Theoretical Studies on the Inhibitory Activity of Levofloxacin-thiadiazole HDACi Conjugates to Histone Deacetylases

Zi-qiang TANG[1,2], Chang-ning LIU[1], Yu-sheng LIU[1], Hui FENG[1] and Chang-jun FENG[1,a,*]

[1]School of Chemistry & Chemical Engineering, Xuzhou University of Technology, Xuzhou, Jiangsu 221111, China

[2]Xuzhou Technical Institute, Xuzhou, Jiangsu 221151, China

[a]fengcj@xzit.edu.cn

*Corresponding author

Keywords: Levofloxacin-thiadiazole HDACi conjugates(LHCc), Histone deacetylases (HDACs), Inhibitory activity (pW_i, i=1, 6), Electrotopological state indices (E_K) of atom types, Quantitative structure-activity relationship (QSAR)

Abstract. Electrotopological state indices (E_K) of atom types is used to describe the structures of 19 conjugates (LHCc) of levofloxacin-thiadiazole HDAC inhibitor (HDACi) and related to the inhibitory activity (pW_i, i=1, 6) of LHCc against histone deacetylases(HDACs, HDAC1 and HDAC6). The optimal two-parameter(E_{17}, E_{18}) quantitative structure-activity relationships(QSAR) models are established by using leaps-and-bounds regression analysis for the inhibitory activities(pW_i) of 19 above compounds to HDAC1 and HDAC6 along with the E_K. The correlation coefficients (R^2) and the leave-one-out (LOO) cross validation R_{cv}^2 for the pW_1 and pW_6 models were 0.942 and 0.918, 0.943 and 0.919, respectively. The QSAR models have favorable correlation, as well as robustness and good prediction capability by R^2, F, R_{cv}^2, A_{IC}, F_{IT}, V_{IF} tests. The results indicate that the molecular structural units: $-NH_2$, $-NH-$ are main factors which can affect the inhibitory activity(pW_i)bioactivities of these compounds directly. Accordingly, it is suggested that the main interactions between HDACs inhibitor and HDACs are hydrophobic, hydrogen bond, and coordination compounds with Zn^{2+}. This is consistent with the molecular docking results of document.

Introduction

Quantitative structure-activity relationship (QSAR) [1-4] can establish the quantitative relationship between the molecular structure and activity of organic compounds, which shows practical significance for predicting the biological activity of new compounds, initial screening and evaluation of activity compounds, and reducing high cost of the activity evaluation. Topological indexes could directly derive from chemical molecular structure independent of the experimental mensuration. Therefore, the appearance of topological indexes in general QSAR models was popular for the development of reliable QSAR models. The electrotopological state (E-state) indexes (E_K) put forward by by Kier and Hall [5-7] is one of the most broadly applied topological indexes.

Histone deacetylases (HDACs) are key enzyme families that regulate gene transcription and expression. HDACs are a new target of anti-tumor research in recent years, and HDACi has become a hot research topic of anti-tumor drugs both at home

and abroad. Eighteen kinds of new, efficient and low toxicity of levofloxacin-thiadiazole HDACi conjugates (referred to as LHCc) were designed and synthesized from levofloxacin by Li Hui and other [8]. The results showed that these new conjugates displayed potent inhibitory activity against HDAC1 and HDAC6. Specifically, conjugate 10 exhibited the most activities, which was more potent than SAHA (Vorinostat). Therefore, it is important to study the levofloxacin-HDACi conjugate by E_K and QSAR method.

Based on the electrotopological state indexes (E_K) [5-7], the robust QSAR model of LHCc inhibitory activity [8] was established by leaps-and-bounds regression, to estimate and predict the inhibitory activity of LHCc and reveal the microstructure that affected its inhibitory activity at the molecular level. It provides theoretical reference a reasonable design and screening of novel and highly effective lead compounds for LHCc.

Material and Methods

Studied Compounds and Their Biological Activity Data

The compounds studied herein are a series of HDACs inhibitors with HDAC1 and HDAC6 inhibitory activity. Their bioactivity data for HDACs were IC_{50}, and drug concentrations resulted in 50% inhibition of HDACs in units of $\mu mol \cdot dm^{-3}$. The inhibitory activity (IC_{50}) [8] of they to HDAC1 and HDAC6 are expressed as W_1 and W_6, respectively.

Molecular Descriptors

In this paper, electrotopological state indices (E_K) [5-7] are used to characterize the molecular structure of different classes of organic compounds. Hall *et al.* have investigated the localization of several famous topologies, and provided E_k which can show the topological, geometrical and electrical characters comprehensively. The E_k of all atoms in the compound molecular skeleton was calculated following the method given by Hall *et al*, and the details can be found in references [5-7].

Statistical Regression Analysis

For QSAR derivation, IC_{50} values [8] act as the dependent variables, and electrotopological state indices descriptors act as the independent variables. The regression analyses are carried out by using multiple linear regressions (MLR), partial least squares (PLS), leaps-and-bounds regression (LBR) program. By using Fischer statistics (F-tests) to eliminate insignificant descriptor when entering a new descriptor, stepwise regression can identify the most important descriptors contributing to the anti-tumor activities of the title compounds.

The correlation between variables in model was estimated by the variance inflation factor (V_{IF}) [9] value. is defined as follows:

$$V_{IF} = 1/(1 - \beta^2) \tag{1}$$

in which β is the correlation coefficient of multiple regressions between one variable and the others in the equation. $V_{IF} = 1.0$ suggests no self-correlation among each variable; if V_{IF} ranges from 1.0 to 5.0, indicating that there is no obvious autocorrelation between variables, the model is stable; when V_{IF} is larger than 5.0, the regression equation is unstable and recheck of variables' correlation coefficient is necessary.

In the QSAR equations, the quality of each regression model was evaluated; using a

squared correlation coefficient (R^2), cross validation squared correlation coefficient ($R_{cv}{}^2$). The "leave-one-out" (LOO) cross-validation coefficient $R_{cv}{}^2$ was considered as an indicator of the predictive performance and stability of a QSAR model. As a rule of thumb, the equations with regression coefficients $R^2 > 0.80$[10] and $R_{cv}{}^2 > 0.50$[11] are considered reasonable. Where parameter $R_{cv}{}^2 > 0.50$ is used as a criterion of both robustness and predictive ability of the model, whereas the conventional correlation coefficient R^2 defines the goodness-of-fit.

Another indicator used to evaluate the quality of model is the standard deviation (S_D). The model is good and the prediction accuracy is acceptable when the ratio of the standard deviation to the value range (the samples between the maximum and minimum values) is less than 10% [12].

The statistic significance of the model was validated by t-tests. If the absolute values of t for all the variables in the validated model are larger than the standard t-value ($t_{\alpha/2}$) at one confidence level α, it will suggests that the model passes t-tests and has obvious statistic significance. We also applied the Akaike's information criterion (A_{IC}; Eq.2; the model that produces the minimum A_{IC} value is considered potentially the most useful) and Kubinyi function (F_{IT}; Eq.3; the best model will present the highest value of this function) [13, 14] to determine if a variable should be included in the model. That is to say, if the Akaike's information criterion decreases in value when adding an additional variable and the Kubinyi function increases in value, then, the introduction of this new variable is justified.

$$A_{IC} = RSS \times \frac{f+b}{(f-b)^2} \tag{2}$$

$$F_{IT} = \frac{R^2(f-b-1)}{(f+b^2)(1-R^2)} \tag{3}$$

Where RSS is the residual sum of squares, f is the number of compounds included in the model, b is the number of variables included in the model, R^2 is the square of the correlation coefficient.

Results and Discussion

According to the balance principle of physical chemistry, relationship between free energy change and the concentration is a logarithmic relationship Therefore, the inhibitory activity (W_i) [8] of HDACs inhibitors to HDAC1 and HDAC6 is taken negative logarithm ($pW_i = -\log W_i$) for modeling.

QSAR Equation of the pW_i

The electrotopological state indices and inhibitory activities(pW_i) of HDACs inhibitors were input into MINITAB14.0 statistical analysis software, and the leaps-and-bounds regression was used to select the best variable combinations, to establish the best QSAR models, the results shown in Table1,2. Where R^2 is the square of the correlation coefficient, $R_{adj}{}^2$ is the square of the adjusted correlation coefficient, $R_{cv}{}^2$ is the LOO cross-validation correlation coefficient, S_D is the standard deviation of the regression, S_{CV} is the standard deviation of the regression of LOO, and F is the Fisher ratio. A_{IC} is the Akaike's information criterion and F_{IT} is the Kubinyi function. Table 1 shows the QSAR models for pW_1 and E_k.

Table 1 Results of electrotopological state indices descriptors and inhibitory activities pW_1 with leaps-and-bounds regression

No.	R^2	R_{adj}^2	R_{cv}^2	AIC	FIT	S_D	S_{CV}	F	Variables
1	0.672	0.6524	0.6035	0.250	1.741	0.444	0.488	34.785	E_{18}
2	0.942	0.9346	**0.9184**	**0.061**	**11.298**	0.193	**0.228**	**129.589**	E_{18}, E_{17}
3	0.946	0.9353	0.9179	0.072	9.385	0.191	0.236	87.712	E_{18}, E_{17}, E_8

Table 1 shows that with the increase in the number of independent variables in the model, in addition to R^2, R_{adj}^2, S_D, the rest of the statistical indicators in the ternary equation has a turning point, which R_{cv}^2, F_{IT}, F has the maximum, A_{IC}, S_{CV} has a minimum. So choose the best binary QSAR model:

$$pW_1 = 2.898(\pm 0.241) - 0.239(\pm 0.028)\ E_{17} + 0.988(\pm 0.066)\ E_{18} \tag{4}$$

$f = 19$, $R^2 = 0.942$, $R_{adj}^2 = 0.935$, $F = 129.589$, $S_D = 0.193$, $R_{cv}^2 = 0.9184$, $S_{CV} = 0.228$

Table 2 shows the QSAR models for inhibitory activities pW_6 and E_k.

Table 2 Results of E_t and pW_6 with leaps-and-bounds regression

No.	R^2	R_{adj}^2	R_{cv}^2	AIC	FIT	S_D	S_{CV}	F	Variables
1	0.655	0.6344	0.5840	0.287	1.614	0.4721	0.519	32.240	E_{18}
2	0.943	**0.9355**	**0.9195**	**0.064**	**11.509**	**0.1983**	**0.235**	**131.512**	E_{18}, E_{17}
3	0.946	0.9354	0.9165	0.079	9.385	0.1984	0.247	87.943	E_{18}, E_{17}, E_8

Table 2 shows that with the increase in the number of independent variables in the model, in addition to R^2, the rest of the statistical indicators in the ternary equation has a turning point, which R_{adj}^2, R_{cv}^2, F_{IT}, F has the maximum, A_{IC}, S_{CV} S_D, has a minimum. So choose the best binary QSAR mode:

$$pW_6 = 2.849(\pm 0.249) - 0.256(\pm 0.029)\ E_{17} + 1.033(\pm 0.068)\ E_{18} \tag{5}$$

$f = 19$, $R^2 = 0.943$, $R_{adj}^2 = 0.9355$, $F = 131.512$, $S_D = 0.1983$, $R_{cv}^2 = 0.9195$, $S_{CV} = 0.235$

Using the QSAR equations (4) and (5), we can predict the inhibitory activities pW_1 and pW_6, respectively.

Validation of the QSAR Equation

The predictive capability of a QSAR model should be tested through model validation. Cross-validation is one of the most often used model validation methods. The R_{cv}^2 values of models(4) and (5) are 0.9184 and 0.9195, respectively, which are well above 0.5, indicating that the model has good robustness and prediction ability.

The V_{IF} values of the variables in model (4), (5) are less than 5.0, indicating that all models are statistically significant and have good stability.

The anti-tumor activities (pW_1) range of the samples between the maximum and minimum values is 2.30(7.51 - 5.21 = 2.30). The ratio of the standard deviation ($S_D = 0.193$) to 2.30 is 8.39%. The standard deviation ratios for pW_6 are 8.00%. They are less than 10%, indicating that the model has acceptable predictive accuracy.

Analysis of the QSAR Equation

Li Hui et al [8] chose the molecular docking of the strongest inhibitory activity between 10 conjugate with of HDAC1 and HDAC6. The results of docking were as follows: the hydrophobic interaction, the hydrogen bond and the coordination compounds formed with Zn^{2+} between the conjugate 10 and HDAC1 and HDAC6.

According to the theory of electrotopological state indices, it can be seen that E_{17} in the models reflects amino nitrogen atom ($-NH_2$), E18 in the models reflects imido

nitrogen atom($-NH-$). Wherein $-NH_2$ and $-NH-$ are highly electronegative polar groups that are capable of forming hydrogen bonds, and can form the coordination compounds formed with Zn^{2+} metal ions in enzyme. The parent contains a benzene ring, which is hydrophobic with the enzyme. This is consistent with the molecular docking results of Document.

In addition, determination coefficient R^2 is also called the reduction error ratio. $R^2 = 0.942$ in model (4), indicating that E_{17}, E_{18} and constant items together show 94.2% of the factors affecting inhibitory activities (pW_1) of HDACs inhibitors to HDAC1, and only 5.8% are random factors; $R^2 = 0.943$ in model (5), indicating that E_{17}, E_{18} and constant items together show 94.3% of the factors affecting inhibitory activities (pW_6) of HDACs inhibitors to HDAC6, and only 5.7% are random factors; which further prove the correctness of the models.

Conclusion

(1) The optimal binary QSAR model of inhibitory activity pW_1 and pW_6 of HDACi to HDACs was constructed by using the leaps-and-bounds regression method. The QSAR models have showed good correlation, as well as robustness and prediction ability by statistical indicators: R^2, R_{cv}^2, S_R, V_{IF}, F_{IT} and A_{IC} tests. To successfully establish QSAR models of two inhibitory activities is rare in QSAR studies by using a class of topological indices.

(2) According to the electrotopological state indices entering the two models, the main molecular structural units that affect their inhibitory activity are $-NH_2$ and $-NH-$.

(3) According to the main molecular structure units that affect inhibitory activity, the interactions between HDACs inhibitor molecules and HDACs are mainly hydrophobic, hydrogen bond and coordination complex with Zn^{2+}. This is consistent with the molecular docking results of Document.

In summary, this study provides theoretical guidance for the further design of novel and efficient HDACs inhibitors.

Acknowledgement

References

[1] Feng C.J., Yang W.H, Linear QSAR regression models for the prediction of bioconcentration factors of chloroanilines in fish by density functional theory. Chin. J. Struct. Chem., 33(2014)830-834.

[2] Wang C., Feng C. J, QSAR study of the action strength of D_{OM} of phenyl-isopropyl-amine dopes using MLR and BP-ANN. Chin. J. Struct. Chem., 36(2017)1720-1728.
[3] Liu B.G., LIU Jiang W., Li J.Q., Geng S., Mo H.Z., Liang G.Z, 3D-QSAR and interaction mechanism of flavonoids as P-glycoprotein inhibitors. Chem. J. Chinese Universities, 38(2017) 41-46

[4] Liao L.M., Li J.F., Lei G...D, Structural characterization and Kovats retention indices (RI) prediction for alkylbenzene compounds. J. Struct. Chem., 57(2016)1627-

1634.

[5] Hall L.H., Kier L.B., Brown B.B, Molecular similarity based on novel atom-type electrotopological state indices. J. Chem. Inf. Comput. Sci., 35(1995)1074-1080.

[6] Hall L.H., Kier L.B., Electrotopological state indices for atom types: A novel combination of electronic, topological, and valence state information. J. Chem. Inf. Comput. Sci., 35(1995)1039-1045.

[7] Hall L.H., Mohney B., Kier L.B., The electrotopological state: structure information at the atomic level for molecular graphs. J. Chem. Inf. Comput. Sci., 31(1991) 76-82.

[8] Li H., Han X., Li D.W., Synthesis and anti-tumor activity of levofloxacin-thiadia-zole histone deacetylase inhibitor conjugates. Acta Pharmaceutica Sinica, 52(2017) 582-591.

[9] Wei D.B., Zhang A.Q., Wu C.D., Han S.K., Wang L.S., Progressive study and robustness test of QSAR model based on quantum chemical parameters for predicting BCF of selected polychlorinated organic compounds. Chemosphere, 44(2001) 1421-1428.

[10] Yao S.W., Lopes V.H., Fernández F., Garcı́a-Mera X., Morales M., Rodrı́guez-Borges J. E., Cordeiro D. S., Synthesis and QSAR study of the anticancer activity of some novel indane carbocyclic nucleosides. Bioorg. Med. Chem., 11(2003)4999-5006.

[11] Xu X., Li X. G., QSAR for predicting biodegradation rates of polycyclic aromatic hydrocarbons in aqueous systems. Chin. J. Struct. Chem., 31(2012) 1212-1221.

[12] Sung-Sau S., Kaprlus M., A comparative study of ligand-receptor complex binding affinity prediction methods based on glycogen phosphorylase inhibitors. Journal of Computer - Aided Molecular Design, 13(1999)243-258.

[13] Urra L.S., Gonza′lez M.P., Teijeira M., 2D-autocorrelation descriptors for predicting cytotoxicity of naphthoquinone ester derivatives against oral human epidermoid carcinoma. Bioorg. Med. Chem., 15(2007)3565-3571.

[14] Urra L.S., Gonza′lez M.P., Teijeira M., QSAR studies about cytotoxicity of benzophenazines with dual inhibition toward both topoisomerases I and II: 3D-MoRSE descriptors and statistical considerations about variable selection. Bioorg. Med. Chem., 14(2006)7347-7358