Scientific paper

Tertiary Structure Prediction of α-Glucosidase and Inhibition Properties of N-(Phenoxydecyl) Phthalimide Derivatives

Maryam Pooyafar and Davood Ajloo*

School of Chemistry, Damghan University, Damghan, Iran

** Corresponding author: E-mail: ajloo@du.ac.ir*
Tel. +98 (0232)-5233051-6; fax: +98 (0232)-5235713; Zipcode : 3671641167

Received: 11-09-2011

Abstract

Due to increasing of population of diabetic patients, identifying factors for disease control has received much attention. α-glucosidase (EC 3.2.1.20) is an essential enzyme that helps to digestion of carbohydrates such as starch and sugar. Carbohydrates are normally converted into simple sugars, which can be absorbed through the intestine. Therefore, α-glucosidase inhibitors can be used to decrease the blood sugar level. We have studied the effect of inhibition of N-(phenoxydecyl) phthalimide derivatives by a computer drug-design protocol involving homology modeling, docking simulation and Quantitative Structure Activity Relationship. The homology modeling of α-glucosidase showed a structure very similar to the crystal structure of oligo-1,6-glucosidase from Saccharomyces cerevisiae. Docking results showed the position of inhibitors binding site is close to active site and the carboxyl oxygen in phthalimide is an effective functional group for binding inhibitors to protein. The equation obtained by QSAR showed that, $\log IC_{50}$ decreases and so inhibition property increases when the size, polarity, geometry and number of halogen factors increase.

Keywords: α-Glucosidase, Inhibition, Homology modeling, Docking, Molecular dynamics simulation, QSAR

1. Introduction

Diabetes is one of the most serious, chronic diseases that is developing with the increase in obesity and ageing in the general population. Persistent hyperglycemia in diabetic patients despite appropriate therapeutic measures leads to several complications including retinopathy, nephropathy, and neuropathy.[1] Some drugs have been developed for diabetes, and the best way to control postprandial plasma glucose level is medication in combination with dietary restriction and an exercise program. One of the therapeutic approaches for decreasing of postprandial hyperglycemia is to retard absorption of glucose by the inhibition of carbohydrate-hydrolysing enzymes, for example α-glucosidase, in the digestive organs.[2] Glucosidases are responsible for the catalytic cleavage of glycosidic bond in the digestive process of carbohydrates with specificity depending on the number of monosaccharides, the position of cleavage site, and the configuration of the hydroxyl groups in the substrate.[3] α-glucosidase (EC 3.2.1.20) has taken a special interest of the pharmaceutical research community because it was shown that the inhibition of its catalytic activity lead to the retardation of glucose absorption and the decrease in postprandial blood glucose level. This indicates that effective α-glucosidase inhibitors may serve as chemotherapeutic agents for clinic use in the treatment of diabetes and obesity. The catalytic role in digesting carbohydrate substrates also makes α-glucosidase a therapeutic target for the other carbohydrate-mediated diseases including cancer, viral infections and hepatitis[4] and therefore many efforts have been made to identify new inhibitors for α-glucosidase[5–8]. On the other hand, traditional and experimental methods that are used in drug design are expensive and time consuming. Quantitative structure-activity relationship (QSAR) information of α-glucosidases has been limited to those of a few bacterial strains only in ligand-free forms.[9] The lack of structural information about the nature of the interactions between α-glucosidases and the inhibitors has thus made it a difficult task to discover good lead compounds based on the structure-based inhibitor design. Consequently docking simulation can be a useful tool for elucidating the ob-

served activity of the identified inhibitors and binding modes.[10]

So far, many studies have been done to investigate the interaction of various inhibitors on the α-glucosidases. In particular, based on pharmacological studies involving thalidomide, it was found that phenylalkyl tetrachlorophthalimide derivatives exhibited potent α-glucosidase inhibition.[11]

In this study, we investigated binding mode of N-(phenoxydecyl)phthalimide derivatives to α-glucosidase, by means of a drug-design protocol involving homology modeling, docking simulations, and quantitative structure-activity relationship.[10] We chose the α-glucosidase from baker's yeast as the target protein in docking because it had been used widely in biological assays to identify new α-glucosidase inhibitors.

2. Methods and Materials

2. 1. Homology Modeling: 3D Structure Prediction Using Computational Methods

Since there is no structural information for α-glucosidase from baker's yeast, we carried out homology modeling to obtain its three-dimensional structure. Primary sequence of the protein includes 584 amino acid residues and was taken from the Swiss-Prot protein sequence data bank (http://www.expasy.org/sprot/;Accession No. P53341).[12] To get a suitable structural template for homology modeling, we searched for the Protein Data Bank (http://www.rcsb.org/pdb/) using BLAST algorithm with the amino acid sequence of the target as input. Oligo-1,6-glucosidase from Saccharomyces cerevisiae has the highest sequence identity with the target. So, its X-ray crystal structure (PDB ID: 3A47)[13] was selected as the template for modeling. The primary structures of model and template share 72% sequence identity. The initial sequence alignment of protein (Uniprot ID: P533411) with structural template (PDB ID: 3A47) was carried out using the ClustalW program with BLOSUM matrices for scoring the alignments. Based on the highest alignment, the structure of α -glucosidase from baker's yeast was constructed using the MODELLER 9V8[14] program. One thousand models were made and the best model was selected based on DOPE score. DOPE (Discrete Optimized Protein Energy) is the total conformational energy of any amino acids which made by MODELLER for all models. The model which has the most negative energy will be the most stable structure and introduced as the best model. This model was validated by a well-established program, PROCHECK[15] for the evaluation of Ramachandran plot.

The molecular dynamics simulation was done by Gromacs 3.3.1[16] to increase the accuracy of the calculated structure and optimization of the final model. Finally, we investigated Root Mean Square Deviation (RMSD) plot during 20 ns simulation.

2. 2. Docking Simulation

AutoDock 3.05 program was used[17] to obtain some energetic and structural insight into the inhibitory mechanisms of the identified inhibitors of α-glucosidase, as well as their binding modes in the active site. The ligands were designed in Hyperchem7 and optimized with AM1 semi-empirical method. Docking simulation was done in three boxes: first box was taken around the whole of protein with dimension 126 × 126 × 126 points and the spacing of 0.703 Å. Second box was taken around the active site of protein, with dimension of 52 × 60 × 74 points and spacing of 0.375 Å. Finally, after blind docking discussed in the first step, the most populated site was determined and the box with dimension of 52 × 58 × 66 points and the spacing of 0.375 Å was located in that site. For each ligand, 250 docking runs were performed with the initial population of 150 individuals. Maximum number of generations and energy evaluations were set to 27,000 and 2.5 × 10^5, respectively.

2. 3. Quantitative Structure-Activity Relationship

QSAR studies were applied to predict logIC_{50} and find the relationship between structure and activity. The ligands were designed by Hyperchem7.0 and then optimized by semi-empirical AM1 method. These ligands were transferred to Dragon-3.0 software and generate 1497 descriptors. All descriptors that had zero values or constant values in the data set were eliminated. The remaining descriptors were used to generate the prediction models using the SPSS 17 software package. Multiple linear regression method (MLR) and principal component analysis (PCA) were used to select descriptors which are responsible for half maximal inhibitory concentration (IC_{50}) parameters of these compounds. PCA involves a mathematical procedure that reduces and classifies descriptors to the new sets of them.

2. 4. Cross-Validation Technique

Since a high-correlation coefficient only indicates how well the equations fit the data, cross-validation procedure was carried out in order to explore the reliability of the proposed models. In this aspect, the well-known "leave-one-out" (LOO) approach was used in which a number of models were developed with one sample ignored each time. Then, the ignored data were predicted by each model and the differences between predicted and observed activity values were evaluated.[18]

The cross-validation parameters (q^2_{cv} and PRESS) are mentioned in the respective equation

$$q^2_{cv} = \frac{(SD-PRESS)}{SD} \quad (1)$$

Which SD is standard deviation and PRESS (predictive residual sum of squares) is the sum of the squared differences between the actual and that predicted. A good q^2_{cv} value should be always smaller than R^2. A model is considered to be significant when $q^2_{cv} > 0.3$.[18]

3. Results and Discussion

3. 1. Protein 3D Structure Prediction Using Homology Modeling Method

Figure 1 shows the sequence alignment between α-glucosidase MAL12 from baker's yeast (EC 3.2.1.20) and crystal structure of oligo-1,6-glucosidase from *saccharomyces cerevisiae* (3A47), (EC 3.2.1.10). According to this alignment, the sequence identity and the similarity amounts are 72% and 85%, respectively. Judging from such a high sequence homology, a high-quality 3D structure of α-glucosidase can be expected in the homology modeling. It is indeed well known that a homology-modeled structure of a target protein can be accurate enough to be used in docking studies. Based on the sequence alignment shown in figure 1, structural models of α-glucosidase were calculated and one that has the lowest value of MODELLER objective function was selected as the final model.

Figure 2 displays the structure of α-glucosidase obtained from the homology modeling in comparison with the X-ray crystal structure of oligo-1,6-glucosidase from saccharomyces that was used as the template. The target and the template possess a very similar structure. The two enzymes also share the catalytic residues that are situated in their respective active sites in a similar fashion. To this way that Asp215, Glu277 and Asp352 residues form the catalytic triad in the template protein while the Asp214, Glu276 and Asp349 form in S. cerevisiae α-glucosidase, respectively. Two more residues, His109 and His348 of oligo-1,6-glucosidase which may be involved in substrate binding are also conserved in α-glucosidase (His111 and His348, respectively). It is because of both enzymes catalyze the hydrolysis of terminal glycosidic bond of carbohydrates. However, α-glucosidase is more extended in

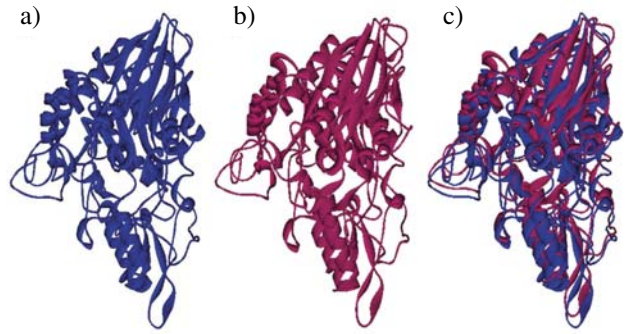


Figure 2: Stereo view of template (A), modeled (B) and both fitted (C) structure of α-glucosidase

```
template  S- - SAHPETEPKWWKEATFYQIYPASFKDSNDDGWGDMKGIASKLEYIKELGADAIWISPFYDSPQDD
query     MTISDHPETEPKWWKEATIYQIYPASFKDSNNDGWGDLKGITSKLQYIKDLGVDAIWVCPFYDSPQQD

template  MGYDIANYEKVWPTYGTNEDCFALIEKTHKLGMKFITDLVINHCSSEHEWFKESRSSKTNPKRDWFFW
query     MGYDISNYEKVWPTYGTNEDCFELIDKTHKLGMKFITDLVINHCSTEHEWFKESRSSKTNPKRDWFFW

template  RPPKGYDAEGKPIPPNNWKSYFGGSAWTFDEKTQEFYLRLFCSTQPDLNWENEDCRKAIYESAVGYWL
query     RPPKGYDAEGKPIPPNNWKSFFGGSAWTFDETTNEFYLRLFASRQVDLNWENEDCRRAIFESAVGFWL

template  DHGVDGFRIDVGSLYSKVVGLPDAPVVDKNSTWQSSDPYTLNGPRIHEFHQEMNQFIRNRVKDGREIM
query     DHGVDGFRIDTAGLYSKRPGLPDSPIFDKTSKLQHPNWGSHNGPRIHEYHQELHRFMKNRVKDGREIM

template  TVGEMQHASDETKRLYTSASRHELSELFNFSHTDVGTSPLFRYNLVPFELKDWKIALAELFRYINGTD
query     TVGEVAHGSDNA- -LYTSAARYEVSEVFSFTHVEVGTSPFFRYNIVPFTLKQWKEAIASNFLFINGTD

template  CWSTIYLENHDQPRSITRFGDDSPKNRVISGKLLSVLLSALTGTLYVYQGQELGQINFKNWPVEKYED
query     SWATTYIENHDQARSITRFADDSPKYRKISGKLLTLLECSLTGTLYVYQGQEIGQINFKEWPIEKYED

template  VEIRNNYNAIKEEHGENSEEMKKFLEAIALISRDHARTPMQWSREEPNAGFSGPSAKPWFYLNDSFRE
query     VDVKNNYEIIKKSFGKNSKEMKDFFKGIALLSRDHSRTPMPWTKDKPNAGFTGPDVKPWFLLNESFEQ

template  GINVEDEIKDPNSVLNFWKEALKFRKAHKDITVYGYDFEFIDLDNKKLFSFTKKYNNKTLFAALNFSS
query     GINVEQESRDDDSVLNFWKRALQARKKYKELMIYGYDFQFIDLDSDQIFSFTKEYEDKTLFAALNFSG

template  DATDFKIPNDDSSFKLEFGNYPKKEVDASSRTLKPWEGRIYISE
query     EEIEFSLPREGASLSFILGNYD- -DTDVSSRVLKPWEGRIYLVK
```

Figure 1: Sequence alignment between α-glucosidase (EC.3.2.1.20) and oligo-1,6-glucosidase (3A47A). The identity and the similarity between the corresponding residues are indicated in red and green, respectively. The active sites characterized by blue color.

amino acid packing than oligo-1,6-glucosidase due to the possession of 5 less amino acid residues in the alignment position.

The results of validation with PROCHECK showed 84.5% of residues of 3D structure are located in favorite region.

Figure 3 shows the DOPE (Discrete Optimized Protein Energy) score profile energy of the homology-modeled α-glucosidase in comparison to that of the X-ray structure of oligo-1,6-glucosidase. In this work, molecular dynamics simulations and energy minimization were performed by GROMACS. Figure 4 shows the corresponding Root Mean Square Deviation (RMSD) plot. As can be seen the system has reached stability during 20 ns.

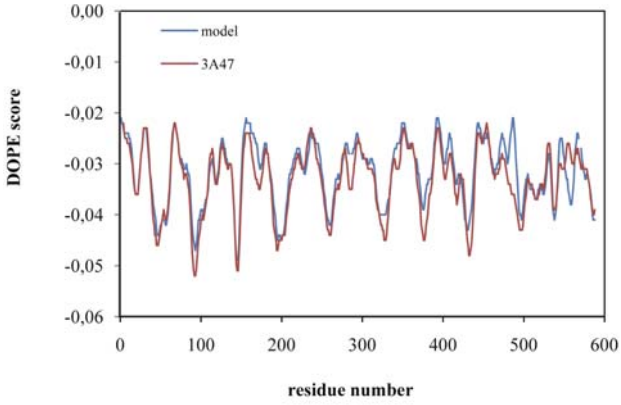


Figure 3: Comparison of the DOPE energy profiles for the homology-modeled structure of α-glucosidase (blue) and the X-ray structure of oligo-1,6-glucosidase (red)

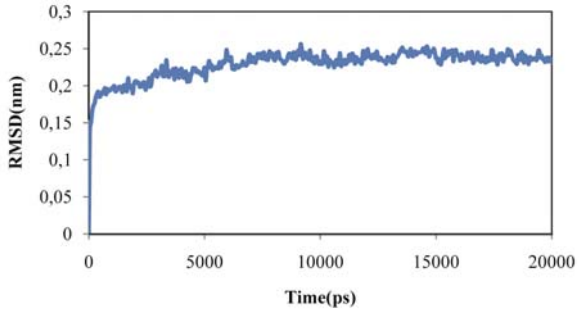


Figure 4: RMSD plot during 20 ns simulation obtained by molecular dynamics

3. 2. Molecular Docking

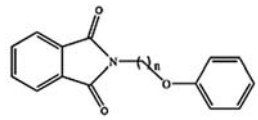
Docking studies were performed to check the most probable binding site of new N-(phenoxydecyl) phthalimide drivatives. The chemical structures and the inhibitory activities of these newly identified inhibitors were shown in Tables 1–3.[20]

In order to survey the correlation between inhibitory activities and binding energy, we made the box that surrounded whole of protein, then the position of the most negative docking energies were analyzed and it was found that the major of them locate near the active. We called that site as docking site. Then we put the box in the binding site. Alternative site was active site that we put the box on active site. So we have three cases; whole of protein, active site and binding site.

Figure 5 shows docking sites (Lys155, Phe157, Glu304, Arg312) and active sites (Asp214, Glu276 and Asp349) of ligands to protein as well as position of two sites relative to each other. These results verify that, the ligands prefer to bind near the active site.

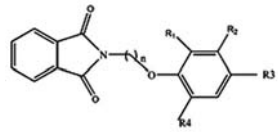
Figure 6 illustrates the calculated binding mode of compounds 20k and 25a (as the samples) in the active site and docking site of α-glucosidase. It is noteworthy that the carboxyl oxygen in phthalimide forms hydrogen bond with Asp349 and Glu276 in both position of active site

Table 1: α-Glucosidase inhibitory activity of substituted N-(phenoxyalkyl) phthalimide derivatives



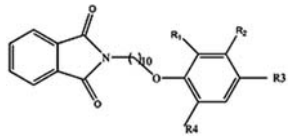
Compound	n	IC_{50}(μM)
8a	2	296 ± 4
9a	3	240 ± 40
25a	4	33 ± 3
10a	5	20.2 ± 0.2
11a	6	10.3 ± 0.1
18a	8	6.5 ± 0.2
19a	9	3.04 ± 0.01
20a	10	2.5 ± 0.2

Table 2: α-Glucosidase inhibitory activity of N-(phenoxyalkyl) phthalimide derivatives



Compound	n	R_1	R_2	R_3	R_4	IC_{50}(μM)
8a	2	H	H	H	H	296 ± 4
8b	2	H	H	Cl	H	94 ± 4
8e	2	CH_3	H	Cl	H	14.8 ± 0.8
8f	2	CH_3	H	Cl	CH_3	13.0 ± 0.1
9a	3	H	H	H	H	240 ± 40
9b	3	H	H	Cl	H	59.0 ± 0.5
9c	3	Cl	Cl	H	H	15.0 ± 0.8
9d	3	H	Cl	Cl	H	7.55 ± 0.25
9e	3	CH_3	H	Cl	H	8.9 ± 0.4
9f	3	CH_3	H	Cl	CH_3	9.5 ± 0.3

Table 3: α-Glucosidase inhibitory activity of N-(phenoxydecyl) phthalimide derivatives



Compound	R_1	R_2	R_3	R_4	IC_{50}(μM)
20a	H	H	H	H	2.5 ± 0.2
20b	H	H	Cl	H	1.2 ± 0.2
20d	H	Cl	Cl	H	1.00 ± 0.01
20f	CH_3	H	Cl	CH_3	1.34 ± 0.04
20g	Cl	H	Cl	H	1.19 ± 0.11
20h	H	CF_3	NO_2	H	0.83 ± 0.05
20i	H	H	CH_3	H	2.10 ± 0.01
20j	H	H	NO_2	H	0.86 ± 0.03
20k	NO_2	H	H	H	1.78 ± 0.01
20l	H	NO_2	H	H	1.075 ± 0.005
20m	CH_3	H	H	CH_3	3.25 ± 0.02
20n	H	Cl	H	H	1.09 ± 0.08
20o	H	Cl	H	CH_3	0.91 ± 0.05
20p	H	H	CF_3	H	0.83 ± 0.12
20q	NO_2	H	Cl	H	0.475 ± 0.05
20r	H	Cl	Cl	NO_2	0.52 ± 0.02
20s	Cl	H	Cl	NO_2	0.75 ± 0.01
20t	NO_2	H	NO_2	H	0.97 ± 0.04
20u	NO_2	H	CF_3	H	1.31 ± 0.01
20v	CH_3	NO_2	Cl	CH_3	3.6 ± 1.5
20w	CH_3	H	NO_2	CH_3	1.32 ± 0.12
20x	H	Cl	NO_2	CH_3	0.65 ± 0.01

Figure 5: Position of active site (blue) and docking site (pink) obtained from docking calculation

and binding site. Since, these two residues play an important role in the most inhibitors of α-glucosidase,[21] it looks like that the carboxyl oxygen can be an effective functional group for binding of inhibitors to protein and consequently phthalimide and its derivatives could be introduced as new inhibitors. The residues of two ligands mentioned above were listed in Table 4. The most negative free energies obtained by Autodock and predicted logIC_{50} were presented in Table 5.

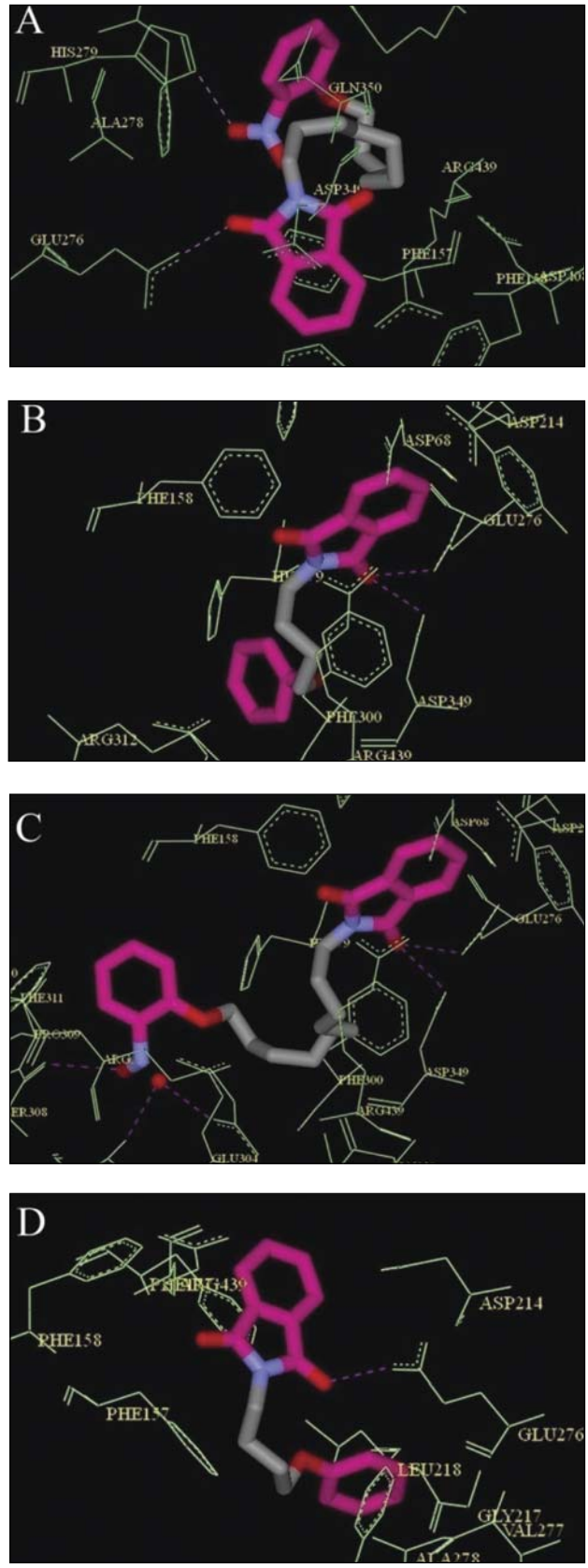


Figure 6: Binding mode of 20k and 25a as the samples in the active site (A,B) and docking site(C,D) respectively, pink dotted lines indicate hydrogen bonds.

Table 4: Amino acids residues in the active site and docking site for compounds 20k and 25a. The common parts in two sites were depicted as bold.

	20k	25a
Active site	Phe157, **Phe158**, **Phe177**, Thr215, **Glu276**, Ala 278, **His279**, **Phe 300**, **Glu304**, **Arg312**, **Asp349**, **Gln350**, Asp408, **Arg439**	**Phe158,Phe177**, **Asp214**, **Glu276**, His279, **Phe300**, Glu304, Arg312, Asp349, Gln350
Docking site	**Phe158, Phe177**, Asp214, **Glu276**, **His279**, **Phe300**, Val303, **Glu304**, Thr307,Ser308,PRO309, Phe311, **Arg312**, **Asp349**, **Gln350, Arg439**	Phe157, **Phe158, Phe177**, **Asp214**, His245, **Glu276**, Ala278, **Phe300**, Arg439

Correlation between docking energies and log IC50 were obtained and correlation coefficient (*R*) for three docking cases; whole of enzyme, active site and binding site are equal to 0.290, 0.185 and 0.801. Figure 6 shows correlation between docking energy and $\log IC_{50}$ only for docking site that has better correlation. Thus, we conclude docking site has higher correlation than others. Thus, more tendency for interaction between ligand and enzyme has less $\log IC_{50}$ and more ligand efficiency.

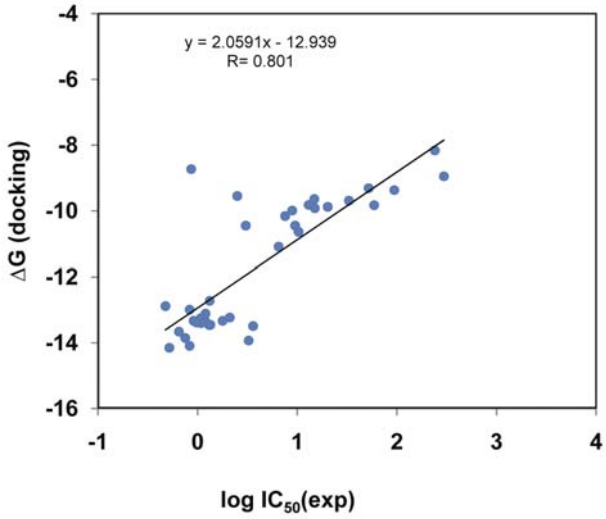


Figure 7: Correlation between docking free energy (ΔG) and inhibition activities (log IC_{50}) for docking site

3. 3. Quantitative Structure-Activity Relationship

The 1497 molecular properties (descriptors) were calculated for 37 ligands using Dragon-3.0. Logarithm of inhibitory activities was employed as dependent variables to find the relationship between structure and activity. Multiple linear regression analysis (MLR) of molecular descriptors was carried out using the stepwise strategy in the SPSS 17. The best obtained regression equation for the logarithm of inhibitory activities of 37 derivatives by MLR model is:

Table 5. The most negative docking free energies and predicted $\log IC_{50}$

Compound	Whole enzyme	Active site	Binding site	Experimental $\log IC_{50}$	Predicted $\log IC_{50}$	
					MLR	PCA
8a	–8.46	–8.56	–8.95	2.47	2.49	1.83
8b	–8.59	–8.98	–9.37	1.97	2.03	1.57
8e	–9.51	–9.05	–9.64	1.17	1.25	1.39
8f	–8.57	–9.55	–9.82	1.11	0.89	1.21
9a	–4.58	–9.15	–8.16	2.38	2.20	1.67
9b	–5.01	–9.40	–9.83	1.77	1.67	1.44
9c	–9.09	–9.73	–9.92	1.18	0.96	1.16
9d	–10.43	–9.97	–10.16	0.88	1.19	1.10
9e	–8.93	–9.91	–9.99	0.95	1.01	1.20
9f	–9.62	–10.01	–10.45	0.98	0.82	1.02
25a	–7.18	–9.49	–9.69	1.52	1.61	1.33
10a	–4.76	–10.06	–9.88	1.30	1.14	1.11
11a	–6.79	–9.92	–10.64	1.01	0.82	0.86
18a	–7.77	–10.21	–11.09	0.81	0.77	0.74
19a	–7.22	–11.31	–10.45	0.48	0.59	0.54
20a	–7.93	–10.92	–9.55	0.40	0.00	0.38
20b	–9.49	–11.41	–13.11	0.08	–0.06	0.033
20d	–9.10	–9.62	–13.38	0.00	0.16	0.02
20f	–7.68	–8.91	–13.45	0.13	–0.15	0.00
20g	–9.93	–10.44	–13.33	0.07	0.11	–0.16
20h	–9.40	–10.16	–14.09	–0.08	–0.01	0.35
20i	–9.43	–10.21	–13.23	0.32	0.14	0.36
20j	–7.53	–11.18	–8.73	–0.07	0.00	0.40
20k	–7.78	–9.68	–13.33	0.25	–0.01	0.34
20l	–9.87	–10.65	–13.25	0.03	0.31	0.16
20m	–10.32	–11.45	–13.93	0.51	0.31	0.16
20n	–6.95	–10.64	–13.4	0.04	–0.10	0.19
20o	–8.81	–11.29	–13.33	–0.04	0.08	0.18
20p	–8.80	–10.57	–12.99	–0.08	–0.07	0.05
20q	–8.55	–10.50	–12.88	–0.32	–0.28	–0.18
20r	–8.15	–10.12	–14.15	–0.28	–0.23	–0.19
20s	–7.79	–10.12	–13.85	–0.12	0.08	0.17
20t	–7.57	–6.12	–13.38	–0.01	0.10	–0.04
20u	–10.47	–12.21	–13.46	0.12	–0.09	–0.27
20v	–9.28	–9.25	–13.49	0.56	0.38	–0.05
20w	–5.78	–5.80	–12.72	0.12	0.08	–0.11
20x	–6.45	–9.08	–13.66	–0.19	–0.05	1.47

$$(\log IC_{50pred} = 18.341 - (0.679 \times IC5) - (7.862 \times MATS7e) - (9.117 \times GATS8m) - (2.371 \times GATS7e) - (4.763 \times MWC09) - (0.070 \times Mor04u) - (3.380 \times BEHm4) \quad (2)$$

R=0.975, F=80.259, sig = 0.0001,

Where IC5 is belong to topological class, MATS7e, GATS8m and GATS7e to electronegativity, (MWC09) to size, Mor04u to 3D-MORSE and BEHm4 to BCUT descriptors. Now we define in detail some of these descriptors such as; Broto-Moreau Autocorrelation Descriptors (labeled as ATS), Moran Autocorrelation Descriptors (labeled as MATS) Geary Autocorrelation Descriptors (labeled as GATS).

2D Autocorrelation Indices are defined as

$$ATSdw = \sum_{i=1}^{n}\sum_{j=1}^{n}\delta_{ij}(w_i w_j) \quad (3)$$

Where w_i and w_j are the weights of the atoms i and j, w∈(*m, p, e, v*). The symbol for each of the autocorrelation descriptors is followed by two indices *d* and *w*, where *d* and *w* stands for the lag and weight, respectively. Thus, for example, ATS4m means the Broto-Moreau Autocorrelation descriptor of lag 4 weighted by mass.The lag is defined as the topological distance *d* between pairs of atoms. The topological distance between a pair of atoms (*i, j*) is given in the *ij*th entry in the Topological Level Matrix. The lags can have any value from the set of {0, 1, 2, 3, 4, 5, 6, 7, 8}. The weights can be *m* (relative atomic mass), *p* (polarizability), *e* (Sanderson electronegativity) and *v* (Van der Waals volume). Relative mass is defined as the ratio of atomic mass of an atom to that of carbon. Similarly, the other three weights *p*, *e* and *v* are scaled by the corresponding values for carbon.

MATS7e: Moran Autocorrelation Descriptors are defined as:

$$MATSdW=(n)(A)/(B), \quad (4)$$

Where *n* is the number of atoms, *A* and *B* are

$$A = (1/\Delta)\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\delta_{ij}(w_i-\bar{w})(w_j-\bar{w})\right)$$

and

$$B = \sum_{i=1}^{n}(w_i-\bar{w})^2$$

w_i and w_j are the weights of the atoms *i* and *j*, w∈(*m, p, e,* v), w is the mean of w_i over the entire molecule, and δ_{ij} is Kronecker delta, that is, $\delta_{ij} = 1$ if the ijth entry in the Topological Level Matrix is = d, and $\delta_{ij} = 0$ otherwise.[22]

GATS8m: GATS7e Geary Autocorrelation Descriptors are defindes as

$$GATAdw = ((n-1)/2)\,(A)/(B)$$

$$A = (1/\Delta)\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\delta_{ij}(w_i-\bar{w})^2\right) \quad (5)$$

and

$$B = \sum_{i=1}^{n}(w_i-\bar{w})^2$$

Other variables are defined same as previous equation [18].

MWC09: Self-returning walk counts are defined as:

$$awc_i^{(k)} = \sum_{j=1}^{n} a_{ij}^{(k)} \quad (6)$$

A walk starting and ending on the same vertex, i.e. closed in itself and called a self-returning walk. In particular, the diagonal elements (*i*, *j*) in the k^{th} power matrix $\mathbf{A}^k$ denote the number of self-returning walks from the i^{th} atom to itself.

Mor04u:MoRSE descriptors

3D MoRSE descriptors (3D Molecule Representation of Structures based on Electron diffraction) are derived from infrared spectra simulation using a generalized scattering function.[23]

$$Mor(s,w) = I(s,w) = \sum_{i=2}^{n}\sum_{j=1}^{n} w_i w_j \sin(sr_{ij})/(sr_{ij}) \quad (7)$$

where, r_{ij} is the Euclidean distance between the atoms *i* and *j*, and w_i and w_j are the weights of the atoms *i* and *j* respectively.[24]

Figure 8A shows correlation between predicted $LogIC_{50}$ by MLR and experimental $LogIC_{50}$ for 37 N-(phenoxydecyl)phthalimide derivatives with a regression coefficient of 0.975.

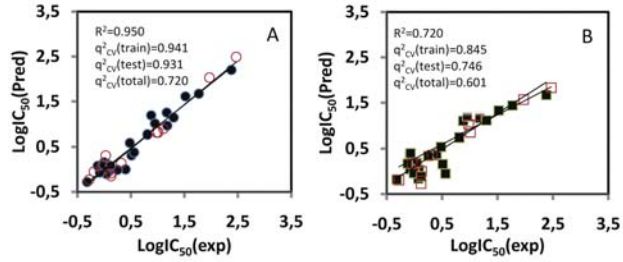


Figure 8: Correlation between predicted $LogIC_{50}$ by (A) MLR (B) PCA and experimental $LogIC_{50}$. Training and test sets were shown by filled and open symbols, respectively.

As we see in the equation (2) we can predict the $logIC_{50}$ for different compounds that we know their descriptor values. On the other hand, these descriptors may be not familiar for all chemists and biologists. So we selected some of more common descriptors that can be discussed the biological activity in detail. Table 6 has listed the 99 familiar descriptors among 1497 descriptors as well as their definition and classification. There are different properties such as size, charge, polarity, aromaticity, geometry or shape and hydrophobicity. Bivariate correlations were obtained between each descriptors and $logIC_{50}$ and correla-

tion coefficients (*R*) values were listed in the last column of Table 6. Positive and negative values of R indicate direct and inverse correlation between $logIC_{50}$ and descriptors. Namely positive values show that $logIC_{50}$ increases with increasing the cited descriptors and negative values show decreases with increasing the descriptor. As table shows, the $logIC_{50}$ decreases with size, polarity, number of functional group, geometry and hydrophobicty and increases with aromaticty.

Pascale, et al[20] reported that the efficacy of the inhibition activity depends on the chain length of the substrate and compound 20a possessing 10 carbons afforded the highest levels of activity. Structure-activity relationship studies indicated a critical role of electron-withdrawing substituents at the phenoxy group for the activity. In addition derivatives having a chlorine atom along with a strong electron-withdrawing group, such as a nitro group, were the most potent of the series.

Thus, they conclude that inhibition activity increases with number of CH_2 group and Cl atoms, and electronegativity of NO_2. On the other hand our findings also confirm their results so that the electronegativity in table 6 (descriptor 4 "sum of atomic Sanderson electronegativities scaled on Carbon atom)" has negative correlation. Also increasing number of nitro group (descriptor 90) causes decreasing the $logIC_{50}$ and increasing biological activity. Number of chlorine (descriptor 24) or halogen atom (25) also has negative correlation with $logIC_{50}$.

Table 6. Symbol, definition and classification of 99 used descriptors for PCA analysis

NO	Symbol	Definition	Class	R
1	MW	molecular weight	constitutional descriptors	–0.598
2	AMW	average molecular weight	constitutional descriptors	0.370
3	Sv	sum of atomic van der Waals volumes (scaled on Carbon atom)	constitutional descriptors	–0.835
4	Se	sum of atomic Sanderson electronegativities (scaled on Carbon atom	constitutional descriptors	–0.823
5	Sp	sum of atomic polarizabilities (scaled on Carbon atom)	constitutional descriptors	–0.834
6	Ss	sum of Kier-Hall electrotopological states	constitutional descriptors	–0.705
7	Mv	mean atomic van der Waals volume (scaled on Carbon atom)	constitutional descriptors	0.638
8	Me	mean atomic Sanderson electronegativity (scaled on Carbon atom)	constitutional descriptors	0.083
9	Mp	mean atomic polarizability (scaled on Carbon atom)	constitutional descriptors	0.594
10	Ms	mean electrotopological state	constitutional descriptors	–0.101
11	nAT	number of atoms	constitutional descriptors	–0.817
12	nSK	number of non-H atoms	constitutional descriptors	–0.827
13	nBT	number of bonds	constitutional descriptors	–0.817
14	nBO	number of non-H bonds	constitutional descriptors	–0.827
15	nBM	number of multiple bonds	constitutional descriptors	–0.489
16	SCBO	sum of conventional bond orders (H-depleted)	constitutional descriptors	–0.814
17	RBN	number of rotatable bonds	constitutional descriptors	–0.818
18	RBF	rotatable bond fraction	constitutional descriptors	–0.839
19	nDB	number of double bonds	constitutional descriptors	–0.489
20	nH	number of Hydrogen atoms	constitutional descriptors	–0.765
21	nC	number of Carbon atoms	constitutional descriptors	–0.810
22	nN	number of Nitrogen atoms	constitutional descriptors	–0.489
23	nO	number of Oxygen atoms	constitutional descriptors	–0.489
24	nCL	number of Chlorine atoms	constitutional descriptors	–0.103
25	nX	number of halogen atoms	constitutional descriptors	–0.281
26	qpmax	maximum positive charge	charge descriptors	–0.527
27	qnmax	maximum negative charge	charge descriptors	–0.326
28	Qpos	total positive charge	charge descriptors	–0.766
29	Qneg	total negative charge	charge descriptors	–0.765
30	Qtot	total absolute charge (electronic charge index - ECI)	charge descriptors	-0.765
31	Qmean	mean absolute charge (charge polarization)	charge descriptors	0.172
32	Q2	total squared charge	charge descriptors	-0.640
33	RPCG	relative positive charge	charge descriptors	0.399
34	RNCG	relative negative charge	charge descriptors	0.792
35	SPP	subpolarity parameter	charge descriptors	-0.524
36	TE1	topographic electronic descriptor	charge descriptors	-0.783
37	TE2	topographic electronic descriptor (bond resctricted)	charge descriptors	-0.769
38	PCWTe	partial charge weighted topological electronic charge	charge descriptors	-0.776
39	LDip	local dipole index	charge descriptors	-0.067
40	HOMA	Harmonic Oscillator Model of Aromaticity index	aromaticity indices	0.368
41	RCI	Jug RC index	aromaticity indices	0.337

NO	Symbol	Definition	Class	R
42	AROM	aromaticity (trial)	aromaticity indices	0.181
43	HOMT	HOMA total (trial)	aromaticity indices	0.362
44	J3D	3D-Balaban index	geometrical descriptors	-0.116
45	H3D	3D-Harary index	geometrical descriptors	-0.805
46	AGDD	average geometric distance degree	geometrical descriptors	-0.811
47	DDI	D/D index	geometrical descriptors	-0.811
48	ADDD	average distance/distance degree	geometrical descriptors	-0.816
49	G1	gravitational index G1	geometrical descriptors	-0.782
50	G2	gravitational index G2	geometrical descriptors	-0.812
51	RGyr	radius of gyration (mass weighted)	geometrical descriptors	-0.817
52	SPAN	span R	geometrical descriptors	-0.776
53	SPAM	average span R	geometrical descriptors	0.052
54	MEcc	molecular eccentricity	geometrical descriptors	-0.169
55	SPH	Spherosity	geometrical descriptors	-0.344
56	ASP	Asphericity	geometrical descriptors	-0.552
57	FDI	folding degree index	geometrical descriptors	-0.283
58	PJI3	3D Petijean shape index	geometrical descriptors	-0.272
59	L/Bw	length-to-breadth ratio by WHIM	geometrical descriptors	-0.499
60	SEig	absolute eigenvalue sum on geometry matrix	geometrical descriptors	-0.821
61	DISPm	d COMMA2 value / weighted by atomic masses	geometrical descriptors	-0.482
62	QXXm	Qxx COMMA2 value / weighted by atomic masses	geometrical descriptors	-0.483
63	QYYm	Qyy COMMA2 value / weighted by atomic masses	geometrical descriptors	-0.307
64	QZZm	Qzz COMMA2 value / weighted by atomic masses	geometrical descriptors	-0.816
65	DISPv	COMMA2 value / weighted by atomic van der Waals volumes	geometrical descriptors	-0.286
66	QXXv	Qxx COMMA2 value / weighted by atomic van der Waals volumes	geometrical descriptors	-0.538
67	QYYv	Qyy COMMA2 value / weighted by atomic van der Waals volumes	geometrical descriptors	-0.550
68	QZZv	Qzz COMMA2 value / weighted by atomic van der Waals volumes	geometrical descriptors	-0.812
69	DISPe	d COMMA2 value / weighted by atomic Sanderson electronegativities	geometrical descriptors	-0.484
70	QXXe	Qxx COMMA2 value / weighted by atomic Sanderson electronegativities	geometrical descriptors	-0.590
71	QYYe	Qyy COMMA2 value / weighted by atomic Sanderson electronegativities	geometrical descriptors	-0.656
72	QZZe	Qzz COMMA2 value / weighted by atomic Sanderson electronegativities	geometrical descriptors	-0.807
73	DISPp	d COMMA2 value / weighted by atomic polarizabilities	geometrical descriptors	-0.425
74	QXXp	Qxx COMMA2 value / weighted by atomic polarizabilities	geometrical descriptors	-0.539
75	QYYp	Qyy COMMA2 value / weighted by atomic polarizabilities	geometrical descriptors	-0.545
76	QZZp	QCOMMA2 value / weighted by atomic polarizabilities	geometrical descriptors	-0.812
77	G(N..N)	sum of geometrical distances between N..N	geometrical descriptors	-0.469
78	G(N..O)	sum of geometrical distances between N..O	geometrical descriptors	-0.519
79	G(N..F)	sum of geometrical distances between N..F	geometrical descriptors	-0.229
80	G(N..Cl)	sum of geometrical distances between N..Cl	geometrical descriptors	-0.333
81	G(O..O)	sum of geometrical distances between O..O	geometrical descriptors	-0.551
82	G(O..F)	sum of geometrical distances between O..F	geometrical descriptors	-0.229
83	G(O..Cl)	sum of geometrical distances between O..Cl	geometrical descriptors	-0.304
84	G(Cl..Cl)	sum of geometrical distances between Cl..Cl	geometrical descriptors	-0.211
85	nCp	number of total primary C(sp3)	functional groups	-0.099
86	nCs	number of total secondary C(sp3)	functional groups	-0.799
87	nCq	number of total quaternary C(sp3)	functional groups	-0.230
88	nCaH	number of unsubstituted aromatic C(sp2)	functional groups	0.429
89	nCaR	number of substituted aromatic C(sp2)	functional groups	-0.429
90	nNO2Ph	number of nitro groups (aromatic)	functional groups	-0.489
91	nRCX3	number of RCX3	functional groups	-0.230
92	nPhX	number of X-C on aromatic ring	functional groups	-0.103
93	nHAcc	number of acceptor atoms for H-bonds (N O F)	functional groups	-0.505
94	Ui	unsaturation index	empirical descriptors	-0.491
95	Hy	hydrophilic factor	empirical descriptors	-0.112
96	ARR	aromatic ratio	empirical descriptors	0.851
97	MR	Ghose-Crippen molar refractivity	Properties	-0.846
98	PSA	fragment-based polar surface area	Properties	-0.489
99	MLOGP	Moriguchi octanol-water partition coeff. (logP)	Properties	-0.817

Number of CH_2 in reference 20 or number of total secondary C(sp3) (descriptor 86) in this study has good and negative correlation with $logIC_{50}$. So our finding not only is similar to reference 20 but also, it gives a quantitative rather than qualitative interpretation for all results. Also number and variety of descriptors in our study is more than cited reference.

Data reduction by PCA method was done on these 99 descriptors to reduce and classify the descriptors. Table 7 shows the actual factors that were extracted. Phrase "Rotation Sums of Squared Loadings," shows only those factors that met cut-off criterion (eigenvalues greater than 1). SPSS software always extracts as many factors initially as there are variables in the dataset, but the rest of these didn't make the grade. The "% of variance" column tells us how much of the total variability (in all of the variables together) can be accounted for by each of these summary scales or factors. Factor 1 accounts for 42.093% of the variability in all variables, factor 2 accounts for 23.882 and so on.

Finally, the Rotated Component Matrix shows the factor loadings for each variable (Table 8). We went across each row, selected the factor that each variable loaded most strongly on higher than absolute 0.500 and removed those less than absolute 0.500. Thus missing values in Table 8 related to descriptors having loading number less than absolute 0.5. The first, second, third and fourth columns loaded strongly on factors 1 to 4, which we called size, polarity, geometry and number of halogen, respectively.

Then, by performing MLR on the resulted PCs (Principal component regression, PCR) we can obtain an equation between $logIC_{50}$ and PCs.

$$\log IC_{50pred} = 0.579 - (0.562 \times \text{size}) - (0.179 \times \text{polarity}) - (0.136 \times \text{geometry}) - (0.191 \times \text{number of halogen}) \quad (8)$$

R=0.849, F=20.708, sig=0.0001

The predicted values of $logIC_{50}$ by PCA were listed in Table 5. Figure 8B shows correlation between predicted and experimental $logIC_{50}$ by PCA analyses.

The correlation coefficient (R) for prediction of inhibitory activation was 0.849 by PCA model.

Table 7. Total variance explained by four principal components (PCs)

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of SquaredLoadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
PC1	51.780	52.303	52.303	51.780	52.303	52.303	41.673	42.093	42.093
PC2	14.867	15.017	67.320	14.867	15.017	67.320	23.643	23.882	65.975
PC3	9.114	9.206	76.526	9.114	9.206	76.526	8.790	8.878	74.854
PC4	6.851	6.920	83.446	6.851	6.920	83.446	8.507	8.593	83.446

Table 8. Rotated component matrixof reduced selected descriptors into 4 factors

Symbol	PC1 (Size)	PC2 (polarity)	PC3 (geometry)	PC4 (number of halogen)
MW	0.792			
AMW	–0.75			0.575
Sv	0.955			
Se	0.954			
Sp	0.962			
Ss	0.603	0.737		
Mv	–0.909			
Me		0.706		
Mp	–0.846			
Ms		0.893		
nAT	0.967			
nSK	0.857			
nBT	0.967			
nBO	0.857			
nBM		0.876		
SCBO	0.828	0.519		
RBN	0.958			
RBF	0.953			
nDB		0.876		
nH	0.988			
nC	0.98			
nN		0.876		
nO		0.876		
nCL				0.926
nX				0.762
Qpmax		0.865		
Qnmax		0.765		
Qpos	0.842	0.529		
Qneg	0.841	0.53		
Qtot	0.841	0.529		
Qmean		0.866		
Q2	0.588	0.801		

Symbol	PC1 (Size)	PC2 (polarity)	PC3 (geometry)	PC4 (number of halogen)
RPCG	–0.712	0.527		
RNCG	–0.924			
SPP		0.874		
TE1	0.857			
TE2	0.838	0.523		
PCWTe	0.871			
LDip		0.945		
HOMA				
RCI				
AROM				
HOMT				
J3D			0.928	
H3D	0.962			
AGDD	0.971			
DDI	0.967			
ADDD	0.976			
G1	0.619	0.58		
G2	0.73	0.554		
RGyr	0.957			
SPAN	0.969			
SPAM			–0.890	
MEcc	0.622			
SPH	0.704			
ASP	0.837			
FDI	0.669		–0.612	
PJI3	0.583		–0.532	
L/Bw	0.753			
SEig	0.968			
DISPm				0.649
QXXm			0.756	
QYYm		0.619		
QZZm	0.888			
DISPv				
QXXv			0.898	

Symbol	PC1 (Size)	PC2 (polarity)	PC3 (geometry)	PC4 (number of halogen)
QXXv			0.898	
QYYv				
QZZv	0.96			
DISPe		0.751		
QXXe			0.828	
QYYe	0.609	0.507		
QZZe	0.951			
DISPp				0.521
QXXp			0.911	
QYYp				
QZZp	0.965			
G(N..N)		0.867		
G(N..O)		0.838		
G(N..F)				
G(N..Cl)				0.889
G(O..O)		0.814		
G(O..F)				
G(O..Cl)				0.905
G(Cl..Cl)				0.802
nCp			0.653	
nCs	0.97			
nCq				
nCaH			–0.581	–0.539
nCaR			0.581	0.539
nNO2Ph		0.876		
nRCX3				
nPhX				0.926
nHAcc		0.919		
Ui		0.876		
Hy		0.844		
ARR	–0.89			
MR	0.922			
PSA		0.876		
MLOGP	0.907			

In order to investigate of model validity we performed a cross validation. We took some of compounds as test set (**8a, 8b, 8f, 9f, 11a, 20f, 20i, 20*l*, 20o, 20r, 20u** and **20x** in Table 5) and the reminder as training set. Then using leave-one-out methods we obtained q^2_{cv} based on equation (1). These values were inserted in the Fig. 8. As we see the q^2_{cv} values are higher than 0.3 and lower than R^2.

4. Conclusions

The purposes of this study were to survey the inhibition effect of N-(phenoxydecyl) phthalimide derivatives on α-glucosidase enzyme, and finding the key features of responsible α-glucosidase inhibitory activity. It was done by a computer drug-design protocol involving homology modeling for target protein, docking simulation and Quantitative Structure Activity Relationship. Firstly the homology modeled structure of *S. cerevisiae* α-glucosidase was built and used for molecular docking to define the interaction mode of the N-(phenoxydecyl) phthalimide derivatives with the protein. The results showed the important role of carboxyl group in binding of ligands with the active site of protein by formation of hydrogen bonds. Results of MLR on total and selected descriptors, showed a realistic correlation between predicted and experimental values. Experimental data showed that the inhibition activity increases with chain length, number of chlorine and electronegativity, while our results not only confirm the experimental data but also introduced new variables such as aromaticity, hydrophobicity, polarity to investigate the data in detail and quantitative manner. The QSAR studies showed that the inhibition increases ($\log IC_{50}$ decreases) with increasing size, polarity, geometry and number of halogen factors and decreasing of aromaticity. Consecutively, these results revealed that N-(phenoxydecyl) phthalimide derivatives bind to site near the active site residues and free energy of binding has relatively good correlation with $\log IC_{50}$.

5. Acknowledgement

Financial support of Damghan University is acknowledged.

6. References

1. S. J. Heo, J. Y. Hwang, J. I. Choi, J. S. Han, H. J. Kim, Y. J. Jeon, *Eur. J. Pharmacol.* **2009,** *615*, 252–256.
2. Y. M. Kim, M. H. Wang, H. I. Rhee, *Carbohydr. Res.* **2004,** *339*, 715–717.
3. A. Kimura, J. H. Lee, I. S. Lee, H. S. Lee, K. H. Park, S. Chiba, D. Kim, *Carbohydr. Res.* **2004,** *339*, 1035–1040.
4. H. Park, K. Y. Hwang, Y. H. Kim, K. H. Oh, J. Y. Lee, K. Kim, *Bioorg. Med. Chem. Lett.* **2008,** *18*, 3711–3715.
5. H. Gao, J. Kawabata, *Bioorg. Med. Chem. Lett.* **2008,** *18*, 812–815.
6. J. Pandey, N. Dwivedi, N. Singh, A. K. Srivastava, A. Tamarkar, R. P. Tripathi, *Bioorg. Med. Chem. Lett.* **2007,** *17*, 1321–1325.
7. G. Tanabe, K. Yoshikai, T. Hatanaka, M. Yamamoto, Y. Shao, T. Minematsu, O. Muraoka, T. Wang, H. Matsuda, M. Yoshikawa, *Bioorg. Med. Chem.* **2007,** *15*, 3926–3937.
8. H. W. Xu, G. F. Dai, G. Z. Liu, J. F. Wang, H. M. Liu, *Bioorg. Med. Chem.* **2007,** *15*, 4247–4255.
9. S. S. Rajan, X. Yang, F. Collart, V. L. Y. Yip, S. G. Withers, A. Varrot, J. Thompson, G. J. Davies, W. F. Anderson, *Struct.* **2004,** *12*, 1619–1629.
10. H. Park, K. Y. Hwang, K. H. Oh, Y. H. Kim, J. Y. Lee, K. Kim, *Bioorg. Med. Chem.* **2008,** *16*, 284–292.
11. S. Sou, S. Mayumi, H. Takahashi, R. Yamasaki, S. Kadoya, M. Sodeoka, Y. Hashimoto, *Bioorg. Med. Chem. Lett.* **2000,** *10*, 1081–1084.
12. P. G. Baker, A. Brass, *Curr. Opin. Biotechnol.* **1998,** *9*, 54–58.
13. K. Yamamoto, H. Miyake, M. Kusunoki, S. Osaki, *FEBS J.* **2010,** *277*, 4205–4214.
14. A. Fiser, A. Sali, In *Methods in Enzymology*, Charles, W., Carter, J., Robert, M. S., Eds. Academic Press: Massachusetts, 2003; Vol. Volume 374, pp 461–491.
15. R. A. Laskowski, M. W. MacArthur, D. S. Moss, J. M. Thornton, *J. Appl. Crystallogr.* **1993,** *26*, 283–291.
16. B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, *J. Chem. Theory Comput.* **2008,** *4*, 435–447.
17. G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, *J. Comput. Chem.* **1998,** *19*, 1639–1662.
18. A. Golbraikh, A. Tropsha, *J. Mol. Graphics Modell.* **2002,** *20*, 269–276.
19. P. V. Khadikar, V. Sharma, S. Karmarkar, C. T. Supuran, *Bioorg. Med. Chem. Lett* **2005,** *15*, 923–930.
20. R. Pascale, A. Carocci, A. Catalano, G. Lentini, A. Spagnoletta, M. M. Cavalluzzi, F. De Santis, A. De Palma, V. Scalera, C. Franchini, *Bioorg. Med. Chem.* **2010,** *18*, 5903–5914.
21. K. Bharatham, N. Bharatham, K. H. Park, K. W. Lee, *J. Mol. Graphics Modell.* **2008,** *26*, 1202–1212.
22. P. A. P. Moran, *Biometrika* **1950,** *37*, 17–23.
23. L. J. Soltzberg, C. L. Wilkins, *J. Am. Chem. Soc.* **1977,** *99*.
24. I. Tetko, J. Gasteiger, R. Todeschini, A. Mauri, D. Livingstone, P. Ertl, V. Palyulin, E. Radchenko, N. Zefirov, A. Makarenko, V. Tanchuk, V. Prokopenko, *J. Comput.-Aided Mol. Des.* **2005,** *19*, 453–463.

Povzetek

Število bolnikov z diabetesom se povečuje, zato je razpozavanje dejavnikov, ki uravnavajo bolezen, zelo pomembno. α-Glukozidaza (EC 3.2.1.20) je esencialni encim, ki sodeluje pri presnovi ogljikovih hidratov, kot je npr. škrob. Ogljikovi hidrati se običajno presnovijo do enostavnih sladkorjev, ki se lahko absorbirajo v tankem črevesu. Zato lahko α-glukozidazne inhibitorje uporabljamo za znižanje ravni krvnega sladkorja. Z računalniškim programom za načrtovanje zdravil, ki vključuje homologno modeliranje, sidranje in QSAR analizo, smo preiskovali učinke inhibicije derivatov N-(fenoksidecil) ftalimida. Homologni model α-glukozidaze kaže na strukturo, ki je zelo podbna kristalni strukturi oligo-1,6-glukozidaze iz kvasovke *Saccharomyces cerevisiae*. Rezultati sidranja so pokazali, da je položaj vezavnega mesta na inhibitorju blizu aktivnega mesta in da je karbonilni kisik na ftalimidu učinkovita funkcionalna skupina za vezavo inhibitorja na protein. Enačba, pridobjena s QSAR analizo, pa je pokazala, da se inhibicijske lastnosti derivatov N-(fenoksidecil) ftalimida zmanjšujejo, če se povečuje velikost, polarnost, geometrija in število halogenskih faktorjev.