QSAR STUDY OF SOME 2-SUBSTITUTED NAPHTHOLS AS 5-LIPOXYGENASE INHIBITORS

*N. H. INDURWADE and R. S. WANARE

S. N. Institute of Pharmacy, PUSAD-445204 (MS) INDIA

ABSTRACT

In the present study, a three–dimensional quantitative structure activity relationship study on series of substituted naphthols as lipoxygenase inhibitors has been performed. Various physicochemical parameters were calculated using Cerius2 molecular modeling software. Quantitative structure activity relationship models were generated for lipoxygenase inhibitory activity using stepwise multiple regression and genetic function approximation analysis. Statistically significant models were obtained in the series, with 38 compounds, gave r² value as 0.857. The studies indicated that in series shape and electronic descriptors dominated the influence on the activity. Cross validation was performed using the leave—on—out method. The so obtained and validated models bring important structural insight to aid the design of novel 5–lipoxygenase inhibitors prior to their synthesis.

Key words: 3D-QSAR, Lipoxygenase inhibitors, Physiochemical parameter, Descriptors.

INTRODUCTION

Non steroidal anti–inflammatory agents are of current interest¹ because there are no drugs of choice for the treatment of most of the diseases like rheumatoid arthritis², allergic rhinitis, psoriasis^{3, 4}, ulcerative colitis and asthama ⁵. The two major approaches for design and synthesis of anti–inflammatory agents are based on the inhibition of two enzymes⁶, cyclooxygenase and lipoxygenase, which are involved in the metabolism of arachidonic acid (AA). Cyclooxygenase has been the common target for most of the anti–inflammatory drugs but due to the association of some side effects such as ulceration and bleeding in gastrointestinal tract with cyclooxygenase inhibitors⁷ and implication of leukotrienes in the above inflammatory and allergic disorders^{6, 9}, the attention is focused on the 5–lipoxygenase enzyme inhibitors, which restrict the synthesis of leukotrienes from AA via peroxidation of AA to 5–hydroperoxysicoteranoic acid (5–HPETE) followed by dehydration to 5,6–epoxy leukotrience A₄(LTA₄). No three–dimension quantitative structure activity relationship (3D QSAR) studies have been attempted so far on series of substituted naphthols derivative, it appeared of interest to perform 3D QSAR analysis employing Cerius2 software. The aim of this study was to find 3D QSAR models with good correlation between molecular structure and biological activity.

Such an effort would facilitate the discovery and development of potent 5-lipoxygenase inhibitors

EXPERIMENTAL

The 5-lipoxygenase inhibitory data were taken from Batt et al 10. All molecular modeling and 3D-QSAR studies were performed on a Silicon Graphics Indigo2 XZ employing Cerius2 software 10 (Version 3.5). The IC₅₀ values originally expressed in micromole have been converted to kilomole and grammole for series for convenience of computation work.. The molecular structure of all compounds (Table 1) were built using molecular sketching facilities provided in modeling environment of Cerius2. The molecular structures were minimized using the steepest descent (SD); conjugate gradient 12 (CG) in sequence followed by truncated Newton-Raphson (N-R) optimization techniques under universal force field 13. The minimization terminates at which the root mean square (RMS) force on the molecule was less than 0.0001 Kcal\mol Å. The conformations were generated and its analysis for each compound was performed using GRID method. To generate the conformation, the energy cut off was set to 5 Kcal/mol. The number of conformation generated for each substrate was limited to a maximum of 150. The conformations produced by the random conformation search were fully optimized and used immediately for further analysis. Among the so constituted conformation space, only the conformers above the lowest energy minima have kept for their geometry reoptimization with the semi-empirical MOPAC package version 6.0 using the Hamiltonian AMI. The lowest energy conformation for each compound was found. The descriptors were calculated using the lowest energy conformation.

Fig.1 2-(Aryl methyl)-1-naphthols

For structures, ^aRat Basophilic Leukemia (RBL-1) cell lysate 5-lipoxygenase inhibitions measured by 5-HETE productions, ^bIC₅₀ values were expressed in terms of grammoles, ^bCalculated (-Log IC₅₀) values using Equation 5.

The following descriptors were calculated for 3D–QSAR studies (values of only those descriptor, which found place in the equation, are given in Table 2. The thermodynamic descriptor such as, desolvation free energy for water $(FH_2 O)^{14,15}$, desolvation free energy for octanol $(FOCT)^{14,15}$.

Log of partition coefficient (AlogP)^{14,15} and molecular reactivity (MR)^{15,16}; Spatial descriptor such as, number of rotable bonds (ROTBONDS)¹⁵, molecular surface area

(AREA)¹⁷, radius of gyration (ROG), density (DENSITY)¹⁷, molecular weight (MW)¹⁷, molecular volume (VM)¹⁷, principle moment of inertia (PMI)¹⁸, principle moment of inertia—X component (PMI–X)¹⁸, principle moment of inertia—Y component (PMI–Y)¹⁸, principle moment of inertia—Z component (PMI–Z)¹⁸; Electronic descriptors such as, sum of atomic polarizabilities (APOL)¹⁹, dipole moment (dipole—mag)^{19,20}, dipole moment—X component (Dipole—X)^{19,20}, dipole moment—Z component (Dipole—X)^{19,20}, energy of highest occupied molecular orbital (HOMO)²¹, energy of lowest unoccupied molecular orbital (LUMO)²¹, superdelocalizability (Sr) and partial atomic charges²² were calculated.

Molecular shape descriptors, common overlap steric volume (COSV), difference volume (DIFFV), common overlap volume ratio (Fo), non–common overlap steric volume (NCOSV), RMS to shape reference (Shape RMS), volume of shape reference (sr.vol.), Vander waal's surface (vdws), were calculated to compare the common properties of the molecule and to measure molecular shape commonality. The conformers were used for computational calculation of different physiochemical properties including atomic charges, electron density, HOMO, LUMO and dipole moments based on atomic contribution using MOPAC 6.0 (MNDO method). Partial charges were calculated using charge (Qeq) equilibrium method 23. Quantum mechanical descriptor such as, lowest unoccupied molecular orbital energy (LUMO_MOPAC), dipole moment (Dipole_MOPAC), highest occupied molecular orbital energy (HOMO_MOPAC), heat formations HF_MOPAC were calculated using MOPAC module. Connoly surface descriptor was derived with the probe radius at 1.40 Å, dot density at 8.0A 2 and VDW scale factor at 1.

To generate 3D–QSAR equations, stepwise multiple regression analysis and genetic function approximation analysis (GFA) method were used. GFA was used since it generates a population of equations for correlation between biological activity and physiochemical properties. GFA developed by Rogers involve combination of Friedman's multiple adaptive regression splines (MARS) algorithm with Holland's genetic algorithm to evolve a population of equation that best fit the data. A distinctive feature of GFA that instead of generating single model, as do most other statical methods, it produces population of models. The range of variation in this population gives added information on the quality of fit and importance of descriptors¹¹. The cross validation was performed using the leave—one—out method procedure. The following statistical measures were used: the number of sample in regression (n), squared correlation coefficient (r²), F–test for statistical significance (F), Friedman's lack of Fitness (LOF), Cross validated squared correlation coefficient (cvr²), boot strapped squared correlation coefficient (bsr²).

Table 1. Structure of compounds

Comp	Ar Ar	X	BA	(-LogIC ₅₀) b	Calculated (-LogIC ₅₀) c 7.133	
No	Territoria, Maria		^a RBL–1, IC50μM	Transport Village		
1	C ₆ H ₅	H	0.019	7.721		
2	4-CH ₃ C ₆ H ₄	Н	0.024	7.619	7.161	
3	2-CH ₃ C ₆ H ₄	Н	0.120	6.920	7.154	
4	4-CH ₃ OC ₆ H ₄	Н	0.075	7.124	7.048	
5	3-CH ₃ OC ₆ H ₄	Н	0.110	6.958	7.110	
6	4-C ₂ H ₅ OC ₆ H ₄	Н -	0.062	7.207	7.099	
7	3,4-(CH ₃ O) ₂ C ₆ H ₃	H	0.051	7.292	7.09	
8	3-C ₆ H ₅ OC ₆ H ₄	Н	0.083	7.080	6.78	
9	4-C ₆ H ₅ CH ₂ OC ₆ H ₄	Н	0.072	7.142	7.30	
10	4-CH ₃ SC ₆ H ₄	Н	0.048	7.318	7.13	
11	4-(CH ₃) ₂ NC ₆ H ₄	Н	0.014	6.853	7.22	
12	4-FC ₆ H ₄	Н	0.130	6.886	7.27	
13	4-CIC ₆ H ₄	H	0.039	7.408	7.84	
14	3-CIC ₆ H ₄	H	0.052	7.283	7.25	
15	3,4-Cl ₂ C ₆ H ₃	H	0.270	6.744	7.00	
16	3-CF ₃ C ₆ H ₄	H	0.048	6.568	6.36	
17	4-(CH ₃ SO ₂)C ₆ H ₄	Н	0.240	6.619	6.78	
18	4-NO ₂ C ₆ H ₄	Н	0.048	7.318	7.24	
19	2-naphthyl	Н	0.058	7.236	7.15	
20	2-thienyl	H	0.060	7.221	6.94	
21	2-furyl	Н	0.037	7.431	6.71	
22	N-CH ₃ -2-pyrryl	H	0.140	6.853	7.06	
23	2-pyridyl	Н	0.390	6.408	6.77	
24	3-pyridyl	Н	0.230	6.638	6.73	
25	4–pyridyl	Н	0.140	6.852	7.00	
26	C ₆ H ₅	5-CH ₃ O	0.100	7.000	7.36	
27	C ₆ H ₅	5,7(CH ₃) ₂	0.062	7.207	7.380	
28	C ₆ H ₅	5,8(CH ₃) ₂	0.120	6.920	7.383	
29	C ₆ H ₅	4-CH ₃	0.020	7.698	7.297	
30	C ₆ H ₅	4-C ₆ H ₅	0.056	7.251	7.466	
31	C ₆ H ₅	4-OCH ₃	0.031	7.508	7.499	
32	C ₆ H ₅	4-COCH ₃	0.510	6.292	6.408	
33	C ₆ H ₅	4-COC ₆ H ₄	0.054	7.267	6.905	
34	C ₆ H ₅	4-COOH	5.800	5.236	5.215	
35	C ₆ H ₅	4-SO ₂ CH ₃	0.160	6.795	6.468	
36	C ₆ H ₅	4-SO ₂ CH ₃	40.00	4.397	4.962	
37	C ₆ H ₅	4-SO ₂ NH ₂	9.400	5.026	4.668	
38	C ₆ H ₅	4-NO ₂	48.00	4.318	4.560	

Table 2. Calculated descriptor values

Comp. No	^a vdws	^b Dipole–Z	cROG	^d PMI–Y	сномо	Fo	gNCOSV.	hHOMO_ MOPAC	ⁱ Dipole–X	JDipole_MOPAC
1	113	2.43	3.68	361	- 10.3	0.92	19.09	-8.19	- 0.46	1.16
2	118	2.54	3.97	445	- 10.3	0.84	39.23	- 8.17	- 0.53	1.35
3	114	2.66	3.78	407	- 10.4	0.81	46.40	-8.18	- 0.44	1.24
4	123	1.67	4.15	544	- 10.3	0.09	48.16	-8.18	- 1.27	1.74
5	123	3.45	3.99	497	-10.4	0.47	133.1	-8.22	-0.21	1.36
6	131	1.63	4.48	656	- 10.3	0.76	63.50	-8.17	-1.14	1.74
7	131	1.76	4.30	662	-10.4	0.73	75.01	-8.21	1.41	2.14
8	146	-0.02	4.99	1096	-10.4	0.49	154.8	-8.37	0.97	1.92
9	152	1.41	5.33	1313	- 10.3	0.62	121.2	-8.19	-0.99	1.50
10	129	2.56	4.18	626	-10.4	0.78	57.71	-8.21	- 1:53	1.67
11	128	1.99	4.35	616	-10.3	0.75	68.24	-8.14	-0.29	1.73
12	118	6.99	3.69	454	-10.4	0.49	116.7	- 8.27	3.34	1.56
13	113	7.95	3.69	542	- 10.5	0.48	124.5	-8.28	4.49	1.48
14	122	5.62	3.69	458	-10.5	0.53	112.2	- 8.25	3.78	1.50
15	133	3.33	3.71	614	-10.5	0.79	51.70	-8.33	-2.68	2.21
16	129	-2.28	3.81	544	- 10.5	0.78	56.32	-8.30	-3.91	2.54
17	133	1.90	4.34	770	-10.5	0.72	76.78	-8.39	-2.71	6.57
18	127	8.90	3.87	597	-10.6	0.40	133.6	-8.52	5.08	6.18
19	130	2.37	4.23	628	-10.3	0.76	64.40	- 8.20	-0.50	1.14
21	112	3.41	3.53	357	-10.3	0.53	101.1	- 8.20	-3.69	1.17
21	106	2.57	3.49	308	-10.3	0.53	95.92	- 8.23	- 3.45	1.01
22	112	2.21	3.75	380	-10.3	0.71	64.91	-8.14	-0.80	3.27
23	113	2.20	3.66	358	- 10.4	0.51	108.4	-8.21	-3.25	2.19
24	112	0.96	3.64	362	-10.3	0.83	38.22	-8.28	-1.35	1.89
25	111	3.27	3.59	362	- 10.4	0.90	22.86	-8.31	- ().62	1.91
26	121	1.31	4.09	497	- 10.1	0.81	47.69	- 7.97	0.05	0.41
27	124	2.28	4.06	459	-10.0	0.79	54.94	- 8.04	-0.53	0.91
28	120	1.69	3.91	426	-10.0	0.78	56.45	-8.00	-0.31	1.13
29	119	2.24	3.77	377	-10.1	0.84	38.32	- 8.09	-0.52	1.12
30	140	3.03	4.20	547	-10.1	0.68	93.49	-8.13	- 0.45	1.13
31	123	1.72	3.83	389	-10.0	0.81	47.26	- 7.93	-1.44	2.23
32	126	-0.92	3.88	422	-9.9	0.78	58.21	-8.38	- 2.46	3.21
33	147	0.92	4.35	623	- 10.3	0.64	113.1	- 8.36	-3.21	3.19
34	124	-5.33	3.79	427	- 10.2	0.49	136.1	- 8.68	0.93	4.43
34	139	0.22	4.37	579	- 10.6	0.70	85.69	- 8.47	0.65	1.13
36	135	-4.84	4.03	495	- 10.5	0.73	74.89	- 8.95	-1.48	6.98
37	133	-6.03	3.97	501	- 10.7	0.74	69.27	- 8.02	-1.42	6.43
38	125	-6.85	3.76	419	-11.0	0.81	46.64	- 8.99	-0.63	6.24

^aVander waals surface, ^bDipole moment–Z component, ^cRadius of gyration, ^d Principle moment of inertia– Y component, ^eEnergy of highest occupied molecular orbital, ^fCommon overlap volume ratio, ^gNon–common overlap steric volume, ^hHighest occupied molecular orbital energy, ⁱDipole moment–X component, ^jDipole moment

RESULTS AND DISCUSSION

The following 3D–QSAR equations 1 to 4 were generated by stepwise multiple regression analysis for series (2–(aryl methyl)–1–naphthols).

- (1) $-\text{Log IC}_{50} = 22.1385 + 0.104179 \text{*Dipole} \text{Z} + 1.865 \text{*HOMO}$ $n = 38, r^2 = 0.841, F = 43.573, \text{ press} = 5.50, \text{ cvr}^2 = 0.772, \text{ bsr}^2 = 0.841$
- (2) $-\text{Log IC}_{50} = 5.37107 0.0779649 \text{*Dipole} \text{X} + 0.178269 \text{*Dipole} \text{Z} + 0.387054 \text{*ROG} 0.174287 \text{*Dipole} \text{MOPAC}$ $n = 38, r^2 = 0.841, F = 43.573, \text{ press} = 5.50, \text{ cvr}^2 = 0.772, \text{ bsr}^2 = 0.841$
- (3) $-\text{Log IC}_{50} = 21.4715 + 0.0106872*\text{vdws} + 0.108456*\text{Dipole} Z + 1.94776*\text{HOMO_MOPAC}$ $n = 38, r^2 = 0.849, F = 63.534, \text{press} = 4.653, \text{cvr}^2 = 0.807, \text{bsr}^2 = 0.848$
- (4) $-\text{LogIC}_{50} = 15.8315 + 0.137275*\text{Dipole-Z} + 0.000607239*\text{PMI-Y} + 0.888898*\text{HOMO-} 0.14235*\text{Dipole_MOPAC}$ $n = 38, r^2 = 0.832, F = 40.989, \text{press} = 5.76, \text{cvr}^2 = 0.761, \text{bsr}^2 = 0.832$

The following 3D-QSAR Equation 5 was obtained by genetic functional approximation analysis,

(5)
$$-\text{LogIC}_{50} = 17.5829 + 0.124249 * \text{Dipole} - Z + 0.013485 * \text{NCOSV} + 1.75508 * \text{HOMO}_{\text{MOPAC}} + 3.67723 * \text{Fo}$$

 $n = 38, r^2 = 0.857, F = 49.754, \text{LOF} = 0.145, \text{press} = 4.81, \text{cvr}^2 = 0.801, \text{bsr}^2 = 0.858$

Equation 4 obtained has good correlation ($r^2 = 0.832$) between the parameter and biological activity. Statistical parameters were obtained for equation (1), viz., $r^2 = 0.841$, F = 43.573, $cvr^2 = 0.772$, $bsr^2 = 0.841$. It shows better statistical significance > 99.9% with $F_{(2, 35)} = 85.25$ against the value for 99.9% significance ($F_{(2, 35)} = 0.001$) = 8.41). The R^2 accounts for 84% variance in the activity values. In equation 1, both parameters such as, Dipole–Z and HOMO contributed positively for activity. By evaluation of this model, we can conclude that electronic parameters play an important role in the activity.

In addition to the above equations, some statistically improved equations (equations 2–4) were also obtained. Equation 4 has good statistical significance > 99.9% with the F-values $F_{(4, 33)} = 40.989$, against the value of 99.9% significance ($F_{(4, 33 \,\alpha\,0.001)} = 5.11$). The F-value obtained for equation 3 is significant at 99.9% level ($F_{(3, 34 \,\alpha\,0.001)} = 6.79$). The $F_{(4, 33)} = 6.79$. The $F_{(4, 33)} = 6.79$ accounts for 84.1%, 84.1% and 83.2% variance in activity values for equations (1), (2) and (4), respectively. Among three equations, equation (3) was found better and indicated that Vander Waals surface (vdws), Dipole–Z and HOMO_MOPAC played an important role for activity.

Among all the 3D–QSAR equations, equation (5) was statistically significant. This model has good predictive power according to the statistical result ($r^2 = 0.857$, $cvr^2 = 0.801$ and $bsr^2 = 0.858$). The goodness of the structure activity correlation was estimated by r^2 ($r^2 = 0.857$). The lack of fitness value is very less (0.145). The F–value has a high statistical significance 99.9% with $F_{(3, 33)} = 49.75$, against the value for 99.9% significance ($F_{(4, 35 \alpha 0.001)} = 5.88$). The $F_{(3, 33)} = 49.75$, variance in the biological activity values. The cross validated ($F_{(2, 35 \alpha 0.001)} = 5.88$) results reflect accuracy of the model. The Dipole–Z, NCOSV, HOMO_MOPAC and Fo contributed positively for activity in the order of 0.03 : 0.003 : 0.47 : 1. From the above analysis, it may be inferred that equation 5 can be used for the theoretical prediction of 5–lipoxygenase inhibitory activities of the new molecules.

ACKNOWLEDGEMENTS

The authors thanks M/s Torrent Research Centre, Bhat, Gandhi Nagar for providing access to molecular modeling software facilities and to Director, S.G.S.I.T.S., Indore, for providing necessary facilities for this work.

REFERENCES

- M. Arockia Babu, S. Neeraj, P. Prathipati and S. G. Kaskhedikar, Bioorg. Med. Chem., 10, 4035 (2002).
- 2. E. M. Davidson, S. A. Rae and M. H. J. Smith, J. Pharm. Pharmacol., 34, 410 (1982).
- 3. S. D. Brain, R. Camp, S. Charleston, P. M. Dowd, A. W. Ford-Hutchinson, M. W. Greaves and B. A. Kobza, Brit. J. Clin. Pharmacol, 17, 65 (1984).
- 4. S. D. Brain, R. Camp, P. M. Dowd, M. W. Greaves and B. A. Kobza, Invest. Derm., 83, 70 (1984).
- 5. S. E. Wenzel, Pharmacotherapy, **35**, 117 (1997).
- 6. C. Hansch, P. G. Sammes, J. B. Taylor, "Enzyme and Other Molecular Target", Vol.II Pargamon Press, NY (1900) p.147.
- 7. M. C. Allison, A. G. Torrance, C. J. Lee and R. I. Russel, N. Engl. J. Med. 20, 749 (1992).
- B. Samueisson, Science, 220, 568 (1983).
- 9. A. W. Ford-Hutchinson, Fed. Proc., 44, 25 (1985).
- D. G. Batt, D. M. George, J. P. Joseph, E. S. Joan, G. William and R. H. Richard, J. Med. Chem., 33, 360 (1990).
- 11. Cerius² Version 3.5 Biosym/Molecular simulation Inc.
- 12. R. Fletcher and C. M. Reeves, Comput. J., 7, 149 (1964)

- A. K. Rappe, C. J. Casewit, K. S. Colwell and W. M. Skiff, J. Amer. Chem. Soc., 114, 10024 (1992).
- J. E. Leffler and E. Grunwald "Rates and Equilibrium Constant of Organic Reactions"
 Vol. I, John Wiley and Sons, NY (1963). p.33
- C. Hansch and A. Leo "Substituent Constant for Correlation Analysis in Chemistry and Biology" John Wiley and Sons, NY (1979) p.18.
- 16. W. J. Dunn, Eur. J. Med. Chem., 12, 109 (1977).
- 17. H. Waterbeemd and B. Testa, Adv. Drug Res., 16, 85 (1987).
- 18. T. L. Hill "Introduction to Statistical Thermodynamics" Addison Wesley (1960) p.189.
- 19. M. Marsali and J. Gasteiger, Croatica Chem. Acta, 53,106 (1980).
- 20. G. Del Re, Biophys. Acta, 75, 153 (1963).
- 21. J. Gastriger and M. Marsali, Tetraedron Lett., 34, 3181 (1966).
- 22. J. A. Pople and G. A. Segal, J. Chem., Phys., 44, 3286 (1991).
- 23. A. K. Rappe and W. A. Goddard, J. Phys. Chem., 95, 3358 (1991).

Accepted: 12.1.2004