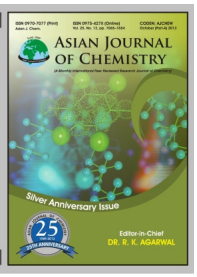
Asian Journal of Chemistry; Vol. 25, No. 17 (2013), 9643-9646

ASIAN JOURNAL OF CHEMISTRY

http://dx.doi.org/10.14233/ajchem.2013.15109



QSPR Study on the Binding Constants of Coumarins and Human Serum Albumin

MEI-PING LI[1,2] and SHENG-WAN ZHANG[1,*]

[1]College of Life Science, Shanxi University, Taiyuan 030006, Shanxi Province, P.R. China
[2]College of Chemistry and Chemical Engineering, Shanxi University, Taiyuan 030006, Shanxi Province, P.R. China

*Corresponding author: E-mail: zswan@sxu.edu.cn

(*Received*: 7 January 2013; *Accepted*: 7 October 2013) AJC-14245

Three-dimensional holographic vector of atomic interaction field (3D-HoVAIF) was used to describe the chemical structures of 20 coumarins and it was employed to the quantitative structure property relationship studies between the binding constants of coumarins and human serum albumin. Here the quantitative structure property relationship model was built by multiple linear regression and partial least square regression. The stability and prediction ability of the established model were strictly examined by leave-one-out cross-validation and external validation. Meanwhile, the model of built pointed out hydrophobic interaction is important especially the hydrophobic interaction between H atoms and sp^3-hybridized O atoms (sp^3O) is in favour of the binding constants. Furthermore, the satisfactory results showed that 3D-HoVAIF could preferably express the structure information of the coumarins.

Key Words: Coumarin, Binding constants, Three-dimensional holographic vector of atomic interaction field, QSPR.

INTRODUCTION

Human serum albumin (HSA) is the most abundant protein in plasma, which is a single peptide chain protein consisting 585 amino acid residues, molecular weight 66500. It is the major storage and transport protein for numerous endogenous and exogenous compounds and is also capable of binding an extraordinarily diverse range of metabolites, drugs and organic compounds. Coumarins drugs are commonly used drugs, many of which exhibit useful pharmaceutical activity, such as anti-bacterial, antithrombotic, anti-inflammatory and antitumor properties[1-4]. In order to make the drugs putting into the body pharmacodynamic effect, the drugs have to reach receptor sites through the storage and transport of plasma. The remarkable binding capacity of drugs and human serum albumin has a significant impact for pharmacodynamics and pharmaco-kinetics. While the binding constant is an important parameter measuring the binding strength, so, construction of quantitative correlation between the molecular structure and the binding constants of coumarins and human serum albumin has an important significance to research drug screening and prediction of protein interactions. This has stimulated a great deal of research on the nature of the drug binding capacity and sites. The quantitative structure-property relationship (QSPR) is a well-established technique for estimation of the physico-chemical properties of a compound based on the descriptors derived from the molecular structure. It is widely used in all kinds of fields,such as the prediction of gas chromatographic relative retention times, drug activity designing and screening, *etc*.[5,6]. To date, many researches about coumarins and protein interaction are mainly focused on experimental and theoretical research is relatively small[7-13]. In this paper, the relation between binding constants of coumarins drugs and human serum albumin molecules were investigated utilizing three-dimensional holographic vector of atomic interaction field (3D-HoVAIF) as the descriptor, with the help of multiple stepwise linear regression, the QSPR model on human serum albumin (HSA) was established. Meanwhile, the influence of molecular structure on the binding was also discussed. The model could predict the binding constants of coumarins and human serum albumin with satisfactory results. The paper provides a new way for the structures characterization of coumarins and derivative and determine which group will help the binding by virtue of their different parameters, the binding of coumarins to human serum albumin shows a relationship with electrostatic and hydrophobic interaction of coumarins molecules, which is useful for the design and screening of medicines.

EXPERIMENTAL

Here 20 coumarins are studied (Table-1), whose molecular structures and the binding constants taken from reference[4,13-16]. The binding constants of the 20 samples were gained in the mimic physiological condition at room temperature, together with a wide concentration range of 1.0×10^{-6}-5.0×10^{-4} mol L^{-1}.

TABLE-1
EXPERIMENT LOGARITHM VALUES OF THE BINDING CONSTANT FOR 20 COUMARINS

Compound	Exp.
7-Anilinocoumarin-4-acetic acid	5.348
7-N-Methylamino-4-methylcoumarin	4.477
7-Amino-4- methylcoumarin	4.00
7-N,N-Diethylamino-4 methylcoumarin	5.909
2,3,6,7-Terahydro-9-trifuoromethyl-1*H*,5*H*,1*H*-[1]benzpyanol[6,7,8-ij]quinolizin-11-one	6.204
7-N-Methylamino-4-trifluoromethylcoumarin	6.447
7-N,N-Diethylamino-4-trifluoromethylcoumarin	6.833
7-N,N-Diethylaminocoumarin	5.491
3-(2'-N-Methylbenzimadazolyl)-7-N,N-diethylaminocoumarin	5.398
7-Butylaminocoumarin-4-acetic acid	5.19
7-Pentylaminocoumarin-4-acetic acid	5.677
7-Hexylaminocoumarin-4-acetic acid	5.716
7-Heptylaminocoumarin-4-acetic acid	6.25
7-Octylaminocoumarin-4-acetic acid	6.449
7-Nonylaminocoumarin-4-acetic acid	6.767
7-Decylaminocoumarin-4-acetic acid	7.155
7-Laurylaminocoumarin-4-acetic acid	7.507
7-Tridecylaminocoumarin-4-acetic acid	7.32
Isofraxidin	5.614
Daphnetin	4.935

Structure characterization: 3D-HoVAIF is proposed based upon 2D structural descriptor developed by Liu *et al.*[17-20]. Proceeding from two spatial invariants, namely atom relative distance and atomic properties on the basis of three common non-bonded (electrostatic, van der and hydrophobic) interaction which are directly associated with bioactivities, 3D-HoVAIF method derives multidimensional vectors to represent molecular steric structural characteristics. According to the defination of 3D-HoVAIF, there are 55 interaction (Table-2) among the 10 atomic types. The binding between the drug molecule and receptor are usually realized through non-covalent bond effect. The different interactions are expressed with three potential energy-static, spatial and hydrophobic, respectively. thus, there will produce $3 \times 55 = 165$ pairs of atom interaction items that are used to characterize the molecular structure information of one organic compound. Among them, V_1-V_{55}, V_{56}-V_{110} and V_{111}-V_{165} correspond to electrostatic, steric and hydrophobic properties in turn.

Three-dimensional molecular structures of the 20 coumarins compounds are automatically generated by software Chemoffice 8.0 and then semi-empirical quantum chemistry software MOPAC6.0 contained in Chem3D is used to obtain final optimized molecular structures at AM1 levels (cut-off value of 0.001 kJ/mol). Simultaneously, atomic partial charges are calculated by Mulliken Method in the form of single-point. Spatial positions for all atoms in a molecule and the atomic charges are put into C program Super-3D.EXE, giving rise to HoVAIF descriptors by taking forms of Cartesian coordinates and partial charges, respectively.

RESULTS AND DISCUSSION

In the present investigation, multiple linear regression method and partial least square analysis were used with leave-one-out cross validation for building the regression model.

Multiple linear regression analysis: Multiple linear regression (MLR)is one of the most used modeling methods in QSPR. In the present study, multiple stepwise linear regression is implemented by SPSS Windows statistical package, version 10.0. In MLR analysis, the number of compounds in sample should be at least five times greater than the numbers of independent. 3 descriptors (including V_{27}, V_{112} and V_{118}) were obtained by the SPSS Windows statistical package, version 10.0. The best equation is selected on the basis of the highest multiple correlation coefficient R^2 and R^2_{CV}. Moreover, a linear regression analysis reveals a fair correlation between experimental binding constant and the structure of coumarins. The 3D-QSPR regression model with 3 variables has good estimation capacity ($R^2 = 0.898$, SD = 0.324) and the best predictive ability ($R^2_{CV} = 0.851$, $SD_{CV} = 0.392$), which is given below:

$$\log K = 0.916V_{112} + 0.584V_{27} + 0.233\ V_{118} \quad (1)$$

$N = 20$, $R^2 = 0.898$ $R^2_{adj} = 0.879$ SD = 0.324 $R^2_{CV} = 0.851$ $SD_{CV} = 0.392$.

According to commonly recognition statistical standard, reliable model about QSAR is $q^2 \geq 0.5$[21]. Therefore, the present model is indeed excellent with a predictive ability of 85.1 %. In eqn. 1 V_{27} is the electrostatic interaction between the third type of atoms (sp^2C) and the tenth type of atoms (F), V_{112} is the hydrophobic interaction between the first type of atoms(H) and the second type of atoms (sp^3C), V_{118} is the hydrophobic interaction between the first type of atoms(H) and the eighth type of atoms (sp^3O). From QSPR model, it can be concluded that hydrophobic interaction is more important than electrostatic interaction. Moreover, it can be seen that the log K of binding constant is high positively correlated with hydrophobic and electrostatic interaction. Especially the hydrophobic interaction between H atoms and sp^3-hybridized C atoms (sp^3C), the addition of alkyl group is helpful to improve binding constant, but alkyl chain should not be too long, because the

TABLE-2
TEN KINDS OF ATOMS AND THEIR FIFTY-FIVE INTERACTION ITEMS IN 3D-HOVAIF

No.	Atom types	H	$C_{(sp3)}$	$C_{(sp2)}$	$C_{(sp)}$	$N/P_{(sp3)}$	$N/P_{(sp2)}$	$N/P_{(sp)}$	$O/S_{(sp3)}$	$O/S_{(sp2)}$	F, Cl, Br, I
1	H	1-1	1-2	1-3	1-4	1-5	1-6	1-7	1-8	1-9	1-10
2	$C_{(sp3)}$	–	2-2	2-3	2-4	2-5	2-6	2-7	2-8	2-9	2-10
3	$C_{(sp2)}$	–	–	3-3	3-4	3-5	3-6	3-7	3-8	3-9	3-10
4	$C_{(sp)}$	–	–	–	4-4	4-5	4-6	4-7	4-8	4-9	4-10
5	$N/P_{(sp3)}$	–	–	–	–	5-5	5-6	5-7	5-8	5-	5-10
6	$N/P_{(sp2)}$	–	–	–	–	–	6-6	6-7	6-8	6-9	6-10
7	$N/P_{(sp)}$	–	–	–	–	–	–	7-7	7-8	7-9	7-10
8	$O/S_{(sp3)}$	–	–	–	–	–	–	–	8-8	8-9	8-10
9	$O/S_{(sp2)}$	–	–	–	–	–	–	–	–	9-9	9-10
10	F, Cl, Br, I	–	–	–	–	–	–	–	–	–	10-10

big bulkness of alkyl chain will cause steric hindrance. The hydrophobic interaction between H atoms and sp^3-hybridized O atoms (sp^3O) is in favour of the binding constants. On the other hand, the electrostatic interaction between sp^2-hybridized C atoms (sp^2C) and halogen atoms (F) is also propitious to binding. For the compounds, the influence mainly reflected in the fourth substituent group. The larger electronegativity substituent group can reduces the electron density between the third and fourth C atoms, resulting in the higher molecular conjugation and lower energy and make the binding be tight. Fig. 1 showed plots of calculated binding constant (log K) against experimental log K values of all samples. From Fig. 1, it can be seen that almost all samples are uniformly distributed around diagonal, not obviously exceptional point has selected.

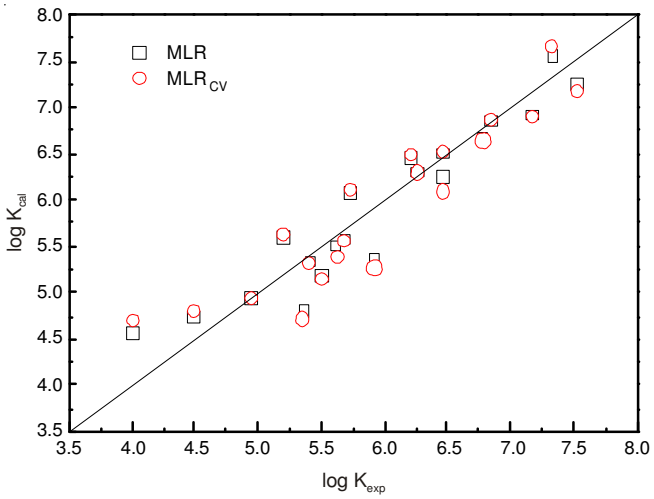


Fig. 1. Plot of log K calculated *vs*. experimental value by MLR model

Partial least square model: Partial least square (PLS) is a novel multivariate data analysis method, can solve many important problems which can not be solved by common multiple regression analysis method. Such as avoiding the harmful effects of multicollinearity and being capable of building the models when the number of observations is less than the number of variables,etc.In this work, PLS was performed in Simca-P 10.0 and the PLS latent variable number for each original variable matrix in PLS was determined by default standards. The selected three variables by SMR were then modeled by PLS, which the two principal components explained 89.1% and 77.2 % variance of Y and cross-validation variance of Y, respectively. Fig. 2 presents different loading contributions of 3 variables to the first two principal components. V_{27} and V_{112} had prominent contributions to PC1 loadings well correlative to Y variables, while V_{27} had prominent contributions to PC2. Meanwhile, dependent variable is further away from PC1 than PC2, this indicated that PC1 had prominent contributions to dependent variable. Besides, variable importance of projection (VIP) index of original variables is presented in Fig. 3. The most contributive top V_{112} indicated an intimate relationship between log K and hydrophobic interaction. The plot of log K calculated by PLS model *versus* those experimental is shown in Fig. 4.

Finally, in order to further verification the reliability of 3D-HoVAIF model, the whole data set was divided into two

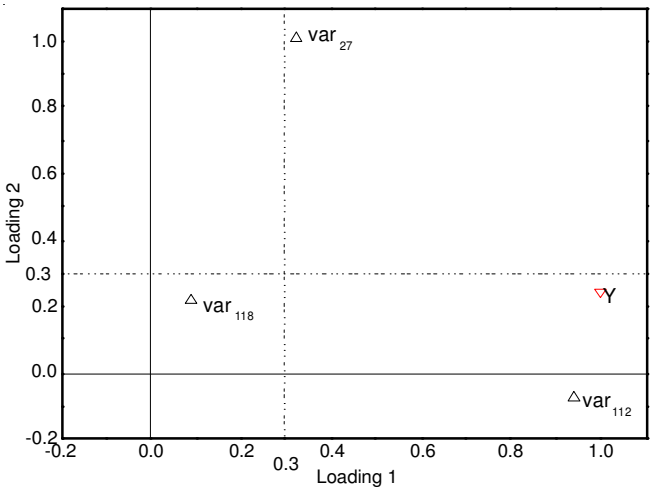


Fig. 2. Loading contributions of original variables to the top two principal components

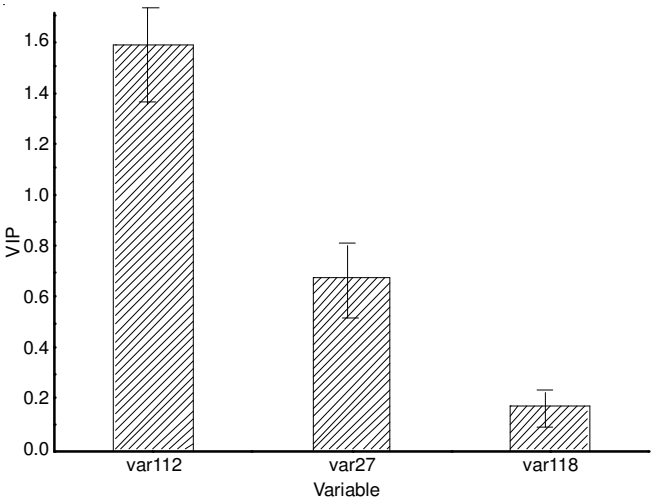


Fig. 3. Plot for variable importance of projection

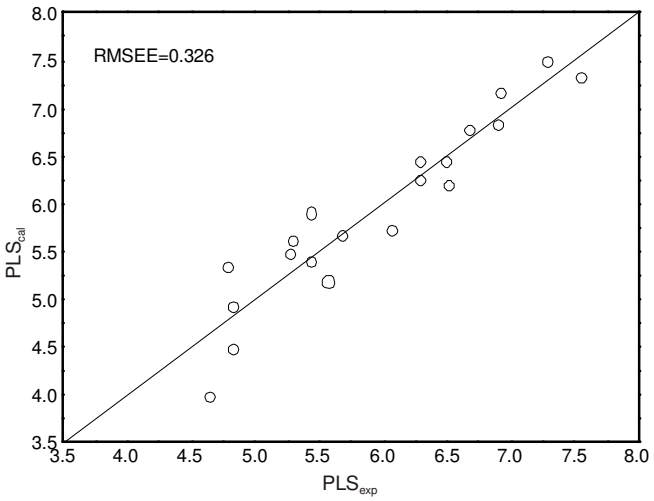


Fig. 4. Plot of log K calculated *vs*. experimental value by PLS model

subsets. In Table-1, the binding constants were rearranged from large to small. The tenth and twentieth sample were chosen to create a test set and the remaining 18 samples were regarded as the training set. Then, 18 training samples were utilized to construct the QSPR model. The other 2 samples were utilized to validate the external prediction power of the model developed.

Consequently, a 3-variable MLR model was constructed for the training set with its R^2cum = 0.894, Q^2_{LOO} = 0.841 and RMSE 0.324. The constructed model was then utilized to predict the test set, with the result of Q^2ext = 0.919. All of these depict the model has favorable simulative for internal samples and predictive ability for external samples.

Conclusion

In this paper, the descriptor-3D-HoVAIF derived solely from chemical structures and easily obtained,involving classical electric, steric and hydrophobic interactions. Furthermore, the obtained model with obvious physicochemical meaning. From QSPR model, it can be concluded that hydrophobic interaction is more important than electrostatic interaction. Moreover, it can be seen that the log K of binding constant is high positively correlated with hydrophobic and electrostatic interaction. In other words, hydrophobic groups will promote the binding constants. Thus it is suggested the 3D-HoVAIF descriptor behaves quite well in structural characterization ability and will be extremely useful in QSAR studies. Mean while, these descriptors have been applied in the QSAR/QSPR studies on many complicated molecular systems,such as steroids and anti HIVdrugs, *etc.*[22-25].

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (21277085) and Features Key Disciplines of Shanxi Province.

REFERENCES

1. U. Kragh-Hansen, *Pharmacol. Rev*., **33**, 17 (1981).
2. A. Dugaiczyk, S.W. Law and O.E. Dennison, *Proc. Natl. Acad. Sci. USA*, **79**, 71 (1982).
3. A.A. El-Bassuony, *Asian J. Chem.*, **19**, 2053 (1999).
4. M.Q. Liu, J.N. Tian, X. Tian, Z.D. Hu and X.G. Chen, *Bioorg. Med. Chem*., **12**, 469 (2004).
5. Z.S. Yi, L.C. Li, A.Q. Zhang and L.S. Wang, *Chin. J. Chem*., **29**, 2495 (2011).
6. S. Zhang, Y. Fan, Z. Shi and S. Cheng, *Chin. J. Chem*., **29**, 623 (2011).
7. J. Shobini, A.K. Mishra, K. Sandhya and N. Chandra, *Spectrochim. Acta A*, **57**, 1133 (2001).
8 M. Markuszewski and R. Kaliszan, *J. Chromatogr B*, **768**, 55 (2002).
9. V. Andrisano, C. Bertucci, V. Cavrini, M. Recanatini, A. Cavalli, L. Varoli, G. Felix and I.W. Wainer, J. *Chromatogr. A*, **876**, 75 (2000).
10. G. Colmenarejo, *Med. Res. Rev*., **23**, 275 (2003).
11. Q.J. Jiang, W.H. Li and H.P. Zeng, *J. Hubei Univ*., **27**, 373 (2005).
12. X.F. Liu, Y.M. Xia, Y.Z. Fang and L. Lu, *Acta Chim. Sinica*, **62**, 1484 (2004).
13. J.Q. Liu, H.L. Zhai, J.Y. Hang, J.N. Tian and Z.D. Hu, *Chinese J. Spectrosc. Lab.*, **23**, 602 (2006).
14. S. Goya, A. Takadate, H. Fujino, M. Otagiri and K. Uekama, *Chem. Pharm. Bull. (Tokyo)*, **30**, 1363 (1982).
15. M. Irikura, A. Takadate, S. Goya and M. Otagiri, *Chem. Pharm. Bull. (Tokyo)*, **39**, 724 (1991).
16. J. Liu, J. Tian, Y. Li, X. Yao, Z. Hu and X. Chen, *Macromol. Biosci*., **4**, 520 (2004).
17. S.S. Liu, C.Z. Cao and Z.L. Li, *J. Chem. Inf. Comput. Sci*., **38**, 387 (1998).
18. S.S. Liu, S.X. Cai, C.Z. Cao and Z.L. Li, *J. Chem. Inf. Comput. Sci*., **40**, 1337 (2000).
19. S.S. Liu, S.H. Cui and L.S. Wang, *Chin. Chem. Lett*., **15**, 467 (2004).
20. P. Zhou, F.F. Tian and Z.L. Li, *Chemom. Intell. Lab. Syst*., **87**, 88 (2007).
21. A. Golbraikh and A. Tropsha, *J. Mol. Graph. Model*., **20**, 269 (2002).
22. H.-Y. Xu, J.-Y. Zhang, J.-W. Zou and X.-S. Chen, *J. Mol. Graph. Model*., **26**, 1076 (2008).
23. J.B. Tong, P. Zhou, S.W. Zhang, Y. Zhou, H. Mei, H. Zeng, M.P. Li and Z.L. Li, *Chin. Sci. Bull*., **51**, 1557 (2006).
24. J.B. Tong, Y.F. Li, S.L. Liu and Y.L. Men, *Chinese J. Struct. Chem.*, **29**, 1893 (2010).
25. J.Y. Sun, S.X. Cai, N. Yan and H. Mei, *Eur. J. Med. Chem*., **45**, 1008 (2010).