Analysis of Tetrahedral Carbon in QSAR Studies. A Case Study Using 3-Hydroxy-3-methylglutaryl-Coenzyme A Reductase Inhibitors†

Yenamandra S. Prabhakar

Medicinal Chemistry Division, Central Drug Research Institute, Lucknow-226001, India

Received April 23, 1998

In this article an attempt has been made to outline a rule-based procedure to address the groups and atoms attached to the tetrahedral carbon in a quantitative manner. For this, computation of a few conceptual parameters, namely, flexibility of rotation, probability of availability, and net detachability, have been defined for the atoms/groups attached to the tetrahedral carbon. This has been used in a case study to explore structure−activity relations in some 3-hydroxy-3-methylglutaryl-coenzyme A reductase inhibitors.

INTRODUCTION

Biological macromolecules, for example enzymes, generally exhibit selectivity in expressing optimum response when approached by a specific optical isomer and minimum or no response when approached by the optical antipode of the active analogue. In organic compounds tetrahedral carbons are the major source of chirality. Except for their response to plane polarized light, optical isomers exhibit the same or even identical physicochemical properties, viz., hydrophobicity, solubility, steric, electronic, etc. Often this creates difficulty in quantitatively expressing the tetrahedral carbon, and atoms/groups attached to it, especially when the center is asymmetric. Consider M−R as the molecule of interest. Schematically molecule M−R can be represented as shown in Figure 1, where R represents the substituent group on the main structural frame (M) of the molecule. In Figure 1b, C_M represents the atom of the main structural frame (M) which is adjacent to R, C_T is the tetrahedral carbon of R which is directly bonded to C_M, and R_a, R_b, and R_c are the atoms/structural units growing out of C_T. When the main frame of this kind of molecule is anchored in a, or approaching a restricted position of the surrounding environment, for example, an enzyme, any rotation of R on the C_M−C_T bond results in the displacement of R_a, R_b, and R_c in space. As a result of this displacement and its ease, they offer different degrees of steric, hydrophobic, and electronic interaction with their surroundings. However, it is very difficult or even impracticable to study this kind of situation in nature, but some explanation and partial solution can be offered by considering the spatial disposition and steric restrictions of the atoms/structural units of the tetrahedral carbon in question. In this article, we attempt to outline a rule-based procedure to address the atoms/structural units (hereafter called structural units) attached to the tetrahedral carbon, whether it is symmetric or asymmetric, in a quantitative manner. As a step in this direction, we define a new parameter, net detachability (ND), to parameterize the ease of rotation and accessibility of R. The parameter ND is a ratio of rotational flexibility of R on the C_M−C_T bond and probability of availability of R_a, R_b, and/or R_c coupled with a constant phase difference between *R*/*S* isomeric centers, wherever necessary. Here, in a case study, the introduced

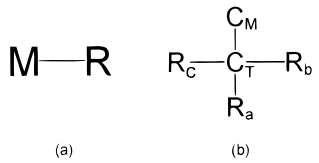

Figure 1. (a) In M−R, M represents the main structural frame of the molecule, and R is the substituent group directly attached to M. (b) C_M is the atom of the main structural frame (M) adjacent to R. C_T is the tetrahedral carbon of R directly bonded to C_M; R_a, R_b, and R_c are the atoms/structural units growing out of C_T.

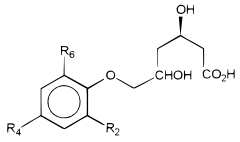

Figure 2. General structure of some HMGR inhibitors used in this study.

concept has been used to parameterize the flexibility of rotation with availability of the varying substituent group, R6, of 6-aryloxy-3,5-dihydroxyhexanoic acids (Figure 2).[1] Further, this has been used as an independent parameter in the correlation analysis of 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGR) inhibitory activity of these compounds.

METHODOLOGY

To start with, the disposition of structural units of the substituent group (R) of molecule M−R is defined below. First, fragment the R into tetrahedral carbon (C_T) and the three structural units, R_a, R_b, and R_c, attached to C_T and represent them as shown in Figure 1, where R_b and R_c are to the right and left side of R_a, respectively. Now, among the three structural units which are part of R, assign R_a the structural unit with the lowest van der Waals volume. After this, assign the remaining two structural units of R (R_b and R_c) using the *R*/*S*-sequence rules to represent the actual

† This article is based on material presented at First Indo-US Workshop on Mathematical Chemistry, January 9−13, 1998, Visva-Bharati, Santiniketan, India, sponsored by the University of Minnesota, USA, CSIR, Visva-Bharati University, and DST, India. C.D.R.I. Communication No. 5810.

10.1021/ci980047s CCC: $18.00
Published on Web 06/10/1999

Table 1. The physicoehemical, ND Values, and HMGR Inhibitory Activity of 6-Aryloxy-3,5-dihydroxyhexanoic Acids (Figure 2)

structure no.	R_2	R_4	R_6	$\Sigma\pi$[a]	VwR_2	VwR_6	NDR_b	NDR_c	Obs[b]	$-\log IC_{50}$ Prd eq 6	Prd eq 8
1	Me	Me	CH(*p*-phF)$_2$	5.03	0.167	1.597	0.439	0.439	7.15	7.27	7.29
2	Cl	Cl	CH(*p*-PhF)$_2$	5.33	0.178	1.597	0.439	0.439	7.30	7.24	7.27
3	Me	Cl	CH(*p*-PhF)$_2$	5.18	0.167	1.597	0.439	0.439	7.15	7.27	7.29
4	Me	Me	CH(*p*-PhF)(*p*-PhOMe)	4.82	0.167	1.768	2.156	2.156	7.00	6.90	6.81
5	Me	Me	CH(*p*-PhF)(*p*-PhCF$_3$)	5.77	0.167	1.843	2.138	2.138	7.05	6.85	6.81
6	Me	Me	CH(*p*-PhF)Me	3.88	0.167	1.011	2.643	2.643	6.80	6.47	6.68
7	Me	Me	CH(*p*-PhF)-*i*-Bu	5.37	0.167	1.374	2.315	2.315	6.85	6.90	6.77
8	Cl	Me	CH(*m*-Me,*p*-F-Ph)$_2$	6.50	0.178	1.905	3.690	3.690	6.57	6.22	6.37
9	Cl	Cl	CH(*m*-Me,*p*-F-Ph)$_2$	6.65	0.178	1.905	3.690	3.690	6.00	6.31	6.50
10	Cl	Cl	CH(*p*-PhF)(C$_2$H$_4$OPh-pF)	5.68	0.178	1.986	2.095	2.095	6.82	6.90	6.83
11	Me	Cl	(CH$_2$)$_3$OPh-pF	3.97	0.167	1.296	0.571	0.856	7.10	6.96	7.26[c]
12	Me	H	(CH$_2$)$_3$OPh-pF	3.26	0.167	1.296	0.571	0.856	6.60	6.54	7.24
13	Me	Cl	(CH$_2$)$_3$OPh-pCl	4.54	0.167	1.415	0.518	0.777	7.30	7.16	7.25
14	c-C$_5$H$_9$	Cl	(CH$_2$)$_3$OPh-pF	5.35	0.679	1.296	0.571	0.856		7.22	7.24
15	c-C$_5$H$_9$	H	(CH$_2$)$_3$OPh-pF	4.64	0.679	1.296	0.571	0.856		7.18	7.24
16	Me	Me	(*S*)-CH(c-hex)O$_2$CCMe$_2$Et	5.41	0.167	1.869	0.558	0.558	7.22	7.22	7.25
17	Me	Me	(*R*)-CH(c-hex)O$_2$CCMe$_2$Et	5.41	0.167	1.869	3.701	3.701	6.30	6.72	6.43
18	Me	Me	CH$_2$-c-C$_6$H$_{11}$	4.58	0.167	0.987	0.776	1.164	7.00	7.14	7.20
19	Cl	Cl	CH$_2$-c-C$_6$H$_{11}$	4.88	0.178	0.987	0.776	1.164	7.57	7.14	7.15
20	Me	Cl	CH$_2$-c-C$_6$H$_{11}$	4.73	0.167	0.987	0.776	1.164	7.12	7.16	7.19
21	Me	Me	(*S*)-CH(c-hex)NHCOCMe$_2$Et	4.19	0.167	1.895	0.550	0.550	6.49	7.13	7.25[c]
22	Me	Me	(*R*)-CH(c-hex)NHCOCMe$_2$Et	4.19	0.167	1.895	3.693	3.693	6.33	6.54	6.42
23	Me	Me	CH-c-C$_6$H$_{11}$(p-PhF)	6.27	0.167	1.677	2.197	2.197	7.02	6.64	6.79
24	i-Pr	Ph	CH$_2$-c-C$_6$H$_{11}$	6.95	0.425	0.987	0.776	1.197	6.40	6.83	7.19[c]
25	Cl	Cl	CF$_2$-c-C$_6$H$_{11}$	4.02	0.178	1.079	0.701	1.051	7.07	6.96	7.22

[a] Ref 2. [b] Ref 1. [c] Not included in deriving the equation.

stereochemistry of C_T, if any. In cases where any two of the three structural units of R become equal to each other, assign R_c the structural unit with the largest van der Waals volume. Now, the rotational and availability factors of R have been defined using the following terms: (1) flexibility of rotation, (2) probability of availability, and (3) net detachability of structural units. Each one is described below.

Flexibility of Rotation of R Rot(R). It is an expression of the freedom of rotation of R along the C_M-C_T bond. It has been defined as numerically equal to the reciprocal of the total van der Waals volume of R_a, R_b, and R_c. In equation form it is expressed as:

$$\mathrm{Rot(R)} = 1/(\mathrm{VwR_a} + \mathrm{VwR_b} + \mathrm{VwR_c}) \qquad (1)$$

Here, higher the value of Rot(R), the greater the flexibility of the rotation of R.

Probability of Availability (PA) of R_a, R_b, R_c. This is an index of relative abundance of R_a, R_b, and R_c for interaction with the probable receptor. It is defined as a ratio of 360° to maximum angular displacement required for a given structural unit to reach an identical position on rotation of R on the C_M-C_T bond. For example, the PA of R_a is computed as:

$$\mathrm{PA(R_a)} = 360°/(\text{maximum angular displacement of } \mathrm{R_a} \text{ to reach an identical position}) \qquad (2)$$

If $R_a = R_b = R_c$, then a rotation of R_6 by a maximum of 120° is necessary for R_a, R_b, and R_c to reach an identical position in space. Accordingly, PA(R_a), PA(R_b), and PA(R_c), respectively, shall become equal to 3. Similarly, if R_a, R_b, and R_c are all different, then a maximum rotation of R_6 by 360° is necessary for R_a, R_b, and R_c to reach an identical position in space; and PA(R_a), PA(R_b), and PA(R_c), respectively, shall become equal to one. In cases where two of the three structural units become equal (e.g., $R_b = R_c$ and $R_a \neq R_b$), the PA shall be 1.5 each for the similar structural units and one for the differing unit.

ND of R_a, R_b, R_c. This is an estimate of detachability of a given structural unit, R_a, R_b, or R_c, from its current position because of the rotation of R along the C_M-C_T bond. It is defined as a ratio of the flexibility of the rotation of R and the PA of the chosen structural unit. In equation form, the ND of R_a is given by,

$$\mathrm{ND(R_a)} = \mathrm{Rot(R)/PA(R_a)} \qquad (3)$$

Here, the higher the value of ND of a given structural unit, the higher its nonavailability to a given receptor pocket. If C_T is a chiral center, a uniform phase difference of $n\pi$ (π in radians, $n = 1, 2...$) is introduced between *R*/*S*-isomeric centers. This has been done keeping in mind that *R* and *S* centers create equal and opposite effects on plane-polarized light.

Data Set. Here, attempts are made to analyze the structure−activity relationship of 6-(2′,4′,6′-substituted phenoxy)-3,5-dihydroxyhexanoic acids in terms of the principal substituent group (R_6, Figure 2) in the light of its flexibility of rotation along the bond joining the substituent group (R_6) and main structural frame of the molecule (Tables 1 and 2). Along with the new parameter ND, Hansch's hydrophobic constant (π)[2] and van der Waals volume (Vw)[3] of the varying groups have been considered as physicochemical descriptors of the varying substituent groups of the compounds under investigation. The reported rat liver HMGR (solubilized and partially purified) inhibitory activity (IC_{50}, 50% inhibitory concentration in nanomoles per liters) of these compounds (Table 1)[1] was considered as a dependent variable after the transformation into logarithmic form of reciprocal inhibitory concentration and expressed as $-\log IC_{50}$. A linear multiple regression analysis by least-square method was applied in deriving the correlations. Table 1 lists the observed and predicted (one-leave-out validation) HMGR inhibitory activi-

Table 2. Assignment of Structural Units of R_6 (Figure 2) to R_a, R_b, and R_c (Figure 1)

structure no.	R_a	R_b	R_c
1	H	*p*-PhF	*p*-PhF
2	H	*p*-PhF	*p*-PhF
3	H	*p*-PhF	*p*-PhF
4[a]	H	*p*-PhOMe	*p*-PhF
5	H	*p*-PhCF$_3$	*p*-PhF
6	H	Me	*p*-PhF
7	H	*i*-Bu	*p*-PhF
8	H	mMe,pF-Ph	mMe,pF-Ph
9	H	mMe,pF-Ph	mMe,pF-Ph
10	H	C_2H_4OPh-pF	*p*-PhF
11	H	H	C_2H_4O-pPhF
12	H	H	C_2H_4O-pPhF
13	H	H	C_2H_4O-pPhCl
14	H	H	C_2H_4O-pPhF
15	H	H	C_2H_4O-pPhF
16	H	O_2CCMe_2Et	c-C_6H_{11}
17	H	c-C_6H_{11}	O_2CCMe_2Et
18	H	H	c-C_6H_{11}
19	H	H	c-C_6H_{11}
20	H	H	c-C_6H_{11}
21	H	$NHCOCMe_2Et$	c-C_6H_{11}
22	H	c-C_6H_{11}	$NHCOCMe_2Et$
23	H	c-C_6H_{11}	*p*-PhF
24	H	H	c-C_6H_{11}
25	F	F	c-C_6H_{11}

[a] In compounds 4−7, 10, and 23, the dispositions of R_a, R_b, and R_c are shown according to the *S*-configuration

ties of the compounds considered in the study along with the hydrophobicity, van der Waals volume of the varying substituent groups, and the ND of R_6 structural units.

RESULTS AND DISCUSSION

For all those compounds listed in Table 1, R_6 is the major varying group and can be represented by M−R_6, similar to the system shown in Figure 1. Here, these compounds have been considered to study the spatial disposition of R_6 vs HMGR inhibitory activity. The following equations show the correlations of the HMGR inhibitory activity of these analogues with the hydrophobicity, van der Waals volume, and/or ND of the varying substituent groups and structural units.

$$-\log IC_{50} = 1.479 + 2.490(\pm 1.293)\Sigma\pi - 0.244(\pm 1.233)(\Sigma\pi)^2 - 0.490(\pm 0.370)VwR_6 \quad (4)$$

$$n = 23,\ r = 0.741,\ s = 0.279,\ F = 7.728$$

$$-\log IC_{50} = 1.802 + 2.404(\pm 1.266)\Sigma\pi - 0.234(\pm 0.121)(\Sigma\pi)^2 - 0.528(\pm 0.391)Vw(R_2 + R_6) \quad (5)$$

$$n = 23,\ r = 0.746,\ s = 0.277,\ F = 7.933$$

$$-\log IC_{50} = 2.417 + 1.914(\pm 1.059)\Sigma\pi - 0.186(\pm 0.102)(\Sigma\pi)^2 - 0.188(\pm 0.092)NDR_b \quad (6)$$

$$n = 23,\ r = 0.822,\ s = 0.237,\ F = 13.241$$

$$-\log IC_{50} = 2.733 + 1.808(\pm 1.095)\Sigma\pi - 0.176(\pm 0.105)(\Sigma\pi)^2 - 0.194(\pm 0.102)NDR_c \quad (7)$$

$$n = 23,\ r = 0.811,\ s = 0.244,\ F = 12.136$$

Here, n is the number of data points, r is the correlation coefficient, s is the standard error of the estimate, and F is the F ratio between the calculated and observed activities. The values given in the parantheses are 95% confidence intervals of the regression coefficients. All F values are significant at more than 99% level. The $\Sigma\pi$ is the sum total hydrophobicity of R_2, R_4, and R_6 groups. NDR_b and NDR_c represent, respectively, the ND of R_b and R_c structural units of R_6. This parameter has been computed as described above. Here, a phase difference of π has been introduced between *R*- and *S*-isomers by adding 3.1428 to the ND value of *R*-isomer. In this data set, compounds 4−7, 10, and 23 are racemic mixtures. The ND values of these compounds have been taken as the average of *R*- and *S*-isomers. Also, in parameterization of compounds 8 and 9 the tetrahedral center of R_6 has been assumed to be present in *R*-configuration. This has been done because their activity is best explained under this situation, otherwise these compounds are outliers of the model. Compared with the rest of the analogues, only in compounds 8 and 9, the R_b and R_c have meta-substitution on the aryl moiety, which may be imposing some kind of steric restrictions in orienting themselves in the desired manner. The regression eqs 6 and 7 are more satisfactory than those represented by eqs 4 and 5. Also, the plots of the activity vs NDR_b, NDR_c, and π revealed that the exclusion of compounds 12, 21, and 24 from the data result in linear equations for the activity in terms of ND parameters as shown below.

$$-\log IC_{50} = 7.395 - 0.267(\pm 0.069)NDR_b \quad (8)$$

$$n = 20,\ r = 0.883,\ s = 0.185,\ F = 63.416$$

$$-\log IC_{50} = 7.453 - 0.284(\pm 0.076)NDR_c \quad (9)$$

$$n = 20,\ r = 0.880,\ s = 0.187,\ F = 61.576$$

These are the best equations that we could derive in terms of ND for the HMGR inhibitory activity of 6-aryloxy-3,5-dihydroxyhexanoic acids. All the equations derived in this study suggest that optimum hydrophobicity and a small NDR_b, NDR_c value would lead to better HMGR inhibitory activity. Further, as ND is a ratio of Rot(R_6) and PA(R_b or R_c) coupled with the *R*/*S* configuration phase difference, smaller magnitudes of ND can be achieved by having small Rot(R_6) and/or large PA(R_b or R_c), and nonchiral or *s*-configured tetrahedral center for C_T. This in turn indicates the necessity of having sterically large R_a, R_b, and R_c structural units. Here, the chirality of tetrahedral carbon C_T of R_6 appears to be optional with a preference for *S*-configured center. This kind of situation is not uncommon in nature. Even though the basic interactions involved in the HMGR inhibitory activity is of the hydrophobic and van der Waals type,[4,5] the advantage of ND is that it offers a complex insight into the possible interactions of the structural units of the tetrahedral carbon with the receptor, which is otherwise not possible with simple steric or hydrophobic parameters. Also, among ND, hydrophobicity, and van der Waals volume of R_6 substituent, very little intercorrelation exists between ND, and πR_6 and VwR_6 (Table 3). From this it is evident that even though the computation of ND consists of van der Waals volume, its information content is beyond the steric features of the substituents. Also, the ND parameter is

Table 3. Correlation Matrix of ND, Hydrophobicity, and van der Waals Volume of Compounds listed in Table 1

	$\Sigma\pi$	πR_6	VwR_6	$ND(R_b)$	$ND(R_c)$
$\Sigma\pi$	1.0				
πR_6	0.844	1.0			
VwR_6	0.345	0.591	1.0		
NDR_b	0.363	0.506	0.485	1.0	
NDR_c	0.346	0.460	0.398	0.992	1.0

designed to study the optically active compounds along with optically inactive compounds in the same data set. For this, the tetrahedral carbon of interest should be heterogeneous in nature. That is, in a given data set the tetrahedral carbon of interest should represent both optically active (*R* and *S*) as well as optically inactive centers. If this feature is absent from the data set, the ND parameter will reduce to flexibility of rotation of the substituent group.

REFERENCES AND NOTES

(1) Jendralla, H.; Granzer, E.; Kerekjarto, B. v.; Krause, R.; Schacht, U.; Baader, E.; Bartmann, W.; Beck, G.; Bergmann, A.; Kesseler, K.; Wess, G.; Chen, L.-J.; Granata, S.; Herchen, J.; Kleine, H.; Schussler, H.; Wagner, K. Synthesis and biological activity of new HMG-CoA reductase inhibitors 3. Lactones of 6-phenoxy-3,5-dihydroxyhexanoic acids. *J. Med. Chem.* **1991**, *34*, 2962−2983.

(2) Hansch, C.; Leo, A. *Substituent Constants for Correlation Analysis in Chemistry and Biology*. John Wiley: New York, 1979.

(3) Moriguchi, I.; Kanada, Y.; Komatsu, K. Van der Waals volume and related parameters for hydrophobicity in structure-activity studies. *Chem. Pharm. Bull.* **1976**, *24*, 1799−1806.

(4) Prabhakar, Y. S.; Saxena, A. K.; Doss, M. J. QSAR study of the role of hydrophobicity in the activity of HMGR inhibitors. *Drug Des. Delivery* **1989**, *4*, 97−108.

(5) Prabhakar, Y. S. QSAR study of HMGR inhibitors: 7-(heteroaryl)-3,5-dihydroxy-6-heptenoic (-heptanoic) acids. *Drug Des. Discovery* **1992**, *9*, 145−154.

CI980047S