www.sphinxsai.com

PharmTech

International Journal of PharmTech Research
CODEN (USA): IJPRIF ISSN : 0974-4304
Vol. 3, No.1, pp 231-236, Jan-Mar 2011

3D-QSAR and insilico study: Modeling parameters for designing new selective 12-LO Enzymes inhibitors

Hemendra Pratap Singh[1]*, Adya Prasad Chaturvedi[2], Chandra Shekhar Sharma[1]

[1]Department of Pharmaceutical Chemistry, B.N. College of Pharmacy, Udaipur -313 001,India
[2]Department of medicinal chemistry, IMS, BHU,India

Corres. Author: hps_medicinalchemistry@yahoo.co.in
Phone: +91-9929378331

Abstract: Selective 5-LO and 12-LO enzyme inhibitors have attracted much attention in recent times in the design of novel anthracenone derivatives, which may be used in cancer and psoriasis. However, not much computational studies has been done to examine specific type of parameters beneficial for anthracenone derivatives against 12-LO inhibitions. In our preliminary study we have done 2D QSAR studies by include a series of 2-arylalkyl substituted anthracenone derivatives having inhibitory action on 12-LO isoforms in epidermal homogenate of mice, bovine platelets and porcine leukocyte. These studies produced good predictive models and give statistically significant correlations with selective 12-LO enzyme inhibition. So these structural activity relationship studies were extended to 3DQSAR model studies utilizing theoretical molecular descriptors that can be calculated directly from molecular structures. All the descriptors were selected by genetic algorithm and multiple linear regression (MLRA) methods. All the QSAR models were validated by leave one out (LOO) method. In addition, *insilico* toxicity studies were also done. These all studies helps in designing some novel anthracenone derivatives with selective 12-LO enzyme inhibition activity with less toxicity.
Key words: QSAR, 12-LO, descriptors, *insilico* toxicity, molecular modeling.

Introduction

Lipoxygenases are a superfamily of enzymes that catalyse the transfer of molecular oxygen into arachidonic acid to yield hydroperoxy fatty acids, which are further metabolised to hydroxy derivatives as end products [1]. These enzymes show not only marked stereospecificity but also differ in the position of dioxygenation in arachidonic acid, and the currently used nomenclature is based on this specificity of the enzyme acting on its substrate [2]. Lipoxygenase proteins have a catalytical domain containing a nonheme iron, and oxidation to the active ferric form is required for catalysis [1]. The physiological role of individual mammalian lipoxygenases is uncertain aside from 5- lipoxygenase, which is the key enzyme of the biosynthesis of leucotrienes [3]. Inhibitors of the 5-lipoxygenase enzyme are used in the therapy of asthma [3] and could provide useful therapy in inflammatory skin diseases such as psoriasis [4, 5]. Different isoforms of 12- lipoxygenase have been described. Based upon biochemical and immunological criteria, 12-lipoxygenases have been characterized into platelet-type (*p*12-LO) and leucocyte-type (*l*12-LO) categories [6]. Furthermore, mouse epidermal lipoxygenase (*e*12- LO) represents a third isoform that is more related to *p*12-LO than to *l*12-LO [7, 8].

Formation of the 12(*S*)-hydroxyeicosatetraenoic acid (12(*S*)-HETE) enantiomer can be accounted for by these isoforms [9–15]. Recently, *R*-lipoxygenases have also been identified [16, 17]. However, for the 12-LO isoforms, their biological roles are far from clear. Increased 12-LO activity has been demonstrated to induce apoptosis in fibroblasts [18], and its metabolite

12(*S*)- HETE has been reported to be critically involved in cancer metastasis [19] and hyperproliferative skin diseases such as psoriasis [4].

Material and Method

The structures of the 21 anthracenone derivatives were chosen for this work and their functional groups are shown in Figure 1. Biological data used in this study is taken from the reference [20]. This is in a good agreement with structure-activity investigations. All original IC_{50} biological activity data has been converted to –log IC_{50} response variables.

Molecular Modeling: The molecular modeling studies (molecular mechanics and semiempirical calculations) were carried out using the Chemoffice Ultra-8 software package. The Molecular Mechanics (MM+) force field was applied for preliminary structure optimization and study of the conformational behaviors of each anthracenone derivatives. Molecular mechanics has been shown to produce more realistic geometry values for the majority of organic molecules owing to the fact of being highly parameterized [21]. The next step was a re-optimization of the MM+ optimized structures by applying AM1 semiempirical method. Quantum mechanical method has been used in order to obtain an accurate charge distribution and quantum-chemical descriptors for each compound in the series.

Molecular descriptors: Descriptors are normally calculated for molecules after a low energy conformation has been found and optimized using any standard optimization technique, e.g. molecular mechanics, *ab-initio*, DFT or semi-empirical methods. The molecular descriptors used in this study have been calculated applying the *DRAGON* program [22]. The program contains scripts for generating descriptors of different types including: constitutional, topological, RDF, GETAWAY, functional groups, WHIM, Randic, 3D-Morse etc [19]. A set of additional quantum-chemical descriptors (energy of heat of formation, HOMO, LUMO, log P, partition coefficient, total energy etc.) has also been obtained for the each molecule after geometry optimization procedure.

Statistical Methods: Preliminary models selection was performed by means of GAMLRA technique as implemented in the B-QSAR [23] program. As mentioned before, this approach allows selection of the models with the following characteristics: high quadratic correlation coefficient R^2, low standard deviation S and the least number of descriptors involved. Next, the openstate-4 professional software package was applied for detailed statistical analysis of the models obtained. Thus, the high Fisher coefficient F, non-collinear descriptors, and the significance level P variable served as additional selection parameters. A final set of QSAR was identified by applying the "leave-one-out" technique with its predicting ability being evaluated and confirmed by cross validation coefficient Q^2 based on predictive error sum of squares (SPRESS).

Result and Discussion

2D QSAR study: 21 Compounds belonging to anthracenone category were taken for the present study. The biological activities data for anthracenone derivatives were taken from literature. The IC_{50} values for inhibitory action on 12-LO isoforms in epidermal homogenate of mice, bovine platelets and porcine leukocyte were transformed into –log IC50. Stepwise regression analysis was performed by taking -log IC_{50} as dependent variable and descriptors as independent variables. From the analysis significant equations were selected which were validated by leave one out method. The significant regression equations are:

-log IC50 = -1.343 (± 0.23) R2π p-1.132 ……...Eq.(1)
n = 17, F = 34.091, R = 0.833, R2 = 0.694, VIF = 1, PRESS = 0.421, Q2 = 0.7137, SDEP= 0.02

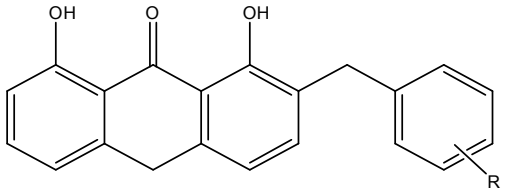
-log IC50 = 1.605 (± 0.281) R2 f-1.133 ……... Eq.(2)
n = 17, F = 32.637, R = 0.828, R2 = 0.685, VIF = 1, PRESS = 0.375, Q2 = 0.684, SDEP =0.242

-log IC50 = 0.443 (±0.08) R2 Ha-1.132 ……...Eq.(3)
n = 17, F = 30.791, R = 0.821, R2 = 0.672, VIF = 1, PRESS = 0.409, Q2 = 0.788, SDEP =0.155

-log IC50 = 3.691 (±0.665) R2 σm -1.132 ……...Eq.(4)
n = 17, F = 30.791, R = 0.820, R2 = 0.672, VIF = 1, PRESS = 0.488, Q2 = 0.5879, SDEP= 0.169

From above equations it was concluded that group at position 2 effects biological activity of the parent compound significantly. Descriptors effects the biological activity significantly, are lipophilicity, field effect, inductive effect and hydrogen acceptor in the pool of descriptors, taken for study. These are also cross validated by PRESS, Q^2, SDEP values. Validation results also concluded that the selected equations are significant. Epidermal homogenate of mice and bovine platelets are having insignificant correlation with physicochemical properties. These initial results are promising for the development of some novel anthracenones, which are selective 12-LO enzyme inhibitors without renal or gastric toxicity.

Table:1 Parent structure and biological activity of the anthracenone series



1.8-dihydroxy-2-phenylalkyl-9(10H)-anthracenone

Com no	R	log IC50 (Porcelein leucocyte)		
		Observed	Predicted	Residual
6a	H	-	-	-
6b	4-Ph	0.734	0.577	0.217
6c	4-Me	0.691	0.721	-0.030
6d	4-CF3	0.83	0.871	-0.041
6e	4-F	0.796	0.936	-0.140
6f	4-Cl	0.706	0.813	-0.107
6g	4-Br	0.698	0.740	-0.042
6h	4-CN	1.107	1.044	0.063
6i	4-OPh	-	-	-
6j	4-On-Pr	0.7966	0.648	0.148
6k	4-Et	0.85	0.964	-0.114
6l	4-OMe	0.745	0.884	-0.139
6m	$3,4\text{-}(OMe)_2$	0.96	0.931	0.029
6n	$2,4\text{-}(OMe)_2$	0.872	1.020	-0.148
6o	$2,5\text{-}(OMe)_2$	0.83	1.060	-0.230
6p	3,4-OCH2O	0.926	0.893	0.033
6r	4-COOMe	0.872	0.869	0.003
6s	4-COOH	0.897	0.968	-0.071
6t	$3,4\text{-}(OH)_2$	1.66	1.368	0.292
6u	$2,4\text{-}(OH)_2$	1.183	1.037	0.146
6v	$2,5\text{-}(OH)_2$	1.285	1.241	0.044
6w	1-naphthyl	-	-	-
6x	2-naphthyl	--	--	--
6y	2-thienyl	0.83	0.742	0.088

3D-QSAR study: As mentioned above, a 2D-QSAR analysis has been performed on anthracenone derivatives aiming at establishing a structure-activity relationship. For the sake of simplicity we have divided the models into two general groups in accordance with nature of the descriptors involved: group one, comprised of 3D descriptors generated by DRAGON, and group two, containing physicochemical descriptors only. While constructing the models, great care was taken in order to avoid inclusion of highly collinear descriptors. The inter correlation matrix for the physicochemical descriptors used in this study is given in Table 1. The table includes only those variables that have comprised the most populated models selected by the variable selection Genetic Algorithm method.

3D Descriptor Containing Models: One of the main advantages of these descriptors is the unambiguity regarding the 3D arrangement of atoms. There are also other properties making them flexible and therefore popular descriptors to be used. One of them is independence from the molecular size resulting in applicability to the large datasets with great structural variance. Another important property of 3D descriptors is their invariance
against translation and rotation of the molecule. Such atom-based descriptors are also suggested to be

applied for the collection of exotic chemicals, since there is a greater chance of physicochemical descriptors giving misleading information [22]. For the model generation we have chosen 3D descriptors of the following type: RDF, 3D-Morse, available within DRAGON. Several runs of GA-MLRA variable selection technique implemented in B-QSAR program have resulted in models containing mainly RDF, and 3D-Morse type descriptors having no significant equation together with the biological activity. Values of 3D descriptors together with the toxicity data are indicated in Table 2.

3D-Morse descriptors were obtained on the basis of the molecular transform equation used in electron diffraction [23]. RDF code is based on the radial distribution function of an ensemble with N atoms, i.e. probability distribution of finding atom on a sphere with radius r [24].

Monoparametric equations

log IC_{50}= -0.18750(±0.08892) par coff +1.98840(±0.515815) ……...Eq.(5)

(n=20; r=0.921; s=0.168; F=39.481; Q^2=0.549; SPRESS=0.196)

Biparametric equations

log IC_{50}= -0.21574(±0.08590)part coffi - 0.00019(±0.00019)Tot Eng +1.28515 (±0.831784).....**Eq.(6)**

(n=20; r=0.889; s=0.154; F=34.045; Q^2=0.555; SPRESS=0.185)

log IC_{50}= -0.22569(±0.08875) part coffi - 0.00002(±0.00002)Ele Eng +1.53776 (±0.636074).....**Eq.(7)**

(n=20; r=0.892; s=0.153; F=34.272; Q^2=0.571; SPRESS=0.182)

Triparametric equations

log IC_{50}= +0.00103(±0.00147)Heat for - 0.25766(±0.10253)part coffi -0.00024(±0.00019) tot ene +1.37921(±0.818404) ……Eq.(8)

(n=20; r=0.878; s=0.148; F=40.756; Q^2=0.565; SPRESS=0.189)

From the above equations it is clear that in monoparametric analysis only the partition coefficient gives the significant equations and in the biparametric analysis the combination of partition coefficient with total energy and partition coefficient with electrical energy gives the significant equation. In triparametric analysis partition coefficient with heat of formation and total energy, partition coefficient with LUMO and electrical energy gives the significant correlation coefficient. On the bases of these significant results it can be concluded the partition coefficient is the descriptor, most prominent one to affect the biological activity studies. But the equation 9 gives the most significant results. By equation 9 it can be concluded that the descriptor LUMO with log P and Ed in trimetric equation having most prominent affect on the 12-LO inhibition activity.

Intercorrelation Matrix

	Hf	HOMO	LUMO	logP	PaCoff	TotEn	Ed	EleE	SEV
Hf	1	.022	.058	.051	.245	.024	.224	.017	.135
HOMO	1	.274	.012	.000	.028	.204	.044	.037	
LUMO	1	.063	.127	.089	.351	.03	.011		
logP	1	.644	0	.007	.000	.049			
PaCoff	1	.102	.149	.168	.393				
Tot En	1	.012	.85	.528					
Ed	1	.023	.043						
Ele E	1	.833							
SEV	1								

Table:2 Physicochemical descriptors of anthracenone derivatives

Com	Hfor	HOMO	LUMO	logP	P Coff	TE	Ed	EE	SEV
6a	-51.05	-9.08	-0.626	4.61	5.65	-3869.49	-0.376	-27514.4	249.16
6b	-7.57	-8.82	-0.662	6.29	7.53	-4692.24	-0.351	-36640.3	310.68
6c	-58.85	-9.05	-0.589	5.1	6.14	-4025.39	-0.42	-29387.3	265.95
6d	-206.67	-9.19	-0.781	5.53	6.53	-5440.16	-0.194	-36596.2	276.85
6e	-96.02	-9.13	-0.696	4.77	5.79	-4340.9	-0.208	-29694.9	252.69
6f	-58.13	-9.12	-0.682	5.17	6.63	-4229.6	-0.213	-29419	263.82
6g	-46.24	-9.11	-0.679	5.44	6.51	-4209.11	-0.222	-29376.5	269.11
6h	-19.39	-9.16	-0.75	4.64	5.08	-4190.01	-0.153	-30005.8	262.15
6i	-48.88	-9.08	-0.663	6.15	7.74	-5012.08	-0.133	-39146.6	321.69
6j	-101.72	-8.99	-0.585	5.31	6.62	-4656.96	-0.346	-35278.1	306.62
6k	-53.42	-8.99	-0.675	4.82	6.09	-4500.69	-0.639	-34025.3	290.81
6l	-89.2	-9.03	-0.599	4.48	5.56	-4345.38	-0.352	-31580.6	272.96
6m	-123.89	-8.97	-0.587	4.36	5.3	-4821.11	-0.554	-36198.7	297.16
6n	-127.81	8.92	-0.486	4.36	5.65	-4821.27	-2.004	-36524.4	296.11
6o	-124.69	-8.92	-0.673	4.36	5.65	-4821.14	-0.524	-36839.8	300.22
6p	64.65	-8.87	-0.751	4.04	7.02	-5485.53	1.65	-44838.4	341.18
6q	-127.13	-9.2	-0.788	4.43	5.61	-4793.93	5.97	-35205.4	291.61
6r	-140.82	-9.14	-0.712	4.17	5.39	-4638.87	1.24	-33004.4	270.59
6s	-95.52	-9.07	-0.607	4.22	4.98	-4190.12	-0.376	-29613.6	253.57
6t	-139.07	-8.89	-0.659	3.83	4.38	-4510.69	-1.665	-31827.9	257.74
6u	-139.74	-9	-0.592	3.83	4.26	-4510.67	-0.331	-31879.7	256.66
6v	-139.24	-8.93	-0.558	3.83	4.26	-4510.69	-1.84	-32155	256.33
6y	-45.59	-8.08	-0.619	4.45	5.29	-3780.1	0.714	-26112	242.9

Hfor=heat of formation, HOMO=highest occupied molecular orbital, LUMO=lowest unoccupied molecular orbitals, P.coff= partition coefficient, Ed= dipole-dipole energy, EE=electrical energy, TE=total energy, SEV= connollys solvent excluded volume.

Table:3 *insilico* toxicity prediction of anthraceone

SN	Toxicity	OverallToxicity	Onco	Muta	Tereto	Irrit	Senti	Immu	neuro
6a	probable	53	0	29	0	53	0	0	29
6b	Probable	53	0	29	0	53	0	0	29
6c	Probable	53	0	29	0	53	0	0	29
6d	Probable	53	0	29	0	53	0	0	29
6e	Probable	53	0	29	0	53	0	0	29
6f	Probable	53	0	29	18	53	0	0	29
6g	Probable	53	0	29	0	53	0	0	29
6h	Probable	53	0	29	0	53	0	0	29
6i	Probable	53	0	29	0	53	0	0	29
6j	Probable	53	0	29	0	53	0	0	29
6k	probable	53	0	29	0	53	0	0	29
6l	Probable	53	0	29	19	53	0	0	29
6m	Probable	53	0	29	19	53	0	0	29
6n	Probable	53	0	29	19	53	0	0	29
6o	Probable	53	0	29	19	53	0	0	29
6p	Probable	53	0	29	0	53	0	0	29
6r	Probable	53	0	29	17	53	0	0	29
6s	Probable	53	0	29	17	53	0	0	29
6t	Probable	53	0	29	0	53	0	0	29
6u	Probable	53	0	29	0	53	0	0	29
6v	Probable	53	0	29	0	53	0	0	29
6w	Probable	53	0	29	0	53	0	0	29

6x	Probable	53	0	29	0	53	0	0	29
6y	probable	53	0	29	0	53	0	0	29

Conclusion

From the results and the discussion above, we conclude that the physicochemical properties can be used successfully for modeling of 12 LO inhibitor activities of anthracenone derivatives and that, for the present set of anthracenone derivatives, the heat of formation, LUMO, log P, partition coefficient, total energy, Ed and Electrical energy are found to be the prominent to affect the biological activity studied. Among these descriptors the LUMO, log P and Ed are found to be the most prominent to affect thelipoxygense inhibition activity. The results also indicate that the combination of these descriptors and molecular (3D) modeling can be used for understanding the structural behavior and selecting the compound with potential activity. All the compounds are free from oncogenicity, sensitivity and immunogenisity and some compounds are free from teretogenisity.

Experimental section

Biological activity *(-logIC50)-* 12-Lipoxygenase Enzyme inhibition activity expressed as -logIC50, was taken from the literature.43

***Molecular descriptors*:** Physicochemical properties HOMO, LUMO, heat of formation, log P, partition coefficient, total energy, electrical energy, Ed, SEV used in proposed investigation, are calculated with help of Chem 3D Ultra 8.

3D-Molecular descriptors: The 3D molecular descriptors as RDF and 3D-Morse used in this study have been calculated applying the *DRAGON* program.

***Regression analysis*:** All the regressions were carried out using maximum R2 method with the help of Openstate-4 software package.

Insilico Toxicity: *Insilico* Toxicity of the anthracenone derivatives were predicted by computational method using Pallas version 3.1.1.2 ADME-Tox prediction software on Pentium-IV computer work station[21]

References

[1] Brash A.R., J. Biol. Chem., 1999, 274, 23679–23682.
[2] Kuhn H., Thiele B.J., FEBS Lett., 1999, 449, 7–11.
[3] Young R.N., Eur. J. Med. Chem., 1999, 34, 671–685.
[4] Ikai K., J. Dermatol. Sci., 1999, 21, 135–146.
[5] Zhu Y.I., Stiller M.J., Skin Pharmacol. Appl. Skin Physiol., 2000, 13, 235–245.
[6] Yamamoto S., Biochim. Biophys. Acta , 1992, 1128, 117–131.
[7] Funk C.D., Keeney D.S., Oliw E.H., Boeglin W.E., Brash A.R., J. Biol. Chem., 1996, 271, 23338–23344.
[8] Kinzig A., Fu¨ rstenberger G., Bu¨ rger F., Vogel S., Mu¨ ller- Decker K., Mincheva A., Lichter P., Marks F., Krieg P., FEBS Lett., 1997, 402, 162–166.
[9] Chang W.-C., Ning C.-C., Lin M.T., Huang J.-D., J. Biol.Chem., 1992, 267, 3657–3666.
[10] Takahashi Y., Reddy G.R., Ueda N., Yamamoto S., Arase S., J. Biol. Chem., 1993,268, 16443–16448.
[11] Hussain H., Shornick L.P., Shannon V.R., Wilson J.D., Funk C.D., Pentland A.P.,Holtzman M.J., Am. J. Physiol., 1994, 266, C243–C253.
[12] Zhao H., Richards-Smith B., Baer A.N., Green F.A., J. Lipid Res., 1995, 36, 2444–2449.
[13] S. T. Shukla and V. V. Sawant, Der Pharma Chemica, 2010, 2, 4, 200-205
[14] Vijay H. Masand, Komalsingh N. Patil, Devidas T. Mahajan, and Rahul D.Jawarkar, Ghulam M. Nazerruddin, Der Pharma Chemica, 2010, 2, 4, 298-308.
[15] M. C. Sharma, Smita Sharma, D. V. Kohli, S. C. Chaturvedi, Der Pharmacia Lettre, 2010, 2, 1, 150-161.
[16] Boeglin W., Kim R.W., Brash A.R., Proc. Natl. Acad. Sci. USA, 1998, 95, 6744–6749.
[17] Krieg P., Siebert M., Kinzig A., Bettenhausen R., Marks F., Fu¨ rstenberger G.,FEBS Lett., 1999, 446, 142–148.
[18] Gu J.L., Liu Y.X., Wen Y.S., Natarajan R., Lanting L., Nadler J.L., J. Cell. Physiol.,2001, 186, 357–365.
[19] Klaus Mu¨llera,, Reinhold Altmannb and Helge Prinza. Eur. J. Med. Chem., 2001,36, 569–575.
[20] Young D., Computational Chemistry: A Practical Guide for Applying Techniques to Real World Problems; John-Wiley & Sons: New York (USA), 2001.
[21]. Todeschini, R.; Consonni, V. DRAGON software for the Calculation of Molecular Descriptors, Web version 3.0 for Windows.
[22]. de Oliveira, D.B. and Gaudio, A.C. Quant. Struct.-Act. Relat. 2000, 19, 599.
[23] Pallas 3.1.1.2, ADME-Tox software, 2001, compu drug international Inc, USA..
[24] Raj K. Prasad, Rajesh Sharma, Der Pharma Chemica, 2010, 2, 2, 241-248.
[25] Fakhr M. Abu-Awwad, Der Pharma Chemica, 2010, 2, 1, 1-13.
[26] Aroop K. Saha, Nitish Gupta, Der Pharma Chemica, 2009, 1, 2, 133-144.

*****