

# A Novel 3D Molecular Structural Characterization Method Applied in QSAR Study

JIANBO TONG<sup>1,2\*</sup>, YANG CHEN<sup>1,2</sup>, SHULING LIU<sup>1,2</sup>, TING CHE<sup>1,2</sup>, XIAMENG XU<sup>1,2</sup>

<sup>1</sup>College of Chemistry and Chemical Engineering, Shaanxi University of Science & Technology, Xi'an 710021, PR China

<sup>2</sup>Key Laboratory of Auxiliary Chemistry & Technology for Chemical Industry, Ministry of Education, Shaanxi University of Science & Technology, Xi'an 710021, PR China

*Three-dimensional holographic vector of atomic interaction field (3D-HoVAIF) was used to describe the chemical structures of 37 anti-Tuberculosis drug. Here quantitative structure activity relationship (QSAR) models were built by partial least square regression (PLS) and multiple linear regression (MLR). The estimation stability and generalization ability of the model was strictly analyzed by both internal and external validations. The correlation coefficient ( $R^2$ ) of established MLR and PLS models, Leave-One-Out (LOO) Cross-Validation (CV), prediction values versus experimental ones of external samples were  $R^2=0.939$ ,  $Q_{LOO}^2=0.906$ ,  $Q_{ext}^2=0.879$  (MLR) and  $R^2=0.848$ ,  $Q_{LOO}^2=0.805$ ,  $Q_{ext}^2=0.892$  (PLS), respectively. The results exhibited both favorable estimation stability and good prediction capabilities. Thus, this developed 3D-HoVAIF could preferably express information related to biological activity of arylamide derivatives.*

**Keywords:** QSAR, three-dimensional holographic vector of atomic interaction field (3D-HoVAIF), arylamide derivatives, partial least square regression (PLS), multiple linear regression (MLR)

Tuberculosis (TB) is a disease mainly caused by the infection of *mycobacterium tuberculosis* affecting approximately ten million people around the world [1]. It is a growing global health problem causing nearly three million deaths with eight million new cases each year [2]. Although TB can be cured with the proper drugs for at least six months, the emergence of multi- and extensively drug resistant has created substantial new challenges for the treatment of the disease [3, 4]. No new antibiotics against TB have been developed in the past 30 years. Accordingly, to address these problems the design of novel and more potent anti-tubercular agents has to be regarded as highly important.

Fortunately, the availability of computational techniques on quantitative structure activity relationships (QSAR) might provide a potential direction for accelerating the drug design process. In fact, QSAR can be viewed as a technique attempting to summarize chemical and biological information in a form that allows one to generate relationships between chemical structure and biological activity [5, 6]. However, 3D-QSAR studies on the anti-TB drug were still rare. Based on the previous works [5, 7] which use this method to do other drugs, a 3D molecular structural characterization method, the three-dimensional holographic vector of atomic interaction field (3D-HoVAIF) was proposed in our laboratory. 3D-HoVAIF method, deriving from multi-dimensional vectors to represent molecular steric structural characteristics, is independent of experiments and does not require conformation alignment, which is used to express drug structure and activity of arylamide derivative analogues and generates a good result.

## Theoretical section

3D-HoVAIF was developed considering three common non-bonding interactions of the biological activities, i.e., electrostatic interaction, steric interaction, and hydrophobic interaction related with atomic relative distance and atomic self-properties. These descriptors neither resort to

any experimental parameters nor consider configuration overlap of samples. Ordinary atoms of organic molecules including H, C, N, P, O, S, F, Cl, Br and I are classified into five types in the Periodic Table of Elements. According to the hybridization state of atoms, the atoms are furthermore subdivided into ten types. Thus, there are 55 interatomic interactions (Supporting table 1) in a molecule. In this paper, electrostatic, steric and hydrophobic potential energies take part in the representation of different interactions, producing  $3 \times 55 = 165$  interaction items for organic compounds.

### Electrostatic Interaction

Electrostatic interaction field is an important non-bonded interaction, which is expressed by classical Coulomb theorem:

$$E_{mn} = e^2 / 4\pi\epsilon_0 \sum_{i \in m, j \in n} (Z_i Z_j / r_{ij}) \quad (1 \leq m \leq 10, m \leq n \leq 10) \quad (1)$$

where  $r_{ij}$  is interatomic Euclid distance (nm),  $e$  the unit electric charge of  $1.6021892 \cdot 10^{-19}$  C,  $\epsilon_0$  the vacuum dielectric constant being  $8.85418782 \cdot 10^{-12}$  C<sup>2</sup>/J·m,  $Z$  the amounts of net electric charges;  $m$  and  $n$  are atomic types. The entire electrostatic interaction items are calculated by formula (1). Fifty-five electrostatic interaction items were calculated according to their attribute.

### Steric Interaction

Steric interaction describes interatomic spatial nondipole-dipole or dipole-induced interactions, which is described by Lennard-Jones formula:

$$S_{mn} = \sum_{i \in m, j \in n} \epsilon_{ij} D [(R_{ij}^* / r_{ij})^{12} - 2(R_{ij}^* / r_{ij})^6] \quad (1 \leq m \leq 10, m \leq n \leq 10) \quad (2)$$

where  $\epsilon_{ij}$  represent potential well [8, 9] (Supporting table 2),  $D$  is 0.01 [10], representing calibration constant of interatomic interaction deduced by experience;  $R_{ij}^* = (C_h \cdot R_{ii}^* + C_h \cdot R_{jj}^*) / 2$  is van der waals radius with its

\* email: jianbotong@yahoo.com.cn; 29-86168315

calibration factor of 1.00 in the case of  $sp^3$  hybridization, 0.95 in the case of  $sp^2$  hybridization and 0.90 in the case of  $sp$  hybridization [10].

### Hydrophobic Interaction

Hydrophobic interaction is frequently important for drug molecules. Indicating information on systematic entropic changes, such an interaction is thus difficult to be uniformly described. In 3D-HoVAIF, hydrophobic interaction force field is defined as interatomic hydrophobic interaction in hint method proposed by Kellogg [11-15].

$$H_{mn} = \sum_{i \in m, j \in n} S_i a_i S_j a_j e^{-r_{ij}} T_{ij} \quad (1 \leq m \leq 10, m \neq n \leq 10) \quad (3)$$

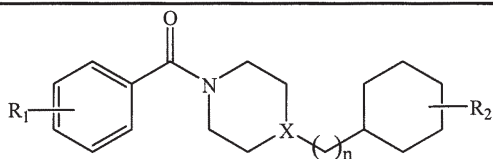
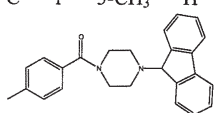
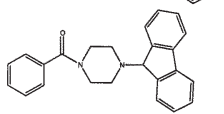
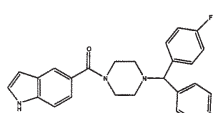
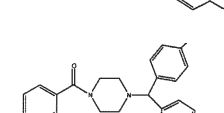
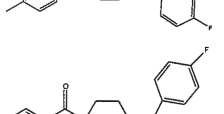
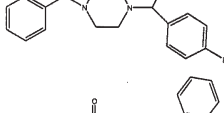
where  $S$  represents solvent accessible surface area of atom (SASA), *i.e.*, surface area formed by a hydrate probe spherically rolling at surface of this atom [16] (Supporting table 3);  $a$  is the hydrophobic constants [17] (Supporting table 4);  $T$  the sign function, indicating entropy change resulting from different types of atomic interactions [11-16] (Supporting table 5).

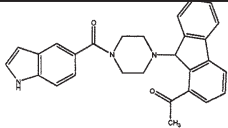
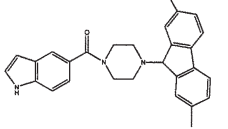
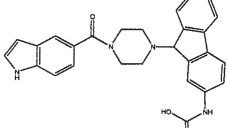
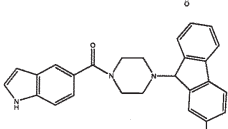
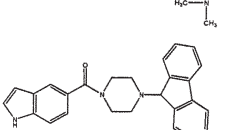
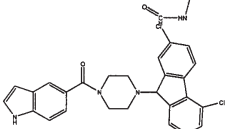
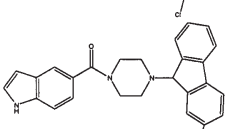
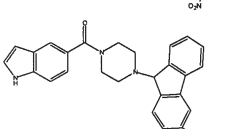
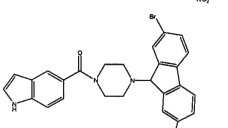
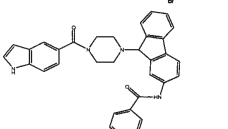
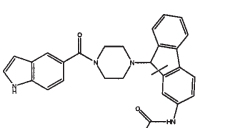
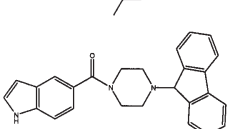
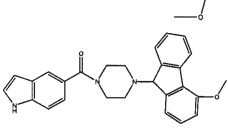
### Results and discussions

#### Structure and $-\log IC_{50}$ of 37 arylamide derivatives

The chemical structures of arylamide derivatives and their biological data are given in table 1 according to [18].  $IC_{50}$  is the drug concentration inhibiting 50% of the cellular growth followed by 1 hour of drug exposure.

**Table 1**  
CHEMICAL STRUCTURE AND BIOLOGICAL ACTIVITIES OF ARYLAMIDE DERIVATIVES

|  |   |   |                     |                   |       |        |        |        |        |
|--|---|---|---------------------|-------------------|-------|--------|--------|--------|--------|
| Compound   | X   | n | R <sub>1</sub>      | R <sub>2</sub>    | Obs   | Cal    |        |        |        |
|  |   |   |                     |                   |       | MLR(a) | MLR(b) | PLS(a) | PLS(b) |
| 1  | N   | 0 | H                   | H                 | 4.410 | 4.640  | 4.736  | 4.939  | 5.105  |
| 2 <sup>b</sup>   | N   | 0 | 4-CH <sub>3</sub>   | H                 | 4.779 | 5.161  | 5.221  | 4.937  | 5.166  |
| 3  | N   | 0 | 4-CH <sub>3</sub>   | 3-CF <sub>3</sub> | 5.203 | 5.26   | 5.159  | 5.514  | 5.019  |
| 4  | N   | 0 | 4-CH <sub>3</sub>   | 3-Cl              | 5.513 | 5.299  | 5.242  | 5.407  | 5.153  |
| 5 <sup>b</sup>   | N   | 0 | 3-CH <sub>3</sub>   | 3-Cl              | 5.025 | 4.887  | 5.300  | 4.904  | 5.320  |
| 6  | N   | 0 | 3-CH <sub>3</sub>   | 4-NO <sub>2</sub> | 4.811 | 4.646  | 4.760  | 4.944  | 5.218  |
| 7  | N   | 0 | 3,4-Me <sub>2</sub> | 3-Cl              | 6.004 | 5.929  | 5.944  | 5.381  | 5.346  |
| 8 <sup>b</sup>   | N   | 0 | 3,4-Me <sub>2</sub> | 3-CF <sub>3</sub> | 5.733 | 6.116  | 5.888  | 5.625  | 5.258  |
| 9  | N   | 0 | 2-F                 | 3-Cl              | 4.858 | 5.192  | 5.119  | 5.442  | 4.979  |
| 10 <sup>a</sup>  | N   | 0 | 4-F                 | 3-Cl              | 5.011 | 4.982  | 4.906  | 5.329  | 4.877  |
| 11   | N   | 0 | 3-Cl                | 3-Cl              | 5.172 | 4.963  | 4.925  | 5.406  | 4.974  |
| 12 <sup>a, b</sup>   | N   | 0 | 3,4-Cl <sub>2</sub> | 3-Cl              | 5.218 | 5.222  | 5.169  | 5.457  | 5.039  |
| 13   | N   | 0 | 3,4-Cl <sub>2</sub> | H                 | 4.754 | 5.047  | 5.123  | 4.968  | 5.022  |
| 14   | N   | 1 | H                   | H                 | 4.502 | 4.514  | 4.655  | 4.731  | 4.907  |
| 15 <sup>a</sup>  | C   | 1 | 3-Cl                | H                 | 5.172 | 4.667  | 4.820  | 4.722  | 4.887  |
| 16   | C   | 1 | 2-F                 | H                 | 4.850 | 4.844  | 4.984  | 4.754  | 4.859  |
| 17   | C   | 1 | 4-CH <sub>3</sub>   | H                 | 5.287 | 4.934  | 5.154  | 4.729  | 5.079  |
| 18   | C   | 1 | 3-CH <sub>3</sub>   | H                 | 5.131 | 5.006  | 5.193  | 4.729  | 5.080  |
| 19   |  |   |                     |                   | 6.398 | 6.651  | 6.618  | 6.493  | 6.618  |
| 20   |  |   |                     |                   | 7.046 | 6.424  | 6.497  | 6.496  | 6.449  |
| 21   |  |   |                     |                   | 5.983 | 5.698  | 5.826  | 5.987  | 5.865  |
| 22 <sup>a</sup>  |  |   |                     |                   | 5.724 | 6.06   | 5.985  | 6.005  | 5.959  |
| 23   |  |   |                     |                   | 5.69  | 6.025  | 5.801  | 5.992  | 5.776  |
| 24   |  |   |                     |                   | 6.796 | 6.492  | 6.604  | 6.495  | 6.554  |

|                 |   |       |       |       |       |       |
|-----------------|---|-------|-------|-------|-------|-------|
| 25              |    | 6.469 | 6.648 | 6.544 | 6.467 | 6.772 |
| 26              |    | 6.886 | 6.782 | 6.896 | 6.468 | 6.896 |
| 27 <sup>b</sup> |    | 6.553 | 6.777 | 6.861 | 6.491 | 6.564 |
| 28              |    | 6.041 | 6.163 | 5.909 | 6.416 | 6.280 |
| 29              |    | 6.745 | 6.727 | 6.873 | 6.486 | 6.593 |
| 30 <sup>a</sup> |    | 6.770 | 6.648 | 6.736 | 6.465 | 6.478 |
| 31 <sup>b</sup> |   | 6.886 | 6.63  | 6.765 | 6.491 | 6.556 |
| 32              |  | 6.886 | 6.852 | 6.876 | 7.129 | 6.802 |
| 33 <sup>a</sup> |  | 6.921 | 6.99  | 7.144 | 6.789 | 6.938 |
| 34 <sup>b</sup> |  | 6.229 | 6.586 | 6.817 | 6.490 | 6.858 |
| 35 <sup>a</sup> |  | 6.337 | 6.335 | 6.322 | 6.136 | 6.463 |
| 36              |  | 6.284 | 6.311 | 6.211 | 6.472 | 6.415 |
| 37              |  | 6.086 | 6.32  | 6.18  | 6.492 | 6.426 |

<sup>a</sup>Chosen as the test set in Ref[18].

<sup>b</sup>Chosen as the test set by author.

### Characterization of arylamide derivatives

Molecular steric structures of arylamide derivatives were firstly autoconstructed by Chemoffice 8.0, and then optimized at AM1 level by MOPAC half-experience

quantum chemistry software in Chem3D. Then net electric charge of atoms was calculated in single-point form by Mülliken methods. After the above two items were input respectively into forms of Descartes coordinates and net

electric charge amounts, 3D-HoVAIF descriptors were produced by applying 3D-HoVAIF.EXE, an applied program written in C language by our laboratory.

### Model validation

An excellent QSAR model should have both favorable estimation ability for any internal sample and outstanding predictive ability for any external sample. The usual method to prove a model to have internal predictive ability is leave-one-out (LOO) cross-validated (CV) method. In the present work, the statistical parameter correlation coefficient ( $Q_{LOO}^2$ ) [19] for internal validation criteria is used. Predictive performance of the model can be assessed by the prediction values of  $Q_{LOO}^2$ . Recently, several novel methods for model validation have been developed, such as leave-several-out (LSO) [20], Y-scrambling [21], self organizing mapping of molecular objects [22], external validation using division of a dataset into training and test sets [21, 23]. The external prediction power of QSAR can be evaluated by  $Q_{ext}^2$  [24, 25] as follows:

$$Q_{ext}^2 = 1 - \frac{\sum_{i=1}^{test} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{test} (y_i - \bar{y}_n)^2} \quad (4)$$

In eq. 4, both  $y_i$  and  $\hat{y}_i$  are the observed and calculated values of the test dataset, and  $\bar{y}_n$  is the mean value of observed values of the training dataset.

In order to prove the validity and stability of the model, the whole dataset is divided into two subsets (table 1), thus some samples were treated as training set which were utilized to construct QSAR model and the remaining samples were regarded as test set in order to validate the predictive power of the model.

### MLR modeling

Multiple linear regression (MLR) is a classic modeling technique. The descriptors are screened before being submitted to MLR analysis. In order to assure their statistical significance, only information-rich descriptors pass the screening step onto regression analysis. The forward stepwise multiple regression (SMR) method was employed for variables screening before the descriptors were submitted to MLR analysis. Statistical analysis was performed using the SPSS statistical package, version 13.0. Figure 1 indicated that the values of the correlation coefficients,  $R^2$  and  $Q_{LOO}^2$ , increased gradually with the increase in the number of the variables ( $n$ ). In addition, based on the past experience, one fine model shall comply with the empirical rule that the ratio number of samples ( $N$ )/number of variables ( $n$ ) shall be larger than 5. The

dataset was divided into a training set (30 molecules, table 1) and a test set (7 molecules, table 1). Therefore MLR analysis was used to construct the QSAR model by using the 6 descriptors. In the model of SMR-MLR (a), the molecules 10, 12, 15, 22, 30, 33 and 35 (in Table 1 marked a) were chosen as the test set in reference [18]. The relationship between all the structural descriptors and the activity ( $-\log IC_{50}$ ) of 30 components of the training set was modeled as follow:

$$\begin{aligned} \text{pMIC} = & 6.865641 - 0.000000 \times V_{121} + 0.000002 \times V_{129} - 5.484250 \times V_{27} \\ & + 0.204938 \times V_{67} - 3.214919 \times V_{12} - 47.678749 \times V_{73} \\ N = & 30 \quad R^2 = 0.939 \quad SD = 0.228 \\ F = & 58.153 \quad Q_{LOO}^2 = 0.906 \quad Q_{ext}^2 = 0.879 \end{aligned} \quad (5)$$

where  $N$  represents the number of samples used for model building,  $R$  is the multiple correlation coefficient,  $SD$  represents the standard deviation,  $F$  is the Fisher statistic. The subscript lowercase letters "LOO" and "ext" stand for the statistical results for training set in leave-one-out cross-validation and test set in the training procedure, respectively.

The relevant results are listed in table 2. According to table 2, the SMR-MLR results were obviously superior to the available reference reports, which indicate that the modeling has obviously a greater effect.

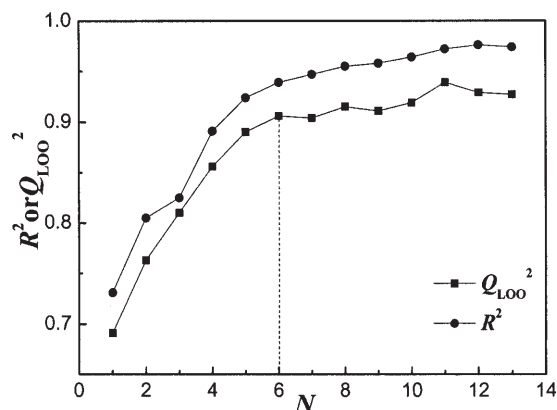


Fig. 1. Results of SMR with  $R^2$  and  $Q_{LOO}^2$  of MLR (a).

We choose another test set which will include the molecules 2, 5, 8, 12, 27, 31 and 34 (table 1 marked b). These supplementary studies were included because many properties calculated for QSAR purposes are critically dependent upon conformation. The same method was used to construct the QSAR model by using all the 6 descriptors. Table 2 lists the model of SMR-MLR (b) results.

| Methods    | Model      | $R_{cum}^2$ | $Q_{LOO}^2$ | $Q_{ext}^2$ |
|------------|------------|-------------|-------------|-------------|
| 3D-HoVAIF  | SMR-MLR(a) | 0.939       | 0.906       | 0.879       |
| 3D-HoVAIF  | SMR-MLR(b) | 0.937       | 0.903       | 0.808       |
| 3D-HoVAIF  | SMR-PLS(a) | 0.848       | 0.805       | 0.892       |
| 3D-HoVAIF  | SMR-PLS(b) | 0.867       | 0.842       | 0.741       |
| CoMFA[18]  | PLS        | 0.989       | 0.663       | 0.882       |
| CoMSIA[18] | PLS        | 0.963       | 0.639       | 0.875       |

**Table 2**  
COMPARING OF MODEL RESULTS  
BY DIFFERENT METHODS

(a) Chosen as the test set in Ref [18].

(b) Chosen as the test set by author.

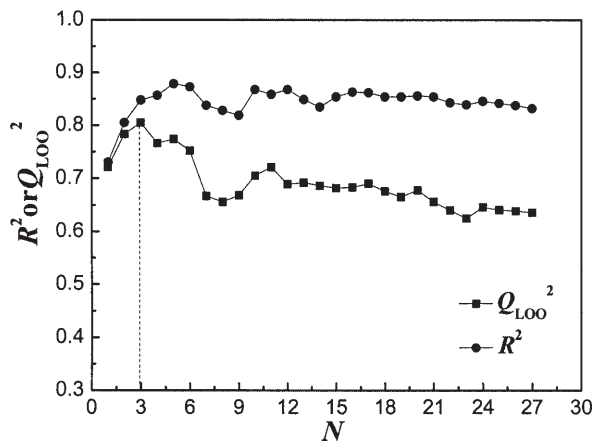


Fig. 2. Results of SMR with  $R^2$  and  $Q_{LOO}^2$  of PLS-(a)

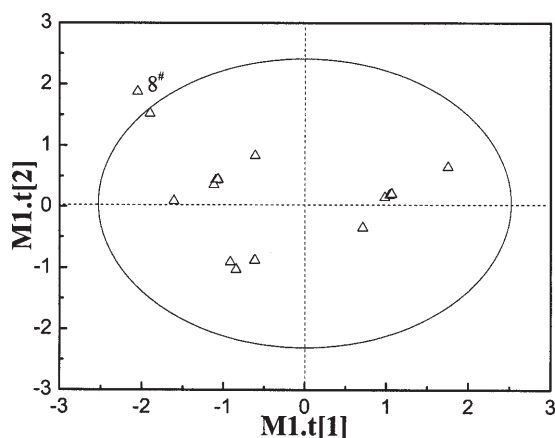


Fig. 3. Score distributions of the top two principal components for 30 samples in training set (a)

The good cross-validation value shows our model has good modeling stabilities.

#### PLS Modeling

PLS regression is mainly used for modeling linear regression between multi-dependent variables and multi-independent variables. Here variables selection is completed by stepwise multiple regression (SMR), which is employed at a given critical value to orderly introduce in terms of their importance and finally chosen out off significant variables.

The ultimate vectors for 37 arylamide derivatives were selected from 165 interaction items. SMR was implemented by SPSS 13.0. The obtained original variable matrix by SMR was then subject to a partial least square (PLS) regression modeling and the optimal model was determined when cross-validation correlative coefficients ( $Q_{LOO}^2$ ) in leave one out cross-validation (LOO-CV) achieved the maximum value (fig. 2). PLS was performed using Simca-12.0.

Table 2 lists reference reports and our SMR-PLS results for the molecules in training set and the rests in predicting, showing 3D-HoVAIF model greatly gains by comparison. The cross-validated correlation coefficient and predictive correlation coefficient were obviously superior to the reference. This study shows that the 3D-HoVAIF model has good modeling stabilities. These results indicate that the 3D-HoVAIF model is acceptable from statistical point of view. The score scatter of 30 samples at the top two principal components (PCs) is presented in figure 3, excluding the 8<sup>#</sup>, the others were in 95% confidence interval of Hotelling  $T^2$  ellipse. The excellent results of this study

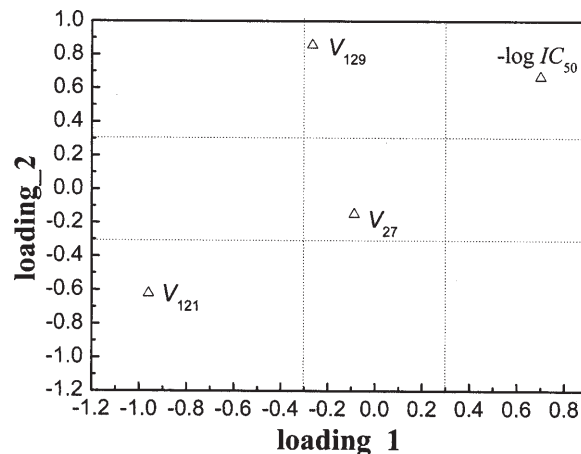


Fig. 4. Loading contributions of original variables to the top two principal components

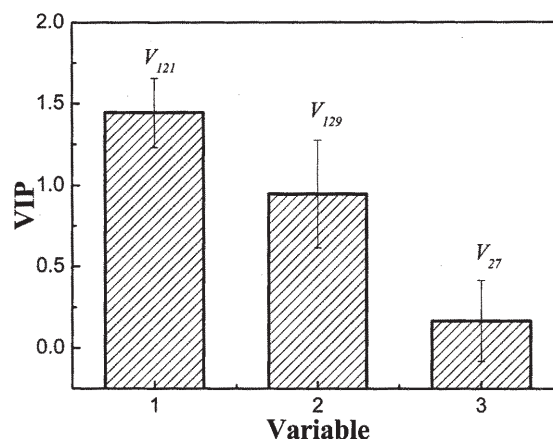


Fig. 5. Plot for variable importance of projection

should be recognized. Figure 4 presents loading contributions of 3 variables to the first two principal components, and it was seen that electrostatic and hydrophobic interactions have different contributions to them. Hydrophobic interactions ( $V_{121}$ ) had prominent contributions to PC1 loadings well correlative to  $Y$  variables, and hydrophobic ( $V_{121}$ ,  $V_{129}$ ) had prominent contributions to PC2. Besides, variable importance of projection (VIP) index of original variables to  $-\log IC_{50}$  is presented in figure 5. The fact that the most contributive top three items ( $V_{121}$ ,  $V_{129}$  and  $V_{27}$ ) including hydrophobic and electrostatic interactions indicated an intimate relationship between  $-\log IC_{50}$  for arylamide derivatives and the three classical interactions, especially the electrostatic and hydrophobic interactions. Therefore, it can be estimated that these interaction items will produce more influence on  $-\log IC_{50}$ , which brings a hopeful idea for designing new molecule and predicting their activities.

#### Conclusions

In this paper, quantitative structure activity relationship model was built to study the relationship between the biological activities of arylamide derivatives and their chemical structure by MLR and PLS methods. The models including classic electrostatic, steric and hydrophobic interactions have favorable stability and good predictive ability, it illustrates that 3D-HoVAIF is an effective description methodology for characterization of the complex interactions of drug molecules. Furthermore, the obtained model with obvious physicochemical meaning and strong structural interpretation is independent of



experimental data and configurationally overlapping required by other popular 3D modeling methods, such as comparative molecular field analysis (CoMFA). It was suggested that 3D-HoVAIF behaves quite well in representation of both molecular structures and biological activities for arylamide derivatives. So 3D-HoVAIF is promising to 3D-QSAR study and is expected to be widely used in the bioactivity prediction of various therapeutical drugs and other diverse substances.

*Acknowledgments: The authors appreciate the National Natural Science Funds of China (21275094), the financial support from the Scientific Research Planning Program of Shaanxi Province of China (2011K07-13), the Scientific Research Planning Program of the Education Department of Shaanxi Province (12JK0629)(11JK0602), the Scientific Research Planning Program of Key laboratory of Shaanxi Province of China(11JS022), the Scientific Research Planning Program of Xianyang(2011K09-09), City the Scientific Research Planning Program of Yulin City and the Graduate Innovation Fund of Shaanxi University of Science and Technology.*

## References

1. World Health Organization in WHO Report 2009 – “Global Tuberculosis Control: Epidemiology, Strategy, Financing”, WHO Press, Geneva. , 2009, p. 6.
2. FURIN, J.J., JOHNSON, J.L., Curr. Opin. Pulm. Med., **11**, 2005, p. 189.
3. VILCHEZE, C., WEISBROD, T.R., CHEN, B., KREMER, L., HAZBON, M.H., WANG F., ALLAND, D., SACCHERRINI, J.C., JACOBS. W.R., Agents Chemother, **49**, 2005, p. 708.
4. ZHANG, Y., Annu Rev Pharmacol Toxicol., **45**, 2005, p. 529.
5. TONG, J.B., LIU, S.L., QSAR & Comb. Sci., **3**, 2008, p. 330.
6. TONG, J.B., LI, Y.F., LIU, S.L., MENG, Y.L., Chinese J. Struct. Chem., **29**, 2010, p. 1893.
7. SUN, J.Y., CAI, S.X., YAN, N., MEI, H., Eur. J. Med. Chem., **45**, 2010, p. 1008.
8. LEVITT, M., J. Mol. Biol., **170**, 1983, p. 723.
9. LEVITT, M., PERUTZ, M.F., J. Mol. Biol., **201**, 1988, p. 751.
10. HAHN, M., J. Med. Chem., **38**, 1995, p. 2080.
11. KELLOGG, G.E., SEMUS, S.F., ABRAHAM, D.J., J. Comput-Aided Mol. Des., **5**, 1991, p. 545.
12. WIREKO, F.C., KELLOGG, G.E., ABRAHAM, D.J., J. Med. Chem., **34**, 1991, p. 758.
13. KELLOGG, G.E., JOSHI, G.S., ABRAHAM, D.J., Med. Chem. Res., **1**, 1992, p. 444.
14. ABRAHAM, D.J., KELLOGG, G.E., J. Mol. Graph., **10**, 1992, p. 212.
15. NAYAK, V.R., KELLOGG, G.E., Med. Chem. Res., **3**, 1994, p. 491.
16. HASEL, W., HENDRIKSON, T.F., STILL, W.C., Tetrahed Comp. Method, **1**, 1988, p. 103.
17. PEI, J.F., WANG, Q., ZHOU, J.J., Proteins: Struct. Funct. Genet., **57**, 2004, p. 651.
18. KUMAR, A., SIDDIQI, M.I., J. Mol. Model, **16**, 2010, p. 877.
19. WOLD, S., Technometrics, **20**, 1978, p. 897.
20. MITTAL, R.R., HARRIS, L., MCKINNON, R.A., SORICH, M.J., J. Chem. Inf. Model, **49**, 2009, p. 704.
21. GHARAGHEIZI, F., ALAMDARI, R.F., GOMBAR, V.K., QSAR & Comb. Sci., **27**, 2008, p. 679.
22. POLANSKI, J., BAK, A., GIELECIK, R., MAGDAIARZ, T., J. Chem. Inf. Model, **46**, 2006, p. 2310.
23. GOLBRIKH, A., TROPSHA, A., J. Mol. Graph. Model, **20**, 2002, p. 269.
24. TONG, J.B., CHE T., Li Y.F., WANG P., XV X.M., CHEN Y., SAR QSAR Environ Res, **22**, 2011, p. 611.
25. TONG, J.B., CHEN, Y., LIU S.L., CHE, T., XU X.M., J. Chemometr., **26**, 2012, p. 549

Manuscript received: 5.11.2012