ACADEMIA ROMÂNĂ
Revue Roumaine de Chimie
http://web.icf.ro/rrch/

Rev. Roum. Chim.,
2018, *63*(3), 217-226

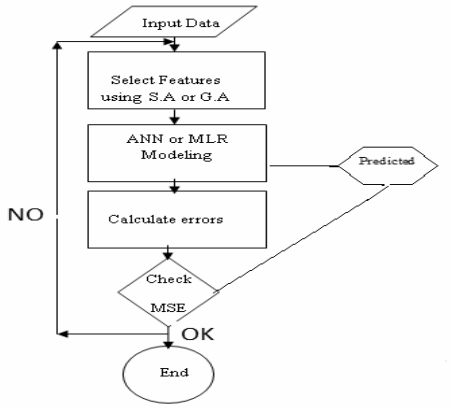
QSAR INVESTIGATION ON THE FEATURE SELECTION OF THE LETROZOLE DERIVATIVES AS ANTICANCER DRUGS

Robabeh SAYYADIKORDABADI,[*,a] Asghar ALIZADEHDAKHEL,[b] Gholam Reza NAJAFI[c] and Mehrnaz MASOMPARAST[c]

[a] Department of Chemistry, Rasht Branch, Islamic Azad University, Rasht, Iran
[b] Department of Chemical Engineering, Rasht Branch, Islamic Azad University, Rasht, Iran
[c] Department of Chemistry, Qom Branch, Islamic Azad University, Qom, Iran

Received September 16, 2017

In the present study, QSAR studies were carried out on twenty seven Letrozole derivatives. Multiple linear regression (MLR) and artificial neural networks (ANN) were used as modelling tools and simulated annealing algorithm (SA) and genetic algorithm (GA) optimization methods were employed to choose the best set of descriptors. The obtained results from four combinations of modelling-optimization methods were compared and GA-ANN combination showed the best performance based on its correlation coefficient and root mean square errors (RMSE). It was concluded that the most effective parameters on the activity of derivatives of Letrozole compounds are nDB, EEig12r, Mor26v, R5m, BIC0, Mor14e, EEig14r descriptors, respectively.



INTRODUCTION

In breast cancer, after surgery, Letrozole is used in hormonal therapy as oral non-steroidal aromatase inhibitor. Estrogens, which are the female sex hormones, play a crucial role in sexual development in women and normal metabolism of bone and lipids, and in the diseases associated with the uterus and ovary.[1-6]

Letrozole are selective aromatase inhibitors and very potent;[7-9] nevertheless, Letrozole[21] is the most active aromatase inhibitor among all, having 99% aromatase inhibition.[10]

Quantitative structure activity relationship (QSAR) studies are employed in molecular design and medicinal chemistry.[11-13] In QSAR[35] models, finding a set of molecular descriptors which have higher impact on the biological activity is the most important step[14-17] and to construct a relationship between chemical structures and biological activities, mathematical equations are used. Multiple linear regression (MLR), genetic algorithm (GA), and simulated annealing algorithm (SA)[15-17, 22, 23] are mostly used in QSAR studies for variable selection.[18,19]

In the current study, Multiple linear regression and artificial neural networks (ANN) as linear and non-linear modelling tools and simulated Annealing (SA), genetic algorithm (GA) optimization methods were applied as to investigate the QSAR in letrozole anticancer drugs. The ability of these methods to predict the inhibitor activity of letrozole anticancer drugs was also compared.

* Corresponding author: sayyadi@iaurasht.ac.ir

THEORY AND COMPUTATIONAL METHODS

Geometry optimizations of letrozole compoundes were carried out using the B3lyp/6-31g at the Gaussian 03W.[24] Dragon program was used for calculation of 3226 molecular descriptors including topological, geometrical, MoRSE,[26,27] RDF,[27,28] GETAWAY,[17,29] auto-correlations[35] and WHIM[30, 31] groups.[25] For each of the 27 compounds and then SPSS[20] program was used to reduce the number of descriptors through an objective feature selection in three steps. These steps include: (i) descriptors having the same value, at least 70% of compounds were removed; (ii) descriptors with correlation coefficient less than 0.25 with the logarithm half maximal inhibitory concentration (-log IC_{50}) as the dependent variables were considered and removed;[32] (iii) by carrying out these two-steps, the number of descriptors was reduced to 1350 and then a stepwise multiple linear regression procedure was employed to select the best descriptors of the 1350 descriptors. Low standard deviation, least numbers of independent variables, high ability for prediction and high F statistic value,[33] high correlation coefficient (R) and RMSE are characteristics of an ideal model, where the RMSE is defined as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - y_o)^2}{n}} \quad (1)$$

where y_i is the desired output, y_o is the predicted value by the model, and n is the number of molecules in the data set.

In QSAR methods including GA-ANN, SA-ANN, MLR-SA, MLR-GA, 1350 descriptors were considered as possible input of the ANN and fed into the input layer of the ANNs. In this study, they were all three-layer and Levenberg-Marquart algorithm[34] was applied for training on the TSET members. (Figure 2). Modelling and optimization calculations were carried out using Matlab. 2014a. These networks were supposed to identify the non-linear relationship between the structural descriptors and inhibitory activity of Letrozole compounds.

RESULTS AND DISCUSSION

Geometrical optimization of Letrozole derivatives was carried out with the B3LYP/6-31G utilizing Gaussian 03.[40] All studied letrozole compounds with the calculated fundamental vibration values are shown in Figure 3. For the whole compounds, it was obtained that the values of NImag are zero and the values of the fundamental vibrations are positive. Therefore, all of the considered compounds are stable.

SPSS [20] and Unscrambler program were used in linear calculation including MLR-MLR, MLR-PCR, MLR-PLS1 methods and are presented in Table 1. The RMSE and the correlation coefficient (R^2) in MLR-MLR for predicted activity were found to be 0.2188, 0.8261 in gas phase, respectively. Also, Table 1 showed that MLR-MLR method was better compared to all the other linear methods (MLR-PLS1 and MLR-PCR).

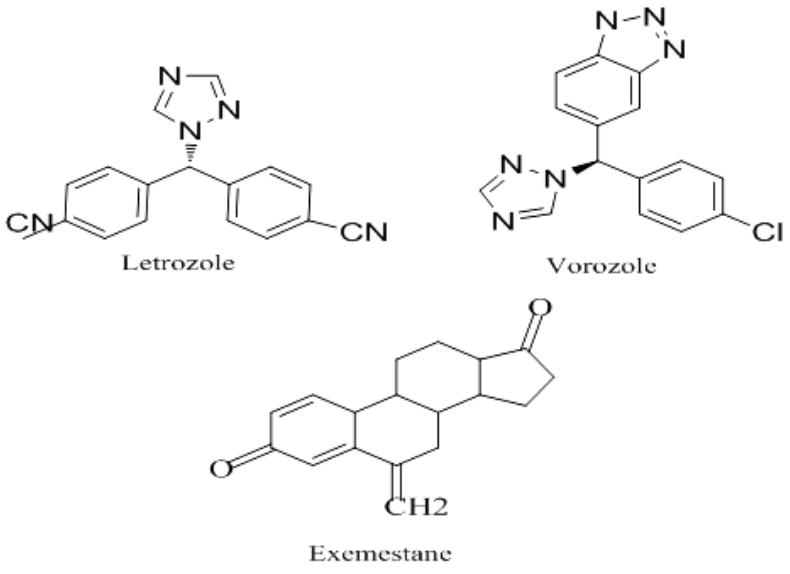


Fig. 1 – Structures of aromatase inhibitors.

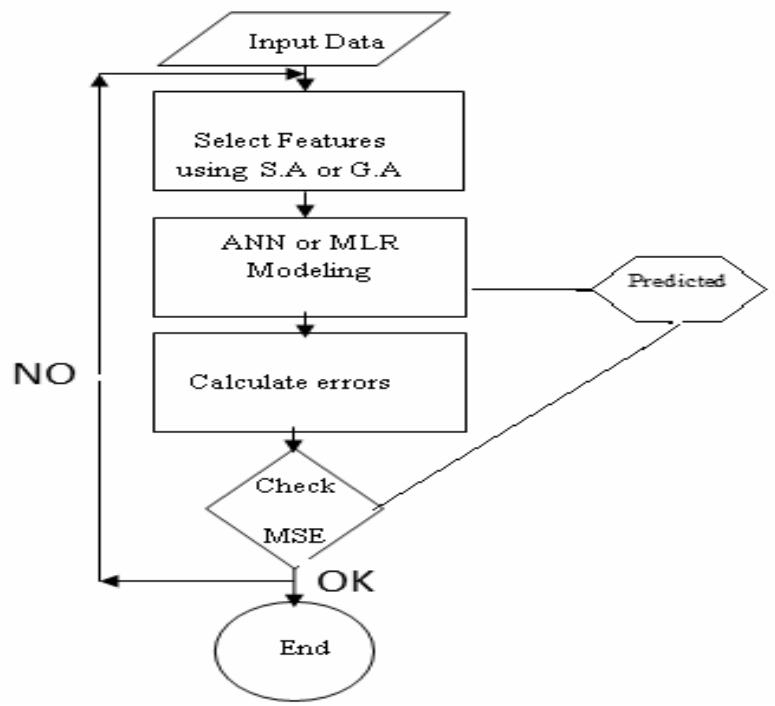
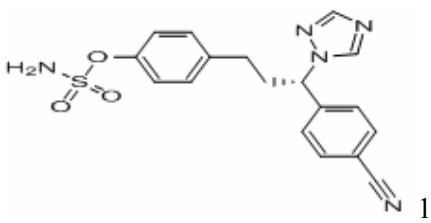
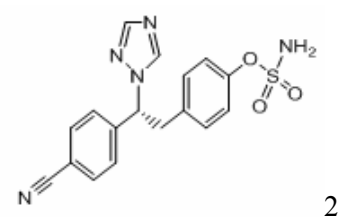


Fig. 2 – The employed procedure for finding optimum descriptors of the nonlinear models.



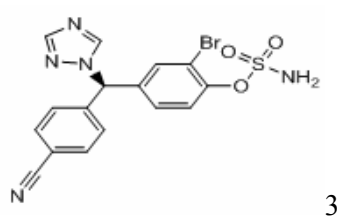
1

fundamental vibration:11.7



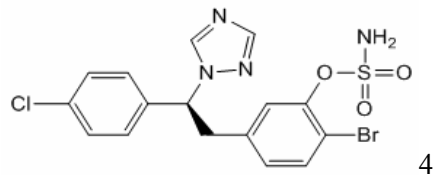
2

Fundamental vibration:15. 1



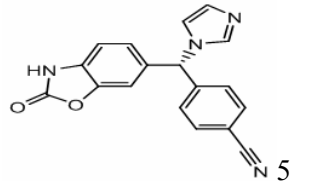
3

Fundamental vibration:10.9



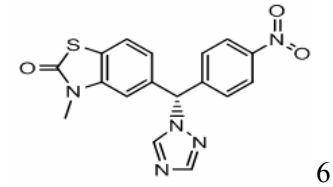
4

Fundamental vibration:10.8



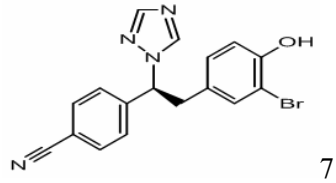
5

Fundamental vibration:16.2



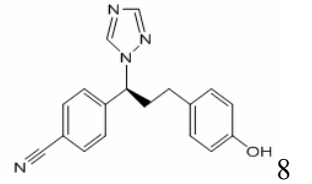
6

Fundamental vibration:12.3



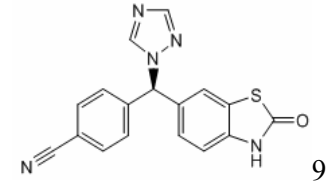
7

Fundamental vibration:12.1



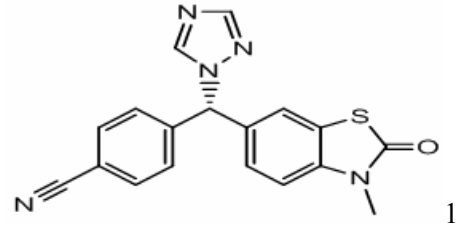
8

Fundamental vibration:13.3



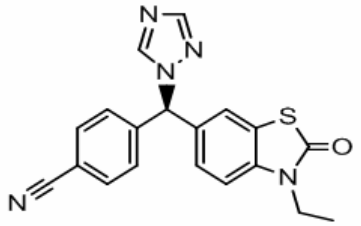
9

Fundamental vibration:18.3



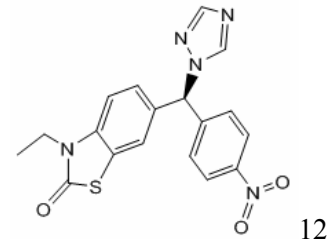
10

Fundamental vibration:14.9



11

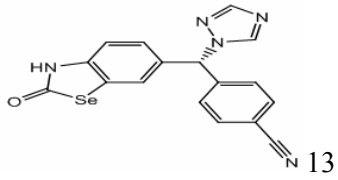
Fundamental vibration:15.8

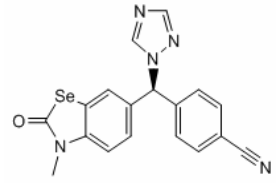


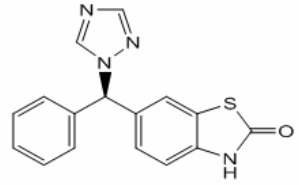
12

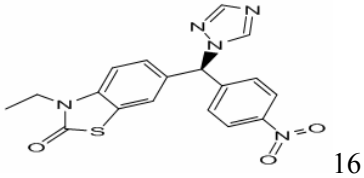
Fundamental vibration:13.7

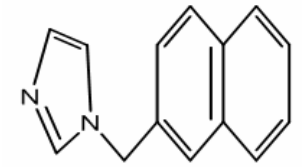
Fig. 3 – Optimized structure of the Letrozole compounds used to build QSAR models with B3lyp/6-311g.

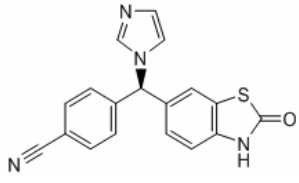

13
Fundamental vibration:11.3

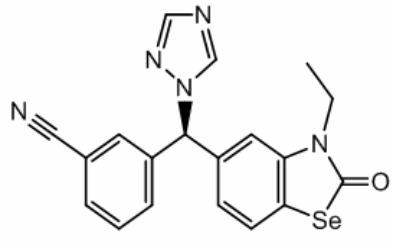

14
Fundamental vibration:14.5

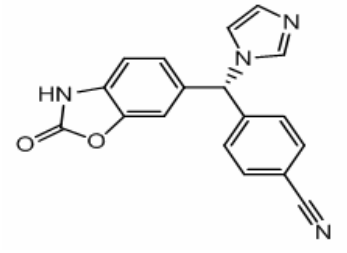

15
Fundamental vibration:14.7

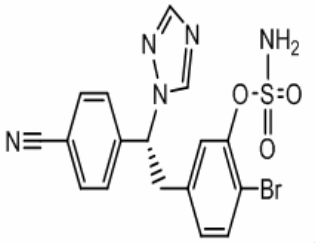

16
Fundamental vibration:13.2

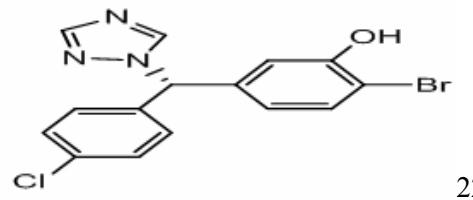

17
Fundamental vibration:16.7

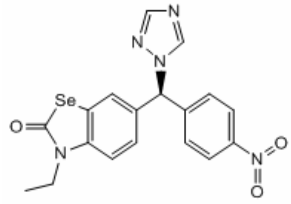

18
Fundamental vibration:18.2

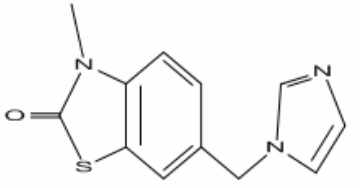

19
Fundamental vibration:9.5

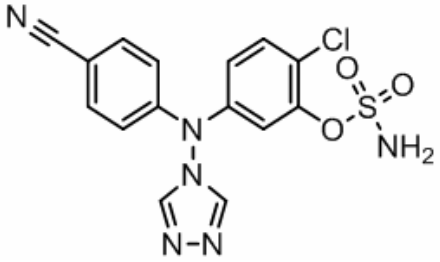

20
Fundamental vibration:16.2

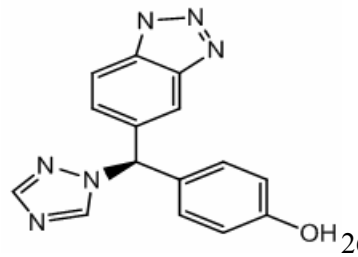

21
Fundamental vibration:12.7


22
Fundamental vibration:13.9


23
Fundamental vibration:11.1


24
Fundamental vibration:13.3


25
Fundamental vibration:12.5


26
Fundamental vibration:11.1

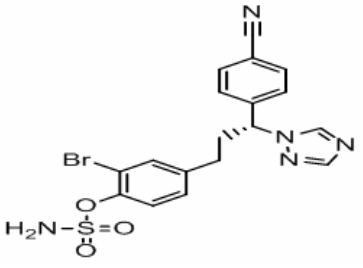

27
Fundamental vibration:10.0

Fig. 3 – Optimized structure of the Letrozole compounds used to build QSAR models with B3lyp/6-311g (continued).

Table 1

Statistical parameters of different linear QSAR models

QSAR Model	RMSE	R^2
MLR-MLR	0.2188	0.8261
MLR-PLS1	0.2189	0.8260
MLR-PCR	0.2874	0.7002

The 1350 descriptors were fed to the MLR-SA, SA-ANN, GA-ANN and MLR-GA models and then the best descriptors were selected (Tables 4, 5). In non-linear methods, 80%, 10% and 10% of data sets were randomly chosen as training, validation and test sets, respectively.

Table 2

Statistical parameters of different non-linear QSAR models

QSAR Models	Predicted		Train	
	R^2	RMSE	R^2	RSME
MLR-SA	0.834	0.0476	0.8394	0.03201
SA-ANN	0.8862	0.0318	0.8687	0.036599
MLR-GA	0.8828	0.0336	0.8813	0.04133
GA-ANN	0.9594	0.0114	0.9604	0.012712

Table 3

Observed and predicted values of –logIC50 by using GA-ANN model

Compound	Observed	Predicted
1	0.658	-0.6320
2	-0.447	-0.49038
3	-0.447	-0.52055
4	-0.146	-0.16584
5	-1.114	-0.99442
6	-0.477	-0.52437
7	-0.678	-0.67593
8	1.000	0.981043
9	-0.875	-0.85587
10	-0.954	-0.86862
11	-0.653	-0.64476
12	-0.602	-0.64177
13	-0.813	-0.88422
14	-0.653	-0.64731
15	-1.141	--0.69689
16	-0.954	-0.92591
17	0.796	-0.771494
18	-0.602	-0.84529
19	-0.602	-0.57834
20	-0.778	-0.79343
21	0.092	0.062941
22	-0.362	-0.37771
23	-0.699	-0.688
24	-0.699	-0.688
25	-0.362	-0.46616
26	-0.580	-0.58776
27	-0.921	-0.9265

Table 4

Selected descriptors using QSAR Methods

MLR-MLR	SA-ANN	MLR-SA	MLR-GA	ANN-GA
BElv1	X3	ISH	Mor19p	R5m
MATS5m	BEHp5	R4v+	IDDM	BIC0
lp1	L2v	SRW04	WA	nDB
G3p	HATs3v	Espm11d	Mor02e	EEig12r
Mor21m	Mor06v	Mor16v	CIC2	Mor14e
H6m	RDF035m	GGI9	H6u	Mor26v
BElv5	EEig05r	Espm01x	RDF080e	EEig14r

Table 5

Definition of the selected descriptors using QSAR Methods

Descriptor	Definition	Type
BElv1	Lowest eigenvalue n.1 of Burden matrix/Weighted by atomic van der waals volumes	Burden eigenvalues
MATS5m	Moran autocorrelation-lag 5/Weighted by atomic masses	2D autocorrelations
lp1	Lovasz-pelikan index/leading eigenvalue	Eigen value-based indices

Table 5 (continued)

G3p	3rd component symmetry directional WHIM index/Weighted by atomic polarizabilities	WHIM descriptors
Mor21m	3-D-MorsE-Signal 21/weighted by atomic masses	3D-MoRSE descriptors
H6m	H autocorrelation of lag 6/Weighted by atomic masses	GETAWAY descriptors
BElv5	Lowest eigenvalue n.5 of Burden matrix/Weighted by atomic van der Waals volumes	Burden eigenvalues
X3	Connectivity index chi-3	Connectivity indices
BEHp5	Highest eigenvalue n.5 of Burden matrix/Weighted by atomic polarizabilities	Burden eigenvalues
L2v	2nd component size directional WHIM index/Weighted by atomic van der Waals volumes	WHIM descriptors
HATs3v	Leverage-Weighted autocorrelation of lag 3/Weighted by atomic van der Waals volumes	GETAWAY descriptors
Mor06v	3D-MoRSE-signal 06/Weighted by atomic van der Waals volumes	3D-MoRSE descriptors
RDF035m	Radial Distribution Function 3.5/Weighted by atomic Sanderson electronegativities	RDF descriptors
EEig05r	Eigenvalue 05 from edge adj.matrix weighted by resonance integrals	Edge adjacency indices
ISH	Standardized information content on the leverage equality	GETAWAY descriptors
R4v+	R maximal autocorrelation of lag4/weighte by atomic van der Waals volumes	GETAWAY descriptors
SRW04	Self-returning walk count of order 04	Walk and path count
Espm11d	Spectra moment 11 from edge adj-matrix weighted by dipole moments	Edge adjacency indices
Mor16v	3D-MoRSE-signal 16/Weighted by atomic van der Waals volumes	3D-MoRSE descriptors
GGI9	Topological charge index of order 9	Topological charge indices
Espm01x	Spectra moment 01 from edge adj-matrix weighted by edge degrees	Edge adjacency indices
Mor19p	3D-MoRSE-Signal 19/Weighted by atomic Polarizabilities/3D-MoRSE descriptors	3D-MoRSE descriptors
IDDM	Mean information content on the distance degree magnitude	Information indices
WA	Mean wiener index	Topological descriptors
Mor02e	3D-MoRSE-Signal 02/Weighted by atomic Sanderson electronegativity's	3D-MoRSE descriptors
CIC2	Complementary information content (neighborhood symmetry of 2-order	Information indices
H6u	H autocorrelation of lag 6/unweighted	GETAWAY descriptors
RDF080e	Radial Distribution Function-8.0/Weighted by atomic Sanderson electronegativity's	RDF descriptors
R5m	R autocorrelation of lag 5/Weighted by atomic masses	GETAWAY descriptors
BIC0	Bond information content (neighborhood symmetry of 0-order	Information indices
nDB	Number of double bonds	Constitutional descriptors
EEig12r	Eigenvalue 12 from edge adj. matrix weighted by resonance integrals	Edge adjacency indices
Mor14e	3D-MoRSE-SIGNAL 14/Weighted by atomic Sanderson electronegativities	3D-MoRSE descriptors
Mor26v	3D-MoRSE-signal 26/Weighted by atomic van der Waals volumes	3D-MoRSE descriptors
EEig14r	Eigenvalue 4 from edge adj. matrix weighted by resonance integrals	Edge adjacency indices

H6m, R5m, H6u, HATS3v, ISH, R4v+ (Table 5) are GETAWAY(Geometry, Topology, and Atom-Weights Assembly) descriptors. These descriptors [27, 29, 17] encoded the geometrical information obtained from the molecular matrix, the topological information obtained from the molecular graph and the information obtained from atomic weights which were specially designed with the aim of matching the 3D-molecular geometry.

IDDM, CIC2, BIC0 (Table 5) are information indices descriptors. The total information content (I) was obtained by multiplying the mean information content by the number of elements.[39]

MATS1m (Table 5) are 2D-autocorrelation descriptors that represent the topological structure of the compounds in nature and are more complex than the classical topological descriptors.[27]

G3p, L2v (Table 5) are WHIM descriptors. The relevant molecular 3D information regarding molecular size, shape, and symmetry and atom distribution with respect to invariant reference frames[27] were built by using WHIM descriptors.

Mor21m, Mor06v, Mor02e, Mor19p, Mor 14e, Mor 26r and Mor 16v (Table 5) are 3D-MoRSE descriptors from which 3D-MoRSE descriptors were obtained through the molecular transformation employed in electron diffraction studies.[36]

BELv1, BELv5, BEHP5 (Table 5) are Burden eigenvalues. Lp1 (Table 5) is eigenvalue-based

indices and EEig05r, EEig14r, EEig12r, Espm11d, Espm01× (Table 5) are edge adjacency indices.

In order to define a new topographic index the edge adjacency relationships were used in molecular graphs. Molecules such as weighted graphs were used for calculation of the novel index, in which the elements of edges set were substituted by the bond orders between connected atoms in the molecule.

The highest eigenvalue of burden matrix was the other group of descriptors which was defined as eigenvalues of Burdun matrix (B). The number of atoms, bond order between two atoms or the electronegativity of the atoms was defined with the B matrix.

RDF035m and RDF080e (Table 5) are RDF descriptors[27,28] and are based on the distance distribution in the molecule[38].

X3 (Table 5) is connectivity index which is used for molecular structure quantitation in which weighted counts of substructure fragments and structural features, such as size, branching, unsaturation, heteroatom content and cyclicity are encoded.[37]

WA (Table 5) is a topological descriptor and nDB (Table 5) is the constitutional descriptor which is equal to the number of non-aromatic double bonds.[27]

GGI9 (Table 5) is the topology charge index that has been proposed to evaluate the charge transfer between pairs of atoms and the global charge transfer in the molecule.[27]

SRW04 (Table 5) is walk and path counts descriptor while self-returning walk count of *k*th order (SRWk) is the total number of self-returning walks of length *k* in the graph and is calculated as follows:

$$SRW k = tr(A^k) \quad (1)$$

where tr is the trace operator (sum of the diagonal elements) and A^k is the *k*th power of the adjacency matrix.[27]

Table 2 shows that RMSE and R^2 for predicted activity in GA-ANN was found to be 0.0216 and 0.9252, respectively. Therefore, GA-ANN model was better than the other nonlinear models. Seven descriptors were chosen from the parameters selected by GA-ANN model. The descriptors selected by the GA-ANN model were employed to build the final model (Table 4) and the selected descriptors presented in Table 4. The observed and predicted values of -logIC50, using GA-ANN method, are shown in Table 3.

The plot showing the variation of observed *versus* predicted -log IC50 values, using the GA-ANN model, is depicted in Figure 4.

Plots of the EEig14r, R5m, BIC0, nDB, Mor14e, EEig12r, Mor26v descriptors *versus* –logIC50 experimental were plotted by using Matlab program (Figure 5).

The charts showed that an increase in EEig14r, R5m descriptors resulted in a corresponding increase in the amount of –log IC50 experimental (R). The –log IC50 experimental (R) reduced when BIC0, nDB, EEig12r and Mor26v descriptors were increased. By increasing the nDB descriptor to values higher than 2, there was a slight increase in the amount of –log IC50 experimental (R) and this increase was almost constant.

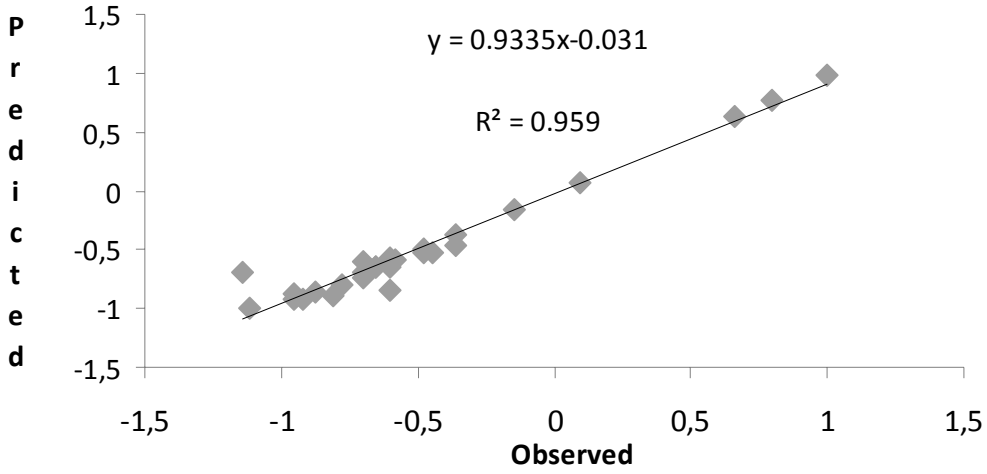


Fig. 4 – Observed *vs*. predicted values of -log IC50 by using GA-ANN method.

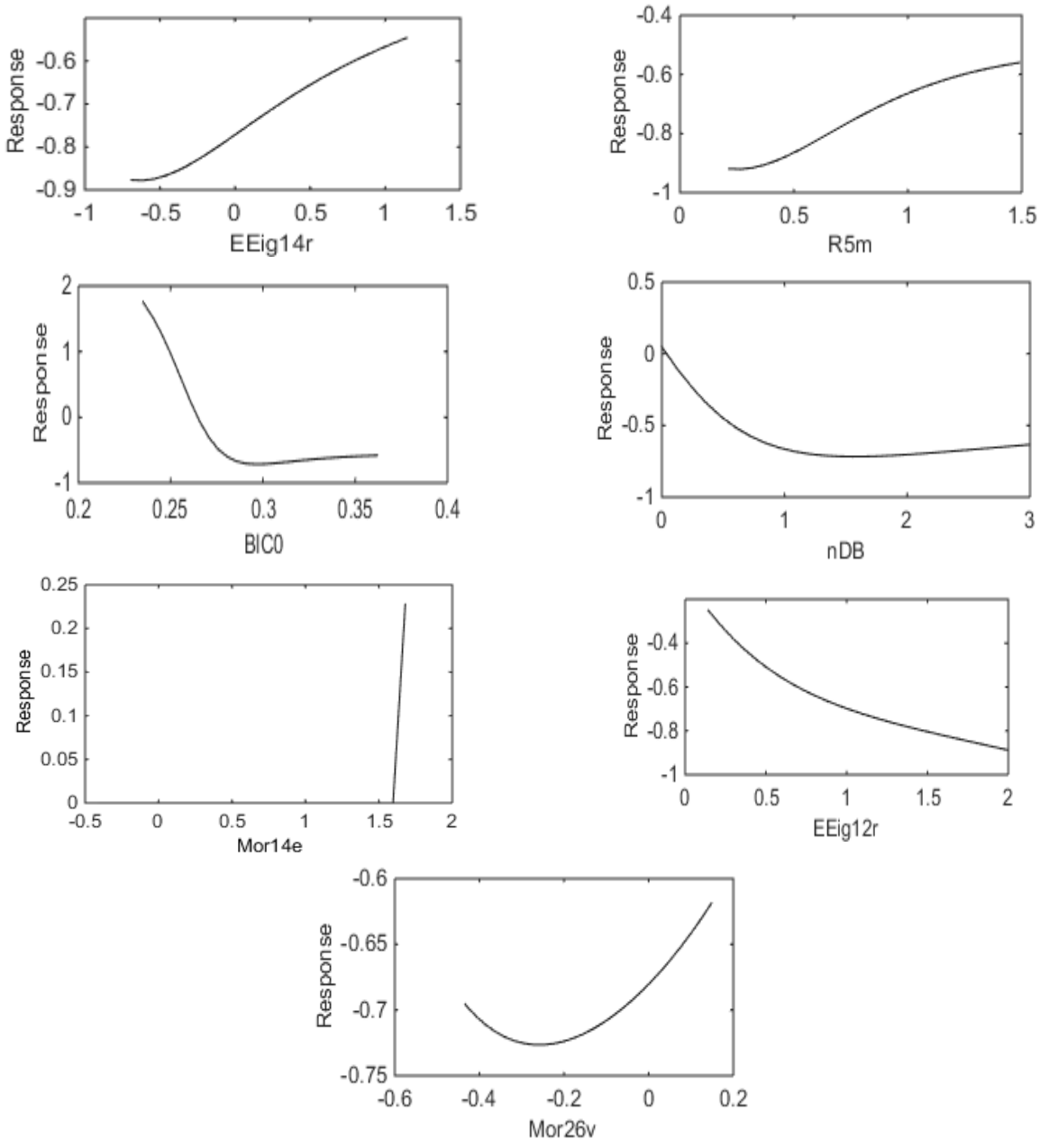


Fig. 5 – Experimental values of -log IC50 *versus* descriptors EEigl4r, R5m, BIC0, nDB, Morl4e, EEigl2r, Mor26v descriptors.

With increasing amounts of -0.2 to a higher value in Mor26v descriptor, an increase was seen in –log IC50 experimental (R). With increasing amounts of Mor14e descriptor to 1.5 no changes were observed in –log IC50 experimental (R). In the case of the Mor14e descriptor, with an amount of 1.5, there was an increase in bar chart in the –log IC50 experimental (R). Therefore, it is evident that the process is independent of Mor14e.

Table 3 showed that compounds No. 8, 17 had high empirical negative logarithm half maximal inhibitory concentration and low empirical half maximal inhibitory concentration (IC_{50}) and can be the best drugs in this study. Descriptors were selected using GA–ANN model as shown in Table 4 and were employed to build the final model.

The graphs of EEigl4r, R5m, BIC0, nDB, Morl4e, EEigl2r, Mor26v descriptors in minimum and maximum values of compounds No. 8, 17 were plotted using excel program (Figure 6).

It was shown that in compound No.8, values for BIC0, nDB descriptors were at the minimum and for EEigl4r, R5m, BIC0, nDB, Morl4e, EEigl2r descriptors the values were at the maximum and for Mor26v descriptor, the value was intermediate between maximum and minimum values.

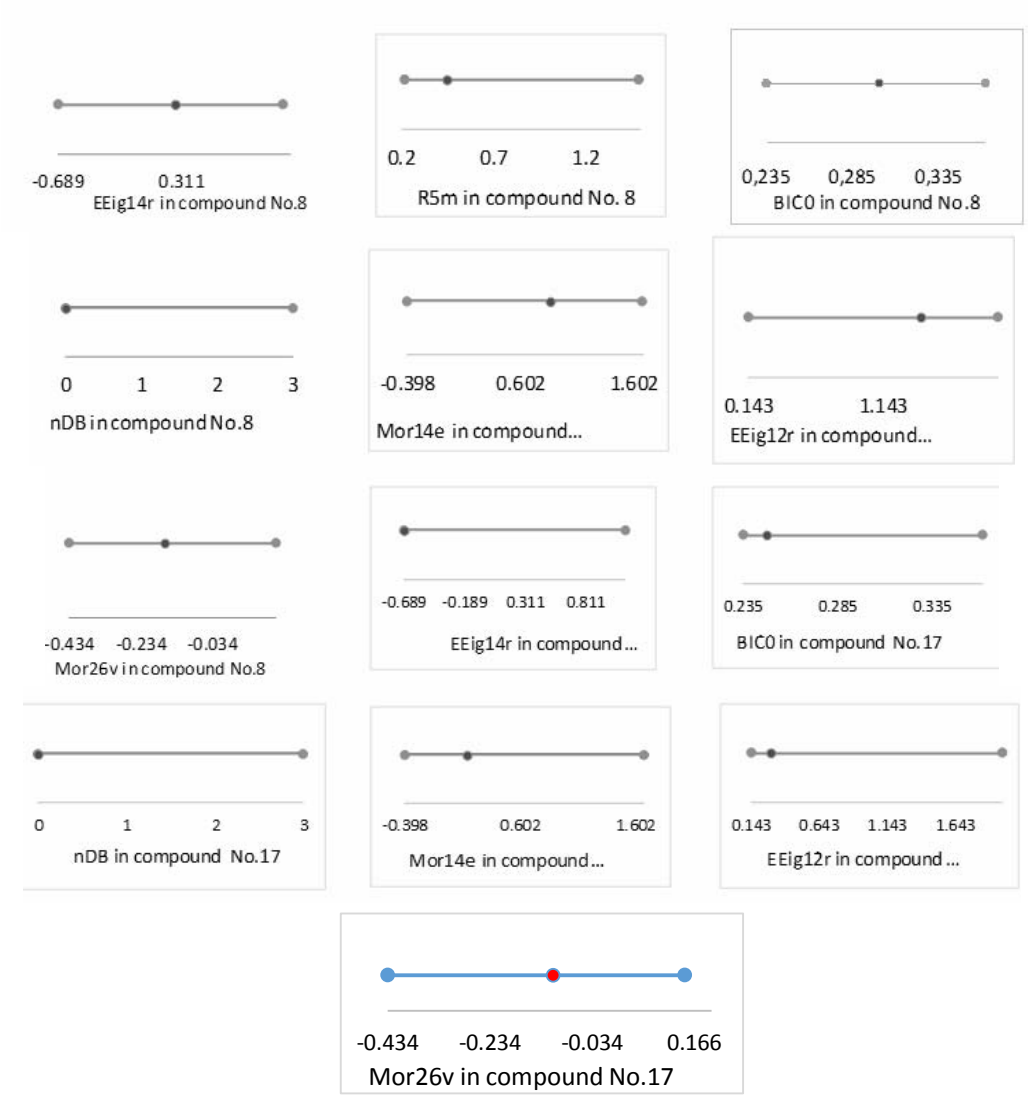


Fig. 6 – Plot EEigl4r, R5m, BIC0, nDB, Morl4e, EEigl2r, Mor26v descriptors in compounds No.8, 17 in GA-ANN method.

However, in compound No. 17, values for EEigl4r, R5m, BIC0, nDB, Morl4e, EEigl2r descriptors were at the minimum and for Mor26v descriptor, the value was intermediate between maximum and minimum values. In compounds No. 17, descriptors have similar variation and it is the best drug in the current study.

Thus this work predicts a new design for this class of drugs and the EEigl4r, R5m, BIC0, nDB, Morl4e, EEigl2r descriptors values are minimum.

CONCLUSIONS

The obtained results from QAR models showed that GA-ANN combination was better than the other models used and also proved that EEig14r, R5m, BIC0, nDB, Mor14e, EEig12r, Mor26v descriptors were more significant than other descriptors in building this QSAR model and predicting biological activity of letrozole substitution patterns.

Acknowledgement: The authors gratefully acknowledge the support provided by the Islamic Azad University of Rasht.

REFERENCES

1. W. R. Miller and S. P. Langdan, Ear. J. Surg. Oncol., **1977**, *23*, 163**.**
2. W. R. Miller, R. A. Hawkins and A. P.Forrest, *Cancer Res*.,**1982**, *42 (8 suppl*), 3365s.
3. L. R. Nelson and S. E. Bulun, *J. Am. Acad. Dermatol.*, **2001**, *45 (3 suppl*), S116.
4. A. M. H. Brodie, W .C. Schwarzel, A. A. Shaikh and H. J. Brodie, *Endocrinology*, **1977**,*100*, 1684.
5. M. Numazawa, T. Sugiyama and M. Nagakama, *Biol. Pharm. Bull*., **1998**, *21*, 289 .

6. N. Zilembo, C. Noberasco and E. Br. Bajetta, *J. Cancer*, **1995**, *72*, 1007.
7. A. Hamilton and M. Piccart, *Ann. Oncol*., **1999**, *10*, 377.
8. A. U. Buzdar, C. L. Jones, J. Vogel, P. Wolter and A. Plourde, *Cancer*, **1997**, *38*, *301*.
9. J. N. Ingle, P. A. Johnoson, V. J. Suman, J. B. Gerstner, J. A. Mailliard, J. K. Camoriano, Jr .D. H. Gesme, C. L. Loprinzi, A. K. Hatfield and L. C. Hartmann, Cancer, 1997, 80, 218.
10. N. Murthy, A. Raghuram and G. Rao, *Curr. Med. Chem. Anticancer Agents*, **2004**, *4*, 395.
11. H. Schmidi, *Chemom. Intell. Lab. Sys.*, **1997**, *37*, *125*.
12. C. Hansch , A. Kurup, R. Garg, H. Gao, *Chem. Rev.*, **2001**, *101*, 619.
13. S. Wold, J. Trygg, A. Berglund and H. Antii, *Chemom. Intell. Lab. Syst*., **2001**, *58*,131.
14. D. Horvath and B. Mao, *QSAR. Comb. Sci.*, **2003**, *22*, 498.
15. S. Pitta, J. Eksterowicz, C. Lemmen and R. Stanton, *J. Chem. Inf. Comput. Sci.*, **2003**, *43*, 1623.
16. S. Gupta, M. Singh and A. K. Madan, *J. Chem. Inf. Comput. Sci.*, **1999**, *39*, 272.
17. V. Consonni, R. Todeschini and M. Pavan, *J. Chem. Inf. Comput. Sci.*, **2002**, *42*, 693.
18. D. A. Winkler, *Briefings in Bioinformatics*, **2002**, *3*, 73.
19. R. Gotha, J. R. Serra and P. C. Jurs, *J. Mol. Graph. Model.*, **2004**, *23*, 1.
20. SPSS, Version 19; (2010**)**; available at http://www.spssscience.com.
21. P. E. Goss*, Breast Cancer Res. Treat.*, **1998**, *49 (Suppl 1)*, S59–65; discussion S73–7.
22. S. Kirkpatrick, Jr. C. D. Gelatt and M. P. Vecchi, *Science*, **1983**, *220*, 671.
23. V. Cerný, *J. Optimiz. Theory Applic.*, **1985**, *45*, 41.
24. E. B. De Melo and M. M. Ferreira, *Eur. J. Med. Chem*., **2009**, *44*, 3577.
25. R. Todechine, Milano Chemometrics, *QSAR Group*, http://www.disat.unimib.it/chem.
26. J. H. Schuur, P. Selzer and J. Gasteiger*, J. Chem. Inf. Comput. Sci.*, **1996**, *36*,334.
27. R. Todeschini and V. Consonni, "Handbook of Molecular Descriptors", Wiley-VCH, 2000.
28. M. C. Hemmer, V. Steinhauer and J. Gasteiger, *Vibr. Spectrosc.*, **1999**, *19*, 151.
29. V. Consonni, R. Todeschini and M. Pavan, *J. Chem. Inf. Comput. Sci.*, **2002***, 42*, 682.
30. P. Gramatica, V. Consonni and R. Todeschini, *Chemosphere*, **1999**, *38*,1371.
31. P. Gramatica, V. Consonni and R. Todeschini*, Chemosphere*, **2000**, *41*, 763.
32. M. H. Fatemi and S. Gharaghani, *Bioorg. Med. Chem.*, **2007**, *15*, 7746.
33. M. Jalali-Heravi and F. Parastar, *J. Chem. Inf. Comput. Sci*., **2000**, *40*, 147.
34. K. Levenberg, *Quarterly of Applied Mathematics*, **1944**, *2*, 164.
35. D. Horvatha and B. Mao, *QSAR. Comb. Sci.*, **2003**, *22*, 498.
36. J. H. Schurz, P. Selzer and J. Gasteiger, *J.Chem.Inform.Comput.Sci*., **1996**, *36*, 334.
37. L. H. Hall and L. B. Kier, *Reviews of Computational Chemistry*, **1991**, *2*, 367.
38. www.strandls.com/sarchitect/.../desctheory
39. J. H. Schuur, P. Selzer and J. Gasteiger, *J. Chem. Inform. Comput. Sci*., **1996**, *36*, 334.
40. E. B. De Melo and M. M. Ferreira, *Eur. J. Med.Chem.*, **2009**, *44*, 3577.