RASĀYAN *J. Chem.*
Vol. 15 | No. 1 |359-368| January - March | 2022
ISSN: 0974-1496 | e-ISSN: 0976-0083 | CODEN: RJCABP
http://www.rasayanjournal.com
http://www.rasayanjournal.co.in

QSAR MODELING AND DESIGN OF A NEW MODEL OF ANTI-HIV DRUG 1-ARYL-TETRAHYDROISOQUINOLINE DERIVED USING THE PM3 SEMIEMPIRICAL METHOD

Yusthinus T. Male✉, I. W. Sutapa, I. B. Kapelle and M. Lopulalan
Department of Chemistry, Pattimura University, Faculty Mathematics, and Natural Sciences, Pattimura University, Jalan. Ir. Putuhena, Ambon, Maluku-Indonesia 97233.
✉Corresponding Author: yusmale@fmipa.unpatti.ac.id

ABSTRACT

Quantitative Structure-Activities Relationship Analysis (QSAR) on the 1-aryl-tetrahydroisoquinoline derivative as a candidate for the anti-HIV drug (Human Immunodeficiency Virus) has been done. The analysis using an electronic descriptor of atomic net charge, dipole moment, polarization, HOMO-LUMO energy, and log P was calculated using the PM3 semiempirical method. The best equation model is determined using multilinear regression analysis. The QSAR analysis show that the best QSAR equation model of 1-aryl-tetrahydroisoquinoline derivatives: pEC_{50}= -23.019-10.867(qC_3)-162.911(qC_4)-107.057(qC_6) + 53.621(qC_8) - 0.047 (μ) + 0.158 (α) + 0.175 (log P); (n = 36; r = 0.935; r^2 = 0.875; SE= 0.135927366; F_{count}/F_{table}= 7.945277618; PRESS = 2.857943). The compound with the higher activity value is the best theoretically derivative compound for the anti-HIV drug candidate. The new model of the 1-aryl-tetrahydroisoquinoline derivative from the design is expected to have better activity than the 1-aryl-tetrahydroisoquinoline derivative compound that already exists is 1-(6-acetylpyridine-2-yl)-7,8-methoxy-quinoline.

Keywords: Anti-HIV, QSAR, Semi-empirical, PM3, Regression.

RASĀYAN *J. Chem.*, Vol. 15, No.1, 2022

INTRODUCTION

The AIDS (Acquired Immune Deficiency Syndrome) plague has been haunting the world for 20 years, and it seems that its spread will continue to increase. AIDS was first classified as a disease in 1981. Meanwhile, HIV (Human Immunodeficiency Virus) was identified as the cause in 1984. AIDS is a disease caused by longstanding HIV infection. Recently, there was no cure or perfect treatment for AIDS, so that AIDS treatment became the most challenging medical problem in the whole world.[1,4,6,7,13,14]

It has been suggested that the compound produced by 1-aryl-tetrahydroisoquinoline can protect cells against HIV. 1-aryl-tetrahydroisoquinoline compounds also have anti-bacterial and anti-convulsion activity.[2,3,20] Many efforts were conducted to get a new anti-HIV compound, and it is necessary to develop a 1-aryl-tetrahydroisoquinoline compound design to obtain a 1-aryl-tetrahydroisoquinoline derivative with higher activity. Computational Chemistry can be used to predict the structure and properties of a compound before it is synthesized. One application of computational chemistry that can be applied is the Study of the Quantitative Structure-Activity Relationship (QSAR). The quantitative correlation between molecular structure and measured biological activity values was studied experimentally.[5,8,11,17] Based on QSAR calculations, it can be predicted that the most active side of a compound provides activity as a drug. This step can reduce the failure of experimental research in the laboratory and can streamline energy and cost.[15,17,18,19,21] In this study was used is the semiempirical method, Parametric Method 3 (PM3). The PM3 is the third parameterization method of the Neglect of Differential Overlap (NDO) method with Modified NDO (MNDO) and Austin Method version 1 (AM1) as the first and second parameterization methods. The PM3 method is a parameterization of the AM1 method so this method is relatively more thorough.[10,12]

Based on the background that has been described, the purpose of this study is QSAR Modelling and Design of a New Model of Anti-HIV Drug 1-aryl-tetrahydroisoquinoline Derived Using the Semi-Empirical Method of PM3. The computational calculations are carried out on a 1-aryl-tetrahydroisoquinoline molecule to obtain a descriptor. The descriptor includes the net charge of atoms,

HOMO energy (Highest Occupied Molecular Orbital), LUMO energy (Low Unoccupied Molecular Orbital), and other parameters, namely dipole moment, polarizability, and the logarithm of the partition coefficient (log P). Then, the calculation was carried out to conduct design a new model of the anti-HIV drug from a derivative of the 1-aryl-tetrahydroisoquinoline compound.

EXPERIMENTAL

Material and Instrumentation

The materials used in this study were molecular structure data and biological activity from a series of 1-aryl-tetrahydroisoquinoline derivatives obtained from the literature. Effective Concentration (EC_{50}) was used as the dependent variable (Table I).[3] The instruments used for the QSAR test were a personal computer with the specification: Intel inside Core i3 processor, DDR3 RAM 4GB, hard drive640GB, Software (OS Windows, HyperChem™ version 8.0 and SPSS® 19.0 for Windows

Procedure

Geometry Optimization Using PM3 Semiempirical Method

One series of compounds used in this study consisted of 36 1-aryl-tetrahydroisoquinoline derivatives, each of which was geometry optimized with the PM3 method. Each 1-aryl-tetrahydroisoquinoline derivative series compound, used as research material with the Hyperchem 8.0 program, was made 2D structure. Furthermore, the structure of the compound was made in a 3D structure, and it is equipped with hydrogen atoms added to each atom.

Multilinear Regression Analysis

Multilinear regression analysis is performed to obtain a mathematical relationship between variables/descriptors with drug activity. The equation is used next to predict the effect of independent variables (descriptors) on non-independent variables (drug activity). The 1-aryl-tetrahydroisoquinoline derivatives were divided into two groups of data, namely 28 compound data used for the fitting process and eight compounds' data used for the validity test process. Compound fitting data is used to find the best QSAR equation model, and the results are tested with test compound data. There are two types of variable variation, the independent variable and the independent variable, with the QSAR equation as follows Eqn.-1.

$$\text{Log}(1/EC_{50}) = \sum_{i}^{n} \left(P_{(qi)} q_{(i)} + P_{(\mu)} \mu + P_{(ELU\quad)} \, ELUMO + P_{(EHOMO)} EHOMO + P_{(\alpha)} \alpha + P_{(\log P)} \log P + D \right) \quad (1)$$

The independent variables were used are q (atomic net charge), μ (dipole moment), E_{HOMO} (highest filled molecular orbital energy), E_{LUMO} (lowest unallocated molecular orbital energy), α (polarizability), and log P (partition coefficient logarithm). The dependent variable is log (1/C), and D is the constant of the regression equation.

RESULTS AND DISCUSSION

Electronic Structure Calculation Results of 1-aryl-tetrahydroisoquinoline

In this study, QSAR anti-HIV studies were carried out for 36 1-aryl-tetrahydroisoquinoline derivatives obtained from the literature[3]. Geometry optimization will produce data of atomic net charge descriptor, dipole moment, HOMO energy, and LUMO energy. In contrast, the calculation of QSAR will produce polarizability descriptor data and logarithm of the partition coefficient (log P) of drug compounds (Table-1). Geometry optimization of 1-aryl-tetrahydroisoquinoline derivatives to get the most stable structure shown in Fig.-1.

The net charge of an atom is usually used to indicate the density of electrons in an atom or molecule. In Table-3, it can be seen that the net charge of atoms for N_2, C_3, C_4, C_6, and C_9 atoms is relatively not much different, whereas the charge of the C_1, C_7, and C_8 atoms shows some changes. This result occurred due to the influence of substituents that are bound and occupy positions on these atoms. The presence of these substituents results in a reduction in the partial charge induction. Following Table-1, it appears that compounds with different structures produce different anti-HIV activities. When Table-2 content

compared to Table-1 can be seen that the position and type of substitutional group used to affect the net charge of the atom.

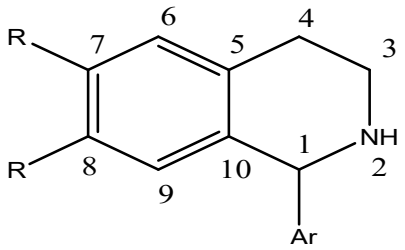


Fig.-1: 1-aryl-tetrahydroisoquinoline Compound Backbone

Based on Table-1, molecular structure data is divided into fitting and test data. The data is divided based on considerations first: the test data, biological activity values must span several times but must not exceed fitting data activity values of more than 10%. Second: the test data must represent an equal number of active and inactive compounds for uniform data sampling[3].

Table-1: Derivatives of the Compound 1-aryl-tetrahydroisoquinoline

Fitting Compound				
Number	R	Ar	EC_{50} (μM)	pEC_{50} (M)
1	-OMethyl	Phenyl	119.6	3.922
2	-OH	Phenyl	67.3	4.172
3	-OMethyl	2-Methylphenyl	42.5	4.372
4	-OH	2-Methylphenyl	61.9	4.210
5	-OMethyl	4-Methylphenyl	67.3	4.172
8	-OH	2-Methoxyphenyl	42.2	4.375
9	-OMethyl	4-Methoxyphenyl	127.2	3.896
10	-OH	4-Methoxyphenyl	34.1	4.467
11	-OMethyl	2,4-Dimethoxy Phenyl	43.8	4.359
12	-OH	2,4-Dimethoxy Phenyl	46.5	4.333
13	-OMethyl	4-Trifluoromethyl phenyl	27.8	4.556
14	-OH	4-Trifluoromethyl phenyl	27.6	4.559
16	-OH	4-Bromophenyl	36.2	4.441
17	-OMethyl	3-Bromophenyl	43.5	4.362
19	-OMethyl	2-Bromophenyl	31.3	4.504
20	-OH	2-Bromophenyl	25.7	4.625
21	-OMethyl	2-Florophenyl	69.0	4.161
22	-OH	2-Florophenyl	58.4	4.234
23	-OMethyl	3-Chlorophenyl	31.8	4.498
25	-OMethyl	2-Chlorophenyl	58.4	4.234
26	-OH	2-Chlorophenyl	51.8	4.286
27	-OMethyl	4-Nitrophenyl	60.1	4.221
28	-OH	4-Nitrophenyl	53.7	4.270
29	-OMethyl	2-Furyl	152.4	3.817
30	-OH	2-Furyl	61.7	4.210
31	-OMethyl	Pyridine-2-yl	171.2	3.766
35	-OMethyl	2-Naphthyl	6.9	5.161
36	-OH	2-Naphthyl	5.3	5.276
Test Compound				
06	-OH	4-Methylphenyl	8.2	5.086
07	-OMethyl	2- Methoxyphenyl	45.9	4.338
15	-OMethyl	4-Bromophenyl	16.9	4.772
18	-OH	3-Bromophenyl	53.5	4.272
24	-OH	3-Chlorophenyl	4.6	5.337

32	-OH	Pyridine-2-yl	48.4	4.315
33	-OMethyl	1-Naphthyl	22.6	4.646
34	-OH	1-Naphthyl	33.2	4.479

Data calculation of the net charge of atoms, the substitution of an atom or group on C_1, C_7, and C_8 atoms reduces the negative value of the C_1, C_7, and C_8 atomic charges (adding to the positive value) relatively. The replacement of H atoms with other groups reduces the induction of positive partial charges. The charge of the C atom that binds to the H atom is more negative than the charge of the C atom, which binds to the substitutional group. This result is due to the C1, C7, and C8 atoms not binding to H, so the effect of the induced positive partial charge caused by the H atom to the C atom is lost. The net charge of a C atom can be positive and negative, depending on the group attached to the C atom. Thus, it can be stated that the substitution of groups with different types and positions will affect the net charge of the atoms where the substituent is located. Differences in molecular structure will result in differences in electronic properties (the net charge of atoms), resulting in different anti-HIV activities.

Other descriptors from the calculation using the PM3 semiempirical method used in this study are HOMO-LUMO energy, dipole moment, polarizability, and P log. HOMO energy is directly related to ionization potential and the susceptibility of molecules in the attack to electrophiles. In contrast, LUMO energy is directly related to electron affinity and the nature of molecular susceptibility in attacking nucleophiles. The HOMO-LUMO energy difference is essential in determining the molecular stability index. Molecules with significant energy differences HOMO-LUMO have high stability.

Table-2 shows the most considerable dipole moment value in compound 29 with a value of 4.352 Dybe and the smallest dipole moment value in compound 18 with a value of 0.989 Dybe. The log P value is related to the distribution of drugs in the body. The more positive the log P value of the compound will tend to be in the nonpolar phase rather than the polar phase. This result has a physical meaning that the compound will interact more with receptors in the body relative to interactions with the blood. For compounds that have a negative log P value, they tend to be in the polar phase. Based on Table-3, the compound with the most positive log P-value is 0.14, and the compound with the most negative log P-value is -2.48.

Results of Correlation Analysis between Variables

The correlation analysis between variables was conducted to know the relationship between variables at a specific position in the structure of compounds on anti-HIV activity. The independent variable consists of net atomic charges (qC_1, qN_2, qC_3, qC_4, qC_6, qC_7, qC_8, and qC_9), dipole moment, HOMO energy, LUMO energy, polarizability, and log P. The non-independent variables are anti-activities HIV (pEC_{50}), all included in one data group and then analyzed with the SPSS 19.0 For Windows program to determine the correlation value between variables. Correlation values can be calculated by first plotting all the variables to be tested for correlation, then analyzed by the bivariate correlation method, which automatically calculates the correlation values between related variables (Table-3).

In Table-2, it is clear that the net atomic charge variables, the dipole moment, the HOMO energy, the LUMO energy, the polarizability, and the P log show the closest correlation. The negative correlation value does not indicate whether or not the influence of the substituents on anti-HIV activity, only shows the direction of its influence. Negative correlation prices show a negative relationship, meaning that the effect of one variable is inversely proportional to other variables.

The effect on the pEC50 value is not strong for other variables because the correlation size is not too large. However, an assessment must still be done to see the effect between these variables because the size of the correlation is not enough to conclude variables that affect the activity. Further assessment of influential variables was carried out by analysis using the multilinear regression method to obtain a linear equation involving all descriptors as independent variables associated with the anti-HIV activity (pEC_{50}) as the dependent variable.

Results of QSAR Multilinear Regression Analysis

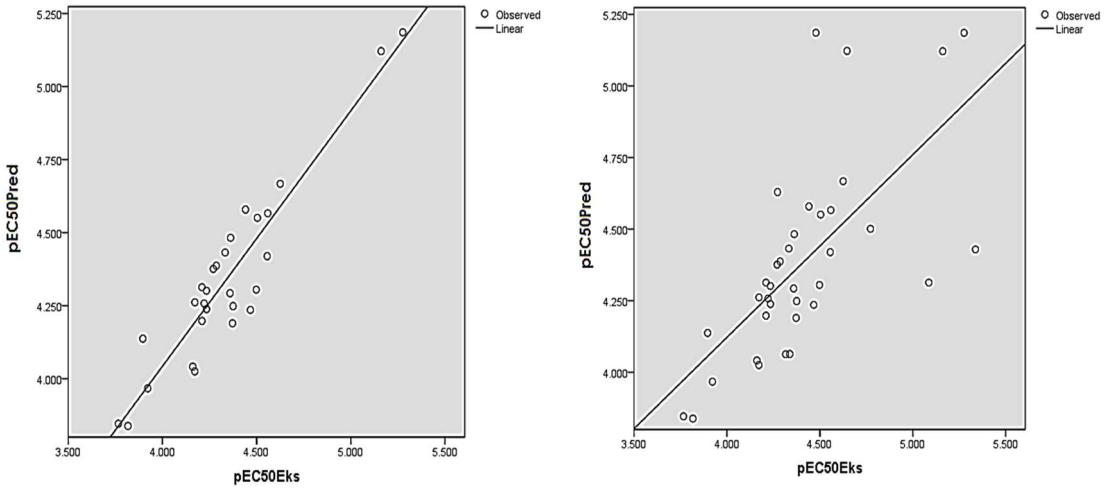
In the analysis QSAR multilinear regression of 1-aryl-tetrahydroisoquinoline derivatives, 28 compounds were used as data compounds, while eight fitting compounds were used for test compounds. The fitting

compound is intended to find the best QSAR equation model, and then its validity is tested using test compound data. The fitting compound descriptors were processed using multilinear regression analysis. The backward method is used to see only the dominant influential descriptors and not to see the percentage of influence of each descriptor. Parameters with little effect will automatically be issued one by the program to produce some acceptable equations. The calculation in this study obtained nine selected QSAR equation models. The constants of each QSAR equation model are worthy of the further study presented in Table-3 and Table-4.

Selection of the Best QSAR Equation Model

The choice of the QSAR equation model is based on several statistical parameters, including r, r^2, F_{count}/F_{table}, SE, and PRESS. The value of r (correlation coefficient) shows the level of relationship between the values of the biological activity of the experiment with the calculated data based on the equation obtained from the regression analysis. The value of r^2 is the square of the correlation coefficient, which is called the coefficient of determination. In statistical tests, the r^2 value is usually used because it has a higher level of accuracy than the value of r. The value of r^2 has a greater interval when compared to the r-value so that a small difference that is hardly observed in r can be observed more clearly on r^2. But the parameter r or r^2 is only a measure of linearity of the related equation model but cannot describe the predicted size of the equation model, so other statistical parameters need to be considered.[9,16]

The QSAR selected equation model was analyzed using 28 compound (internal PRESS) data and tested using eight compound external PRESS data. Based on Table-5, the best QSAR equation model is Eqn.-7 which is supported by the PRESS parameter with the relatively minimum compared to the other eight equations. After obtaining the best QSAR equation model, the prediction of the predicted activity (pEC_{50}) for 28 fitting compounds and 8 test compounds is carried out to determine the relationship between experimental activity and predictive activity. A comparison of the price of the predicted activity of the QSAR model 7 equation with the experimental activity is shown in Table-5 and the correlation graph is presented in Fig.-2 (a and b). The complete QSAR model 7 equation can be written as follows: pEC_{50} = - 23,019 - 10,867 (qC_3) - 162,911 (qC_4) - 107,057 (qC_6) + 53,621 (qC_8) - 0.047 (μ) + 0.158 (α) + 0.175 (log P) (n = 36; r = 0.935; r^2 = 0.875; SE = 0.135927366; F_{count}/F_{table}= 7.945277618; PRESS = 2.857943).



(a) (b)

Fig.*-2:*(a). The correlation graph between experimental anti-HIV activity (pEC_{50exp}) and predictive anti-HIV activity ($pEC_{50pred.}$) for the best internal and external QSAR equation model, (b). The correlation graph between experimental anti-HIV activity (pEC_{50exp}.) and predictive anti-HIV activity (pEC_{50pred}.) for the best QSAR equation model.

Based on the selected QSAR model 7 equation, it can be predicted that the active region of the 1-aryl-tetrahydroisoquinoline derivative is C_3, C_4, C_6, and C_8, which are active centers to provide anti-HIV effects (Fig.-3). If the synthesis of the 1-aryl-tetrahydroisoquinoline derivative of these four atoms is lost, the activity will decrease. Therefore, to design a new 1-aryl-tetrahydroisoquinoline derivative compound,

the substitution modification should focus on the atomic region, but it does not rule out substitution being carried out in other atomic regions.

Design of New Compound Models

The new molecular structure model of 1-aryl-tetrahydroisoquinoline derivative compounds needs to design. It is necessary to consider the modification of the position substituents and the ease of synthesis of the new compounds. Following the variations in the position of the substituents, the study focused on the position of the previous group. The new compound model design does not use the best QSAR equation results obtained from the QSAR analysis but uses an organic approach. The structural model of the 1-aryl-tetrahydroisoquinoline derivative compound can be seen in Table-6.

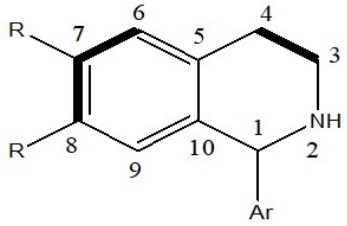


Fig.-3: Prediction of active central regions in the basic structure of 1-aryl-tetrahydroisoquinoline compounds

The result of the electronic structure descriptors calculated using the PM3 semiempirical method is presented in Table-9. The results of the electronic structure calculations were used to calculate the theoretical activity values (pEC_{50}) of the new compound model design by entering the descriptor values in the best selected QSAR equation model (Table-7).

Table-2: Results of Electronic Structure Calculations of 1-aryl-tetrahydroisoquinoline Derivatives Using the PM3 Semiempirical Method

Comp. No.	Net Atomic Charge (Coulomb)								Dipole Moment	E_{HOMO}	E_{LUMO}	Polarizabilities	Log P
	qC_1	qN_2	qC_3	qC_4	qC_6	qC_7	qC_8	qC_9	(Dybe)	(eV)	(eV)	(A3)	
1	0.03055	-0.05677	-0.08262	-0.04661	-0.09529	0.02582	0.06580	-0.17086	1.773	-8.87384	0.17181	31.12	-0.43
2	0.03574	-0.05651	-0.08222	-0.0498	-0.12506	0.07758	0.00918	-0.1459	2.116	-8.69807	0.22860	27.45	-0.49
3	0.02027	-0.05558	-0.08226	-0.04672	-0.09374	0.02362	0.06816	-0.17668	3.234	-8.84625	0.18800	32.95	-0.27
4	0.02635	-0.05893	-0.07996	-0.0512	-0.12191	0.07712	0.01120	-0.15894	2.179	-8.70507	0.23886	29.28	-0.34
5	0.03198	-0.05687	-0.08276	-0.04649	-0.09548	0.02544	0.06556	-0.17068	2.004	-8.85677	0.17692	32.95	-0.27
6	0.03714	-0.0566	-0.08236	-0.04967	-0.12518	0.07719	0.00892	-0.14578	2.257	-8.68024	0.21542	29.28	-0.34
7	0.03563	-0.0613	-0.08111	-0.04759	-0.09418	0.02509	0.06408	-0.17099	2.86	-8.85688	0.17372	33.59	-1.42
8	0.04602	-0.06557	-0.07722	-0.05093	-0.12286	0.07531	0.01218	-0.16259	3.167	-8.59741	0.24680	29.92	-1.48
9	0.03703	-0.05836	-0.08271	-0.04633	-0.09541	0.02496	0.06580	-0.17096	2.059	-8.85418	0.19274	33.59	-1.42
10	0.04207	-0.05797	-0.08242	-0.04949	-0.12527	0.07691	0.00902	-0.14592	1.515	-8.67895	0.25314	29.92	-1.48
11	0.04181	-0.06275	-0.08124	-0.04717	-0.09456	0.02441	0.06408	-0.17049	2.056	-8.83313	0.19653	36.06	-2.41
12	0.05221	-0.06673	-0.07747	-0.05046	-0.12304	0.07456	0.01222	-0.16259	2.728	-8.57452	0.28909	32.39	**-2.48**
13	0.02499	-0.0574	-0.08149	-0.04793	-0.09333	0.03004	0.06836	-0.173	2.289	-9.06731	-0.56038	32.68	**0.14**
14	0.03031	-0.05731	-0.08098	-0.05124	-0.1239	0.08187	0.01189	-0.14727	3.15	-8.90012	-0.54414	29.01	0.08
15	0.02882	-0.05732	-0.08213	-0.0472	-0.0945	0.02774	0.06707	-0.1718	1.031	-8.95692	-0.0953	33.75	-0.37
16	0.03391	-0.05697	-0.08179	-0.05042	-0.12462	0.07955	0.01051	-0.14642	1.916	-8.78541	-0.0781	30.08	-0.44
17	0.02987	-0.05789	-0.08197	-0.04712	-0.09418	0.02708	0.06788	-0.17277	1.304	-8.94644	-0.12205	33.75	-0.37
18	0.03498	-0.05751	-0.08168	-0.05036	-0.12451	0.07924	0.01107	-0.14701	**0.989**	-8.77803	-0.11366	30.08	-0.44
19	0.02768	-0.06036	-0.08004	-0.04788	-0.09253	0.02497	0.07112	-0.17691	2.026	-8.88894	-0.22919	33.75	-0.37
20	0.03243	-0.06122	-0.07957	-0.05113	-0.12304	0.07833	0.01314	-0.14976	1.414	-8.73115	-0.23481	30.08	-0.44
21	0.03855	-0.06396	-0.07734	-0.04846	-0.09069	0.02557	0.07458	-0.19855	2.386	-8.88919	-0.06584	31.03	-1.03
22	0.04382	-0.06414	-0.07663	-0.05185	-0.12099	0.07780	0.01806	-0.17305	1.722	-8.72485	-0.02941	27.36	-1.09
23	0.02944	-0.05738	-0.08218	-0.04702	-0.09446	0.02687	0.06725	-0.17211	1.315	-8.93169	-0.00202	33.05	-0.65
24	0.03454	-0.05701	-0.08185	-0.05024	-0.12462	0.07889	0.01053	-0.14661	1.162	-8.76087	0.017709	29.38	-0.71
25	0.02228	-0.05485	-0.08226	-0.04713	-0.09427	0.02528	0.06790	-0.17261	2.099	-8.89068	-0.01486	33.05	-0.65
26	0.02683	-0.05594	-0.08182	-0.05021	-0.12443	0.07811	0.01057	-0.14568	1.587	-8.72347	-0.02399	29.38	-0.71
27	0.02702	-0.05774	-0.08156	-0.04778	-0.0934	0.02945	0.06794	-0.17298	2.713	-9.06173	-0.64713	32.97	-0.7
28	0.03224	-0.05745	-0.08112	-0.05106	-0.12383	0.08140	0.01135	-0.14717	3.628	-8.89265	-0.63097	29.3	-0.77
29	0.06110	-0.05471	-0.07929	-0.04892	-0.09116	0.03008	0.07046	-0.18216	**4.352**	-9.05409	-0.38822	32.03	-2.14
30	0.06239	-0.05496	-0.0806	-0.05163	-0.12497	0.08329	0.01087	-0.14236	2.843	-8.89458	-0.43903	28.36	-2.2
31	0.02724	-0.06238	-0.08075	-0.04783	-0.09336	0.02562	0.06698	-0.17103	2.33	-8.92099	-0.08797	30.41	-0.52
32	0.03249	-0.06222	-0.08022	-0.05119	-0.12387	0.07848	0.00942	-0.14414	1.374	-8.75256	-0.06095	26.74	-0.59
33	0.02881	-0.05653	-0.08358	-0.04779	-0.09377	0.02582	0.06793	-0.16904	1.546	-8.75881	-0.44362	37.3	-0.35
34	0.03384	-0.05628	-0.08324	-0.05095	-0.12364	0.07762	0.0113	-0.14399	1.951	-8.69931	-0.42279	33.63	-0.41
35	0.02881	-0.05653	-0.08358	-0.04779	-0.09376	0.02582	0.06794	-0.16906	1.545	-8.75864	-0.44346	37.3	-0.35
36	0.03384	-0.05628	-0.08324	-0.05095	-0.12364	0.07762	0.01130	-0.14399	1.95	-8.69931	-0.42282	33.63	-0.41

Table-3: Selected QSAR Equation Models and related Statistical Parameters r, r^2, SE and F_{count}/F_{table}.

Model	Variable	r	r^2	SE	F_{count}/F_{table}
1	qC_1, qN_2, qC_3, qC_4, qC_6, qC_7, qC_8, qC_9, Dipole moment, E_{HOMO}, E_{LUMO}, polarizability, Log P	0.940	0.884	0.15622	3.27387
2	qC_1, qN_2, qC_3, qC_4, qC_6, qC_7, qC_8, qC_9, Dipole moment, E_{HOMO}, polarizability, Log P	0.940	0.884	015096	3.84406
3	qC_1, qN_2, qC_3, qC_4, qC_6, qC_8, qC_9, Dipole moment, E_{HOMO}, polarizability, Log P	0.940	0.884	0.14647	4.48830
4	qC_1, qC_3, qC_4, qC_6, qC_8, qC_9, Dipole moment, E_{HOMO}, polarizability, Log P	0.940	0.883	0.14259	5.22619
5	qC_1, qC_3, qC_4, qC_6, qC_8, qC_9, Dipole moment, polarizability, Log P	0.939	0.882	0.13927	6.05378
6	qC_3, qC_4, qC_6, qC_8, qC_9, Dipole moment, polarizability, Log P	0.938	0.880	0.13646	7.02349
7	qC_3, qC_4, qC_6, qC_8, Dipole moment, polarizability, Log P	0.935	0.875	0.13593	7.94528
8	qC_4, qC_6, qC_8, Dipole moment, polarizability, Log P	0.935	0.873	0.13337	9.38868
9	qC_4, qC_6, qC_8, polarizability, Log P	0.931	0.867	0.13350	10.78670

Table-4: Coefficient of Independent Variables on Nine QSAR Equation Models selected.

Model	Independent variable coefficient													Constant (D)
	qC_1	qN_2	qC_3	qC_4	qC_6	qC_7	qC_8	qC_9	μ	E_{HOMO}	E_{LUMO}	α	Log P	
1	2.202	-12.559	-68.434	-152.871	-85.692	23.652	63.774	-3.030	-0.022	0.519	0.055	0.149	0.186	-22.999
2	2.631	-9.830	-62.823	-151.704	-88.157	16.528	57.925	-4.367	-0.028	0.481		0.148	0.189	-22.500
3	3.320	-7.477	-62.532	-190.037	-10.,079		54.054	-6.727	-0.033	0.190		0.150	0.199	-28.510
4	2.936		-42.405	-178.670	-105.231		51.911	-5.574	-0.040	0.221		0.154	0.200	-24.992
5	3.165		-40.281	-163.241	-103.287		48.255	-7.354	-0.047			0.159	0.204	-26.108
6			-43.833	-176.111	-110.667		52.713	-7.521	-0.045			0.157	0.163	-27.898
7			-10.867	-162.911	-107.057		53.621		-0.047			0.158	0.175	-23.019
8				-148.506	-109.075		53.782		-0.043			0.162	0.187	-21.763
9				-137.752	-121.663		59.608					0.166	0.199	-23.049

Table-5: PRESS Values from the selected QSAR Equation Models

Compound	pEC_{50}	pEC_{50} Prediction								
	Experiment	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Fitting compounds										
1	3.922	3.984042	3.987309	3.989482	3.988129	3.980931	3.969738	3.966798	3.963814	3.955821
2	4.172	4.067607	4.068036	4.06088	4.059051	4.047373	4.033489	4.025084	4.01932	4.021351
3	4.372	4.207334	4.203331	4.206099	4.213942	4.201787	4.216812	4.189517	4.199103	4.256534
4	4.21	4.321536	4.3205	4.329069	4.339408	4.342396	4.346452	4.312823	4.313541	4.284228
5	4.172	4.276875	4.279478	4.288354	4.289318	4.289448	4.263156	4.26127	4.267778	4.281162
8	4.375	4.24993	4.253463	4.257202	4.257796	4.244856	4.232948	4.248403	4.277098	4.300525
9	3.896	4.159981	4.157407	4.155081	4.146167	4.1486	4.150152	4.137244	4.135594	4.142906
10	4.467	4.266816	4.266449	4.257566	4.253139	4.258341	4.255305	4.235364	4.22733	4.206824
11	4.359	4.253385	4.251387	4.261371	4.252733	4.259368	4.272371	4.29247	4.289624	4.265552
12	4.333	4.429973	4.430525	4.425096	4.424937	4.428435	4.427473	4.431869	4.46051	4.470974
13	4.556	4.392134	4.393062	4.389359	4.388479	4.409657	4.401303	4.419191	4.419592	4.422713
14	4.559	4.549352	4.546126	4.550604	4.542511	4.552006	4.549581	4.566117	4.566481	4.611609
16	4.441	4.573431	4.569142	4.571615	4.573715	4.585526	4.57852	4.579027	4.577331	4.57731
17	4.362	4.46178	4.461833	4.463683	4.468296	4.486957	4.474254	4.482237	4.485346	4.460942
19	4.504	4.510497	4.509145	4.511002	4.519563	4.510473	4.512669	4.550608	4.563239	4.560035
20	4.625	4.625186	4.624109	4.615316	4.62424	4.620639	4.61894	4.667128	4.674862	4.640913
21	4.161	4.09735	4.095872	4.092889	4.082177	4.08186	4.08639	4.040941	4.055553	4.040065
22	4.234	4.27316	4.273992	4.278353	4.27239	4.272916	4.277607	4.238234	4.2482	4.204184
23	4.498	4.288664	4.286655	4.28674	4.291076	4.303128	4.30331	4.304815	4.301911	4.272949
25	4.234	4.255811	4.258695	4.269826	4.287116	4.279504	4.307697	4.301161	4.298061	4.303093
26	4.286	4.355397	4.356337	4.34603	4.362527	4.358673	4.38189	4.386789	4.380671	4.360737
27	4.221	4.222776	4.228788	4.220774	4.21498	4.234648	4.25131	4.257755	4.25231	4.265287
28	4.27	4.359355	4.360325	4.354441	4.342394	4.350345	4.372204	4.375971	4.369899	4.425087
29	3.817	3.865799	3.864889	3.856624	3.873584	3.875525	3.85207	3.83808	3.822161	3.859863
30	4.21	4.164136	4.164237	4.175109	4.179526	4.174039	4.169347	4.197528	4.168641	4.174953
31	3.766	3.835514	3.834413	3.835519	3.814441	3.793387	3.799258	3.845961	3.840438	3.821972
35	5.161	5.19047	5.190261	5.176553	5.172681	5.147788	5.141724	5.121207	5.109544	5.097749
36	5.276	5.22071	5.223234	5.234364	5.224684	5.220392	5.213032	5.185404	5.171048	5.173662
Test compounds										
6	5.086	4.351774	4.353221	4.351681	4.353202	4.349732	4.320363	4.312983	4.316274	4.334752
7	4.338	4.045057	4.043317	4.057702	4.04805	4.037327	4.039222	4.063595	4.061016	4.063833
15	4.772	4.466116	4.466546	4.473211	4.481371	4.503188	4.49329	4.500838	4.500637	4.462798
18	4.272	4.608134	4.610842	4.614805	4.620403	4.638309	4.626032	4.629285	4.626534	4.588974
24	5.337	4.418243	4.417743	4.416865	4.422237	4.43262	4.432057	4.428841	4.421029	4.384517
32	4.315	3.984846	3.988537	3.991785	3.979689	3.961701	3.964476	4.062911	4.008327	3.945642
33	4.646	5.191066	5.190913	5.177479	5.173579	5.148704	5.142722	5.122205	5.110574	5.099005
34	4.479	5.220944	5.223446	5.234607	5.224911	5.220583	5.213245	5.185619	5.171246	5.173876
	PRESS internal	0.341671	0.341824	0.343248	0.34562	0.349127	0.353825	0.369525	0.373549	0.392071
	PRESS external	2.632149	2.634661	2.629111	2.610861	2.584317	2.606126	2.488418	2.496614	2.567334
	PRESS total	2.97382	2.976485	2.972358	2.956481	2.933444	2.959951	**2.857943**	2.870162	2.959405

Table-6: Predicted Activity Values (pEC_{50}) of the best QSAR Equation

Compounds No.	pEC_{50} Experiment	pE_{50} Prediction
Fitting Compound		
1	3.922	3.966798
2	4.172	4.025084
3	4.372	4.189517
4	4.21	4.312823
5	4.172	4.26127
8	4.375	4.248403
9	3.896	4.137244
10	4.467	4.235364
11	4.359	4.29247
12	4.333	4.431869
13	4.556	4.419191
14	4.559	4.566117
16	4.441	4.579027
17	4.362	4.482237
19	4.504	4.550608
20	4.625	4.667128
21	4.161	4.040941
22	4.234	4.238234
23	4.498	4.304815
25	4.234	4.301161
26	4.286	4.386789
27	4.221	4.257755
28	4.27	4.375971
29	3.817	3.83808
30	4.21	4.197528
31	3.766	3.845961
35	5.161	5.121207
36	5.276	5.185404
Test compound		
6	5.086	4.312983
7	4.338	4.063595
15	4.772	4.500838
18	4.272	4.629285
24	5.337	4.428841
32	4.315	4.062911
33	4.646	5.122205
34	4.479	5.185619

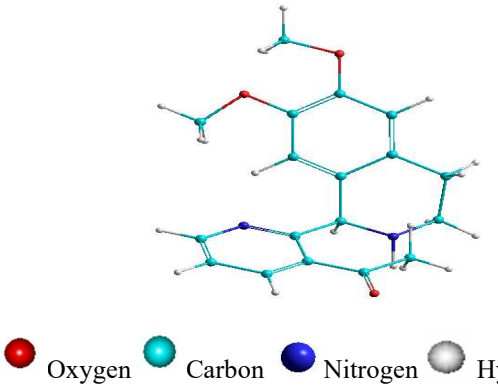


Fig.-4: Compounds 1- (6-acetylpyridine-2-yl) -7, 8-dimethoxyisoquinoline

Table-3 showed the new compound model of 1-aryl-tetrahydroisoquinoline derivative, which has anti-HIV activity (pEC_{50}) obtained from the higher QSAR equation model or has a smaller EC_{50} is designed the compound number 40, 1- (6- acetylpyridine-2-yl)-7,8-dimethoxyisoquinoline (Figure 4). EC_{50} is a sufficient concentration that can protect cells against HIV as much as 50% obtained through the regression equation. The parameter is expected to get the smallest EC_{50} value because the smaller the EC_{50} value of a compound design model, the compound model is more active as a protective cell against HIV or anti-HIV viruses.

Table-7: Predictions of Theoretical Activity Values (pEC_{50}) of New Compounds of 1-aryl-tetrahydroisoquinoline Derivatives

Compound No.	R	Ar	pEC_{50}	EC_{50}
37	-OMethyl	3-acetylpyridine-2-yl	4.378647005	41.82
38	-OMethyl	4- acetylpyridine-2-yl	4.369956395	42.66
39	-OMethyl	5- acetylpyridine-2-yl	4.309194746	49.07
40	-OMethyl	6- acetylpyridine-2-yl	5.601409924	2.5
41	-OMethyl	3- triflorometilpiridin-2-yl	4.135987085	73.11
42	-OMethyl	4- triflorometilpiridin-2-yl	4.224612048	59.62
43	-OMethyl	5- triflorometilpiridin-2-yl	4.328419163	46.95
44	-OMethyl	6- triflorometilpiridin-2-yl	4.905651272	12.43
45	-OMethyl	4-methoxypyridine-2-yl	3.998226618	100.41
46	-OMethyl	4-aminophenyl	3.819192991	151.64
47	-OMethyl	4- ethoxyphenyl	4.475821636	33.43
48	-OMethyl	2,3-ethylene-dioxy-phenyl	4.058111494	87.48
49	-OMethyl	4-nitrofenyl	3.775391774	167.73
50	-OMethyl	4-aminometilfenyl	4.181124576	65.9
51	-OMethyl	4-aminodimetilfenyl	4.565107059	27.22
52	-OMethyl	4-hydroxy phenyl	3.854536163	139.79
53	-OMethyl	4-acetylaminophenyl	4.417623942	38.23
54	-OMethyl	4-ethoxy-pyridin-2-yl	4.348081568	44.87
55	-OMethyl	4-acetyloxyphenyl	4.383062730	41.39

Compounds 1-(6-acetylpyridine-2-yl)-7,8-dimethoxyisoquinoline have better activity because the 6-acetylpyridine-2-yl substituent group, which acts as an electron-driving group that affects the aromatic substitution reactivity. The 6-acetylpyridine-2-yl group, as an activator, causes the aromatic ring to become more reactive so that it will increase its activity as an anti-HIV drug. The design compound model has a theoretically better activity value than the 1-aryl-tetrahydroisoquinoline derivative compound. Thus, compounds 1- (6-acetylpyridine-2-yl) -7, 8-dimethoxyisoquinoline can be proposed as a new model of anti-HIV drug compounds that can be synthesized and tested in laboratory activities.

CONCLUSION

Based on the research results obtained can be concluded that the results of the analysis of the QSAR study of the derivative series of 1-aryl-tetrahydroisoquinoline compounds using atomic charges, and other molecular properties as descriptors using the PM3 semi-empirical calculation method produce the best QSAR equation model as follows: pEC_{50} = -23.019 – 10.867 (qC_3)– 162.911 (qC_4) – 107.057 (qC_6) + 53.621 (qC_8) - 0.047 (μ) + 0.158 (α) + 0.175 (log P) (n = 36; r = 0.935; r^2) = 0.875; SE = 0.135927366; F_{count}/F_{table} = 7.945277618; PRESS = 2.857943). The new model of the 1-aryl-tetrahydroisoquinoline derivative from the design which is expected to have better activity than the 1-aryl-tetrahydroisoquinoline derivative compound that already exists is 1- (6-acetylpyridine-2-yl) -7,8-dimethoxyquinoline.

REFERENCES

1. J. W. Aunsborg, B. L. Hønge, S., Jespersen, F. Rudolf, C. Medina, F. G. Correira, I. S. Johansen, and C. Wejse, *International Journal of Infectious Diseases*, **92**, S78 **(**2020), https://doi.org/10.1016/j.ijid.2020.03.012
2. F. Berhal, Z. Wu, Z. Zhang, T. Ayad, and V. Ratovelomanana-Vidal, *Organic Letters*, **14**, 3308(2012), https://doi.org/10.1021/ol301281s

3. P. Cheng, N. Huang, Z.Y. Jiang, Q. Zhang, Y.T. Zheng, J.J. Chen, X. M. Zhang, and Y. B. Ma, *Bioorganic & Medicinal Chemistry Letters*, **18**, 2475(2008), https://doi.org/10.1016/j.bmcl.2008.02.040
4. T. D. Frank, A. Carter, D. Jahagirdar, M. H. Biehl, D. Douwes-Schultz, S. L. Larson, M. Arora, L. Dwyer-Lindgren, K. M. Steuben, H. Abbastabar, L. J. Abu-Raddad, D. M. Abyu, M. Adabi, O. M. Adebayo, V. Adekanmbi, O. O. Adetokunboh, A. Ahmadi, K. Ahmadi, Ahmadian, and C. J. L. Murray, *The Lancet HIV*, **6**, e831(2019), https://doi.org/10.1016/S2352-3018(19)30196-1
5. F. Galimberti, A. Moretto, and E. Papa, *Water Research*, **174**, 115583(2020), https://doi.org/10.1016/j.watres.2020.115583
6. J. C. Gamble-George, C. T. Longenecker, A. R. Webel, D. H. Au, A. F. Brown, H. Bosworth, K. Crothers, W. E. Cunningham, K. A. Fiscella, A. B. Hamilton, C. D. Helfrich, J. A. Ladapo, A. Luque, J. N. Tobin, and G. E. Wyatt, *Progress in Cardiovascular Diseases*, **63**, 79(2020), https://doi.org/10.1016/j.pcad.2020.03.006
7. M. J. Gartner, M. Roche, M. J. Churchill, P. R. Gorry, and J. K. Flynn, *EBioMedicine*, **53**, 102682(2020), https://doi.org/10.1016/j.ebiom.2020.102682
8. R. Hadanu, *Indonesian Journal of Pharmacy*, **30**, 167(2019) http://dx.doi.org/10.14499/indonesianjpharm30iss3pp167
9. R. Hadanu, S. Idris, and I. W. Sutapa, *Indonesian Journal of Chemistry*, **15**, 86(2015), https://doi.org/10.22146/ijc.21228
10. M. Haghighatlari, and J. Hachmann, *Current Opinion in Chemical Engineering*, **23**, 51(2019), https://doi.org/10.1016/j.coche.2019.02.009
11. G. S. Hassan, H. H. Georgey, E. Z. Mohammed, and F. A. Omar, *European Journal of Medicinal Chemistry*, **184**, 111747(2019), https://doi.org/10.1016/j.ejmech.2019.111747
12. G. Hautier, *Computational Materials Science*, **163,** 108(2019), https://doi.org/10.1016/j.commatsci.2019.02.040
13. S. Ickowicz, E. Wood, H. Dong, P. Nguyen, W. Small, T. Kerr, J. S. G. Montaner, and M. J. Milloy, *Drug and Alcohol Dependence*, **180**, 33(2017), https://doi.org/10.1016/j.drugalcdep.2017.07.016
14. N. Irmayati, S. Yona, and A. Waluyo, *Enfermería Clínica*, **29**, 546(2019), https://doi.org/10.1016/j.enfcli.2019.04.084
15. S. Marquina, M. Maldonado-Santiago, J. N. Sánchez-Carranza, M. Antúnez-Mojica, L. González-Maya, R. S. Razo-Hernández, and L. Alvarez, *Bioorganic & Medicinal Chemistry*, **27**, 43(2019), https://doi.org/10.1016/j.bmc.2018.10.045
16. I. Miladiyah, I. Tahir, Jumina, S. Mubarika, and M. Mustofa, *Molekul*, **11**, 143(2016), http://dx.doi.org/10.20884/1.jm.2016.11.1.203
17. Pudjono, S. S. Susilowati, and Rehana, *Indonesian Journal of Pharmacy*, **22**, 144(2011), http://dx.doi.org/10.14499/indonesianjpharm0iss0pp144-150
18. Y. Tian, S. Zhang, H. Yin, and A. Yan, *Chemometrics and Intelligent Laboratory Systems*, **196,** 103888(2020), https://doi.org/10.1016/j.chemolab.2019.103888
19. L. Yang, Y. Wang, W. Hao, J. Chang, Y. Pan, J. Li, and H. Wang, *Ecotoxicology and Environmental Safety*, **193,** 110352(2020), https://doi.org/10.1016/j.ecoenv.2020.110352
20. H. Zhang, D. E. Zembower, Z. Chen, *Bioorganic & Medicinal Chemistry Letters*, **7**, 2687(1997), https://doi.org/10.1016/S0960-894X(97)10057-9
21. T.H. Sugara, Jumina, E.N. Solikhah and H.D. Pranowo, *Rasayan Journal of Chemistry*, **14**, 1927 (2021), http://doi.org/10.31788/RJC.2021.1436214

[RJC-6479/2021]