QSAR, Molecular Docking and Dynamics Studies of Quinazoline Derivatives as Inhibitor of Phosphatidylinositol 3-Kinase

Muhammad Arba¹*, Ruslin¹, Waode Umi Kalsum², Armid Alroem², Muhammad Zakir Muzakkar², Ida Usman³, Daryono Hadi Tjahjono⁴

¹Faculty of Pharmacy, Halu Oleo University, Kendari, 93232, Indonesia.  
²Department of Chemistry, Halu Oleo University, Kendari, 93232, Indonesia.  
³Department of Physics, Halu Oleo University, Kendari, 93232, Indonesia.  
⁴School of Pharmacy, Institut Teknologi Bandung, Bandung, 40132, Indonesia.

ARTICLE INFO
Article history: 
Received on: 18/12/2017  
Accepted on: 22/03/2018  
Available online: 30/05/2018

Key words:  
Docking,  
MM-PBSA, Molecular dynamics simulation,  
Phosphatidylinositol 3-kinase, QSAR,  
Quinazoline.

ABSTRACT
Phosphatidylinositol 3-kinase (PI3K) plays a prominent role in regulating various crucial cellular functions. Many studies have indicated the involvement of PI3K in tumorigenesis. In the current study, thirty-one quinazoline derivatives were utilized to build a Quantitative Structure-Activity Relationship (QSAR) model which correlates structural feature with PI3K inhibition. The statistically robust QSAR model is $\text{pIC}_{50} = 2.515 + 0.000005 \text{(AM1}_\text{Eele}) + 0.004 \text{(AM1}_\text{HF}) + 1.170 \text{(AM1}_\text{LUMO}) - 0.117 \text{(apol)} + 0.003 \text{(ASA}_\text{H})$ with a leave-one-out cross-validation coefficient ($q^2$) of 0.6058 and external validation ($R^2_{\text{pred}}$) of 0.7725. A novel compound (SC25) was proposed based on the validated QSAR model. Molecular docking of the ligand on PI3K revealed the similar binding mode of SC25 and parent compound ((S)-C5) as well as native ligand (2NQ). Molecular dynamics simulation of 40 ns confirmed the conformational stability of each SC25, (S)-C5, and 2NQ, complexed with PI3K. Prediction of affinity using MM-PBSA method revealed that SC25 has a comparable affinity with that of (S)-C5 and better than that of 2NQ.

INTRODUCTION
The phosphatidylinositol 3-kinase (PI3K) signaling pathway is an important signaling pathway, which controls the cycle, survival, metabolism, motility, and survival of cells. The pathway is commonly activated transduction cascade in human cancers as its aberrations are found in up to 40% of all tumor types; thus, targeting the PI3K pathway was regarded as a potential strategy in the treatment of various tumor types (Polivka and Janku, 2014; Miled et al., 2007). To date, numerous efforts have been performed to develop compounds targeting the PI3K/AKT/mTOR pathway; some of which are in clinical development including PF-04691502 (Pfizer), BEZ235 (Novartis), XL765 (Exelixis/Sanofi-Aventis) and GSK2126458 (GlaxoSmithKline) (Figure 1) (Yap et al., 2015; Bauer et al., 2015; Britten et al., 2014; Papadopoulos et al., 2015). As the development of dual inhibitors of the PI3K/mTOR has not been straightforward, novel compounds with reduced toxicity and increased efficacy are still needed.

With the growing appreciation of quinazoline as important scaffolds in the development of a PI3K inhibitor, Zhang et al. (2015) synthesized and evaluated the biological activity of $N$-(2-methoxy-5-(3-substituted quinazolin-4(3H)-one-6-yl)-pyridin-3-yl)phenylsulfonamide as a PI3K inhibitor. They found that two compounds, (S)-C5 and (S)-C8, displayed potent inhibitory activity against PI3Kαs and mTOR, especially against PI3Kα. To aid in the design of quinazoline derivatives as a PI3K inhibitor, in the current study, we performed a quantitative structure-activity relationship (QSAR), molecular docking, and molecular dynamics study on a series of quinazoline derivatives as a PI3K inhibitor. Our goal is to find a novel compound of...
the quinazoline derivatives with better affinity than the existing compound by means of the computational method.

![Structure of some PI3K inhibitors](image)

Fig. 1: Structure of some PI3K inhibitors.

<table>
<thead>
<tr>
<th>No.</th>
<th>Compound</th>
<th>R₁</th>
<th>R₂</th>
<th>IC₅₀ (µM) HCT-116</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A1</td>
<td>4-F</td>
<td>CH₂CH₃</td>
<td>4.06</td>
</tr>
<tr>
<td>2</td>
<td>A2</td>
<td>4-F</td>
<td>CH₃(CH₂)₂</td>
<td>2.69</td>
</tr>
<tr>
<td>3</td>
<td>A3</td>
<td>4-F</td>
<td>CH₃N(CH₂CH₃)</td>
<td>1.2</td>
</tr>
<tr>
<td>4</td>
<td>A4</td>
<td>4-F</td>
<td>CH₃N(CH₂CH₃)O</td>
<td>1.82</td>
</tr>
<tr>
<td>5</td>
<td>A5</td>
<td>4-F</td>
<td>CH₃N(CH₂CH₃)NCH₂</td>
<td>1.45</td>
</tr>
<tr>
<td>6</td>
<td>A6</td>
<td>4-F</td>
<td>CH₃N(CH₂)</td>
<td>2.64</td>
</tr>
<tr>
<td>7</td>
<td>A7</td>
<td>4-F</td>
<td>CH₃N(CH₂)</td>
<td>2.75</td>
</tr>
<tr>
<td>8</td>
<td>A8</td>
<td>4-Cl</td>
<td>CH₃N(CH₂CH₃)O</td>
<td>2.02</td>
</tr>
<tr>
<td>9</td>
<td>A9</td>
<td>4-CH₃</td>
<td>CH₃N(CH₂CH₃)O</td>
<td>1.35</td>
</tr>
<tr>
<td>10</td>
<td>A10</td>
<td>2,4-di-F</td>
<td>CH₃N(CH₂CH₃)</td>
<td>1.03</td>
</tr>
<tr>
<td>11</td>
<td>A11</td>
<td>F</td>
<td>Ph</td>
<td>4.74</td>
</tr>
<tr>
<td>12</td>
<td>A12</td>
<td>F</td>
<td>Ph-4-OCH₃</td>
<td>2.80</td>
</tr>
<tr>
<td>13</td>
<td>A13</td>
<td>F</td>
<td>Py-4</td>
<td>2.95</td>
</tr>
<tr>
<td>14</td>
<td>A14</td>
<td>F</td>
<td>Py-3</td>
<td>1.29</td>
</tr>
<tr>
<td>15</td>
<td>B1</td>
<td>F</td>
<td>Py-3</td>
<td>1.29</td>
</tr>
<tr>
<td>16</td>
<td>B2</td>
<td>F</td>
<td>CH₃N(CH₂CH₃)O</td>
<td>1.42</td>
</tr>
<tr>
<td>17</td>
<td>(S)-C1</td>
<td>F</td>
<td>Boc</td>
<td>0.66</td>
</tr>
<tr>
<td>18</td>
<td>(S)-C2</td>
<td>F</td>
<td>H</td>
<td>1.46</td>
</tr>
<tr>
<td>19</td>
<td>(S)-C3</td>
<td>F</td>
<td>Ac</td>
<td>1.25</td>
</tr>
<tr>
<td>20</td>
<td>(R)-C4</td>
<td>F</td>
<td>Ac</td>
<td>4.72</td>
</tr>
<tr>
<td>21</td>
<td>(S)-C5</td>
<td>F</td>
<td>COPr-c</td>
<td>0.71</td>
</tr>
<tr>
<td>22</td>
<td>(R)-C6</td>
<td>F</td>
<td>COPr-c</td>
<td>6.09</td>
</tr>
<tr>
<td>23</td>
<td>(S)-C7</td>
<td>F</td>
<td>COCH(CH₂CH₃)O</td>
<td>0.60</td>
</tr>
<tr>
<td>24</td>
<td>(S)-C8</td>
<td>F</td>
<td>CH₃Pr-c</td>
<td>0.67</td>
</tr>
<tr>
<td>25</td>
<td>(R)-C9</td>
<td>F</td>
<td>CH₃Pr-c</td>
<td>4.88</td>
</tr>
<tr>
<td>26</td>
<td>D1</td>
<td>F</td>
<td>Boc</td>
<td>1.53</td>
</tr>
<tr>
<td>27</td>
<td>D2</td>
<td>F</td>
<td>H</td>
<td>1.15</td>
</tr>
<tr>
<td>28</td>
<td>D3</td>
<td>F</td>
<td>Ac</td>
<td>1.30</td>
</tr>
<tr>
<td>29</td>
<td>D4</td>
<td>F</td>
<td>COPr-c</td>
<td>1.35</td>
</tr>
<tr>
<td>30</td>
<td>D5</td>
<td>F</td>
<td>COCH(CH₂CH₃)O</td>
<td>0.96</td>
</tr>
<tr>
<td>31</td>
<td>D6</td>
<td>F</td>
<td>CH₃Pr-c</td>
<td>1.22</td>
</tr>
</tbody>
</table>

Table 1: The quinazoline derivatives in the data set.

A computational method such as QSAR, molecular docking, and molecular dynamics simulation, has been extensively applied in the rational drug design (Kumar et al., 2016; Shinde et al., 2017). It has been a powerful approach to improve potency, selectivity and/or ADMET properties of lead compounds in the development of novel chemical entities. The combined QSAR, molecular docking, and molecular dynamics simulations are widely used to predict the biological activity of potential new drugs and the binding mode of a drug in the active site of a protein target, as well as to evaluate the conformational stability of drug-protein complex during a period of time (El-Sawy et al., 2017; Abdalsalam, 2017; Khan et al., 2017; Arba et al., 2017a).

**COMPUTATIONAL METHOD**

**Dataset**

Thirty-one compounds of quinazoline derivatives were taken from a reference (Zhang et al., 2015) as shown in Figure 2 and Table 1. The IC₅₀ values (concentration that needs for inhibition the 50% of enzyme activity) of the 31 compounds were converted to the pIC₅₀ (− Log IC₅₀) values. The pIC₅₀ values were in the range of 0.60 µM to 6.09 µM. For all compounds, the values of studentized deleted residual of >2 or <−2 were used to identified outlier compound(s). Furthermore, the dataset of biological activities of different compounds were grouped into a training set (75% of compounds) and test set (25% of compounds). The test set was randomly selected which span the entire activity range.

**Geometry optimization and descriptor calculation**

Each compound was built and geometrically optimized by using Gaussian 09 software (Frisch et al., 2009). The semi-empirical Austin Model-1 (AM1) method was used for the optimization. Using the optimized structure of the molecule, various descriptors were calculated using Molecular Operating Environment (MOE 2009.10) software. The following molecular descriptors were used: total energy (AM1_E), electronic energy (AM1_Eele), dipole moment (AM1_Dipol), formation heat
(AM1_HF), HOMO energy (AM1_HOMO), LUMO energy (AM1_LUMO), polarity (Apol), hydrophobic surface area (ASA_H), water solubility (Log S), partition coefficient (Log P), globularity (Glob), van der Waals volume (Vol), and molar refractivity (Mr).

Fig. 2: Structure of parent compound.

**QSAR model calculation and validation**

The QSAR model development was carried out by using the multiple linear regression analysis to find the linear relationship between the independent variable (a set of descriptor) and the dependent variable (biological activity) by means of SPSS software (version 19; SPSS Inc., Chicago, IL, USA). The reliability of the QSAR model was statistically adjudged based on the following criteria: $R^2$ (squared correlation coefficient), $F$ test (Fischer’s statistic value), $R_{adj}^2$ (adjusted squared correlation coefficient), and SEE (standard error of estimation) (Dearden et al., 2009). The QSAR model was also checked for its predictive potential by using leave-one-out (LOO) cross-validation coefficient ($q^2$), which is a reliable method for testing the significance of the model (Golbraikh and Tropsha, 2002). In the LOO cross-validation, every single compound of the training set was removed and its biological activity was predicted using the model built from the rest of compounds. The following equation was utilized to calculate $q^2$ value:

$$q^2 = 1 - \frac{\sum(y_i - \tilde{y}_i)^2}{\sum(y_i - \bar{y})^2}$$

where $y_i$, $\tilde{y}_i$, and $\bar{y}$ are measured, predicted, and average measured activities of the compound in the training set, respectively. High $q^2$ value ($q^2 > 0.5$) is considered as a proof of the predictive QSAR model (Golbraikh et al. 2003; Tropsha et al., 2003). However, Tropsha et al. (2003) explained that it is necessary to validate the model with the test compounds in addition to the internal cross-validation. In that scheme, the biological activities of the test set compounds were predicted using the developed model and the validity of the model was evaluated in terms of external cross-validation coefficient ($R_{pred}^2$). The $R_{pred}^2$ value higher than 0.6 indicates that the model was valid.

**Design for new molecule and molecular docking**

Designing and prediction of the biological activity of the novel compound of quinazoline derivative were performed using the validated QSAR model. As a reference compound, (S)-C5 was used as it was the most potent compound in the enzyme-based activity test (Zhang et al., 2015). Furthermore, to predict the binding mode of the novel compound, molecular docking to PI3Kγ was performed, in which the structure of PI3Kγ was extracted from protein data bank with PDB ID 3S2A and X-ray resolution of 2.55 Å (Nishimura et al., 2011). The binding site of the protein was located following the native ligand (2NQ) binding conformation. The grid point spacing of 0.375 Å with a dimension of 50 in each x, y, and z direction was used, while other docking parameters were left as default (Arba et al., 2017b).
Molecular dynamics simulation

Molecular dynamics (MD) simulation was performed using AMBER16 (Case et al., 2015) on each complex of top-ranked docking conformation of the parent compound ((S)-C5), novel designed compound (SC25), and native ligand (2NQ). Protein and ligand parameterization was performed by employing the ff14SB, GAFF force fields and AM1-BCC (Maier et al., 2015; Wang et al., 2004; Jakalian et al., 2002). Each system was then solvated using water model of a truncated octahedron TIP3P with a minimum distance of 10 Å around the complex. The Na\(^+\) counterions addition was done to keep the system neutral. The minimization, heating, and equilibration were performed using Sander module of Amber 16. The minimization step was performed sequentially during 3 three steps; each consisted of 500 steps of steepest descents and 5500 steps of conjugate gradients. The restraint (\(k = 500 \text{ kcal mol}\(^{-1}\) Å\(^{-2}\)) was applied on whole protein and backbone atoms of protein during first and second minimization, respectively, while the third minimization was performed without any restraint.

Each system was heated gradually from 0 to 100, 100 to 200, and 200 to 300 K during every 50 ps in NVT ensemble with a time step of 0.0005 ps and restraints (\(k\)) of 5 kcal mol\(^{-1}\) Å\(^{-2}\). System relaxation was performed using three 100 ps equilibration steps in NPT ensemble from 5 to 3 and 0 kcal mol\(^{-1}\) Å\(^{-2}\). The production step of 40 ns was performed by using pmemd.cuda module in Amber 16 in an isothermal-isobaric ensemble (Salomon-Ferrer et al., 2013). To keep the system in 300 K thermal bath, the Langevin thermostat was used with a collision rate of 1.0 ps\(^{-1}\). The SHAKE algorithm was used to restrain all covalent bonds involving hydrogen atoms (Ryckaert et al., 1977). The PME method was used to treat the long-range electrostatic interactions with an integration step of 2 fs (Darden et al., 1993). The long-range non-bonded interactions were calculated by applying periodic boundary conditions with a cutoff distance of 9.0 Å. The CPPTRAJ module was utilized to analyze and extract structural snapshots of MD trajectory (Roe and Cheatham, 2013), while Visual Molecular Dynamics and Discovery Studio Visualizer were used for visualization (Humphrey et al., 1996).

Binding free energy calculations

The prediction of binding free energy for each system was performed by using Molecular Mechanics Poisson-Boltzmann solvent accessible surface area (MM-PBSA) method (Kollman et al., 2000; Arba et al., 2017b). The last 5 ns MD trajectory of 200 snapshots was utilized to calculate the binding free (Miller et al., 2012).

RESULTS AND DISCUSSION

The present investigation aims to provide the 2D QSAR study which correlates descriptors as an independent variable with biological activity as a dependent variable. Firstly, calculation of the values of studentized deleted residual was performed, which then identified six compounds as outliers (Table 2). The six compounds were then removed from the dataset. Furthermore, 25 compounds were left as data set which was divided randomly to be training set and a test set of 19 and 6 compounds, respectively (Table 2).

Furthermore, multiple linear regression analysis was applied to generate QSAR models using 19 compounds as the training set. The best QSAR model as shown below, which was chosen based on statistical criteria, has good statistical parameters such as correlation coefficient (\(R\)), determination coefficient (\(R^2\)), and Fischer’s value (\(F\)) of 0.926, 0.857, and 16.452, respectively.

\[
pIC_{50} = 2.515 + 0.000005 (AM1_{Eele}) + 0.004 (AM1_{HF}) + 1.170 (AM1_{LUMO}) - 0.117 \text{ (apol)} + 0.003 \text{ (ASA\_H)}.
\]

The accuracy of the model was also assured by the low standard error (SE) of 0.1112. Table 3 shows descriptors and statistical parameters of the statistically significant QSAR model.

As shown in Table 3, the validity of QSAR model was indicated by leave-one-out cross-validation coefficient (\(q^2\)) of
It can be inferred from the model that more hydrophobic groups are preferable for increasing the biological activity as indicated by the positive sign of the coefficient of ASA_H. Meanwhile, the negative sign of the coefficient of apol indicates that less polar groups were favorable for the activity. In addition, the positive sign of the coefficient of AM1_Eele, AM1_HF, and AM1_LUMO implies the positive contribution of those parameters to biological activity. It is noted that AM1_LUMO is the most influencing descriptor as indicated by the highest coefficient value. The validity of the model was also supported by external validation of the test compounds ($R^{2}_{\text{pred}}$) of 0.7725, which fulfills the condition, $R^{2}_{\text{pred}} > 0.5$ (Golbraikh and Tropsha, 2002). Figure 3 shows the relationship between observed and predicted pIC$_{50}$.

**Table 3:** Statistical results of QSAR model for quinazoline derivatives.

<table>
<thead>
<tr>
<th>Descriptors</th>
<th>$q^2$</th>
<th>$R$</th>
<th>$R^{2}$</th>
<th>Adjusted $R^{2}$</th>
<th>Standard Error (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AM1_Eele, AM1_HF, AM1_LUMO, apol, ASA_H</td>
<td>0.6058</td>
<td>0.926</td>
<td>0.857</td>
<td>0.802</td>
<td>0.1112</td>
</tr>
</tbody>
</table>

Fig. 3: The relationship between observed pIC$_{50}$($X$) and predicted pIC$_{50}$($Y$) of quinazoline derivatives.

The design of new compound and molecular docking

Designing and prediction of the biological activity of novel compound were carried out by using the best QSAR model. The results showed that a new compound (SC25) has predicted IC$_{50}$ lower (IC$_{50\text{pred}} = 0.5364$ µM) than that of parent compound ((S)-C5) (IC$_{50\text{pred}} = 0.71$ µM). The structure of (S)-C5 and SC25 is depicted in Figure 4, which differs by ethyl substituent in meta-position.

Next, molecular docking of ((S)-C5 and SC25 on human PI3Kγ (PDB code 3S2A) was performed. For docking validation, the native ligand (2NQ) was docked on the PI3Kγ. The result showed that the root-mean-square-deviation (RMSD) between

![Fig. 4: Structures of (a) (S)-C5, (b) SC25.](image-url)
crystallographic and docking conformations is 1.1.6 Å, indicating that the docking protocol was valid (Jones et al., 1997; Morris et al., 1998). The docking conformation of 2NQ was able to reproduce important interactions such as hydrogen bond between nitrogen quinoline with Val882. The 2D structure and interaction of 2NQ with human PI3Kγ were depicted in Figure 5.

The docking conformation of SC25 showed that several hydrogen bonds were formed in the binding of SC25: two oxygen atoms of sulfonamido with Lys833; nitrogen atom of pyridine ring with Tyr867; nitrogen atom of quinazolin-4(3H)-one moiety with Val882; and the oxygen atom of cyclopropylcarbonyl with Thr886. On other hand, docking conformation of (S)-C5 showed less number of hydrogen bonds: only one oxygen atom of sulfonamido with Lys833; nitrogen atom of quinazolin-4(3H)-one moiety with Val882, and nitrogen atom of pyridine ring with Asp964, instead of Tyr867 (Zhang et al., 2015; Nishimura et al., 2011). Figure 6 depicts docking conformations of (S)-C5 and SC25 in the binding site of PI3K.

![Figure 5: The 2D structure (a) and docked conformation (b) of 2NQ.](image-url)

![Figure 6: The docked conformations of (a) (S)-C5 and (b) SC25 in the binding pocket of PI3K.](image-url)

**Molecular dynamics simulation**

Molecular dynamics simulation was performed to evaluate the conformational stability of three compounds, i.e. (S)-C5, SC25, and 2NQ, each complexed with PI3K. The stability of each complexed was required before performing energetics analysis, which was measured by the values of Root-Mean-Squared-Deviation (RMSD). Figure 7 shows the RMSD plot of heavy atoms of PI3K with respect to simulation time for each
complex. The average values of RMSD during 40 ns dynamics simulation of the three complexes are around 3 Å, in which the RMSD of SC25 is lower than that of S-C5, indicating a more stable structure.

In addition, the fluctuation of amino acid residues due to ligand binding was measured by the root-mean-square-fluctuation (RMSF) values. Figure 8 shows RMSF of protein versus residue number. The similar pattern of RMSF was observed in all three complexes, which indicate the similar binding mode of the three ligands. It was shown that high fluctuation was recorded in the amino acid residue Asp515 (Asp758) and carbonyl end of the protein.

![Graph showing RMSD values for each complex](image1)

**Fig. 7:** The RMSD values of each 2NQ (red), (S)-C5 (green), and SC25 (blue), each complexed with PI3K.

![Graph showing RMSF values versus residue number](image2)

**Fig. 8:** The RMSF values versus residue number of 2NQ (red), (S)-C5 (green), and SC25 (blue), each complexed with PI3Kγ.

### Binding energy prediction by MM-PBSA

Table 4 shows the binding free energies of each compound bound to the PI3K. The MM-PBSA prediction showed that the binding free energies of both (S)-C5 ($\Delta G_{\text{PBTOT}} = -33.78 \text{ kcal/mol}$) and SC25 ($\Delta G_{\text{PBTOT}} = -33.74 \text{ kcal/mol}$) were almost the same, indicating that both compounds have similar affinities. It is worth to note that both compounds have better affinities than that of native ligand (2NQ) ($\Delta G_{\text{PBTOT}} = -20.17 \text{ kcal/mol}$). Furthermore, in each complex, both van der Waals ($\Delta E_{\text{VDW}}$) and electrostatic energies ($\Delta E_{\text{ELE}}$) were favorable for ligand binding. It was also the case for the contribution of the nonpolar energy of desolvation. In the meantime, the polar energy of desolvation ($\Delta E_{\text{PBCAL}}$) was all unfavorable for ligand binding.
CONCLUSION

The current study shows that good correlation of 13 descriptors which represent the chemical structures of the compound and the biological activities result in valid QSAR equation. The QSAR model was used to design novel compound (SC25) which has lower predicted IC_{50} than the parent compound ((S)-C5). Molecular dynamics simulation of 40 ns was used to assure the stability of ligand in the binding cavity of PI3K. Prediction of binding free energy using MM-PBSA method shows that the novel compound has a comparable affinity with that of the parent compound. It is noted that the affinity of novel compound is much lower than that of native ligand.

REFERENCES


Molecular Operating Environment (MOE), Version 2009.10, Chemical Computing Group Inc., Montreal, Quebec, Canada.


Molecular Operating Environment (MOE), Version 2009.10, Chemical Computing Group Inc., Montreal, Quebec, Canada.


Polivka J Jr, Janku F. Molecular targets for cancer therapy in the PI3K/AKT/mTOR pathway. Pharmacol & Ther. 2014; 142:164-175.


How to cite this article: