Pergamon

Bioorganic & Medicinal Chemistry 9 (2001) 2885–2893

BIOORGANIC & MEDICINAL CHEMISTRY

Comparative QSAR Studies on Bibenzimidazoles and Terbenzimidazoles Inhibiting Topoisomerase I

Suresh Babu Mekapati and Corwin Hansch*

Department of Chemistry, Pomona College, Claremont, CA 91711, USA

Received 16 February 2001; accepted 7 May 2001

Abstract—Terbenzimidazoles that inhibit topoisomerase are of interest as anticancer drugs. We have reviewed the literature and have developed 13 quantitative structure–activity relationships (QSARs) on cleaving DNA or inhibiting the growth of tumor cell cultures. The results are correlated with octanol/water partition coefficients or molecular refractivity. Suggestions have been made for the development of improved derivatives.

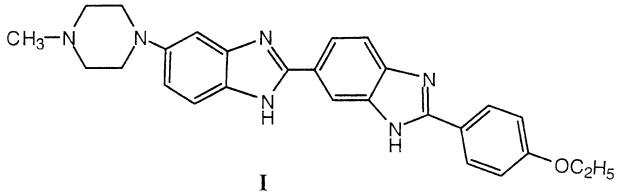
Introduction

Since its advent,[1,2] the subject of quantitative structure–activity relationships (QSARs) has been expanding at an exponential rate. There are now almost 7000 web sites devoted to QSAR. Citations to C. Hansch in the period 1979–1989 were only 8788. In the period 1989 through August 2000 the rate expanded greatly to 102,000.

Bioinformatics is clearly called for to get some kind of grasp on the enormous amount of data that is appearing at an ever faster rate. On this point, we are building a computerized system of QSAR that now contains 16,000 equations, and attendant data, of which 7600 are for chemical–biological reactions and 8400 are from mechanistic organic chemistry for comparison. In the course, of this work we find from time to time specific subjects that are worthy of an organized review that we publish[3–8] to illustrate the value of the system. It is most surprising that with the large number of SAR studies being published that very few attempts are made by the authors to formulate any kind of a QSAR.

In this report, we consider the anticancer activity of bibenzimidazoles and terbenzimidazoles. As usual, we find that of 13 sets of data, only one attempt[24] was made by the authors to formulate QSAR.

The chemotherapeutic action of several anticancer agents has been linked to their ability to inhibit nuclear DNA topoisomerases. These are nuclear enzymes involved in generating the necessary topological and conformational changes in DNA critical to many cellular processes such as replication and transcription.[9–11] Recent studies suggest that topoisomerases are also involved in controlling template supercoiling during RNA transcription and helicase movement.[12–14] Topoisomerases have been classified into type I and II depending on their ability to produce transient protein-bridged single-strand or double-strand DNA breaks. Inhibitors that interfere with the breaking and rejoining reactions of these enzymes by trapping an abortive enzyme–DNA cleavable complex have been termed topoisomerase poisons.[2] Such inhibition of mammalian topoisomerases has been recognized as an effective approach for cancer chemotherapy.[15,16] It has been demonstrated that Hoechst 33342 (I) represents a structurally unique class of topo I poisons. This agent, that binds to the minor groove of DNA, traps the reversible cleavable complex formed from DNA and topo I and produces a limited number of highly specific single-stranded DNA breaks.[17,18] A limitation of Hoechst



I

*Corresponding author.

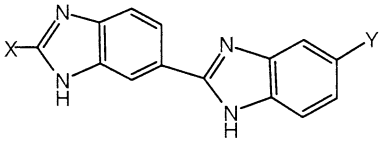
0968-0896/01/$ - see front matter
PII: S0968-0896(01)00174-2

33342 as an effective anticancer drug is that it is not effective against certain tumor cell lines. Several analogues of bibenzimidazoles and terbenzimidazoles have been identified as topo I poisons and have exhibited good cytotoxicity against RPMI8402 a human lymphoblastoma cell line.[19–26] In every instance, we have taken data from the literature, as noted, and formulated QSAR. In all the equations, *n* is the number of data points, r^2 is the square of correlation coefficient, q^2 is the measure of quality of fit and s is the standard deviation.

Results

The relative potency of various substituted bibenzimidazoles and terbenzimidazoles has been determined by assessing their ability to induce DNA cleavage in the presence of the enzyme and from this data we have derived the following QSARs.

i. Relative effective concentration of bibenzimidazoles (**II**) causing 50% cleavage of DNA in the presence of calf thymus topoisomerase I[21] (Table 1).



II

$$\text{LogRBR} = 10.67(\pm 7.40)\text{ClogP} - 1.06(\pm 0.69)(\text{ClogP})^2 - 1.58(\pm 0.58)\text{I} - 26.70(\pm 19.84)$$

$$n = 13, \quad r^2 = 0.871, \quad s = 0.321, \quad q^2 = 0.724 \qquad (1)$$

Opt. ClogP = 5.03 (4.29–5.27)

Outliers: X = 2-tolyl, Y = CN; X = phenyl, Y = $CONH_2$; X = 2-tolyl, Y = 4-methyl piperazinyl. r^2 for ClogP versus CMR = 0.385. The indicator variable I takes the value of 1 for instances where $CONH_2$ is present as Y. Its negative coefficient shows the such a group is detrimental to activity.

ii. Relative effective concentrations of terbenzimidazoles (Fig. 1) causing 50% cleavage of DNA in the presence of *Escherichia coli* topoisomerase I[23] (Table 2).

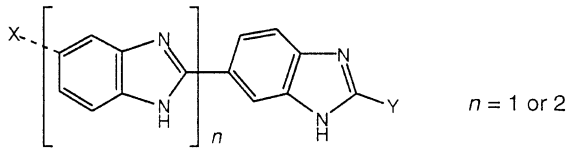
$$\text{LogRBR} = 21.90(\pm 7.90)\text{CMR} - 0.78(\pm 0.28)(\text{CMR})^2 - 153.55(\pm 55.82)$$

$$n = 10, \quad r^2 = 0.869, \quad s = 0.208, \quad q^2 = 0.787 \qquad (2)$$

Opt. CMR = 14.1 (13.9–14.2)

See Table 2 for three outliers.

iii. Relative effective concentrations of benzimidazoles (**III**) causing 50% cleavage of DNA in the presence of calf thymus topoisomerase I[20] (Table 3)



III

$$\text{LogRBR} = 7.56(\pm 1.21)\text{ClogP} - 0.83(\pm 0.13)(\text{ClogP})^2 - 17.32(\pm 2.74)$$

$$n = 9, \quad r^2 = 0.976, \quad s = 0.222, \quad q^2 = 0.844 \qquad (3)$$

Outliers: X = C_6H_5, Y = $C_7H_4N_2$. Opt. ClogP = 4.44 (4.44–4.66).

Table 1. Relative effective concentrations of bibenzimidazoles (**II**) causing 50% cleavage of DNA in the presence of calf thymus topoisomerase I[21]

Compound	X	Y	LogRBR			ClogP	I
			Obsd	calcd [eq (1)]	Δ		
1	3-Tolyl	CN	−2.00[a]	0.03	−2.03	5.35	0
2	4-Tolyl	CN	0	0.03	−0.03	5.35	0
3	1-Naphthyl	CN	−1.00	−0.85	−0.15	6.02	0
4	2-Naphthyl	CN	−1.00	−0.85	−0.15	6.02	0
5	Phenyl	$CONH_2$	−1.00[a]	−2.22	1.22	4.50	1
6	2-Tolyl	$CONH_2$	−2.00	−1.91	−0.09	4.69	1
7	3-Tolyl	$CONH_2$	−2.00	−1.60	−0.40	4.99	1
8	4-Tolyl	$CONH_2$	−1.00	−1.60	0.60	4.99	1
9	1-Naphthyl	$CONH_2$	−1.30	−1.60	0.30	5.67	1
10	2-Naphthyl	$CONH_2$	−2.00	−1.60	−0.40	5.67	1
11	Phenyl	4-Methyl-piperazinyl	0	−0.10	−0.10	4.91	0
12	2-Tolyl	4-Methyl-piperazinyl	−1.00[a]	0.08	−1.08	5.11	0
13	3-Tolyl	4-Methyl-piperazinyl	0	−0.12	0.12	5.41	0
14	4-Tolyl	4-Methyl-piperazinyl	0	−0.12	0.12	5.41	0
15	1-Naphthyl	4-Methyl-piperazinyl	−1.30	−1.25	−0.50	6.08	0
16	2-Naphthyl	4-Methyl-piperazinyl	−1.00	−1.25	0.25	6.08	0

[a]Data points not included in deriving the equation.

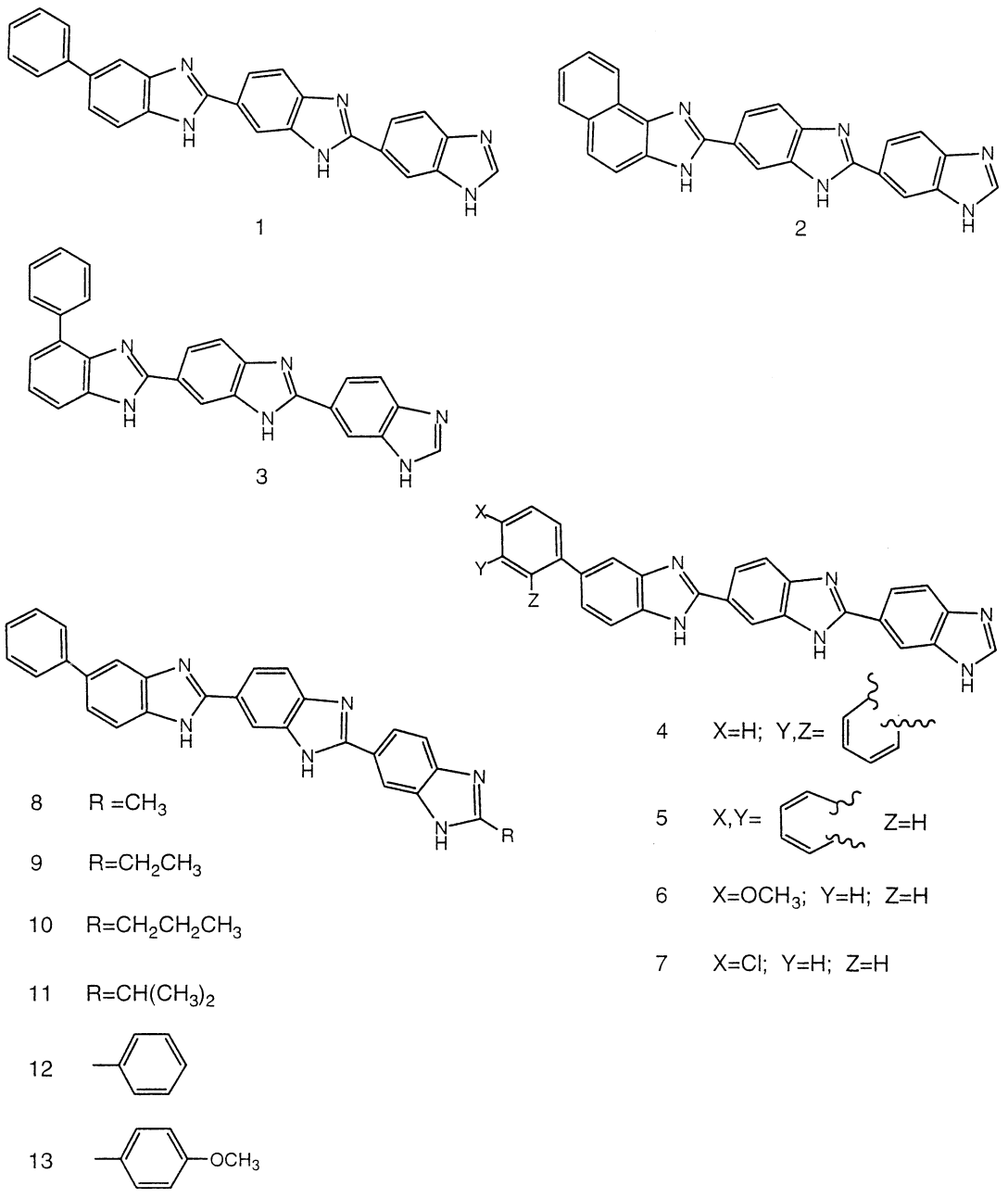
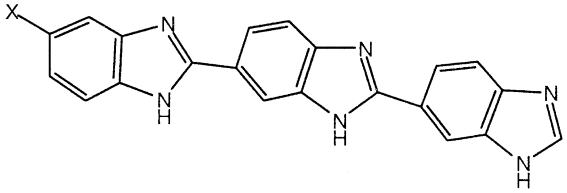


Figure 1.

iv. Relative effective concentrations of 5-substituted terbenzimidazoles (**IV**) causing 50% cleavage of DNA in the presence of *E. coli* topoisomerase[24] (Table 4)



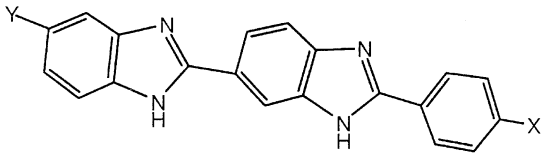
IV

$$\text{LogRBR} = 3.76(\pm 1.83)\text{CMR} - 0.16(\pm 0.07)(\text{CMR})^2 - 21.95(\pm 11.33)$$
$$n = 11, \quad r^2 = 0.902, \quad s = 0.168, \quad q^2 = 0.835 \qquad (4)$$

Outliers: F, OCH_3, CN. Opt. CMR = 11.81 (11.11–12.13).

A reasonably good correlation is also observed with ClogP, with an optimum ClogP 5.29 (4.50–5.61) ($r^2 = 0.86$). ClogP and CMR are mutually correlated ($r^2 = 0.899$).

v. Relative concentrations of bisbenzimidazoles (**V**) causing 50% cleavage of DNA in presence of calf thymus topoisomerase[19] (Table 5).



V

$$\text{LogRBR} = -1.63(\pm1.09)\text{ClogP} + 4.08(\pm1.25)\text{MgVol} - 5.24(\pm5.65)$$
$$n = 9, \quad r^2 = 0.916, \quad s = 0.417, \quad q^2 = 0.823 \qquad (5)$$

MgVol is the molar volume calculated by the method of McGowan.[27]

Cytotoxicity studies of benzimidazoles

Cytotoxicity for QSARs 6–12 of substituted bibenzimidazoles and terbenzimidazoles were determined against human lymphoblastoma cell line RPMI8402.

Table 2. Relative effective concentrations of terbenzimidazoles (Fig. 1) causing 50% cleavage of DNA in the presence of *E. coli* topoisomerase I[23]

Compound	LogRBR			CMR
	Obsd	calcd [eq (2)]	Δ	
1	0.00[a]	−0.78	0.78	12.94
2	0.00[a]	−2.74	2.74	12.12
3	−0.70	−0.78	0.08	12.94
4	−0.30	−0.05	−0.25	14.63
5	0.00	−0.05	0.05	14.63
6	0.30	−0.00	0.30	13.56
7	−0.30	−0.11	−0.19	13.43
8	−0.30	−0.13	−0.16	13.41
9	0.00	0.16	−0.17	13.87
10	0.30	0.13	0.16	14.33
11	0.30	0.13	0.16	14.33
12	−1.30	−1.31	0.01	15.45
13	−1.30[a]	−2.94	1.64	16.07

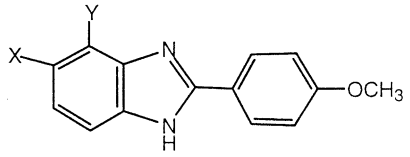
[a]Data points not included in deriving the equation.

Table 3. Relative effective concentrations of benzimidazoles (**III**) causing 50% cleavage of DNA in the presence of calf thymus topoisomerase I[20]

Compound	X	Y	*n*	LogRBR			ClogP
				Obsd	calcd [eq (3)]	Δ	
1	H	H	2	−0.04	−0.13	0.09	4.61
2	CN	H	2	0	−0.22	0.22	4.22
3	n-C_3H_7	H	2	−2.0	−2.30	0.30	6.16
4	Phenyl	H	2	−0.30[a]	−3.29	2.99	6.50
5	2-Pyridyl	H	2	−0.52	−0.49	−0.03	5.21
6	3-Pyridyl	H	2	−0.30	−0.30	−0.00	5.00
7	4-Pyridyl	H	2	−0.30	−0.30	−0.00	5.00
8	CN	H	1	−3.0	−2.97	−0.03	2.70
9	CN	OCH_3	1	−0.52	−0.19	−0.33	4.82
10	CN	4-OCH_3–C_6H_4	2	−3.0	−2.80	−0.20	6.34

[a]Data point not included in deriving the equation.

i. IC_{50} values (concentration of drug to reduce cell number to 50% of control) of 4,5-substituted 2-(4-methoxyphenyl)-1H-benzimidazoles (**VI**) against human lymphoblastoma cell line RPMI8402[22] (Table 6).



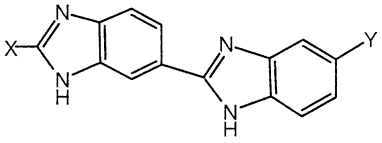
VI

$$\text{Log1/C} = 0.79(\pm0.27)\ \text{ClogP} + 1.93(\pm0.93)$$
$$n = 8, \quad r^2 = 0.893, \quad s = 0.20, \quad q^2 = 0.846 \qquad (6)$$

Outliers X = H, Y = H.

r^2 for ClogP versus CMR = 0.028.

ii. IC_{50} values of 2,5′-substituted-bi-1H-benzimidazoles (**VII**) against human lymphoblastoma cell line RPMI8402[21] (Table 7).



VII

$$\text{Log1/C} = 0.88(\pm0.31)\text{ClogP} - 0.55(\pm0.33)I_1 - 0.36(\pm0.31)I_2 + 1.57(\pm1.66)$$
$$n = 15, \quad r^2 = 0.909, \quad s = 0.187, \quad q^2 = 0.840 \qquad (7)$$

Indicator variable I_1 takes the value of 1 for the presence of 1-naphthyl groups in Y, I_2 takes the value of 1 for instances where $CONH_2$ is present in X. The negative coefficient of I_1 and I_2 indicates such groups are detrimental to cytotoxicity.

Table 4. Relative effective concentrations of 5-substituted terbenzimidazoles (**IV**) causing 50% cleavage of DNA in the presence of *E. coli* topoisomerase I[24]

Compound	X	LogRBR			CMR
		Obsd	calcd [eq (4)]	Δ	
1	1-Naphthyl	−1.00	−0.98	−0.02	14.63
2	2-Naphthyl	−1.00	−0.98	−0.02	14.63
3	Phenyl	0.00	0.08	−0.08	12.94
4	Propyl	0.30	0.28	0.02	11.82
5	Br	0.00	0.23	−0.22	11.21
6	Piperidnyl	0.30	0.08	0.22	12.94
7	Cl	0.00	0.16	−0.16	10.92
8	F	1.30[a]	−0.01	1.32	10.45
9	H	0.00	−0.02	0.02	10.43
10	OCH_3	1.30[a]	0.19	1.11	11.05
11	NO_2	0.30	0.19	0.11	11.04
12	CN	1.00[a]	0.15	0.85	10.91
13	OH	0.30	0.04	0.26	10.58
14	NH_2	0.00	0.12	−0.12	10.80

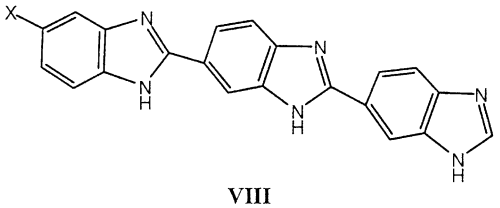
[a]Data points not included in deriving the equation.

Table 5. Relative effective concentrations of bisbenzimidazoles (**V**) causing 50% cleavage of DNA in the presence of calf thymus topoisomerase I[19]

Compound	X	Y	LogRBR			ClogP	MgVol
			Obsd	calcd [eq (5)]	Δ		
1	OC_2H_5	N–(4-CH_3)-piperazinyl	0	0.11	−0.11	5.36	3.45
2	OCH_3	$(CH_2)_3N(CH_3)_2$	−0.30	−0.75	0.45	5.56	3.31
3	OCH_3	$(CH_2)_2N(CH_3)_2$	−0.70	−0.71	0.01	5.18	3.17
4	OCH_3	$CH_2N(CH_3)_2$	−0.40	−1.05	0.66	5.04	3.03
5	OCH_3	$N(CH_3)_2$	−3.00	−2.84	−0.16	5.78	2.89
6	OCH_3	NO_2	−3.00	−2.60	−0.40	5.12	2.68
7	OCH_3	NH_2	−2.00	−2.15	0.15	4.65	2.61
8	OCH_3	N–(4-CH_3)-piperazinyl	0	0.40	−0.40	4.83	3.31
9	OCH_3	O–[(4-N–CH_3)-piperidinyl]	−0.30	−0.11	−0.19	5.40	3.41

r^2 for ClogP versus CMR = 0.267.

iii. IC_{50} values of 5-substituted terbenzimidazoles (**VIII**) against human lymphoblastoma cell line RPMI8402[24] (Table 8).



VIII

$$\text{Log}1/C = 1.55(\pm 0.51)\text{ClogP} - 0.95(\pm 0.49)B5_X - 0.81(\pm 1.83)$$
$$n = 13, \quad r^2 = 0.854, \quad s = 0.448, \quad q^2 = 0.773 \qquad (8)$$

r^2 for ClogP versus CMR = 0.884. Outliers: OCH_3.

iv. IC_{50} values of 2″-substituted-5-phenyl terbenzimidazoles (**IX**) against human lymphoblastoma cell line RPMI8402[25] (Table 9).

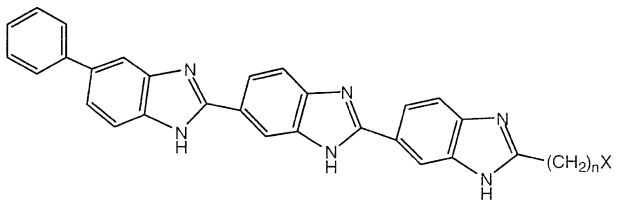
Table 6. IC_{50} values of 4,5-substituted 2-(4-methoxy-phenyl)-1H-benzimidazoles (**VI**) against human lymphoblastoma cell line RPMI8402[22]

Compound	X	Y	Log1/C			ClogP
			Obsd	calcd [eq (6)]	Δ	
1	CN	H	4.80	4.54	0.25	3.29
2	CHO	H	4.44	4.70	−0.26	3.49
3	CH_2OH	H	4.01	4.03	−0.02	2.65
4	CH_2NH_2	H	3.92	4.02	−0.10	2.64
5	$CONH_2$	H	4.18	4.06	0.12	2.68
6	H	H	4.35[a]	4.86	−0.51	3.69
7	Br	H	5.64	5.60	0.04	4.62
8	NO_2	H	4.55	4.78	−0.22	3.59
9	H	NO_2	4.96	4.78	0.18	3.59

[a]Data points not included in deriving equation.

Table 7. IC_{50} values of 2,5′ substituted-bi-1H-benzimidazoles (**VII**) against human lymphoblastoma cell line RPMI8402[21]

Compound	X	Y	Log1/C			ClogP	I_1	I_2
			Obsd	calcd [eq (7)]	Δ			
1	2-Tolyl	CN	5.92	5.97	−0.05	5.05	0	0
2	3-Tolyl	CN	6.07	6.23	−0.16	5.35	0	0
3	4-Tolyl	CN	6.24	6.23	0.01	5.35	0	0
4	1-Naphthyl	CN	6.19	6.28	−0.09	6.02	1	0
5	2-Naphthyl	CN	6.89	6.82	0.06	6.02	0	0
6	2-Tolyl	$CONH_2$	5.09	5.08	0.01	4.69	0	1
7	3-Tolyl	$CONH_2$	5.09	5.34	−0.25	4.99	0	1
8	1-Naphthyl	$CONH_2$	5.43	5.39	0.05	5.67	1	1
9	2-Naphthyl	$CONH_2$	6.13	5.93	0.20	5.67	0	1
10	Phenyl	4-Methyl-piperazinyl	6.31	5.95	0.36	4.91	0	0
11	2-Tolyl	4-Methyl-piperazinyl	6.23	6.13	0.10	5.11	0	0
12	3-Tolyl	4-Methyl-piperazinyl	6.44	6.39	0.05	5.41	0	0
13	4-Tolyl	4-Methyl-piperazinyl	6.08	6.39	−0.31	5.41	0	0
14	1-Naphthyl	4-Methyl-piperazinyl	6.48	6.44	0.04	6.08	1	0
15	2-Naphthyl	4-Methyl-piperazinyl	6.96	6.99	−0.03	6.08	0	0



IX

$$\text{Log}1/C = 0.69(\pm0.23)\text{ClogP} - 0.58(\pm0.19)\text{CMR} + 9.82(\pm2.94)$$
$$n = 12, \quad r^2 = 0.902, \quad s = 0.281, \quad q^2 = 0.790 \qquad (9)$$

Outliers: $CH_2CH_2OCH_3$; $CH_2CH_2NHCOCH_3$.

r^2 for ClogP versus CMR = 0.013.

v. IC_{50} values of bisbenzimidazoles (**X**) against human lymphoblastoma cell line RPMI8402[19] (Table 10).

Table 8. IC_{50} values of 5-substituted terbenzimidazoles (**VIII**) against human lymphoblastoma cell line RPMI8402[24]

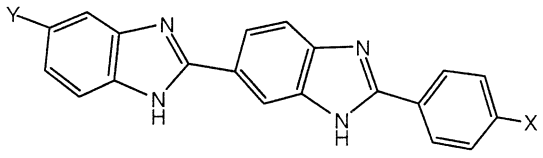
Compound	X	Log1/C			ClogP	$B5_X$
		Obsd	calcd [eq (8)]	Δ		
1	1-Naphthyl	5.90	5.85	0.05	7.67	5.50
2	2-Naphthyl	6.80	6.98	−0.18	7.67	4.31
3	Phenyl	6.72	6.29	0.43	6.50	3.11
4	Propyl	4.82	5.42	−0.61	6.16	3.49
5	Br	5.79	5.92	−0.14	5.55	1.95
6	Piperidnyl	6.20	5.75	0.45	6.38	3.49
7	Cl	5.89	5.83	0.05	5.40	1.80
8	F	5.77	5.38	0.39	4.83	1.35
9	H	5.30	5.37	−0.07	4.61	1.00
10	OCH_3	6.10[a]	3.88	2.22	4.91	3.07
11	NO_2	3.94	3.87	0.08	4.52	2.44
12	CN	3.88	4.20	−0.33	4.22	1.60
13	OH	3.82	4.56	−0.74	4.66	1.93
14	NH_2	4.21	3.60	0.61	4.06	1.97

[a]Data point not included in deriving the equation.

Table 9. IC_{50} values of 2″-substituted-5-phenyl terbenzimidazoles (**IX**) against human lymphoblastoma cell line RPMI8402[25]

Compound	X	Log1/C			ClogP	CMR
		Obsd	calcd [eq (9)]	Δ		
1	H	6.57	6.84	−0.27	6.50	12.94
2	OH	6.70	6.95	−0.25	6.78	13.09
3	SH	6.92	6.96	−0.04	7.34	13.75
4	Cl	7.16	7.12	0.03	7.31	13.43
5	CF_3	7.16	7.28	−0.13	7.56	13.45
6	CH_3	6.96	6.76	0.20	6.76	13.41
7	CH_2OH	6.27	5.77	0.50	5.46	13.56
8	$CH_2NHCOC_6H_5$	5.02	5.10	−0.08	7.18	16.79
9	CH_2CH_3	6.96	6.86	0.10	7.29	13.87
10	$CH_2CH_2OCH_3$	6.54[a]	5.63	0.91	6.03	14.49
11	CH_2CH_2OH	5.59	5.50	0.09	5.46	14.02
12	$CH_2CH_2NH_2$	5.03	5.45	−0.42	5.56	14.24
13	$CH_2CH_2NHCOCH_3$	7.00[a]	4.61	2.39	5.14	15.20
14	$CH_2CH_2CH_3$	7.22	6.96	0.27	7.82	14.33

[a]Data points not included in deriving equation.



X

$$\text{Log}1/C = 0.94(\pm0.61)\text{ClogP} + 1.44(\pm3.16)$$
$$n = 5, \quad r^2 = 0.889, \quad s = 0.156, \quad q^2 = 0.802 \qquad (10)$$

Outliers: $X = OC_2H_5$, Y = N-(4-CH_3–piperazine) and $X = OCH_3$, $Y = (CH_2)_3N(CH_3)_2$.

r^2 for ClogP versus CMR = 0.134.

vi. IC_{50} values of terbenzimidazoles (Fig. 1) against human lymphoblastoma cell line RPMI8402[23] (Table 11).

$$\text{Log}1/C = 13.05(\pm5.41)\text{CMR} - 0.45(\pm0.19)(\text{CMR})^2 - 88.43(\pm38.18)$$
$$n = 11, \quad r^2 = 0.855, \quad s = 0.385, \quad q^2 = 0.714 \qquad (11)$$

Table 10. IC_{50} values of bis benzimidazoles (**X**) against human lymphoblastoma cell line RPMI8402[19]

Compound	X	Y	Log 1/C			ClogP
			Obsd	calcd [eq (10)]	Δ	
1	OC_2H_5	N-(4-CH_3)-piperazinyl	8.30[a]	8.17	0.13	5.36
2	OCH_3	$(CH_2)_3N(CH_3)_2$	5.92[a]	7.66	−1.74	5.56
3	OCH_3	$(CH_2)_2N(CH_3)_2$	6.22	7.30	−1.08	5.18
4	OCH_3	$CH_2N(CH_3)_2$	6.42	6.94	−0.52	5.04
5	OCH_3	$N(CH_3)_2$	6.89	6.57	0.31	5.78
6	OCH_3	NO_2	6.19	6.04	0.15	5.12
7	OCH_3	NH_2	5.77	5.85	−0.08	4.65

[a]Data points not included in deriving the equation.

Table 11. IC_{50} values of terbenzimidazoles (Fig. 1) against human lymphoblastoma cell line RPMI8402[23]

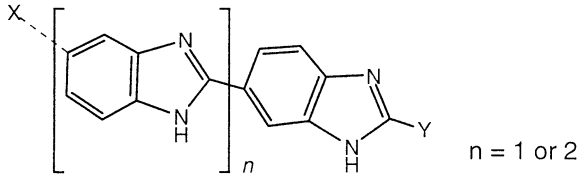
Compound	Log1/C			CMR
	Obsd	Cald [eq (11)]	Δ	
1	7.05[a]	6.00	1.05	12.94
2	4.60	4.43	0.17	12.12
3	5.15	6.00	−0.85	12.94
4	5.68[a]	7.33	−1.65	14.63
5	7.10	7.33	−0.24	14.63
6	7.05	6.78	0.27	13.56
7	7.05	6.65	0.40	13.43
8	6.96	6.62	0.34	13.41
9	6.96	7.05	−0.09	13.87
10	7.22	7.28	−0.06	14.33
11	7.22	7.28	−0.06	14.33
12	7.22	7.06	0.16	15.45
13	6.42	6.46	−0.04	16.07

[a]Data points not included in deriving the equation.

See Table 11 for two outliers.

r^2 for ClogP versus CMR = 0.919.

vii. IC_{50} values bis and terbenzimidazoles (**XI**) against human lymphoblastoma cell line RPMI8402[20] (Table 12)

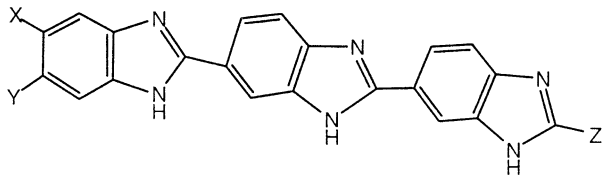


XI

$$\text{Log}1/C = 1.11(\pm 0.52)\text{CMR} - 7.09(\pm 0.30)$$
$$n = 7, \quad r^2 = 0.855, \quad s = 0.534, \quad q^2 = 0.741 \qquad (12)$$

r^2 for ClogP versus CMR is 0.232.

viii. IC_{50} values of 5-X, 6-Y, 2″-Z-substituted terbenzimidazoles (**XII**) against human lymphoblastoma cell line RPMI8402[26] (Table 13).



XII

$$\text{Log}1/C = 0.75(\pm 0.38)\pi_Z + 0.86(\pm 0.47)\text{I} + 5.86(\pm 0.34)$$
$$n = 10, \quad r^2 = 0.851, \quad s = 0.309, \quad q^2 = 0.669 \qquad (13)$$

Outliers: X = 4-Cl-C_6H_4, Y = H, Z = CH_3; X = Br, Y = Br, Z = H and X = Y = C_6H_5, Z = H.

The indicator variable I takes the value of 1 for instances where C_6H_5 is present in X. Its positive coefficient indicates that presence of phenyl group is more beneficial than Br.

Table 12. IC_{50} values of bis and ter-benzimidazoles (**XI**) against human lymphoblastoma cell line RPMI8402[20]

Compound	X	Y	*n*	Log 1/C			ClogP
				Obsd	calcd [eq (12)]	Δ	
1	H	H	2	4.85	4.45	0.41	4.61
2	*n*-C_3H_7	H	2	5.12	5.99	−0.87	6.16
3	Phenyl	H	2	7.05	7.23	−0.18	6.50
4	2-Pyridyl	H	2	6.80	6.99	−0.20	5.21
5	3-Pyridyl	H	2	7.46	6.99	0.46	5.00
6	4-Pyridyl	H	2	7.46	6.99	0.46	5.00
7	CN	OCH_3	1	4.57	4.66	−0.09	4.82

Discussion

CMR (calculated MR) is defined as molecular refractivity = $(n^2-1/n^2+2)MW/d$, where n = refractive index, M_r = molecular weight and d = density. 'n' accounts for the polarizability. Unless there is a significant variation in 'n', CMR is a measure of molecular volume. Inspection of the 13 QSARs shows that the two parameters of greatest significance are ClogP (calculated octanol/water partition coefficient) and CMR. If care is not taken in preparation of the data sets collinearity can be a serious problem. Hence, in each example we have listed the degree of collinearity for these two variables. Serious collinearity occurs in eqs (2), (3), (4), (8), and (11). This information must be kept in mind in the data analysis. There is a clear need to test a few more well designed congeners for these sets to be certain of the role of hydrophobicity. From QSARs (1) and (3), we find optimum ClogP of 5.3 and 4.5, respectively. The similarity is a bit surprising considering the difference in the parent structure. This is also seen in eq (3), where n can be either 1 or 2 but no account of this was needed to form the QSAR.

However gross bulk and polarizability are important if the variation is large enough as seen in QSAR (2).

Although we have used CMR in eq (4), the collinearity between ClogP and CMR is so high that ClogP might well be the significant parameter. QSAR (5) also shows the importantance of bulk in terms of volume but with negative coefficent of ClogP.

For QSARs (1)–(5), activity is defined as the relative concentration required to cleave 50% of DNA. QSARs (6)–(13) are based on cell inhibition.

QSAR (6) is clearly ClogP-dependent and this would suggest the presence of a hydrophobic site in the receptor that would not accommodate larger molecules. QSAR (7) has a ClogP term but needs help from two indicator variables. I_1 accounts for a bulky naphthyl group in the X-position and its negative coefficient sug-

Table 13. IC_{50} values of 5,6,2″-substituted terbenzimidazoles (**XII**) against human lymphoblastoma cell line RPMI8402[26]

Compound	X	Y	Z	Log 1/C			π_Z	I
				Obsd	calcd [eq (13)]	Δ		
1	C_6H_5	H	H	7.05	6.72	0.33	0	1
2	C_6H_5	H	Cl	7.16	7.25	−0.09	0.71	1
3	C_6H_5	H	CF_3	7.40	7.38	0.02	0.88	1
4	Br	H	H	5.80	5.86	−0.06	0	0
5	Br	H	OH	5.19	5.36	−0.17	−0.67	0
6	Br	H	$CH_2CH_2CH_3$	6.68	7.03	−0.35	1.55	0
7	4-Cl–C_6H_4	H	CF_3	6.48[a]	7.38	−0.90	0.88	1
8	Br	Br	H	6.59[a]	5.86	0.73	0	0
9	Br	Br	Cl	6.96	6.39	0.57	0.71	0
10	Br	Br	CF_3	6.60	6.52	0.08	0.88	0
11	C_6H_5	C_6H_5	H	5.80[a]	6.72	−0.92	0	1
12	Br	OCH_3	H	5.80	5.86	−0.06	0	0
13	C_6H_5	OCH_3	H	6.46	6.72	−0.26	0	1

[a]Data points not included in deriving the equation.

gests a steric effect. It is of interest that only 1-naphthyl is so parameterized, 2-naphthyl is accommodated by ClogP.

Eq (8) is also highly ClogP-dependent, however, bulky variations of X have a deleterious effect brought out by $B5_X$.

QSAR (9) is derived from a very large parent structure; nevertheless, with a small CMR correction, ClogP is quite significant. In this instance, substitution is only at one site on the molecule.

QSAR (10) is clearly ClogP determined after removing two of most basic congeners. There is essentially no variation in X.

In the case of QSAR (11), very large terbenzimidazoles with rather large substituents must be considered. The collinearity between ClogP and CMR is extreme. Despite this, r^2 using ClogP is only 0.690. Although CMR is considerably better, one can not be sure that both effects are to some degree present.

QSAR (12) clearly demonstrates the role of CMR. Only one example where $n=1$ was tested and it is well fit without special parameterization.

In example (13), the use of π constants (hydrophobicity of the substituent) clearly points to a lipophilic pocket.

It is of course of great interest to see which are the most potent compounds. Most of the end points are relative biological responses and are not good for comparison. In six examples [(6), (7), (8), (9), (11), and (13)] we can determine the most potent congener as shown in Table 14.

The agreement between sets (7)–(13) is remarkable and would suggest that we are at a dead end. Of course, we must keep in mind that some of the correlations are not as sharp as one would like. Only sets (7) and (12) have no outliers. Set (8) has 1, others have 2 or 3. In a way, this is amazing when one considers the complexity of the structures involved. The receptor site would seem to be relatively flexible. Clearly, there is considerable hydrophobic space. There is almost a one to one correspondence between Log 1/C and ClogP. However, these results are from cell studies. It is quite likely that in animal or human studies, a lower LogP would be desirable for better bioavailability.[28] Very hydrophobic compounds tend to be retained by lipophilic compartment of the body. They may also induce P450 enzymes that could metabolize molecules.[29] A place to start looking for increased potency would with set (7). Compound **5** is not covered by the indicator variable. Hence replacing CN group ($\pi=-0.57$) with SO_2NHNH_2 ($\pi=-2.04$) would lower Log P 1.5 units. Many other possibilities are available.[30] Over 1000 π values have been published.

Table 14.

Compound	Highest Log 1/C	ClogP
6	4.8	3.29
7	7.0	6.08
8	6.8	7.67
9	7.2	7.56
11	7.2	7.43
13	7.4	7.20

There are not many obvious points where one might start to make more potent compounds. In QSARs (1), (2), (3), (4), (9), and (11), where the optimums have been established for either ClogP or CMR, there are no obvious paths to more potent compounds. In none of the QSARs could we find an electronic role for substituents. The one general opportunity would be in the structures for QSARs (6), (10), (12), and (13) where optimum values have not been established for CMR or hydrophobic terms. Hence, increases in these parameters should yield more potent compounds.

References and Notes

1. Hansch, C.; Maloney, P. P.; Fujita, T.; Muir, R. M. *Nature* **1962**, *194*, 178.
2. Hansch, C. *Acc. Chem. Res.* **1969**, *2*, 232.
3. Hansch, C.; Hoekman, D.; Gao, H. *Chem. Rev.* **1996**, *96*, 1045.
4. Gao, H.; Hansch, C. *Drug Metab. Rev.* **1996**, *18*, 513.
5. Hansch, C.; Gao, H. *Chem. Rev.* **1997**, *97*, 2995.
6. Gao, H.; Katzenellenbogen, J. A.; Garg, R.; Hansch, C. *Chem. Rev.* **1999**, *99*, 723.
7. Garg, R.; Gupta, S. P.; Gao, H.; Babu, M. S.; Debnath, A. K.; Hansch, C. *Chem. Rev.* **1999**, *99*, 3525.
8. Kurup, A.; Garg, R.; Hansch, C. *Chem. Rev.* **2000**, *100*, 909.
9. Wang, J. C. *Annu. Rev. Biochem.* **1985**, *54*, 665.
10. Liu, L. F. *Annu. Rev. Biochem.* **1989**, *58*, 351.
11. D'Arpa, P.; Liu, L. F. *Biochem. Biophys. Acta* **1989**, *989*, 163.
12. Liu, L. F.; Wang, J. C. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7024.
13. Tsao, Y.-P.; Wu, H.-Y.; Liu, L. F. *Cell* **1989**, *56*, 111.
14. Koo, H. S.; Classen, L.; Grossman, L.; Liu, L. F. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 1212.
15. Schneider, E.; Hsiang, Y.-H.; Liu, L. F. *Adv. Pharmacol.* **1990**, *21*, 149.
16. Chen, A. Y.; Liu, L. F. *Annu. Rev. Pharmacol. Toxicol.* **1994**, *34*, 191.
17. Chen, A. Y.; Yu, C.; Bodley, A. L.; Peng, L. F.; Liu, L. F. *Cancer Res.* **1993**, *53*, 1332.
18. Chen, A.; Yu, C.; Gatto, B.; Liu, L. F. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8131.
19. Sun, Q.; Gatto, B.; Yu, C.; Liu, A.; Liu, L. F.; LaVoie, E. J. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2871.
20. Sun, Q.; Gatto, B.; Yu, C.; Liu, A.; Liu, L. F.; LaVoie, E. J. *J. Med. Chem.* **1995**, *38*, 3638.
21. Kim, J. S.; Gatto, B.; Yu, C.; Liu, A.; Liu, L. F.; LaVoie, E. J. *J. Med. Chem.* **1996**, *39*, 992.
22. Kim, J. S.; Sun, Q.; Gatto, B.; Yu, C.; Liu, A.; Liu, L. F.; LaVoie, E. J. *Bioorg. Med. Chem.* **1996**, *4*, 621.
23. Kim, J. S.; Yu, C.; Liu, A.; Liu, L. F.; LaVoie, E. J. *J. Med. Chem.* **1997**, *40*, 2818.
24. Kim, J. S.; Sun, Q.; Yu, C.; Liu, A.; Liu, L. F.; LaVoie, E. J. *Bioorg. Med. Chem.* **1998**, *6*, 163.
25. Rangarajan, M.; Kim, J. S.; Jin, S.; Sim, S.-P.; Liu, A.; Pilch, D. S.; Liu, L. F.; LaVoie, E. J. *Bioorg. Med. Chem.* **2000**, *8*, 1371.

26. Rangarajan, M.; Kim, J. S.; Sim, S.-P.; Liu, A.; Liu, L. F.; LaVoie, E. J. *Bioorg. Med. Chem.* **2000**, *8*, 2591.
27. Abraham, M.; McGowan, J. J. *Chromatographica* **1987**, *23*, 243.
28. Hansch, C.; Björkroth, J. P.; Leo, A. *J. Pharm. Sci.* **1987**, *76*, 663.
29. Hansch, C.; Sinclair, J. F.; Sinclair, P. R. *Quant. Struct.-Act. Relat.* **1990**, *9*, 223.
30. Hansch, C., Leo, A., Hoekman, D., *Exploring QSAR. Hydrophobic, Electronic and Steric Constants*; American Chemical Society, Washington D.C., 1995.