Pergamon

Bioorganic & Medicinal Chemistry Letters 12 (2002) 2281–2285

BIOORGANIC &
MEDICINAL
CHEMISTRY
LETTERS

Quantum Chemical- and 3-D-QSAR (CoMFA) Studies of Benzalacetones and 1,1,1-Trifluoro-4-phenyl-3-buten-2-ones

Chisako Yamagami,[a,*] Noriko Motohashi[a] and Miki Akamatsu[b]

[a] *Kobe Pharmaceutical University, Motoyamakita-machi, Higashinadaku, Kobe 658-8558, Japan*
[b] *Graduate School of Agriculture, Koyto University, Kitashirakawa, Sakyoku, Kyoto 606-8502, Japan*

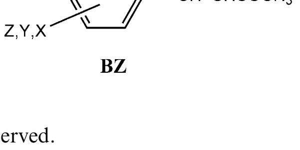
Received 26 April 2002; accepted 7 June 2002

Abstract—The inhibitory effect (IC_{50}) of the title compounds on UV-induced mutagenesis in *Escherichia coli* WP2*uvrA* was analyzed quantitatively by using various quantum chemical descriptors and also by the CoMFA method: both approaches provided results of similar quality. The activity was shown to be increased by electron-withdrawing substituents and also by hydrogen-bonding between 2-hydroxy group and the bio-component. © 2002 Elsevier Science Ltd. All rights reserved.

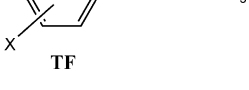
The quantitative structure–activity relationship (QSAR) study is a useful tool for rational search for bioactive compounds.[1] In a previous work,[2] we performed QSAR analyses of the antimutagenic activities of benzalacetones (**BZ**) and 1,1,1-trifluoro-4-phenyl-3-buten-2-ones (**TF**) given in Table 1, yielding the following equation,

$$\begin{aligned} -\log IC_{50} = {} & \underset{(0.183)}{1.096\Sigma\sigma^\circ} + \underset{(0.217)}{0.722HB_{2\mathrm{OH}}} + \underset{(0.335)}{0.511\Sigma\sigma^\circ_{\mathrm{TF}}} \\ & + \underset{(0.166)}{1.330I_{\mathrm{TF}}} + \underset{(0.066)}{3.089} \qquad (1) \\ & n = 26, r = 0.978, s = 0.138, F = 172 \end{aligned}$$

where IC_{50} represents the dose required to inhibit mutagenicity by 50% toward UV-induced mutagenesis in *Escherichia coli* WP2*uvrA*. The $\Sigma\sigma^\circ$ and $\sigma^\circ_{\mathrm{TF}}$ are electronic substituent terms applied separately to the **BZ** and **TF** series. $HB_{2\mathrm{OH}}$ is a hydrogen-bond parameter that takes the value of 1 for **20** and **24** (2-OH-derivatives incapable of intramolecular hydrogen-bonding with the 3-substituent); intramolecular hydrogen-bonding with the adjacent alkoxy group in the other 2-OH-derivatives, **25** and **26**, may well occur, impeding intermolecular hydrogen-bonding to the bio-phase that may enhance the potency. I_{TF} is an indicator variable that takes the value of 1 for the **TF** series and 0 for the **BZ** series. In eq 1 and throughout this paper, n is the number of compounds used for calculations, r is the correlation coefficient and s is the standard deviation. F is the value of the F ratio between the variances of the observed and calculated values. The figures in parentheses are the 95% confidence intervals of the regression coefficients and the intercept. A positive coefficient of $\Sigma\sigma^\circ(\sigma^\circ_{\mathrm{TF}})$ has been considered to be indicative of nucleophilic attack by a bio-component on the α,β-unsaturated carbonyl moiety in the side chain of **BZ** and **TF**. We have speculated that the I_{TF} term is also associated with the electronic factor, reflecting a large difference between the **BZ** and **TF** series in electron density at the reaction center. However, these concepts are still to be proved. Another problem we often encounter in performing classical QSAR analyses is unavailability of the required parameters. In application of the above equation to a greater number of compounds with a wider range of substituents, it would be necessary for us to find an appropriate method for predicting inaccessible σ° values or to use alternative electronic descriptors. Although many QSAR studies in terms of various quantum chemical parameters have been made to analyze activities of mutagenic and/or carcinogenic compounds, little is known about those for mutagenic compounds.[3–5]



BZ



TF

---

*Corresponding author. Tel.: +81-78-441-7547; fax: +81-78-441-7547; e-mail: yamagami@kobepharma-u.ac.jp

0960-894X/02/$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S0960-894X(02)00428-6

Table 1. Inhibitory potency toward UV-induced mutagenesis in *E. coli* WP2*uvrA* and quantum chemical descriptors for the **BZ** and **TF** series of compounds

			AM1					PM3				
No.	Substituent(s)	$-\log IC_{50}$[a]	E_{LUMO} (eV)	$q_{L,C7}$	$q_{L,C9}$	$F_{L,C7}$	$R_{C7/C9}$	E_{LUMO} (eV)	$q_{L,C7}$	$q_{L,C9}$	$F_{L,C7}$	$R_{C7/C9}$
					BZ							
1	H	3.30	−0.647	0.1962	0.08237	0.303	2.38	−0.804	0.1989	0.08237	0.244	2.42
2	3-Me	2.82	−0.615	0.1995	0.08468	0.324	2.36	−0.771	0.2027	0.08462	0.258	2.40
3	4-Me	2.80	−0.635	0.1892	0.07896	0.298	2.40	−0.787	0.1930	0.07958	0.240	2.43
4	4-Et	2.92	−0.622	0.1907	0.08009	0.307	2.38	−0.773	0.1944	0.08043	0.247	2.42
5	2-F	3.47	−0.836	0.1610	0.06300	0.193	2.56	−0.999	0.1658	0.06371	0.161	2.60
6	3-F	3.59	−0.865	0.1746	0.07076	0.202	2.47	−1.025	0.1774	0.07038	0.171	2.52
7	4-F	3.20	−0.872	0.1647	0.06554	0.189	2.51	−1.036	0.1699	0.06641	0.159	2.56
8	2-Cl	3.28	−0.826	0.1655	0.06554	0.200	2.53	−0.948	0.1758	0.06917	0.174	2.54
9	3-Cl	3.37	−0.836	0.1756	0.07129	0.211	2.46	−0.954	0.1820	0.07295	0.185	2.50
10	4-Cl	3.49	−0.879	0.1653	0.06605	0.188	2.50	−0.989	0.1720	0.06802	0.167	2.53
11	2-Br	3.40	−0.862	0.1624	0.06401	0.188	2.54	−0.960	0.1724	0.06750	0.169	2.55
12	3-Br	3.39	−0.853	0.1721	0.06970	0.202	2.48	−0.972	0.1798	0.07166	0.177	2.51
13	4-Br	3.28	−0.941	0.1585	0.06300	0.168	2.52	−1.025	0.1711	0.06750	0.154	2.54
14	3-CF_3	3.66	−1.053	0.1499	0.05808	0.142	2.58	−1.221	0.1506	0.05688	0.123	2.65
15	4-CF_3	3.96	−1.206	0.1348	0.05198	0.112	2.59	−1.374	0.1341	0.04995	0.098	2.69
16	3-OMe	3.15	−0.698	0.1991	0.08294	0.285	2.40	−0.839	0.2030	0.08375	0.237	2.42
17	4-OMe	2.80	−0.597	0.1941	0.08066	0.325	2.41	−0.741	0.1993	0.08220	0.262	2.43
18	3-CN	3.55	−1.017	0.1466	0.05712	0.145	2.57	−1.198	0.1481	0.05617	0.123	2.64
19	4-CN	3.77	−1.207	0.1219	0.04666	0.101	2.61	−1.373	0.1219	0.04511	0.089	2.70
20	2-OH	3.64	−0.614	0.1642	0.06605	0.267	2.49	−0.743	0.1727	0.06922	0.221	2.50
21	4-OH	2.92	−0.651	0.1892	0.07784	0.290	2.43	−0.776	0.1963	0.08026	0.244	2.45
22	3-NO_2	3.82	−1.373	0.0668	0.02465	0.049	2.71	−1.454	0.0854	0.03063	0.046	2.79
23	4-NO_2	4.08	−1.654	0.0787	0.02924	0.048	2.69	−1.742	0.0913	0.03276	0.045	2.79
24	2-OH-3-Me	3.62	−0.579	0.1648	0.06656	0.285	2.48	−0.707	0.1736	0.07017	0.233	2.47
25	2-OH-3-OMe	3.17	−0.619	0.1586	0.06350	0.257	2.50	−0.731	0.1692	0.06776	0.218	2.50
26	2-OH-3-OEt	2.92	−0.597	0.1603	0.06452	0.268	2.49	−0.716	0.1703	0.06838	0.224	2.49
27	3,4-diOH	3.01	−0.752	0.1890	0.07618	0.251	2.48	−0.847	0.1987	0.08026	0.223	2.48
28	3,5-diMe-4-OH	2.96	−0.587	0.1941	0.08123	0.331	2.39	−0.713	0.2020	0.08393	0.272	2.41
					TF							
29	H	4.52	−1.311	0.2507	0.15210	0.192	1.65	−1.455	0.2490	0.15272	0.173	1.63
30	2-Cl	4.55	−1.447	0.2261	0.12888	0.156	1.75	−1.563	0.2315	0.13462	0.145	1.72
31	3-Cl	4.62	−1.465	0.2364	0.13913	0.161	1.70	−1.572	0.2379	0.14205	0.150	1.68
32	4-Cl	4.80	−1.495	0.2274	0.13104	0.152	1.74	−1.594	0.2298	0.13462	0.142	1.71
33	3-CN	4.77	−1.608	0.2201	0.12461	0.137	1.77	−1.771	0.2187	0.12369	0.124	1.77
34	4-CN	4.60	−1.755	0.1900	0.10304	0.108	1.84	−1.900	0.1889	0.10227	0.100	1.85
35	3-NO_2	4.77	−1.851	0.1742	0.09303	0.094	1.87	−1.943	0.1834	0.09872	0.090	1.86
36	4-OMe	4.13	−1.246	0.2501	0.14977	0.201	1.67	−1.382	0.2498	0.15202	0.181	1.64

[a] IC_{50} (mmol/mL), taken from ref 2.

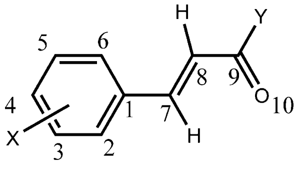
Keeping these in mind, we attempted a computational approach to analyze the same activity data used in eq 1. Firstly, correlation analyses were performed in terms of quantum chemical descriptors[6] obtained by semi-empirical molecular orbital (MO) calculations. Then, we carried out 3-D-QSAR studies by using the comparative molecular field analysis (CoMFA) model. The results were compared with the classical QSAR equation.

Activity Data

Antimutagenic activity was evaluated by the inhibitory effect of these compounds on UV-induced mutagenesis in *E. coli* WP2*uvrA*. The IC_{50} values were taken from our previous work.[2]

Quantum Chemical QSARs

Quantum chemical calculations were carried out by the AM1[7] and PM3[8] methods in the MOPAC93 program package incorporated in an ANCHOR II modeling system.[9] The initial geometry for optimization and the numbering of atoms are shown in Figure 1. With the optimized conformations, we calculated various descriptors such as the energy of the highest occupied molecular orbital (E_{HOMO}), that of the lowest unoccupied molecular orbital (E_{LUMO}), the net charge on an atom A (Q_A), and the electron density on an atom A in the HOMO ($q_{H,A}$) or LUMO ($q_{L,A}$). According to the resonance theory, nucleophilic attack would take place either on the C7 or C9 atom; parameters related to the frontier electron density were calculated for C7 and C9. The parameters necessary for discussion are listed in Table 1.



BZ: Y=CH_3, **TF**: Y=CF_3

Figure 1. The initial geometry and numbering of **BZ** and **TF**.

Regression analyses were carried out for **BZ** by using these parameters; statistically satisfactory correlations are summarized in Table 2. In all cases, the AM1 and PM3 provided essentially equivalent results. Among parameters examined, E_{LUMO} was found to be the most useful descriptor (eqs 2 and 8), and addition of a hydrogen-bonding parameter, $HB_{2\mathrm{OH}}$, always improved the correlations. The electron-attracting reactivity index for C7, $F_{\mathrm{L,C7}}$, which is defined as the LUMO density at C7 normalized by the E_{LUMO} as shown by $F_{\mathrm{L,C7}} = (q_{\mathrm{L,C7}}/-E_{\mathrm{LUMO}})$,[6,10] also provided correlations (eqs. 3 and 9) of significance equivalent to those of eqs. 2 and 8, while the use of $q_{\mathrm{L,C7}}$ in combination with $HB_{2\mathrm{OH}}$ resulted in poorer correlations (not shown). Although analyses using $q_{\mathrm{L,C9}}$ instead of $q_{\mathrm{L,C7}}$ gave correlations statistically comparable to eqs 3 and 9, higher values of $q_{\mathrm{L,C7}}$ compared to the corresponding $q_{\mathrm{L,C9}}$ values predict that the reaction would take place at C7 in preference to C9 according to the frontier electron reactivity theory.[11] Superiority of the LUMO-related parameters over the HOMO-related ones supports the previous view that the compounds tested would be attacked by a nucleophile in the critical step.

Furthermore, the validity of using σ°[12] as the electronic parameter in eq 1 could be justified by the fact that E_{LUMO}, $q_{\mathrm{L,C7}}$ and $F_{\mathrm{L,C7}}$ are correlated better with σ° than other kinds of electronic substituent constants such as σ and σ^{+}.

Next, analyses were extended to the whole set of compounds (**BZ**+**TF**). Addition of the I_{TF} term to each correlation obtained for **BZ** provided excellent correlations (eqs 5, 6, 11 and 12) in which none of the coefficients were significantly modified by inclusion of the **TF** series compounds.

The slopes of I_{TF} seemed to vary depending on which parameter was used as the electronic descriptor in combination with the I_{TF} term. This would mean that the electronic difference between **BZ** and **TF** is incompletely expressed by using E_{LUMO} or $F_{\mathrm{L,C7}}$. In order to find alternative parameters replacing the I_{TF} term, attempts were made to derive an appropriate parameter that may reasonably differentiate between the electronic structures of **BZ** and **TF**. Inspection of Table 1 shows that replacement of CH_3 on the C9 atom by CF_3 induces a considerable increase in values for $q_{\mathrm{L,C7}}$ and $q_{\mathrm{L,C9}}$. Such a tendency is more distinct for $q_{\mathrm{L,C9}}$ relative to $q_{\mathrm{L,C7}}$. To express this feature we derived a new index, $R_{\mathrm{C7/C9}}$, that is defined by the equation: $R_{\mathrm{C7/C9}} = q_{\mathrm{L,C7}}/q_{\mathrm{L,C9}}$. As shown in Table 1, the $R_{\mathrm{C7/C9}}$ values thus obtained, ranging from 2.3 to 2.8 for **BZ** but 1.6 to 1.9 for **TF**, are observed to decrease on going from **BZ** to **TF**, suggesting that the relative reactivity of the C9 atom becomes greater with decreasing $R_{\mathrm{C7/C9}}$ values. Using this parameter instead of the I_{TF} parameter provided excellent correlations (eqs. 7 and 13) of significance equivalent to those of eqs. 5 and 11 (Fig. 2). These results are thought to support our argument that the I_{TF} term in eq. 1 can be attributed to an electronic factor.

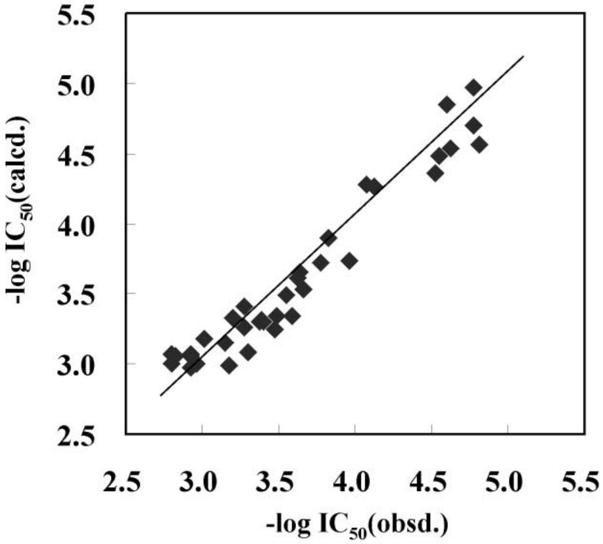


Figure 2. Plot of $-\log \mathrm{IC}_{50}$ values calculated by eq 7 against the observed values.

Table 2. Regression coefficients and constant terms in correlation analyses using various parameters (X_i) by $-\log \mathrm{IC}_{50} = a_i \Sigma X_i + \mathrm{const}$

Series	E_{LUMO}	$F_{\mathrm{L,C7}}$	$R_{\mathrm{C7/C9}}$	$HB_{2\mathrm{OH}}$	I_{TF}	Const.	n	r	s	F	Eq. no.
				AM1							
BZ	−1.240 (0.238)[a]			0.643 (0.240)		2.248 (0.213)	28	0.912	0.154	61.8	2
		−4.036 (0.712)		0.568 (0.221)		4.176 (0.163)	28	0.924	0.143	72.9	3
BZ + TF	−1.599 (0.200)			0.685 (0.329)		1.992 (0.217)	36	0.943	0.215	132	4
	−1.163 (0.229)			0.623 (0.249)	0.511 (0.202)	2.313 (0.206)	36	0.969	0.162	165	5
		−4.037 (0.699)		0.568 (0.223)	1.024 (0.129)	4.176 (0.161)	36	0.975	0.146	204	6
	−1.341 (0.184)		−0.519 (0.210)	0.660 (0.250)		3.459 (0.614)	36	0.969	0.163	162	7
				PM3							
BZ	−1.270 (0.229)			0.657 (0.229)		2.052 (0.234)	28	0.921	0.146	69.8	8
		−5.171 (0.887)		0.557 (0.215)		4.247 (0.170)	28	0.927	0.140	76.8	9
BZ + TF	−1.624 (0.199)			0.700 (0.323)		1.753 (0.240)	36	0.945	0.211	138	10
	−1.189 (0.227)			0.635 (0.243)	0.504 (0.197)	2.133 (0.233)	36	0.971	0.158	173	11
		−5.130 (0.879)		0.555 (0.220)	1.063 (0.125)	4.240 (0.169)	36	0.975	0.145	207	12
	−1.387 (0.179)		−0.459 (0.185)	0.664 (0.246)		3.101 (0.573)	36	0.970	0.160	169	13

[a] The figures in parentheses are the 95% confidence intervals.

CoMFA Studies

In the regression analyses described above, no steric effect was involved. To ascertain the validity of this result, we performed 3-D-QSAR studies with the same data set using the Advanced CoMFA module of Sybyl 6.7.[13] The full geometry optimization and the calculations of atomic charges (Q) were performed by using the AM1 method as noted above. The five atoms 1, 7, 8, 9, 10 in Figure 1 were superposed. The aligned set of active conformers was placed in an automatically created lattice of 16.5×20.2×12.6 Å ($X = -6.4$ to 10.1, $Y = -9.8$ to 10.4, $Z = -6.5$ to 6.1) with 2 Å spaces. The steric and electrostatic field energies were calculated using an sp^3 carbon probe atom with a +1 charge at all lattice intersections. The data matrix was analyzed by the partial least squares (PLS) method. We initially selected the number of compounds in the set as the number of cross-validations and then performed the analysis using the optimum number of latent variables, deduced from the cross-validation tests without actual cross-validation. Analyses were done with and without the CoMFA steric ($\mathrm{CoMFA_{st.}}$) field parameter, in addition to the CoMFA electrostatic ($\mathrm{CoMFA_{el.}}$) field parameter, and eqs. 14 and 15 were obtained when combined with the same hydrogen-bonding parameter, $HB_{2\mathrm{OH}}$, used in eq. 1, as the most reasonable correlations for **BZ** and **BZ + TF** (all compounds), respectively.

$$-\log \mathrm{IC}_{50} = 3.377 + 0.587\ HB_{2\mathrm{OH}} + [\mathrm{CoMFA_{el.}}\ \text{field terms}] \tag{14}$$

$$n = 28,\ \mathrm{CN} = 3,\ s = 0.142,\ r^2 = 0.863,\ s_{\mathrm{CV}} = 0.222,\ q^2 = 0.662,\ RC_{\mathrm{el.}} = 0.812,\ RC_{\mathrm{HB}} = 0.188$$

$$-\log \mathrm{IC}_{50} = 3.922 + 0.558\ HB_{2\mathrm{OH}} + [\mathrm{CoMFA_{el.}}\ \text{field terms}] \tag{15}$$

$$n = 36,\ \mathrm{CN} = 4,\ s = 0.145,\ r^2 = 0.952,\ s_{\mathrm{CV}} = 0.208,\ q^2 = 0.902,\ RC_{\mathrm{el.}} = 0.894,\ RC_{\mathrm{HB}} = 0.106$$

Here, CN is the number of components, and s_{CV} and q^2 represent the standard error and the correlation coefficient obtained from the leave-one-out cross-validation, respectively. $RC_{\mathrm{el.}}$ and RC_{HB} are relative contributions (%) of electrostatic and hydrogen-bonding effects, respectively. Figure 3 shows an overlay of highly potent compound **34** with the major electrostatic contour maps drawn according to eq 15. Addition of the CoMFA steric field to eq 14 ($\mathrm{CN} = 2$, $s_{\mathrm{CV}} = 0.247$, $q^2 = 0.565$) and eq 15 ($\mathrm{CN} = 3$, $s_{\mathrm{CV}} = 0.249$, $q^2 = 0.855$) failed to improve the correlation, indicating that the steric effects are of no or minor importance in the present series of compounds in accordance with the conclusion drawn by the classical QSAR studies. Blue positive electrostatic regions appear near the side chain, demonstrating that the α,β-unsaturated group is susceptible to attack by a nucleophile. Red negative regions near the terminal carbon in the side chain represent that introduction of a strongly electron-withdrawing substituent (CF_3) tends to enhance the activity, suggesting again that the nature of the I_{TF} term is electronic.

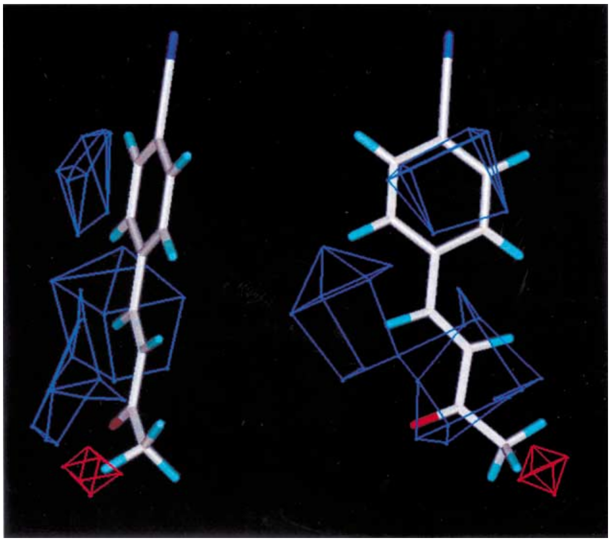


Figure 3. Orthogonal views of contour diagrams of the electrostatic filed with compound **34**. Red: negative charge favored for the activity; blue: positive charge favored for the activity.

The present work has shown that different QSAR approaches to analyze antimutagenic activities of **BZ** and **TF** have produced correlations of similar quality. The activity was shown to be increased by electron-withdrawing substituents and also by hydrogen-bonding between 2-hydroxy group and the bio-component. The quantum chemical descriptors such as frontier energy and frontier electron density can be effectively used for expressing antimutagenic activity, yielding correlations of quality equivalent to that of eq. 1. Although classical QSAR using various physicochemical substituent constants affords more straightforward information about the reaction mechanism, the successful prediction of activity by a quantum chemical QSAR using parameters obtained by rapid semi-empirical MO calculations will permit their wider application in the search for antimutagenic compounds.

Acknowledgements

This work was partly supported by the Kobe Pharmaceutical University Collaboration Fund.

References and Notes

1. Hansch, C.; Leo, A. *Exploring QSAR — Fundamentals and Applications in Chemistry and Biology*; Heller, S. R., Ed.; American Chemical Society: Washington, DC, 1995.
2. Yamagami, C.; Motohashi, N. *Eur. J. Med. Chem.* **2002**, *37*, 127.
3. Shusterman, A. J.; Debnath, A. K.; Hansch, C.; Horn,

G. W.; Fronczek, F. R.; Greene, A. C.; Watkins, S. F. *Mol. Pharmacol.* **1990**, *36*, 939.
4. Tuppurainen, K.; Lõtjõnen, S.; Laatikainen, R.; Vartiainen, T.; Maran, U.; Strandberg, M.; Tamm, T. *Mutation Res.* **1991**, *247*, 97.
5. Debnath, A. K.; Shusterman, A. J.; Lopez de Compadre, R. L.; Hansch, C. *Mutation Res.* **1994**, *305*, 63.
6. Karelson, M.; Lobanov, V. S.; Katritzky, A. R. *Chem. Rev.* **1996**, *96*, 1027.
7. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
8. Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 210.
9. Stewart, J. J. P. Fujitsu Ltd.: Tokyo, 1993.
10. Nakao, K.; Shimizu, R.; Kubota, H.; Yasuhara, M.; Hashimura, Y.; Suzuki, T.; Fujita, T.; Ohmizu, H. *Bioorg. Med. Chem.* **1998**, *6*, 849.
11. Fukui, K.; Fujimoto, H. *Bull. Chem. Soc. Jpn.* **1969**, *42*, 3399.
12. Taft, R. W. *J. Phys. Chem.* **1960**, *64*, 1805.
13. *SYBYL Molecular Modeling Software*; Tripos Associates: St. Louis, MO, USA.