INTERNATIONAL RESEARCH JOURNAL OF PHARMACY
www.irjponline.com ISSN 2230 – 8407
Research Article

QSAR MODELING OF GELATINASE INHIBITORS: MLR APPROACH

Ajeet[1]*, Kumar Vipul[2], Singh Brajpal[1]
[1]S. D. College of Pharmacy and Vocational Studies, Muzaffarnagar, U.P., India
[2]Department of Pharmacy, Meerut Institute of Engineering and Technology, Meerut, U.P., India

Article Received on: 15/01/12 Revised on: 28/02/12 Approved for publication: 11/03/12

*E-mail: ajeet_pharma111@rediffmail.com

ABSTRACT
Matrix metalloproteinases are regulated by growth factors, hormones, cytokines etc. and endogenous inhibitors control them. Over expression of matrix metalloproteinases can lead to a variety of pathological disorders which results because of imbalance between the activity of matrix metalloproteinase inhibitors and tissue inhibitors of matrix metalloproteinase. So, it is the necessity to develop the Phenylalanine analogues with impact inhibitive concentration. Here Phenylalanine analogues have been used to correlate the inhibiting activity with the Eccentric Connectivity index (ECI), Fragment Complexity (FC) and Topological Polar Surface Area (TPSA) for studying the Quantitative Structure Activity Relationship (QSAR). Correlation may be an adequate predictive model which can help to provide guidance in designing and subsequently yielding greatly specific compounds that may have reduced side effects and improved pharmacological activities. We have used Multiple Linear Regression (MLR), one of the best methods for developing the QSAR model. Results from this QSAR study have suggested that ECI, FC and TPSA are the important descriptors for inhibitory activities of endopeptidase inhibitors. For the validation of the developed QSAR model, statistical analysis such as data point-descriptor ratio, fraction of variance, cross validation test, standard deviation, quality factor, fischers test; and internal validation such as Y-randomization test have been performed and all the tests validated this QSAR model.
KEYWORDS - Eccentric Connectivity index, Fragment Complexity, Topological Polar Surface Area, endopeptidase inhibitor-Gelatinase inhibitors, QSAR, Multiple Linear Regression.

INTRODUCTION

Matrix metalloproteinase is a family of calcium-dependent zinc-containing endopeptidase, which is responsible for the remodeling of tissue and extracellular matrix degradation, including gelatin, matrix glycoprotein, proteoglycan etc. These are excreted by a variety of connective tissue and some pro-inflammatory cells like lymphocytes, osteoblasts, macrophages, endothelial cells, neutrophils, and fibroblasts. Matrix metalloproteinases have been considered as a target for cancer therapy and a number of natural and synthetic Matrix metalloproteinase inhibitors have been identified as cytostatic and anti-angiogenic agents. Matrix metalloproteinase consist of four domains, these are hinge region, N-terminal prodomain, catalytic domain, and C-terminal hemopexin like domain. It may be responsible for interaction with tissue inhibitors of Matrix metalloproteinase as well as for the macromolecular substrate recognition[1].

In the present study, we developed a QSAR model on a series of Phenylalanine analogues with respect to their inhibitory activity towards gelatinase inhibition. The QSAR studies are perfect tool for understanding the drug design process in terms of their chemical-pharmacological activity interaction, along with it is also used in toxicology and pesticide research. QSAR studies can focus on mechanism of action of ligands with human, bacteria, virus, membranes, enzymes etc. It can also be used for the evaluation of the metabolism, absorption, distribution and excretion phenomena. The QSAR methodology comprises of computationally derived descriptors to correlate with pharmacological activities. These descriptors are principally of four types such as electronic, stearic, hydrophobic and topological indices[2]. The descriptors used by us for developing the QSAR model are Eccentric Connectivity Index (ECI)[3], Fragment complexity (FC)[4] and Topological Polar Surface Area (TPSA)[5].

MATERIALS AND METHOD

All the bioactivity values and information about 2D structure of Phenylalanine derivatives were taken from literature[1]. IC_{50} is referred as the molar concentration of a compound that inhibits 50% growth of bacteria[2]; $\log 1/IC_{50}$ is subsequent variable that comprises the bioactivity parameter for the QSAR model. In order to calculate the 2D molecular descriptors, PaDEL descriptor software[6] which incorporate CDK library for descriptor calculation have been used. For the development of QSAR model, Multiple Linear Regression has been employed[2].

Molecular Descriptors

Eccentric connectivity index denoted by ξ is defined as the summation of the product of eccentricity and the degree of each vertex in the hydrogen suppressed molecular graph having n vertices.

$\xi = \sum(E_i - V_i)$

Where V_i is the degree of vertex i, E_i is the eccentricity of the vertex i and n is the number of vertices in graph G. The eccentricity E_i of a vertex i in a graph G is the path length from vertex i to vertex j that is farthest from I (E_i = max d (ij); j ϵ G); Eccentric connectivity index takes into consideration the eccentricity as well as valency of the vertices in a hydrogen suppressed graph[3].

Topological Polar Surface Area is calculated from molecular bonding information only. Hou et. al. has used the solvent-accessible surface area for calculating the topological polar surface area. Different procedures of surface calculations, even different calculation parameters may generate different TPSA. It may be calculated as-

$TPSA = \sum n_i s_i$

Where n_i is the frequency of fragment i in the molecule, s_i is the surface contribution of type i[4].

Statistical Parameters

In the QSAR model, number of data points is denoted as n, squared correlation coefficient as r^2 (fraction of variance), cross-validated r^2 is denoted as q^2, s is standard deviation. Q is quality factor, where Q= r/s (here r is correlation coefficient and s is standard deviation). Fischer statistics is denoted by F.

Model Validation

The QSAR model validation was carried with statistical analysis and with internal validation.

RESULT

The 2D structure of Phenylalanine pharmacophore for which the QSAR model has been developed is shown in Figure 1

From the data in Table 1, QSAR equation have been developed where number of data point (n) is 14, is given below, here 95% confidence intervals are given in parantheses.

$\log(1/IC_{50})$=26.25678(±3.961666)+0.0092199(±0.0069291) (ECI)-0.0024298(±0.0014742)(FC)-0.1690361(±0.0376384) (TPSA)

Validation Of QSAR Model

A quantitative assessment of model robustness has been performed through model validation. All the statistical results of model validation have been given in Table 2.

Statistical Analysis

1. **n/p ratio**: $\frac{n}{p} \geq 4$, where n is the number of data points and p is the number of descriptors used in the QSAR model. The model obeys the condition.
2. **Fraction of variance (r^2):** The value of fraction of variance may vary between 0 (means model without explanatory power) and 1 (means perfect model). QSAR model having $r^2 > 0.6$ will only be considered for validation[2]. The value for this QSAR model is 0.9265.
3. **Cross-Validation Test (q^2):** A QSAR model must have $q^2 > 0.5$ for the predictive ability[2]. The value of q^2 for this QSAR model is 0.9264.
4. **Standard deviation (s):** The smaller s value is always required for the predictive QSAR model. The value of s for this QSAR model is 1.16.
5. **r^2-q^2 < 0.3:** The difference between r^2 and q^2 should never be exceed by 0.3. A large difference suggests the following: presence of outliers, over-fitted model, and presence of irrelevant variables in data[2]. The value of r^2-q^2 for this QSAR model is 0.0001.
6. **Quality Factor (Q):** Over fitting and chance correlation, due to excess number of descriptors, can be detected by Q value. Positive value for this QSAR model suggests its high predictive power and lack of over fitting.
7. **Fischer Statistics (F):** The F value of QSAR model was compared with their literature value at 95% level. The F value of this QSAR model is 24.3786 (where $F > F_{lit}$) suggests that the QSAR model is statistically significant at 95% level.

Internal Validation

Y-Randomization Test*:* To establish the QSAR model robustness, this technique is being used widely. For this test, the dependent variable vector is randomly shuffled, and a new QSAR model is developed using the unchanged independent variable. This process was repeated for five times. The statistical data of r^2 for five runs are given in Table 3. The values $r^2<0.6$ in Y-randomization test confirm the robustness of this QSAR model[2].

DISCUSSION

According to the developed QSAR model, the Phenylalanine derivatives must have more Eccentric Connectivity Index for enhanced endopeptidase inhibitory action at X and Y subtituents. A negative coefficient of Fragment Complexity containing X and Y substituents decreases the activity of Phenylalanine derivatives towards its inhibitory action of endopeptidase. Moving towards the effects of the Topological Polar Surface Area on the bioactivity of derivatives of Phenylalanine, the developed QSAR model suggest that an decrement in TPSA at subtituents X and Y will definitely be favourable to the activity, as discussed by R. P. Verma and Corwin Hansch[2]. A comparison (multiple linear regression plots) of observed values and predicted values of $\log(1/IC_{50})$ for Phenylalanine derivatives used for development of QSAR equation is shown in Figure 2 and multiple linear graph is shown in Figure 3.

REFERENCES

1. Verma Rajeshwar P, Hansch Corwin. Matrix metalloproteinases (MMPs): Chemical–biological functions and (Q)SARs. Bioorg Med Chem 2007; 15: 2223-2268.
2. Verma Rajeshwar P, Hansch Corwin. QSAR modeling of taxane analogues against colon cancer. Eur J Med Chem 2010; 45: 1470-1477.
3. Sardana S, Madan AK. Topological models for prediction of antihypertensive activity of substituted benzylimidazoles. J Mol Struct (Theochem) 2003; 638: 41-49.
4. Gregg Siegal, Eiso AB, Jan Schultz. Integration of fragment screening and library design. Drug Discovery Today 2007; 12(23/24) 1032-1039.
5. Hou TJ, Xu XJ. ADME Evaluation in Drug Discovery. 3. Modeling Blood-Brain Barrier Partitioning Using Simple Molecular Descriptors. J Chem Inf Comput Sci 2003; 43: 2137-2152.
6. Yap CW. PaDEL-descriptor: open source software to calculate molecular descriptors and fingerprints. J Comput Chem 2011; 32(7): 1466-1474.

Table 1 Descriptors Used To Derive QSAR Equation Along With Bioactivities For The Inhibition Of Gelatinase By Phenylalanine Analogues With Substituent X and Y

No.	X	Y	log 1/IC_{50}			ECI	FC	TPSA
			Observed	Predicted	Residuals			
1	$(CH_2)_3C_6H_5$	CONHOH	7.82	7.75455	0.06545	1246	5699.06	95.5
2	C_8H_{17}	COOH	6.7	6.675272	0.024728	629	2965.07	107.53
3	C_9H_{19}	COOH	6.22	6.471816	-0.25182	691	3284.07	107.53
4	$C_{10}H_{21}$	COOH	6	6.358563	-0.35856	767	3619.07	107.53
5	$C_{12}H_{25}$	COOH	6.3	6.224874	0.075126	845	3970.07	107.53
6	$C_{15}H_{31}$	COOH	6.1	5.596395	0.503605	1189	5534.07	107.53
7	C_6H_{13}	CONHOH	8	7.993079	0.006921	710	2730.07	107.53
8	C_7H_{15}	CONHOH	9	9.107399	-0.1074	546	2486.06	95.5
9	C_8H_{17}	CONHOH	9.22	8.948677	0.271323	606	2779.06	95.5
10	C_9H_{19}	CONHOH	9	8.769518	0.230482	668	3088.06	95.5
11	$C_{10}H_{21}$	CONHOH	8.7	8.662123	0.037877	742	3413.06	95.5
12	$C_{12}H_{25}$	CONHOH	9	8.534291	0.465709	818	3754.06	95.5
13	$C_{14}H_{29}$	CONHOH	7.52	8.2542	-0.7342	980	4484.06	95.5
14	$C_{16}H_{33}$	CONHOH	7.7	7.929243	-0.22924	1154	5278.06	95.5

Table 2 Results Of Statistical Validation

n/p ≥ 4	r^2	q^2	s	$r^2 - q^2 < 0.3$	Q	F
4.66	0.9265	0.9264	1.16	0.0001	0.829741	33.6145

Table 3 Results Of Internal Validation: Y-Randomization Test (5 Runs)

No. of Y-randomization	First	Second	Third	Fourth	Fifth
r^2	0.275072	0.090387	0.307335	0.198623	0.187889

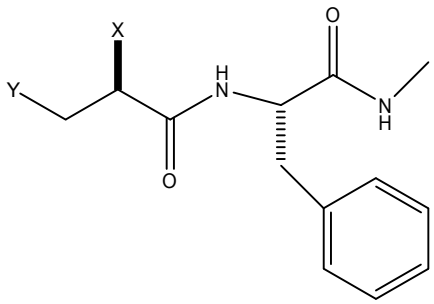

Figure 1 Pharmacophore of Phenylalanine used for model development

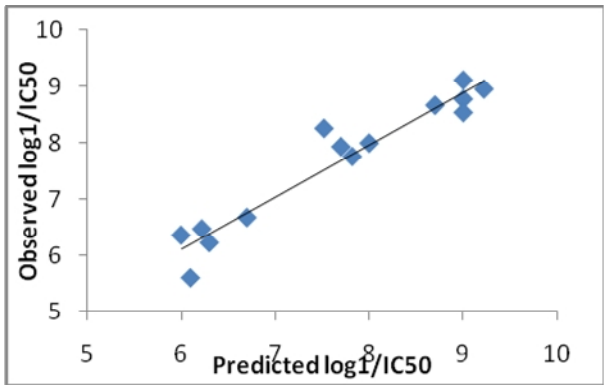

Figure 2 Multiple linear regression plot for QSAR study

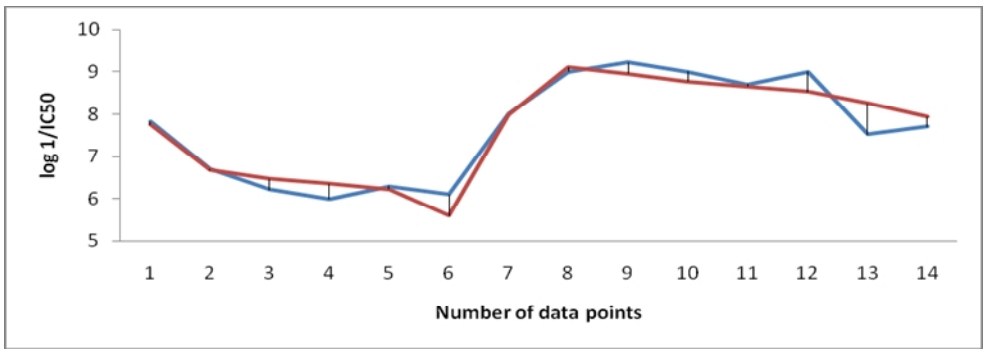

Figure 3 Multiple linear graph between No. of data points and bioactivities

Source of support: Nil, Conflict of interest: None Declared