Indian Journal of Chemistry
Vol. 41B, February 2002, pp. 436-439

Note

QSAR studies on anti HIV-1 *N*-substituted betulinic acid amides

Arun K Srivastava*, Mohd Shakeel & Arbab A Khan
Department of Chemistry, University of Allahabad, Allahabad-211002,India.

Received 19 July 1999, accepted (revised) 22 January 2001

QSAR studies, using structural parameters such as equalized electronegativity 'Xeq', molecular connectivity '$^{1}\chi^{b}$' and hydrophobicity 'logP' on a novel series of ω-amino alkanoic acid derivatives of betulinic acid, have been discussed. Anti HIV-1 activity of these compounds both in CEM-4 and MT-4 cell cultures is found to correlate well with steric $^{1}\chi^{b}$ as well as electronic Xeq parameters. The presence of an amide function is found to be important for activity.

HIV has drawn global attention of both medical practitioners and medicinal chemists. Glycyrrhetinic acids[1] have shown limited activity against a wide spectrum of viruses including HIV-1 type-1. Salaspermic acid[2] and suberol[3] are known to inhibit HIV-1 in H9 cells. Bile acid derivatives[4] were found slightly active (at 10^{-4} *M*) against HIV in MT-4 cells. Francoise Soler and coworkers[5,6] have shown that the derivatives of betulinic acid are potent inhibitors of the cytopathicity of HIV-1 both in CEM-4 and MT-4 cells. For the synthesis of these derivatives, betulinic acid (**Figure 1**) has been used as the common starting material. This new class of anti HIV-1 compounds seems to block a post binding event involved in the virus-cell fusion.

For the purpose of present study the betulinic acid derivatives have been devided into two series:

(a) ω-amino alkanoic acid amides of betulinic acids (**Table I**).

(b) (Betulinylamino)- octanoic acid amides of α,β and γ amino acids (**Table II**).

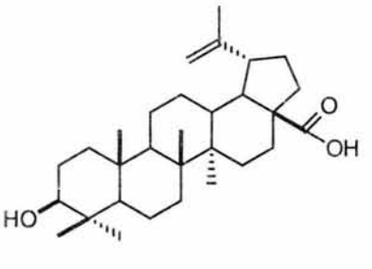


Figure 1 – Betulinic acid

The electronic parameter i.e. equalized electronegativity 'Xeq' was evaluated using Pauling [7] formula, the steric parameter first order molecular connectivity $^{1}\chi^{b}$ was calculated as defined by Kupchik[8,10] and the partition coefficient 'logP' was calculated by the fragmental method developed by Nys and Rekker[11].

Results and Discussion

First QSAR studies on ω-amino alkanoic acid amides of betulinic acid derivatives are discussed.

All the eighteen analogues in this series are listed in **Table I** along with their physicochemical data and biological activity. In this **Table I** $\log(1/IC_{50})$ i.e. $p(I_1)$ represents the anti HIV-1 activity against CEM-4 cells while $p(I_2)$ represents activity against MT-4 cells. IC_{50} (50% cell culture inhibitory concentrations) are defined as those which inhibit HIV-1 induced cytopathicity by 50%.

Table III contains the correlation matrix with $p(I_1)$activity showing correlation among the structural parameters in this series of compounds and it is clear that auto correlation exists between Xeq and logP.

Regression analysis of $p(I_1)$ with Xeq, $^{1}\chi^{b}$ and logP gave the following significant correlation,

$$p(I_1) = 0.587\ (\pm 0.178)\ {}^{1}\chi^{b} - 8.819 \quad \ldots (1)$$

n = 18, EV = 40.4%, SEE = 0.674, $F_{(1,16)} = 10.840$

where 'n' represents the number of data points, 'r' the correlation coefficient, 'F' the variance ratio, 'SEE' the standard error of estimate and data within the parentheses represents the confidence interval of regression constant at 95% level.

Introduction of an indicator parameter 'Ind' whose value was taken 1 for the presence of -NHCO group at Y and zero otherwise, resulted in the following equations.

$$p(I_1) = 0.493\ (\pm\ 0.162)\ {}^{1}\chi^{b} + 0.918\ (\pm 0.387)\ \text{Ind} - 7.197 \quad \ldots (2)$$

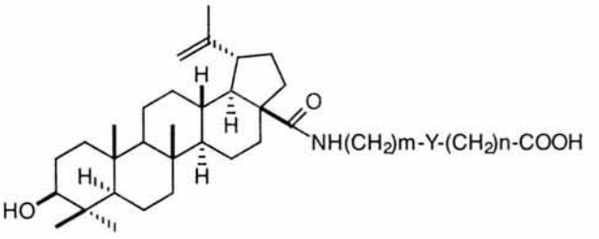
n = 18, EVE = 56.6%, SEE = 0.593, $F_{(2,15)} = 9.80$

$$p(I_1) = 0.456\ (\pm 0.158)\ {}^{1}\chi^{b} + 0.169\ (\pm 0.119)\ \text{logP} + 1.011\ (\pm 0.380)\ \text{Ind} - 7.221 \quad \ldots (3)$$

n= 18, EVE = 62.1%, SEE = 0.574, $F_{(2,15)} = 7.656$

correlation of $p(I_1)$, with $^{1}\chi^{b}$ and LogP, gave a statistically acceptable regression equation (3) with 62.1%

Table I—Biological activities and physico-chemical data for ω-amino alkanoic acid amide of betulinic acid derivatives



Compd	m	n	Y	Xeq	$^1\chi^b$	logP	$p(I_1)$	$p(I_2)$	$p(I_1)$ Calculated		$p(I_2)$ Calculated	
									Eq(2)	Eq(3)	Eq(4)	Eq(6)
A1	5	0	—	2.240	16.900	2.700	1.639	0.620	1.140	1.088	1.078	0.561
A2	6	0	—	2.239	17.400	3.240	0.921	0.367	1.386	1.412	1.053	0.706
A3	7	0	—	2.239	17.900	3.780	2.125	1.333	1.633	1.736	1.053	0.869
A4	8	0	—	2.238	18.400	4.320	2.377	0.469	1.880	2.059	1.028	1.015
A5	9	0	—	2.237	18.900	4.860	2.222	1.377	2.126	2.383	1.033	1.160
A6	10	0	—	2.237	19.400	5.400	2.639	1.354	2.373	2.707	1.033	1.324
A7	11	0	—	2.236	19.900	5.940	2.260	1.603	2.620	3.031	0.978	1.469
A8	3	5	CONH	2.281	19.003	1.976	1.260	—	2.177	1.942	—	—
A9	4	4	CONH	2.281	19.003	1.490	1.347	—	2.177	1.859	—	—
A10	5	3	CONH	2.281	19.003	1.872	1.523	—	2.177	1.924	—	—
A11	6	2	CONH	2.281	19.003	2.483	1.347	—	2.177	2.028	—	—
A12	7	1	CONH	2.281	19.003	3.094	2.284	1.807	2.177	2.132	2.105	1.985
A13	7	2	CONH	2.280	19.503	3.023	3.000	2.207	2.424	2.352	2.080	2.130
A14	7	3	CONH	2.278	20.003	2.952	3.357	1.890	2.671	2.572	2.030	2.258
A15	7	4	CONH	2.277	20.503	3.110	3.215	—	2.917	2.831	—	—
A16	7	1	NHCO	2.281	19.003	3.094	3.302	2.469	3.095	3.143	2.105	1.985
A17	7	2	NHCO	2.280	19.503	3.023	3.023	2.276	3.341	3.363	2.080	2.130
A18	7	3	NHCO	2.278	20.003	2.570	3.699-	—	3.588	3.518	—	—

Note: Cal. pI_1 and Cal pI_2 corresponding to the activities calculated from eqns.2, 3, 4 & 6 respectively.

explained variance. It seems that the topological parameter $^1\chi^b$ and logP together play an important role in determining the activity of this class of anti HIV-1 agents against CEM-4 cells. It is noteworthy that in all these equation Ind parameter has positive regression coefficients and therefore it may be suggested that the presence of a -NHCO group will be favourable for increasing the activity of the drug. Solar *et al.* have also reported that increasing the amide moiety with the chain were found to be somewhat potent as compared to those having -CONH.

Table IV contains the correlation matrix between structural parameters Xeq, $^1\chi^b$ and logP for the twelve compounds whose activity $p(I_2)\{p(IC_{50})$ against MT-4 cell cultures} is known and it can be seen that some autocorrelation exists between Xeq and logP parameters.

QSAR on the data given in the **Table II** gave the following regression equations,

$$p(I_2) = 25.051\ (\pm 6.235)\ Xeq - 55.036 \quad \ldots (4)$$

$n = 12$, EV = 61.8%, SEE = 0.448, $F_{(1,10)} = 16.141$

$$p(I_2) = 0.516\ (\pm 0.151)\ ^1\chi^b - 8.242 \quad \ldots (5)$$

$n = 12$, EV = 53.9%, SEE = 0.492, $F_{(1,10)} = 11.707$

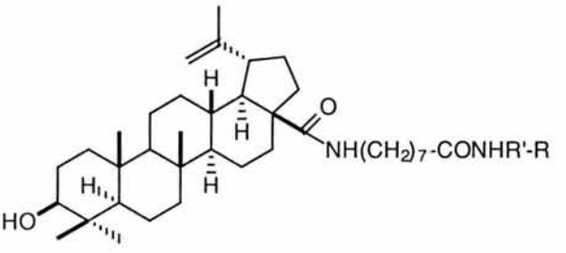
Correlation level was found to improve when both Xeq and $^1\chi^b$ were considered together.

$$p(I_2) = 17.969\ (\pm 5.613)\ Xeq + 0.327\ (\pm 0.124)\ ^1\chi^b - 45.217 \quad \ldots (6)$$

$n = 12$, EV = 78.5%, SEE = 0.354, $F_{(2,9)} = 16.391$

Equations (4), (5) and (6) make it clear that the first order molecular connectivity $^1\chi^b$ and equalized electronegativity Xeq play an important role in determining the activity. Incremental lengthening of the chain in compounds (A1-A7) also shows that significant activity was observed for the compounds positioned between A3 to A7 i.e. between the betulinylamino octanoic and (betulinylamino) dodecanoic acids.

Table II—Biological activities and physico-chemical data for (betulinylamino)-octanoic acid amides of α- amino acids, β-amino acids and γ-amino acids:



Compd	RNHR'	Xeq				pI_3 Calculated	
			$^1\chi^b$	Log p	pI_3	Eq.7	Eq.8
B1	Ala	2.280	19.454	-1.366	1.203	1.287	1.183
B2	$NH_2C(Me)2COOH$	2.278	19.804	3.080	1.051	1.271	1.169
B3	Ser	2.189	19.601	-2.225	0.603	0.529	0.534
B4	Sarcosin	2.189	19.360	-1.097	0.478	0.529	0.534
B5	Pro	2.280	20.600	-1.706	1.237	1.287	1.183
B6	$raCH_2CH(C_6H_5)CH_2COOH$	2.283	21.547	-0.197	1.328	1.312	1.204
B7	$(S,S)NH_2CH(i\text{-}Bu)CHOHCH_2COOH$	2.278	21.911	-0.434	1.398	1.271	1.403
B8	$raCNH_2CH_2CHOHCH_2COOH$	2.285	20.077	-2.564	1.482	1.329	1.453
B9	$(R)\text{-}NH_2CH(i\text{-}Bu)CH_2CH_2COOH$	2.278	21.820	-0.468	1.071	1.271	1.169
B10	$(S,S)NH_2CH(BZ)CHOHCH_2COOH$	2.286	22.611	-0.484	1.482	1.337	1.460
B11	$(3R,YS)NH_2CH(i\text{-}Bu)CHOHCH_2COOH$	2.278	21.911	-0.434	1.357	1.271	1.403

Note: Cal. pI_3 correspond to the activities calculated from eqns.7 & 8 respectively.

Table III—Correlation matrix of 18 compounds set with activity pI_1

	Xeq	$^1\chi^b$	logP
Xeq	1.000		
$^1\chi^b$	0.500	1.000	
logP	-0.751	0.087	1.000

Table IV—Correlation matrix of 12 Compounds set with activity pI_2

	Xeq	$^1\chi^b$	logP
Xeq	1.000		
$^1\chi^b$	0.477	1.000	
logP	-0.644	0.316	1.000

For the purpose of performing QSAR studies on the (betulinylamino)octanoic acid amides of α-, β- and γ- amino acids, the calculated physicochemical data and anti HIV-1 activity $p(I_3)$\{$p(\log 1/IC_{50})$ against MT-4 cells\} have been listed in **Table II**. The correlation matrix for the structural parameters in **Table V** shows the absence of any autocorrelation between $^1\chi^b$, Xeq & logP.

The regression analysis gave the following statistically significant relationship between $p(I_3)$ and Xeq.

Table V—Correlation matrix of 11 Compounds set with activity pI_3

	Xeq	$^1\chi^b$	logP
Xeq	1.000		
$^1\chi^b$	0.549	1.000	
logP	0.282	0.150	1.000

$$p(I_3) = 8.334\ (\pm 1.179)\ Xeq - 17.715 \qquad \ldots (7)$$

$n = 11$, $EV = 84.7\%$, $SEE = 0.139$, $F_{(1,9)} = 49.904$

This correlation explains 84.7% variation in the activity.The standard error of estimate is low and F-ratio is fairly high. The 't-test' also suggest that this relationship is significant at 95% confidence level [t cal.7.064 > t 0.05 = 2.262].

When for the compounds in which $CH(OH)CH_2COOH$ group is present in variant position (RNHR') of the substituent (compds. B7, B8, B10 & B11) an indicator parameter Ind1 was assigned a value 1 otherwise it was treated as 0, a highly significant regression equation was obtained.

$$p(I_3) = 7.132\ (\pm 0.752)\ Xeq + 0.234\ (\pm 0.055)\ Ind._1 - 15.078 \qquad \ldots (8)$$

$n = 11$, $EV = 95.3\%$, $SEE = 0.082$, $F_{(2,8)} = 80.635$

For this equation explained variance is very high at 95.3%, standard error of estimate SEE is very low and F-ratio is significant at 95% confidence level.

It can therefore be suggested that Xeq plays a dominant role in determining the activity pI_3 against MT-4 cells. In future designing of more potent drugs in this class one may look for the compounds with higher equalized electronegativity having $-CH(OH)CH_2COOH$ group in variant portion (RNHR') of the substituent.

Acknowledgement

This research was supported by Grant (1648)/00/EMR II awarded by CSIR, New Delhi.

References

1 (a) Ishida S, Sakiya Y, Ichikawa T & Awasu S, *Chem Pharm Bull*, 37, **1989**, 2413. (b) Panpei R, Flore O, Marccialis M A, Pani A & Loddo B, *Nature*, 218, **1979**, 689. (c) Ito M, Nakashima H, Baba M, Pauwels R, De clercq E, Shigeta S & Yamamoto N, *Antiviral Res*, 7, **1987**, 127. (d) Ito M, Sato A, Hirabayashi K, Tanabe F, Shigeta, Baba M, De clercq E, Nakashima H & Yamamoto N, *Antiviral Res*, 10, **1988**, 289.

2 Chen K, Shiq, Kashiwada Y, Hang D C, Hu CQ, Jin JQ, Nozaki H, Kilkuskiere, Tramontano E, Cheng Y C, McPhail D R & Lee K H, *J Nat Prod*, 55, **1992**, 340.

3 LiHY, Sun N J, Kashiwada Y, Sun L, Snider J V, Cosentino L M & Lee K H, *J Nat Prod*, 56, **1993**, 1130.

4 Baba M, Schols D, Nakashima H, Pauwels R, Parentier G, Meijer D K F & Declercq E, *J Acquired Immune Defic Syndr*, 2, **1989**, 264.

5 Soler F, Poujade C, Ever M, Carry J C, Henin Y, Boussean A, Huet T, Pauwels R, De clercq E, Mayaux J F, LePecq J B & Dereu N, *J Med Chem*, 39, **1996**, 1069.

6 Evers M, Poujade C, Soler F, Ribeill Y, James C, Lelievre Y, Guegnen J C, Reisdorf D, Morize I, Pauwels R, De clercq E, Henin Y, Bousseau A, Mayaux J F, Le Pecq J B & Deren N, *J Med Chem*, 39, **1996**, 1056.

7 Pauling L, The nature of chemical bond (Cornell University Press, Itacha, N Y), Chapter 2, **1939**.

8 Kupchick E J, *Quant Struct Act Relat*, 4, **1985**, 123.

9 Kupchick E J, *Quant Struct Act Relat*, 5, **1986**, 95.

10 Kupchick E J, *Quant Struct Act Relat*, 7, **1988**, 57.

11 Nys G C & Rekker R F, *Chem Therap*, 8, **1973**, 521.