Journal of Pharmaceutical Research International

31(5): 1-24, 2019; Article no.JPRI.52361
ISSN: 2456-9119
(Past name: British Journal of Pharmaceutical Research, Past ISSN: 2231-2919, NLM ID: 101631759)

QSAR, Docking Study of Isatin Analogues as Anti-bacterial Agents for Design New Compounds

Razieh Sabet[1*] and Isar Jafroudi[1]

[1]Department of Medicinal Chemistry, Faculty of Pharmacy, Shiraz University of Medical Sciences, Shiraz, Iran.

Authors' contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JPRI/2019/v31i530314
Editor(s):
(1) Dr. Jongwha Chang, University of Texas, College of Pharmacy, USA.
Reviewers:
(1) Vesna Rastija, Josip Juraj Strossmayer University of Osijek, Croatia.
(2) Toluwase Hezekiah Fatoki, Federal University, Oye Ekiti, Nigeria.
(3) Karen Cordovil, Oswaldo Cruz Foundation, Brazil.
Complete Peer review History: https://sdiarticle4.com/review-history/52361

Original Research Article

Received 03 September 2019
Accepted 07 November 2019
Published 16 November 2019

ABSTRACT

Computational chemistry is unique method in drug discovery which reduce cost. In this study 86 molecules containing isatin core were subjected to quantitative structure-activity relationship analysis and docking study to find the structure requirements for ligand binding. The structures were sketched and optimized in Hyperchem. The structural invariants used in this study were those obtained from whole molecular structures: by both hyperchem and dragon software (16 types of descriptors). Four chemometrics methods including MLR, FA-MLR, PCR and GA-PLS were employed to make connection between structural parameters and anticancer effects. MLR analysis explained the positive effect of the number of urea derivatives, thio urea, amide, thioamide, hydrazone, thiocarbohydrazone, nBnz with the halogen substitution on 5 position of isatin ring on the antimicrobial activity. It also shows nArCN, nPyridines have negative effects on the antimicrobial activity of studied compound.
The FA-MLR describes the effect of 3D-MORSE and Galvez Topological charge descriptors and on antimicrobial activity of the studied compounds. The quality of PCRA equation is better than those derived from FA-MLR. A comparison between the different statistical methods employed revealed

**Corresponding author: E-mail: Sabet_r@sums.ac.ir;*

that GA-PLS represented superior results and it could explain and predict 73% and 68% of variances in the –LogMIC data, respectively. Comparison between QSAR and docking analysis revealed that by decreasing in number of ring and lipophilicity (also Logp) for design of new compounds can have better activity. Substitutions such as urea, thiourea, thiocarbohydrazone, benzhydrazide as on isatin ring, can cause better interaction with receptor.

Keywords: Isatin; QSAR; docking; antibacterial.

1. INTRODUCTION

FabI was confirmed to be the only enoyl-acyl carrier protein (ACP) reductase required for the synthesis of fatty acids. FabI is required for the elongation of both long-chain saturated and unsaturated fatty acids in *Escherichia coli*. Thus, the activity of this enzyme plays a determinant role in completing cycles of fatty acid biosynthesis in *E. coli*. Indole compounds can inhibit FabI [1]. One of the best promising new indole compounds having many interesting activity profiles are isatin and isatin derivatives. The isatin (1*H*-indole-2,3-dione) moiety is responsible for a wide spectrum of biological property such as antibacterial, antifungal, antiviral, anticancer, anticonvulsant, anti-HIV and antiparkinsonian activity in many synthetically versatile molecules [2–9]. Among these properties antibacterial activities (against *Escherichia coli*) of this moiety was of our interest to study the quantitative structure-activity relationships of a series of 86 isatin derivatives reported in literature.

Synthesis and evaluation of biological activity of novel compounds usually time consuming and take large amounts of money. Today, the use of computational methods for designing newly biologically active compounds have opened a new window to modern drug discovery study. Computational methods can accelerate the procedure of discovering new drugs by designing new compounds and predicting potency or activity of them. Quantitative structure activity relationship (QSAR) studies provide pharmaceutical chemists valuable information that is useful for drug design and prediction of drug activity [10-14]. QSAR studies, as one of the most important areas in chemometrics, give information that is useful for molecular design and medicinal chemistry [8-12]. QSAR models are mathematical equations constructing a relationship between chemical scaffold and biological property. These models have another ability, which is providing a deeper knowledge about molecule design.

Linear and nonlinear QSAR models are mathematical equations that display us enough information about the mechanism of biological activity of compounds by constructing a relationship between chemical structures and biological activities. The first step in constructing QSAR models is the proper representation of the structural and physicochemical features of chemical compounds [15-18]. These features named molecular descriptors that represent variation in the structural features property of the molecules by numerical and have high effect on the biological property of compound [19-22]. Molecular descriptors have been classified into different categories such as physiochemical, constitutional, geometrical, topological, and quantum chemical descriptors. Dragon and hyperchem are two famous computational softwares provide us more than 4000 of these descriptors [23,24].

Different QSAR methods including multiple linear regression (MLR), partial least squares combined with genetic algorithm for variable selection (GA-PLS), factor analysis–MLR (FA-MLR), principal component regression analysis (PCR) were used to make connections between structural descriptors and antibacterial activity of studied compounds [25-28]. An important approach of the researchers in modification of the isatin moiety has been to establish a comprehensive structure–activity relationship (SAR), for this class of antibacterial agents (against *Escherichia coli)*.

(Our research show that our studied series of compounds didn't evaluate for QSAR studies). Our different QSAR analysis establishes mathematical relationship between biological activities and computable parameters such as chemical, topological, physicochemical, stereochemical or geometrical and so on indices.

The molecular docking studies help us to understand the different interactions between the ligands and enzyme active sites (FabI) in detail and also help to design novel potent structure.

Molecular docking simulation technique was also performed on eighty six compounds to show the details molecular binding models for these compounds interacting with the key active site of protein.

2. METHODS

2.1 Data Set

The biological data used in this study were antimicrobial activity against Ecoli, (in terms of –log MIC), of a set of 86 isatin derivatives [29-35]. Firstly, the outlier data was removed from the data set according to principle component analysis, it was shown in Fig. 1. The data set was classified into calibration and prediction set by kenardston algorithm (26) of the 16 prediction molecules from the spaces of the calculated descriptors. The structural features and biological activity of these compounds are listed in Table 1.

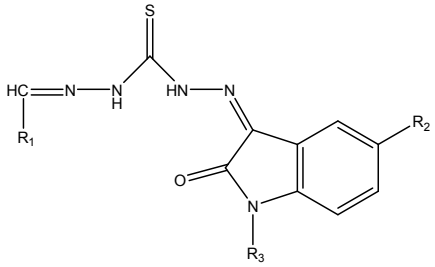
2.2 Descriptor Generation

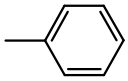
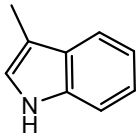
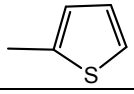
The structural features of the studied compounds are listed in Table 1. The two-dimensional structures of molecules were drawn by Hyperchem 8.0 software (Hypercube Inc.) to calculate whole molecular structure-based descriptors. The final geometries were obtained with semi-empirical AM1 calculations in Hyperchem program. The molecular structures were optimized using the Polak-Ribiere algorithm until the root mean square gradient was 0.01 $kcal^{-1}$ mol [23]. Some physicochemical parameters including molecular volume (V), molecular surface area (SA), hydrophobicity (Log P), hydration energy (HE) and molecular polarizability (MP) were calculated using Hyperchem Software. In order to calculate some molecular descriptors including topological, constitutional and functional group descriptors the optimized molecules were transferred into the Dragon package, developed by the Milano chemometrics and QSAR Group [24]. The calculated descriptors from whole molecular structures are briefly described in Table 2.

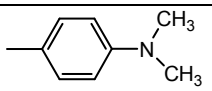
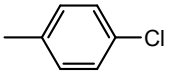
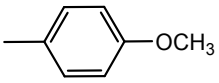
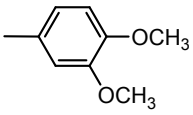
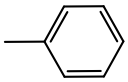
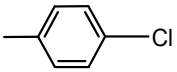
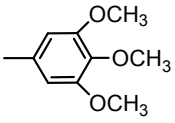
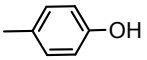
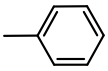
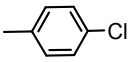
2.3 Data Screening & Model Building

The selected descriptors from each class and the experimental data were analyzed by the stepwise regression SPSS (version 22.0) software. The calculated descriptors were collected in a data matrix whose number of rows and columns were the number of molecules and descriptors, respectively. Multiple linear regressions (MLR) and partial least squares (PLS) were used to derive the QSAR equations and feature selection was performed by the use of genetic algorithm (GA). MLR with factor analysis as the data pre-processing step for variable selection (FA-MLR) and principal component regression analysis (PCRA) methods were also used to explore the QSAR equations.

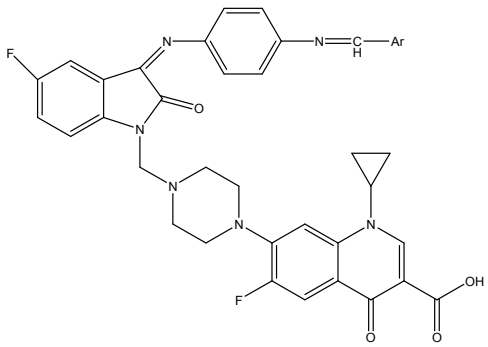
Table 1. Chemical structure of isatin derivatives used in this study

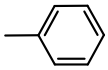
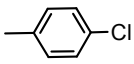
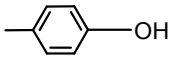
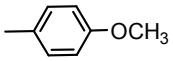
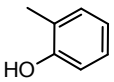


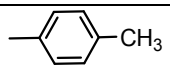
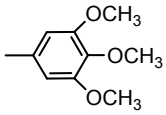
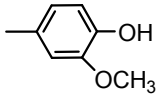
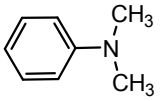
Compound	R_1	R_2	R_3	PMIC
1		H	H	4.00
2		H	H	4.60
3**		H	H	3.70

Compound	R_1	R_2	R_3	PMIC
4		H	H	4.00
5		H	H	4.60
6		H	H	3.70
7**		H	H	4.60
8		CH3	H	4.00
9		CH_3	H	4.30
*		CH_3	H	
*		H	CH3	
*		H	CH3	
*		H	CH3	

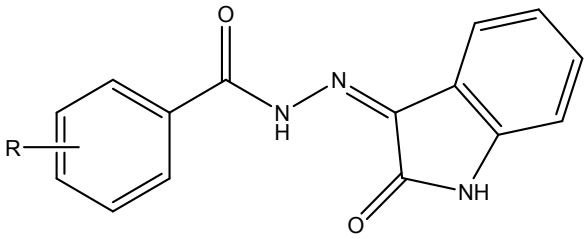
**outlier data; **test set*



Compound	Ar	PMIC
*		
*		
10**		4.81
*		
*		

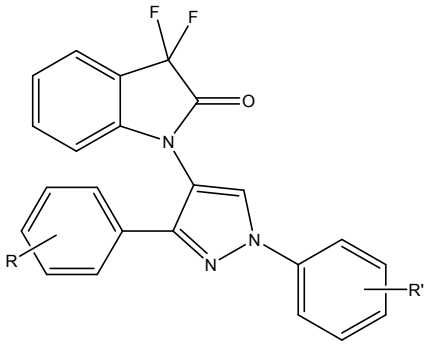
*		
11**		4.51
12		4.51
13		4.51

**outlier data; **test set*



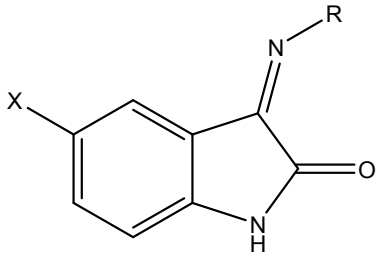
Compound	R	PMIC
14	H	5.18
15	2-F	6.74
16	3-Cl	5.95
17	3-CH3	5.31
18	2-Br	5.61
19**	3-F	6.17

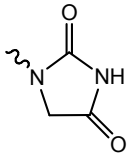
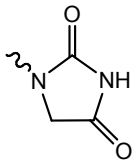
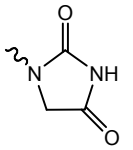
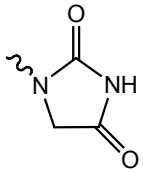
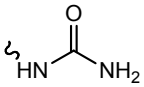
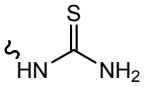
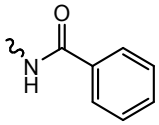
***test set*

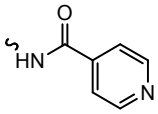
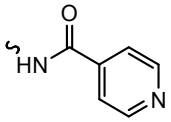


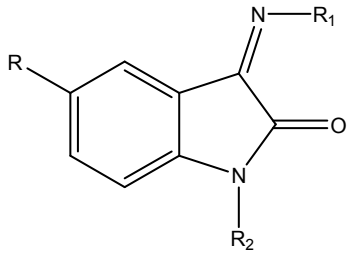
Compound	R	R’	PMIC
20	H	H	3.60
21	H	4-CH3	3.60
22	H	4-CN	3.70
23	4-OCH3	H	3.60
24	4-OCH3	4-CH3	3.70
25	4-OCH3	4-CN	3.82
26	4-Cl	H	3.60
27	4-Cl	4-CH3	3.70
28	4-Cl	4-CN	3.60
29	4-CH3	H	3.70

Compound	R	R'	PMIC
30	4-CH3	4-CH3	3.60
31	4-CH3	4-CN	3.70
32	4-Br	H	3.30
33	4-Br	4-CH3	3.60
34	4-Br	4-CN	3.70
35	4-F	H	3.82
36	4-F	4-CH3	3.82
37	4-F	4-CN	4.10
38	2-Cl,4-Cl	H	3.60
39**	2-Cl,4-Cl	4-CN	3.82



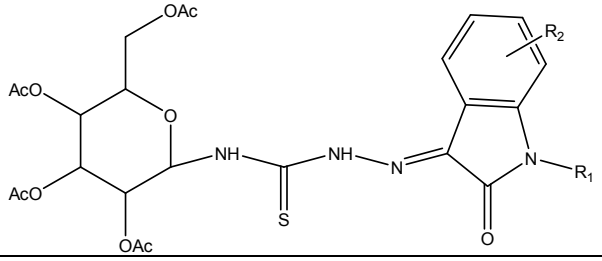
Compound	R	X	PMIC
40		H	4.90
41**		F	4.90
42		Cl	4.60
43		CH3	4.60
44		H	4.90
45**		F	4.30
46		F	4.90

Compound	R	X	PMIC
47		F	4.60
48		Cl	4.90



Compound	R	R1	R2	PMIC
49	H	1-naphthyl	_____	3.78
50	H	4-Chloro phenyl	_____	4.92
51	H	4-Bromo phenyl	_____	3.78
52	H	4-Methyl phenyl	_____	3.94
53	H	Phenyl hydrazino	_____	3.88
54	H	Thiosemicarbazino	_____	4.17
55	CH3	1-Naphthyl	_____	3.85
56	CH3	4-Chloro phenyl	_____	3.86
57	NO2	1-Naphthyl	_____	4.39
58	NO2	4-Bromo phenyl	_____	4.18
59	NO2	4-Chloro phenyl	_____	3.78
60	CH3	4-Methoxy phenyl	_____	3.58
61	CH3	Thiosenicarbazino	_____	4.40
62	Cl	4-chloro phenyl	_____	4.01
63	Cl	4-Methyl phenyl	H	4.85
64	Cl	Thiosemicarbazino	H	4.00
65	Br	1-Naphtyl	H	3.78
66	Br	4-Methoxy phenyl	H	3.79
67	H	4-Methyl phenyl	CH2-N(C6H5)2	3.89
68	Cl	4-Bromo phenyl	CH2-N(C6H5)2	3.79
*	Br	4-Methoxy phenyl	CH2-N(C6H5)2	

*outlier data; ** test set*



Compound	R1	R2	PMIC
*	H	H	
*	H	5'-NO2	
69**	H	5'-Cl,7'-NO2	4.60
70	H	5'-Br,7'-NO2	4.60

Compound	R1	R2	PMIC
71**	H	5',7'-diBr	4.30
72**	H	5'-F	4.30
73**	H	5'-Cl	4.30
74	H	5'-Br	4.30
75**	H	5'-I	4.30
76**	H	5'-Me	4.60
77	H	7'-Me	4.60
78	H	5'-iPr	4.30
79	Me	H	4.90
80	Et	H	4.90
81	n-Pr	H	4.90
82**	n-Bu	H	4.90
83**	i-Bu	H	4.90
84	Allyl	H	4.00
85	Bn	H	4.00
86	Phen	H	4.00

*outlier data/ **test set

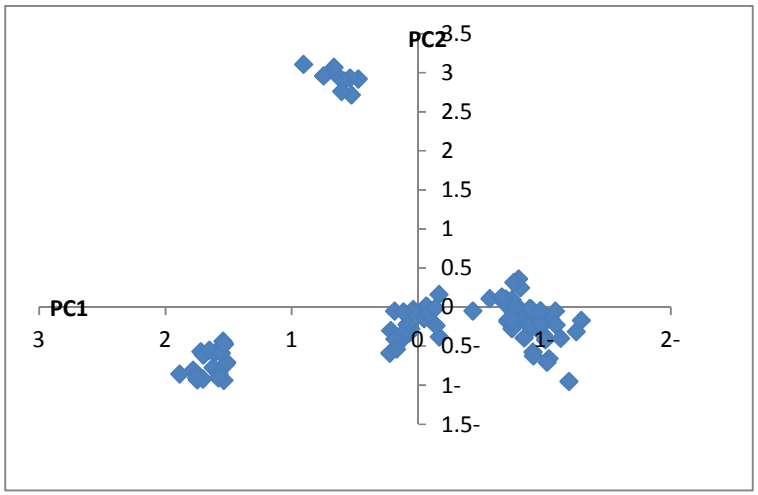


Fig. 1. Outlier data by principle component analysis before QSAR analysis

The resulted models were validated by leave-one out cross-validation procedure (using MATLAB software) to check their predictability and robustness.

A key step in QSAR modeling is evaluating model's stability and prediction ability. We used cross-validation and external test set for these proposes. Cross-validation has different variants such as leave-one-out (LOO), leave-group-out (LGO) and ν-fold. It was shown previously that LOO can leads to chance and overfitted models whereas LGO is more sensitive to chance variables [36]. Therefore, we used LGO for model-validation utilizing correlation coefficient and root mean square error of cross-validation (*q2* and *RMSECV*, respectively) as scoring function. In addition, an external test set composed of 6 molecules was also used. The molecules in this set did not have contribution in the model step and thus their predicted values can give a final prediction power of the models as measured by correlation coefficient, root mean square errors of prediction, relative error of prediction (R^2_P, $RMSE_P$ and *REP*, respectively).

The PLS regression method used in this study was the NIPALS-based algorithm existed in the chemometrics toolbox of MATLAB software (version 12 Math work Inc.). Leave-one-out cross-validation procedure was used to obtain the optimum number of factors based on the Haaland and Thomas F-ratio criterion [37].

2.4 Docking Procedures

An in house batch script (DOCK-FACE) for automatic running of AutoDock 4.2 was used to

Table 2. Brief description of some descriptors used in this study

Descriptor type	Molecular Description
Chemical	LogP (Octanol-water partition coefficient), Hydration Energy (HE), Polarizability (Pol), Molar refractivity (MR), Molecular volume (V), Molecular surface area (SA).
Constitutional	mean atomic vander Waals volume (MV), no. of atoms, no. of non-H atoms, no. of bonds, no. of heteroatoms, no. of multiple bonds (nBM), no. of aromatic bonds, no. of functional groups (hydroxyl, amine, aldehyde, carbonyl, nitro, nitroso, etc.), no. of rings, no. of circuits, no of H-bond donors, no of H-bond acceptors, no. of Nitrogen atoms (NN), chemical composition, sum of Kier-Hall electrotopological states (Ss), mean atomic polarizability (Mp), number of rotable bonds (RBN), mean atomic Sanderson electronegativity (Me), number of Chlorine atoms (NCl), number of 9-membered rings (NR09), etc.
Topological	Molecular size index, molecular connectivity indices (X1A, X4A, X2v, X1Av, X2Av, X3Av, X4Av), information content index (IC), Sum of topological distances between F..F (T(F..F)), Ratio of multiple path count to path counts (PCR), Mean information content vertex degree magnitude (IVDM), Eigenvalue sum of Z weighted distance matrix (SEigZ), reciprocal hyper-detour index (Rww), Eigenvalue coefficient sum from adjacency matrix (VEA1), radial centric information index, 2D petijean shape index (PJI2), mean information index on atomic composition(AAC), Kier symmetry index(S0K), mean information content on the distance degree equality (IDDE), structural information content (neighborhood symmetry of 3-order) (SIC3), Randic-type eigenvector-based index from adjacency matrix (VRA1), sum of topological distances between N..N (T(N..N)), sum of topological distances between O..O(T(O..O)),etc.
Geometrical	3D-Balaban index (J3D), span R (SPAN), length-to-breadth ratio by WHIM (L/BW), sum of geometrical distances between N..N (G(N..N)), sum of geometrical distances between N..O (G(N..O)), sum of geometrical distances between O..O (G(O..O)), ect.
Mol-Walk	molecular walk count of order 08 (MWC08), self-returning walk count of order 05 (SRW05), total walk count (TWC), etc.
Burden matrix	highest eigenvalue n. 1 of Burden matrix / weighted by atomic masses (BEHM1), highest eigenvalue n. 7 of Burden matrix / weighted by atomic masses (BEHM7), lowest eigenvalue n. 1 of Burden matrix / weighted by atomic masses (BELM1), highest eigenvalue n. 1 of Burden matrix / weighted by atomic van der Waals volumes (BELV1), highest eigenvalue n. 2 of Burden matrix / weighted by atomic Sanderson electronegativities (BEHE2), etc.
Galvez	topological charge index of order 1 (GGI1), topological charge index of order 6 (GGI6),topological charge index of order 7 (GGI7), global topological charge index (JGT), etc.
2D autocorrelation	Broto-Moreau autocorrelation of a topological structure - lag 7 / weighted by atomic Sanderson electronegativities (ATS7E), Moran autocorrelation -lag 4 / weighted by atomic Sanderson electronegativities (MATS4E), Broto-Moreau autocorrelation of a topological structure - lag 3 / weighted by atomic Sanderson electronegativities (ATS3E), Broto-Moreau autocorrelation of a topological structure - lag 3 / weighted by atomic van der Waals volumes (ATS3V), etc.
Charge	maximum positive charge (QPOS), partial charge weighted topological electronic charge (PCWTE), etc.
Aromaticity	HOMA Harmonic Oscillator Model of Aromaticity index,RCI;Jug RC index aromaticity indices,HOMT;HOMA total (trial) , etc.
Randic	DP0;molecular profile, SP0;shape profile; SHP;average shape profile index , etc.

Descriptor type	Molecular Description
RDF	Radial Distribution Function - 7.0 / unweighted(RDF070U),Radial Distribution Function - 13.5 / unweighted(RDF135U),Radial Distribution Function - 1.0 / weighted by atomic masses(RDF010M),Radial Distribution Function - 3.0 / weighted by atomic masses(RDF030M),Radial Distribution Function - 4.5 / weighted by atomic masses(RDF045M),Radial Distribution Function - 12.5 / weighted by atomic masses(RFD125M),Radial Distribution Function - 2.0 / weighted by atomic van der Waals volumes(RDF020V),Radial Distribution Function - 8.5 / weighted by atomic van der Waals volumes(RDF085V),Radial Distribution Function - 1.0 / weighted by atomic Sanderson electronegativities(RDF010E), etc.
3D-MoRSE	3D-MoRSE - signal 01 / unweighted (MOR01U)(01U,02U,…,32U), 3D-MoRSE - signal 01 / weighted by atomic van der Waals volumes (MOR01V)(01V,02V,…,32V), ect.
WHIM	1st component symmetry directional WHIM index / weighted by atomic polarizabilities (G1P), 2st component symmetry directional WHIM index / weighted by atomic electrotopological states (G2S), D total accessibility index / weighted by atomic van der Waals volumes (DV), etc.
GETAWAY	H autocorrelation of lag 1 / lag2/ lag3 weighted by atomic Sanderson electronegativities (H1E,H2E,H3E), total information content on the leverage equality (ITH), R maximal autocorrelation of lag 3 / lag4 unweighted (R3U+,R4U+), R maximal autocorrelation of lag 6 / weighted by atomic masses (R6M+), R maximal autocorrelation of lag 5 / weighted by atomic van der Waals volumes (R5V+), R maximal autocorrelation of lag 1 / lag 4 weighted by atomic Sanderson electronegativities (R1E+), R maximal autocorrelation of lag 3 / weighted by atomic polarizabilities (R3P+), etc.
Functional	number of total secondary C(sp3) (NCS), number of ring tertiary C(sp3) (NCRHR), number of secondary C(sp2) (n=CHR), number of tertiary amines (aliphatic) (NNR2), number of N hydrazines (aromatic) (nN-NPH), number of nitriles (aliphatic) (NCN), number of phenols (NOHPH), number of ethers (aromatic) (NRORPH), number of solfures (NRSR), etc.
Atom-Centred	CHR3 (C-003), CR4 (C-004), X--CR..X (C-034), Ar-C(=X)-R (C-039), R-C(=X)-X / R-C#X / X-=C=X (C-040), X--CH..X (C-042), H attached to C1(sp3) / C0(sp2) (H-047), RCO-N< / >N-X=X (N-072),R2S / RS-SR (S-107), etc.
connectivity indices	X0(connectivity index chi-0), connectivity index chi-1(x1), average connectivity index chi-0(XOA)
information indices	Uindex(Balaban U index), IC0(information content index), TIC0(total information content index)
edge adjacency indices	EEig01x(Eigenvalue 01),EEig01r(Eigenvalue 01 from edge)
eigenvalue-based indices	Eig1v(Leading eigenvalue from van der Waals weighted distance matrix), SEigm Eigenvalue sum from mass weighted distance matrix eigenvalue-based indices

carry out the docking simulations [38] in a parallel mode [39]. To prepare the receptor structure, the three dimensional crystal structure of isatin (PDB:1Lx6) was acquired from Protein Data Bank (PDB data base; http://www.rcsb.org) [40] and water molecules and co-crystal ligand were removed from the structure. The PDB were then checked for missing atom types with the python script as implemented in MODELLER 9.17 [41]. The ligand structures were made by Hyper Chem software package (Version 7, Hypercube Inc). For geometry optimization, Molecular Mechanic (MM^+), followed by semi empirical AM1 method was performed. The prepared Ligands were given to 100 independent genetic algorithm (GA) runs. 160 population size, a maximum number of 2,600,000 energy evaluations and 26,000 maximum generations were used for Lamarckian GA method. The grid points of 40, 40, and 40 in x-, y-, and z directions 2.299, 19.177 and 136.7 were usedAll visualization of protein ligand interaction was evaluated using VMD software [42]. Cluster analysis was performed on the docked results using a root mean square deviation (RMSD) tolerance of 1.76 Å.

3. RESULTS AND DISCUSSION

3.1 MLR Analysis

In the first step, separate stepwise selection-based MLR analyses were performed using different types of descriptors, and then, an MLR equation was obtained utilizing the pool of all calculated descriptors. The resulted QSAR models from different types of descriptors for the compounds (86 molecules as calibration and 16 molecules as prediction sets) are listed in Table 3.

The equation E1 of Table 3 shows among chemical descriptors, the negative effect of log P of the molecules on the antimicrobial activity. This equation shows the hydrophilic molecules shows better antimicrobila effect. The second equation of Table 3 demonstrated the effect of constitutional descriptors on the activity of these compounds. It also explain the negative effects nCIC (number of rings), and nR10 (number of 10-membered rings) on activity (such as molecule series 20-39, 49, 55, 65 have intermediate activity and inactive compounds).

The effect of topological group counts parameter on antimicrobial activity of the studied compounds has been described by equation E_3 of Table 3. It shows that among topological descriptors spanning tree number (STN) has the negative effect on cytotoxic activity of the compounds.

The equation E_4 of Table 3 was found by using Mol-Walk descriptors (E_4), which explains the negative effect of PIPC09 of studied compounds on the activity of the compounds. The equation E_5-E_{16} and E_{16} of Table 3 demonstrated the effect of posititive and negative effects of BCUT, Galvz topological Charge indices, 2D autocorrelations, Charge, Burden eigenvalues, RDF, 3D MoRSE, WHIM, GETAWAY and charge descriptors on the anti-cancer activity of these compounds.

The MLR equation of Table 3 obtained from the pool of functional groups descriptors, E_{17}, explained the positive effect of the number of urea derivatives (nCONN), thiourea (nCSNN) (such as molecules of 1-9, 69-86), amide(nCONH2), thioamide (nRCONHR), hydrazone (nC=N-N), thiocarbohydrazone (such as molecules of 1-9) on the antimicrobial activity. It also indicate the positive effect of nBnz (number of benzene rings) with the halogen substitution on 5 position of isatin ring F > Cl > Br (nArX) (molecules series 14-19). This equation also shows nArCN (Aromatic nitrile such as compounds 22, 25, 28, 31, 34, 37, 39), nPyridines (pyridine derivetives) have negative effects on the antimicrobial activity. May be there isn't any electron with drawing group in the receptor site, thus for design of new compounds it's better to don't use these substitution on the backbone of compounds. The negative sign of this group proposed that a decrease in the number of these descriptors resulted in an activity enhancement. This equation, which has a high statistical quality (R^2 = 0.63, Q^2 = 0.59).

The statistical parameters of prediction, listed in Table 4, indicate the suitability of the proposed QSAR model based on MLR analysis of molecular descriptors. The correlation coefficient of prediction is 0.62, which means that the resulted QSAR model could predict 62% of variances in the antimicrobial activity data. It has root mean square error of 0.21.

3.2 GA-PLS Model

Multicolinearity is a real problem in MLR analysis. This problem in the descriptors is omitted by PLS analysis. In fact, in PLS analysis, the descriptors data matrix is decomposed to orthogonal matrices with an inner relationship

between the dependent and independent variables. This modeling method coincides with noisy data better than MLR, because a minimal number of latent variables are used for modeling in PLS. In GA-PLS analysis a variable selection method is used to find the more convenient set of descriptors because redundant variables degrade the performance of PLS analysis, similar to other regression methods. In the present study, GA was used as variable selection method. The data set (n = 86) was divided into two groups: calibration set (n = 70) and prediction set (n = 16). Given 70 calibration samples; cross-validation procedure was used to find the optimum number of latent variables for each PLS model. In this work, in each run of GA-PLS method a large number of acceptable models were created. GA produces apopulation of acceptable models in each run. In this work, many different GA-PLS runs were conducted using different initial set of populations (50-250) and therefore a large number of acceptable models were created. The most convenient GA-PLS model that resulted in the best fitness contained 8 descriptors including, one constitutional descriptor (nCIC), one 3D MoRSE descriptors (MOR30M) parameter, two WHIM descriptor (P1P, E1U) and four functional descriptors (nRCONHR, nCONN, nArCN, nPyridines). The majority of these descriptors are functional indices All of them being those obtained by different MLR-based QSAR models. The PLS estimate of the regression coefficients are shown in Fig. 2.

Table 3. The results of MLR analysis with different types of descriptors

Equation	Descriptors	(+) effect	(-) effect	R^2	F	Q^2	SE
E1	chemical	--	Logp	0.72	5.3	0.67	0.21
E2	constitutional	--	NCIC,nR10	0.58	18.52	0.52	0.32
E3	topological	--	STN	0.214	18.53	0.17	0.43
E4	Walk and path counts	--	PIPC09	0.283	26.81	0.21	0.53
E5	Connectivity indices	X4A	--	0.63	18.99	0.58	0.34
E6	Information indices	BIC5	--	0.38	19.685	0.32	0.36
E7	2D autocorrelation	GATS1M	MATS3E	0.64	10.417	0.58	0.17
E8	Edge adjacency indices	--	EEIG03X	0.231	20.45	0.18	0.55
E9	Burden eigenvalues	--	BEHm1	0.63	28.562	0.58	0.28
E10	Topological charge indices	--	JGI5	0.68	7.50	0.54	0.34
E11	Eigenvalue-based indices	--	LP1	0.25	23.358	0.17	0.54
E12	Geometrical descriptors	--	G(N..F),DISPV	0.46	8.704	0.39	0.41
E13	RDF descriptors	RDF030M	RDF020M	0.65	10.05	0.59	0.23
E14	3D MoRSE descriptors	--	MOR30M,MOR24U	0.69	17.67	0.62	0.38
E15	WHIM descriptors	E1U,G1M	E3U,P1P	0.61	15.25	0.56	0.46
E16	GETAWAY descriptors	R4M	R2V,HOU,HTM	0.59	10.88	0.53	0.39
E17	Fuctional group counts	nBnz, nCONN, nRCONHR, nArX, nC=N-N	nArCN, nPyridines	0.63	15.29	0.59	0.34
E18	Atom-centred fragments	C-039,C-034	--	0.56	13.51	0.49	0.24
E19	Charge descriptors	QMEAN, QPOS		0.58	13.52	0.52	0.54

This model not only has a high cross-validation statistics, but also represents a high ability for modeling external test samples. It could explain and predict about 75% of variances in the antimicrobial activity (against *Ecoli*) of the studied molecules. There is a close agreement between the experimental and predicted values of antimicrobial activity data.

To measure the significance of the 8 selected PLS descriptors in the protein tyrosine kinase inhibitory activity; In order to investigate the relative importance of the variable appeared in the final model obtained by GA-PLS method, variable important in projection (VIP) was employed [43]. VIP values reflect the importance of terms in PLS model. According to Erikson et al. X-variables (predictor variables) could be classified according to their relevance in explaining y (predicted variable), so that VIP > 1.0 and VIP < 0.8 mean highly or less influential, respectively, and 0.8 < VIP< 1.0 means moderately influential. The VIP analysis of PLS equation is shown in Fig. 3. As it is observed, nRCONHR, P1P and E1U indices represent the most significant contribution in the resulted QSAR model. In addition, parameters such as nArCN and MOR30M have been found to be moderately influential parameters.

3.3 FA-MLR and PCRA

FA-MLR was performed on the dataset. Factor analysis (FA) was used to reduce the number of

Table 4. Statistical parameters for testing prediction ability of the MLR, GA-PLS, PCR, and FA-MLR models

Model	R^2	R^2_{LOOCV}	RMSEcv	R^2p	RMSEp
MLR	0.67	0.62	0.21	0.74	0.18
GA-PLS	0.81	0.75	0.19	0.87	0.23
PCR	0.73	0.68	0.16	0.78	0.22
FA-MLR	0.64	0.59	0.15	0.70	0.14

R^2: Regression Coefficient for Calibration set
R^2_{LOOCV}: Regression Coefficient for Leave One Out Cross Validation
$RMSE_{cv}$: Root Mean Square Error of cross validation
R^2p: Regression Coefficient for prediction set; RMSEp: Root Mean Square Error of prediction set

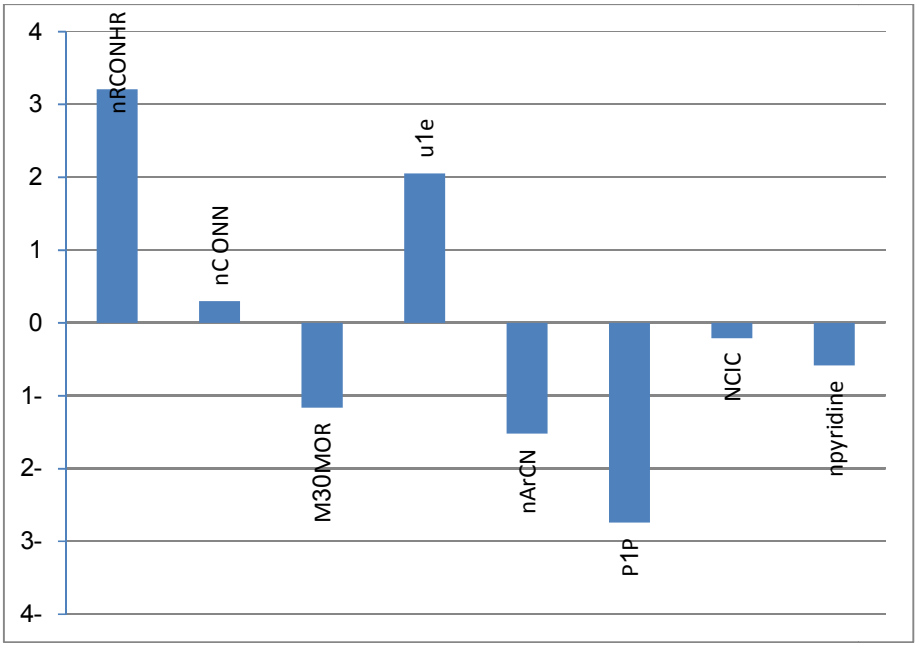


Fig. 2. PLS regression coefficients for the variables used in GA-PLS model

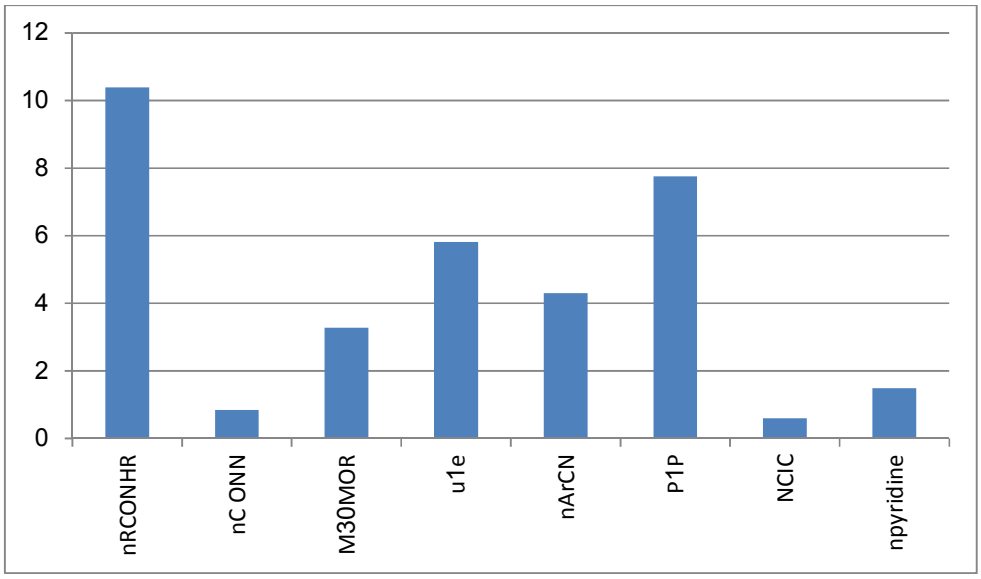


Fig. 3. Plot of variables important in projection (VIP) for the descriptors used in GA-PLS model

variables and to detect structure in the relationships between them. This data-processing step is applied to identify the important predictor variables and to avoid collinearities among them [44]. Principle component regression analysis, PCRA, was tried for the dataset along with FA-MLR. With PCRA collinearities among **X** variables are not a disturbing factor and the number of variables included in the analysis may exceed the number of observations [45]. In this method, factor scores, as obtained from FA, are used as the predictor variables [44]. In PCRA, all descriptors are assumed to be important while the aim of factor analysis is to identify relevant descriptors.

Table 5 shows the nine factor loadings of the variables (after VARIMAX rotation) for the compounds tested for cytotoxic activity. As it is observed, about 81.2% of variances in the original data matrix could be explained by the selected nine factors.

Based on the procedure explained in the experimental section, the following two-parametric equation was derived (Table 6).

Y= 5.766(±0.547) -0.031(±0.012) MOR30m -2.336(±0.721) JGI5

R^2 = 0.64 *S.E* = 0.24 F = 14.69 Q^2 = 0.59 *RMScv* = 0.12

This equation could explain about 59% of the variance and predict 64% of the variance in pMIC data. It has a root mean square error of 0.12. This equation describes the effect of 3D-MORSE and Galvez Topological charge descriptors (MOR30m and JGI5) and on antimicrobial activity of the studied compounds.

When factor scores were used as the predictor parameters in a multiple regression equation using forward selection method (PCRA), the following equation was obtained (Table 7):

Y=4.285(±0.57)-0.258(±0.031)F1-.189(±0.021) F7+0.124(±0.027)F6+0.123(0.057)F8

R^2 = 0.73 S.E. = 0.34 F = 15.54 Q^2 = 0.68 RMScv = 0.15

This equation could explain and predict 68% and 73% of the variances in pMIC data, respectively. The root mean square error of PCRA analysis was 0.15. Since factor scores are used instead of selected descriptors, and any factor-score contains information from different descriptors, loss of information is thus avoided and the quality of PCRA equation is better than those derived from FA-MLR. Whilst the data of this analysis show acceptable prediction, we see that the predicted values of some molecules are near to each other.

As it is observed from Table 5, in the case of each factor, the loading values for some descriptors are much higher than those of the others. These high values for each factor indicate that this factor contains higher information about

which descriptors. It should be noted that all factors have information from all descriptors but the contribution of descriptor in different factors are not equal. For example, factors 1 and 2 have higher Constitutional, Charge, WHIM, Atom-center, Connectivity, Functional, MORSE and GETAWAY whereas information about RDF, MORSE, burden eigenvalues 2Dautocorrelations and functional descriptors are highly incorporated in factor 3 and 4. Factor score 5, 6, 7 and 8 signify the importance of functional chemical and Atom-center descriptors.

Table 5. Numerical values of factor loading numbers 1–9 for descriptors after VARIMAX rotation

	Component									
	F1	**F2**	**F3**	**F4**	**F5**	**F6**	**F7**	**F8**	**F9**	**extraction**
LOGP	.190	.152	.038	-.175	.002	.091	.724	-.083	-.123	.646
NCIC	.897	.188	-.017	-.184	-.060	-.107	.013	.137	-.030	.909
NR10	.774	.073	-.091	-.021	.210	.398	.064	.254	-.063	.887
X4A	-.655	-.587	-.004	-.101	-.173	-.085	.006	.283	.105	.912
BIC5	-.304	-.314	.059	.025	-.444	.461	-.225	.373	-.035	.795
MATS3E	.456	.541	-.085	-.100	.145	-.301	.304	.094	-.286	.812
GATS1M	-.293	.306	-.664	.155	-.050	.184	-.002	.004	-.106	.691
BEHM1	-.006	.183	.851	.034	-.011	-.074	-.017	.001	-.147	.785
JGI3	.004	-.125	-.106	.094	.009	.932	.078	-.006	-.017	.910
JGI5	.325	.150	-.009	.105	.818	.107	-.040	-.074	.049	.830
L/Bw	-.319	-.787	.151	-.090	-.008	-.197	.129	.042	.059	.813
DISPV	-.059	-.107	.286	-.194	.547	-.357	.240	-.047	-.179	.653
RDF020M	-.176	.316	.653	.314	-.214	-.022	.245	-.027	-.144	.783
RDF030M	.019	.893	.188	-.021	.007	.026	.182	.003	-.023	.869
RDF025P	.485	.724	-.104	-.190	.170	-.219	.190	.045	-.078	.927
MOR24U	.230	-.099	-.006	-.719	.283	-.156	.092	.312	.046	.792
MOR30M	.631	.100	-.125	.102	-.282	-.142	.221	-.394	.014	.738
MOR32V	-.002	.245	.067	.757	.268	.076	-.113	.104	.011	.739
E1U	-.736	-.493	-.028	-.006	-.211	-.136	.002	.148	.011	.872
E3U	.128	.543	.089	.457	-.215	-.196	.215	-.259	.278	.803
G1M	-.073	-.677	.092	-.079	-.148	.237	-.312	-.159	.293	.764
P1P	-.013	-.898	-.035	-.285	.090	.011	.048	.159	-.065	.930
H0U	.150	.257	-.114	.322	.354	-.172	.537	-.259	.226	.767
HTM	.020	.747	.581	.092	.172	-.120	.073	.039	.015	.956
R4M	.079	.709	.444	.218	.029	.062	.052	-.032	.259	.829
R2V	.773	.204	.144	.154	.031	-.063	.225	.018	.401	.900
nCb	.728	-.199	.158	-.150	.180	-.237	-.069	-.010	-.096	.720
NRCONHR	-.592	-.406	.136	.290	-.107	-.205	-.411	.066	.006	.845
NCONN	-.708	.282	-.109	.211	.162	.125	.290	.300	-.136	.872
NARCN	.617	.033	.028	.203	.179	.036	.107	.112	-.012	.480
nC=N-N	-.687	.035	-.107	-.158	.158	-.006	.139	.540	.080	.851
NArX	.140	-.046	.652	-.024	.229	.173	-.350	-.160	.260	.746
nPyridines	-.062	-.069	-.037	-.007	.000	-.009	-.084	-.025	.876	.784
C-034	.008	-.037	-.064	-.052	-.151	.010	-.119	.839	-.051	.752
C-039	-.868	-.138	-.051	.082	-.260	-.087	-.092	.070	.016	.870
QPOS	-.367	.881	-.057	.111	.027	-.161	.082	.029	-.038	.962
QMEAN	-.732	.348	-.053	.222	.085	.161	-.304	.123	.048	.853
% variance	21.769	19.668	7.805	5.894	5.753	5.520	5.368	5.277	4.153	81.208

Table 6. The results of FA-MLR analysis with different types of descriptors

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	R^2	F	Q^2	SE
	B	Std. Error	Beta						
(Constant)	5.766	.547		10.537	.000	0.64	14.691	0.59	0.24
MOR30m	-0.031	.012	-.303	-2.619	.000				
JGI5	-2.336	0.721	-.442	-3.240	.002				

Table 7. The results of PCR analysis

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	R^2	F	Q^2	SE
	B	Std. Error	Beta						
(Constant)	4.285	.057		75.368	.000				
F1	-.258	.031	-.409	-4.510	.000	0.73	17.16	0.68	0.34
F7	-.189	.021	-.300	-3.305	.001				
F6	.124	.027	.197	2.176	.033				
F8	.123	.057	.196	2.159	.034				

3.4 Robustness and Applicability Domain of the Models

Leverage is one of standard methods for this purpose. Warning leverage (h^*) is another criterion for interpretation of the results. The warning leverage is, generally, fixed at $3k/n$, where n is the number of training compounds and k is the number of model parameters. A leverage greater than warning leverage h^* means that the predicted response is the result of substantial extrapolation of the model and therefore may not be reliable [46]. The calculated leverage values of the test set samples for different models and the warning leverage, as the threshold value for accepted prediction, are listed in Table 8. As seen, the leverages of all test samples are lower than h^* for all models. This means that all predicted values are acceptable.

3.5 Molecular Docking Studies

All the eighty six isatin derivatives were docked into the active site of the enzymes bacterial enoyl-ACP reductase (FabI) (PDB:1Ix6). All the docking protocols were done on validated structures, with RMSD values below 2 Å. The conformation with the lowest ones was considered as the best docking result. Docking binding energies of these active compounds were summarized in Table 9.

Table 8. Leverage (h) of the external test set molecules for different models. The last row (h^*) is the warning leverage

Molecule .no	MLR	GA-PLS	PCR	FA-MLR
3	0.117418	0.084802	0.026426	0.028202
7	0.058532	0.067963	0.03644	0.034729
10	0.087443	0.157524	0.131804	0.035335
11	0.071099	0.093302	0.092915	0.021066
19	0.054337	0.08314	0.03296	0.037432
39	0.081619	0.077263	0.136844	0.040156
41	0.097168	0.134119	0.13121	0.056011
45	0.158855	0.144921	0.152167	0.036003
69	0.045048	0.101806	0.06149	0.068055
71	0.109807	0.13409	0.023009	0.022631
72	0.102708	0.297308	0.041009	0.060281
73	0.105906	0.198805	0.022111	0.063121
75	0.087529	0.127991	0.018691	0.025659
76	0.04769	0.084609	0.021734	0.045611
82	0.081846	0.058078	0.022526	0.016686
83	0.077447	0.07017	0.016547	0.014426
h*	0.214286	0.342857	0.191429	0.108571

Docking analysis showed that Compounds 1-9 with thiocarbazone moiety, was good inhibitor for FabI, because of good interaction between enzyme and cofactor. With suitable orientation of thiocarbazone group, hydrogen and hydrophobic bounds can occur. An electron rich group such as NH_2 substitution on phenyl ring that increase electron charge can create better interaction with receptor and has low binding energy. Also compounds have benzhydrazide substitution hydrogen binding interaction between tyrosine 156 and benzhydrazide. Halogen and methyl substitution on isatin ring of this series (1-9) can cause better interaction with receptor. Benzyl amide, methyl and halogen groups on isatin ring of compound 10-19 showed good docking score. But Halogen and methyl groups at C-5 isatin ring of series (40-43) can cause bad interaction with receptor. Urea moiety can show good interaction between tyrosine 156 and coenzyme. Our results indicated Benzamide group show good interaction but OH group has bad interaction with receptors. Compounds 47 and 48 with isonicotine amide group on isatin ring showed good interaction with receptor. The interaction modes of 2, 14 and 27, 40 those with the best docking score are shown in Fig. 4.

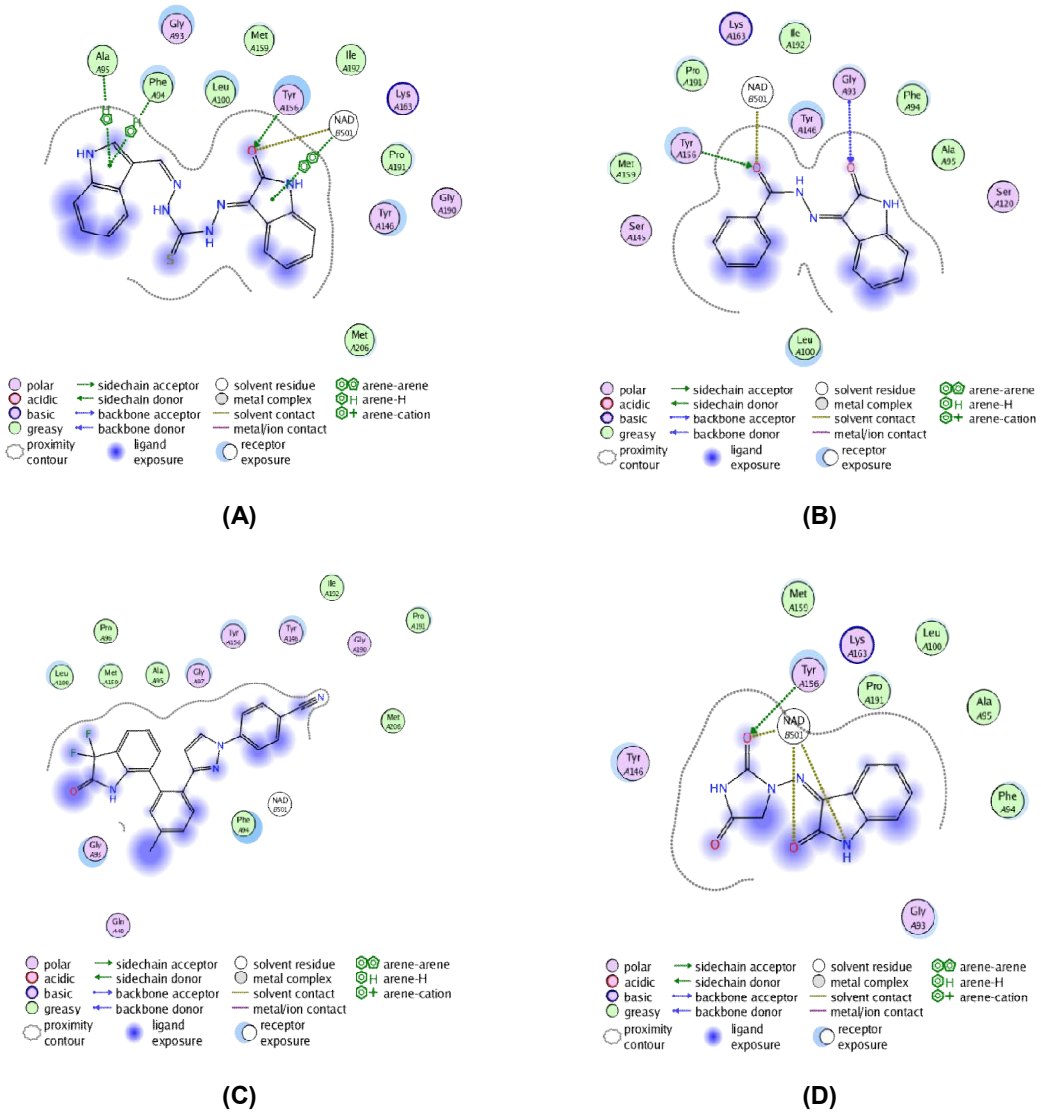


Fig. 4. The docked configuration of 2 (A), 14(B), 27(C) and 40 (D) in the binding site of FabI

Table 9. Binding interaction of some studied compounds in active site of enzyme

Compounds	ΔG kcal/mol	K_i (μM)	Atom of the ligand	Receptor/Coenzyme	Interaction	Distance(A°)
1	-8.46	631.45	C=O (isatine)	Tyr 156	Hydrogen bond	1.80
			C=O (isatine)	2'-OH (ribose)NAD^+	Hydrogen bond	1.85
2	-9.01	248.47	C=O (isatine)	Tyr 156	Hydrogen bond	1.82
			C=O (isatine)	2'-OH (ribose)NAD^+	Hydrogen bond	1.83
			Isatine	NAD^+ -Pyridine	Pi-Pi	
			Indole	Phe 94	H-Pi	
			Indole	Ala 95	H-pi	
3	-8.58	512.32	C=O (isatine)	Tyr 156	Hydrogen bond	1.81
			C=O (isatine)	2'-OH (ribose)NAD^+	Hydrogen bond	1.89
			Isatine	NAD^+ -Pyridine	Pi-Pi	
4	-8.86	322.58	C=O (isatine)	Tyr 156	Hydrogen bond	1.77
			C=O (isatine)	2'-OH (ribose)NAD^+	Hydrogen bond	1.94
			C=S	NH_2- NAD^+	Hydrogen bond	2.66
			Isatine	Pyridine- NAD^+	Pi-Pi	
			Indole	Phe 94	H-Pi	
5	-8.67	443.90	C=O (isatine)	Tyr 156	Hydrogen bond	1.78
			C=O (isatine)	2'-OH (ribose)NAD^+	Hydrogen bond	1.89
			Isatine	NAD^+ -Pyridine	Pi-Pi	
			Indole	Phe 94	H-Pi	
6	-8.24	912.17	C=O (isatine)	Tyr 156	Hydrogen bond	1.85
			C=O (isatine)	2'-OH (ribose)NAD^+	Hydrogen bond	1.88
			Isatine	NAD^+ -Pyridine	Pi-Pi	
7	-8.51	577.86	C=O (isatine)	Tyr 156	Hydrogen bond	1.95
			C=O (isatine)	2'-OH (ribose)NAD^+	Hydrogen bond	2.09
			Isatine	NAD^+ -Pyridine	Pi-Pi	
8	-8.46	630.94	C=O (isatine)	Tyr 156	Hydrogen bond	1.89
			C=O (isatine)	2'-OH (ribose)NAD^+	Hydrogen bond	1.95
9	-8.47	614.00	C=O (isatine)	2'-OH (ribose)NAD^+	H-Pi	
			NH-thiocarboydrazone	Met 159	Hydrogen bond	1.87
			N-thiocarboydrazone	Tyr 156	Hydrogen bond	2.35
			Isatine	Ala 95	H-Pi	

Compounds	ΔG kcal/mol	K_i (μM)	Atom of the ligand	Receptor/Coenzyme	Interaction	Distance(A°)
14	-7.44	3.52	C=O(benzohydrazide)	Tyr 156	Hydrogen bond	2.22
			C=O(benzohydrazide)	2′-OH-ribose(NAD^+)	Hydrogen bond	2.17
			C=O(isatine)	Gly 93	Hydrogen bond	2.66
15	-7.94	1.52	C=O(benzohydrazide)	Tyr 156	Hydrogen bond	1.88
			C=O(benzohydrazide)	2′-OH-ribose(NAD^+)	Hydrogen bond	1.52
			C=O(isatine)	Gly 93	Hydrogen bond	2.02
16	-7.87	1.66	C=O(benzohydrazide)	Tyr 156	Hydrogen bond	2.19
			C=O(benzohydrazide)	2′-OH-ribose(NAD^+)	Hydrogen bond	2.03
			C=O(isatine)	Gly 93	Hydrogen bond	2.36
			Phenyl	Tyr 146	H-Pi	
17	-7.79	1.95	C=O(benzohydrazide)	Tyr 156	Hydrogen bond	1.97
			C=O(benzohydrazide)	2′-OH-ribose(NAD^+)	Hydrogen bond	2.12
			C=O(isatine)	Gly 93	Hydrogen bond	2.92
			Phenyl	Tyr 146	H-Pi	
18	-7.81	1.88	C=O(benzohydrazide)	Tyr 156	Hydrogen bond	1.98
			C=O(benzohydrazide)	2′-OH-ribose(NAD^+)	Hydrogen bond	1.63
			C=O(isatine)	Gly 93	Hydrogen bond	3.15
19	-7.86	1.72	C=O(benzohydrazide)	Tyr 156	Hydrogen bond	1.79
			C=O(benzohydrazide)	2′-OH-ribose(NAD^+)	Hydrogen bond	2.23
			C=O(isatine)	Gly 93	Hydrogen bond	2.46
20	-6.57	15.30	Pyrazole	Tyr 156	H-Pi	
21	-6.42	19.80	Pyrazole	Tyr 156	H-Pi	
22	-6.21	28.20				
23	-6.28	25.03				
24	-6.30	26.91				
25	-6.79	10.57	Pyrazole	Tyr 156	H-Pi	
26	-6.33	27.55				
28	-6.25	26.00				
29	-6.46	18.25				

Compounds	ΔG kcal/mol	K_i (μM)	Atom of the ligand	Receptor/Coenzyme	Interaction	Distance(A°)
30	-6.03	38.14				
31	-6.37	21.35				
32	-6.35	22.18				
33	-6.12	32.63				
34	-6.30	24.20				
35	-6.02	23.42				
36	-6.26	26.17				
37	-6.65	13.38				
38	-6.65	13.38	2-Cl	Leu100	Hydrogen bond	2.90
39	-6.68	12.61	2-Cl	Leu100	Hydrogen bond	2.83
40	-8.47	618.68	C=O (isatine)	NH_2(NAD^+)	Hydrogen bond	2.18
		nano	N-H (isatine)	Phosphate (NAD^+)	Hydrogen bond	2.33
			C=O (imidazolidine)	Tyr 156	Hydrogen bond	1.81
			C=O (imidazolidine)	2'-OH (ribose)NAD^+	Hydrogen bond	1.83
41	-8.30	828.33 nano	C=O (isatine)	Tyr 156	Hydrogen bond	1.70
			C=O (isatine)	2'-OH (ribose)NAD^+	Hydrogen bond	1.76
			Isatine	NAD^+ -Pyridine	Pi-Pi	
42	-8.08	1.19	C=O (isatine)	Tyr 156	Hydrogen bond	1.81
			C=O (isatine)	2'-OH (ribose)NAD^+	Hydrogen bond	1.89
			Isatine	NAD^+ -Pyridine	Pi-Pi	
43	-7.93	1.53	C=O (isatine)	Tyr 156	Hydrogen bond	1.77
			C=O (isatine)	2'-OH (ribose)NAD^+	Hydrogen bond	1.94
			Isatine	NAD^+ -Pyridine	Pi-Pi	
44	-8.39	709.17 nano	C=O (isatine)	Tyr 156	Hydrogen bond	1.91
			C=O (isatine)	2'-OH (ribose)NAD^+	Hydrogen bond	1.64
			NH_2 (urea)	Phosphate (NAD^+)	Hydrogen bond	2.81
			NH_2 (urea)	Phosphate (NAD^+)	Hydrogen bond	3.74
46	-8.24	912.17	C=O (isatine)	Gly 93	Hydrogen bond	2.44
			C=O (benzamide)	Tyr 156	Hydrogen bond	1.86
			C=O (benzamide)	2'-OH (ribose)NAD^+	Hydrogen bond	2.30
47	-8.33	778.55 nano	C=O (isatine)	Tyr 156	Hydrogen bond	1.81
			C=O (isatine)	2'-OH (ribose)NAD^+	Hydrogen bond	1.81
			C=O (nicotinamide)	Gly 93	Hydrogen bond	2.10
48	-8.22	943.86 nano	C=O (isatine)	Tyr 156	Hydrogen bond	1.97

Compounds	ΔG kcal/mol	K_i (μM)	Atom of the ligand	Receptor/Coenzyme	Interaction	Distance(A°)
			C=O (isatine)	2'-OH (ribose)NAD^+	Hydrogen bond	1.63
			C=O (nicotinamide)	Gly 93	Hydrogen bond	2.80
49	-8.57	519.34	C=O (isatine)	Tyr 156	Hydrogen bond	1.83
			C=O (isatine)	2'-OH (ribose)NAD^+	Hydrogen bond	1.68
50	-8.47	618.68	C=O (isatine)	Tyr 156	Hydrogen bond	2.01
			C=O (isatine)	2'-OH (ribose)NAD^+	Hydrogen bond	1.76
52	-7.58	2.79 (μM)	C=O (isatine)	Tyr 156	Hydrogen bond	1.81
			C=O (isatine)	2'-OH (ribose)NAD^+	Hydrogen bond	1.85
53	-7.80	1.90	C=O (isatine)	Tyr 156	Hydrogen bond	1.87
		(μM)	C=O (isatine)	2'-OH (ribose)NAD^+	Hydrogen bond	2.05
			Isatine	NAD^+	Pi-Pi	

4. CONCLUSION

Quantitative relationships between molecular structure and antibacterial activity of isatin derivatives were discovered by four chemometrics methods: MLR, GA-PLS, PCR and FA-MLR. MLR analysis explained the positive effect of the number of urea derivatives (nCONN), thio urea (nCSNN), amide (nCONH2), thioamide (nRCONHR), hydrazone (nC=N-N), thiocarbohydrazone on the antimicrobial activity. It also indicate the positive effect of nBnz (number of benzene rings) with the halogen substitution on 5 position of isatin ring F > Cl > Br (nArX). This equation also shows nArCN, nPyridines (pyridine derivetives) have negative effects on the antimicrobial activity of studied compound. The FA-MLR describes the effect of 3D-MORSE and Galvez Topological charge descriptors (MOR30m and JGI5) and on antimicrobial activity of the studied compounds. The quality of PCRA equation is better than those derived from FA-MLR. Whilst the data of this analysis show acceptable prediction, we see that the predicted values of some molecules are near to each other. Factors 1 and 2 have higher Constitutional, Charge, WHIM, Atom-center, Connectivity, Functional, MORSE and GETAWAY whereas information about RDF, MORSE, burden eigenvalues 2Dautocorrelations and functional descriptors are highly incorporated in factor 3 and 4. Factor score 5, 6, 7 and 8 signify the importance of functional chemical and Atom-center descriptors. A comparison between the different statistical methods employed revealed that GA-PLS represented superior results and it could explain and predict 73% and 68% of variances in the –LogMIC data, respectively. Comparison between QSAR and docking analysis revealed that by decreasing in number of ring and lipophilicity (also logp) for design of new compounds can have better activity. Substitutions such as urea, thiourea, thiocarbohydrazone, benzhydrazide as on isatin ring, can cause better interaction with receptor.

CONSENT

It is not applicable.

ETHICAL APPROVAL

It is not applicable.

ACKNOWLEDGEMENTS

Financial support of Shiraz University of Medical Sciences (Shiraz, I.R. Iran) through the Grant No. 97-01-36-17343 is highly appreciated.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Heath RJ, Rubin JR, Holland DR, Zhang E, Snow ME, Rock CO. Mechanism of triclosan inhibition of bacterial fatty acid synthesis, J Biol Chem. 1999;16;274(16): 11110-4.
2. Pandeya SN, Smitha S, Jyoti M, Sridhar S. Acta Pharm. 2005;55:27–46.
3. Sharma VM, Prasanna P, Seshu VA, Renuka B, Rao VL, Kumar GS, Narasimhulu CP, Babu PA, Puranik RC, Subramanyam D, Venkateswarlu A, Rajagopal S, Kumar KBS, Rao CS, Mamidi NVSR, Deevi DS, Ajaykumar R, Rajagopalan R. Bioorg. Med. Chem. Lett. 2002;12:2303–2307.
4. Verma M, Nath PS, Nandsingh K, Stables JP. Acta Pharm. 2004;54:49-54.
5. Pandeya SN, Sriram D, Nath G, De Clercq E. Eur. J. Med. Chem. 2000;35:249-255.
6. Igosheva N, Lorz C, O'Connor E, Glover V, Mehmet H. Neurochem. Int. 2005;47:(216-224.
7. Moon MJ, Lee SK, Lee J-W, Song WK, Kim SW, Kim JI, Cho C, Choi SJ,. Kim Y-C, Bioorg. Med. Chem.2006;14:237–246.
8. Abadi AH, Abou-Seri SM, Abdel-Rahman DE, Klein C, Lozach O, Meijer L. Eur. J. Med. Chem. 2006;41:296–305.
9. Gursoy N. Karali. Eur. J. Med. Chem. 2003;38:633–643.
10. Schmidi H. Multivariate prediction for QSAR, Chemom. Intell. Lab. Syst. 1997; 37:125-134.
11. Hansch C, Kurup A, Garg R, Gao H, Chem-bioinformatics and QSAR: A review of QSAR lacking positive hydrophobic terms, Chem. Rev. 2001;101:619-672.
12. Wold S, Trygg J, Berglund A, Antii H, Some recent developments in PLS modeling, Chemom. Intell. Lab. Syst. 2001; 58:131-150.
13. Sabet R, Fassihi A, Hemmateenejad B, Saghaie L, Miri R, Gholami M, Computer-aided drug design of novel antibacterial 3-hydroxypyridine-4-ones: application of QSAR methods based on the MOLMAP approch. Journal of Computer-Aided Molecular Design. 2012;26:349-361.

14. Sabet R, Fassihi A, Moeinifard B. QSAR study of PETT Derivatives as Potent HIV-Reverse Transcriptase Inhibitors. J. Mol. Graph & Model. 2009;28:146-155.
15. Hansch C, Fujita T. ρ-σ-π Analysis. A method for the correlation of biological activity and chemical structure. J. Am. Chem. Soc. 1964;86:1616-1626.
16. Wang J, Zhang L, Yang G, Zhan CG, Quantitative structure-activity relationship for cyclic imide derivatives of protoporphyrinogen oxidase inhibitors: A study of quantum chemical descriptors from density functional theory, J. Chem. Inf. Comput. Sci. 2004;44:2099-2105.
17. Hansch C, Hoekman D, Gao H. Comparative QSAR: Toward a deeper understanding of chemicobiological interactions, Chem. Rev. 1996;96:1045-1075.
18. Todeschini R, Consonni V. Handbook of Molecular Descriptors. Wiley-VCH, Weinheim; 2000.
19. Horvath D, Mao B. Neighborhood behavior. Fuzzy molecular descriptors and their influence on the relationship between structural similarity and property similarity, QSAR Comb. Sci. 2003;22:498-509.
20. Putta S, Eksterowicz J, Lemmen C, Stanton R. A novel subshape molecular descriptor, J. Chem. Inf. Comput. Sci. 2003;43:1623-1635.
21. Gupta S, Singh M, Madan AK. Superpendentic index: A novel topological descriptor for predicting biological activity. J. Chem. Inf. Comput. Sci. 1999;39:272-277.
22. Consonni V, Todeschini R, Pavan M. Structure/response correlations and similarity/diversity analysis by GETAWAY descriptors. 2. Application of the novel 3D molecular descriptors to QSAR/QSPR studies, J. Chem. Inf. Comput. Sci. 2002; 42:693-705.
23. Hyper Chem, Release 8.0 for Windows, Molecular Modeling System: Hyper Cube.
24. Todeschini R. Milano Chemometrics and QSAR Group. Available:http://michem.disat.unimib.it/
25. Fassihi A, Sabet R. QSAR study of p56lck protein tyrosine Kinase Inhibitory Activity of Flavonoid Derivatives Using MLR and GA-PLS. Int. J. Mol. Sci. 2008;9:1876-1892.
26. Sabet R, Fassihi A. QSAR study of antimicrobial 3-Hydroxypyridin-4-one and 3-Hydroxypyran-4-one Derivatives Using Different Chemometric Tools. Int. J. Mol. Sci. 2008;9:2407-2423.
27. Fassihi A, Abedi D, Saghaie L, Sabet R, Fazeli H, Bostaki GH. Deilami O, Sadinpour H. Synthesis, antimicrobial evaluation and QSAR study of some 3-hydroxypyridine-4- one and 3-hydroxypyran-4-one Derivatives. Eur. J. Med. Chem. 2009;44:2145-2157.
28. ConsonniV, Todeschini R, Pavan M. J. Chem. Inf. Comput. Sci. 2002;42:693-705.
29. Thanh ND, Giang NTK, Quyen TH, Huong DT, Toan VN. Synthesis and evaluation of in vivo antioxidant, in vitro antibacterial, MRSA and antifungal activity of novel substituted isatin N-(2, 3, 4, 6-tetra-O-acetyl-β-d-glucopyranosyl) thiosemicarbazones. Eur J Med Chem. 2016;123:532-43.
30. Tehrani KH, Hashemi M, Hassan M, Kobarfard F, Mohebbi S. Synthesis and antibacterial activity of Schiff bases of 5-substituted isatins. Chin Chem Lett. 2016; 27(2):221-5.
31. Sridhar SK, Saravanan M, Ramesh A. Synthesis and antibacterial screening of hydrazones, Schiff and Mannich bases of isatin derivatives. Eur J Med Chem. 2001; 36(7):615-25.
32. Šekularac GM, Nikolić JB, Petrović P, Bugarski B, Đurović B, Drmanić SŽ. Synthesis, antimicrobial and antioxidative activity of some new isatin derivatives. Journal of the Serbian Chemical Society. 2014;79(11):1347-54.
33. Prakash CR, Raja S. Synthesis, characterization and in vitro antimicrobial activity of some novel 5-substituted Schiff and Mannich base of isatin derivatives. Journal of Saudi Chemical Society. 2013; 17(3):337-44.
34. Lian Z-M, Sun J, Zhu H-L. Design, synthesis and antibacterial activity of isatin derivatives as FtsZ inhibitors. J Mol Struct. 2016;1117:8-16.
35. Kiran G, Maneshwar T, Rajeshwar Y, Sarangapani M. Microwave-assisted synthesis, characterization, antimicrobial and antioxidant activity of some new isatin derivatives. J Chem. 2012;2013.
36. Leardi R. Genetic Algorithms in Chemometrics and Chemistry: A Review. J. Chemometrics. 2001;15:559-569.
37. Sabet R, Fassihi A, Saghaie L. Octanol-water partition coefficients determination and QSPR study of some 3-hydroxy pyridine-4-one derivatives, Journal of

Pharmaceutical Research International. 2018;22(4):1-15.
38. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. Journal of Molecular Graphics. 1996;14(1):33-8.
39. Fereidoonnezhad M, Faghih Z, Mojaddami A, Sakhteman A, Rezaei Z. A Comparative Docking Studies of Dichloroacetate Analogues on Four Isozymes of Pyruvate Dehydrogenase Kinase in Humans. Indian J Pharm Educ. 2016;50(2):S32-S8.
40. BF Mirjalili, L Zamani, K Zomorodian, S Khabnadideh, Z Haghighijoo, Z Malakotikhah. Synthesis, antifungal activity and docking study of 2-amino-4H-benzochromene-3-carbonitrile derivatives. Journal of Molecular Structure. 2016; 1116:102-8.
41. Li Z, Gu J, Zhuang H, Kang L, Zhao X, Guo Q. Adaptive molecular docking method based on information entropy genetic algorithm. Applied Soft Computing. 2015;26:299-302.
42. Feng J, Ablajan K, Sali A. 4-Dimethylaminopyridine-catalyzed multi-component one-pot reactions for the convenient synthesis of spiro[indoline-3,4′-pyrano[2,3-c]pyrazole] derivatives. Tetrahedron. 2014;70(2):484-9.
43. Olah M, Bologa C, Oprea TI. An Automated PLS Search for Biologically Relevant QSAR Descriptors. J. Comput. Aided Mol. Des. 2004;18:437-449.
44. Franke R, Gruska A. Chemometrics Methods in molecular design, in: H. van Waterbeemd, (Ed.), Methods and Principles in Medicinal Chemistry, VCH, Weinheim. 1995;2:113–119.
45. Kubinyi H. The quantitative analysis of structure-activity relationships, in: M.E. Wolff, (Ed.), Burger's Medicinal Chemistry and Drug Discovery, 5th Ed.; Wiley, New York. 1995;1:506-509.
46. Brereton R. Chemometrics Data Analysis for the Laboratory and Chemical Plant. Wiley. 2004;47–54.

Peer-review history:
The peer review history for this paper can be accessed here:
https://sdiarticle4.com/review-history/52361