

## Original article

## 3D-QSAR studies of Checkpoint Kinase Weel inhibitors based on molecular docking, CoMFA and CoMSIA

Ping Yi <sup>a,b</sup>, Xin Fang <sup>a,b</sup>, Minghua Qiu <sup>a,\*</sup><sup>a</sup> State Key Laboratory of Phytochemistry and Plant Resources in West China, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650204, China<sup>b</sup> Graduate School of the Chinese Academy of Sciences, Beijing 100039, China

Received 6 February 2007; received in revised form 18 June 2007; accepted 21 June 2007

Available online 14 August 2007

## Abstract

Three-dimensional quantitative structure-activity relationship (3D-QSAR) studies were performed on 97 4-phenylpyrrolo[3,4-c]carbazole-1,3(2*H*,6*H*)-dione inhibitors, based on molecular docking scores obtained by using GOLD 3.1, comparative molecular field analysis (CoMFA) and comparative molecular similarity indices (CoMSIA). The docking results provided a reliable conformational alignment scheme for the 3D-QSAR model. Based on the docking conformations and alignments, highly predictive CoMFA and CoMSIA were obtained with cross-validated  $q^2$  value of 0.828 and 0.796, respectively, and non-cross-validated partial least-squares (PLS) analysis with the optimum components of five showed a conventional  $r^2$  of 0.962 and 0.949, respectively. The predictive ability was validated by compounds that were not included in the training set. Furthermore, the CoMFA and CoMSIA model plots were mapped back to the binding sites of Checkpoint Kinase Weel, to get a better understanding of vital interactions between the inhibitors and Weel kinase. As a result, we have identified some key features in the 4-phenylpyrrolo[3,4-c]carbazole-1,3(2*H*,6*H*)-diones responsible for the Weel inhibitory activity that may be used to design more potent 4-phenylpyrrolo[3,4-c]carbazole-1,3(2*H*,6*H*)-diones and predict their activity prior to synthesis.

© 2007 Elsevier Masson SAS. All rights reserved.

**Keywords:** 3D-QSAR; GOLD; CoMFA; CoMSIA; Checkpoint Kinase Weel

## 1. Introduction

Of the diverse ways of treating cancer, the combination of a Weel inhibitor with a conventional DNA-damaging cytotoxic agent could be a useful one. It is well known that the lack of a functional p53 signaling pathway in many cancer cells leads to the malfunction of the important DNA damage-sensitive checkpoint, G1/S. Therefore, inhibitors of the protein tyrosine kinase enzyme Weel, which play an important role in regulation of the DNA damage-sensitive G2/M checkpoint in the eukaryotic cell cycle by the inhibitory phosphorylation of cdc2 on tyrosine, should abrogate the G2/M checkpoint and preferentially enhance the cytotoxic effects of DNA-damaging agents on p53-negative cells. It is well known that by allowing

them to bypass the checkpoint where damaged cells normally arrest to allow time for DNA repair, the inhibitors succeed in achieve that goal [1–6].

In recent years, a number of Weel inhibitors have appeared in the patent and primary literatures [7–9]. Among them, Palmer et al. found that the 4-phenylpyrrolocarbazole **1** was a potent inhibitor of Weel ( $IC_{50} = 0.097 \mu M$ ) in a high-throughput screening program. Moreover, they synthesized a series of 4-phenylpyrrolocarbazole derivatives as potent, selective Weel inhibitors [10]. The binding modes of this series of Weel inhibitors have been determined by X-ray crystallography, which provided not only insights into the interaction mechanisms of Weel with the inhibitors, but also valuable clues for designing new inhibitors [11–13]. In this study, by using molecular docking and 3D-QSAR analyses it is possible to get new insights into the relationship between the structural information of the series of 97 4-phenylpyrrolocarbazole

\* Corresponding author. Tel./fax: +86 871 5223255.

E-mail address: [mhchiu@mail.kib.ac.cn](mailto:mhchiu@mail.kib.ac.cn) (M. Qiu).

derivatives and their inhibitory potency, with the aim of identifying structural features in Weel that can be used to find new inhibitors.

## 2. Computational details

### 2.1. Biological data and molecular structures

A series of Weel 4-phenylpyrrolo[3,4-*c*]carbazole-1,3(2*H*,6*H*)-dione inhibitors, published by Palmer et al. [10], were divided into a training set and a test set as shown in Tables 1 and 2. The training set consists of 73 compounds and the test set, selected randomly, is comprised of 14 compounds. The Weel IC<sub>50</sub>s were converted to pIC<sub>50</sub> (−log IC<sub>50</sub>) and used as dependent variables in the CoMFA and CoMSIA calculations.

The 3D structures of these compounds were constructed using the molecular modeling software package Sybyl6.9 [14]. Partial atomic charges were calculated by the Gasteiger–Huckel method and energy minimizations were performed using the Tripos force field with a distance-dependent dielectric and the Powell conjugate gradient algorithm (convergence criterion of 0.001 kcal/mol/Å) [15]. Atomic coordinates for the Weel complex with ligand **824** (phenylpyrrolo[3,4-*c*]carbazole-1,3(2*H*,6*H*)-dione) that was used for our modeling study have been deposited in the Brookhaven Protein DataBank with a resolution of 1.81 Å (PDB ID:1X8B) [13].

### 2.2. Docking studies

To locate the appropriate binding orientations and conformations of the 4-phenylpyrrolo[3,4-*c*]carbazole-1,3(2*H*,6*H*)-dione inhibitors interacting with Checkpoint Kinase Weel, a powerful computational searching method was needed. The advanced molecular docking program GOLD (version 3.1), with a powerful genetic algorithm (GA) method for conformational search and docking programs [16], was employed to generate an ensemble of docked conformations. The structures of Weel and the inhibitors were built using the Sybyl6.9 molecular modeling software. The original ligand was removed from the coordinated set of Weel, and 9 ordered water molecules and a coordinated magnesium atom within 6 Å of the active pocket were taken into account during the docking procedures. The genetic operators were 100 for the population size, 1.1 for the selection, five for the number of subpopulations, 100,000 for the maximum number of genetic applications, and two for the size of the niche used to increase population diversity. The weights were chosen so that cross-over mutations were applied with equal probability (95/95 for the values) and migration was applied 5% of the time.

### 2.3. Choice of docking fitness functions

The ChemScoring function described by Eldridge et al. [17,18] and encoded in GOLD was applied to predict binding positions between the Weel and 97 inhibitors. ChemScore

estimates the total free energy change that occurs on ligand binding as:

$$\Delta G_{\text{binding}} = \Delta G_0 + \Delta G_{\text{hbond}} + \Delta G_{\text{metal}} + \Delta G_{\text{lipo}} + \Delta G_{\text{rot}}$$

The final ChemScore value is obtained by adding in a clash penalty and internal torsion terms, which militate against close contacts in docking and poor internal conformations. Covalent and constraint scores may also be included. *P* terms represent the various types of physical contributions to binding.

$$\text{ChemScore} = \Delta G_{\text{binding}} + P_{\text{clash}} + C_{\text{internal}}P_{\text{internal}} + (c_{\text{covalent}}P_{\text{covalent}} + P_{\text{constraint}})$$

The fitness score is taken as the negative of the sum of the component energy terms, so that larger fitness scores are better [19]. However, there is not a statistically significant relationship between the GOLD ChemScore and biological activity for these inhibitors. Since the GOLD fitness function was designed to discriminate between different binding modes of the same molecule, maybe extra terms such as entropic loss are required to compare different molecules.

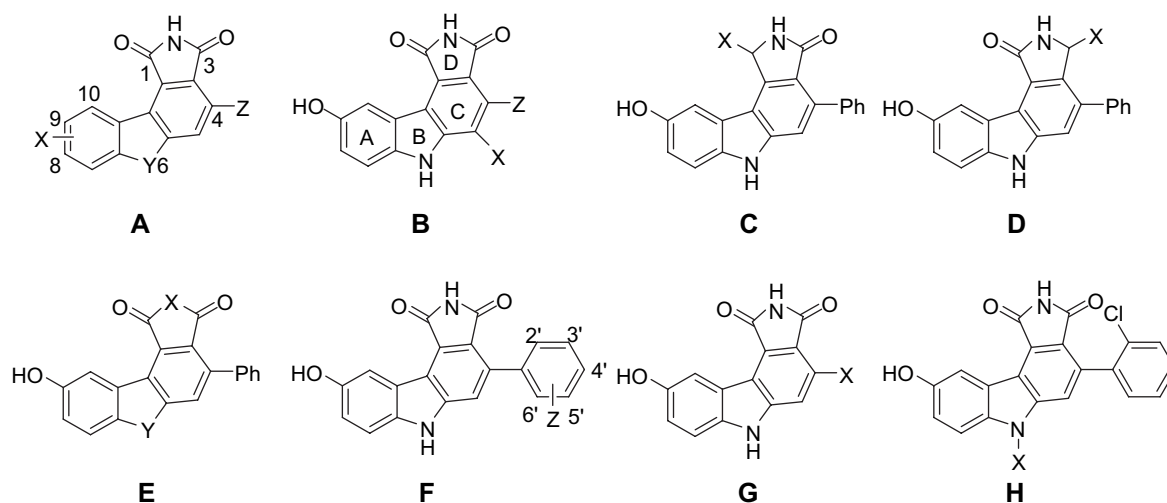
### 2.4. Structural alignment

Ten conformations were obtained using GOLD for each ligand, in which the training set contained the conformations with the highest ChemScores while the test set contained the conformations with the lowest residues between actual and predicted pIC<sub>50</sub> predicted by the CoMFA model of the training set. The conformations and their alignment, and the relative binding positions of the conformations in Weel, all obtained from GOLD, were used directly in CoMFA and CoMSIA to explore the specific contributions of electrostatic, steric, hydrophobic and hydrogen bonding effects to the molecular bioactivities. The result was a CoMFA model with a cross-validated *q*<sup>2</sup> value of 0.828 for five components. To make a comparison, another common alignment strategy for CoMFA and CoMSIA was tried, in which the most active compound **23** was used as a template for superimposition. After superimposing the 1, 2 and 3-position of the inhibitors, the result was a second CoMFA model with a cross-validated *q*<sup>2</sup> value of 0.577 for 6 components, which was worse than that of the first model. In addition, the alignment of the first CoMFA model represents the binding positions of the docked compounds in Weel. Therefore, we adopted the first alignment strategy.

### 2.5. CoMFA

Steric and electrostatic interactions were calculated using the Tripos force field with a distance-dependent dielectric constant at all intersections in a regular space (2 Å) grid taking a sp<sup>3</sup> carbon atom as steric probe and a +1 charge as electrostatic probe. The cutoff was set to 30 kcal/mol. With standard options for scaling of variables, the regression analysis was carried out using the full cross-validated partial least-squares (PLS) method of LOO (leave-one-out). The minimum-sigma (column filtering) was set to 2.0 kcal/mol to improve the

Table 1  
Inhibitory activity of miscellaneous pyrrolo[3,4-c]carbazole-1,3(2*H*,6*H*)-dione derivatives



| Compound | Structure | X                 | Y   | Z                                  | IC <sub>50</sub> (μM) | pIC <sub>50</sub> |
|----------|-----------|-------------------|-----|------------------------------------|-----------------------|-------------------|
| 1        | A         | 9-OH              | NH  | Ph                                 | 0.097                 | 7.013             |
| 2        | A         | 9-OH              | NH  | H                                  | 4                     | 5.398             |
| 3        | A         | 9-OH              | NH  | I                                  | 2.3                   | 5.638             |
| 4        | A         | 8-OH              | NH  | Ph                                 | 0.31                  | 6.509             |
| 5        | A         | 9-OH              | O   | Ph                                 | 0.43                  | 6.367             |
| 6        | A         | 9-OH              | S   | Ph                                 | 0.078                 | 7.108             |
| 7        | A         | 9-OH              | NMe | Ph                                 | 0.26                  | 6.585             |
| 8        | B         | Me                |     | Ph                                 | 0.13                  | 6.886             |
| 9        | B         | Et                |     | Ph                                 | 1.6                   | 5.796             |
| 10       | B         | Ph                |     | Me                                 | 9.7                   | 5.013             |
| 11       | B         | Ph                |     | Ph                                 | 2.3                   | 5.638             |
| 12       | B         | Ph                |     | H                                  | 4                     | 5.398             |
| 13       | C         | OMe               |     |                                    | 20                    | 4.699             |
| 14       | C         | H                 |     |                                    | 37                    | 4.432             |
| 15       | D         | OH                |     |                                    | 2.8                   | 5.553             |
| 16       | E         | N–NH <sub>2</sub> | NH  |                                    | 3.9                   | 5.409             |
| 17       | A         | 9-OH              | NH  | 2'-ClPh                            | 0.011                 | 7.959             |
| 18       | A         | 9-OMe             | NH  | 2'-ClPh                            | 0.64                  | 6.194             |
| 19       | A         | 9-OH              | NMe | 2'-ClPh                            | 0.057                 | 7.244             |
| 20       | A         | 9-OH              | O   | 2'-ClPh                            | 0.033                 | 7.481             |
| 21       | F         |                   |     | 2'-F                               | 0.33                  | 6.481             |
| 22       | F         |                   |     | 2'-Br                              | 0.023                 | 7.638             |
| 23       | F         |                   |     | 2'-I                               | 0.013                 | 7.886             |
| 24       | F         |                   |     | 2'-Me                              | 0.15                  | 6.824             |
| 25       | F         |                   |     | 2'-Et                              | 0.51                  | 6.292             |
| 26       | F         |                   |     | 2'-CF <sub>3</sub>                 | 0.58                  | 6.237             |
| 27       | F         |                   |     | 2'-CH <sub>2</sub> OH              | 0.45                  | 6.347             |
| 28       | F         |                   |     | 2'-CN                              | 0.19                  | 6.721             |
| 29       | F         |                   |     | 2'-COMe                            | 0.83                  | 6.081             |
| 30       | F         |                   |     | 2'-CONH <sub>2</sub>               | 0.16                  | 6.796             |
| 31       | F         |                   |     | 2'-Ph                              | 0.57                  | 6.244             |
| 32       | F         |                   |     | 2'-OH                              | 0.06                  | 7.222             |
| 33       | F         |                   |     | 2'-OMe                             | 0.024                 | 7.620             |
| 34       | F         |                   |     | 2'-OEt                             | 0.26                  | 6.585             |
| 35       | F         |                   |     | 2'-SMe                             | 0.033                 | 7.481             |
| 36       | F         |                   |     | 2'-SOMe                            | 0.22                  | 6.658             |
| 37       | F         |                   |     | 2'-NO <sub>2</sub>                 | 0.047                 | 7.328             |
| 38       | F         |                   |     | 2'-NH <sub>2</sub>                 | 0.21                  | 6.678             |
| 39       | F         |                   |     | 3'-F                               | 0.22                  | 6.658             |
| 40       | F         |                   |     | 3'-Cl                              | 0.055                 | 7.260             |
| 41       | F         |                   |     | 3'-Me                              | 0.23                  | 6.638             |
| 42       | F         |                   |     | 3'-CH <sub>2</sub> OH              | 0.87                  | 6.060             |
| 43       | F         |                   |     | 3'-CH <sub>2</sub> NH <sub>2</sub> | 4.4                   | 5.357             |

(continued on next page)

Table 1 (continued)









| Compound | Structure | X   | Y | Z                          | IC <sub>50</sub> (μM) | pIC <sub>50</sub> |
|----------|-----------|---|---|----------------------------|-----------------------|-------------------|
| 44       | F         |   |   | 3'-CN                      | 0.18                  | 6.745             |
| 45       | F         |   |   | 3'-COMe                    | 4.3                   | 5.367             |
| 46       | F         |   |   | 3'-Ph                      | 40                    | 4.398             |
| 47       | F         |   |   | 3'-OH                      | 0.089                 | 7.051             |
| 48       | F         |   |   | 3'-OMe                     | 0.62                  | 6.208             |
| 49       | F         |   |   | 3'-NO <sub>2</sub>         | 0.3                   | 6.523             |
| 50       | F         |   |   | 3'-NH <sub>3</sub>         | 0.07                  | 7.155             |
| 51       | F         |   |   | 4'-F                       | 16                    | 4.796             |
| 52       | F         |   |   | 4'-Cl                      | 0.73                  | 6.137             |
| 53       | F         |   |   | 4'-Me                      | 3.3                   | 5.481             |
| 54       | F         |   |   | 4'-CH <sub>2</sub> OH      | 1.2                   | 5.921             |
| 55       | F         |   |   | 4'-CN                      | 1.8                   | 5.745             |
| 56       | F         |   |   | 4'-COMe                    | 3.6                   | 5.444             |
| 57       | F         |   |   | 4'-OH                      | 0.067                 | 7.174             |
| 58       | F         |   |   | 4'-OMe                     | 12                    | 4.921             |
| 59       | F         |   |   | 4'-SMe                     | 29                    | 4.538             |
| 60       | F         |   |   | 4'-SO <sub>2</sub> Me      | 1.1                   | 5.959             |
| 61       | F         |   |   | 4'-NH <sub>2</sub>         | 0.15                  | 6.824             |
| 62       | F         |   |   | 2'-Cl, 3'-Cl               | 0.028                 | 7.553             |
| 63       | F         |   |   | 2'-Cl, 3'-OH               | 0.012                 | 7.921             |
| 64       | F         |   |   | 2'-Cl, 3'-NH <sub>2</sub>  | 0.024                 | 7.678             |
| 65       | F         |   |   | 2'-Cl, 5'-Cl               | 0.023                 | 7.638             |
| 66       | F         |   |   | 2'-Cl, 4'-NH <sub>2</sub>  | 0.024                 | 7.620             |
| 67       | F         |   |   | 2'-Cl, 5'-Cl               | 0.49                  | 6.310             |
| 68       | F         |   |   | 2'-Cl, 5'-OH               | 0.042                 | 7.377             |
| 69       | F         |   |   | 2'-Cl, 5'-NH <sub>2</sub>  | 0.02                  | 7.699             |
| 70       | F         |   |   | 2'-Cl, 6'-Cl               | 0.028                 | 7.553             |
| 71       | F         |   |   | 2'-Cl, 6'-OH               | 0.045                 | 7.347             |
| 72       | F         |   |   | 2'-Cl, 6'-OMe              | 0.015                 | 7.824             |
| 73       | F         |   |   | 2'-Br, 4'-NH <sub>2</sub>  | 0.02                  | 7.699             |
| 74       | F         |   |   | 2'-Br, 6'-Br               | 0.035                 | 7.456             |
| 75       | F         |   |   | 2'-Me, 3'-Me               | 0.27                  | 6.569             |
| 76       | F         |   |   | 2'-Me, 5'-Me               | 0.96                  | 6.018             |
| 77       | F         |   |   | 2'-Me, 6'-Me               | 0.075                 | 7.125             |
| 78       | F         |   |   | 2'-OMe, 4'-NH <sub>2</sub> | 0.019                 | 7.721             |
| 79       | F         |   |   | 2'-OMe, 5'-NH <sub>2</sub> | 0.11                  | 6.959             |
| 80       | F         |   |   | 2'-OMe, 6'-OMe             | 0.027                 | 7.569             |
| 81       | F         |   |   | 2'-OMe, 6'-F               | 0.029                 | 7.538             |
| 82       | F         |   |   | 2',6'-diCl, 3'-OH          | 0.018                 | 7.745             |
| 83       | F         |   |   | 2',6'-diCl, 4'-OH          | 0.049                 | 7.310             |
| 84       | G         |  |   |                            | 0.14                  | 6.854             |
| 85       | G         |  |   |                            | 0.042                 | 7.377             |
| 86       | G         |  |   |                            | 0.18                  | 6.745             |
| 87       | G         |  |   |                            | 0.038                 | 7.420             |
| 88       | G         |  |   |                            | 13                    | 5.886             |

Table 1 (continued)

| Compound | Structure | X   | Y | Z | IC <sub>50</sub> (μM) | pIC <sub>50</sub> |
|----------|-----------|---|---|---|-----------------------|-------------------|
| 89       | G         |  |   |   | 0.82                  | 6.086             |
| 90       | G         |  |   |   | 0.58                  | 6.237             |
| 91       | G         |  |   |   | 10                    | 5.000             |
| 92       | H         | Et  |   |   | 0.057                 | 7.301             |
| 93       | H         | <i>n</i> -Pr  |   |   | 0.05                  | 7.201             |
| 94       | H         | <i>i</i> -Pr  |   |   | 0.063                 | 7.276             |
| 95       | H         | <i>n</i> -Bu  |   |   | 0.059                 | 7.229             |
| 96       | H         | (CH <sub>2</sub> ) <sub>2</sub> <i>i</i> -Pr                                      |   |   | 0.15                  | 6.824             |
| 97       | H         | <i>n</i> -pent  |   |   | 0.17                  | 6.770             |

signal-to-noise ratio by omitting those lattice points whose energy variation was below this threshold. The final model, a non-cross-validated conventional analysis, was developed with the optimum number of components to yield a non-cross-validated  $r^2$  value.

## 2.6. CoMSIA

The steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor potential fields were calculated at each lattice intersection of a regularly spaced grid of 2.0 Å. A probe atom with radius 1.0 Å and +1.0 charge with hydrophobicity of +1.0 and hydrogen bond donor and hydrogen bond acceptor properties of +1.0 was used to calculate steric, electrostatic, hydrophobic, donor and acceptor fields. Gaussian type distance dependence and the default value of the attenuation factor ( $\alpha = 0.3$ ) were used.

## 2.7. Partial least-squares (PLS) analysis

PLS method was used to linearly correlate the CoMFA and CoMSIA fields to the inhibitory activity values. The cross-validation analysis was performed using the leave-one-out (LOO) method in which one compound is removed from the data set and its activity is predicted using the model derived from the rest of the dataset [20,21]. The cross-validated values of  $r^2$  that resulted in optimum number of components and lowest standard error of prediction were considered for further analysis. Equal weights were assigned to steric and electrostatic fields using the COMFA\_STD scaling option. To speed up the analysis and reduce noise, a minimum filter value  $\sigma$  of 2.0 kcal/mol was used. In the final analysis, conventional  $r^2$  was calculated using the optimum number of components. To further assess the robustness and statistical confidence of the derived models, bootstrapping analysis for 100 runs was performed.

All cross-validated results were analyzed by considering the fact that a value of cross-validated  $r^2$  above 0.3 indicates that the probability of chance correlation is less than 5% [22].

## 3. Results and discussion

### 3.1. Binding conformations of 4-phenylpyrrolo[3,4-*c*]carbazole-1,3(2*H*,6*H*)-diones

In order to determine the probable binding conformations of these 4-phenylpyrrolo[3,4-*c*]carbazole-1,3(2*H*,6*H*)-diones, GOLD was used to dock all compounds into the active sites of Weel. The docking reliability was validated using the known X-ray structure of Weel in complex with the molecular ligand **824** (phenylpyrrolo[3,4-*c*]carbazole-1,3(2*H*,6*H*)-dione). The ligand **824** was redocked to the binding sites of Weel and the docked conformation corresponding to the highest ChemScore was selected as the most probable binding conformation. The root-mean-square deviation (RMSD) between the conformations of cocrystallized **824** and redocked **824** was equal to 0.264 Å, suggesting that a high docking reliability of GOLD in reproducing the experimentally observed binding mode for Weel inhibitors and the parameter set for the GOLD simulation was reasonable to reproduce the X-ray structure. Just as shown in Fig. 1(a), the cocrystallized **824** and redocked **824** are almost at the same position in the active sites of Weel. Therefore, the GOLD method and the parameter set could be extended to search the Weel binding conformations for other inhibitors.

Fig. 1(b) shows the 3D model of 4-phenylpyrrolo[3,4-*c*]carbazole-1,3(2*H*,6*H*)-diones at the active sites of Weel and Fig. 1(c) illustrates the probable binding conformational alignment for the 97 4-phenylpyrrolo[3,4-*c*]carbazole-1,3(2*H*,6*H*)-diones chosen from the docked conformations. All of the inhibitors bonded in the active sites of Weel in

Table 2

Actual and predicted inhibitory activities (pIC<sub>50</sub>) and residuals of the training set molecules (using CoMFA and best model from CoMSIA analysis)

| Compound numbers | ChemScore | Actual pIC <sub>50</sub> | Predicted pIC <sub>50</sub> |                     | Residuals |                     |
|------------------|-----------|--------------------------|-----------------------------|---------------------|-----------|---------------------|
|                  |           |                          | CoMFA                       | CoMSIA <sup>a</sup> | CoMFA     | CoMSIA <sup>a</sup> |
| 1                | 48.24     | 7.013                    | 6.773                       | 6.781               | 0.240     | 0.232               |
| 2                | 42.67     | 5.398                    | 5.520                       | 5.820               | −0.122    | −0.422              |
| 3                | 44.37     | 5.638                    | 5.656                       | 5.661               | −0.018    | −0.023              |
| 4                | 48.49     | 6.509                    | 6.497                       | 6.462               | 0.012     | 0.047               |
| 6                | 49.65     | 7.108                    | 7.123                       | 7.070               | −0.015    | 0.038               |
| 7                | 49.65     | 6.585                    | 6.727                       | 6.853               | −0.142    | −0.268              |
| 8                | 49.37     | 6.886                    | 6.842                       | 6.744               | 0.044     | 0.142               |
| 9                | 48.28     | 5.796                    | 6.022                       | 6.186               | −0.226    | −0.390              |
| 10               | 43.71     | 5.013                    | 5.284                       | 5.159               | −0.271    | −0.146              |
| 11               | 46.84     | 5.638                    | 5.604                       | 5.532               | 0.034     | 0.106               |
| 12               | 43.44     | 5.398                    | 5.311                       | 5.123               | 0.087     | 0.275               |
| 15               | 46.76     | 5.553                    | 5.564                       | 5.691               | −0.011    | −0.138              |
| 16               | 48.36     | 5.409                    | 5.427                       | 5.393               | −0.018    | 0.016               |
| 17               | 50.89     | 7.959                    | 7.780                       | 7.483               | 0.179     | 0.476               |
| 19               | 50.45     | 7.244                    | 7.474                       | 7.336               | −0.230    | −0.092              |
| 20               | 50.93     | 7.481                    | 7.459                       | 7.556               | 0.022     | −0.075              |
| 21               | 50.68     | 7.638                    | 7.680                       | 7.556               | −0.042    | 0.082               |
| 23               | 49.57     | 7.886                    | 7.697                       | 7.737               | 0.189     | 0.149               |
| 24               | 50.69     | 6.824                    | 7.218                       | 7.149               | −0.394    | −0.325              |
| 26               | 40.56     | 6.237                    | 6.278                       | 6.403               | −0.041    | −0.166              |
| 27               | 50.66     | 6.347                    | 6.036                       | 6.355               | 0.311     | −0.008              |
| 28               | 48.92     | 6.721                    | 6.674                       | 6.969               | 0.047     | −0.248              |
| 29               | 45.45     | 6.081                    | 6.166                       | 6.059               | −0.085    | 0.022               |
| 30               | 45.50     | 6.796                    | 6.613                       | 6.922               | 0.183     | −0.126              |
| 32               | 49.44     | 7.222                    | 6.840                       | 6.754               | 0.382     | 0.468               |
| 33               | 48.52     | 7.620                    | 7.678                       | 7.611               | −0.058    | 0.009               |
| 35               | 48.46     | 7.481                    | 7.571                       | 7.499               | −0.090    | −0.018              |
| 36               | 46.34     | 6.658                    | 6.584                       | 6.620               | 0.074     | 0.038               |
| 37               | 42.48     | 7.328                    | 7.157                       | 7.213               | 0.171     | 0.115               |
| 38               | 49.42     | 6.678                    | 6.822                       | 6.647               | −0.144    | 0.031               |
| 39               | 47.56     | 6.658                    | 6.731                       | 6.589               | −0.073    | 0.069               |
| 40               | 48.62     | 7.260                    | 6.966                       | 7.155               | 0.294     | 0.105               |
| 43               | 52.10     | 5.357                    | 5.648                       | 5.319               | −0.291    | 0.038               |
| 44               | 46.60     | 6.745                    | 6.919                       | 6.745               | −0.174    | 0.000               |
| 45               | 50.49     | 5.367                    | 5.295                       | 5.184               | 0.072     | 0.183               |
| 46               | 47.64     | 4.398                    | 4.394                       | 4.595               | 0.004     | −0.197              |
| 48               | 47.36     | 6.208                    | 6.120                       | 6.159               | 0.088     | 0.049               |
| 49               | 47.13     | 6.523                    | 6.691                       | 6.665               | −0.168    | −0.142              |
| 50               | 46.30     | 6.137                    | 5.918                       | 6.023               | 0.219     | 0.114               |
| 53               | 46.75     | 5.481                    | 5.841                       | 6.003               | −0.360    | −0.522              |
| 54               | 44.35     | 5.921                    | 6.029                       | 5.974               | −0.108    | −0.053              |
| 56               | 45.02     | 5.444                    | 5.409                       | 5.454               | 0.035     | −0.010              |
| 58               | 45.89     | 4.921                    | 4.595                       | 4.962               | 0.326     | −0.041              |
| 59               | 46.03     | 4.538                    | 4.547                       | 4.453               | −0.009    | 0.085               |
| 60               | 41.21     | 5.959                    | 5.911                       | 6.094               | 0.048     | −0.135              |
| 62               | 48.83     | 7.553                    | 7.655                       | 7.637               | −0.102    | −0.084              |
| 65               | 46.06     | 7.639                    | 7.775                       | 7.574               | −0.136    | 0.065               |
| 66               | 46.41     | 7.620                    | 7.308                       | 7.569               | 0.312     | 0.051               |
| 67               | 45.77     | 6.310                    | 6.040                       | 5.932               | 0.270     | 0.378               |
| 69               | 48.00     | 7.699                    | 7.622                       | 7.620               | 0.077     | 0.079               |
| 70               | 50.21     | 7.553                    | 7.633                       | 7.484               | −0.080    | 0.069               |
| 71               | 50.08     | 7.347                    | 7.492                       | 7.487               | −0.145    | −0.140              |
| 72               | 49.45     | 7.824                    | 7.674                       | 7.819               | 0.150     | 0.005               |
| 73               | 46.70     | 7.699                    | 7.769                       | 7.899               | −0.070    | −0.200              |
| 74               | 49.46     | 7.456                    | 7.522                       | 7.676               | −0.066    | −0.220              |
| 76               | 45.67     | 6.018                    | 6.374                       | 6.080               | −0.356    | −0.062              |
| 77               | 50.27     | 7.125                    | 7.238                       | 7.239               | −0.113    | −0.114              |
| 80               | 46.02     | 7.569                    | 7.384                       | 7.653               | 0.185     | −0.084              |
| 81               | 48.05     | 7.538                    | 7.585                       | 7.746               | −0.047    | −0.208              |
| 82               | 45.78     | 7.745                    | 7.613                       | 7.513               | 0.132     | 0.232               |
| 83               | 43.95     | 7.310                    | 7.318                       | 7.151               | −0.008    | 0.159               |
| 84               | 48.69     | 6.854                    | 6.971                       | 6.835               | −0.117    | 0.019               |

Table 2 (continued)

| Compound numbers | ChemScore | Actual pIC <sub>50</sub> | Predicted pIC <sub>50</sub> |                     | Residuals |                     |
|------------------|-----------|--------------------------|-----------------------------|---------------------|-----------|---------------------|
|                  |           |                          | CoMFA                       | CoMSIA <sup>a</sup> | CoMFA     | CoMSIA <sup>a</sup> |
| 85               | 47.44     | 7.377                    | 7.143                       | 6.777               | 0.234     | 0.600               |
| 86               | 46.89     | 6.745                    | 6.443                       | 6.329               | 0.302     | 0.416               |
| 87               | 47.74     | 7.420                    | 7.253                       | 7.297               | 0.167     | 0.123               |
| 88               | 47.13     | 5.887                    | 6.055                       | 5.921               | −0.168    | −0.034              |
| 90               | 46.10     | 6.237                    | 6.306                       | 6.056               | −0.069    | 0.181               |
| 92               | 51.01     | 7.301                    | 7.327                       | 7.300               | −0.026    | 0.001               |
| 93               | 51.82     | 7.201                    | 7.195                       | 7.320               | 0.006     | −0.119              |
| 94               | 51.17     | 7.276                    | 7.259                       | 7.301               | 0.017     | −0.025              |
| 95               | 51.53     | 7.229                    | 7.408                       | 7.177               | −0.179    | 0.052               |
| 96               | 51.89     | 6.824                    | 6.939                       | 7.108               | −0.115    | −0.284              |
| 97               | 51.08     | 6.770                    | 6.795                       | 6.971               | −0.025    | −0.201              |

<sup>a</sup> Best model for CoMSIA.

a similar conformation to **824** in the X-ray structure cocrystallized with Weel (Fig. 1(b) and (c)), and the common chain structures superimposed rather well. Based on the set of binding conformations and their alignment, CoMFA and CoMSIA were performed.

### 3.2. Docking study

All 97 inhibitors were docked in the binding sites of Weel using GOLD. All orientations of 97 docking solutions were selected to perform the CoMFA and CoMSIA studies, and the ChemScores of the inhibitors are shown in Tables 2 and 3.

Fig. 2(a) and (b) represents the interaction model of the docked inhibitor **1** with Weel. Inhibitor **1** binds to the active sites and makes several interactions with the hinge-binding region of the enzyme. The docking graphs provide additional information concerning the position of ordered water molecules in and around the active sites and also identified a coordinated magnesium atom in the binding sites. While several of the ordered water molecules are sufficiently close to the ligand to provide useful targets for drug improvement, the magnesium atom and its coordination sphere are not. The pyrrolocarbazole binds with the pendant 4-phenyl ring twisted out of the plane of the chromophore and situated in the back of the ATP-binding pocket (Fig. 2(a) and (d)). The docking study showed that substituents at the 2'- and 3'-position of the 4-phenyl ring would determine the twisting direction of the ring. The docking gestures of the 4-phenyl ring could turn the inhibitors with substituents in the 2'- and 3'- position into two groups. One group includes compounds **7–25**, **33–35**, **40**, **62–66**, **68–71**, **73–75**, **77**, **80**, **81**, **83**, **85** and **92–97**, which are similar to compound **17** (Fig. 2(c)), whose lipophilic 2'- and 3'- substituents such as Cl, Br, I, OMe, and SMe may interact well with the nearest lipophilic side chain of Val313, Ala326, Lys328, Ile374 and Asn376. The other group, including compounds **26–32**, **36–39**, **41–50**, **67**, **72**, **76**, **78**, **79**, **82**, **84** and **86**, are similar to compound **26** (Fig. 2(d)), whose 2'- and 3'- substituents may not interact with the lipophilic side chain of above residues well and twist away from those residues. The 9-hydroxyl group in the A-ring of the chromophore is placed at the entrance to the pocket and forms an H-bond contact



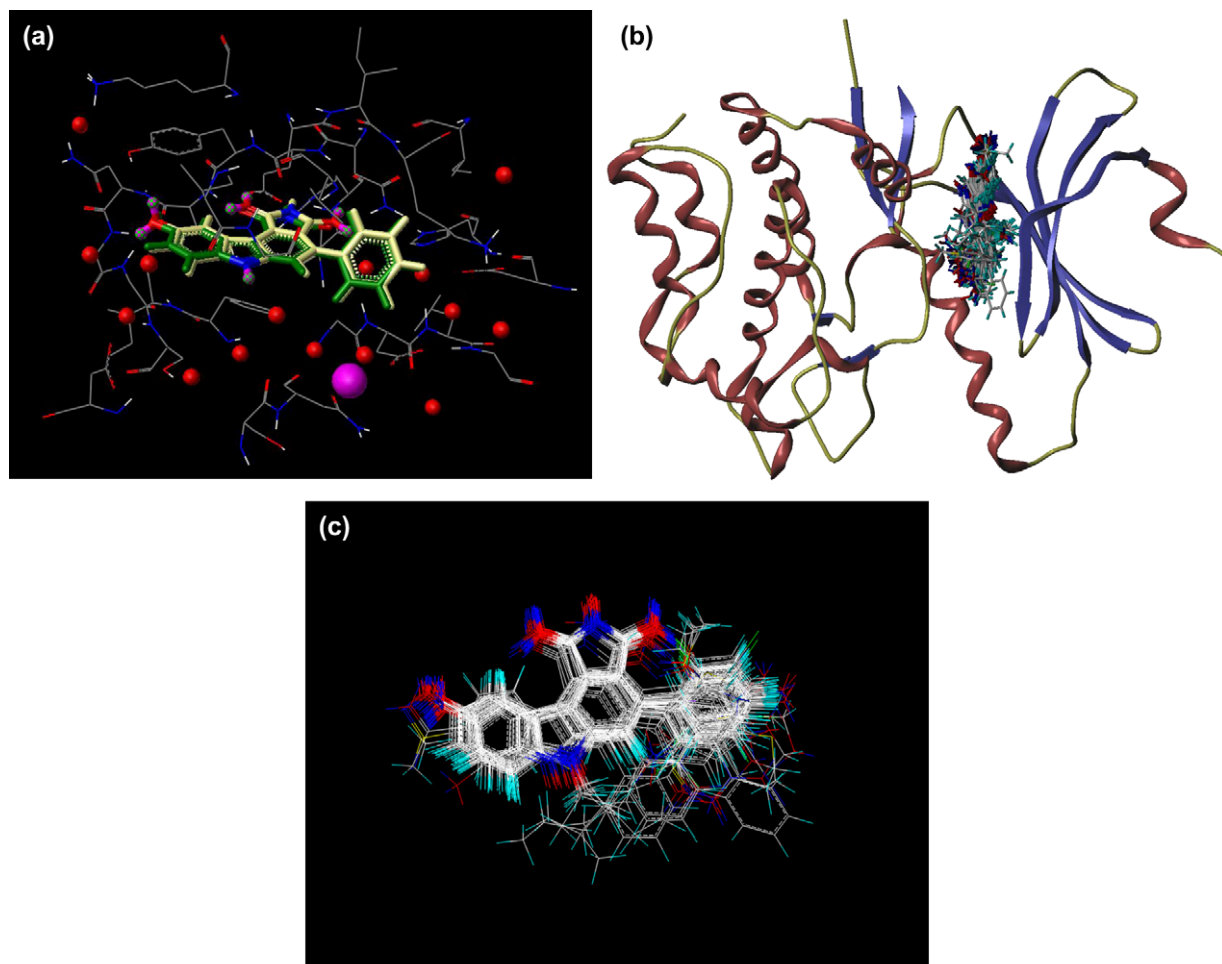


Fig. 1. (a) Binding conformations of the cocrystallized **824** and redocked **824** at the active sites of Weel. (Green compound represents cocrystallized **824** and yellow compound represents redocked **824**. Red balls represent water molecules and the pink ball represents a coordinated magnesium atom.) (b) Binding conformations of docked compounds at the active sites of Weel. (c) Superimposition of 97 4-phenylpyrrolo[3,4-c]carbazole-1,3(2H,6H)-diones for 3D-QSAR studies. (For interpretation of the references to color in this figure legend and text citation, the reader is referred to the web version of this article.)

Table 3

Actual and predicted inhibitory activities (pIC<sub>50</sub>) and residuals of the test set molecules (using CoMFA and best model from CoMSIA analysis)

| Compound numbers | ChemScore | Actual pIC <sub>50</sub> | Predicted pIC <sub>50</sub> |                     | Residuals |                     |
|------------------|-----------|--------------------------|-----------------------------|---------------------|-----------|---------------------|
|                  |           |                          | CoMFA                       | CoMSIA <sup>a</sup> | CoMFA     | CoMSIA <sup>a</sup> |
| <b>5</b>         | 48.37     | 6.367                    | 6.992                       | 7.026               | −0.625    | −0.659              |
| <b>31</b>        | 42.44     | 6.244                    | 5.692                       | 5.419               | 0.552     | 0.825               |
| <b>34</b>        | 44.19     | 6.585                    | 6.539                       | 6.439               | 0.046     | 0.146               |
| <b>41</b>        | 48.25     | 6.638                    | 6.332                       | 6.141               | 0.306     | 0.497               |
| <b>47</b>        | 45.67     | 7.051                    | 6.813                       | 6.515               | 0.238     | 0.536               |
| <b>51</b>        | 46.11     | 4.796                    | 5.467                       | 5.624               | −0.671    | −0.828              |
| <b>55</b>        | 41.04     | 5.745                    | 5.453                       | 5.469               | 0.292     | 0.276               |
| <b>61</b>        | 44.68     | 6.824                    | 6.359                       | 7.042               | 0.465     | −0.218              |
| <b>63</b>        | 45.19     | 7.921                    | 7.237                       | 6.961               | 0.684     | 0.960               |
| <b>64</b>        | 48.09     | 7.678                    | 7.461                       | 7.085               | 0.217     | 0.593               |
| <b>68</b>        | 45.56     | 7.377                    | 7.407                       | 7.412               | −0.030    | −0.035              |
| <b>75</b>        | 48.66     | 6.569                    | 7.130                       | 6.837               | −0.561    | −0.268              |
| <b>78</b>        | 43.41     | 7.721                    | 7.710                       | 7.580               | 0.011     | 0.141               |
| <b>89</b>        | 47.10     | 6.086                    | 6.227                       | 5.507               | −0.141    | 0.579               |

<sup>a</sup> Best model for CoMSIA.

with Cys379. The 1-carbonyl group accepts an H-bond from the same Cys379, and the 3-carbonyl group accepts two H-bonds from an ordered water and Asn376 (Fig. 2(b)). The planar A, B and C rings of these inhibitors are located near hydrophobic side chains of Ile305, Val313, Lys315 and Tyr378 (Fig. 2(c) and (d)). Phe433 forms a  $\pi$ -stacking interaction with the B and C rings of the planar pyrrolocarbazole chromophore (Fig. 2(b) and (d)). There is a room at the 6-position to accommodate moderately sized side chains, while side chains of the 8-position should protrude beyond the protein into solvent-accessible space (Fig. 2(a) and (d)). In addition, there are two interesting docking results on compounds **56** and **60**, whose docking gestures rotated about 180° from the others as shown in Fig. 2(e).

### 3.3. CoMFA model

Based on the results of QSAR studies from partial least-squares (PLS) analysis, 10 molecules (compounds **13**, **14**, **18**, **21**, **25**, **42**, **50**, **57**, **79** and **91**) of the training set with high residual values were omitted. The actual inhibition

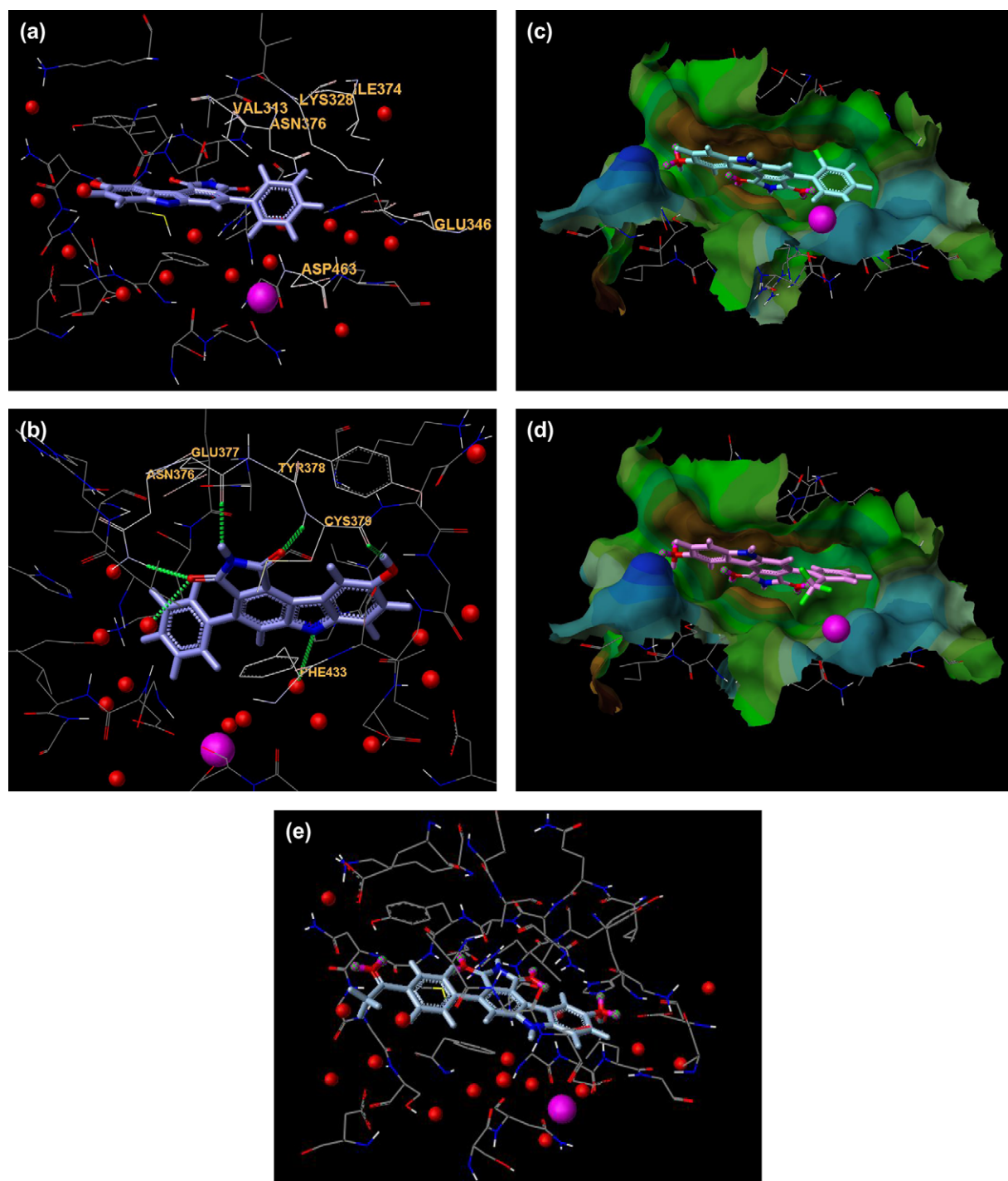


Fig. 2. (a) Docking structure of **1** bound in the ATP-binding sites of a construct of Weel kinase. The phenyl ring pocket residues were labeled. (Red balls represent water molecules and the pink ball represents a coordinated magnesium atom.) (b) Docking structure of **1** bound in the ATP-binding sites of a construct of Weel kinase. Important hydrogen bonds among the ligand, protein, and ordered water molecules in the active sites are indicated with dotted green lines. (Red balls represent water molecules and the pink ball represents a coordinated magnesium atom.) (c) Docking structure of **17** bound in the ATP-binding sites of a construct of Weel, with the sites color-coded by lipophilic potential. Blue areas are hydrophilic, dark yellow areas are lipophilic, and neutral areas are in green. (d) Docking structure of **26** bound in the ATP-binding sites of a construct of Weel, the sites color-coded by lipophilic potential. Blue areas are hydrophilic, dark yellow areas are lipophilic, and neutral areas are in green. (e) Docking structure of **56** bound in the ATP-binding sites of a construct of Weel kinase. (Red balls represent water molecules and the pink ball represents a coordinated magnesium atom. For interpretation of the references to color in this figure legend and text citation, the reader is referred to the web version of this article.)



abilities of all compounds were determined by using a functional construct of human Weel kinase (aa 215–647), while the docking study was performed by using a construct containing residues 291–575 of the human Weel enzyme [10]. The docking construct contained the substrate and ATP-binding sites of Weel and was thus useful for identifying the key features of ligand binding, though it did not retain the full catalytic activity of the larger construct of Weel enzyme that was used in the IC<sub>50</sub> enzyme measurements. Since the conformation of the construct of human Weel kinase (aa 215–647) was unknown, it was hard to decide whether the two constructs have the same influences when they bind the same inhibitor. It was acceptable to omit the 10 molecules whose actual pIC<sub>50</sub> values were not in good agreement with the predicted data within a tolerable error range. Finally, 73 4-phenylpyrrolo[3,4-c]carbazole-1,3(2*H*,6*H*)-dione inhibitors were randomly included in a training set for constructing CoMFA models, and the remaining 14 were used as a test set for model validation.

PLS analysis was carried out for the 73 binding conformations, and the results are listed in Table 4, which showed that a CoMFA model with a cross-validated  $q^2$  of 0.828 for five components was obtained. The non-cross-validated PLS analysis with the optimum components of five resulted in a conventional  $r^2$  of 0.962,  $F = 340.454$ , and an estimated standard error of 0.178. The steric field descriptors explain 55.7% of the variance, while the electrostatic descriptors explain 44.3%. The predicted activities for the 87 inhibitors versus their experimental activities with their residues are listed in Tables 2 and 3, and the correlation between the predicted activities and the experimental activities are depicted in Fig. 3(a). Tables 2–4 demonstrate that the activities predicted by the constructed CoMFA model are in good agreement with the experimental data, suggesting that a reliable CoMFA model was successfully constructed.

CoMFA steric and electrostatic contours are shown in Fig. 3(b) and (c). In addition, the CoMFA model plots were

mapped back to the binding sites of Weel, to get a better understanding of vital interactions between the inhibitors and the Weel enzyme. The steric interaction is represented by green and yellow contours, in which green-colored regions indicate areas where increased steric bulk is associated with the enhanced Weel enzyme inhibitory activity, and yellow regions suggest areas where increased steric bulk is unfavorable to activity. At the same time, electrostatic interaction is denoted by red and blue contours, among which blue-colored regions show areas where more positively charged groups are favorable to Weel enzyme inhibitory activity, and red regions represent areas where increased negatively charged groups are favorable to activity. In order to obtain clear insight into the electrostatic interaction between the inhibitors and Weel, the docking structure of **17** bound in the ATP-binding sites of a construct of Weel kinase was color-coded according to electrostatic potential (Fig. 3(d)), blue areas being negatively charged, red areas denoting positive charge, and green and cyan areas representing neutral.

A large green contour is found above the 2'- and 3'-position of 4-phenyl ring in compound **17** (Fig. 3(b)), which is surrounded by Val313, Ala326, Lys328, Ile374 and Asn376 that form a moderately lipophilic binding pocket in Weel enzyme. Compounds **17**, **19**, **20**, **22**, **23**, **33**, **35**, **40**, **62–66**, **68–74**, **77** and **80–83** which contain small hydrophobic 2'-substituents or 3'-substituents such as Cl, Br, I, OMe and SMe may well have hydrophobic interaction with these residues. This is a possible reason why these compounds have higher potency than compound **1**. In addition, one big and one small yellow contour are close to the 5'-substituent of the 4-phenyl ring in compound **17** (Fig. 3(b)), which is blocked by the polar side chain of Asp463. That may explain why compounds **42–46** and **48** are less potent than compound **1**. Another small sterically disfavored yellow region is located near the 9-OH group on the A-ring in compound **17**. The area is blocked by Cys379 and Asn380, and the 4'-COMe group on the 4-phenyl in compound **56** and the 4'-SO<sub>2</sub>Me group on the 4-phenyl in compound **60** are close to the yellow contour, which can explain the low activity of compounds **56** and **60**.

The CoMFA electrostatic contour plots are displayed in Fig. 3(d). One blue contour is found below the 6'-position of the 4-phenyl ring in compound **17**, indicating that negatively charged substituents in the area are unfavorable. The X-ray structure of Weel enzyme showed that there were two ordered water molecules in this area and the nearest residue was Asn431. Compounds **26–30**, **36** and **38** which contain 6'-electron-rich groups such as CF<sub>3</sub>, COMe, and SOME in the 4-phenyl ring may not interact with water and Asn431 well, so they show lower activity than compound **77** with 6'-Me in the 4-phenyl ring. Another blue contour with the positively charged hydrogen atoms of the 4'-OH and 4'-NH<sub>2</sub> groups lies in the region near the 4'-position of the 4-phenyl in compound **17**. The docking study showed that the hydrogen atoms of the 4'-OH and 4'-NH<sub>2</sub> groups formed H-bonds with the near ordered water molecule and Glu346, which explains why the activity of compounds **65**, **73**, **78** and **83** was higher than compound **1**.

Table 4  
Comparison of statistical parameters for two QSAR based models

| Parameter                          | CoMFA   | CoMSIA <sup>a</sup> |
|------------------------------------|---------|---------------------|
| PLS statistics                     |         |                     |
| Cross-Validated (LOO)              |         |                     |
| $q^2$ (CV correlation coefficient) | 0.828   | 0.796               |
| $N$ (number of components)         | 5       | 5                   |
| SEP (standard error of prediction) | 0.380   | 0.414               |
| Non-cross-validated                |         |                     |
| $r^2$ (correlation coefficient)    | 0.962   | 0.949               |
| SEE (standard error of estimate)   | 0.178   | 0.208               |
| $F$ (F-ratio)                      | 340.454 | 247.257             |
| Field distribution(%)              |         |                     |
| Steric                             | 55.7    | 17.9                |
| Electrostatic                      | 44.3    | 25.4                |
| Hydrophobic                        |         | 25.3                |
| Donor                              |         | 22.7                |
| Acceptor                           |         | 8.7                 |
| Testing set                        |         |                     |
| $r^2$ (correlation coefficient)    | 0.904   | 0.783               |
| $S$ (standard error of prediction) | 0.393   | 0.487               |

<sup>a</sup> Best model for CoMSIA.

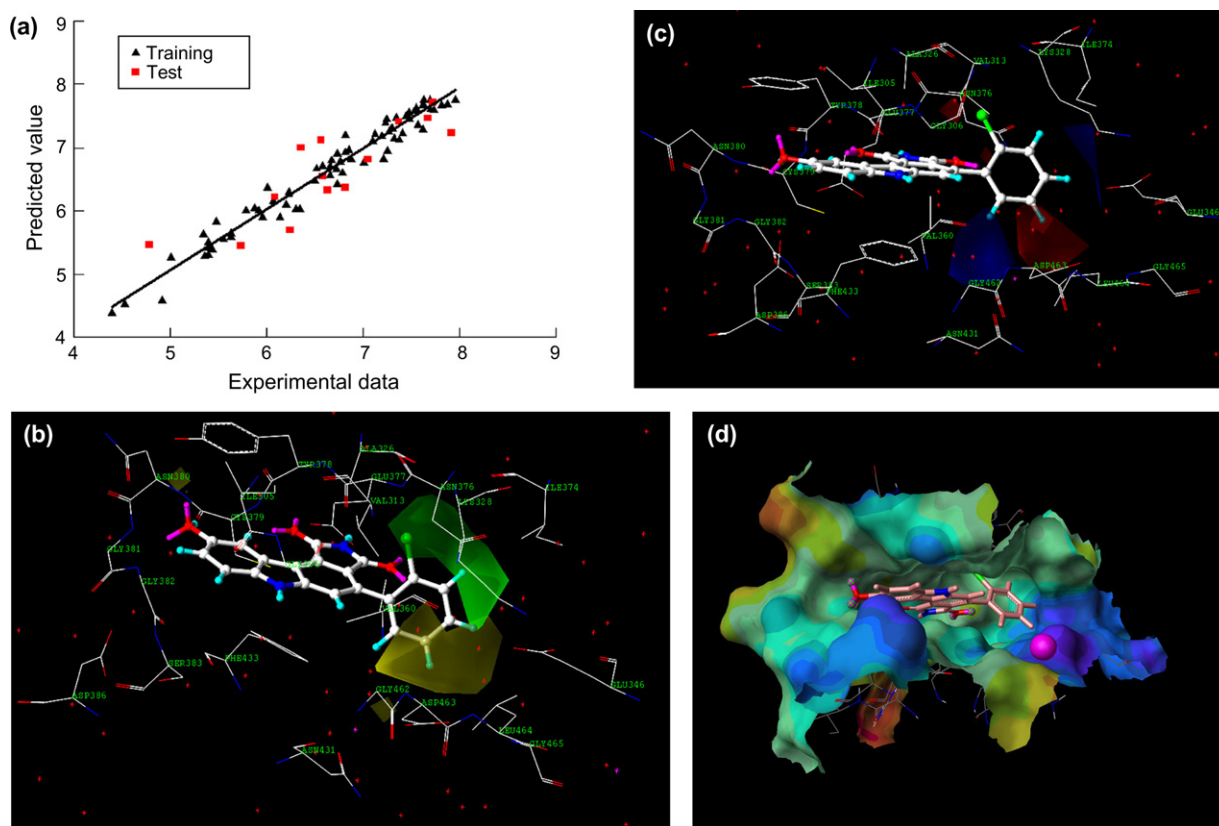


Fig. 3. (a) Correlation between predicted activities (PA) by CoMFA models and the experimental  $\text{pIC}_{50}$  values of training and test sets. Black triangles represent predictions for the training set, while red filled rectangles represent predictions for the test set. (b) CoMFA steric field contour maps in combination with inhibitor **17** within the active site of the complex structure of Weel. Sterically favored areas are in green, and sterically unfavored areas are in yellow. (Red dots represent water molecules and pink dots represent the coordinated magnesium atoms.) (c) CoMFA electrostatic field contour maps in combination with inhibitor **17** within the active site of the complex structure of Weel. Positive potential favored areas are in blue; positive potential unfavored areas are in red. (Red dots represent water molecules and pink dots represent coordinated magnesium atoms.) (d) Docking structure of **17** bound in the ATP-binding site of a construct of Weel, the sites color-coded by electrostatic potential. Blue areas are negatively charged, red areas denote positive charge, and neutral areas are in green and cyan. (For interpretation of the references to color in this figure legend and text citation, the reader is referred to the web version of this article.)

One big red contour is found near the 5'-position of the 4-phenyl in compound **17**, indicating a preference for negatively charged substituents in this area. The OH group is a strong negatively charged group, whereas the Me, OMe, and COMe substituents are weaker ones. The docking study found that the 3'-OH on the 4-phenyl ring in compounds **47**, **68** and **69** formed H-bonds with the residue Asp463 and two ordered water molecules, respectively, causing compounds **47**, **68** and **69** to display higher activity than compound **1**. The other two small red contours are found near 3-carbonyl in compound **17**, showing that negatively charged 3-carbonyl is favored in this area. The docking study also found that 3-carbonyl was located between Ala326 and Asn376 and formed two H-bonds with an ordered water molecule and Asn376, respectively, but the 3'-OH in compound **15** did not form H-bonds with any nearby residue. This is the reason why compound **15** is less potent than compound **1**.

### 3.4. CoMSIA model

A total of 16 CoMSIA models were generated using either single or combined fields. The alignment used in the CoMSIA

study was performed as in the CoMFA study. The statistical parameters are summarized in Tables 4 and 5. Most of the models including hydrophobic and steric fields showed a high cross-validated  $q^2$  value, which indicates the importance of lipophilicity and steric factor for the present series of molecules. Among the different field combinations, the model 15 with five fields — steric, electrostatic, donor, acceptor and hydrophobic descriptors — had the highest cross-validated  $q^2$  value of 0.796 with five components, and the non-cross-validated PLS analysis with the optimum components of five had a conventional  $r^2$  of 0.949,  $F = 247.257$ , and an estimated standard error of 0.208. So the model 15 was considered as the best CoMSIA model and was used to generate contour plots. Furthermore, the CoMSIA model plots were mapped back to the binding sites of Weel enzyme (1X8B) [13] for better understanding of vital interactions between the inhibitors and the kinase. Since the steric contours of CoMSIA are very similar to those of the CoMFA (Fig. 3(b)), only hydrophobic and hydrogen bonding interaction fields will be described. The predicted activities for the 87 inhibitors versus their experimental activities with their residues are listed in Tables 2 and 3, and the correlation between the predicted activities and the experimental activities are depicted in

Table 5  
Summary of CoMSIA analysis

| Model | Field                          | $q_{cv}^2$ | SEP   | <i>N</i> | $r^2$ | SEE   | <i>F</i> (F-ratio) |
|-------|--------------------------------|------------|-------|----------|-------|-------|--------------------|
| 1     | H <sup>a</sup>                 | 0.728      | 0.489 | 8        | 0.920 | 0.265 | 91.853             |
| 2     | D <sup>b</sup>                 | 0.485      | 0.648 | 3        | 0.599 | 0.572 | 34.360             |
| 3     | A <sup>c</sup>                 | 0.350      | 0.728 | 3        | 0.519 | 0.627 | 24.774             |
| 4     | S <sup>d</sup>                 | 0.707      | 0.482 | 1        | 0.751 | 0.445 | 213.703            |
| 5     | E <sup>e</sup>                 | 0.389      | 0.701 | 2        | 0.563 | 0.593 | 45.147             |
| 6     | D + A                          | 0.457      | 0.666 | 3        | 0.628 | 0.551 | 38.797             |
| 7     | S + E                          | 0.712      | 0.488 | 4        | 0.885 | 0.309 | 130.539            |
| 8     | H + D + A                      | 0.789      | 0.431 | 8        | 0.957 | 0.195 | 176.388            |
| 9     | H + S + E                      | 0.748      | 0.453 | 3        | 0.875 | 0.319 | 161.614            |
| 10    | S + E + D                      | 0.731      | 0.472 | 4        | 0.898 | 0.290 | 150.018            |
| 11    | S + E + A                      | 0.727      | 0.475 | 4        | 0.892 | 0.299 | 140.654            |
| 12    | D + A + E                      | 0.517      | 0.632 | 4        | 0.771 | 0.435 | 57.293             |
| 13    | D + A + S                      | 0.721      | 0.477 | 3        | 0.846 | 0.354 | 126.762            |
| 14    | S + E + D + A                  | 0.736      | 0.467 | 4        | 0.900 | 0.288 | 152.789            |
| 15    | S + E + D + A + H <sup>f</sup> | 0.796      | 0.414 | 5        | 0.949 | 0.208 | 247.257            |
| 16    | All                            | 0.791      | 0.416 | 4        | 0.925 | 0.249 | 209.238            |

<sup>a</sup> Hydrophobic field.

<sup>b</sup> Hydrogen bond donor field.

<sup>c</sup> Hydrogen bond acceptor field.

<sup>d</sup> Steric field.

<sup>e</sup> Electrostatic field.

<sup>f</sup> Best model for CoMSIA.

Fig. 4(a). Tables 2–4 and Fig. 4(a) demonstrate that the activities predicted by the constructed CoMSIA model are in good agreement with the experimental data, suggesting that a reliable CoMSIA model was successfully constructed.

Fig. 4(b) displays the hydrophobic contour maps. Yellow polyhedra indicate regions where hydrophobic substituents are favored and white polyhedra indicate disfavored regions. One big white contour penetrating the plane of 4-phenyl ring in compound **17** is blocked by a few ordered water molecules and the polar residues Lys328, Asp463 and Asn376, showing that compounds **10–12**, **45**, **46**, **48** and **59** are actively lower than compound **1** in this area. Also, two small yellow contours are found near the 2'- and 3'-positions of the 4-phenyl ring in compound **17**. These yellow contours are surrounded by the residues Val313, Ala326, Lys328, Ile374 and Asn376, which formed a moderately lipophilic binding pocket in Weel kinase. This is a possible reason why compounds **17**, **19**, **20**, **22**, **23**, **33**, **35**, **40**, **62–66**, **68–74**, **77** and **80–83** with small hydrophobic 2'- or 3'-substituents such as Cl, Br, I, OMe and SMe on the 4-phenyl rings have higher potency than compound **1**.

Fig. 4(c) displays the hydrogen bond donor plot represented by cyan and purple contours. Cyan contours indicate regions where hydrogen bond donor substituents on ligands are favored and purple contours represent areas where hydrogen bond donors of inhibitors are disfavored. There are clearly two cyan contours and two purple contours in the hydrogen bond donor maps. One big cyan contour is near the 4'- and 5'-position on the 4-phenyl ring, indicating that hydrogen bond donor functionalities such as OH or NH<sub>2</sub> in this region will enhance the activity. The docking study showed that the OH or NH<sub>2</sub> groups in this region formed H-bonds with both

the carbonyl group of Glu346 or Asp463 and the near ordered water molecules, which explained well why compounds **57**, **65**, **66**, **68**, **73**, **78**, **82**, and **83** with hydrogen bond donor groups in this region have higher activity than compound **1**. The second tiny cyan contour is observed near the N6 position of compound **17**. An ordered water molecule is located in almost the same region in the Weel enzyme (1X8B). The docking study showed that a few compounds containing 6-NH moiety near the tiny cyan contour formed weak H-bonds with the ordered water molecule. This may be one reason why compounds **1**, **17**, **23**, **33**, **37**, **40**, **47**, **57**, **62**, **66**, **68**, **70**, **71**, **74**, **77**, **78**, **80–83**, **85** and **87** with 6-NH moiety near the tiny cyan contour have high activities. Conversely, according to the docking results, most other compounds whose 6-NH moiety are not near the tiny cyan contour did not form H-bonds with the ordered water molecule at their 6-positions. It is not strange that there is a big purple contour on the left side of the tiny cyan contour located near the 6-positions of these compounds (Fig. 4(c)). The second purple contour is observed below the 2-NH imide group in compound **17**. The inhibitors with their 2-NH imide group near the second purple contour have low activity. The docking study showed that their 2-NH imide moiety did not donate an H-bond to Glu377. This may be why the second purple contour appears in this region.

Fig. 4(d) displays the CoMSIA hydrogen bond acceptor contour plot within the active site of the complex structure of Weel kinase. Magenta contours represent regions where hydrogen bond acceptor groups on ligands enhance activity, while red contours indicate areas where hydrogen bond acceptor groups on inhibitors are unfavorable for the activity. A large magenta contour is found near the 4-phenyl of

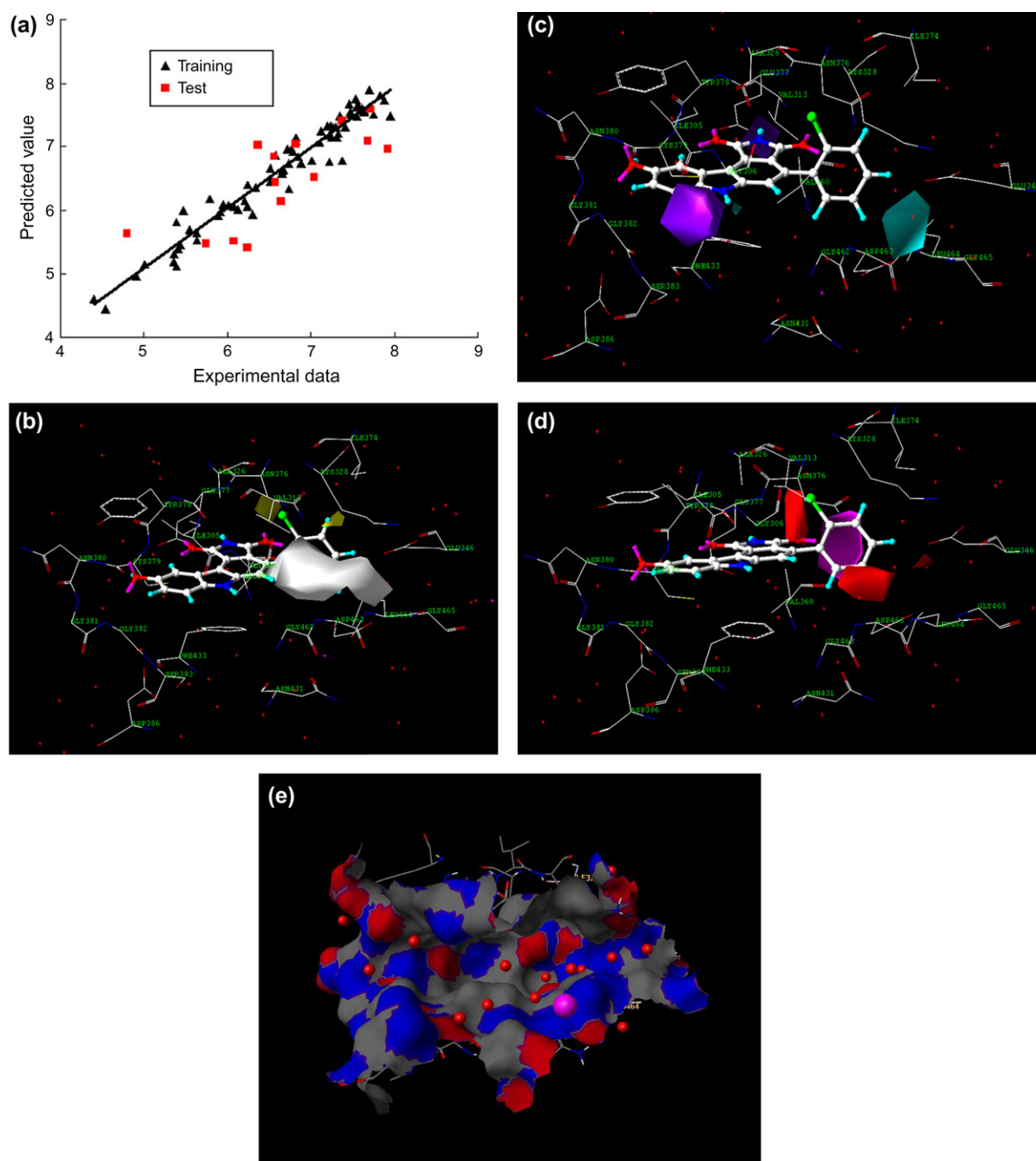


Fig. 4. (a) Correlation between predicted activities (PA) by CoMSIA models and the experimental  $pIC_{50}$  values of training and test sets. Black triangles represent predictions for the training set, while red filled rectangles represent predictions for the test set. (b) CoMSIA hydrophobic fields in combination with inhibitor **17** within the active sites of the complex structure of Wee1. Yellow polyhedra indicate regions where hydrophobic substituents are favored and white polyhedra indicate disfavored regions. (Red dots represent water molecules and pink dots represent coordinated magnesium atoms.) (c) CoMSIA hydrogen bond donor contour plots in combination with inhibitor **17** within the active site of the complex structure of Wee1. Cyan contours represent areas where hydrogen bond donor groups on the ligand enhance activity whereas purple contours indicate area where hydrogen bond donor groups on the ligand decrease activity. (Red dots represent water molecules and pink dots represent coordinated magnesium atoms.) (d) CoMSIA hydrogen bond acceptor contour plots in combination with inhibitor **17** within the active sites of the complex structure of Wee1. Magenta contours represent regions where hydrogen bond acceptor groups on the ligand enhance activity while red contours indicate areas where hydrogen bond acceptor groups on the ligand decrease activity. (Red dots represent water molecules and pink dots represent coordinated magnesium atoms.) (e) The ATP-binding sites of a construct of Wee1 kinase which are color-coded by H-bonding regions. Red represents H-bonding donors in the protein, and blue represents H-bond acceptors. (Red balls represent water molecules and the pink ball represents a coordinated magnesium atom. For interpretation of the references to color in this figure legend and text citation, the reader is referred to the web version of this article.)



compound **17** and located near Asn376 and one ordered water molecule, suggesting that hydrogen bond acceptor features of ligands in this area will enhance inhibition potency. The docking study showed that the 3-carbonyl moiety formed H-bonds with the ordered water molecule and a hydrogen atom in the side chain amide group of Asn376. Compound **15** showed low activity with its 3-carbonyl group being replaced by –OH and not forming an H-bond with Asn376. In addition, there are three red contours in the hydrogen bond acceptor maps suggesting that hydrogen bond acceptor features of ligands in this area may reduce activity. The first one is just above the 3-carbonyl group of compound **17**, and is located near the carbonyl group (hydrogen bond acceptor group) of Asn376. Compounds **43** and **45** with their 3-carbonyl group near this contour may not interact well with the carbonyl group of Asn376. As for the other two red contours, one big contour is found between 5'-phenyl of compound **17** and the side chain carboxyl group of Asp463, and one small contour is located between the 4'-phenyl of compound **17** and the side chain carboxyl group of Glu346. The docking study showed that compounds containing hydrogen bond donor groups in these regions might form H-bonds with the side chain carboxyl group of Asp463 and Glu346, but hydrogen bond acceptor groups near the contours might have poor interaction with the side chain carboxyl group of Asp463 and Glu346, which could explain the low activity of compounds **45**, **56** and **60** with hydrogen bond acceptor groups near the contours.

### 3.5. Lipophilic potential, electrostatic potential, and H-bonding regions of Weel enzyme

In order to better understand the structural and chemical features of the Weel enzyme and inhibitors, the active site of the Weel enzyme was color-coded by lipophilic potential, electrostatic potential and H-bonding regions.

Fig. 2(c) and (d) shows the active site of the Weel enzyme as color-coded by lipophilic potential. Hydrophilic areas are in blue, lipophilic areas are in dark yellow, and neutral areas are in green. The region above the plane of A, B and C rings of the compounds **17** and **26** is in dark yellow, indicating the region is lipophilic. The hydrophobic side chain of Ile305, Val313, Lys315, Ala326 and Tyr378 may form hydrophobic interactions with the hydrophobic A, B and C rings. Just below the plane of the B and C rings, the region is also dark yellow. The aromatic side chain of Phe433 forms a  $\pi$ -stacking interaction with the B and C rings and the hydrophobic side chain Val360 forms a hydrophobic interaction with the B and C rings. The 4-phenyl binding pocket was considered hydrophobic by Palmer et al. [10], but in fact except that the region near the 1'- and 2'-position of the 4-phenyl ring of compound **17** is hydrophobic and the area on the edge of the 4'- and 5'-position of the 4-phenyl ring of compound **17** is hydrophilic, the 4-phenyl binding pocket is rather neutral as shown in Fig. 2(c) and (d). Since the aromatic 4-phenyl ring is rather hydrophobic, whereas the side chains of the phenyl bind pocket residues Val313, Ala326, Ile327, Lys328, Leu464, Glu346, His350, Val360, Arg361, Gly462, Asp463,

Ile374 and Asn376 are mostly polar, it can be guessed that a neutral substructure with a size similar to phenyl would be more suitable than a phenyl ring in the 4-phenyl binding pocket.

Fig. 3(d) illustrates the docking structure of compound **17** bound in the ATP-binding site of a construct of Weel sites color-coded by electrostatic potential. Blue areas are negatively charged, red areas denote positive charge, and neutral areas are in green and cyan. The whole binding sites are nearly neutral except that the areas blue color-coded at the edge of the 4'-, 5'- and 6'-position of the 4-phenyl ring of compound **17** are negatively charged, for they are near the negatively charged side chain carboxyl groups of Glu346, Asp463 and Asn431, indicating that positively charged group near the blue region would increase binding affinity.

Fig. 4(e) displays the ATP-binding sites of a construct of Weel kinase color-coded by H-bonding regions, in which red denotes H-bonding donors in the protein and blue represents H-bond acceptors. In the A, B, C and D rings' binding pocket, there are four H-bonding regions in the inner side which have an alternate blue and red pattern. As discussed above, the backbone carbonyl of Cys379 forms an H-bond contact with the 9-hydroxyl group in the A-ring; the 1-carbonyl group is hydrogen-bonded to the Cys379 NH group; the 2-NH group in the D ring forms an H-bond with the backbone carbonyl of Glu377; and the 3-carbonyl group accepts an H-bond from the amide group of Asn376. The above four residues form the H-bonded hinge-binding regions of the enzyme. Furthermore, in the inner 4-phenyl bind pocket there are at least four red regions representing four H-bond donors, including two amide hydrogens of the side chain of Asn376, one amide hydrogen of the backbone of Asp463 and one amide hydrogen of the backbone of Leu464. Also, there are at least two blue regions in the inner top 4-phenyl binding pocket, and two backbone carbonyl groups of Ala326 and Ile374 just near the regions, indicating that they can act as potential H-bond acceptors. However, the region near the 2'- and 3'-position of the 4-phenyl of compound **17** is hydrophobic, so a hydrophilic H-bond donor group may not appear in this region and interact with Ala326 and Ile374. Still, there is a large blue area on the bottom edge of the 4-phenyl pocket, where the two side chain carboxyl groups of Glu346 and Asp463 are nearby. The docking study showed that H-bond donors near this region would form H-bonds with the side chain carboxyl groups.

From the study above, we can speculate that a neutral 4-substituent with an H-bond acceptor inside the 4-phenyl pocket and H-bond donor groups outside the bottom 4-phenyl pocket may have high affinity to the Weel enzyme. Compounds **84–91** were evaluated with the 4-phenyl substituent replaced with a heterocyclic ring. Since most of these heterocyclic rings are rather hydrophobic and have no H-bond acceptor inside the 4-phenyl binding pocket and no H-bond donor near the bottom edge of the 4-phenyl binding pocket, it is not strange that most of these compounds have low affinity to the Weel enzyme.



### 3.6. Validation of the 3D-QSAR models

The 14 randomly selected compounds (Table 3) were used as the test set to verify the constructed CoMFA and CoMSIA models. The calculated results are listed in Table 3 and displayed in Figs. 4(b) and 5(a) (red rectangles). The predicted  $\text{pIC}_{50}\text{'s}$  with the QSAR models were in good agreement with the experimental data within a statistically tolerable error range. In addition, in the CoMFA model, a correlation coefficient of  $r^2 = 0.904$  was found, while in the CoMSIA model, the correlation coefficient  $r^2$  was 0.783 (Table 4). The test results indicated that the CoMFA and CoMSIA models would be reliable for use in new Checkpoint Kinase Weel inhibitor design for developing drug leads against cancer.

### 4. Conclusion

By using the alignment scheme generated from the docking study, highly predictive CoMFA and CoMSIA models were developed for Weel inhibitors. The satisfactory models of CoMFA and CoMSIA were obtained with leave-one-out (LOO) cross-validated  $q^2$  values of 0.828 and 0.796, respectively, and the non-cross-validated PLS analyses with the optimum components of five had conventional  $r^2$  values of 0.962 and 0.949, respectively. CoMFA calculates steric fields using a Lennard–Jones potential, and electrostatic fields using a Coulombic potential [23]; while CoMSIA calculates steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor fields using Gaussian functions [24]. Both the CoMFA and the CoMSIA models obtained very high cross-validated  $q^2$  values, indicating the strong reliability of the two models. Moreover, the predictive ability was validated by some compounds not included in the training set. In addition, the 3D-QSAR results suggested that the hydrophobic group is the favored group at the 2'-position of the 4-phenyl substituent and the 4'- or 5'-H-bond donor group may enhance the inhibitory activity of Weel enzyme. It can even be concluded that 4-phenyl may be a good pattern substituent for Weel inhibitors. Further studies suggested that the neutral 4-position substituents with H-bond donor and acceptor groups are worth trying. These results demonstrated the power of a combined docking/QSAR approach to explore the probable binding conformations of compounds at the active sites of the protein target, and further provided useful information in understanding the structural and chemical features of Weel for designing and finding new potential inhibitors.

### Acknowledgements

The authors are grateful to the State Key Laboratory of Phytochemistry & Plant Resources in West China, and Xibuzhiguang Program of CAS for financial support. And we also thank Visiting Scholar, Ph.D. Laura Ediger for helping the revision of this article.

### References

- [1] L.L. Parker, P. Worms, *Science* 257 (1992) 1955–1957.
- [2] C.H. McGowan, P. Russell, *EMBO J.* 12 (1993) 75–85.
- [3] C.H. McGowan, P. Russell, *EMBO J.* 14 (1995) 2166–2175.
- [4] T. Kawabe, *Mol. Cancer Ther.* 3 (2004) 513–519.
- [5] A. Tenzer, M. Pruschy, *Curr. Med. Chem.* 3 (2003) 35–46.
- [6] I. Collins, M.D. Garrett, *Curr. Opin. Pharmacol.* 5 (2005) 366–373.
- [7] B.D. Palmer, J.B. Smaill, G.W. Rewcastle, E.M. Dobrusin, A. Kraker, C.W. Moore, R.W. Steinkampf, W.A. Denny, *Bioorg. Med. Chem. Lett.* 15 (2005) 1931–1935.
- [8] O.A. Mizenina, M.M. Moasser, *Cell Cycle* 3 (2004) 796–803.
- [9] J. Wissing, K. Godl, D. Brehmer, S. Blencke, M. Weber, P. Habenberger, M. Stein-Gerlach, A. Missio, M. Cotton, S. Muller, H. Daub, *Mol. Cell. Proteomics* 3 (2004) 1181–1193.
- [10] B.D. Palmer, A.M. Thompson, R. J Booth, E.M. Dobrusin, A.J. Kraker, H.H. Lee, E.A. Lunney, L.H. Mitchell, D.F. Ortwine, J.B. Smaill, L.M. Swan, W.A. Denny, *J. Med. Chem.* 49 (2006) 4896–4911.
- [11] B. Zhao, M.J. Bower, P.J. McDevitt, H. Zhao, S.T. Davis, K.O. Johanson, S.M. Green, N.O. Concha, B.–B.S. Zhou, *J. Biol. Chem.* 277 (2002) 46609–46615.
- [12] P. Chen, C. Luo, Y. Deng, K. Ryan, J. Register, S. Margosiak, A. Tempczyk-Russell, B. Nguyen, P. Myers, K. Lundgren, C.-C. Kan, P.M. O'Connor, *Cell* 100 (2000) 681–692.
- [13] C. Squire, J.M. Dickson, I. Ivanovic, E.N. Baker, *Structure* 13 (2005) 541–550.
- [14] Sybyl Version 6.9, Tripos Associates, St. Louis (MO), 2001.
- [15] M. Clark, R.D. Cramer, N.V. Opdenbosch, *J. Comput. Chem.* 10 (1989) 982–1012.
- [16] G. Jones, P. Willett, R.C. Glen, A.R. Leach, R.J. Taylor, *Mol. Biol.* 267 (1997) 727–748.
- [17] M.D. Eldridge, C.W. Murray, T.R. Auton, G.V. Paolini, R.P. Mee, J. Comput. –Aided Mol. Des. 11 (1997) 425–445.
- [18] C.A. Baxter, C.W. Murray, D.E. Clark, D.R. Westhead, M.D. Eldridge, *Proteins* 33 (1998) 367–382.
- [19] M.L. Verdonk, J.C. Cole, M.H. Hartshorn, C.W. Murray, R.D. Taylor, *Proteins* 52 (2003) 609–623.
- [20] T. Halgren, *Am. J. Chem. Soc.* 112 (1990) 4710–4723.
- [21] B.L. Podlogar, D.M. Fergusson, *Drug Des. Discov.* 17 (2000) 4–12.
- [22] M. Clark, R.D. Crammer III, D.M. Jones, D.E. Patterson, P.E. Simeroth, *Tetrahedron Comput. Methodol.* 3 (1990) 47–59.
- [23] R.D. Cramer III, D.E. Patterson, J.D. Bunce, *J. Am. Chem. Soc.* 110 (1988) 5959–5967.
- [24] G. Klebe, U. Abramam, *J. Comput. Aided Mol. Des.* 13 (1999) 1–10.