

Quantitative Structure Activity Relationship Studies of Piperazinyl Phenylalanine Derivatives as VLA-4/VCAM-1 Inhibitors

Dinesh Bhargava, C. Karthikeyan, N.S.H.N. Moorthy* and Piyush Trivedi

School of Pharmaceutical Sciences, Rajiv Gandhi Proudyogiki Vishwavidyalaya, Airport Bypass Road, Gandhi Nagar, Bhopal-462036, Madhya Pradesh, India

Abstract: QSAR study was carried out for a series of piperazinyl phenylalanine derivatives exhibiting VLA-4/VCAM-1 inhibitory activity to find out the structural features responsible for the biological activity. The QSAR study was carried out on V-life Molecular Design Suite software and the derived best QSAR model by partial least square (forward) regression method showed 85.67% variation in biological activity. The statistically significant model with high correlation coefficient ($r^2=0.85$) was selected for further study and the resulted validation parameters of the model, cross validated correlation coefficient ($q^2=0.76$ and $\text{pred}_r^2=0.42$) show the model has good predictive ability. The model showed that the parameters SaaNEindex, SsClcount SlogP, and 4PathCount are highly correlated with VLA-4/VCAM-1 inhibitory activity of piperazinyl phenylalanine derivatives. The result of the study suggests that the chlorine atoms in the molecule and fourth order fragmentation patterns in the molecular skeleton favour VLA-4/VCAM-1 inhibition shown by the title compounds whereas lipophilicity and nitrogen bonded to aromatic bond are not conducive for VLA-4/VCAM-1 inhibitory activity.

Key Words: QSAR, Piperazinyl phenylalanine, V-life, VLA-4/VCAM-1.

INTRODUCTION

VLA-4 (very late activating antigen-4, CD49d/CD29) is a member of the integrin superfamily, composed of $\alpha 4$ and $\beta 1$ -heterodimers. It is mainly expressed on eosinophils and lymphocytes. VCAM-1 (vascular cell adhesion molecule-1), expressed in vascular endothelial cells and fibronectins, ingredients of extra-cellular matrix, is the ligand for VLA-4. Binding of VLA-4 and VCAM-1 is one of the most important mechanisms for tight adhesion of leukocytes with vascular endothelial cells and is strongly related to infiltration of inflammatory cells, such as lymphocytes and eosinophils into inflammation sites. As VLA-4 also functions as a sub-signal in the activation and multiplication of T-cells, agents that inhibit adhesion of VLA-4 and VCAM-1 interaction are likely to be useful in the treatment of inflammatory disease [1, 2].

AntiVLA-4 antibody has been demonstrated to be effective in various inflammation models of asthma [3], rheumatoid arthritis [4], multiple sclerosis (MS) [5] and inflammatory bowel diseases [6], with the humanized monoclonal antibody, Tysabri [7] (natalizumab), recently relaunched for the treatment of relapsing forms of multiple sclerosis but its marketing and clinical trials have been suspended due to cases of progressive multifocal leukoencephalopathy (PML). Therefore, non-peptide small molecule VLA-4 inhibitors are also an attractive target for the treatment of chronic inflammatory disease, with several reports on orally active inhibitors published in recent years. Preclinical tests of peptide

mimetic BIO-1272 (Biogen) derived from connecting segment 1 (CS-1) of fibronectins were performed for the first time as a small molecule inhibitor, showing a marked effect in inhibiting antigen induced respiratory hypersensitivity in sheep models of asthma [8]. Following compounds like Bio-1272, R-411, TR-14035, SB-683699, Merk-1, Merk-2, TBC-3342 and CT5219 are under clinical trials as VLA-4 inhibitory activity.

Quantitative structure activity relationship (QSAR) is one of the major tools in drug discovery to explore ligand-receptor/enzyme interactions, especially when the structural details of the target are not known. 2D-QSAR does not involve complex alignment or assumptions on conformations, therefore they can easily be applied to large compound sets, both in model building and in model application to new compounds. In such methods one has the choice among a wide variety of molecular descriptors independent on 3D conformation, for example, topological descriptors, simple molecular properties such as the molecular weight and easily calculated physicochemical properties such as ClogP or atomic partial charges. In any case, the resulting models are described in the form of equations and therefore are easily portable. Considering the relevance of piperazinyl phenylalanine as VLA-4 inhibitors, a 2D QSAR analysis was performed on the series of piperazinyl phenylalanine reported by Osamu Saku *et al.* [1]. Acylated L-phenylalanine derivatives have been reported as highly potent antagonists of VLA-4 binding to VCAM-1, such as SB-683698 (TR-14035) or R-411. The report of piperazinyl phenylalanine as VLA-4 antagonists shows that the optimization at the 4-position of piperazine produced highly potent compounds and also improves the pharmacokinetic behavior of the molecules [1, 9].

In this present research work, a series of 31 piperazinyl phenylalanine derivatives were subjected to QSAR analysis.

*Address correspondence to this author at the School of Pharmaceutical Sciences, Rajiv Gandhi Proudyogiki Vishwavidyalaya, Airport Bypass Road, Gandhi Nagar, Bhopal-462036, Madhya Pradesh, India; Tel: +91-755-2678883, Fax: +91-755-2742006; E-mail: nshnm06@yahoo.co.in

The goal of our research was to gain further insight into the structural features related to the VLA-4 inhibitory activity of the title compounds.

MATERIAL AND METHODS

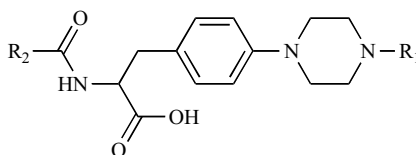
A training set of 21 piperazinyl phenylalanine exhibiting potent anti-inflammatory activity reported by Osamu Saku *et al.* was taken for the study [1]. The activity data given as IC_{50} values, where IC_{50} refers to the experimentally determined nanomolar concentration of the piperazinyl phenylalanine required to inhibit VLA-4 derived from Jurkat cells to compete for bound human VACM-1 by 50%. The biological activity values [IC_{50} (nM)] reported in literature were con-

verted to their molar units and then further to negative logarithmic scale and subsequently used as the dependent variable for the QSAR analysis. The $-\log$ values of IC_{50} along with the structure of the 31 compounds in the series is presented in Table 1.

COMPUTER SOFTWARE

All the computational studies were performed on a HP Compaq PC running on intel Pentium-D processor. The molecular structure of the training set and test were sketched using V-life MDS (Molecular Design Suite)TM 3.5 [10, 11] software supplied by V-life Sciences Technologies Pvt. Ltd., Pune, India 2006. Each compound was energy minimized

Table 1. Biological Activity Data and Structures of the Compounds in the Series Piperazinyl Phenylalanine [1]


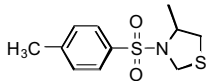

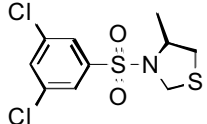

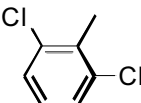

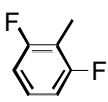

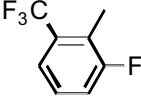

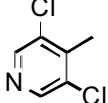

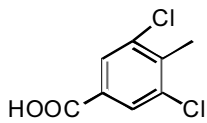


Compound	R ₁	R ₂	IC ₅₀ (nM/L)	-Log IC ₅₀
1	-C ₆ H ₅		123	7.91
2	-o C ₆ H ₄ OMe		135	7.87
3	-m C ₆ H ₄ OMe		170	7.77
4	-p C ₆ H ₄ OMe		152	7.82
5	-p C ₆ H ₄ Me		353	7.45
6	-o,m C ₆ H ₅ Me ₂		327	7.49
7	-p C ₆ H ₄ C ₆ H ₅		450	7.35
8	-p C ₆ H ₄ NO ₂		143	7.84
9	-m C ₆ H ₄ CN		99	8.00
10	-2 pyrimidine		239	7.62

(Table 1. Contd....)

Compound	R ₁	R ₂	IC ₅₀ (nM/L)	-Log IC ₅₀
11	-2 pyrazine		75	8.12
12	-2 pyridine		179	7.75
13	-Me		235	7.63
14	-CH ₂ C ₆ H ₅		222	7.65
15	-CH ₂ CH ₂ COOEt		130	7.89
16	-C ₆ H ₅		76	8.12
17	-o C ₆ H ₄ NO ₂		69	8.16
18	-o C ₆ H ₄ CN		67	8.17
19	-p C ₆ H ₄ CN		101	8.00
20	-o,o C ₆ H ₃ (CN) ₂		19	8.72
21			67	8.17
22	-2 pyrazine		124	7.91
23			80	8.10
24			200	7.70

(Table 1. Contd....)

Compound	R ₁	R ₂	IC ₅₀ (nM/L)	-Log IC ₅₀
25			196	7.71
26			28	8.55
27			1122	6.95
28			6807	6.17
29			4318	6.36
30			510	7.29
31			42	8.38

O, m, p are ortho, meta, para.

and batch optimized by using Merck Molecular Force Field (MMFF) fixing Root Mean Square Gradients (RMS) to 0.01 Kcal/mol Å. The optimized batch of molecules was selected for calculation of the physiochemical descriptors. The descriptor pool was reduced by eliminating out the descriptors with constant and near constant values. Further reduction in the descriptor pool was done by ousting the descriptors that are highly degenerate and difficult to interpret. A correlation analysis was performed between biological data and remaining descriptors, most of which were molecular and electrotopological descriptors and the descriptors those were showing very low correlation with inhibitory activity were also removed. The descriptors selected for modeling inhibitory activity of the piperazinyl phenylalanine derivatives are summarized in Table 2.

The random selection method was used for training data selection at 70% of total number of compounds and variable selection was performed by simulated annealing methodology. The QSAR model was generated by using partial least square (PLS) method by using V-life molecular design suite (MDS). The program computes the best model on the basis of squared correlation coefficient r^2 , crossed validated q^2 , F-test and pred_r^2 . The calculated value of F-test when compared with tabulated value of F-test shows the level of statistical significance (99.99%) of the QSAR model. The low

standard error of Pred_r^2 , q^2 , q^2_{se} and r^2_{se} shows absolute quality of fitness of the model. The correlation matrix was developed by using OPENSTAT software ver. March 2008 [12].

The generated QSAR models were validated for predictive ability inside the model by using cross validation (Jack-Knife method or leave one out) for q^2 and external validation, which is more robust alternative method by dividing the data into training set and test set and calculating pred_r^2 . The high pred_r^2 and low $\text{pred}_r^2_{se}$ were show high predictive ability of the model.

RESULT AND DISCUSSION

QSAR study of a series of piperazinyl phenylalanine was performed by using -log of biological activity and various physiochemical descriptors as dependent and independent variable respectively and correlations were established using partial least square analysis. Among the models generated, the following model was selected on the basis of its statistical significance for further study.

Model:

$$-\log\text{IC}_{50} = [2.93] + \text{SaaNEindex} [-0.037] + \text{SsClcount} [0.77] + 4\text{path count} [0.064] + \text{SlogP} [-0.69]$$

Table 2. Description of Descriptor Used in the QSAR Study

S. No.	Descriptor	Sub-Descriptor	Type
1	Individual	Mol.Wt, H-Acceptor count, H-donarcount, Rotatable bond count, Polarizability AHP, XLogP, SlogP, SMR and Volume	Spatial
2	Chi	Chi0, Chi1, Chi2, Chi3, Chi4 and Chi5	Mol. Con. Index
3	ChiV	ChiV0, ChiV1, ChiV2, ChiV3, ChiV4 and ChiV5	Val. Mol. Con. Index
4	Path count	0Pathcount, 1Pathcount, 2Pathcount, 3Pathcount, 4Pathcount and 5Pathcount	High order Con.index
5.	ChiChain	Chi5Chain and Chi6Chain	High order Con. Index
6	ChiVChain	ChiV5Chain and ChiV6Chain	High order Con.index
7	Chainpathcount	5ChainCount and 6ChainCount	High order Con. Index
8	Cluster	Chi3cluster, ChiV3cluster and 3clusterCount	High order Con.index
9	Pathcluster	Chi4pathcluster, ChiV4path cluster and 4pathcluster Count	High order Con. Index
10	Kapa	Kappa1, Kappa2, Kappa3, K1alpha, K2alpha and K3alpha	Mol. Shape index
11	Element Count	H-count, C-count, S-count, O-count, N-count, Cl-count and F-count	Atom count
12	Estate number	SsCH3count, SssCH2count, SdssCcount, StsCcount, SaaCHcount, StNcount, SsssCHcount, StNcount, SaaNcount, SsssNcount, SdOcount, SssOcount, SddssS(sulfate)count and SsClcount	Electrotopological
13	Estate contribution	SsCH3E-index, SssCH2E-index, SaaCHE-index, SsssCHE-index, StsCE-index, SdssCEindex, SaasCE-index, SssNHEindex, StNEindex, SaaNE-index, SsssNEindex, SdOEindex, SsOHE-index, SddssS(sulfate)E-index and SsClE-index	Electrotopological
14	Polar surface area	Polar surface AreaExcludingPandS Polar surface AreaIncludingPandS	Electro-topological

$n=21$, $r=0.93$, $r^2=0.8567$, $r^2_{se}=0.25$, $F=33.89$, $F_{tabulated}=7.46$, $q^2=0.76$, $q^2_{se}=0.33$, $pred_r^2=0.42$, $pred_r^2_{se}=0.24$, degree of freedom=17

The derived model shows good correlation between biological activity and parameters SaaNEindex, SsClcount SlogP, and 4PathCount as the correlation coefficient $r = 0.93$ and the model explains about 85% variance in VLA-4/VCAM-1 inhibitory activity exhibited by piperazinyl phenylalanine derivatives. The low standard error of $r^2_{se}=0.25$ demonstrates accuracy of the model. The model show overall significance level better than 99.99%, with $F_{(4,17)} = 33.89$ against values of 99.99% significance $F_{(4,17, p0.001)} = 7.46$. The leave-one-out procedure was used for internal validation of the model. In this procedure high cross validated r^2 ($q^2=0.76$) and low $q^2_{se}=0.33$ value, reflects the very good internal predictive power of the model. However, a high q^2 value does not necessarily give a suitable representation of the real predictive power of the model for VLA-4 inhibitory ligands. So, an external validation was also carried out in the present study. The external predictive power of the model was assessed by predicting pIC_{50} value of the 10 test set molecules (Table 3), which were not included in the QSAR model de-

velopment. The maximum residual value for test set compounds is 0.48 or 6.06% which is showing good correlation between calculated activities and experimental activities. Another parameter for predictivity of test set compound is high $Pred_r^2 = 0.42$ and low $Pred_r^2_{se}=0.24$, which is showing good external predictive power of the model. The results are listed in Tables 3 and 4 and the experimental versus predicted activity of both test set and training set are depicted in Figs. (1) and (2).

The inter-correlation among the selected descriptors was very less due to auto scaling and cross correlation limit permitted was 0.6. The correlation matrix was shown in Table 5 which shows good correlation of selected parameters with biological activity.

The model incorporates four parameters SaaNEindex, SsClcount SlogP, and 4PathCount and their corresponding values for each molecule in the selected series were listed in Table 6. The descriptor SaaNE-index in the model represents the electro-topological state indices for number of nitrogen atoms connected with two aromatic double bonds. It shows presence of lone pair of electrons on nitrogen atom which is

Table 3. Experimental and Predicted Activities (pIC₅₀) with Residual Values for the Test Set Compounds

Compound	Experimental pIC ₅₀	Predicted pIC ₅₀	Residual ^a
01	8.16	8.12	0.04
02	8.17	7.93	0.24
04	7.99	7.93	0.06
10	7.91	7.43	0.48
11	7.65	7.48	0.17
14	7.86	8.00	-0.14
17	7.81	7.75	0.06
18	7.62	7.95	-0.33
19	8.12	7.95	0.17
30	7.29	7.53	-0.24

^aThe difference between experimental and predicted values.**Table 4. Experimental and Predicted Activities (pIC₅₀) with Residual Values for the Training Set Compounds**

Compound No.	Experimental pIC ₅₀	Predicted pIC ₅₀	Residual ^a
03	7.77	7.69	0.08
05	7.45	7.35	0.10
06	7.49	7.52	-0.03
07	7.35	7.50	-0.15
08	7.84	8.01	-0.17
09	8.00	8.10	-0.10
12	7.75	7.69	0.06
13	7.63	7.44	0.19
15	7.89	7.86	0.03
16	8.12	7.52	0.6
20	8.72	8.91	-0.19
21	8.17	7.79	0.38
22	7.91	8.04	-0.13
23	8.10	8.20	-0.10
24	7.70	8.09	-0.39
25	7.71	7.50	0.21
26	8.55	8.50	0.05
27	6.95	7.27	-0.32
28	6.17	6.42	-0.45
29	6.36	6.46	-0.10
31	8.38	8.11	0.27

^aThe difference between experimental and predicted values.

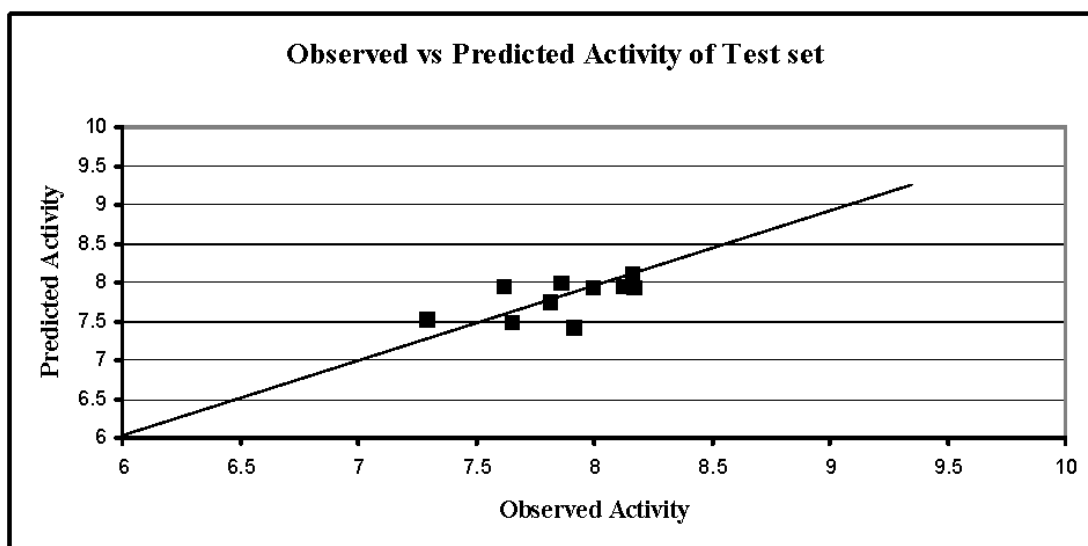


Fig. (1). Calculated versus experimental pIC_{50} for test set of VLA-4 inhibitors.

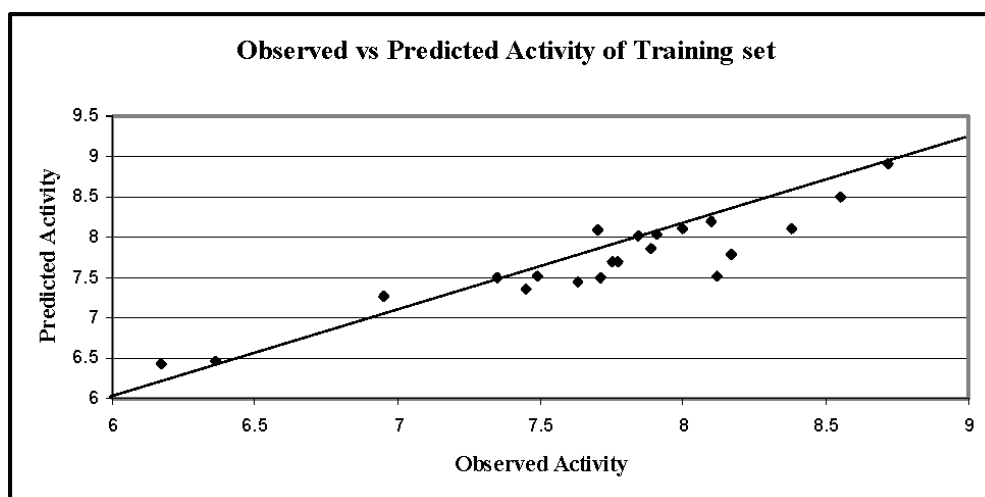


Fig. (2). Calculated versus experimental pIC_{50} for training set of VLA-4 inhibitors.

Table 5. Correlation Matrix for the Selected Descriptors

Variable	SsCl count	4path count	Slogp	SaaNEindex
SsCl count	1.00	-0.38	0.47	-0.018
4path count	-0.38	1.00	0.23	0.011
SlogP	0.47	0.23	1.00	-0.395
SaaNEindex	-0.01	-0.11	-0.39	1.00
Activity	0.022	0.48	-0.27	0.24

representing the interaction between drug and receptors. The negative correlation of the descriptor in the model indicates that electro-topological properties of the nitrogen atoms present in aromatic rings negatively influences VLA-4/VCAM-1 inhibitory activity shown by piperazinyl phenylalanine derivatives.

The SsClcount is an electro-topological parameter which can define the total number of chlorine atoms connected with

one single bond. The descriptor shows highest correlation among the all four parameter selected for the derived QSAR model. The positive coefficient of the descriptor suggests that VLA-4/VCAM-1 inhibitory activity of piperazinyl phenylalanine derivatives may be increased by increasing the number of chlorine atoms present in the nucleus.

The 4Pathcount is topological parameter which can signify the total number of fragments of fourth order (four bond

Table 6. Descriptors Used in QSAR Model with Value

Compound No	SaaNEindex	SsClcount	4path count	slogP
01	0	0	111	3.7170
02	0	0	120	3.7256
03	0	0	115	3.7256
04	0	0	116	3.7256
05	0	0	113	4.0254
06	0	0	119	4.3339
07	0	0	130	5.3840
08	0	0	119	3.6252
09	0	0	120	3.5887
10	8.6396	0	111	2.5070
11	8.5077	0	111	2.5070
12	4.4309	0	111	3.1120
13	0	0	94	2.1422
14	0	0	110	3.5607
15	0	0	104	2.4656
16	0	0	109	3.4086
17	0	0	117	3.2803
18	0	0	114	3.2803
19	0	0	114	3.2803
20	0	0	128	3.1519
21	0	0	116	3.6727
22	8.5023	0	109	2.1986
23	0	0	106	2.1406
24	0	0	102	1.9335
25	0	0	111	3.6275
26	0	2	114	4.7154
27	0	2	95	4.7457
28	0	0	95	3.7171
29	0	0	104	4.4884
30	4.1539	2	95	4.1407
31	0	2	105	4.4439

path) in compound. It is positively correlated with inhibitory activity so, it may be inferred that increasing the branching of compound is favorable for VLA-4 inhibitory activity.

The last descriptor SlogP in model represents to signify log of the octanol/water partition coefficient (including implicit hydrogen) which is a measure of the lipophilicity of the molecule. The descriptor is negatively correlated with bio-

logical activity in the QSAR models which suggests that decreasing the lipophilicity of the compounds will lead to increase in the VLA-4 inhibitory activity.

CONCLUSION

Summarizing the above, it may be concluded that a QSAR model with reliable predictive power for has been

successfully generated for the VLA-4/VCAM-1 inhibitory activity exhibited by a series of piperazinyl phenylalanine derivatives. The good correlation between experimental and predicted biological activity for 10 compounds in the test set further highlights the reliability of the constructed QSAR model. The result of the study suggests that the chlorine atoms in the molecule and fourth order fragmentation patterns in the molecular skeleton favours VLA-4/VCAM-1 inhibition shown by the title compounds whereas lipophilicity and nitrogen bonded to aromatic bond are not conducive for VLA-4/VCAM-1 inhibitory activity. The findings of the study will be helpful in the design of the potent VLA-4 inhibitors which will be potent anti-inflammatory agents of clinical utility.

ACKNOWLEDGEMENTS

With great pleasure and a sense of obligation, the authors wish express their heartfelt gratitude to Vice Chancellor, RGPV, Bhopal for providing experimental facilities and V-Life Sciences, Pune for providing VLifeMDS (Molecular Design Suite software) for the present study. The authors Mr. C. Karthikeyan wishes to thank CSIR, New Delhi for providing Senior Research Fellowship and Mr. Dinesh Bhargava thankful to AICTE, New Delhi for providing Scholarship.

REFERENCES

- [1] Osamu, S.; Kiminori, O.; Eri, A.; Yuji, N.; Hiroko, M.; Hiroaki, N.; Eiichi, F.; Yoshisuke, N. Synthetic study of VLA-4/VCAM-1 inhibitors: synthesis and structure-activity relationship of piperazinyl phenylalanine derivatives. *Bioorg. Med. Chem. Lett.*, **2008**, *18*, 1053-1057.
- [2] George, A.D.; Theodore, K.; Ermenegilda, M.C.; Gail, V.R.; Richard, A.M.; Sharon, T.; William, K.H. N-aryl 2,6-dimethoxybiphenylalanine analogues as VLA-4 antagonists. *Bioorg. Med. Chem. Lett.*, **2002**, *12*, 729-731.
- [3] Pretolani, M.; Ruffie, C.; Lapo e Silva, J.R.; Joseph, D.; Lobb, R.R.; Vargaftig, B.B. Antibody to very late activation antigen 4 prevents antigen-induced bronchial hyper-reactivity and cellular infiltration in the guinea pig airways. *J. Exp. Med.*, **1994**, *180*, 795-805.
- [4] Ianaro, A.; Cicala, C.; Calignano, A.; Koteliansky, V.; Gotwals, P.; Bucci, M.; Gerli, R.; Santucci, L.; Fiorucci, S.; Cirino, G. Anti-very late antigen-1 monoclonal antibody modulates the development of secondary lesion and T-cell response in experimental arthritis. *Lab. Invest.*, **2000**, *80*, 73-80.
- [5] Keszthelyi, E.; Karlik, S.; Hydruk, S.; Rice, G.P.; Gordon, G.; Yednock, T.; Horner, H. Evidence for a prolonged role of alpha sub 4 integrin throughout active experimental allergic encephalomyelitis. *Neurology*, **1996**, *47*, 1053-1059.
- [6] Podolsky, D.K.; Lobb, R.R.; King, N.; Benjamin, C.D.; Pepinsky, B.; Sehgal, P.; de Beaumont, M. Attenuation of colitis in the cotton-top tamarin by anti- α 4 integrin monoclonal antibody. *J. Clin. Invest.*, **1993**, *92*, 372-380.
- [7] Miller, D.H.; Khan, O.A.; Sheremata, W.A.; Blumhardt, L.D.; Rice, G.P.; Libonati, M.A.; Willmer-Hulme, A.J.; Dalton, C.M.; Miszkiel, K.A.; O'Connor, P.W.N. A controlled trial of natalizumab for relapsing multiple sclerosis. *Eng. J. Med.*, **2003**, *348*, 15-23.
- [8] Abraham, W.M.; Gill, A.; Ahmed, A.; Sielczak, M.W.; Laredo, I.T.; Botinnikova, Y.; Lin, K.C.; Pepinsky, B.; Leone, D.R.; Lobb, R.R.; Adams, S.P. A small-molecule, tight-binding inhibitor of the integrin α -4 β -1 blocks antigen-induced airway responses and inflammation in experimental asthma in sheep. *Am. J. Respir. Crit. Care Med.*, **2000**, *162*, 603-611.
- [9] Marie-Agne's, L.; Fabienne, B.; Philippe, C.; Laurent, D.; Francoise de, L.D.; Thierry, D.; Abdoulaye, G.; Ge'rald, G.; Jean-Claude, H.; Laszlo, K.; Laurent, K.; Jean-Marie, N.; Ste'phanie, N.; Luc, Q.; Sylvain, R.; Vale'rie, V.; Laurent, P. 2,6-Quinolinylnyl derivatives as potent VLA-4 antagonists. *Bioorg. Med. Chem. Lett.*, **2007**, *17*, 142-146.
- [10] www.Vlifesciences.com
- [11] V-Life, molecular design suite TM 3.5, Developed by V-Life Technologies, Pune, India, **2006**.
- [12] OPENSTAT software ver. 4, developed by Bill Miller from Iowa State University, USA, March, **2008**.