ISSN: 1410-8917
Jurnal Kimia Sains & Aplikasi
e-ISSN: 2597-9914

Jurnal Kimia Sains dan Aplikasi 27 (7) (2024): 316-327

Jurnal Kimia Sains dan Aplikasi
Journal of Scientific and Applied Chemistry

Journal homepage: http://ejournal.undip.ac.id/index.php/ksa

QSAR, Molecular Docking, and Molecular Dynamic of Novel Coumarin Derivatives as α-Glucosidase Inhibitor

Mutista Hafshah [1,*], Irvan Maulana Firdaus [2], Ika Nur Fitriani [3], Lutfiah Rahmawati [1]

[1] Department of Chemistry, Faculty of Science and Technology, Universitas Islam Negeri Walisongo, Semarang, 50181, Indonesia
[2] Al Irsyad Al Islamiyyah Boarding School Purwokerto, Banyumas 53113, Indonesia
[3] School of Chemistry, Faculty of Science, The University of Sydney, Camperdown NSW 2050, Australia

* Corresponding author: mutista.hafshah@walisongo.ac.id

https://doi.org/10.14710/jksa.27.7.316-327

Article Info

Article history:

Received: 18th January 2024
Revised: 13th June 2024
Accepted: 19th June 2024
Online: 31st July 2024

Keywords:
Alpha-glucosidase inhibitor; Coumarin; Molecular dynamic; Molecular docking; QSAR

Abstract

Diabetes mellitus (DM) is a chronic metabolic disorder posing a significant health risk. Effective treatments are continually sought. Coumarin derivatives with oxime ester groups have shown potential as antidiabetic agents by inhibiting the α-glucosidase enzyme, a key player in glycoprotein metabolism and postprandial hyperglycemia control. This makes lysosomal α-glucosidase a promising therapeutic target. A study used 28 coumarin derivatives with known α- glucosidase inhibitory IC_{50} values for computer-assisted drug design (CADD) through quantitative structure-activity relationship (QSAR) analysis, yielding a statistically robust equation for guiding new compound development. Subsequently, eleven new coumarin derivatives with oxime ester groups were synthesized, showing enhanced α-glucosidase inhibitory activity compared to the initial set. Molecular docking assays indicated that compounds 32, 37, 38, and 39 had lower free energy values than the native ligand, suggesting higher stability in target protein interactions. Notably, compound 38 had the lowest free energy (-8.351) and demonstrated lower root mean square deviation (RMSD) and root mean square fluctuation (RMSF) than the original ligand, indicating greater stability. The QSAR equation derived is:

$$\text{Log } IC_{50} = 2.886 - 0.054\,(LUMO) + 0.073\,(\mu) - 0.148\,(\alpha) - 0.046\,(RD) + 0.046\,(BM) + 0.001\,(V_{vdw}) - 0.421\,(qC2) + 1.138\,(qC8) - 0.092\,(qC9) + 2.61\,(qC10) + 1.354\,(qN1) \quad \text{(Eq 1)}$$

n=28; R=0.918; R^2=0.843; SD=0.196; F calc/F tab=3.169; Sig =<0.01; PRESS = 1.376.

Compound 38's SMILES notation is:

C\C(=N/OC(=O)\C=C/C1=CC=C(Br)C=C1)C1=CC2=CC(O)=C(CC(O)=O)C=C2OC1=O).

1. Introduction

Diabetes mellitus (DM) causes hyperglycemia due to hormonal abnormalities influencing insulin production [1, 2, 3]. The pancreas regulates blood glucose using insulin [4]. DM has two types: Type I, which kills pancreatic beta cells and requires insulin injections, and Type II, which reduces insulin release and raises blood glucose [2, 4]. High hyperglycemia can induce kidney failure, heart attack, and eyesight loss. The International Diabetes Federation (IDF) estimated 9.3% of 483 million persons aged 20–79 had DM in 2019. This number will climb to 578 million by 2030 and 700 million by 2045. Indonesia has 10.7 million diabetics, placing it among the top 10 countries with the greatest frequency. Better prevention and treatment strategies are needed.

Treatment varies by diabetes type. Insulin treats Type I diabetes, while blood glucose-lowering medicines treat Type II. α-Glucosidase inhibitors that convert carbs into glucose treat Type II diabetes. Coumarin derivatives are promising natural and synthetic α-glucosidase inhibitors [5, 6, 7, 8]. Coumarins, found in Umbelliferae/Apiaceae and Rutaceae [9]. The molecular structure of a chemical has an impact on its biological action [10, 11].

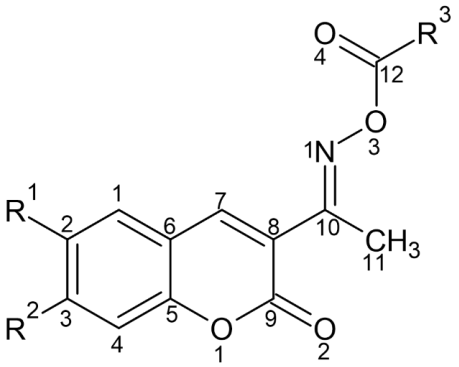


Figure 1. Coumarin compounds with oxime esters atomic numbering

Modifying the chemical structure of coumarins enhances their effectiveness [12]. In order to identify potential drugs, one can alter naturally occurring molecules that have already demonstrated activity and employ quantitative structure-activity relationship (QSAR) studies [13] to establish a connection between chemical structures and biological effects. Coumarin and its derivatives exhibit interaction with α-glucosidase, indicating its potential inhibitory effect on it. A total of 28 coumarin derivatives with oxime ester groups were synthesized and evaluated for their ability to block α- glucosidase [14]. Fourteen compounds exhibited IC_{50} values lower than 10 µM, indicating significant potency.

QSAR equation, produced by optimization of molecular modeling geometry using quantum mechanics and semiempirical methods [15, 16, 17, 18, 19, 20, 21, 22, 23, 24], predicts IC_{50} values, and helps design new coumarin derivative antidiabetic medicines. Data pertaining to the compound's physical and chemical properties, commonly referred to as descriptors, are important for deriving the QSAR equation.

It focused on coumarin compounds' unusual structure, including ester oxime groups, which may boost their biological activity, unlike earlier investigations [25, 26]. In 2022, Zhang *et al.* [14] synthesized a coumarin derivative containing an oxime ester group, demonstrating potential as an α-glucosidase inhibitor with an IC_{50} of 2.540 µ new coumarin compounds containing ester oxime groups were created and tested via molecular docking with the Homo sapiens hydrolase enzyme (PDB: 2OLE). Molecular docking analysis determines compound-target protein interactions, with reduced free energy indicating stability. RMSD (≤ 3 Å) [27, 28] and RMSF values are essential for molecular docking validation and protein-ligand binding stability. Molecular dynamics simulations can monitor protein and ligand binding conformational changes alongside molecular docking [29, 30, 31]. This unique approach uses molecular docking and molecular dynamics analysis to improve in silico analysis data and identify α-glucosidase inhibitors from coumarin derivative compounds with oxime ester groups.

2. Experimental

2.1. Tools and Materials

The research utilized an Intel Core i5 computer processor, 8 GB of Random Access Memory (RAM), and the Windows 10 operating system. The software tools employed were Hyperchem version 8.0, SPSS version 25, JASP, PLANTS (Protein-Ligand ANTSystem), Autodock Vina, and SketchChem. The study included 28 coumarin compounds (Figure 1), and the IC_{50} values were derived from Zhang *et al.* [14]. The compounds used are categorized as 21 fitting compounds and 7 test compounds (Table 1). The selection of 7 test compounds is based on the steric representation of 28 coumarin derivative compounds with ester oxime groups.

A selection of representatives with varying degrees of steric hindrance, including mild, medium, and substantial hindrances, were chosen as test sets [32]. Fitting compounds are used to determine the equation, whereas test compounds are used to validate the resulting equation. The equation is validated by calculating the predicted residual error sum of squares (PRESS) value from the experimental log IC_{50} data and the predicted log IC_{50} data of 7 test compounds.

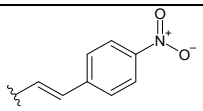
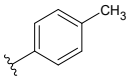
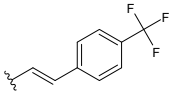
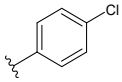
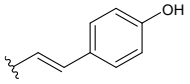
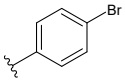
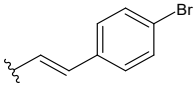
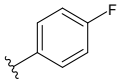
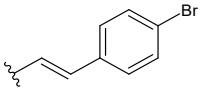
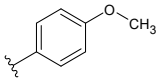
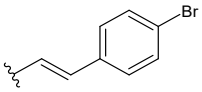
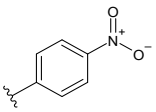
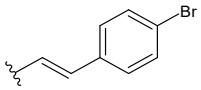
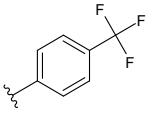
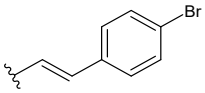
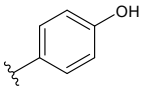
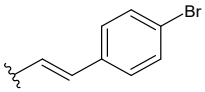
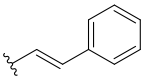
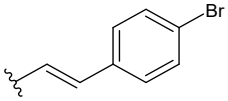
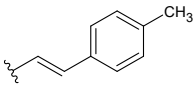
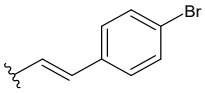
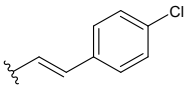
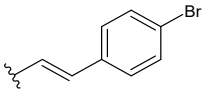
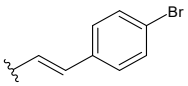
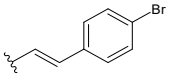
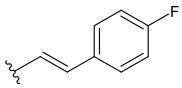
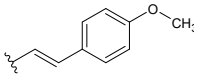
2.2. Geometric Optimization

HyperChem 8.0.8 was used to create a two-dimensional model of the molecular structure of molecules that are derivatives of coumarins. Then, hydrogen atoms were added to each of the structure's constituent atoms. Geometric principles were applied to construct each structure to minimize molecular energy and produce the most stable conformation possible. The convergence limit was established at 0.001 kcal/mol using the Polak-Ribiere algorithm, and the optimization calculations were performed using the PM3 semiempirical technique. The atomic charge (qC1-qC12, qO1-qO4, and qN1), dipole moment (µ), log P, polarizability (α), molecular weight, van der Waals area (A_{vdw}), van der Waals volume (V_{vdw}), refractive index (RD), and the HOMO and LUMO energies were among the descriptors used. Descriptors were included in Table 2, along with instructions on how to retrieve them [24].

2.3. Determining Descriptors and Correlation Analysis

The analysis involved a bivariate correlation utilizing the two-tailed technique and Pearson's coefficient. At this point, each descriptor was carefully examined to evaluate its degree of association with log IC_{50} activity. Next, the atomic charge (qC1-qC12, qO1-qO4, and qN1), dipole moment (µ), log P, polarizability (α), molecular weight, van der Waals Area (A_{vdw}), van der Waals volume (V_{vdw}), refractive index (RD), and the HOMO and LUMO energies were selected. A strong relationship indicates a significant impact on log IC_{50} activity. Afterward, the chosen descriptors were utilized as independent variables in Multiple Linear Regression (MLR) studies [33].

Table 1. Coumarin compound derivatives

Compound number	R^1	R^2	R^3	IC_{50}	Log IC_{50}	Compound number	R^1	R^2	R^3	IC_{50}	Log IC_{50}
1	H	H		93.230	1.969	16	H	H		8.710	0.940
2	H	H		15.910	1.201	17	H	H		19.66	1.293584
3	H	H		10.950	1.039	18	H	H		21.540	1.333
4	H	H		6.460	0.810	19	H	Cl		3.420	0.534
5	H	H		29.210	1.466	20	H	Br		2.540	0.405
6	H	H		19.340	1.286	21	H	F		3.920	0.593
7	H	H		26.120	1.417	22	H	CH_3		5.240	0.719
8	H	H		26.160	1.418	23	H	OCH_3		6.430	0.808
9	H	H		26.140	1.417	24	Cl	H		4.250	0.628
10	H	H		34.350	1.536	25	Br	H		3.810	0.581
11	H	H		13.390	1.127	26	F	H		3.730	0.572
12	H	H		6.340	0.802	27	CH_3	H		6.170	0.790
13	H	H		5.060	0.704	28	OCH_3	H		7.840	0.894
14	H	H		11.550	1.063						
15	H	H		15.090	1.179						

2.4. Determining the QSAR Equation

The most efficient QSAR equation was determined using multilinear regression statistical analysis. The JASP application was utilized to analyze 21 compounds. These compounds were determined to be appropriate for the dependent variable log IC_{50}, while the independent variable was produced via the correlation test [34].

2.5. Development of Novel Coumarin Derivatives

The satisfactory completion of the molecular design was achieved by modifying the substituents at positions R1, R2, and R3. Afterward, the molecules that have been generated undergo geometry optimization and descriptor computation calculations. We employed the most precise QSAR equation to calculate the theoretical activity of the chemical, represented by the notation log IC_{50}. The compounds exhibiting the lowest log IC_{50} values had the highest levels of anti-diabetic efficacy [35].

Table 2. Descriptors and methods of determination

No	Symbol	Descriptors	Unit	Methods of determination
1	q	Net atomic charge of atoms C1-C12 and atoms O1-O4 and N1	Coulomb	Semiempirical PM3, Hyperchem, optimization of coumarin compounds with oxime ester groups
2	μ	Dipole moment	Debye	
3	Log P	n-octanol/water partition coefficient	-	
4	α	Molecular polarizability	$Å^3$	
5	BM	Molecular weight	s.m.a	QSAR properties, Hyperchem
6	A_{vdw}	Van der Waals surface area	$Å^2$	
7	V_{vdw}	Van der Waals volume	$Å^3$	
8	RD	Refractive index	Å	
9	E_{HOMO}	HOMO energy	eV	Orbitals, Hyperchem

2.6. Molecular Docking

Molecular docking involves several steps: protein and ligand preparation, addition of hydrogen atoms and charges to proteins and ligands, determination of grid box, and actual molecular docking [36]. The protein and ligand preparation process commences with The protein code utilized in this investigation, 2OLE, obtained from the website https://www.rcsb.org/structure/2OLE and stored in PDB format. Dehydration was performed to obtain a protein molecule, and the initial ligand was extracted, resulting in the preservation of a protein molecule. This molecule was then recorded under the file name 2OLE_protein.pdb. Subsequently, a preparatory procedure was conducted to isolate the native ligand molecule exclusively. Next, the file (2OLE_ligand.pdb) in the current director was stored. Hydrogen atoms and charges were introduced to the proteins and ligands. Specifically, the protein molecule (2OLE_protein.pdb) and the ligand molecule (2OLE_ligand.pdb) were modified by adding hydrogen atoms.

The protein and ligand were assigned Gasteiger charges and saved in pdbqt format, leading to the generation of a protein file named 2OLE_protein.pdbqt and a ligand file named 2OLE_ligand.pdbqt. Grid box determination The AutodockTools software was used to determine the dimensions of the grid box by changing the spacing to 1 Å. The measurements of the grid box were stored under the file name "Grid.txt". The docking procedure was executed in the current directory, including the 2OLE_protein.pdbqt, 2OLE_ligand.pdbqt, and Grid.txt files. Generate a Conf.txt file that includes the specified details: receptor = 2OLE_protein.pdbqt, ligand = 2OLE_ligand.pdbqt, center x,y,z based on the grid box dimensions, size_x, y, z based on the grid box dimensions, num_modes = 10, and exhaustiveness = 32. Launch Windows Power Shell in the current directory and input the command vina.exe --config conf.txt --cpu 4. Save the docking results in a text format and then separate the ligand from the docking results using the command vina_split.exe –2OLE_ligand_out.pdbqt. Choose a ligand with the highest stability level in terms of affinity value.

2.7. Molecular Dynamics

A molecular dynamics simulation was performed on the compound, which was predicted to be the best in each activity. The process involves preparing the ions.mdp, em.mdp, nvt.mdp, npt.mdp, and md.mdp files. Then, the preparation and merging of receptor and ligand topologies, addition of ions, energy minimization, temperature and pressure equilibration, production simulation, analysis, and visualization of results (RMSD, RMSF) are performed [36].

3. Results and Discussion

This study examines the quantitative correlation between coumarin derivative architectures and their antidiabetic effectiveness. Using Hyperchem 8.0.8, 28 coumarin oxime ester compounds were geometrically optimized to find the most stable molecular structure by minimizing potential energy. The semi-empirical PM3 method was chosen for this optimization, as it uses experimental data parameterization and simplified integral computations for faster and more accurate predictions compared to the DFT method for complex organic compounds.

The DFT approach requires a large database since optimizing geometry and performing computations is time-consuming. Therefore, the semi-empirical method is a very efficient strategy for developing new organic compounds to improve their capabilities [37]. Multiple studies have been carried out to analyze the QSAR of organic molecules. This research utilized the semi-empirical PM3 approach as a geometric optimization technique to calculate the descriptor values of the molecules, among other factors. The following studies were conducted by [29, 30, 38, 39, 40, 41, 42, 43, 44, 45]. The semi-empirical PM3 approach optimizes geometry using the Polak-Ribiere algorithm at the ground state with an RMS gradient of 0.001 kcal/A·mol. The optimized compounds have larger structures due to the inclusion of oppositely charged groups, resulting in increased spatial separation from the charge distribution around atoms and groups.

Table 3. Results of correlation test analysis with JASP

Descriptor	Correlation value	Significance value
BM (amu)	-0.851	0.001
α (Å)	-0.825	0.001
RD (Å)	-0.809	0.001
qN1	-0.67	0.001
qO3	0.607	0.001
HOMO	-0.6	0.001
V_{vdw} (Å)	-0.593	0.001
qC8	-0.588	0.001
LUMO	0.577	0.001
μ (Debye)	0.527	0.001
A_{vdw} (Å) grid	-0.484	0.009
qC10	0.48	0.01
qC1	0.46	0.014
qC2	-0.39	0.04
qC9	0.383	0.044
qO1	0.37	0.053
qC6	-0.319	0.098
Log P	-0.304	0.116
qO4	-0.218	0.265
qC5	-0.2	0.308
qC11	-0.172	0.383
qC7	-0.165	0.401
qC4	0.159	0.418
qC3	0.097	0.623
qO2	-0.065	0.744
A_{vdw} (Å) approx.	0.041	0.837
qC12	0.01	0.959

Table 4. Six equations from backward analysis

Model	Equation
1	8.885 + 0.289 (HOMO) – 0.925 (LUMO) + 0.04 (μ) – 0.167 (α) – 0.081 (RD) – 0.003 (BM) + 0.012 (V_{vdw}) + 1.261 x 10^{-5} (A_{vdw}) + 5.982 (qC1) – 1.657 (qC2) + 13.803 (qC8) + 3.289 (qC9) + 6.471 (qC10) + 6.487 (qN1) + 9.331 (qO3)
2	8.885 + 0.289 (HOMO) – 0.925 (LUMO) + 0.04 (μ) – 0.167 (α) – 0.081 (RD) – 0.003 (BM) + 0.012 (V_{vdw}) + 5.98 (qC1) – 1.657 (qC2) + 13.803 (qC8) + 3.288 (qC9) + 6.471 (qC10) + 6.485 (qN1) + 9.329 (qO3)
3	4.659 – 0.655 (LUMO) + 0.045 (μ) – 0.138 (α) – 0.057 (RD) – 0.003 (BM) + 0.009 (V_{vdw}) + 3.434 (qC1) – 1.572 (qC2) + 9.932 (qC8) + 2.285 (qC9) + 4.944 (qC10) + 5.463 (qN1) + 6.862 (qO3)
4	3.892 – 0.469 (LUMO) + 0.067 (μ) – 0.123 (α) – 0.052 (RD) – 0.002 (BM) + 0.008 (V_{vdw}) – 2.228 (qC2) + 7.142 (qC8) + 1.667 (qC9) + 4.223 (qC10) + 3.874 (qN1) + 3.834 (qO3)
5	3.741 - 0.298 (LUMO) + 0.062 (μ) – 0.139 (α) – 0.036 (RD) – 0.002 (BM) + 0.006 (V_{vdw}) – 2.038 (qC2) + 5.067 (qC8) + 1.339 (qC9) + 3.799 (qC10) + 3.174 (qN1)
6	3.946 – 0.288 (LUMO) + 0.052 (μ) – 0.132 (α) – 0.048 (RD) + 0.006 (V_{vdw}) -2.066 (qC2) + 5.261 (qC8) + 1.523 (qC9) + 3.615 (qC10) + 3.038 (qN1)

Table 5. Value R, R^2, SD, F calc/F table, and PRESS from six equations

Equation	R	R^2	SD	F calc/F table	PRESS
1	0.991	0.982	0.084	1.923	24.826
2	0.991	0.982	0.076	3.160	24.843
3	0.987	0.974	0.084	3.278	17.739
4	0.982	0.964	0.093	3.241	7.677
5	0.980	0.960	0.094	3.782	5.227
6	0.978	0.956	0.094	4.453	6.758

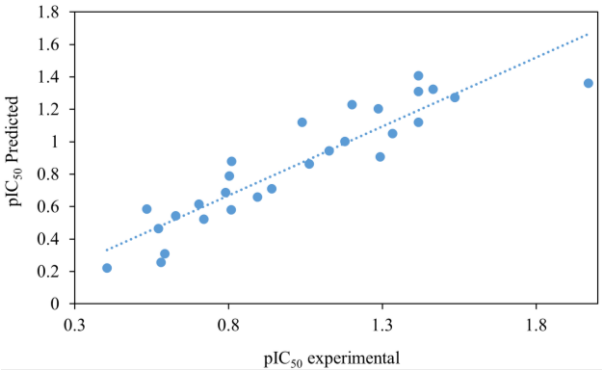


Figure 2. Correlation of experimental log IC_{50} with predicted log IC_{50}

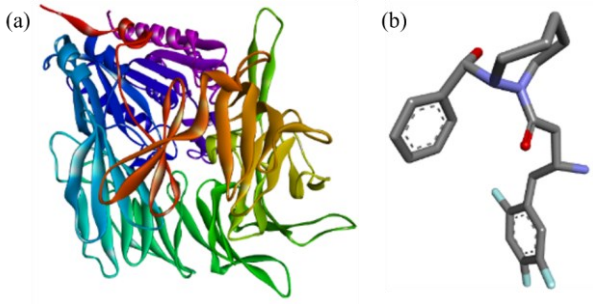


Figure 3. (a) Protein and (b) native ligand

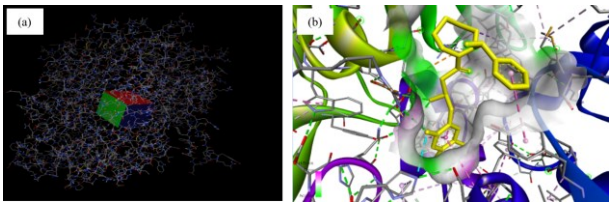


Figure 4. (a) Grid box, (b) docking protein and native ligand

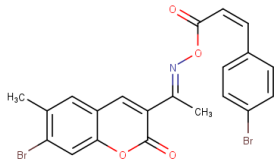
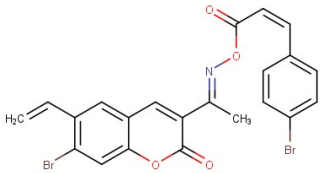
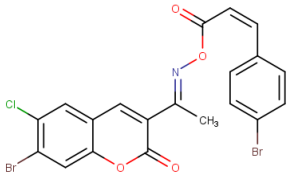
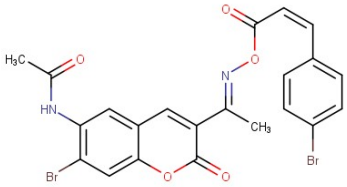
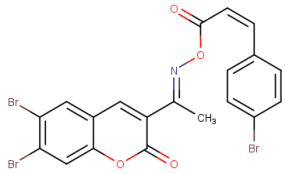
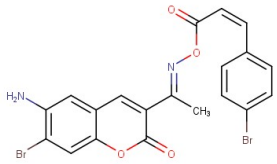
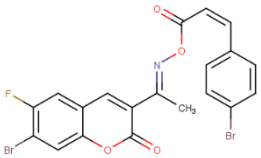
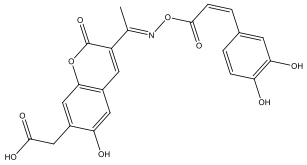
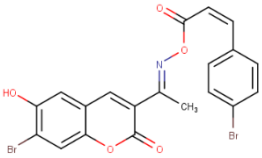
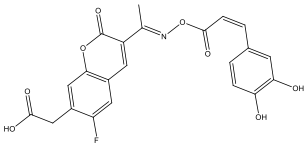
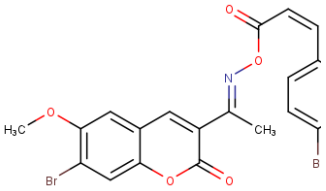
The optimized structure was assessed for its physicochemical properties using the descriptors (Table 2). A statistical correlation approach is employed to evaluate the 27 descriptors. Correlation studies investigated the relationship between descriptors and their anti-diabetic activity. The test allows for identifying the characteristics that exert the greatest influence on the IC_{50} value of coumarin derivative compounds containing oxime ester groups, which are used as anti-diabetic agents. Correlation analysis was conducted using the two-tailed approach and Pearson correlation coefficient. Correlation analysis results are utilized to select descriptors as independent variables in MLR statistical calculations. The results of the correlation test between variables and log IC_{50} are presented in Table 3.

Based on the correlation study, only 15 of the 27 descriptors had a significance value below 0.05. Therefore, a total of 15 descriptors that exhibited the highest correlation values and met the criteria for acceptable significance were selected as independent variables to conduct a multi-linear regression analysis to establish the QSAR equation. The variables used as descriptors for the MLR test, based on the JASP results, are BM, α, RD, qN1, qC3, E HOMO, V_{vdw}, qC8, E LUMO, μ, A_{vdw}, qC1, qC2, and qC9. Correlation data is presented in Table 3. For the multilinear regression analysis of the log IC_{50} value, 15 specific descriptors were used as independent variables. Utilizing the backward technique, multilinear regression analysis was conducted on 21 fitting compounds, deriving 6 equation models as displayed in Table 4. The statistical values employed for the test compound (7 compounds) included the partition coefficient, correlation coefficient, standard deviation, estimated F/F table, and PRESS value.

Upon analyzing the data shown in Table 5, it is evident that the estimated values of R, R^2, SD, and F and F table values exhibit little differences. However, the PRESS value stands out as significantly divergent. Consequently, this research selects equation model 5, which corresponds to the lowest PRESS value. The MLR calculation results obtained using the backward method (model Equation 5) will serve as a fundamental reference for the MLR calculation using the Enter method. This will be done to fit the compounds used for training and the test compounds (28 compounds). The resulting data will then be analyzed for statistical values, including correlation coefficient, partition coefficient, standard deviation, Fcount/Ftable, and PRESS for all the compounds. The ideal QSAR equation is derived from the findings of the Enter approach, and Figure 2 demonstrates the comparison between observed log IC_{50} and predicted log IC_{50}.

The correlation coefficient exceeds 0.8 according to the optimal QSAR equation and the correlation image (Figure 4) between experimental log IC_{50} and anticipated log IC_{50}. This suggests the predicted log IC_{50} values are widely scattered over the experimental log IC_{50} correlation line. This suggests that the equation is highly effective and can be utilized in creating innovative molecules using Oxime Ester Coumarin derivatives.

Table 6. Design of new compounds and their predicted log IC_{50}

Compound number	Structure	Log IC_{50} Pred.	Compound number	Structure	Log IC_{50} Pred.
29		-0.00018	35		0.314485
SMILES:C\C(=N/OC(=O)\C=C/C1=CC=C(Br)C=C1)C1=CC2=C(OC1=O)C=C(Br)C(C)=C2			SMILES:C\C(=N/OC(=O)\C=C/C1=CC=C(Br)C=C1)C1=CC2=CC(C=C)=C(Br)C=C2OC1=O		
30		0.377253	36		-0.06001
SMILES:C\C(=N/OC(=O)\C=C/C1=CC=C(Br)C=C1)C1=CC2=C(OC1=O)C=C(Br)C(Cl)=C2			SMILES:CC(=O)NCC1=C(Br)C=C2OC(=O)C(=CC2=C1)C(\C)=N\OC(=O)\C=C/C1=CC=C(Br)C=C1		
31		0.140997	37		0.250982
SMILES:C\C(=N/OC(=O)\C=C/C1=CC=C(Br)C=C1)C1=CC2=C(OC1=O)C=C(Br)C(Br)=C2			SMILES:C\C(=N/OC(=O)\C=C/C1=CC=C(Br)C=C1)C1=CC2=CC(N)=C(Br)C=C2OC1=O		
32		0.285733	38		0.320
SMILES:C\C(=N/OC(=O)\C=C/C1=CC=C(Br)C=C1)C1=CC2=C(OC1=O)C=C(Br)C(F)=C2			SMILES:C\C(=N/OC(=O)\C=C/C1=CC=C(Br)C=C1)C1=CC2=CC(O)=C(CC(O)=O)C=C2OC1=O		
33		0.334779	39		0.355
SMILES:C\C(=N/OC(=O)\C=C/C1=CC=C(Br)C=C1)C1=CC2=C(OC1=O)C=C(Br)C(O)=C2			SMILES:C\C(=N/OC(=O)\C=C/C1=CC=C(Br)C=C1)C1=CC2=CC(F)=C(CC(O)=O)C=C2OC1=O		
34		0.344635			
SMILES:COC1=CC2=C(OC(=O)C(=C2)C(\C)=N\OC(=O)\C=C/C2=CC=C(Br)C=C2)C=C1Br					

The novel product was designed by modifying substituents at the R1, R2, and R3 positions in the original coumarin oxime ester compound (Table 6). The resulting compound was further refined utilizing the PM3 method's geometric optimization. Based on the calculation results of log IC_{50} prediction using Equation 1, it is known that eleven newly modified compounds had lower predicted log IC_{50} values compared to 28 synthesized coumarin derivative compounds. This demonstrates that modifying the structure produces compounds that are more potent in inhibiting the activity of the α-glucosidase enzyme, as predicted.

Table 7. Docking results of Novel Coumarin derivatives

No.	Compound	Binding energy (kcal/mol)
1	**Ligand native**	**-7.283**
2	Compound 29	-6.667
3	Compound 30	-6.882
4	Compound 31	-6.307
5	**Compound 32**	**-7.374**
6	Compound 33	-7.05
7	Compound 34	-5.678
8	Compound 35	-6.554
9	Compound 36	-6.844
10	**Compound 37**	**-7.505**
11	**Compound 38**	**-8.351**
12	**Compound 39**	**-7.45**

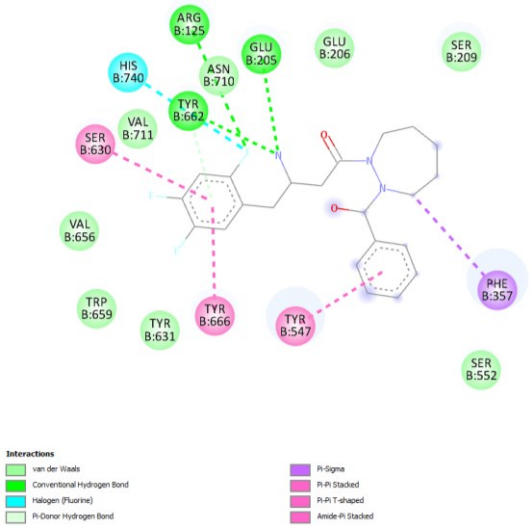


Figure 5. Docking interaction native ligand and amino acids from protein 2OLE

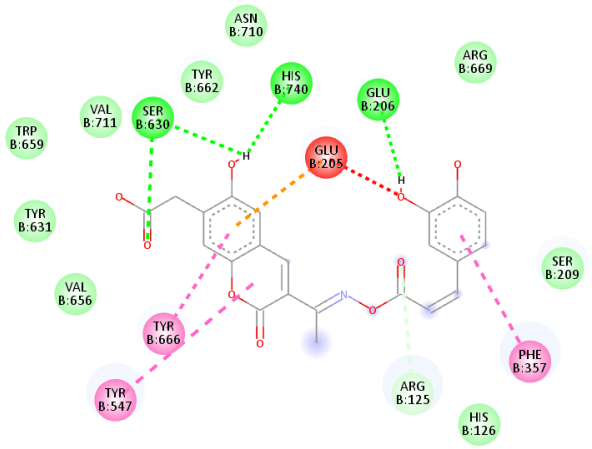


Figure 6. Docking interaction compound 38 with amino acids from Protein 2OLE

In addition to predicting IC_{50} values, the determination of the potential of a new compound can also be studied by in silico analysis. In silico analysis can be conducted using the molecular docking approach, utilizing proteins targeted based on their biological activity. Molecular docking analysis is used to observe and compare the interactions between the target protein, the predicted compound, and its native ligand.

Molecular docking analysis was performed on the protein 2OLE using eleven new compounds derived from coumarin with ester oxime groups. The protein 2OLE, downloaded from https://www.rscb.org, was prepared by removing water molecules and separating them from their native ligands. The separated protein and ligand, which have lost their water molecules, are then supplemented with hydrogen atoms and charges to adjust the docking environment to approximate the pH conditions inside the body and to restore hydrogen atoms in the macromolecule, allowing the formed hydrogen bonds to be observed [46]. Figure 4 presents the native protein and ligand resulting from the preparation. Before conducting docking, the molecular docking method must be validated. The validity of the method is determined by redocking the native ligand onto the target protein using the AutoDockTools 1.5.6 tool.

In the validation of molecular docking methods, the configuration of a grid box is also performed to serve as the space for the native ligand to form conformations when docked with the target protein, which is done after the protein. The grid box (Figure 5a) serves as the location for the ligand to interact with the amino acid residues at the binding site of the target protein. The grid box is determined to ascertain the coordinates of the binding site of a protein. The grid box configuration involves setting the grid center coordinates and size. RMSD (Root Mean Square Deviation) validates molecular docking by comparing the native ligand conformation from docking with that from crystallographic measurements [31].

The acceptable range for the RMSD value is < 3 Å [28]. The RMSD value of protein 2OLE for the native ligand is 1.64 Å, and the affinity energy is -7.28 kcal/mol. Based on these results, the method used is valid, and the docking process for the new compound derivative of coumarin with an ester oxime group can be performed. The visualization of the grid box and the interaction between the target protein and the native ligand is presented in Figures 5a and b. Figure 6 also illustrates the interaction of native ligands with the amino acids of the target protein. The amino acids bound to the native ligand include TYRB:666, TYRB:547, SERB:630, HISB:740, TYRB:662, ARGB:125, ASNB:710, GLUB:205, and PHEB:357.

Eleven novel compounds, coumarin derivatives containing ester oxime groups, were utilized for molecular docking on a specific protein. The docking procedure employed was identical to the validation approach. By employing consistent grid box size and coordinates throughout validation, the test compounds are accurately docked into the active region of the target protein, as confirmed. The chemicals exhibit flexibility while the target protein remains rigid, enabling it to adopt its ideal shape for binding to its active site.

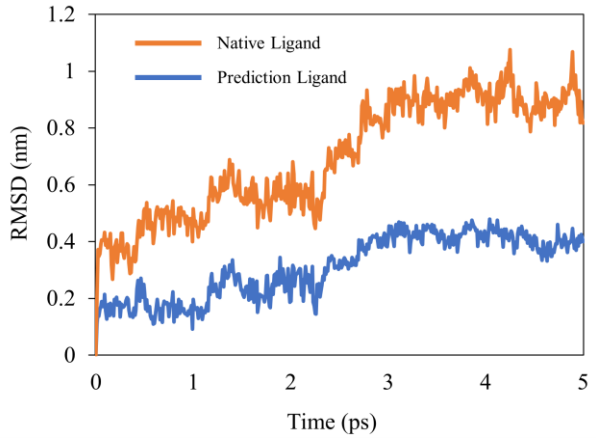


Figure 7. RMSD of the protein-ligand complex

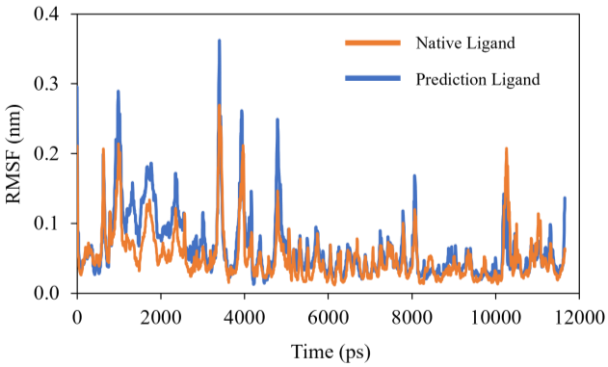


Figure 8. RMSF of the protein-ligand complex

The docking technique determines the binding energy between the test chemical and the target protein. The bond energy quantifies the attraction between the new chemical, coumarin derivative, and the oxime ester group on the target protein. The bond stability established between the novel derivative compounds of coumarin and the ester oxime group with the target protein increases as the acquired bond energy decreases. The energy values for the binding interactions are displayed in Table 7.

Table 7 displays data showing the existence of 4 recently altered compounds with lower binding energies than their original ligands. The four newly discovered compounds are 32, 37, 38, and 39. Their corresponding binding energies are -7.374, -7.505, -8.351, and -7.45 kcal/mol, respectively. The values are lower than the binding energy of the native ligand, which is -7.283 kcal/mol. According to the results of this investigation, it is possible to assume that the four compounds have the potential to operate as α-glucosidase inhibitors since they have lower binding energy with the target protein compared to the natural ligand. In addition to the binding energy data, Figure 7 illustrates the interaction between chemical 38 and the amino acids of the 2OLE protein. The compound 38 is associated with the following amino acids: TYRB:666, TYRB:547, PHEB:357, GLUB:205, SERB:630, HISB:740, and GLUB:206.

The molecular docking results of the compound derivative of coumarin with oxime ester group and standard ligand were further analyzed using molecular dynamics. Molecular dynamics analysis aims to determine the stability of ligand-receptor interactions, while molecular docking does not provide information regarding the stability of these interactions over space and time. The parameters used to evaluate the stability of ligand-enzyme interactions in molecular dynamics simulations are RMSD and RMSF.

During the simulation, the stability of the genuine protein ligand and the predicted protein ligand (compound 38) were carefully studied. Both interactions have an average RMSD value below 0.4 nm. The interaction between native and predicted protein-ligand has shown a stable RMSD with an average of 0.385 nm and 0.311 nm. Based on these values, it is evident that the RMSD of the predicted protein-ligand complex is lower compared to the RMSD of the native protein-ligand complex. The graph illustrating the RMSD fluctuations is presented in Figure 8. This indicates that the prediction of protein-ligand interactions is more stable than native protein-ligand interactions.

The parameter RMSF (Root Mean Square Fluctuation) is evaluated to determine the fluctuation of ligand-amino acid interactions in the enzyme throughout the simulation. Unlike RMSD, RMSF is calculated for each constituent residue of the protein, measuring the extent of fluctuation in the movement of each residue during the simulation. Generally, the RMSF value describes the flexibility of ligand interactions with each amino acid residue. The residual amino acid protein is measured after each fluctuation with RMSF. The protein-ligand fluctuation prediction shows a lower RMSF value than the native protein-ligand. The RMSF residue projection is below 0.45 nm (Figure 8), indicating that the protein is highly stable, which supports the RMSD results. This indicates that the predicted ligand or compound 38 is more stable than its native ligand in terms of bonding and interactions with the amino acids in protein 2OLE.

The interaction between coumarin derivatives and the oxime ester group with the amino acid acting as the catalytic site of the hydrolase 2OLE enzyme is mediated by residues TYRB:666, TYRB:547, PHEB:357, GLUB:205, SERB:630, HISB:740, and GLUB:206. The lower the value of RMSF, the more stable the interaction between the ligand and the amino acid [47]. The residual amino acid protein is measured after each fluctuation with RMSF. The protein-ligand fluctuation prediction shows a lower RMSF value than the native protein-ligand. The RMSF residue projection is below 0.45 nm, indicating that the protein is highly stable, which supports the RMSD results. This indicates that the predicted ligand or compound 38 is more stable than its native ligand in terms of bonding and interactions with the amino acids in protein 2OLE.

4. Conclusion

The best QSAR equation (Equation 1) has been obtained for the derivative compound of coumarin with an ester oxime group. Based on the equation, eleven new, more potent compounds have been designed. An in silico study was conducted on eleven new compounds with the hydrolase protein from Homo sapiens (PDB:2OLE) using molecular docking and molecular dynamics. This study identified compound 38 (SMILES: C\C(=N/OC(=O)\C=C/

C1=CC=C(Br)C=C1)C1=CC2=CC(O)=C(CC(O)=O)C=C2OC1) as the most potent compound, with the lowest binding energy value. in addition to having smaller RMSD and RMSF values than the native ligand.

Acknowledgment

The author thanks the BOPTN UIN Walisongo Semarang (980/Un.10.0/R/HK.01.14/4/2023) from the Ministry of Religion of Indonesia for funding this work through the Applied Research Grant program.

References

[1] Muhammad Rizqi Fahriza, Faktor Mempengaruhi yang Penyebab Kejadian Diabetes Mellitus (DM), *OSF Preprints*, (2019), https://doi.org/10.31219/osf.io/v82ea

[2] Chinmay D. Deshmukh, Anurekha Jain, B. Nahata, Diabetes Mellitus: A Review, *International Journal of Pure & Applied Bioscience*, 3, 3, (2015), 224-230

[3] Md Saidur Rahman, Khandkar Shaharina Hossain, Sharnali Das, Sushmita Kundu, Elikanah Olusayo Adegoke, Md. Ataur Rahman, Md. Abdul Hannan, Md Jamal Uddin, Myung-Geol Pang, Role of Insulin in Health and Disease: An Update, *International Journal of Molecular Sciences*, 22, 12, (2021), 6403 https://doi.org/10.3390/ijms22126403

[4] Pia V. Röder, Bingbing Wu, Yixian Liu, Weiping Han, Pancreatic regulation of glucose homeostasis, *Experimental & Molecular Medicine*, 48, 3, (2016), e219-e219 https://doi.org/10.1038/emm.2016.6

[5] Hanbing Li, Yuanfa Yao, Linghuan Li, Coumarins as potential antidiabetic agents, *Journal of Pharmacy and Pharmacology*, 69, 10, (2017), 1253-1264 https://doi.org/10.1111/jphp.12774

[6] Sara Ranđelović, Robbert Bipat, A Review of Coumarins and Coumarin-Related Compounds for Their Potential Antidiabetic Effect, *Clinical Medicine Insights: Endocrinology and Diabetes*, 14, (2021), https://doi.org/10.1177/11795514211042023

[7] Yinbo Pan, Teng Liu, Xiaojing Wang, Jie Sun, Research progress of coumarins and their derivatives in the treatment of diabetes, *Journal of Enzyme Inhibition and Medicinal Chemistry*, 37, 1, (2022), 616-628 https://doi.org/10.1080/14756366.2021.2024526

[8] Shashank M. Patil, Reshma Mary Martiz, A. M. Satish, Abdullah M. Shbeer, Mohammed Ageel, Mohammed Al-Ghorbani, Lakshmi Ranganatha, Saravanan Parameswaran, Ramith Ramu, Discovery of Novel Coumarin Derivatives as Potential Dual Inhibitors against α-Glucosidase and α-Amylase for the Management of Post-Prandial Hyperglycemia via Molecular Modelling Approaches, *Molecules*, 27, 12, (2022), 3888 https://doi.org/10.3390/molecules27123888

[9] Paul M. Dewick, The Mevalonate and Deoxyxylulose Phosphate Pathways: Terpenoids and Steroids, in: *Medicinal Natural Products: A Biosynthetic Approach*, John Wiley & Sons, 2001, https://doi.org/10.1002/0470846275.ch5

[10] Vladimir K. Mukhomorov, Biological activity of chemical compounds and their molecular Structure-information approach, *Journal of Chemical Engineering and Chemistry Research*, 1, 1, (2014), 54-65

[11] Włodzimierz Lewandowski, Hanna Lewandowska, Aleksandra Golonko, Grzegorz Świderski, Renata Świsłocka, Monika Kalinowska, Correlations between molecular structure and biological activity in "logical series" of dietary chromone derivatives, *PLoS ONE*, 15, 8, (2020), e0229477 https://doi.org/10.1371/journal.pone.0229477

[12] Hugo Hernandez, Livy Shivraj, *In Silico* Toxicity Prediction Using an Integrative Multimodel Approach, *ForsChem Research Reports*, 5, (2020),

[13] Samuel Genheden, Anna Reymer, Patricia Saenz-Méndez, Leif A. Eriksson, Computational Chemistry and Molecular Modelling Basics, in: S. Martín-Santamaría (Ed.) *Computational Tools for Chemical Biology*, The Royal Society of Chemistry, 2017, https://doi.org/10.1039/9781788010139-00001

[14] Xin Zhang, Ying-Ying Zheng, Chun-Mei Hu, Xiao-Zheng Wu, Jing Lin, Zhuang Xiong, Kun Zhang, Xue-Tao Xu, Synthesis and biological evaluation of coumarin derivatives containing oxime ester as α-glucosidase inhibitors, *Arabian Journal of Chemistry*, 15, 9, (2022), 104072 https://doi.org/10.1016/j.arabjc.2022.104072

[15] Anjar Purba Asmara, Mudasir Mudasir, Dwi Siswanta, Analisis Hubungan Kuantitatif Struktur dan Aktivitas Senyawa Turunan Triazolopiperazin Amida Menggunakan Metode Semiempirik AM1, *Elkawnie: Journal of Islamic Science Technology*, 1, 2, (2015), 125-138

[16] Anjar Purba Asmara, Mudasir Mudasir, Dwi Siswanta, Studi Qsar Senyawa Turunan Triazolopiperazin Amida Sebagai Inhibitor Enzim Dipeptidil Peptidase-IV (DPP IV) Menggunakan Metode Semiempirik AM1, *BIMIPA*, 23, 3, (2013), 288-296

[17] Ruslin Hadanu, Salim Idris, I Wayan Sutapa, QSAR Analysis of Benzothiazole Derivatives of Antimalarial Compounds Based on AM1 Semi-Empirical Method, *Indonesian Journal of Chemistry*, 15, 1, (2015), 86-92 https://doi.org/10.22146/ijc.21228

[18] Mutista Hafshah, Lilis Karlina, Desain Turunan Kalkon Baru Sebagai Antikanker Payudara Berdasarkan Molecular Docking, *Walisongo Journal of Chemistry*, 2, 2, (2019), 57-63 https://doi.org/10.21580/wjc.v2i2.6025

[19] Mudasir Mudasir, Iqmal Tahir, Ida Puji Astuti Maryono Putri, Quantitative Structure and Activity Relationship Analysis of 1, 2, 4-Thiadiazoline Fungicides Based on Molecular Structure Calculated by Am1 Method, *Indonesian Journal of Chemistry*, 3, 1, (2003), 39-47

[20] Mustofa Mustofa, Iqmal Tahir, Jumina Jumina, QSAR study of 1, 10-phenanthroline derivatives as the antimalarial compounds using electronic descriptors based on semiempirical AM1 calculation, *Indonesian Journal of Chemistry*, 2, 2, (2010), 91-96

[21] Khusna Arif Rakhman, Nur Asbirayani Limatahu, Hasbul Budiman Karim, Muhammad Ikhlas Abdjan, Kajian Senyawa Turunan Benzopirazin sebagai Antimalaria Menggunakan Metode HKSA dan MLR, *EduChemia*, 4, 2, (2019), 112-126 https://dx.doi.org/10.30870/educhemia.v4i2.4989

[22] T. Thalheim, D. Wondrousch, S. Stöckl, D. Mulliner, R. U. Ebert, R. Kühne, G. Schüürmann, Diagnostic of tautomer behaviour on QSAR models and AM1 optimisation, *Journal of Cheminformatics*, 3, (2011), P24 https://doi.org/10.1186/1758-2946-3-S1-P24

[23] Mutista Hafshah, Irvan Maulana Firdaus, Suratno Suratno, Quantitative Relationships Between Structure and Activity of Gamma-Carboline Derivative Compounds as Anti-Bovine Viral Diarrhea Virus (BVDV) Using Semi-Empirical AM1 Method, *Walisongo Journal of Chemistry*, 5, 2, (2022), 182-193 https://doi.org/10.21580/wjc.v5i2.13409

[24] Ponco Iswanto, Isnaeni Tatik Rosdiyana, Iqmal Tahir, Hubungan kuantitatif struktur dan aktivitas antikanker senyawa turunan estradiol hasil perhitungan metode semiempiris PM3, *BIMIPA*, 17, 1, (2007), 12-20

[25] Akhilesh Kumar Maurya, Viswajit Mulpuru, Nidhi Mishra, Discovery of Novel Coumarin Analogs against the α -Glucosidase Protein Target of Diabetes Mellitus: Pharmacophore-Based QSAR, Docking, and Molecular Dynamics Simulation Studies, *ACS Omega*, 5, 50, (2020), 32234-32249 https://doi.org/10.1021/acsomega.0c03871

[26] Thi-Hong-Truc Phan, Kowit Hengphasatporn, Yasuteru Shigeta, Wanting Xie, Phornphimon Maitarad, Thanyada Rungrotmongkol, Warinthorn Chavasiri, Designing Potent α -Glucosidase Inhibitors: A Synthesis and QSAR Modeling Approach for Biscoumarin Derivatives, *ACS Omega*, 8, 29, (2023), 26340-26350 https://doi.org/10.1021/acsomega.3c02868

[27] Muhammad Sulaiman Zubair, Saipul Maulana, Alwiyah Mukaddas, Penambatan molekuler dan simulasi dinamika molekuler senyawa dari genus nigella terhadap penghambatan aktivitas enzim protease HIV-1, *Jurnal Farmasi Galenika*, 6, 1, (2020), 132-140 https://doi.org/10.22487/j24428744.2020.v6.i1.14982

[28] Ming-Hui Li, Quan Luo, Xiang-Gui Xue, Ze-Sheng Li, Molecular dynamics studies of the 3D structure and planar ligand binding of a quadruplex dimer, *Journal of Molecular Modeling*, 17, 3, (2011), 515-526 https://doi.org/10.1007/s00894-010-0746-0

[29] Ranjan K. Mohapatra, Kuldeep Dhama, Amr Ahmed El–Arabey, Ashish K. Sarangi, Ruchi Tiwari, Talha Bin Emran, Mohammad Azam, Saud I. Al-Resayes, Mukesh K. Raval, Veronique Seidel, Mohnad Abdalla, Repurposing benzimidazole and benzothiazole derivatives as potential inhibitors of SARS-CoV-2: DFT, QSAR, molecular docking, molecular dynamics simulation, and in-silico pharmacokinetic and toxicity studies, *Journal of King Saud University - Science*, 33, 8, (2021), 101637 https://doi.org/10.1016/j.jksus.2021.101637

[30] Emmanuel Israel Edache, Adamu Uzairu, Paul Andrew Mamza, Gideon Adamu Shallangwa, Structure-based simulated scanning of rheumatoid arthritis inhibitors: 2D-QSAR, 3D-QSAR, docking, molecular dynamics simulation, and lipophilicity indices calculation, *Scientific African*, 15, (2022), e01088 https://doi.org/10.1016/j.sciaf.2021.e01088

[31] Lalu Sanik Wahyu Fadil Amrulloh, Nuraini Harmastuti, Andri Prasetiyo, Rina Herowati, Analysis of Molecular Docking and Dynamics Simulation of Mahogany (*Swietenia macrophylla* King) Compounds Against the PLpro Enzyme SARS-COV-2, *Jurnal Farmasi dan Ilmu Kefarmasian Indonesia*, 10, 3, (2023), 347-359 https://doi.org/10.20473/jfiki.v10i32023.347-359

[32] Devendra K. Dhaked, Jitender Verma, Anil Saran, Evans C. Coutinho, Exploring the binding of HIV-1 integrase inhibitors by comparative residue interaction analysis (CoRIA), *Journal of Molecular Modeling*, 15, 3, (2009), 233-245 https://doi.org/10.1007/s00894-008-0399-4

[33] Ponco Iswanto, Eva Vaulina Yulistia Delsy, Ely Setiawan, Fiandy Aminullah Putra, Quantitative Structure-Property Relationship Analysis Against Critical Micelle Concentration of Sulfonate-Based Surfactant Based on Semiempirical Zindo/1 Calculation, *Molekul*, 14, 2, (2019), 78-83 https://doi.org/10.20884/1.jm.2019.14.2.467

[34] Ponco Iswanto, Irvan Maulana Firdaus, Ahmad Fawwaz Dafaulhaq, Ahmad Ghifari Ramadhani, Maylani Permata Saputri, Heny Ekowati, Quantitative Structure-Activity Relationship of 3-Thiocyanate-1H-Indoles Derived Compounds as Antileukemia by AM1, PM3, and RM1 Methods, *Jurnal Kimia Sains dan Aplikasi*, 26, 3, (2023), 109-117 https://doi.org/10.14710/jksa.26.3.109-117

[35] Eva Vaulina, Ponco Iswanto, Model QSAR Senyawa Fluorokuinolon Baru Sebagai Zat Antibakteri *Salmonella Thypimurium*, *Molekul*, 1, 1, (2006), 10-18 http://dx.doi.org/10.20884/1.jm.2006.1.1.17

[36] Purnawan Pontana Putra, *Teori dan Tutorial Molecular Docking Menggunakan AutoDock Vina*, Wawasan Ilmu, Banyumas, 2022,

[37] Chan Kyung Kim, Soo Gyeong Cho, Chang Kon Kim, Mi-Ri Kim, Hai Whang Lee, Prediction of Physicochemical Properties of Organic Molecules Using Semi-Empirical Methods, *Bulletin of the Korean Chemical Society*, 34, 4, (2013), 1043-1046 https://doi.org/10.5012/bkcs.2013.34.4.1043

[38] Muneerah Mogren Al Mogren, Enfale Zerroug, Salah Belaidi, Ahlam BenAmor, Sarah Dhaif Allah Al Harbi, Molecular structure, drug likeness and QSAR modeling of 1,2-diazole derivatives as inhibitors of enoyl-acyl carrier protein reductase, *Journal of King Saud University - Science*, 32, 4, (2020), 2301-2310 https://doi.org/10.1016/j.jksus.2020.03.007

[39] Samia Boudergua, Salah Belaidi, Muneerah Mogren AlMogren, Aouda Bounif, Mohamed Bakhouch, Samir Chtita, QSAR modeling using the Gaussian process applied for a series of flavonoids as potential antioxidants, *Journal of King Saud University - Science*, 35, 8, (2023), 102898 https://doi.org/10.1016/j.jksus.2023.102898

[40] Daratu Eviana Kusuma Putri, Harno Dwi Pranowo, Anugrah Ricky Wijaya, Novia Suryani, Maisari Utami, Artania Adnin Tri Suma, Woo Jin Chung, Saeedah Musaed Almutairi, Dina S. Hussein, Rabab Ahmed Rasheed, Venkatalakshmi Ranganathan, The predicted models of anti-colon cancer and anti-hepatoma activities of substituted 4-anilino coumarin derivatives using quantitative structure-activity relationship (QSAR), *Journal of King Saud University - Science*, 34, 3, (2022), 101837 https://doi.org/10.1016/j.jksus.2022.101837

[41] Giuseppe Floresta, Agostino Cilibrizzi, Vincenzo Abbate, Ambra Spampinato, Chiara Zagni, Antonio Rescifina, FABP4 inhibitors 3D-QSAR model and isosteric replacement of BMS309403 datasets, *Data in Brief*, 22, (2019), 471-483 https://doi.org/10.1016/j.dib.2018.12.047

[42] Natalia Piekuś-Słomka, Mariusz Zapadka, Bogumiła Kupcewicz, Methoxy and methylthio-substituted trans-stilbene derivatives as CYP1B1 inhibitors – QSAR study with detailed interpretation of molecular descriptors, *Arabian Journal of Chemistry*, 15, 11, (2022), 104204 https://doi.org/10.1016/j.arabjc.2022.104204

[43] Alicia Rosell-Hidalgo, Anthony L. Moore, Taravat Ghafourian, Prediction of drug-induced mitochondrial dysfunction using succinate-cytochrome c reductase activity, QSAR and molecular docking, *Toxicology*, 485, (2023), 153412 https://doi.org/10.1016/j.tox.2022.153412

[44] Emmy Yuanita, Sudirman, Ni Komang Tri Dharmayani, Maria Ulfa, Jufrizal Syahri, Quantitative structure–activity relationship (QSAR) and molecular docking of xanthone derivatives as anti-tuberculosis agents, *Journal of Clinical Tuberculosis and Other Mycobacterial Diseases*, 21, (2020), 100203 https://doi.org/10.1016/j.jctube.2020.100203

[45] M. Yusuf, N. Dari, E. Utama, PM3 Semi-Empirical Method on the Ring Opening Polymerization of ε-Caprolactone Using Bis (Benzoyltrifluoroacetone)Zirconium(IV) Chloride as catalyst, *Journal of Physics: Conference Series*, 1462, (2020), 012054 https://doi.org/10.1088/1742-6596/1462/1/012054

[46] Gordon Lemmon, Jens Meiler, Towards Ligand Docking Including Explicit Interface Water Molecules, *PLoS ONE*, 8, 6, (2013), e67536 https://doi.org/10.1371/journal.pone.0067536

[47] Shailima Rampogu, Gihwan Lee, Jun Sung Park, Keun Woo Lee, Myeong Ok Kim, Molecular Docking and Molecular Dynamics Simulations Discover Curcumin Analogue as a Plausible Dual Inhibitor for SARS-CoV-2, *International Journal of Molecular Sciences*, 23, 3, (2022), 1771 https://doi.org/10.3390/ijms23031771