.5 Hz, 1H), 7.46 (t, J = 7.5 Hz, 1H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 163.4, 148.3, 142.2, 136.8, 135.1, 134.2, 133.8, 132.4, 132.2, 131.7, 130.1, 128.9, 128.1, 127.6, 126.7, 125.2, 124.7, 124.0, 119.6, 113.6; ESI-MS, m/z (%) 483 ([M + H]$^+$, 100). HRMS (ESI): calcd for $C_{22}H_{16}ClN_4O_5S$ ([M + H])$^+$, 483.0524; found, 483.0531.

Data for ***V-7****:* yield = 89%; white solid; mp = 242−243 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.60 (s, 1H), 8.30 (d, J = 8.0 Hz, 1H), 8.27 (s, 1H), 7.92−7.93 (m, 4H), 7.78 (s, 1H), 7.62 (t, J = 7.0 Hz, 1H), 7.55 (t, J = 7.0 Hz, 2H), 7.42 (d, J = 8.5 Hz, 2H), 7.23 (d, J = 8.0 Hz, 1H), 2.47 (s, 3H), 2.32 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 163.4, 146.2, 142.7, 136.0, 135.6, 134.3, 133.9, 132.1, 130.8, 129.8, 128.9, 128.0, 127.3, 126.1, 125.2, 123.4, 118.7, 113.5, 21.9, 21.5; ESI-MS, m/z (%) 454 ([M + Na]$^+$, 100). HRMS (ESI): calcd for $C_{24}H_{22}N_3O_3S$ ([M + H])$^+$, 432.1376; found, 432.1385.

Data for ***V-8****:* yield = 82%; white solid; mp = 224−225 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.60 (s, 1H), 8.30 (d, J = 8.0 Hz, 1H), 8.26 (s, 1H), 8.00 (d, J = 9.0 Hz, 2H), 7.95 (d, J = 7.5 Hz, 2H), 7.79 (s, 1H), 7.61 (t, J = 7.5 Hz, 1H), 7.55 (t, J = 7.5 Hz, 2H), 7.23 (d, J = 8.5 Hz, 1H), 7.11 (d, J = 9.0 Hz, 2H), 3.79 (s, 3H), 2.48 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 164.4, 142.7, 136.0, 135.6, 134.0, 133.9, 132.1, 129.8, 129.7, 128.9, 128.5, 128.0, 126.0, 125.2, 123.4, 118.5, 115.5, 113.5, 56.3, 22.0; ESI-MS, m/z (%) 448 ([M + H]$^+$, 100). HRMS (ESI): calcd for $C_{24}H_{22}N_3O_4S$ ([M + H])$^+$, 448.1325; found, 448.1330.

Data for ***V-9****:* yield = 84%; white solid; mp = 234−235 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.59 (s, 1H), 8.28−8.31 (m, 2H), 8.08 (d, J = 8.5 Hz, 2H), 7.94 (d, J = 7.5 Hz, 2H), 7.78 (s, 1H), 7.71 (d, J = 8.5 Hz, 2H), 7.62 (t, J = 7.0 Hz, 1H), 7.55 (t, J = 7.5 Hz, 2H), 7.25 (d, J = 8.0 Hz, 1H), 2.48 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 163.4, 142.5, 140.4, 136.3, 135.9, 135.6, 133.8, 132.2, 130.6, 129.6, 129.2, 128.9, 128.0, 126.3, 125.2, 123.5, 119.1, 113.4, 21.9; ESI-MS, m/z (%) 452 ([M + H]$^+$, 100). HRMS (ESI): calcd for $C_{23}H_{19}N_3O_3SCl$ ([M + H])$^+$, 452.0830; found, 452.0832.

Data for ***V-10****:* yield = 82%; yellow solid; mp = 189−190 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.70 (s, 1H), 8.59 (s, 1H), 8.52 (dd, J = 8.0 Hz, 1.5 Hz, 2H), 8.40 (s, 1H), 8.32 (d, J = 8.0 Hz, 1H), 7.91−7.94 (m, 3H), 7.85 (s, 1H), 7.62 (d, J = 7.0 Hz, 1H), 7.55 (t, J = 7.5 Hz, 2H), 7.26 (d, J = 8.0 Hz, 1H), 2.49 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 163.4, 148.6, 142.4, 138.4, 136.6, 135.6, 133.8, 133.0, 132.5, 132.2, 129.8, 129.7, 128.9, 128.0, 126.5, 125.3, 123.6, 122.0, 119.6, 113.5, 21.9; ESI-MS, m/z (%) 463 ([M + H]$^+$, 100). HRMS (ESI): calcd for $C_{23}H_{19}N_4O_5S$ ([M + H])$^+$, 463.1070; found, 463.1078.

Data for ***V-11****:* yield = 79%; white solid; mp = 219−220 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.83 (s, 1H), 8.64 (s, 1H), 8.59 (s, 1H), 8.17 (d, J = 8.5 Hz, 1H), 8.00 (d, J = 8.0 Hz, 2H), 7.96 (d, J = 7.0 Hz, 2H), 7.87 (d, J = 9.5 Hz, 1H), 7.63 (t, J = 7.5 Hz, 1H), 7.57 (t, J = 7.5 Hz, 2H), 7.46 (d, J = 8.0 Hz, 2H), 2.35 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 163.5, 146.9, 141.7, 137.0, 133.8, 133.6, 132.3, 132.1, 131.0, 129.3, 128.9, 128.5, 128.1, 127.5, 119.5, 118.1, 114.8, 107.4, 21.5; ESI-MS, m/z (%) 443 ([M + H]$^+$, 100). HRMS (ESI): calcd for $C_{24}H_{19}N_4O_3S$ ([M + H])$^+$, 443.1172; found, 443.1174.

Data for ***V-12****.* yield = 93%; yellow solid; mp = 250−251 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.82 (s, 1H), 8.75 (s, 1H), 8.70 (s, 1H), 8.62 (d, J = 2.5 Hz, 1H), 8.56 (t, J = 6.5 Hz, 2H), 8.23 (d, J = 8.5 Hz, 1H), 7.88−7.95 (m, 4H), 7.62 (d, J = 7.0 Hz, 1H), 7.56 (t, J = 7.5 Hz, 2H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 163.6, 148.7, 141.4, 138.0, 137.0, 133.6, 133.2, 132.7, 132.3, 132.0, 130.3, 129.7, 128.9, 128.5, 128.1, 127.7, 122.4, 119.4, 118.9, 114.9, 107.8; ESI-MS, m/z (%) 474 ([M + H]$^+$, 100). HRMS (ESI): calcd for $C_{23}H_{16}N_5O_5S$ ([M + H])$^+$, 474.0866; found, 474.0878.

Data for ***V-13****:* yield = 74%; yellow solid; mp = 157−158 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.30 (s, 1H), 8.67 (s, 1H), 8.63 (s, 1H), 8.32 (dd, J = 9.5, 2.0 Hz, 1H), 8.21 (d, J = 9.0 Hz, 1H), 8.02 (d, J = 8.0 Hz, 2H), 7.96 (d, J = 7.5 Hz, 2H), 7.63 (t, J = 7.5 Hz, 1H), 7.96 (t, J = 7.5 Hz, 2H), 7.47 (d, J = 8.5 Hz, 2H), 2.35 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 162.9, 146.9, 146.5, 144.2, 142.1, 141.3, 137.5, 133.2, 133.1, 131.7, 130.5, 128.4, 128.0, 127.5, 127.0, 126.8, 121.1, 118.4, 21.0; ESI-MS, m/z (%) 463 ([M + H]$^+$, 100). HRMS (ESI): calcd for $C_{23}H_{19}N_4O_5S$ ([M + H])$^+$, 463.1070; found, 463.1081.

Data for ***V-14****:* yield = 93%; white solid; mp = 205−206 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.74 (d, J = 4.0 Hz, 1H), 8.34 (s, 1H), 8.09 (d, J = 7.5 Hz, 2H), 7.92−7.97 (m, 3H), 7.72 (t, J = 7.0 Hz, 1H), 7.58−7.63 (m, 3H), 7.53−7.54 (m, 2H), 7.36−7.41 (m, 2H), 2.47 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 163.9, 150.7, 136.6, 134.7, 133.9, 131.4, 129.8, 128.2, 127.8, 127.4, 126.7, 125.4, 124.6, 124.1, 121.7, 112.7, 14.7; ESI-MS, m/z (%) 418 ([M + H]$^+$, 100). HRMS (ESI): calcd for $C_{23}H_{20}N_3O_3S$ ([M + H])$^+$, 418.1219; found, 418.1228.

Data for ***V-15****:* yield = 84%; white solid; mp = 201−202 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.72 (d, J = 5.5 Hz, 1H), 8.33 (s, 1H), 8.09 (d, J = 7.5 Hz, 2H), 7.98 (d, J = 7.0 Hz, 1H), 7.72 (t, J = 7.5 Hz, 3H), 7.63 (t, J = 7.5 Hz, 2H), 7.36−7.40 (m, 4H), 2.47 (s, 3H), 2.40 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 163.9, 150.6, 140.3, 137.5, 136.6, 134.7, 133.8, 132.0, 129.8, 128.1, 127.8, 127.4, 126.7, 125.4, 125.0, 124.1, 121.7, 112.7, 20.8, 14.7; ESI-MS, m/z (%) 432 ([M + H]$^+$, 100). HRMS (ESI): calcd for $C_{24}H_{22}N_3O_3S$ ([M + H])$^+$, 432.1376; found, 432.1382.

Data for ***V-16****:* yield = 90%; white solid; mp = 208−210 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.66 (s, 1H), 8.73 (s, 1H), 8.32 (s, 1H), 8.09 (d, J = 7.5 Hz, 2H), 7.97 (d, J = 8.5 Hz, 1H), 7.91 (s, 2H), 7.73 (t, J = 7.5 Hz, 1H), 7.64 (t, J = 7.5 Hz, 2H), 7.39−7.41 (m, 1H), 7.34 (s, 1H), 7.08 (d, J = 8.0 Hz, 2H), 3.85 (s, 3H), 2.47 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 162.3, 137.3, 135.3, 135.2, 130.3, 128.2, 128.0, 127.3, 126.6, 126.5, 125.9, 125.1, 124.5, 122.4, 114.0, 113.3, 55.9, 15.1; ESI-MS, m/z (%) 470 ([M + Na]$^+$, 100). HRMS (ESI): calcd for $C_{24}H_{22}N_3O_4S$ ([M + H])$^+$, 448.1325; found, 448.1320.

Data for ***V-17****:* yield = 87%; white solid; mp = 196−198 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.93 (s, 1H), 8.72 (d, J = 7.5 Hz, 1H), 8.36 (s, 1H), 8.10 (d, J = 7.0 Hz, 2H), 7.97 (s, 2H), 7.89 (d, J = 7.0 Hz, 1H), 7.74 (t, J = 7.5 Hz, 1H), 7.68 (d, J = 7.5 Hz, 1H), 7.64 (t, J = 7.5 Hz, 2H), 7.59 (t, J = 7.5 Hz, 1H), 7.41−7.42 (m, 1H), 7.36−7.38 (m, 1H), 2.49 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 163.1, 151.9, 137.2, 136.4, 135.3, 133.5, 131.8, 130.8, 130.4, 128.7, 128.1, 128.0, 127.3, 127.2, 126.0, 125.1, 124.7, 122.1, 113.3, 15.4; ESI-MS, m/z (%) 452 ([M + H]$^+$, 100). HRMS (ESI): calcd for $C_{23}H_{19}N_3O_3SCl$ ([M + H])$^+$, 452.0830; found, 452.0828.

Data for ***V-18****:* yield = 92%; white solid; mp = 220−221 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.72 (d, J = 4.0 Hz, 1H), 8.30 (s, 1H), 7.92−7.95 (m, 5H), 7.59 (d, J = 7.0 Hz, 1H), 7.54 (d, J = 7.0 Hz, 2H), 7.40 (d, J = 8.0 Hz, 3H), 7.36 (d, J = 6.5 Hz, 1H), 2.46 (s, 3H), 2.31 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 163.8, 150.7, 145.6, 139.7, 137.7, 134.7, 133.7, 131.4, 130.2, 128.2, 127.8, 127.4, 126.8, 125.3, 124.5, 124.0, 121.5, 112.7, 20.9, 14.7; ESI-MS, m/z (%) 432 ([M + H]$^+$, 100). HRMS (ESI): calcd for $C_{24}H_{22}N_3O_3S$ ([M + H])$^+$, 432.1376; found, 432.1379.

Data for ***V-19****:* yield = 83%; white solid; mp = 189−190 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.74 (d, J = 3.0 Hz, 1H), 8.31 (s, 1H), 7.96 (d, J = 7.5 Hz, 3H), 7.74 (s, 2H), 7.37−7.40 (m, 6H), 2.47 (s, 3H), 2.40 (s, 3H), 2.30 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 163.9, 150.6, 145.6, 137.5, 134.7, 133.7, 132.0, 130.1, 128.1, 127.8, 127.4, 126.8, 125.3, 124.9, 124.5, 123.9, 121.6, 112.7, 20.9, 20.8, 14.7; ESI-MS, m/z (%) 446 ([M + H]$^+$, 100). HRMS (ESI) calcd for $C_{25}H_{24}N_3O_3S$ ([M + H])$^+$, 446.1532; found, 446.1541.

Data for ***V-20****:* yield = 86%; white solid; mp = 191−192 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.74 (d, J = 5.0 Hz, 1H), 8.31 (s, 1H), 8.02 (d, J = 8.0 Hz, 2H), 7.93−7.96 (m, 3H), 7.59 (d, J = 7.0 Hz, 1H), 7.54 (d, J = 7.5 Hz, 2H), 7.35−7.42 (m, 2H), 7.10 (d, J = 9.0 Hz, 2H), 3.78 (s, 3H), 2.47 (s, 3H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 163.8, 150.9, 134.7, 133.9, 131.5, 129.3, 128.2, 127.9, 127.8, 127.4, 127.0, 125.3, 124.6, 123.9, 121.4, 115.1, 114.9, 112.7, 55.7, 14.7; ESI-MS, m/z (%) 448 ([M + H]$^+$, 100). HRMS (ESI): calcd for $C_{24}H_{22}N_3O_4S$ ([M + H])$^+$, 448.1325; found, 448.1321.

Assay of Nematicidal Activity.[36] Acetone solutions of compounds **V-1**−**V-54** and emamectin benzoate (used as a positive control) were first prepared at concentrations of 5, 10, 25, 50, and 100 mg/L, respectively. Then 10 μL of the above solutions was added to the aqueous suspension (90 μL) containing approximately 2500 living nematodes (third-instar and fourth-instar larvae of *B. xylophilus*) per milliliter. The blank control group was prepared in the same way but lacked the tested compound. Three replicates in each trial were made and kept at 25 °C for 24 h. Finally, the activities of five concentrations of the tested compounds were monitored under a microscope by recording the death rate of the tested nematodes. Nematodes that did not move when prodded with a needle were considered to be dead. The LC_{50} values of tested compounds were calculated using the probit method.

QSAR Model Development. *Data Set.* The experimental data used in this work contained 54 compounds (**V-1**−**V-54**). The nematicidal activity of 54 compounds was expressed as LC_{50} values (μmol/L) and used as the dependent variable in the following QSAR study.

Molecular Descriptor Calculation. To obtain a QSAR model, the compound was represented by structural descriptors. The molecular descriptors were calculated by the following process. All of the compound structures were sketched in the HyperChem[37] program and preoptimized with the MM+ molecular mechanics force filed. To obtain more precise optimization, the semiempirical quantum chemistry method AM1[38] was used. The resulting minimum energy conformations of the 54 compounds were input into DRAGON 5.4[39] software to calculate molecular descriptors.

In DRAGON, 1664 molecular descriptors were calculated, including (a) 0D-constitutional descriptors; (b) 1D-functional groups counts, atom-centered fragments; (c) 2D-topological descriptors, walk and path counts, connectivity indices, information indices, 2D autocorrelations, edge adjacency indices, Burden eigenvalues, topological charge index, eigenvalue-based index; (d) 3D-Randic molecular profiles, geometrical descriptors, RDF descriptors, 3D-MoRSE descriptors, WHIM descriptors,[40] GETAWAY descriptors;[41] (e) charge descriptors; and (f) molecular properties. The *Handbook of Molecular Descriptors*[42] details the calculation procedure. The list of the above-mentioned descriptors and corresponding meanings could be found in the literature references of the DRAGON package.

To obtain nonredundant information, constant or near-constant variables and two descriptors found to be correlated pairwise (one of any two descriptors with a correlation coefficient >0.99 was removed) were exlcuded. After the prereduction step, 851 molecular parameters were obtained. Thus, 851 structural descriptors were retained for subsequent subvariable selection.

Splitting Data set into Training Set and Test Set. To build and validate the QSAR model, the studied data set were divided into a training set, used to develop the model, and a test set, used to validate the external predictive ability of the proposed model. In this study, the Kennard and Stone (KS) method[43] was used to split the data set into a training set and a test set due to its good performance in other studies. The KS method can be used to rationally select training and test sets based on the descriptor space.

Feature Selection and QSAR Construction by Genetic Algorithm−Multiple Linear Regression (GA-MLR). In this work, the relationship between bioactivity and structural descriptors was built by the GA-MLR method. Genetic algorithm[44] was performed to search the descriptors pool and select the descriptors relevant to the bioactivity. Multiple linear regression is a classical linear regression method, the model constructed by which is simple and could be interpreted easily. In the present work, the GA-MLR procedure was performed by MobyDigs software[45] using the correlation coefficient of leave-one-out cross-validation (LOO) as fitness function. When increasing the number of the descriptors did not increase the cross-validated correlation coefficient (Q^2_{LOO}) value to any significant degree, the GA selection was stopped. The corresponding parameters used in the model-building process can be found as follows: population size, 100; maximum allowed descriptors in a model, 8; and reproduction/mutation trade-off, 0.5; the other parameters were set as default values.

Performance and Applicability Domain Evaluation of the QSAR Model. Several statistical parameters were adopted to assess the quality of the developed QSAR models, such as the correlation coefficient (R^2) for fitness ability, Q^2_{LOO} for internal predictive ability, and root-mean-square error (RMSE). Moreover, the 7-fold cross-validation correlation coefficient ($Q^2_{7\text{-fold}}$) was also employed to check for reliability and robustness. The external predictive power of the QSAR model was estimated by the external cross-validation correlation coefficient (Q^2_{ext}) defined as

$$Q^2_{ext} = 1 - \frac{\sum_{i=1}^{m}(y_i - y_{pred})^2}{\sum_{i=1}^{m}(y_i - \bar{y}_{tr})^2}$$

where y_i and y_{pred} are the experimental and predicted values of the bioactivity of the compounds in the test set, respectively; $\bar{y}_{tr}$ is the averaged value of the dependent variable for the training set; and m is the number of compounds in the test set.

The applicability domain is important for a proposed QSAR model, which is defined by the nature of the chemicals in the data set and can be characterized in a different way. The leverage (h) approach[46] is the commonly used methodology, which is defined as

$$h_i = x_i(X^TX)^{-1}x_i^T \quad (i = 1, ..., n)$$

where x_i is the descriptor row-vector of the query chemical and X is the $n \times k$ matrix of the data set (k is the number of model descriptors and n is the number of query compounds). The warning leverage h^* was calculated by $3k'/n$, where k' is the number of variables used in the QSAR model plus one. If the leverage value of a compound is higher than h^*, the predicted activity was the result of extrapolation of the model and may be unreliable. The Williams plot (leave-one-out cross-validated standardized errors versus leverage values) could provide an efficient way for verifying the presence of Y outliers (i.e., compounds with cross-validated standardized residuals greater than three standard deviation units, >3σ) and X outliers (i.e., compounds with leverage values greater than h^*).

RESULTS AND DISCUSSION

Synthesis. As shown in Scheme 1, 3-formylindoles (**II-1**−**II-4**) were easily obtained by Vilsmeier−Haack formylation reaction of indoles (**I-1**−**I-4**) with DMF in the presence of $POCl_3$. Subsequently, **II-1**−**II-4** reacted with arylsulfonyl chlorides to afford 3-formyl-*N*-arylsulfonylindoles (**IV-1**−**IV-13**). 3-Acyl-*N*-arylsulfonylindoles (**IV-14**−**IV-33**) were prepared as shown in Scheme 2. Starting from **I-1**−**I-4**, the arylsulfonyl substituents were first introduced at their N-1 position to afford *N*-arylsulfonylindoles (**III-1**−**III-17**). Then introduction of the different acyl groups at the C-3 position of **III-1**−**III-17** gave 3-acyl-*N*-arylsulfonylindoles (**IV-14**−**IV-33**). Finally, as described in Scheme 3, *N*-arylsulfonyl-3-acylindole

dx.doi.org/10.1021/jf400536q | J. Agric. Food Chem. 2013, 61, 5696−5705

arylcarbonyl hydrazone derivatives (**V-1**−**V-54**) were smoothly prepared by the reaction of **IV-1**−**IV-33** with the corresponding hydrazides. The structures of all target compounds were well characterized by ^{1}H NMR, ^{13}C NMR, HRMS, MS, and mp. Additionally, to confirm the three-dimensional structural information of **V-1**−**V-54**, the single-crystal structure of **V-33** was determined by X-ray crystallography as illustrated in Figure 1.

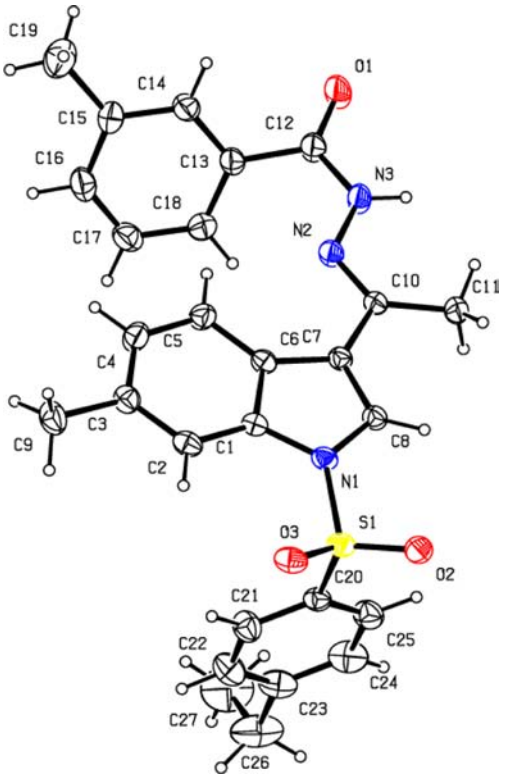


Figure 1. X-ray crystal structure of **V-33**.

This demonstrated that the substituents on the C═N bond of **V-33** adopted a *trans* configuration. If the substituents on the C═N bond of **V-33** adopted a *cis* configuration, big steric effects could be observed between the indolyl ring and the arylcarbonylamino group. Crystallographic data (excluding structure factors) for the structure of **V-33** have been deposited with the Cambridge Crystallographic Data Centre as supplementary publication CCDC 915720. Copies of the data can be obtained, free of charge, on application to CCDC, 12 Union Road, Cambridge CB2 1EZ, UK [fax, +44 (0)1223 336033; e-mail, deposit@ccdc.cam.ac.uk].

Nematicidal Activity. The nematicidal activity of **V-1**−**V-54** against *B. xylophilus* is indicated in Table 1. Some compounds such as **V-6**, **V-12**, **V-19**, **V-20**, **V-27**, **V-38**, **V-39**, and **V-49** showed potent nematicidal activity with LC_{50} values ranging from 1 to 2 mg/L. Especially **V-12** and **V-39** displayed the best promising nematicidal activity, with LC_{50} values of 1.0969 and 1.2632 mg/L, respectively. On the other hand, some interesting results of the structure−activity relationships of **V-1**−**V-54** were also observed: (a) When R^3 = H and R^4 = Ph, introduction of R^1 and R^2 together as the electron-withdrawing substituents could lead to the pronounced compound (e.g., **V-12** vs **V-1**−**V-11** and **V-13**). For example, the LC_{50} of **V-12** (containing R^1 = 5-CN and R^2 = 3-NO_2) was 1.0969 mg/L. It is noteworthy that introduction of R^2 as a two-electron-withdrawing substituent (such as NO_2 and Cl) could result in the more potent compound **V-6** relative to those containing R^2 as a one-electron-withdrawing substituent (e.g., **V-4** and **V-5**). That is, electron deficiency of the indolyl ring and the phenyl ring of the *N*-arylsulfonyl group could favor their nematicidal activity. (b) When R^1 = H and R^3 = Me, introduction of R^2 and R^4 together as electron-donating substituents could generally afford a promising compound (e.g., **V-19** vs **V-14**−**V-18** and **V-20**−**V-26**). For example, **V-19** contains R^2 as 4-Me and R^4 as (3-Me)Ph with the LC_{50} value of 1.5955 mg/L. (c) When R^1 = 6-Me and R^3 = Me (**V-27**−**V-45**), introduction of R^2 as H, 4-OMe, or 4-Cl and R^4 as Ph could produce promising compounds (e.g., **V-27**, **V-38**, and **V-39**). (d) When R^3 = Me and R^4 = Ph (**V-46**−**V-51**), introduction of R^1 as 5-NO_2 and R^2 as H could lead to the pronounced compound (e.g., **V-49**). (e) Interestingly, the proper chain length of R^3 was essential for nematicidal activity. For example, the LC_{50} values of **V-52** (R^3 = Et) and **V-53** (R^3 = *n*-pentyl) were 3.7058 and 6.6758 mg/L, respectively, whereas the LC_{50} value of **V-14** (R^3 = Me) was 2.3985 mg/L. All in all, introduction of R^1 and R^2 together as electron-withdrawing substituents, R^3 as the methyl group, and R^4 as the phenyl with the electron-donating substituents could be taken into account for further preparation of this kind of compounds as nematicidal agents.

QSAR Model. Through the KS method, a training set containing 40 compounds and a test set containing 14 compounds were obtained. To select the molecular parameters that are most relevant to the LC_{50} values of the compounds, 851 structural descriptors calculated by DRAGON 5.4 were used as inputs for GA selection procedure. When adding another variable did not improve the performance of the model significiently, the optimal subset size was believed to obtain. In the current work, the LOO cross-validation was used to evaluate the proposed QSAR models. On the basis of this principle, the six-variable model was selected as the best model. The corresponding regression equation and the statistical parameters were

$$Y = -0.377\mathrm{EEig05x} - 0.421\mathrm{EEig13d} + 0.467\mathrm{BEHm8} + 0.967\mathrm{E1m} + 0.693\mathrm{R1m+} - 0.694\mathrm{R3m+} - 17.276$$

where Y is the LC_{50} values (μmol/L).

$$n_{\text{training set}} = 40\ R^2_{\text{training set}} = 0.791\ \mathrm{RMSE}_{\text{training set}} = 3.279$$

$$Q^2_{\mathrm{LOO}} = 0.701\ \mathrm{RMSE}_{\mathrm{LOO}} = 3.929\ Q^2_{\text{7-fold}} = 0.715$$

$$\mathrm{RMSE}_{\text{7-fold}} = 0.370\ n_{\text{test set}} = 14\ Q^2_{\text{ext}} = 0.774$$

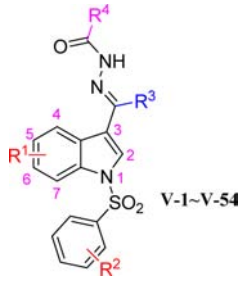
$$\mathrm{RMSE}_{\text{test set}} = 3.412$$

The best six-parameter model gave the correlation coefficient ($R^2_{\text{training set}}$), Q^2_{LOO}, and $Q^2_{\text{7-fold}}$ as 0.791, 0.701, and 0.715, respectively. The prediction ability of a QSAR model is very important, and statistical parameters for the test set were Q^2_{ext} = 0.774 and $\mathrm{RMSE}_{\text{test set}}$ = 3.412, which are satisfactory. From the statistical parameters discussed above, the proposed model is stable, robust, and predictive. The predicted LC_{50} values by the derived model are listed in Table 2, and the regression plot of the developed best model is shown in Figure 2.

From a deep analysis of the descriptors used in the proposed model, we could gain some insight into the factors that would influence the bioactivity of the compounds. The relative importance of the descriptors is weighted by the standardized regression coefficient value of the descriptor. The most important descriptor is E1m (first component accessibility directional

dx.doi.org/10.1021/jf400536q | J. Agric. Food Chem. 2013, 61, 5696−5705

Table 1. Nematicidal Activity of Compounds V-1−V-54 against *B. xylophilus*



V-1~V-54

compd	R^1	R^2	R^3	R^4	LC_{50} (mg/L)
V-1	H	4-Me	H	Ph	7.6852
V-2	H	4-OMe	H	Ph	6.7337
V-3	H	4-NHAc	H	Ph	3.3856
V-4	H	4-Cl	H	Ph	14.8897
V-5	H	3-NO_2	H	Ph	6.4652
V-6	H	3-NO_2, 4-Cl	H	Ph	1.8956
V-7	6-Me	4-Me	H	Ph	3.0984
V-8	6-Me	4-OMe	H	Ph	4.3049
V-9	6-Me	4-Cl	H	Ph	10.0989
V-10	6-Me	3-NO_2	H	Ph	5.4508
V-11	5-CN	4-Me	H	Ph	2.4815
V-12	5-CN	3-NO_2	H	Ph	1.0969
V-13	5-NO_2	4-Me	H	Ph	2.7624
V-14	H	H	Me	Ph	2.3985
V-15	H	H	Me	(*m*-Me)Ph	4.5374
V-16	H	H	Me	(*p*-OMe)Ph	3.7563
V-17	H	H	Me	(*m*-Cl)Ph	15.6279
V-18	H	4-Me	Me	Ph	3.8666
V-19	H	4-Me	Me	(*m*-Me)Ph	1.5955
V-20	H	4-OMe	Me	Ph	1.8840
V-21	H	4-Cl	Me	Ph	2.9612
V-22	H	3-NO_2	Me	Ph	5.0187
V-23	H	3-NO_2	Me	(*m*-Me)Ph	2.9816
V-24	H	3-NO_2	Me	(*p*-OMe)Ph	4.4858
V-25	H	3-NO_2	Me	(*m*-Cl)Ph	4.3052
V-26	H	3-NO_2	Me	(*p*-NO_2)Ph	6.6147
V-27	6-Me	H	Me	Ph	1.4112
V-28	6-Me	H	Me	(*p*-OMe)Ph	2.9201
V-29	6-Me	H	Me	(*p*-NO_2)Ph	8.8139
V-30	6-Me	4-Me	Me	Ph	4.2817
V-31	6-Me	4-Et	Me	2-thienyl	6.0769
V-32	6-Me	4-Et	Me	3-pyridyl	6.2622
V-33	6-Me	4-Et	Me	(*m*-Me)Ph	6.7608
V-34	6-Me	4-Et	Me	(*p*-OMe)Ph	4.8998
V-35	6-Me	4-Et	Me	(*m*-Cl)Ph	5.9591
V-36	6-Me	4-Et	Me	(*p*-NO_2)Ph	10.7303
V-37	6-Me	4-Et	Me	(*p*−OH)Ph	6.5078
V-38	6-Me	4-OMe	Me	Ph	1.8538
V-39	6-Me	4-Cl	Me	Ph	1.2632
V-40	6-Me	3-NO_2	Me	Ph	4.2262
V-41	6-Me	3-NO_2	Me	2-thienyl	6.3488
V-42	6-Me	3-NO_2	Me	3-pyridyl	3.6371
V-43	6-Me	3-NO_2	Me	(*m*-Cl)Ph	4.7920
V-44	6-Me	3-NO_2	Me	(*p*-NO_2)Ph	3.2280
V-45	6-Me	3-NO_2	Me	(*p*-OH)Ph	2.7766
V-46	5-CN	H	Me	Ph	5.2533
V-47	5-CN	4-Me	Me	Ph	2.6172
V-48	5-CN	4-Cl	Me	Ph	5.7650
V-49	5-NO_2	H	Me	Ph	1.8339
V-50	5-NO_2	4-Me	Me	Ph	3.6960
V-51	5-NO_2	4-Cl	Me	Ph	5.5885
V-52	H	H	Et	Ph	3.7058
V-53	H	H	*n*-pentyl	Ph	6.6758
V-54	H	4-Me	*n*-pentyl	Ph	4.7448
emamectin benzoate					0.4102

dx.doi.org/10.1021/jf400536q | *J. Agric. Food Chem.* 2013, 61, 5696−5705

Table 2. Experimental and Predicted Activity (LC_{50}, μmol/L) by the Developed QSAR Model

no.	compd	status	experimental activity	predicted activity
1	**V-1**	training	18.41	15.34
2	**V-2**	training	15.53	19.38
3	**V-3**	training	7.35	13.07
4	**V-4**	training	34.00	28.68
5	**V-5**	training	14.42	14.09
6	**V-6**	training	3.93	4.98
7	**V-7**	test	7.18	7.87
8	**V-8**	test	9.62	10.41
9	**V-9**	training	22.35	24.93
10	**V-10**	test	11.79	8.74
11	**V-11**	training	5.61	3.64
12	**V-12**	training	2.32	2.3
13	**V-13**	training	5.97	5.2
14	**V-14**	training	5.75	6.73
15	**V-15**	test	10.52	12.3
16	**V-16**	training	8.39	8.5
17	**V-17**	training	34.58	27.43
18	**V-18**	test	8.96	9.11
19	**V-19**	training	3.58	8.91
20	**V-20**	test	4.21	12.48
21	**V-21**	test	6.55	10.99
22	**V-22**	test	10.85	7.05
23	**V-23**	test	6.26	4.21
24	**V-24**	training	9.11	7.6
25	**V-25**	training	8.66	12.14
26	**V-26**	training	13.03	13.54
27	**V-27**	test	3.27	8.04
28	**V-28**	training	6.33	6
29	**V-29**	training	18.5	19.42
30	**V-30**	test	9.61	6.4
31	**V-31**	training	13.05	14.64
32	**V-32**	test	13.60	11.52
33	**V-33**	training	14.28	8.37
34	**V-34**	training	10.00	9.15
35	**V-35**	training	12.06	15.81
36	**V-36**	training	21.27	22.06
37	**V-37**	training	13.68	14
38	**V-38**	training	4.02	10.48
39	**V-39**	training	2.71	7.23
40	**V-40**	training	8.87	2.27
41	**V-41**	training	13.16	10.84
42	**V-42**	training	7.62	7.88
43	**V-43**	training	9.38	9.7
44	**V-44**	training	6.19	7.43
45	**V-45**	training	5.64	5.01
46	**V-46**	training	11.87	5.65
47	**V-47**	test	5.73	5.67
48	**V-48**	training	12.09	8.29
49	**V-49**	training	3.97	6.8
50	**V-50**	test	7.76	5.46
51	**V-51**	training	11.25	9.59
52	**V-52**	training	8.59	10.22
53	**V-53**	training	14.1	15.54
54	**V-54**	training	9.73	8.49

WHIM index/weighted by atomic masses), which is a WHIM descriptor. WHIM descriptors are built in such a way to capture relevant molecular 3D information regarding molecular size, shape, symmetry, and atom distribution with respect to

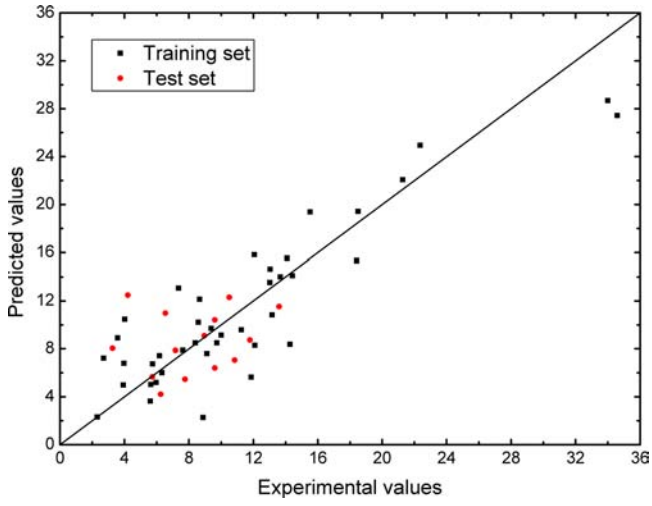


Figure 2. Plot of experimental and predicted biological activity values (LC_{50}, μmol/L) of 54 compounds by GA-MLR model for the training and test sets.

invariant reference frames. R1m+ is a GETAWAY descriptor, which represents R maximal autocorrelation of lag 1/weighted by atomic masses. Another important descriptor is R3m+ (R maximal autocorrelation of lag 3/weighted by atomics masses), which is also a GETAWAY descriptor. GETAWAY (geometry, topology, and atom-weights-assembly) is derived from the leverage matrix, which is deduced by the centering of all atomic coordinates. BEHm8 is a Burden eigenvalues descriptor weighted by atomic masses. The above four descriptors all are weighted by atomic masses, which represents atomic masses having an important correlation with the bioability. EEig05x represents eigenvalue 05 from the edge-adjacency matrix weighted by edge degrees, which belongs to edge-adjacency indices. EEig13d is an edge-adjacency index descriptor similar to EEig05x, and the difference is that EEig13d is weighted by dipole moments. It can be seen from the above discussion that the descriptors that are structural factors are likely to govern the activities of these compounds, including molecular masses (a negative correlation with the bioactivity) and molecular polarity (a positive correlation with bioactivity).

The QSAR model should be verified by chemical domain applicability. In this work, as shown in Figure 3, the applicability domain (AD) of the model and the reliability of the prediction were evaluated by the leverage approach expressed as Williams plot. It is obvious that only one compound (**V-6**) in the training set has a hat value higher than the warning h^* value of 0.525 and, thus, is regarded as a structural outlier. The predicted value of this compound would be more reliable when regarded as an extrapolation of the model, but the compound has a small residual, so it is a "good leverage" compound. There is no response outlier for either the training set or the test set.

In conclusion, 54 novel *N*-arylsulfonyl-3-acylindole arylcarbonyl hydrazone derivatives (**V-1**−**V-54**) were prepared and tested for their nematicidal activity against *B. xylophilus* in vivo. Among all of the compounds, **V-12** and **V-39** displayed the best promising nematicidal activity with LC_{50} values of 1.0969 and 1.2632 mg/L, respectively. It is generally suggested that introduction of R^1 and R^2 together as the electron-withdrawing substituents, R^3 as the methyl group, and R^4 as the phenyl with the electron-donating substituents can be considered for future

dx.doi.org/10.1021/jf400536q | J. Agric. Food Chem. 2013, 61, 5696−5705

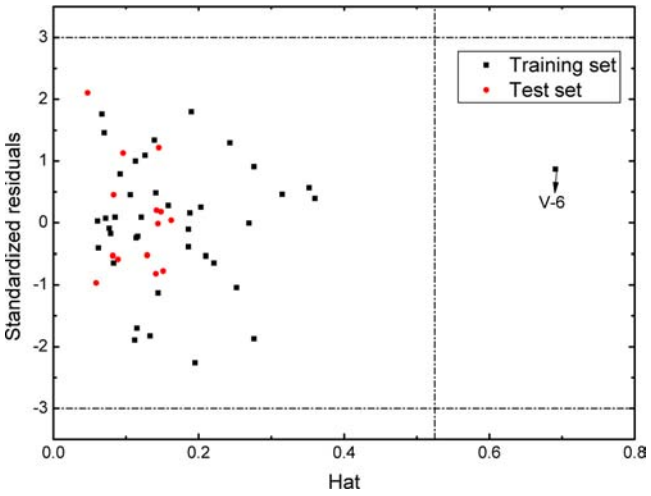


Figure 3. Williams plot for the training and test sets by GA-MLR model with six descriptors. The dotted lines are the 3σ limit and the warning value of hat (h^* = 0.525).

preparation of these kinds of compounds as nematicidal agents. The QSAR model demonstrated that the structural factors, such as molecular masses (a negative correlation with the bioactivity) and molecular polarity (a positive correlation with bioactivity), are likely to govern the nematicidal activities of these compounds.

■ ASSOCIATED CONTENT

Supporting Information

^{1}H NMR, HRMS, MS, mp, or ^{13}C NMR data for compounds **III-3**−**III-17**, **IV-3**−**IV-13**, **IV-16**−**IV-33**, and **V-21**−**V-54**. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +86-029-87091952. Fax: +86-029-87091952. E-mail: orgxuhui@nwsuaf.edu.cn.

Funding

This research was supported by the National Natural Science Foundation of China (No. 31171896), and the Special Funds of Central Colleges Basic Scientific Research Operating Expenses (YQ2013008).

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Park, I. K.; Park, J. Y.; Kim, K. H.; Choi, K. S.; Choi, I. H.; Kim, C. S.; Shin, S. C. Nematicidal activity of plant essential oils and components from garlic (*Allium sativum*) and cinnamon (*Cinnamomum verum*) oils against the pine wood nematode (*Bursaphelenchus xylophilus*). *Nematology* **2005**, *7*, 767−774.

(2) Qui, S. J.; Can, J. Y.; Liu, W. P.; Becker, J. O. Degradation and absorption of fosthiazate in soil. *J. Agric. Food Chem.* **2004**, *52*, 6239−6242.

(3) Meher, H. C.; Gajbhiye, V. T.; Chawla, G.; Singh, G. Virulence development and genetic polymorphism in *Meloidogyne incognita* (Kofoid & White) Chitwood after prolonged exposure to sublethal concentrations of nematicides and continuous growing of resistant tomato cultivars. *Pest Manag. Sci.* **2009**, *65*, 1201−1207.

(4) Ntalli, N. G.; Caboni, P. Botanical nematicides: a review. *J. Agric. Food Chem.* **2012**, *60*, 9929−9940.

(5) Seo, S.-M.; Kim, J.; Kim, E.; Park, H.-M.; Kim, Y.-J.; Park, I.-K. Structure−activity relationship of aliphatic compounds for nematicidal activity against pine wood nematode (*Bursaphelenchus xylophilus*). *J. Agric. Food Chem.* **2010**, *58*, 1823−1827.

(6) Faizi, S.; Fayyaz, S.; Bano, S.; Iqbal, E. Y.; Lubna; Siddiqi, H.; Naz, A. Isolation of nematicidal compounds from *Tagetes patula* L. yellow flowers: structure−activity relationship studies against cyst nematode *Heterodera zeae* infective stage larvae. *J. Agric. Food Chem.* **2011**, *59*, 9080−9093.

(7) Caboni, P.; Ntalli, N. G.; Aissani, N.; Cavoski, I.; Angioni, A. Nematicidal activity of (*E*,*E*)-2,4-decadienal and (*E*)-2-decenal from *Ailanthus altissima* against *meloidogyne javanica*. *J. Agric. Food Chem.* **2012**, *60*, 1146−1151.

(8) Caboni, P.; Sarais, G.; Aissani, N.; Tocco, G.; Sasanelli, N.; Liori, B.; Carta, A.; Angioni, A. Nematicidal activity of 2-thiophenecarboxaldehyde and methylisothiocyanate from caper (*Capparis spinosa*) against Meloidogyne incognita. *J. Agric. Food Chem.* **2012**, *60*, 7345−7351.

(9) Ntalli, N. G.; Ferrari, F.; Giannakou, I. O.; Menkissoglu-Spiroudi, U. Synergistic and antagonistic interactions of terpenes against *Meloidogyne incognita* and nematicidal activity of essential oils from 7 plants indigenous in Greece. *Pest Manag. Sci.* **2011**, *67*, 341−351.

(10) Zasada, I. A.; Meyer, S. L. F.; Morra, M. J. Brassicaceous seed meals as soil amendments to suppress the plant-parasitic nematodes *Pratylenchus penetrans* and *Meloidogyne incognita*. *J. Nematol.* **2009**, *41*, 221−227.

(11) Campbell, B. E.; Tarleton, M.; Gordon, C. P.; Sakoff, J. A.; Gilbert, J.; McCluskey, A.; Gasser, R. B. Norcantharidin analogues with nematocidal activity in *Haemonchus contortus*. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 3277−3281.

(12) Isman, M. B. Botanical insecticides, deterrents, and repellents in modern agriculture and an increasingly regulated world. *Annu. Rev. Entomol.* **2006**, *51*, 45−66.

(13) Chitwood, D. J. Phytochemical based strategies for nematode control. *Annu. Rev. Phytopathol.* **2002**, *40*, 221−249.

(14) Chen, J. T.; Su, H. J.; Huang, J. W. Isolation and identification of secondary metabolites of *Clitocybe nuda* responsible for inhibition of zoospore germination of *Phytophthora capsici*. *J. Agric. Food Chem.* **2012**, *60*, 7341−7344.

(15) Zhou, Y. Y.; Luo, S. H.; Yi, T. S.; Li, C. H.; Luo, Q.; Hua, J.; Liu, Y.; Li, S. H. Secondary metabolites from *Glycine soja* and their growth inhibitory effect against *Spodoptera litura*. *J. Agric. Food Chem.* **2011**, *59*, 6004−6010.

(16) Isman, M. B. Botanical insecticides: for richer, for poorer. *Pest Manag. Sci.* **2008**, *64*, 8−11.

(17) Regnault-Roger, C.; Philogène, B. J. R. Past and current prospects for the use of botanicals and plant allelochemicals in integrated pest management. *Pharm. Biol.* **2008**, *46*, 41−52.

(18) Dayan, F. E.; Cantrell, C. L.; Duke, S. O. Natural products in crop protection. *Bioorg. Med. Chem.* **2009**, *17*, 4022−4034.

(19) Xu, H.; Lv, M. Developments of indoles as anti-HIV-1 inhibitors. *Curr. Pharm. Des.* **2009**, *15*, 2120−2148.

(20) Silvestri, R.; Artico, M. Indolyl aryl sulfones (IASs): development of highly potent NNRTIs active against wt-HIV-1 and clinically relevant drug resistant mutants. *Curr. Pharm. Des.* **2005**, *11*, 3779−3806.

(21) Chen, K. X.; Vibulbhan, B.; Yang, W. Y.; Sannigrahi, M.; Velazquez, F.; Chan, T. Y.; Venkatraman, S.; Anilkumar, G. N.; Zeng, Q. B.; Bennet, F.; Jiang, Y. H.; Lesburg, C. A.; Duca, J.; Pinto, P.; Gayalas, S.; Huang, Y. H.; Wu, W. L.; Selyutin, O.; Agrawal, S.; Feld, B.; Huang, H. C.; Li, C.; Cheng, K. C.; Shih, N. Y.; Kozlowski, J. A.; Rosenblum, S. B.; Njoroge, F. G. Structure-activity relationship (SAR) development and discovery of potent indole-based inhibitors of the hepatitis C virus (HCV) NS5B polymerase. *J. Med. Chem.* **2012**, *55*, 754−765.

(22) Singh, P.; Verma, P.; Yadav, B.; Komath, S. S. Synthesis and evaluation of indole-based new scaffolds for antimicrobial activities-identification of promising candidates. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 3367−3372.

(23) Grella, B.; Adams, J.; Berry, J. F.; Delahanty, G.; Ferraris, D. V.; Majer, P.; Ni, C. Y.; Shukla, K.; Shuler, S. A.; Slusher, B. S.; Stathis, M.; Tsukamoto, T. The discovery and structure-activity relationships of indole-based inhibitors of glutamate carboxypeptidase II. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 7222−7225.
(24) Xu, H.; Fan, L. Synthesis and antifungal activities of novel 5,6-dihydro-indolo[1,2-*a*]quinoxaline derivatives. *Eur. J. Med. Chem.* **2011**, *46*, 1919−1925.
(25) Guo, Y.; Yan, Y.; Yang, C.; Yu, X.; Zhi, X.; Xu, H. Regioselective synthesis of fraxinellone-based hydrazone derivatives as insecticidal agents. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 5384−5387.
(26) Ran, J. Q.; Huang, N.; Xu, H.; Yang, L. M.; Lv, M.; Zheng, Y. T. Anti HIV-1 agents 5: synthesis and anti-HIV-1 activity of some *N*-arylsulfonyl-3-acetylindoles *in vitro*. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 3534−3536.
(27) Guo, Y.; Yan, Y.; Yu, X.; Wang, Y.; Xu, H. Synthesis and insecticidal activity of some novel fraxinellone-based esters. *J. Agric. Food Chem.* **2012**, *60*, 7016−7021.
(28) Wang, Y.; Shao, Y.; Wang, Y.; Fan, L.; Yu, X.; Zhi, X.; Yang, C.; Qu, H.; Yao, X.; Xu, H. Synthesis and quantitative structure-activity relationship (QSAR) study of novel isoxazoline and oxime derivatives of podophyllotoxin as insecticidal agents. *J. Agric. Food Chem.* **2012**, *60*, 8435−8443.
(29) Ge, Y. H.; Wu, Y. M.; Xue, Z. J. Synthesis of substituted indole-3-carboxaldehyde derivatives. *Chin. J. Org. Chem.* **2006**, *26*, 563−567.
(30) Nishimura, T.; Yamada, T.; Takebe, T.; Yokoshima, S.; Fukuyama, T. (1-Nosyl-5-nitroindol-3-yl)methyl ester: a novel protective group for carboxylic acids. *Org. Lett.* **2008**, *10*, 2601−2604.
(31) Arisawa, M.; Terada, Y.; Takahashi, K.; Nakayawa, M.; Nishida, A. Development of isomerization and cycloisomerization with use of a ruthenium hydride with *N*-heterocyclic carbene and its application to the synthesis of heterocycles. *J. Org. Chem.* **2006**, *71*, 4255−4261.
(32) Abid, M.; Teixeira, L.; Torok, B. Triflic acid controlled successive annelation of aromatic sulfonamides: an efficient one-pot synthesis of *N*-sulfonyl pyrroles, indoles and carbazoles. *Tetrahedron Lett.* **2007**, *48*, 4047−4050.
(33) Ranjana, A.; Fabio, B.; Federico, B.; Sabrina, B.; Alfonso, C.; Francesca, D.; Vinicio, G.; Stefano, N. A catalytic antibody programmed for torsional activation of amide bond hydrolysis. *Chem.−Eur. J.* **2003**, *9*, 3132−3142.
(34) Ketcha, D. M.; Gribble, G. W. A convenient synthesis of 3-acylindoles via Friedel Crafts acylation of 1-(phenylsulfonyl)indole. A new route to pyridocarbazole-5,11-quinones and ellipticine. *J. Org. Chem.* **1985**, *50*, 5451−5457.
(35) Kurihara, T.; Fujimoto, T.; Harusawa, S.; Yoneda, R. Simple *N*-alkylation and *N*-acylation of 3-acetylindole and 3-indolecarbaldehyde. *Synthesis* **1987**, 396−397.
(36) Kohno, T.; Togashi, K.; Fukamiya, N. The nematocidal activity and the structure−activity relationships of stilbenes. *Nat. Prod. Res.* **2007**, *21*, 606−615.
(37) Hypercube, Inc. HyperChem 7.0; Gainesville, FL, 2002.
(38) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. Development and use of quantum-mechanical molecular models. 76. AM1 − a new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(39) Talete srl. DRAGON for windows (software for molecular descriptor calculations), version 5.4, 2006; http://www.talete.mi.it.
(40) Todeschini, R.; Gramatica, P. New 3D molecular descriptors: the WHIM theory and QSAR applications. *Perspect. Drug Discov. Des.* **1998**, *9*−*11*, 355−380.
(41) Consonni, V.; Todeschini, R.; Pavan, M. Structure/response correlations and similarity/diversity analysis by GETAWAY descriptors. 1. Theory of the novel 3D molecular descriptors. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 682−692.
(42) Todeschini, R.; Consonni, V. *Handbook of Molecular Descriptors: Methods and Principles in Medicinal Chemistry*; Wiley-VCH: Weinheim, Germany, 2000.
(43) Kennard, R. W.; Stone, L. A. Computer aided design of experiments. *Technometrics* **1969**, *11*, 121−134.
(44) Leardi, R.; Boggia, R.; Terrile, M. Genetic algorithms as a strategy for feature-selection. *J. Chemom.* **1992**, *6*, 267−281.
(45) Todeschini, R.; Consonni, V.; Pavan, M. MOBY DIGS, version 1.2 for Windows, Software for Multilinear Regression Analysis and Variable Subset Selection by Genetic Algorithm; Talete Srl, Milan, Italy, 2002.
(46) Tropsha, A.; Gramatica, P.; Gombar, V. K. The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models. *QSAR Comb. Sci.* **2003**, *22*, 69−77.