Molecular modeling of 4-fluoropyrrolidine-2-carbonitrile and octahydrocyclopenta[b]pyrrole-2-carbonitrile as a dipeptidyl peptidase IV (DPP4) inhibitor

Muhammad Arba*, Ruslin Ruslin1, Nur Illiyyin Akib1, Yamin Yamin1, Sabarudin Ombe1, Jessi Jessi2, Muhammad Zakir Muzakkar3, Daryono Hadi Tjahjono1

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

ARTICLE INFO
Received on: 01/02/2018
Accepted on: 30/08/2018
Available online: 31/10/2018

Key words:
Dipeptidyl peptidase, carbonitrile, QSAR, docking, molecular dynamic simulation, MM-PBSA.

ABSTRACT
Research on the quantitative structure–activity relationship (QSAR) of the 4-fluoropyrrolidine-2-carbonitrile and octahydrocyclopenta[b]pyrrole-2-carbonitrile as dipeptidyl peptidase IV (DPP4) inhibitor was performed. The molecular descriptors were calculated and the best QSAR model was developed, which satisfied statistical parameters such as correlation coefficient \( R = 0.912 \) and leave-one-out validation coefficients \( q^2 = 0.608 \). The predictive quality of the model was tested against test set compounds with \( R^2_{\text{pred}} \) value of 0.7057. A novel compound (ND1) was designed and its predicted IC\(_{50}\) was predicted, which was lower compared with that of the parent compound (S24). Molecular docking and molecular dynamics simulation of 40 ns showed the stability of binding orientation of ND1, the parent compound, and native ligand of DPP4. Prediction of affinity using molecular mechanics/Poisson-Boltzmann/surface area method revealed that the ND1 has a comparable affinity with the parent and natural ligands.

INTRODUCTION
Type 2 diabetes mellitus (T2DM) is a chronic metabolic disease characterized by insulin resistance and insulin deficiency. It affects more than 400 million adults worldwide in 2014, and its global prevalence is estimated to reach 330 and 640 million in 2030 and 2040, respectively (Ahren, 2009; Drucker et al., 2010; Reusch and Manson, 2017).

One of the molecular targets for curing T2DM is dipeptidyl peptidase IV (DPPIV/DPP4). DPP4 is a serine protease which deactivates intestinally derived hormones glucagon-like peptide 1 (GLP-1) and glucose-dependent insulinotropic polypeptide (GIP). GLP-1 is an important incretin hormone that helps in insulin secretion and suppresses glucagon formation. However, its half-life is very short due to DPP4 catalytic activity. Therefore, inhibiting DPP4 is considered as a novel therapeutic strategy for restoring glucose homeostasis in diabetic patients by leaving more GLP-1 and GIP in the blood circulation (Chahal and Chowdhury, 2007; Demuth et al., 2005; Sneha and Doss, 2016).

Several DPP4 inhibitors have recently been approved by Food and Drug Administration (FDA), such as sitagliptin (Merck), vildagliptin (Novartis), saxagliptin (BMS), alogliptin (Takeda), and linagliptin (Lilly) (Cox et al., 2016). However, their use in the clinical application is not devoid of problems such as severe adverse effects, including hypoglycemia, edema, weight gain, and gastrointestinal distress (Li et al., 2016). Therefore, the need for novel and potent antidiabetic agents with minimum side effects is indispensable. Ji et al. (2014) designed and synthesized novel \( \beta \)-amino pyrrolo-2-carbonitrile derivatives and found that compound 9l showed excellent DPP4 inhibitory activity resulting in decreased blood glucose in vivo.
To further reveal and develop the structural features of novel derivatives of β-amino pyrrolo-2-carbonitrile having more potent DPP4 inhibitory activity, 2D-quantitative structure–activity relationship (QSAR) study was carried out to build a linear correlation between physicochemical properties (descriptors) and biological activity of the compounds. The molecular docking was employed to investigate the binding interaction of ligand to the DPP4 enzyme (Abdalsalam, 2017; Arba et al., 2018). The molecular dynamics (MD) simulation coupled with molecular mechanics/Poisson-Boltzmann/surface area (MM-PBSA) free energy prediction was applied to evaluate the ligand-enzyme dynamics during 40 ns and to predict the binding free energy of the ligand-enzyme complex (Ruslin et al., 2017).

**COMPUTATIONAL METHOD**

**Data set**

The data set of 24 4-fluoropyrrolidine-2-carbonitrile and octahydrocyclopenta[b]pyrrole-2-carbonitrile derivatives (Table 1) as reported by Ji et al. (2014) were selected. The inhibition activity data (IC$_{50}$) in negative logarithmic scale (pIC$_{50}$) was used, which span from −1.6761 up to 2. Each compound was built and geometrically optimized on semi-empirical Austin Model-1 (AM1) method by using Gaussian 09 software (Frisch et al., 2009). The molecular descriptors were generated for each built structure using molecular operating environment (MOE, 2009.10), which includes total energy (AM1_E), electronic energy (AM1_Eele), dipole moment (AM1_Dipol), formation heat (AM1_HF), highest occupied molecular orbital energy (AM1_HOMO), lowest unoccupied molecular orbital 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).

**Recognition of outlier**

Outliers are defined as unrelated values in the normal distribution values. In the QSAR model, they must be eliminated to avoid invalid prediction (Zakariazadeh et al., 2015). In the present study, outliers were identified by calculating the value of $Z$-SCORE, in which a compound is considered as an outlier when its $Z$-SCORE value $≥2.5$ (Hamerton et al., 2013). The remaining

**Table 1.** The 4-fluoropyrrolidine-2-carbonitrile and octahydrocyclopenta[b]pyrrole-2-carbonitrile derivatives in the data set.

<table>
<thead>
<tr>
<th>No.</th>
<th>Compound</th>
<th>R</th>
<th>IC$_{50}$ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A1</td>
<td>H</td>
<td>0.45</td>
</tr>
<tr>
<td>2</td>
<td>A2</td>
<td>2-Cl</td>
<td>0.22</td>
</tr>
<tr>
<td>3</td>
<td>A3</td>
<td>2-Me</td>
<td>0.16</td>
</tr>
<tr>
<td>4</td>
<td>A4</td>
<td>3-F</td>
<td>0.22</td>
</tr>
<tr>
<td>5</td>
<td>A5</td>
<td>4-F</td>
<td>0.32</td>
</tr>
<tr>
<td>6</td>
<td>A6</td>
<td>4-I</td>
<td>0.43</td>
</tr>
<tr>
<td>7</td>
<td>A7</td>
<td>4-CF$_3$</td>
<td>0.64</td>
</tr>
<tr>
<td>8</td>
<td>A8</td>
<td>4-OMe</td>
<td>0.77</td>
</tr>
<tr>
<td>9</td>
<td>A9</td>
<td>2,4-di-Cl</td>
<td>0.04</td>
</tr>
<tr>
<td>10</td>
<td>A10</td>
<td>3,4-di-Cl</td>
<td>0.53</td>
</tr>
<tr>
<td>11</td>
<td>A11</td>
<td>3,5-di-F</td>
<td>1.58</td>
</tr>
<tr>
<td>12</td>
<td>A12</td>
<td>2,4,5-tri-F</td>
<td>0.05</td>
</tr>
<tr>
<td>13</td>
<td>A13</td>
<td>H</td>
<td>1.94</td>
</tr>
<tr>
<td>14</td>
<td>B14</td>
<td>H</td>
<td>0.44</td>
</tr>
<tr>
<td>15</td>
<td>B15</td>
<td>2-Cl</td>
<td>0.07</td>
</tr>
<tr>
<td>16</td>
<td>B16</td>
<td>2-Me</td>
<td>0.36</td>
</tr>
<tr>
<td>17</td>
<td>B17</td>
<td>3-F</td>
<td>0.13</td>
</tr>
<tr>
<td>18</td>
<td>B18</td>
<td>4-F</td>
<td>0.27</td>
</tr>
<tr>
<td>19</td>
<td>B19</td>
<td>4-I</td>
<td>8.80</td>
</tr>
<tr>
<td>20</td>
<td>B20</td>
<td>4-CF$_3$</td>
<td>1.98</td>
</tr>
<tr>
<td>21</td>
<td>B21</td>
<td>2,4-di-Cl</td>
<td>0.07</td>
</tr>
<tr>
<td>22</td>
<td>B22</td>
<td>3,4-di-Cl</td>
<td>47.44</td>
</tr>
<tr>
<td>23</td>
<td>B23</td>
<td>3,5-di-F</td>
<td>9.65</td>
</tr>
<tr>
<td>24</td>
<td>B24</td>
<td>2,4,5-tri-F</td>
<td>0.01</td>
</tr>
</tbody>
</table>
compounds were then randomly grouped into a training set (18 compounds) and a test set (five compounds) considering their structural diversity and biological activities.

**QSAR model**

The calculation of the QSAR model was achieved by using the multiple linear regression approach using the SPSS for Windows version 19 to establish the linear relationship between a set of descriptors and biological activity. The best QSAR model was identified using several statistical parameters, such as squared correlation coefficient ($R^2$), Fischer’s value for statistical significance ($F$), adjusted squared correlation coefficient ($R^2_{adj}$), and standard error of estimation (Arba et al., 2018; Dearden et al., 2009). Besides, the QSAR equation was also validated by internal validation of leave-one-out (LOO) cross-validation coefficient ($q^{2}$) (Golbraikh and Tropsha, 2002). The LOO cross-validation works by eliminating each compound in the training set and predicting the biological activity using the model of remaining compounds. The calculation of LOO cross-validation coefficient ($q^{2}$) was described elsewhere (Arba et al., 2018), in which a value of $q^{2}$ which is higher than 0.5 is necessary to assure the predictability of the built QSAR equation (Golbraikh et al., 2003; Tropsha et al., 2003). Furthermore, in addition to internal cross-validation, the reliability of the built model was also evaluated externally, in which the model was used to predict the biological activity of the test set compounds (Tropsha et al., 2003). In that scheme, the validity of the model was evaluated using the external cross-validation coefficient ($R^2_{ext}$), in which the $R^2_{ext}$ value higher than 0.6 is necessary to assure the validity of the built model.

**Design for new molecule and molecular docking**

Based on built QSAR equation, a new compound (ND1) was proposed and its predicted biological activity (IC_{50pred}) was calculated. The biological activity of the ND1 was then compared with that of the parent compound (S24), which has the lowest observed IC_{50} (IC_{50} = 0.01 μM). Furthermore, molecular docking of ND1 and S24 to DPP4 protein was performed to reveal the interaction mechanism of those compounds in the binding site of DPP4 by using AutoDock 4.2 (Morris et al., 1998). The crystallographic structure of DPP4 from the Protein Data Bank (PDB) with PDB code 2AJL and resolution 2.5 Å was selected (Qiao et al., 2006). The grid box was located following the catalytic cavity with the dimension of 50 in each xyz direction. The default values were used for other docking parameters (Arba et al., 2017a).

**MD simulation and prediction of binding free energy**

MD simulation was performed on each S24, ND1, and native ligand 1-[2-(s)-amino-3-biphenyl-4-yl-propionyl]-pyrrolidine-2-(s)-carbonitrile (JNH), each complexed with DPP4 using Amber 16 package (Case et al., 2015; Salomon-Ferrer et al., 2013). In the present study, the ff14SB force field was used to describe the protein, while GAFF (General Amber force field) and AM1-BCC were used to describe ligand (Jakalian et al., 2002; Maier et al., 2015; Wang et al., 2004). Each system was neutralized by the addition of Na+ ions and then solvated using the TIP3P box water model with a distance of 1 nm around the complex. The minimization, heating, and equilibration were performed with the aid of Sander module following our previous procedure (Arba et al., 2018). The production step was performed using GPU version of the PMEMD engine of Amber 16 package for 40 ns in NPT ensemble without restraint using Langevin thermostat at 1.0 ps⁻¹ random collision frequency to maintain the system in 300 K thermal bath. The SHAKE algorithm was used to restrain bonds involving hydrogen atoms (Ryckaert et al., 1977). The long-range electrostatic interactions were treated with the particle-mesh Ewald method with an integration step of 2 fs (Darden et al., 1993) by applying periodic boundary conditions with a cutoff distance of 9.0 Å. The CPTRAJ module (Roe and Cheatham, 2013) was used to perform analysis, including root mean square deviations (RMSD) and root mean square fluctuation (RMSF), while visual molecular dynamics was used for visualization (Humphrey et al., 1996).

The prediction of free energy of binding was achieved by performing MM/PBSA following our previous protocol (Arba et al., 2017b; 2018; Kollman et al., 2000; Miller et al., 2012).

**RESULTS AND DISCUSSION**

The QSAR study was performed to reveal structure–activity relationship of β-amino pyrrolo-2-carbonitrile as a DPP4 inhibitor. The outlier of the data set was first determined using SZ-SCORE. Calculation of the values of SZ-SCORE identified S22 and S24 as outliers with SZ-SCORE values of 3.10848 and 2.7395, respectively. Theoretically, both compounds should be removed from the data set to improve the QSAR model. However, in the present study, only S22 was removed, while compound S24 was kept in the data set since it is the most active compound experimentally. Furthermore, 23 compounds of the data set were divided randomly into a training set (18 compounds), which was used to build QSAR model, and a test set (five compounds), which was used to test the predictive ability of the built model (Table 2). The test set was selected by considering the distribution of biological activity in the whole data set.

Furthermore, multiple linear regression analysis was applied to build the QSAR models using 18 compounds of a training set. The resulted QSAR model contains five molecular descriptors, i.e., dipole moment (AM1_dipol), HOMO energy (AM1_HOMO), LUMO energy (AM1_LUMO), partition coefficient (Log P), and molar refractivity (Mr). The following equation shows the best QSAR model:

$$
\text{pIC}_{50} = 75.842 - 0.579 \text{AM1_dipol} + 5.359 \text{AM1_HOMO} - 5.297 \text{AM1_LUMO} + 2.278 \text{Log P} - 3.070 \text{Mr}
$$

The above equation fulfills statistical criteria such as the correlation coefficient ($R$), determination coefficient ($R^2$), and Fischer’s value ($F$) of 0.912, 0.831, and 11.820, respectively. The quality of the model was also indicated by the low standard error (SE) of 0.3290. The value of LOO cross-validation coefficient ($q^2$) of 0.608 indicated that the model was valid. Table 3 shows molecular descriptors and statistical parameters of the built QSAR model.

The QSAR model indicates that the biological activity would increase with more lipophilic groups as indicated by the positive sign of coefficient of Log P. On the other hand, less steric groups were favorable for increasing the biological activity as indicated by the negative sign of the coefficient of Mr. Meanwhile,
the negative contribution of AM1_Dipole and AMI_LUMO to biological activity was noted as implied by the negative sign of their coefficients. Compared to the other four descriptors, AM1_HOMO is the most influencing descriptor as indicated by the highest coefficient value. Furthermore, Tropsha et al. (2003) indicated that a QSAR model must be validated externally by test set compounds, in which $R^2_{pred}$ value of training and test set must be higher than 0.6. Our QSAR analysis revealed the value of $R^2_{pred} = 0.7057$, indicating the validity of the built model. The relationship between observed and predicted $pIC_{50}$ both training and test set is depicted in Figure 1.

**The design of new compound and molecular docking**

With the aim of finding the new potent β-amino pyrrole-2-carbonitrile derivative, a ND1 was designed by using the built QSAR model. Table 4 shows the structures of S24 and ND1. The results of predicting biological activity revealed that the predicted

### Table 2. The value of $Z$-SCORE and data set division. Compounds assigned as * and ** are outlier and test set, respectively.

<table>
<thead>
<tr>
<th>No.</th>
<th>Compound</th>
<th>Code</th>
<th>$pIC_{50}$</th>
<th>$Z$-SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A1</td>
<td>S1</td>
<td>0.34679</td>
<td>0.02637</td>
</tr>
<tr>
<td>2</td>
<td>A2</td>
<td>S2</td>
<td>0.65758</td>
<td>0.34277</td>
</tr>
<tr>
<td>3</td>
<td>A3</td>
<td>S3</td>
<td>0.79588</td>
<td>0.63559</td>
</tr>
<tr>
<td>4</td>
<td>A4</td>
<td>S4</td>
<td>0.65758</td>
<td>0.28256</td>
</tr>
<tr>
<td>5</td>
<td>A5</td>
<td>S5**</td>
<td>0.49485</td>
<td>0.05558</td>
</tr>
<tr>
<td>6</td>
<td>A6</td>
<td>S6</td>
<td>0.36653</td>
<td>0.06973</td>
</tr>
<tr>
<td>7</td>
<td>A7</td>
<td>S7</td>
<td>0.19382</td>
<td>0.6272</td>
</tr>
<tr>
<td>8</td>
<td>A8</td>
<td>S8</td>
<td>0.11351</td>
<td>0.3386</td>
</tr>
<tr>
<td>9</td>
<td>A9</td>
<td>S9</td>
<td>1.39794</td>
<td>1.288</td>
</tr>
<tr>
<td>10</td>
<td>A10</td>
<td>S10**</td>
<td>0.27572</td>
<td>0.25733</td>
</tr>
<tr>
<td>11</td>
<td>A11</td>
<td>S11</td>
<td>−0.1987</td>
<td>1.0489</td>
</tr>
<tr>
<td>12</td>
<td>A12</td>
<td>S12**</td>
<td>1.30103</td>
<td>1.14483</td>
</tr>
<tr>
<td>13</td>
<td>A13</td>
<td>S13**</td>
<td>−0.2878</td>
<td>0.95978</td>
</tr>
<tr>
<td>14</td>
<td>B14</td>
<td>S14</td>
<td>0.35655</td>
<td>0.24665</td>
</tr>
<tr>
<td>15</td>
<td>B15</td>
<td>S15</td>
<td>1.1549</td>
<td>1.34364</td>
</tr>
<tr>
<td>16</td>
<td>B16</td>
<td>S16</td>
<td>0.4437</td>
<td>0.45808</td>
</tr>
<tr>
<td>17</td>
<td>B17</td>
<td>S17</td>
<td>0.88606</td>
<td>0.82957</td>
</tr>
<tr>
<td>18</td>
<td>B18</td>
<td>S18</td>
<td>0.56864</td>
<td>0.3929</td>
</tr>
<tr>
<td>19</td>
<td>B19</td>
<td>S19</td>
<td>−0.9445</td>
<td>1.69089</td>
</tr>
<tr>
<td>20</td>
<td>B20</td>
<td>S20</td>
<td>−0.2967</td>
<td>1.18939</td>
</tr>
<tr>
<td>21</td>
<td>B21</td>
<td>S21</td>
<td>1.1549</td>
<td>1.25829</td>
</tr>
<tr>
<td>22</td>
<td>B22</td>
<td>S22*</td>
<td>−1.6761</td>
<td>3.10848</td>
</tr>
<tr>
<td>23</td>
<td>B23</td>
<td>S23**</td>
<td>−0.9845</td>
<td>1.90469</td>
</tr>
<tr>
<td>24</td>
<td>B24</td>
<td>S24</td>
<td>2</td>
<td>2.7395</td>
</tr>
</tbody>
</table>

### Table 3. Statistical parameters of best QSAR equation for β-amino pyrrole-2-carbonitrile derivatives.

<table>
<thead>
<tr>
<th>Descriptors</th>
<th>$q^2$</th>
<th>$R$</th>
<th>$R^2$</th>
<th>Adjusted $R^2$</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>AM1_Dipole, AM1_HOMO, AM1_LUMO, Log P(o/w), Mr</td>
<td>0.608</td>
<td>0.912</td>
<td>0.831</td>
<td>0.761</td>
<td>0.329</td>
</tr>
</tbody>
</table>

### Table 4. The structures of S24 and ND1.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Structure</th>
<th>Predicted $IC_{50}$ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent compound (S24)</td>
<td><img src="image" alt="Parent compound" /></td>
<td><strong>0.015279802</strong> (experiment)</td>
</tr>
<tr>
<td>Novel compound (ND1)</td>
<td><img src="image" alt="Novel compound" /></td>
<td><strong>0.000340859</strong> (calculation)</td>
</tr>
</tbody>
</table>
IC\textsubscript{50} of ND1 was lower (IC\textsubscript{50pred} = 0.000340859 µM) than that of S24 (IC\textsubscript{50pred} = 0.015279802 µM).

Next, molecular docking of S24 and ND1 on DPP4 was performed to examine the preferred orientation of ligand to the protein. Molecular docking was commenced by redocking native ligand (JNH) on the protein to examine the reliability of docking protocol. In the docked conformation, JNH established hydrogen bonds (H-bonds) with Glu206 and Tyr547 as well as hydrophobic interactions with Phe357 and His740. The H-bond with Glu205 and hydrophobic interactions with Phe357 and His740 was also observed in the X-ray crystallographic conformation. Figure 2 shows the docked and X-ray crystallographic conformations of JNH with the RMSD of 1.21 Å, indicating that the docking protocol was valid (Jones et al., 1997; Morris et al., 1998).

The molecular docking poses showed that both S24 and ND1 were able to interact with active site residues of DPP4. Several H-bonds were formed with Glu205, Glu206, and Arg669 in the binding of ND1. Interestingly, fluor atom interacted through H-bonds with Arg125, Tyr631, Asn710, and His740. Moreover, pi-pi stacking interactions between ND1 with Tyr662 and Tyr666 were also established. On the other hand, binding of S24 was maintained by H-bonds with Glu205, Glu206, and Asn710, as well as by pi-pi stacking interactions with Tyr662 and Tyr666. It is noted that H-bonds with Glu205 and Asn710 were also detected in X-ray crystallographic pose. The hydrophobic interaction including with Phe357 was also observed. It was clear from the docked poses that ND1 established more interactions that that of S24. Figure 3 depicts the binding mode of S24 and ND1 in the binding cavity of DPP4.

Figure 2. The docked (green) and experimental (blue) poses of JNH and the interaction of JNH in the binding site of DPP4.

Figure 3. The docked conformation of (a) S24 and (b) ND1. The hydrogen bond and pi-pi stacking interactions are represented in green and pink colored dashed lines, respectively.

Figure 4. The RMSD plot of each JNH (red), S24 (green), and ND1 (blue) complexed to DPP4.
Molecular dynamics simulation

The docked conformation of each S24, ND1, and JNH was monitored for their conformational stability through MD simulation of 40 ns. The conformation stability was checked by using the values of RMSD. Figure 4 shows the RMSD plot of heavy atoms of DPP4 with respect to simulation time for each complex. As Figure 4 shows, JNH and ND1 reached stability after about 7 ns. On the other hand, the complex of the parent compound (S24) shows slight fluctuation after 20 ns. However, considering its fluctuation, which is about 3 Å, it can be inferred that the complex was sufficiently stable. The dynamics simulation was also used to monitor the fluctuation of amino acid residues due to ligand binding during 40 ns. Figure 5 shows the RMSF pattern of protein versus residue number. All complexes show a similar pattern of RMSF, which indicates the similar binding mode. The highest fluctuation was recorded in the amino acid residue Ser205 (Ser245) which was due to beta helix end.

Binding energy prediction by MM/PBSA

Table 5 shows the calculated binding free energy of each S24, ND1, and JNH bound to the DPP4. The MM-PBSA prediction showed that the binding free energy is slightly lower in parent compound (S24) (Δ\(E_{\text{PBTOT}}\) = −26.73 kcal/mol) and ND1 (Δ\(E_{\text{PBTOT}}\) = −22.14 kcal/mol) compared to JNH (Δ\(E_{\text{PBTOT}}\) = −19.96 kcal/mol). The better affinities of S24 and ND1 were also reflected by the total number of nonbond interactions during simulation period of 40 ns (Fig. 6), in which both S24 and ND1 have more number of interactions than that of JNH. The S24-DPP4 complex also displayed a slightly lower electrostatic energy (Δ\(E_{\text{ELE}}\) = −20.44 kcal/mol) compared to the ND1 and JNH complexes (Δ\(E_{\text{ELE}}\) = −17.92 and −2.17 kcal/mol, respectively). However, van der Waals energy was slightly lower in ND1 (Δ\(E_{\text{VDW}}\) = −41.46 kcal/mol) compared to S24 and JNH (Δ\(E_{\text{VDW}}\) = −38.39 and −35.03 kcal/mol, respectively). Meanwhile, the contribution of the nonpolar energy of desolvation (Δ\(E_{\text{PRESUR}}\)) was almost the same for all complexes. The contribution of the polar energy of desolvation is higher in ND1 (Δ\(E_{\text{PRESUR}}\) = 41.04 kcal/mol) compared to S24 and JNH (Δ\(E_{\text{PRESUR}}\) = 37.48 and 21.84 kcal/mol, respectively). In the meantime, the total electrostatic contribution was lower in S24 (Δ\(E_{\text{PBELE}}\) = 17.04 kcal/mol) compared to the ND1 and JNH (Δ\(E_{\text{PBELE}}\) = 23.12 and 19.67 kcal/mol, respectively), resulting in the lowest total binding free energy in S24 compared to those in ND1 and JNH.

CONCLUSION

In the present study, a QSAR model of five descriptors which correlate structure and inhibitory activity of DPP4 of \(\beta\)-amino pyrrolo-2-carbonitrile derivatives was developed. It was then used to design ND1 which has lower predicted IC\(_50\) than the parent compound. The ND1 interacted with the active site of the DPP4 protein and its complex with DPP4 protein was stabilized during 40 ns MD simulation. This study identifies ND1, with binding affinity amenable to the further study of the discovery of DPP4 inhibitor.

REFERENCES


How to cite this article: