Research Journal of Chemical Sciences ISSN 2231-606X
Vol. **4(6),** 18-22, June **(2014)** *Res. J. Chem. Sci.*

Comparative QSAR study of Anthracycline Analogues

Chincholikar P.M.[1], Amlathe S.[1] and Joshi S.[2]
[1]Department of Applied Chemistry, BUIT, Barkatullah University, Bhopal, MP, INDIA
[2]Department of Chemistry, SNGGPG College, Bhopal, MP, INDIA

Available online at: www.isca.in, www.isca.me
Received 19th March 2014, revised 3rd May 2014, accepted 10th June 2014

Abstract

Anthracycline, a potent anticancer molecule being used extensively in chemotherapy, though known for its efficacy - but known to cause adverse effects when administered. We are trying to analyse the variations in its analogues structurally. To study the QSAR of anthracycline analogues by using graph theoretical indices and physicochemical properties efforts have been made to derive a best mathematical model for the calculation of biological activity of the analogues in multivariate form. We have used multiple linear regression method. Two equations were developed separately, 1. Containing graph theoretical indices with r value 0.7038 and 2. Containing physicochemical properties as predictor variable with r value 0.8162 both models are significant to understand QSAR of anthracycline analogues.

Keywords: Anthracycline analogues, QSAR, drug designing.

Introduction

Anthracyclines (viz., Doxorubicin, Daunorubicin, Epirubicin, valirubicin, etc) are the anti-cancer drugs used in chemotherapy[1]. The use of these drugs have the common side-effects that were observed, in majority they are: nausea, alopecia, bone marrow suppression, and vomiting. Doxorubicin-induced bone marrow suppression can now be reduced by the use of hematopoietic growth factors. The *n*-octanol/water partition coefficient is the ratio of the concentration of a chemical in *n*-octanol to that in water in a two phase system at equilibrium. The logarithm of this coefficient, log *P*, has been shown to be one of the key parameters in quantitative structure activity/property relationship (QSAR/QSPR) studies[2,3]. The hydrophobicity and hydrophilicity of a substance is measured by *n*- octanol–water partition coefficient. It gives the idea of the tendencies of hydrophobic molecules or parts of molecules that avoid water, because they are not readily accommodated in the highly ordered hydrogen bonded structure of water[2]. Thermodynamically Hydrophobic interaction is favoured because of it increases the entropy of the water molecules that accompanies the association of non-polar molecules, which squeeze out water. There are some reports about the applications of MLR[3-6] and artificial neural network[7-10] modeling to predict the *n*-octanol/water partition coefficient of anti-cancer drugs. In some previous papers, reports on the application of QSAR techniques in the development of a new, simplified approach to prediction of compounds properties[11–17]. Study of log $P_{o/w}$ and its experimental determination is often complex and time consuming that can be done only for already synthesized compounds[11]. Hence, computational methods were applied for the prediction of this parameter.

The family composed of 23 substituted anthracyclines have been studied, and activity analyzed is the logIC_{50} (50% inhibitory concentration). The topological and physicochemical descriptors were studied for the investigation of QSAR of the compound. Multiple linear regressions were performed to obtain correlation for systematic study on substituents of anthracycline (table 1).

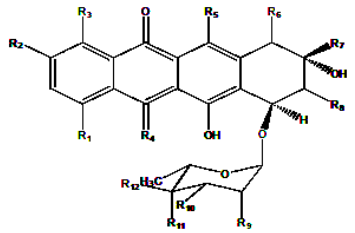


Figure-1
Parent structure of anthracyclines

Methodology

Two methodologies are adopted to develop mathematical models. In the first method graph theoretical descriptor are used for correlation with biological activity (logIC_{50}) and in the second method physicochemical properties are used for correlation with biological activity (log IC_{50}).

The multivariate regression analysis (MRA) was used to obtain a model used in QSAR studies. Typically in such studies, after selecting the compounds and the activity to be analysed, one considers selection of potentially useful descriptors. The major descriptor used in QSAR are divided into two classes: i. Graph theoretical descriptors, this class includes Electrotopological state(S), Balaban J Index(J), Szeged Index(Sz), Schultz molecular topological Index(MTI), Wiener Index(W), Randic connectivity Index(χ), etc. and ii. Physicochemical properties (descriptor), this includes logP, Molar refractivity (MR), Molar

volume(MV), Parachor(Pc), Index of refraction(IR), Surface tension(ST) etc. Indicator parameters are also used for the presentation of substitutent effect.

Descriptors	Type	Abbreviation
Graph theoretical descriptors	Wiener index	W
	Randic connectivity Index	χ
	Balaban J Index	J
	Szeged index	Sz
	Schultz molecular topological index	MTI
	Electro-topological state	S
Physicochemical descriptors	Partition Coefficient	Log P
	Molar refractivity	MR
	Molar volume	MV
	Parachor	Pc
	Index of refraction	IR
	Surface tension	ST

The topological indices in the study included: Wiener index (W)[16], Randic connectivity index (χ)[17], Balaban J index (J)[18], Szeged index (Sz)[19], Electro topological state (S)[20], Schultz Molecular topological index (MTI)[21]. Calculation of these indices were done using computer program Chemsketch 5.0 (from ACD labs).

The physicochemical properties were estimated using Chemsketch 5.0 (from ACD labs), while the logP values were calculated using program ChemlogP. Multiple Linear Regression method[22] was used for regression analysis.

Results and Discussion

The present study deals with the QSAR on benzodiazepines. The very first step is to calculate the desired topological indices, and the physicochemical properties and then investigate their utility.

The topological indices viz. logW, physicochemical properties, Electrotopological state (S) and Schultz Molecular topological index (MTI), were calculated using software ACD lab (Chemsketch 5.0) and are presented in the table 2. The indicator parameters are also given in table 2.

Table-1
Substituents on anthracyclines (ANTHs)

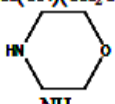
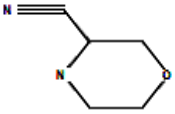
N	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}	R_{11}	R_{12}	exp
1	OCH_3	H	H	O	OH	H	$COCH_2OH$	H	H	NH_2	OH_{ndo}	H	1.27
2	OCH_3	H	H	O	OH	H	$COCH_3$	H	H	NH_2	OH	H	1.83
3	H	H	H	O	OH	H	$COCH_3$	H	H	OH	OH	H	0.9
4	OCH_3	H	H	O	OH	H	$COCH_2OCH_3$	H	H	NH_2	OH	H	1.37
5[a]	OCH_3	H	H	O	OH	H	$COCH_3$	H	H	$N(CH_3)_2$	OH	H	1.405
6	OCH_3	H	H	O	OH	H	$COCH_2OH$	H	H	NH_2	OCH_3	H	0.94
7	OCH_3	H	H	O	OH	H	$COCH_2OH$	H	H	NH_2	H	H	1.83
8	OCH_3	H	H	O	OH	H	$COCH_2OH$	H	H	NH_2	H	H	0.68
9	OCH_3	H	H	O	OH	H	$COCH_3$	H	H	OH	OH	H	1.34
10	OCH_3	H	H	O	OH	H	$COCH_2OH$	H	H	$N(CH_3)_2$	OH	H	1.56
11	OC_6H_5	H	H	O	OH	H	$COCH_3$	H	H	NH_2	OH	H	2.02
12	OCH_3	H	H	O	OH	H	$CHCH_3OH$	H	H	NH_2	OH	H	0.62
13[a]	OCH_3	H	H	O	OH	H	$COCH_2OH$	H	H	OH	OH	H	1.75
14	OCH_3	H	H	NH	OH	H	$COCH_3$	H	H	NH_2	OH	H	0.8
15	OCH_3	H	H	O	OH	H	$COCH_2OH$	H	H	$NH(CH(CN)(CH_2OCH_3))$	OH	H	0.92
16	OCH_3	H	H	O	OH	H	$COCH_2OH$	H	H		OH	H	0.42
17	OCH_3	H	H	O	OH	H	$COCH_2F$	H	H	NH_2	OH	H	0.72
18	OCH_3	H	H	O	OH	H	$C(NOH)(CH_3)$	H	H	NH_2	OH	H	0.479
19[a]	OCH_3	H	H	O	OH	H	$COCH_2OH$	H	H				0.286
20	H	H	H	O	OH	H	$COCH_2Br$	H	F	OH	OH	H	2.5007
21[a]	OH	H	H	O	OH	$COOCH_3$	CH_2CH_3	H	H	$N(CH_3)_2$	OH	H	2.234
22[a]	OCH_3	H	H	O	OH	H	$C(NNHCOC_6H_5)(CH_3)$	H	H	NH_2	OH	H	1.13
23	OCH_3	H	H	O	OH	H	$COCH_2OCO(CH_2)_3CH_3$	H	H	$NHCOCF_3$	OH	H	2.2

Table-2
Topological indices, physicochemical properties and indicator parameters of anthracyclines (ANTHs)

ANTHs	MTI	S	logW	logP	ST	Pol	I_2	I_1	$I_{3,8}$	$I_{1,7}$
1	3173	48.35	2.8609	1.176	45.4	28.09	0	0	0	0
2	3173	49.85	2.8609	0.832	42.3	30.20	0	1	0	1
3	3543	56.02	2.9133	1.424	42.9	28.03	1	0	0	0
4	3543	57.52	2.9133	1.08	40.2	30.14	1	1	0	1
5	3173	44.46	2.8609	1.479	49.6	29.96	0	0	0	0
6	3173	49.85	2.8609	1.231	46.1	32.07	0	1	0	1
7	3543	52.13	2.9133	1.823	47.0	29.91	1	0	0	0
8	3543	48.24	2.9133	2.222	51.1	31.78	1	0	0	0
9	3543	51.64	2.9133	2.126	47.5	35.21	1	1	0	1
10	3543	65.6	3.0298	1.936	40.4	30.02	0	0	0	0
11	4173	45.95	2.9795	0.310	60.6	30.38	0	0	0	0
12	4523	53.62	3.0265	0.902	57.2	30.33	1	0	0	0
13	4523	55.12	3.0265	0.558	52.5	32.44	1	1	0	1
14	4523	49.73	3.0265	1.301	61.8	32.20	1	0	0	0
15	4523	51.23	3.0265	0.957	56.5	34.32	1	1	0	1
16	6119	70.87	3.1658	1.662	50.0	32.26	1	0	0	0
17	5111	58.02	3.0824	-0.58	54.4	32.16	1	1	0	1
18	4173	44.35	2.9795	-0.889	49.3	29.30	0	0	0	0
19	4173	45.85	2.9795	-0.71	48.7	31.26	0	1	0	1
20	4523	53.52	3.0265	-0.462	47.2	30.96	1	1	0	1
21	4523	48.13	3.0265	0.281	55.7	30.72	1	0	0	0
22	3629	50.35	2.9238	-0.058	48.5	32.83	0	1	0	1
23	4029	58.02	2.7931	0.19	46.1	32.78	1	1	0	1

I_1 = Indicator Parameter 1 If substituents on R_1, I_2 = Indicator Parameter 1 If substituents on R_2', $I_{3,8}$ = Indicator Parameter 1 If substituents on R_3 or R_8, $I_{1,7}$ = Indicator Parameter 1 If substituents on R_1, R_7 or on both.

The correlations that are low in values of R (<0.50) are not considered being statistically insignificant. Individually indices were correlated in lesser manner with the biological activity ($logIC_{50}$) and same was the case with physicochemical properties of observed biological activity.

The results show that the topological, physicochemical properties and biological activity when studied for univariate correlation was insufficient to describe the SAR (structure activity relationship) in quantitative manner. All the univariate correlation gives very low correlation coefficient. Out of all univariate correlations the best statistics are shown by S (r = 0.2105, s = 0.7321).

The correlation coefficient in the case of bivariate correlation showed little higher but insufficient to explain structure activity relationship quantitatively and the best statistics are shown by MTI and I2 (r = 0.6567, s = 0.5631), so is the case with trivariate correlation and best statistics are shown by MTI, log W and I2 (r =0.7038, s = 0.5492). In case of tetravariate combination, the best statistics are shown by MTI, logW, I2 and I1,7 (r =0.8162, s = 0.5468). In case of pentavariate combination, the best statistics are shown by MTI, logW, S, I2 and I1,7 (r = 0.7019, s = 0.549). In cases of hexavariate correlation the results are quite good and the correlation coefficient r (0.7063) is obtained as a best one from the correlation of MTI, S I2, I1, I1,7 and logW with the BA (biological activity). The low value of standard deviation (0.5479) and high value of F-ratio (8.129) supports the findings.

The mathematical model from above variables is shown in equation (1).

$$logIC_{50} = 7.3644\ 10^{-4}\ MTI - 0.0173\ S - 0.9519\ I_1 - 0.3466\ I_2 + 0.5664\ I_{1,7} - 3.508$$
$$LogW + 10.0314 \qquad (1)$$

The estimation of $logIC_{50}$ values were essential to confirm our findings between the best suited model and the observed values. The observed and calculated $logIC_{50}$ values obtained from equation (1) are presented in table III.

Using the equation (1) we can describe the effect of substituents. It is observed that the substituents at 1 and 7 positions and together increase the biological activity, while the substituents at 1 and 2 positions have a negative impact on the activity quantitatively.

Similarly, in case of physicochemical properties best results of univariate correlation are shown by logP (r = –0.574, s = 0.6056). Similarly for bivariate correlation the best correlation coefficient is shown by logP and I_2 (r = 0.6776, s = 0.5491) and

the best results for trivariate correlation are shown by logP, Pol and I_2(r = 0.7186, s = 0.5242). In the case of tetravariate correlation compatible results are obtained by the correlation of logP, ST, Pol, and I_2 (r = 0.7695, s = 0.4860). The pentavariate correlation shows more compatibility than the tetravariate correlation by logP, ST, Pol, I_2, and $I_{3,8}$ (r = 0.7948, s = 0.4665). For the hexavariate correlation the results are encouraging but the better results are obtained with I_2, $I_{3,8}$, $I_{1,7,}$ logP, surface tension (ST) and polarizability (Pol) with the r value of 0.8153. The high value of F-ratio (16.192) and the low value of standard error (0.4496) supports the above finding.

The mathematical model obtained from above variables can be given by equation (2):

$$\log IC_{50} = -0.5520 \log P - 0.0628\ ST + 0.1792\ Pol - 0.6754\ I_2 + 0.4947\ I_{3,8} - 0.4340\ I_{1,7} - 0.1839 \quad (2)$$

The observed and calculated values of $\log IC_{50}$ obtained from equation, (2) are presented in table 3.

Using the equation (2) we can describe the effect of substituent. It is observed that the substituent at 3 or 8 position increases the biological activity, whereas the substitutions at position-2 and together at 1 and 7-position have a negative impact on biological activity.

The results are little contrary to the one observed with the graph theoretical descriptors but can very well be seen that the negative impact is comparatively very low in case of substituent at positions-1 and 7.

Conclusion

The indicator parameters were successfully used to explore the substituent effect. On the basis of equation (2) we conclude that the substituent at 2nd position has a negative impact on the biological activity quantitatively, i.e, presence of substituent at R_2 position decreases the biological activity ($\log IC_{50}$) and the substituent at 3rd or 8th position show the direct relationship and will help increase in biological activity quantitatively, i.e., the presence of the substituent at positions-3 and 8 play a significant role to enhance the biological activity ($\log IC_{50}$). Based on the results and discussion made above conclusion may be drawn - no single molecular descriptors or physicochemical properties could be used individually for QSAR studies. In multivariate correlations more than the graph theoretical descriptors the physicochemical properties have shown significant correlation. Thus, physicochemical property studies are most suited for understanding QSAR of anthracyclines.

Table-3
Observed and calculated $\log IC_{50}$ of substituted anthracyclines (ANTHs)

ANTHs	$\log IC_{50}$(Obs.)	$\log IC_{50}^{a}$	Residual	$\log IC_{50}^{b}$	Residual
1	1.602	1.4956	0.1063	1.4366	0.1653
2	1.23	1.0841	0.1458	1.7131	-0.4836
3	0.869	1.105	-0.236	0.7767	0.0922
4	0.708	0.6935	0.0144	1.0548	-0.3468
5	0.973	1.5629	-0.5899	1.4095	-0.4365
6	0.908	1.0841	-0.1761	1.6672	-0.7592
7	0.301	1.1722	-0.8712	0.721	-0.42
8	0.255	1.2395	-0.9845	0.6805	-0.4255
9	0.462	0.7952	-0.3332	0.6001	-0.1381
10	1.114	0.8771	0.2368	1.5741	-0.4601
11	1	1.8575	-0.8575	1.3394	-0.3394
12	0.176	1.4711	-1.2951	0.5934	-0.4174
13	0.58	1.0596	-0.4796	0.9646	-0.3846
14	0.255	1.5384	-1.2834	0.5339	-0.2789
15	0.342	1.1269	-0.7849	0.9296	-0.5876
16	0.544	1.8593	-1.3153	0.8485	-0.3045
17	1.982	1.2464	0.7355	1.3508	0.6311
18	2.587	1.8852	0.7017	2.3358	0.2511
19	2.663	1.4737	1.1892	2.2361	0.4268
20	1.813	1.0873	0.7256	1.4195	0.3934
21	1.875	1.566	0.3089	1.0361	0.8388
22	2.58	1.1906	1.3893	2.276	0.3039
23	1.477	1.4644	0.0125	1.6139	-0.1369

Reference

1. Shaul P., Frenkel M., Breiner-Goldstein E., Mittelman L., Grunwald A., Ebenstein Y., Tsarfaty I., Fridman M., The structure of anthracycline derivatives determines their sub-cellular localization and cytotoxic activity, *ACS Med. Chem. Lett.*, **4,** 323–328 **(2013)**

2. Breiner-Goldstein E., Evron Z., Frenkel M., Cohen K., Meiron K.N., Peer D., Roichman Y., Flescher E. and Fridman M., Targeting Anthracycline-Resistant Tumor Cells with Synthetic Aloe-Emodin Glycosides, *ACS Med. Chem. Lett.***, 2,** 528–531 **(2011)**

3. Ajeet Brajpal Singh, Vipul Kumar, qsar modeling of 2-[ch(oh)x]-5,8-(oy)2-1,4-naphthoquinines against l1210 cells using multiple linear regression, *Indonesian J. Pharm*., **23(3)**, 171–176 **(2012)**

4. Salum L., Andricopulo A., Fragment-based QSAR: Perspectives in drug design, *Mol. Divers.,* **13,** 277–285 **(2009)**

5. Anthracycline Chemistry and Biology: Biological Occurence and Biosynthesis, Synthesis and Chemistry: No. *1*. 1 (9 July **2008**) ed. Topics in Current Chemistry, ed. K. Krohn: Springer **(2008)**

6. Esposito E.X., Hopfinger A.J. and Madura J.D., Methods for applying the quantitative structure-activity relationship paradigm, *Methods Mol. Biol.,* **275**, 131–213 **(2004)**

7. Minotti G., et al., Anthracyclines: molecular advances and pharmacologic developments in antitumor activity and cardiotoxicity, *Pharmacol Rev,* **56(2)**, 185-229 **(2004)**

8. Perkins R., Fang H., Tong W. and Welsh W., Quantitative structure-activity relationship methods: perspectives on drug discovery and toxicology, *Environ. Toxicol. Chem.,* **22**, 1666–1679 **(2003)**

9. Studzian K., et al., Inhibition of RNA synthesis in vitro and cell growth by anthracycline antibiotics, *Neoplasma*, **48(5)**, 412-8 **(2001)**

10. Rudolph U. Koop C and Keist R, *Nature*, 401, 796 **(1999)**

11. Winkler D.A., Burden F.R. and Watkins J.R., *Quant Struct-Act Relat*, 17, 14 **(1998)**

12. Sigel E. and Buhr A., *TIPS,* 18, 425 **(1997)**

13. Khadikar P.V. and Gutman I., *J Serb Chem Soc*, **62**, 235 **(1997)**

14. McKernan R.M. and Whiting P.J., *Trends in Neurosci*, **19**, 139 **(1996)**

15. Chaires J.B., et al., Parsing the free energy of anthracycline antibiotic binding to DNA*, Biochemistry*, **35(7)**, 2047-53 **(1996)**

16. Macdonald R.L. and Olsen R.W., *Ann Rev Neurosci*, **17,** 569 **(1994)**

17. Randic M., *Croat Chem Acta,* **66,** 289 **(1993)**

18. Villar H.O., Davies M.F., Loew G.H. and Maguire P.A., *Life Sci*, **48**, 593 **(1991)**

19. Schultz H.P., Schultz E.B. and Schultz T.P., *J Chem Inf Comput Sci*, 30, 27 **(1990)**

20. Olsufyeva E.N. and Rozynov B.V., [Synthesis and massspectrometric analysis of N-cycloalkyl derivatives of carminomycin, daunorubicin and their analogs], Bioorg Khim(Russ), **16**, 847-853 **(1990)**

21. Vincenzo Rizzo1, Carlo Battistini1,et al, Association of anthracyclines and synthetic hexanucleotides, Structural factors influencing sequence specificity, *Journal of Molecular Recognition,* **2(3),** 132–141 **(1989)**

22. Balaban A.T., *J Mol Struct (Theochem),* **165**, 243 **(1988)**

23. Kier L.B. and Hall L.H., Molecular Connectivity in Chemistry and Drug Research*,* (Academic Press, New York), **(1976)**

24. Krueger W.C., et al., The interaction of nogalamycin and analogs with DNA and other biopolymers, *Chem Biol Interact*, **36(1)**, 1-18 **(1981)**

25. Wiener H., *J Am Chem Soc,* **69,** 17 **(1947)**