PREDICTION OF PHYSICOCHEMICAL PROPERTIES AND ANTICANCER ACTIVITY OF SIMILAR STRUCTURES OF FLAVONES AND ISOFLAVONES

Bui Thi Phuong Thuy[(1)], Pham Van Tat[(2)], Le Thi Dao[(3)]

(1) University of Hue Science, (2) Industrial University of Ho Chi Minh City, (3) Thu Dau Mot University

ABSTRACT

The reliability of Quantitative Structure – Activity or Property Relationships for prediction of physico-chemical properties and anticancer activity of flavone and isoflavone derivatives was improved by using the quantitative relationships between structurally similar flavone and isoflavone structures (QSSRs). The targeted-compound method was developed by a training set, which contains only similar compounds structurally to target compound. The structural similarity is presented by multidimensional correlation between the dimensions of atomic-charge descriptors of target compound and those of predictive compounds with $R^2_{fitness}$ = 0.9999 and R^2_{test} = 0.9999. The available physicochemical properties and anticancer activities of predictive substances in training set were used in the usual manner for predicting the unknown physicochemical properties and anticancer activity of target substances. Preliminary results show that the targeted - compound method yields the predictive results within the uncertain extent of experimental measurements.

Keywords: *QSSR models; physicochemical property; anticancer activity.*

*

1. Introduction

Physicochemical properties and biological activity of pure substances deriving from experimental measurements are serviceable only for a small portion referring to chemistry and pharmaceutical engineering and environmental impact assessment [[1],[2]]. Consequently, the development of targeted-compound method for accurately prediction of physicochemical property and biological activity are very necessary. In particular, the physicochemical properties for instance the boiling and critical temperature are very important for chemical industrial techni-ques. In recent years, the use of quantitative structure property relationships (QSPRs) has been interesting for using structural descriptors to predict the several physico-chemical properties.

One of the last attempts Dearden proposed a QSPR model for predicting vapour pressure [[1]]. The models QSPR were developed recently by Shacham et al. [[2]] and Cholakov et al. [[3],[4],[5]] for prediction of tem-perature-dependent properties. The linear structure - structure relationships were derived from the similar substances with QSPR model proposed by Schacham [[2]]. For a specified property of target substance, a structure-structure correlation has to be esta-blished by using the structural descriptors of predictive substances. The

molecular desc-riptors are resulted by quantum chemical calcu-lations. This suggested for the develop-ment of the structure-structure correlations for complex structures proposed by Cholakov et al. [[3]].

In this work, the quantitative structure – structure relatioships (QSSR) are developed for predicting the physicochemical properties and anticancer activity of similar flavones and isoflavones. The physicochemical properties and anticancer activities of target flavones and isoflavones resulting from multivariable linear regression techniques are compared with experimental data and those from reference data.

2. Methodology

2.1. Data and software

The physicochemistry properties selected are in Table 3 for pure flavones and isoflavones. Those are the major important properties for a pure substance. In this case, they are obtained from the empirical correlation equation of package ChemOffice [[9]]. The anticancer activity GI_{50} (μM) (drug molar concentration causing 50% cell growth inhibition) of structurally similar flavones and isoflavones are taken from a source of Wang [[6],[7]], as given in Figure 1 and Table 1. The programs BMDP new system 2.0 [[8],[10][10]] are used for constructing multivariate linear regression models. The experimental structures of flavones and isoflavones, and the molecular descriptors as the atomiccharge descriptors are optimized and calculated by MM+ molecular mechanics and semiempirical quantum chemical calculations PM3 SCF using package HyperChem [[11]]. For convenient calculation the original anticancer activity values GI_{50} (μM) are transformed into negative logarithm of values GI_{50} (pGI_{50}) in this study.

2.2. Multiple linear modeling

For quantitative structure–structure rela-tionships (QSSR), the predictive substances (X) correlated with target substance (Y). This relationship is well represented by a model that is linear in regressed predictors as

$$Y = \sum_{i=1}^{k} b_i X_i + C \qquad (1)$$

Where parameters, b_i are unknown regression coefficients; C is constant.

Multiple linear regression analysis based on leastsquares procedure is very frequent used for estimating the regression coefficients. The multiple linear models QSSR were constructed by using programs BMDP and Regress [[8],[10]].

The QSSR models are constructed by using the linear regression. The goodness-of-fit quality of these was expressed as the fit R^2, respectively; the predictability of models was also validated by the test R^2:

$$R^2 = \left(1 - \frac{\sum_{i=1}^{N}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{N}(Y_i - \bar{Y})^2}\right)100 \qquad (2)$$

Where Y, $\bar{Y}$ and $\hat{Y}$ are the experimental, mean and predicted properties or anticancer activity of target substance.

Figure 1*. Molecular skeleton: a) flavone and b) isoflavone*

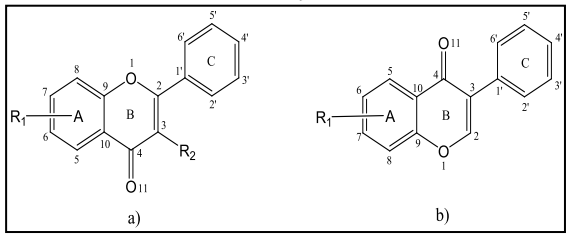


Table 1. *Anticancer activity pGI_{50} and experimental structure flavone and isoflavone [[6],[7]].*

Substance	Skeleton	Position	Substitutional group	pGI_{50}
fla-A_1	flavone	C_3-R_2	-OCH_2CCH_3=NOH	5.699
fla-A_2	flavone	C_6-R_1	-OCH_2CCH_3=NOH	5.921
fla-A_3	flavone	C_7-R_1	-OCH_2CCH_3=NOH	5.699
isofla-A_4	isoflavone	C_7-R_1	-OCH_2CCH_3=NOH	5.009
fla-A_5	flavone	C_3-R_2	-OCH_2CCH_3=$NOCH_3$	5.699
fla-A_6	flavone	C_3-R_2	-OCH_2CCH_3=$NOCH_3$	6.046
fla-A_7	flavone	C_7-R_1	-OCH_2CCH_3=$NOCH_3$	5.658
isofla-A_8	isoflavone	C_7-R_1	-OCH_2CCH_3=$NOCH_3$	5.071
fla-A_9	flavone	C_3-R_2	-$OCH_2CC_6H_5$=NOH	5.745
fla-A_{10}	flavone	C_3-R_2	-OCH_2C(4-F-C_6H_4)=NOH	5.678
fla-A_{11}	flavone	C_3-R_2	-OCH_2C(4-CH_3O-C_6H_4)=NOH	5.699
fla-A_{12}	flavone	C_6-R_1	-$OCH_2CC_6H_5$=NOH	6.097
fla-A_{13}	flavone	C_6-R_1	-OCH_2C(4-F-C_6H_4)=NOH	5.796
fla-A_{14}	flavone	C_6-R_1	-OCH_2C(4-CH_3O-C_6H_4)=NOH	6.000
fla-A_{15}	flavone	C_7-R_1	-$OCH_2CC_6H_5$=NOH	5.699
fla-A_{16}	flavone	C_7-R_1	-OCH_2C(4-F-C_6H_4)=NOH	5.699
fla-A_{17}	flavone	C_7-R_1	-OCH_2C(4-CH_3O-C_6H_4)=NOH	5.699
isofla-A_{18}	isoflavone	C_7-R_1	-OCH_2C(C_6H_5)=NOH	5.046
isofla-A_{19}	isoflavone	C_7-R_1	-OCH_2C(4-F-C_6H_4)=NOH	5.108
isofla-A_{20}	isoflavone	C_7-R_1	-OCH_2C(4-CH_3O-C_6H_4)=NOH	5.119
fla-A_{21}	flavone	C_3-R_2	-OCH_2C(C_6H_5)=$NOCH_3$	5.796
fla-A_{22}	flavone	C_3-R_2	-OCH_2C(4-F-C_6H_4)=$NOCH_3$	5.699
fla-A_{23}	flavone	C_3-R_2	-OCH_2C(4-CH_3O-C_6H_4)=$NOCH_3$	5.699
fla-A_{24}	flavone	C_6-R_1	-OCH_2C(C_6H_5)=$NOCH_3$	5.620
fla-A_{25}	flavone	C_6-R_1	-OCH_2C(4-F-C_6H_4)=$NOCH_3$	5.638
fla-A_{26}	flavone	C_6-R_1	-OCH_2C(4-CH_3O-C_6H_4)=$NOCH_3$	5.699
fla-A_{27}	flavone	C_7-R_1	-OCH_2C(C_6H_5)=$NOCH_3$	5.180
fla-A_{28}	flavone	C_7-R_1	-OCH_2C(4-F-C_6H_4)=$NOCH_3$	5.569
fla-A_{29}	flavone	C_7-R_1	-OCH_2C(4-CH_3O-C_6H_4)=$NOCH_3$	5.602
isofla-A_{30}	isoflavone	C_7-R_1	-OCH_2C(C_6H_5)=$NOCH_3$	5.086
isofla-A_{31}	isoflavone	C_7-R_1	-OCH_2C(4-F-C_6H_4)=$NOCH_3$	5.194
Isofla-A_{32}	isoflavone	C_7-R_1	-OCH_2C(4-CH_3O-C_6H_4)=$NOCH_3$	5.137

3. Results and discussion

3.1. Molecular modeling and atomic charge

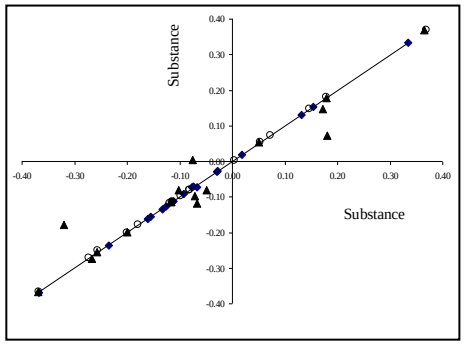
In order to calculate the atomic-charge descriptors, the experimental structures in Table 1 were optimized by MM+ molecular mechanics method at gradient level of 0.05 using HyperChem program [[11]]. After optimizing the molecular geometries of flavones and isoflavones the atomic charges of each structure were calculated by using semi-empirical quantum chemical calculation PM3 SCF in package HyperChem [[11]].

3.2. Building linear model

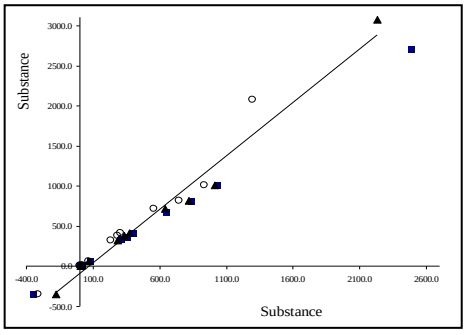
As a first step, the linear model QSSR was searched through exploring regression models, with the purpose of incorporating the representative predictive substances with target substance. The QSSR models in Table 2 including important predictive substances were founded by multivariate regression techniques. Furthermore, these are clear that predictive substances are able to lead to the best regression statistical parameters. The substance group is partly considered during the modeling construction.

The multivariate linear regression technique was used for constructing the linear relationship between the similar compounds structurally. These linear relationships were built by using the atomic-charge descriptors of predictive substances and those of target substance. All the atomic-charge descriptors consist of the atomic charges on atoms O_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8, C_9, C_{10}, O_{11}, $C_{1'}$, $C_{2'}$, $C_{3'}$, $C_{4'}$, $C_{5'}$ and $C_{6'}$. These aligned along a line with the correlation coefficient values for linear correlation between substances

using the atomic charges and physicochemical properties, as are shown in Figure 1.



a) Using atomic charges



b) Using physicochemical properties

Figure 2. *Correlation between substances symbol: ■: fla-A_{23} vs. fla-A_{11}; ▲: fla-A_{15} vs. isofla-A_{32}; ○: isofla-A_{32} vs. isofla-A_4.*

The predictive substances in Table 1 were selected randomly to evaluate the correlation magnitudes between substances. The correlation coefficients between the selected substances are given in Table 2. The similar substances structurally turn out to be a good correlation with each other. The linear regression models with the statistical parameters for target substances flavones and isoflavones were built from the atomic-charge descriptors [[8],[10]], as are given in Table 3. These linear QSSR models turn out to be in very good fit values $R^2_{fitness}$ = 0.9999 and R^2_{test} = 0.9999. The Table 3 shows that 10 models of 32 QSSR models resulting from 32 target substances in Table 1 represented for predictability of the quantitative relationships between flavones and isoflavones.

From the correlation coefficients between substances in Table 2, the similar substances structurally exhibited in higher correlation than others. Therefore, the construction of QSSR models based on the incorporation of predictive substances, as is depicted in equation (1). The correlation coefficients can be used to identify their important communion. Furthermore, the molecular structural descriptors of each substance have also to be considered prudentially to establish the QSSR models, as are exhibited in Figure 2.

Table 2: *Correlation of predictive substances using the atomic-charge descriptors*

	fla-A_{23}	fla-A_6	fla-A_{15}	fla-A_{22}	isofla-A_{32}	fla-A_{28}	fla-A_5	isofla-A_4
fla-A_{23}	1.0000							
fla-A_6	0.8664	1.0000						
fla-A_{15}	0.9220	0.8254	1.0000					
fla-A_{22}	0.9984	0.8548	0.9132	1.0000				
isofla-A_{32}	0.9247	0.7565	0.9659	0.9254	1.0000			
fla-A_{28}	0.9222	0.8259	1.0000	0.9134	0.9656	1.0000		
fla-A_5	0.9986	0.8696	0.9267	0.9983	0.9261	0.9270	1.0000	
isofla-A_4	0.9250	0.7560	0.9659	0.9257	1.0000	0.9657	0.9264	1.0000
fla-A_{11}	0.9999	0.8668	0.9225	0.9981	0.9236	0.9227	0.9986	0.9239

Table 3*. Physicochemical properties and anticancer activity pGI_{50} of target substances derived from QSSR models and predictive substances, respectively.*

Physicochemical properties and activity pGI_{50}	method		*ARE*%
	QSSR model	Ref. values [[6],[9]]	
QSSR model for flavone fla-A_1 with $R^2_{fitness}$ = 0.9999; R^2_{test} = 0.9999; SE = 0.00020159			

fla-A_1 = 0.00015 + 1.018 (fla-A_5) - 0.513 (fla-A_{21}) + 0.497 (fla-A_{22})			
Polar Surface Area	68.4533	68.1200	0.4893
pGI_{50}	5.699	5.663	0.638
QSSR model for flavone fla-A_2 with $R^2_{fitness}$ = 0.9999; R^2_{test} = 0.9999; SE = 0.00035399 fla-A_2 = -0.00020 + 1.260 (fla-A_6) + 0.871 (fla-A_{14}) - 1.134 (fla-A_{24})			
Melting point in K (T_m) at 1 atm	741.521	745.496	0.533
Critical temperature in K (T_C)	931.125	934.452	0.356
Mol. Refractivity	8.711	8.715	0.053
Boiling point in K (T_b) at 1 atm	978.789	980.510	0.176
pGI_{50}	5.921	6.473	9.321
QSSR model for flavone fla-A_3 with $R^2_{fitness}$ = 0.9999; R^2_{test} = 0.9999; SE = 0.00010411 fla-A_3 = 0.00002 + 0.935 (fla-A_7) + 0.582 (fla-A_{16}) - 0.517 (fla-A_{28})			
Melting point in K (T_m) at 1 atm	737.884	745.496	1.021
Critical temperature in K (T_C)	932.899	934.452	0.166
Heat of Formation in KJ/mol	-318.085	-313.160	1.573
Henry's Law constant	7.266	7.240	0.355
pGI_{50}	5.699	5.726	0.469
QSSR model for isoflavone isofla-A_4 with $R^2_{fitness}$ = 0.9999; R^2_{test} = 0.9999; SE = 0.00013747 isofla-A_4 = -0.000002 + 0.980 (isofla-A_8) - 0.233 (isofla-A_{18}) + 0.252 (isofla-A_{19})			
Melting point in K (T_m) at 1 atm	718.146	745.496	3.669
Critical temperature in K (T_C)	914.478	934.452	2.138
Henry's Law constant	7.237	7.240	0.042
pGI_{50}	5.009	5.0837	1.495
QSSR model for flavone fla-A_5 with $R^2_{fitness}$ = 0.9999; R^2_{test} = 0.9999; SE = 0.00019793 fla-A_5 = -0.00015 + 0.982 (fla-A_1) + 0.499 (fla-A_{21}) - 0.483 (fla-A_{22})			
Critical temperature in K (T_C)	936.289	913.478	2.497
Mol. Refractivity	8.731	9.179	4.884
Boiling point in K (T_b) at 1 atm	977.737	933.630	4.724
Henry's Law constant	7.034	7.110	1.073
logP	8.731	9.179	4.884
pGI_{50}	5.699	5.734	0.618
QSSR model for flavone fla-A_6 with $R^2_{fitness}$ = 0.9999; R^2_{test} = 0.9999; SE = 0.00026038 fla-A_6 = 0.00019 + 0.682 (fla-A_2) - 0.587 (fla-A_{14}) + 0.907 (fla-A_{24})			
Melting point in K (T_m) at 1 atm	730.455	717.167	1.853
Critical temperature in K (T_C)	927.997	914.743	1.449
Mol. Weigh	324.833	323.343	0.461
pGI_{50}	6.046	5.772	4.533
QSSR model for flavone fla-A_7 with $R^2_{fitness}$ = 0.9999; R^2_{test} = 0.9999; SE = 0.00013549 fla-A_7 = -0.00003+1.037 (fla-A_3) - 0.041 (fla-A_{16}) + 0.004 (fla-A_{27})			
Melting point in K (T_m) at 1 atm	743.221	717.167	3.633
Critical temperature in K (T_C)	932.252	914.743	1.914
Heat of Formation in KJ/mol	-309.816	-313.790	1.267
Henry's Law constant	7.228	7.240	0.171
pGI_{50}	5.658	5.700	0.750
QSSR model for isoflavone isofla-A_8 with $R^2_{fitness}$ = 0.9999; R^2_{test} = 0.9999; SE = 0.00119054 isofla-A_8 = 0.0000051 + 1.006 (isofla-A_4) + 0.253 (isofla-A_{18}) - 0.259 (isofla-A_{19})			
Melting point in K (T_m) 1 atm	746.066	717.167	4.030
Critical temperature in K (T_C)	936.202	914.743	2.346
Henry's Law constant	7.243	7.240	0.038
pGI_{50}	5.071	4.9944	1.503
QSSR model for flavone fla-A_9 with $R^2_{fitness}$ = 0.9999; R^2_{test} = 0.9999; SE = 0.00018592 fla-A_9 = 0.000004 + 0.047 (fla-A_5) + 1.025 (fla-A_{11}) - 0.072 (fla-A_{23})			
Melting point in K (T_m) at 1 atm	836.779	817.055	2.414
Critical temperature in K (T_C)	1029.858	1011.888	1.776
Henry's Law constant	7.052	7.050	0.026
logP	4.663	4.537	2.772
pGI_{50}	5.745	5.698	0.810
QSSR model for flavone fla-A_{10} with $R^2_{fitness}$ = 0.9999; R^2_{test} = 0.9999; SE = 0.00042716 fla-A_{10} = 0.00012 + 0.977 (fla-A_9) - 1.055 (fla-A_{21}) + 1.079 (fla-A_{22})			
Melting point in K (T_m) at 1 atm	815.011	814.381	0.077
Critical Pressure in Bar (P_C)	18.820	18.692	0.683
Critical temperature in K (T_C)	1003.621	1004.806	0.118

Heat of Formation in KJ/mol	-404.221	-387.410	4.339
Mol. Refractivity	10.963	10.930	0.305
logP	3.766	3.740	0.694
Henry's Law constant	7.063	7.050	0.190
pGI_{50}	5.678	5.652	0.448

The results in Table 3 pointed out that the linear relationship models QSSR between flavones and isoflavones using atomic-charge descriptors of target compound and those of predictive compounds are reliable and accurate. The linear models QSSR for target substances can be also applied for prediction of their physicochemical properties and anticancer activity of flavones and isoflavones, respectively. ANOVA single factor analysis also showed that the predicted physicochemical properties and anticancer activities of flavones and isoflavones resulting from the QSSR models are not different from the reference physicochemical values and experimental activities [[6]] with ($F = 0.0010 < F_{0.05} = 3.9423$).

The physicochemical properties and anticancer activity for target flavones and isoflavones were predicted by using the QSSR models are given in Table 3. The results turn out to be very good agreement with experimental data and those from empirical correlation calculated by ChemOffice [[9]]. This is illustrated in Figure 3.

The absolute relative errors (*ARE%*) are calculated by using the equation:

$$ARE\% = 100\left|Y_{i,\exp} - \hat{Y}_{i,\text{cal}}\right| / Y_{i,\exp} \qquad (3)$$

The values *ARE%* resulting from the linear models QSSR are in uncertainty extent of experimental measurements. The discrepancies between calculated and experimental properties and anticancer activity are insignificant.

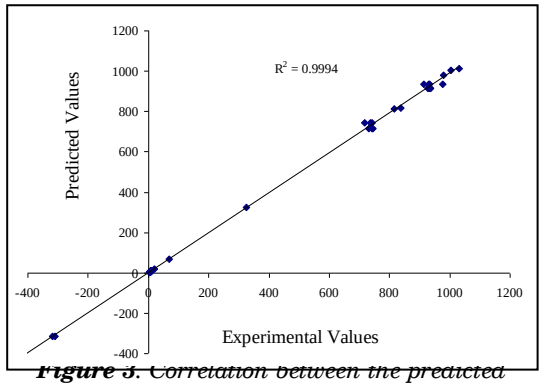


Figure 3. *Correlation between the predicted physicochemical and experimental data.*

4. Conclusion

This work exhibits the predictive approach for physicochemical properties of anticancer activity using the group of structurally similar flavones and isoflavones. But the most importance success is predictability of anticancer activity of flavones and isoflavones by using QSSR models. The atomic-charge matrix of flavones and isoflavones was used to construct effectively the QSSR models. This shows a promising technique and a good way for having physicochemical property data and biological activity by using similar compounds structurally.

DỰ ĐOÁN TÍNH CHẤT HÓA LÍ VÀ HOẠT TÍNH KHÁNG UNG THƯ CỦA CÁC CẤU TRÚC TƯƠNG TỰ NHAU CỦA CÁC FLAVONE VÀ ISOFLAVONE

Bùi Thị Phương Thúy[(1)], Phạm Văn Tất[(2)], Lê Thị Đào[(3)]

(1) Trường Đại học Khoa học Huế, (2) Trường Đại học Công nghiệp thành phố Hồ Chí Minh, (3) Trường Đại học Thủ Dầu Một

TÓM TẮT

Độ tin cậy của các mối quan hệ định lượng cấu trúc – hoạt tính hoặc tính chất để dự đoán các tính chất hóa lí và hoạt tính kháng ung thư của các dẫn xuất flavone và isoflavone

được cải thiện bằng các mối quan hệ định lượng giữa cấu trúc tương tự nhau của các chất flavon và isoflavon (QSSRs). Phương pháp chất đích được phát triển bằng nhóm luyện, mà chỉ chứa các hợp chất có cấu trúc tương tự với chất đích. Sự giống nhau về cấu trúc được thể hiện bằng sự tương quan đa chiều giữa các chiều tham số mô tả điện tích của chất đích và các chất dự báo với $R^2_{fitness}$ *= 0,9999 và* R^2_{test} *= 0,9999. Các tính chất hóa lý đã có và các hoạt tính kháng ung thư của các chất dự báo trong nhóm luyện được sử dụng trong trường hợp dự đoán các tính chất hóa lý chưa biết và hoạt tính kháng ung thư của các chất đích. Các kết quả ban đầu cho thấy phương pháp hợp chất đích cho kết quả dự đoán nằm trong vùng không chắc chắn của các phép đo thực nghiệm.*

REFERENCES

[1] J. C. Dearden, *Quantitative structure-property relationships for prediction of boiling point, vapor pressure, and melting point*. Environmental toxicology and Chemistry, Vol. 22, pp 1696-1709, (2003).

[2] M. Shacham, N. Brauner, H. Shore and D. Benson-Karhi, *Predicting Temperature-Dependent Properties by Correlations Based on Similarity of Molecular Structures Application to Liquid Density,* Ind. Eng. Chem. Res. 47, 4496-4504 (2008).

[3] G. St. Cholakov, R. P. Stateva, N. Brauner and M. Shacham, *Estimation of Properties of Homologous Series with Targeted Quantitative Structure Property Relationships (TQSPRs),* Journal of Chemical and Engineering Data, 53, 2510-2520, (2008).

[4] N. Brauner, G. St. Cholakov, O. Kahrs, R.P. Stateva and M. Shacham, *Linear QSPRs for Predicting Pure Compound Properties in Homologous Series,* AIChE J, 54, 978-990 (2008).

[5] Pham Van Tat, *Prediction of thermodynamic properties of similar organic compounds using artificial neural network,* Vietnamese Journal of Chemistry, P. 611-616, No.4A, 2009.

[6] T.C. Wang, I.L. Chen, P.J. Lu, C.H. Wong, C.H. Liao, K.C. Tsiao, K.M. Chang, Y.L. Chen, C.C. Tzeng, Bioorg. Med. Chem., *Synthesis, antiproliferative, and antiplatelet activities of oxime-and methyloxime-containing flavone and isoflavone derivatives,* Bioorganic & Medicinal Chemistry, Vol. 13, 6045–6053, (2005).

[7] Si Yan Liao, Jin Can Chen, Li Qian, Yong Shen, Kang Cheng Zheng, *QSAR., action mechanism and molecular design of flavone and isoflavone derivatives with cytotoxicity against HeLa,* European Journal of Medicinal Chemistry, Vol. 43, 2159-2170, (2008).

[8] D. D. Steppan, J. Werner, P. R. Yeater, *Essential Regression and Experimental Design for Chemists and Engineers,* (2006).

[9] CS Chem3D Ultra 2008, CambridgeSoft Corporation, USA, (2008).

[10] BMDP new system 2.0, Statistical Solutions Ltd., USA (2003).

[11] Hyper hem Release 8.03, Hypercube, Inc., USA (2008)