STUDY ON ANTI-HIV ACTIVITY OF DIARYLANILINE DERIVATIVES USING QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIP (QSAR)

Ihsanul Arief*, Ria Armunanto, and Bambang Setiaji

Department of Chemistry, Faculty of Mathematics and Natural Sciences, Universitas Gadjah Mada, Jl. Sekip Utara, Yogyakarta 55281

Received February 19, 2013; Accepted June 20, 2013

ABSTRACT

Study on anti-HIV activity of diarylaniline derivative compounds by using quantitative structure-activity relationship (QSAR) has been done. The compounds structure and their anti-HIV activities were obtained from literature. Molecular and electronic parameters were calculated by Austin Model 1 (AM1), Parameterized Model 3 (PM3), Hartree-Fock (HF), and density functional theory (DFT) methods. QSAR analysis was performed using multilinear regression method. The result shows that HF method can produce the best model as follows:
log EC_{50} = 46.418 + (99.360 × qC4) – (67.189 × qC9) – (278.869 × qC15) + (782.466 × qC19) – (127.463 × qO7)
n = 20; r^2 = 0.815; SEE = 0.393; F_{cal}/F_{tab} = 4.185; PRESS = 2.160
Those model can predict a good inhibitory activity (log EC_{50}) value of -0.3359 to compound N^1-(4'-Cyanophenyl)-5-(4"-cyanovinyl-2",6"-dimethyl-phenoxy)-4-hydroxyethylbenzene-1,2-diamine).

***Keywords**: diarylaniline; anti-HIV; QSAR*

ABSTRAK

Telah dilakukan kajian terhadap aktivitas anti-HIV dari senyawa turunan diaril anilina menggunakan model hubungan kuantitatif struktur-aktivitas (HKSA). Struktur senyawa dan aktivitas anti-HIV diperoleh dari literatur. Perhitungan parameter molekul dan parameter elektronik dilakukan dengan menggunakan metode Austin Model 1 (AM1), Parameterized Model 3 (PM3), Hartree-Fock (HF), dan density functional theory (DFT). Analisis HKSA dilakukan dengan menggunakan metode regresi multilinier. Diperoleh hasil bahwa metode HF dapar menghasilkan model terbaik sebagai berikut:
log EC_{50} = 46, 418 + (99,360 × qC4) – (67,189 × qC9) – (278,869 × qC15) + (782,466 × qC19) – (127,463 × qO7)
n = 20; r^2 = 0,815; SEE = 0,393; F_{hit}/F_{tab} = 4,185; PRESS = 2,160
Model tersebut dapat memprediksi nilai aktivitas penghambatan (log EC_{50}) sebesar -0,3359 dari senyawa N^1-(4'-sianofenil)-5-(4"-sianovinil-2",6"-dimetil-fenoksi)-4-hidroksietilbenzena-1,2-diamina).

***Kata Kunci**: diaril anilina; anti-HIV; HKSA*

INTRODUCTION

Since reported firstly on 1981, *acquired immunodeficiency syndrome* (AIDS) has been spread to humans in whole of the world. Like as other viruses, HIV also has been mutated along the time. Johnson et al. [1] reported that combination of E138K and M184I showed a 6.7-fold reduced phenotypic susceptibility to rilpirivine, in comparison with a 2.8-fold reduction for E138K alone. Those facts show the necessary of studies about anti-HIV compounds which effectively handle the growth of AIDS.

Numbers of studies have found that several organic compounds have anti-HIV activities, with their own mode of actions. Those organic compounds are derivatives of benzilpirimidine [2], arylurasil [3], lactam [4], bevirimat [5] or betulinic acid [6]. Sun et al. [7] reported that a series of diarylaniline (DAAN) compounds that is potent as anti-HIV against wild-type and HIV-1 RT-resistant viral strains have been discovered.

Currently, molecular modeling and computational chemistry are the inseparable parts in drug design. Computational methods result in saving in time and money for discovering a new drug. Quantitative structure-activity relationship (QSAR) method could predict the activities of compounds using their molecular and electronic properties. Those properties can be obtained from calculation in quantum-mechanical methods [8]. Studies to discover anti-HIV compounds using QSAR methods also developed, i.e. QSAR studies on anti-HIV activity of derivatives of

* Corresponding author. Tel/Fax : +62-85245644507
Email address : ihsanul.arief@mail.ugm.ac.id

Table 1. Molecular structures of DAAN compounds and their inhibitory activities [7]

No	R_1	R_2	EC_{50} (nM)	log EC_{50}
13a[a]	CF_3	$CH=CHCOCH_3$	10.90	1.0374
13b[a]	CF_3	$CH_2CH_2COCH_3$	6.85	0.8357
13c[a]	CF_3	CN	3.79	0.5786
13d[a]	CF_3	NH_2	13.80	1.1399
13e[b]	CH_2OH	CH_2OH	10.00	1.0000
14a[a]	CF_3	CH=CHCN	9.38	0.9722
14b[b]	SO_2NH_2	CH=CHCN	93.70	1.9717
14c[a]	$COOCH_3$	CH=CHCN	2.74	0.4378
14d[a]	COOH	CH=CHCN	2.30	2.3617
14e[b]	$CONH_2$	CH=CHCN	0.87	-0.0605
14f[b]	$CONHCH_3$	CH=CHCN	5.72	0.7574
14h[a]	CH_2OH	CH=CHCN	0.53	-0.2757
15a[a]	CF_3	CH_2CH_2CN	1.13	0.0531
15b[a]	SO_2NH_2	CH_2CH_2CN	380.00	2.5798
15c[a]	$COOCH_3$	CH_2CH_2CN	4.32	0.6355
15d[a]	COOH	CH_2CH_2CN	96.00	1.9823
15e[a]	$CONH_2$	CH_2CH_2CN	1.39	0.1430
15f[b]	$CONHCH_3$	CH_2CH_2CN	2.73	0.4362
15g[a]	$CONHNH_2$	CH_2CH_2CN	19.10	1.2810
15h[a]	NH_2	CH_2CH_2CN	7.30	0.8633
rilpirivine	H	CH=CHCN	0.52	-0.2840

[a]Training set
[b]Test set

phenyl ethyl thiourea [9], peptide [10], cyanoguanidine [11], and diaryl pirimidine [12].

Based on the data of effective concentration (EC_{50}, nM) of DAAN derivatives reported by Sun et al. [7] (Table 1), we have predicted a new compounds of diarylaniline derivative using QSAR procedure. To get a good prediction, the properties have been calculated using quantum-mechanical methods. In this study, the electronic and molecular properties were obtained from semi empirical (AM1 and PM3), HF level of theory, and DFT methods. The chosen QSAR model obtained from this study was then used to design *in silico* new compounds of anti-HIV of DAAN derivatives with better activity.

COMPUTATIONAL DETAILS

Hardware and Software

This study used a PC with Intel® Core™ i3 CPU M 350 4.54 GHz; RAM 5.00 GB. The programs used were Gaussian® 09W [14], HyperChem™ 8.0.10 [15], statistical programs IBM® SPSS® Release 19.0.0 [16].

Data Set

The structures and anti-HIV activities data of DAAN derivative compounds (Table 1) were divided into a training set (15 compounds) for generating QSAR models and a test set (5 compounds) for validating the quality of the models. Selection of the compounds in the training set and test is a key and important feature of any QSAR model. Therefore, the care was taken in such a way that biological activities of all compounds in test set lie within the maximum and minimum value range of biological activities of training set of compounds. The Uni-Column Statistics of test and training sets further reflected the correct selection of test and training sets.

The maximum and minimum values in training and test set were compared in a way that (i) the maximum value of log EC_{50} of test set should be less than or equal to maximum value of log EC_{50} of training set, (ii) the minimum value of log EC_{50} of test set should be higher than or equal to minimum value of log EC_{50} of training set [13]. In vitro effective concentrations (EC_{50}) of the compounds were converted into corresponding log EC_{50} values and used as dependent variables in QSAR calculations.

Computational Validation and Descriptor Calculation

To obtain the most suitable method of calculation, compound 13a was first computationally modeled using either AM1, PM3, Hartree-Fock (HF), and density functional theory (DFT) on Gaussian package to calculate chemical shif of the compound using H-NMR calculation method. Basis set of STO-3G was used for HF and DFT calculation. The calculated chemical shift data of the compound was then compared to the ones available from experimental H-NMR measurement [7]. The method of calculation (AM1, PM3, HF or DFT) giving smallest difference (PRESS value) between calculated and experimental data was chosen as the

Table 2. Uni-column statistics of the training and test sets for QSAR models

Set	Number of Data	Average	Max	Min	St. Dev	Sum
Training	15	0.975	2.580	-2.756	0.817	14.626
Test	5	0.821	1.972	-0.060	0.756	4.105

Table 3. Comparison of calculated and experimental H-NMR chemical shift data (δ, ppm)

Atomic number	Exp. [7]	AM1	DFT	HF	PM3
H42	2.14	3.23	2.63	2.21	2.56
H47	2.37	3.04	2.41	2.09	2.40
H56	5.71	4.46	5.20	7.11	3.90
H35	6.21	7.63	7.49	7.01	7.42
H45	6.65	7.05	6.59	6.50	7.39
H46	7.43	8.13	8.42	7.48	7.70
PRESS		5.8819	3.1200	2.7141	5.5344

Table 4. Atomic charge of C14 atom

No	qC14			
	AM1	PM3	HF	DFT
14b	-0.048	-0.058	0.000	-0.009
13e	-0.087	-0.108	-0.002	-0.019

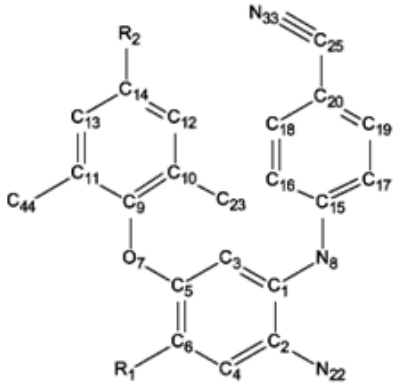

Fig 1. Lead structure of diarylaniline derivative

most suitable method and was used for further calculation in this study.

A large number of theoretical molecular descriptors such as surface area, volume, hydration energy, log P, refractivity, polarisability, molecular mass, HOMO-energy, LUMO-energy, ΔE (E_{HOMO}-E_{LUMO}) were calculated with HyperChem package; and electronic descriptors as atomic charge like described in Fig. 1 have been computed using Gaussian package. Those descriptors then used to develop structure–activity relationship of DAAN derivatives against the HIV.

Model Development

The QSAR model was generated by Multiple Linear Regression (MLR) Backward method by using SPSS package. It relates the dependent variable ŷ (biological activity) to a number of independent variables x_i (molecular and electronic descriptors) by using linear equations. This method of regression estimates the values of the regression coefficients by applying least square curve fitting method.

The model was chosen based on some statistical parameters such as r^2, standard estimation of error (SEE), F-ratio between the variance of predicted and observed activity, and PRESS [17], where:

$$\text{PRESS} = \Sigma (\text{predicted value-observed value})^2 \quad (1)$$

Model Validation

The best chosen model was used to predict log EC_{50} values of test set. The model are validated using criteria $r^2_{pred} > 0.5$, $r^2_m > 0.5$, where [18]:

$$r^2_m = r^2\left(1-\sqrt{\left|r^2-r^2_0\right|}\right) \quad (2)$$

The r^2 is linearity coefficient of test set with intercept and r^2_0 is linearity coefficient of test set without intercept.

Design and Activities Prediction of New Compounds

Based on the validated model, have been designed and predicted inhibitory activities (EC_{50}) of new DAAN derivative compounds. Then it will be chosen a compound with value of EC_{50} which is lower than rilpirivine.

RESULT AND DISCUSSION

Training Set and Data Set

Training set and test set were checked using a uni-column statistics as listed in Table 2. Table 2 shows that average and standard deviation values of training and test set are not different significantly, indicating a similar data distribution in both.

Table 5. Statistical parameters of 4 QSAR models of DAAN derivatives

Model	Descriptors	r^2	adjusted-r^2	F_{cal}/F_{tab}	PRESS
1	qC4, qC9, qC10, qC15, qC19, qC44, qO7, delta-E	0.914	0.799	1.922	956.031
2	qC4, qC9, qC15, qC19, qC44, qO7, delta-E	0.901	0.802	2.398	794.003
3	qC4, qC9, qC15, qC19, qO7, delta-E	0.868	0.768	2.438	448.627
4	qC4, qC9, qC15, qC19, qO7	0.847	0.762	2.859	1.430

Table 6. Comparison of observed and predicted values of log EC_{50}

Compound	Log EC_{50} observed	Predicted log EC_{50}			
		Model 1	Model 2	Model 3	Model 4
14b	1.972	9.957	9.254	6.918	1.415
13e	1.000	8.577	7.977	6.068	0.709
14f	0.757	7.916	7.160	5.437	0.044
15f	0.436	8.494	7.716	6.009	0.537
14e	-0.060	7.794	7.036	5.312	-0.082
	PRESS	299.054	246.047	131.965	0.915

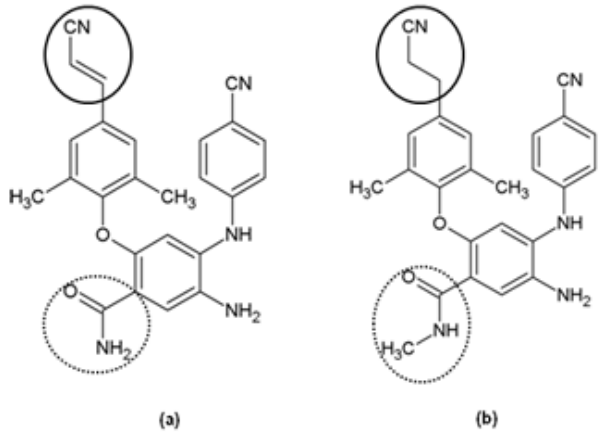

Fig 2. Structures of molecules 14e (a) and 15f (b)

Computational Validation

The results of the calculation were compared to those obtained from experimental measurement (1H-NMR, 400 MHz) [7] as listed in Table 3. Table 3 shows that chemical shift data obtained from H-NMR calculation using HF have a better agreement with those resulted from experimental measurement (lowest PRESS value) as compared to those calculated by AM1, PM3 and DFT methods, suggesting that HF method describe the chemical conformation of DAAN derivatives more accurately than the other methods. Therefore, HF method has been selected as calculation method for further modeling of DAAN derivatives in this study.

All of electrons were included in HF calculation. This causes HF method more accurate than AM1 and PM3 which only includes the valence-electrons and some empirical parameter represented the core-electrons [20]. In the other hand, DFT method does not use electron wave-function to determine the energy, but it uses the density of electron that distributed in molecule. Therefore, DFT fails to produce good electronic properties of the DAAN compounds.

Model Development

The descriptors were used to develop QSAR models, showing physico-chemical properties of the molecules. Log P values of compound 15f (4.54) was higher than 14e (4.23), indicating a more hydrophobic properties of 15f. Compound 15f become more non-polar due to the existence of $-CH_3$ group (dashed-line circle in Fig. 2) which can decrease the polarity of a compound. In contrary, the compound 14e has a double-bond (full-line circle in) which can increase the polarity.

Table 4 shows the atomic charge of C14 atom in compound 14b (-0.048) is more positive than 13e (-0.087). This is due to compound 14b has $–CH=CHNH_2$ group which has higher electronegativity than $-CH_2OH$ group in compound 13e. Therefore, electron withdrawing in compound 14b is stronger than compound 13e.

In first step of model development, all of the descriptors (11 moleculars and 25 electronics) were included. Based on statistical backward analysis, the non-significant descriptors were excluded from the model. The results were listed in Table 5. In Table 5, descriptor qO7 always appears in all of the models, indicating the important descriptor in anti-HIV activity of DAAN derivatives. This result is in a good agreement with Johnson et al. [19] showing DAAN derivatives bind to reverse transcriptase by hydrogen bonding.

Model 1, model 2, and model 3 have higher r^2-value than model 4, but their PRESS-value were significantly higher than model 4. It due to the first three models have over-fitting in their predictions, so the predicted values were over-estimate than the observed

Table 7. New designed DAAN derivative compounds and their predicted log EC_{50} and EC_{50} using HF model

No	R1	R2	Predicted log EC_{50}	Predicted EC_{50} (nM)
1	CH_2CH_2OH	CH=CHCN	-0.3359	0.46
2	CH_2OCH_3	CH=CHCN	-0.2249	0.60
3	$CONHCH_2CH_3$	CH=CHCN	-0.1861	0.65
4	$COOCH_2CH_3$	CH=CHCN	-0.1598	0.69
5	$CH_2CH_2CH_2OH$	CH=CHCN	-0.1293	0.74
6	CH_2F	CH=CHCN	-0.1100	0.78
7	$COOCH_2CH_2OH$	CH=CHCN	0.1233	1.33
8	CH_2Br	CH=CHCN	0.2029	1.60
9	CH_2CH_2OH	CH_2CH_2CN	0.3553	2.27
10	$CONHCH_2CH_3$	CH_2CH_2CN	0.4223	2.64

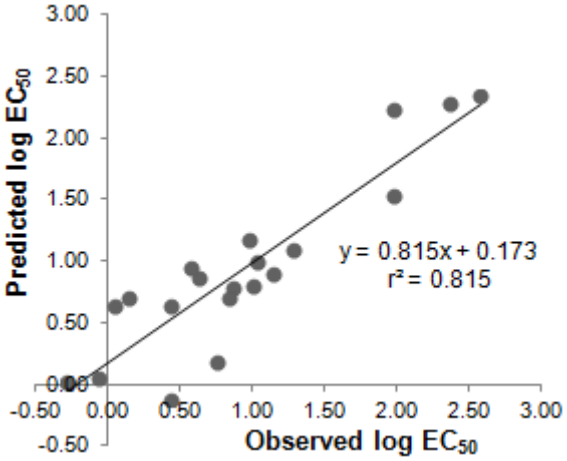

Fig 3. Plot of observed versus predicted log EC_{50} values of DAAN derivatives calculated by HF model

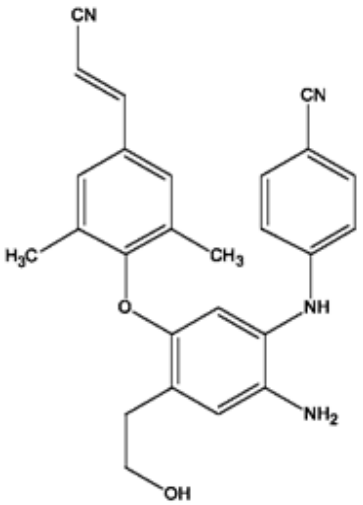

Fig 4. The chosen designed diarylaniline derivative compound (predicted log EC_{50} = -0.3359)

values. Based on the r^2 criteria ($r^2 > 0.6$), all of the model were passed to validation step.

Model Validation

The passed models were then used to predict the value of log EC_{50} of the data in test set. The results were shown in Table 6. From Table 6, model 4 can predict log EC_{50} value that has negative value with a small difference (± 0.02). It show that model 4 was good enough, because the new compound is expected to has negative value of log EC_{50}. The negative value indicates those new compound has better activity than synthesized compound and standard compound rilpirivine.

The values of observed and predicted log EC_{50} then plotted and the values of r^2_{pred} were evaluated. The values of r^2_{pred} from 4 model is 0.818; 0.836; 0.785; 0.803 for model 1, model 2, model 3, and model 4 respectively. It indicates that all of the models were passed the next criteria ($r^2_{pred} > 0.5$).

To evaluate the external predictability, values of r^2_m were determined from the plot of observed and predicted log EC_{50} values. It was obtained that model 1, 2, and 3 have negative r^2_0 values because they were over-estimate in prediction, thus their r^2_m-values can not be determined. The over-estimate were also detected in PRESS-values. Only r^2_m-value of model 4 that can be determined, it was 0.742. Those values are greater than 0.5, indicating a good external predictability of model 4. In the next step, model 4 was used to design and predict the new DAAN derivatives.

The all descriptors obtained from HF calculation (training and test set) were applied to Enter procedure, to obtained a new equation as presented in Equation (3):

$$log\ EC_{50} = 46.418 + (99.360 \times qC4) - (67.189 \times qC9) - (278.869 \times qC15) + (782.466 \times qC19) - (127463 \times qO7) \quad (3)$$

n = 20; r^2 = 0.815; SEE = 0.393; F_{cal}/F_{tab} = 4.185; PRESS = 2.160

Log EC_{50} values calculated from Equation (3) were compared with the observed values showing a slope of 0.815. These slopes indicate that the model 4 can predict 81.5% anti-HIV activities of DAAN derivatives.

Design and Activities Prediction of New Compounds

In the design of new compounds, substituent was replaced with other groups or chemical species. The selection of those new species or groups, was based on the possibility to be synthesized and the materials availability. It was expected also, the step of the synthesis will be done only in single or double steps to keep the rendemen in a good term.

Predictions of log EC50 values have been done using the descriptors from designed compounds with Equation (3). Table 7 shows 10 designed compounds with their predicted log EC_{50} and also predicted EC_{50} values. In Table 7, generally, it can be seen that the chosen new compounds have substituent groups which potentially to make hydrogen bonding. Those groups have F, O, N atoms which are electronegative atoms, so they can bind to hydrogen atom in reverse transcriptase residues and resulting inhibitory effects.

In the other side, most of the best compounds have R2 groups with a double-bond. This indicates, double-bond in R2 also has a significant effect on binding mechanism between DAAN derivatives to reverse transcriptase enzyme of HIV.

Compound number 1 has predicted log EC_{50} value of -0.3359, which is lower than rilpirivine's (-0.2840M). This compound assumed as more active as an anti-HIV agent and has a systematic name of N^1-(4′-Cyanophenyl)-5-(4″-cyanovinyl-2″,6″-dimethyl-phenoxy)-4-hydroxyethylbenzene-1,2-diamine) (the molecular structure is shown in Fig. 4).

CONCLUSION

From the results of this study, it can be concluded that Hartree-Fock method give the best QSAR model to describes relationship between descriptors of diarylaniline derivatives to its anti-HIV activities. The best designed compound is N^1-(4′-Cyanophenyl)-5-(4″-cyano vinyl-2″,6″-dimethyl-phenoxy)-4-diethyletherbenzene1,2-diamine) with predicted log EC_{50} value of -0.3359.

REFERENCES

1. Johnson, V.A., Calvez, V., Günthard, .F.H., Paredes, R., Pillay, D., Shafer, R., Wensing, A.M., and Richman, D.D., 2011*, Top. Antivir. Med.*, 19, 4, 156–164.
2. Rotili, D., Samuele, A., Tarantino, D., Ragno, R., Musmuca, I., Ballante, F., Botta, G., Morera, L., Pierini, M., Cirilli, R., Nawrozkij, M.B., Gonzalez, E., Clotet, B., Artico, M., Esté , J.A., Maga, G., and Mai A., 2012, *J. Med. Chem.*, 55, 7, 3558–3562.
3. Wang, X., Zhang, J., Huang, Y., Wang, R., Zhang, L., Qiao, K., Li, L., Liu, C., Ouyang, Y., Xu, W., Zhang, Z., Zhang, L., Shao, Y., Jiang, S., Ma, L., and Liu, J., 2012, *J. Med. Chem.*, 55, 5, 2242–2250.
4. Wu, X., Öhrngren, P., Joshi, A.A., Trejos, A., Persson, M., Arvela, R.K., Wallberg, H., Vrang, L., Rosenquist, Å., Samuelsson, B.B., Unge, J., and Larhed, M., 2012, *J. Med. Chem.*, 55, 6, 2724–2736.
5. Qian, K., Bori, I.D., Chen, C-H., Huang, L., and Lee K-H., 2012, *J. Med. Chem.,* 55, 18, 8128–8136.
6. Dang, Z., Qian, K., Ho, P., Zhu, L., Lee, K., Huang, L., and Chen, C-H., 2012, *Bioorg. Med. Chem. Lett.*, 22, 16, 5190–5194.
7. Sun, L-Q., Zhu, L., Qian, K., Qin, B., Huang, L., Chen, C-H., Lee, K-H., and Xie, L., 2012, *J. Med. Chem.*, 55, 16, 7219–7229.
8. Veerasamy, R., Chean, O.C., Subramaniam, D.K., Ying, N.M., Sivadasan, S., Rajak, H., and Rasheed, A., 2013, *Med. Chem. Res.*, 22, 35-44.
9. Ravichandran, V., Mourya, V.K., and Agrawal, R.K., 2009, *Dig. J. Nanomat. Biostruct.,* 4, 1, 213–221.
10. Singh, R.K., Sahu, V.K., Khan, S.A., and Singh, P.P., 2011, *J. Appl. Chem. Res.*, 18, 35–50.
11. Deeb, O., and Jawabreh, M., 2012, *Adv. Chem. Eng. Sci.*, 2, 82–100.
12. Hemmateenejad, B., Javidnia, K., Nematollahi, M., and Elyasi, M., 2009, *J. Iran. Chem. Soc*., 6, 2, 420–435.
13. Jain, S.V., Ghate, M., Bhadoriya, K.S., Bar, S.B., Chaudhari, A., and Borse, J.S., 2012, *Org. Med. Chem. Lett.,* 2, 1, 22–34.
14. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G.A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H.P., Izmaylov, A.F., Bloino, J., Zheng, G., Sonnenberg, J.L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, Jr., J.A., Peralta, J.E., Ogliaro, F., Bearpark, M., Heyd, J.J., Brothers, E., Kudin, K.N., Staroverov, V.N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J.C., Iyengar, S.S., Tomasi, J., Cossi, M., Rega, N., Millam, J.M., Klene, M., Knox, J.E., Cross, J.B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R.E., Yazyev, O., Austin, A.J., Cammi, R., Pomelli, C., Ochterski, J.W., Martin, R.L., Morokuma, K., Zakrzewski, V.G., Voth, G.A., Salvador, P., Dannenberg, J.J., Dapprich, S., Daniels, A.D., Farkas, Ö., Foresman, J.B., Ortiz, J.V., Cioslowski, J., Fox, D.J., 2009, *Gaussian 09, Revision A.02*, Gaussian, Inc., Wallingford CT.

15. HyperCube, Inc, HyperChem^{TM} 8.0.10 for Windows, http://www.hyper.com., 2011.
16. IBM® SPSS® Release 19.0.0, SPSS, Inc., 2010.
17. Podunavac-Kuzmanović, S.O., Cvetković, D.D., and Barna, D.J., 2009, *Int. J. Mol. Sci.,* 10, 4, 1670–1682.
18. Hu, R., Doucet, J., Delamar, M., and Zhang, R., 2009, *Eur. J. Med. Chem.*, 44, 5, 2158–2171.
19. Johnson, B.C., Pauly, G.T., Rai, G., Patel, D., Bauman, J.D., Baker, H.L., Das, K., Schneider, J.P., Maloney, D.J., Arnold, E., Thomas, C.J., and Hughes, S.H., 2012, *Retrovirology,* 9, 99–109.
20. Cramer, C.J., 2004, *Essentials of Computational Chemistry: Theories and Models*, 2nd ed., John Wiley & Sons Ltd., Chichester, West Sussex, England.