Antitumor Structure-Activity Relationships. Nitrosoureas vs. L-1210 Leukemia[1]

Corwin Hansch,* A. Leo, Charles Schmidt, P. Y. C. Jow,

Department of Chemistry, Pomona College, Claremont, California 91711

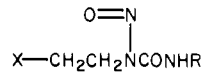
and John A. Montgomery*

Kettering-Meyer Laboratory, Southern Research Institute, Birmingham, Alabama 35205. *Received April 17, 1980*

A quantitative structure–activity relationship (QSAR) for 90 nitrosoureas acting against L-1210 leukemia in mice has been formulated. This QSAR is compared with one correlating the LD_{10} of 96 nitrosoureas. The results indicate that neutral nitrosoureas with octanol/water partition coefficients in the range of −1.5 to −2.5 might have better therapeutic indices than those currently in use.

The nitrosoureas are an extremely active class of antitumor agents which are effective against solid tumors as well as leukemias. They are particularly interesting to those who are working with QSAR, since it has been shown that they yield satisfactory QSAR.[2] Enough experience has now accumulated in the area of cancer QSAR to provide some guidelines for the design of new antitumor agents.[3] One of the most useful parameters which comes from correlation analysis is log P_0, the ideal lipophilic character for a set of congeners acting under uniform conditions in a given biological system.

In a first analysis of nitrosoureas (I) acting against L-



I

1210 leukemia in BDF_2 mice inoculated intracerebrally with 10^4 cells of leukemia and with drugs administered intraperitoneally, eq 1 was obtained.[2a] In I, X = halogen

$$\log 1/C = -0.057\ (\pm 0.04)\ (\log P)^2 - 0.069\ (\pm 0.17)\ \log P + 4.53\ (\pm 0.17) \quad (1)$$

$$n = 22;\ r = 0.922;\ s = 0.163;\ \log P_0 = -0.60\ (-7.6 \text{ to } 0.5)$$

(usually Cl), while R has been varied greatly (see Table I for examples). C in eq 1 is the molar concentration (mol/kg) required to produce a T/C of 175 (i.e., a 75% increase in life span of the mice), P is the octanol/water partition coefficient, and the figures in parentheses are the 95% cofidence limits. The number of data points is given by n, r is the correlation coefficient, and s is the standard deviation.

Equation 2 was formulated for the LD_{10} of I in mice.

$$\log 1/C = -0.069\ (\pm 0.05)\ (\log P)^2 + 0.059\ (\pm 0.02)\ \log P + 4.07\ (\pm 0.21) \quad (2)$$

$$n = 28;\ r = 0.809;\ s = 0.206;\ \log P_0 = 0.43\ (-3.7 \text{ to } 1.18)$$

Although the confidence limits on log P_0 are large, the values from eq 1 and 2 suggest that there might be a significant difference between the log P optima for T/C 175 and LD_{10}; from this one would expect hydrophilic drugs to be somewhat less toxic.

In a subsequent study of nitrosoureas on ip-injected L-1210 leukemia and ip-injected drug using an expanded set of about 80 compounds, log P_0 for ED_{50} for L-1210 leukemia was found to be 0.63 and log P_0 for LD_{10} was 0.75.[2c] The correlations were rather poor, and from an inspection of the data it seemed that the use of indicator variables might sharpen the analysis and produce better defined values for log P_0. Also, more test data are now in hand, and it is of interest to see how well our early results predict the activity of the new congeners.

Since hydrophobicity is the most important variable in correlating the activity of the nitrosoureas, we have measured more P values so that 43 of the compounds of Table I have experimentally determined log P. These new values also mean that we have a better basis set for the calculation of other log P for nitrosoureas. The calculated values in Table I have been obtained using the improved technique employing fragment constants.[4]

Results and Discussion

We have formulated eq 3 and 4 from the data in Table I for the action of congeners I on ip-implanted L-1210 leukemia in mice. Equation 4 is based on all of the data points in Table I, while eq 3 is based on five less. Dropping these five points, whose predicted activity is off by more than 3 standard deviations, gives an equation similar to eq 4, except that eq 3 contains a marginal term in $(\log P)^2$ (note development of eq 3 in Table II). However, we cannot place confidence limits on log P_0. These two equations suggest that making more hydrophilic congeners should yield more active drugs.

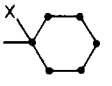
$$\log 1/C = -0.13\ (\pm 0.07)\ \log P - 0.014\ (\pm 0.015)\ (\log P)_2 - 0.76\ (\pm 0.15)\ I_1 + 0.33\ (\pm 0.17)\ I_2 - 0.24\ (\pm 0.11)\ I_3 + 1.78\ (\pm 0.09) \quad (3)$$

$$n = 90;\ r = 0.868;\ s = 0.206;\ \log P_0 = -4.4\ (\pm\infty)$$

$$\log 1/C = -0.17\ (\pm 0.04)\ \log P - 0.66\ (\pm 0.18)\ I_1 + 0.38\ (\pm 0.21)\ I_2 - 0.26\ (\pm 0.14)\ I_3 + 1.76\ (\pm 0.11) \quad (4)$$

$$n = 95;\ r = 0.791;\ s = 0.260$$

In eq 3 and 4, C is the molar concentration (mol/kg) of I producing 3 log kill of leukemia cells (i.e., a reduction of 1000-fold in tumor cells).[5] The indicator variable I_1 is assigned the value of 1 for nine cases where α substitution occurs on R:



where X = COOEt, CH_3, COOH, or −C(X)HR′ and where

(1) This research was supported in part by Grant CA-11110 from the National Cancer Institute, National Institutes of Health.

(2) (a) Hansch, C; Smith, R. N.; Engle, R.; Wood, H. *Cancer Chemother. Rep.* **1972**, *56*, 443. (b) Montgomery, J. A.; Mayo, J. G.; Hansch, C. *J. Med. Chem.* **1974**, *17*, 477. (c) Montgomery, J. A. *Cancer Treatment Rep.* **1976**, *60*, 651. (d) Wheeler, G. P.; Bowdon, B. J.; Grimsley, J. A.; Lloyd, H. H. *Cancer Res.* **1974**, *34*, 194.

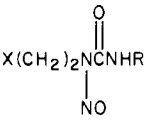
(3) (a) Hansch, C.; Hatheway, G. J.; Quinn, F. R.; Greenberg, N. *J. Med. Chem.* **1978**, *21*, 574. (b) Hansch, C. *Farmaco, Ed. Sci.* **1979**, *34*, 89.

(4) Hansch, C.; Leo, A. "Substituent Constants for Correlation Analysis in Chemistry and Biology", Wiley-Interscience: New York, 1979; Chapter IV.

(5) Johnston, T. P.; McCaleb, G. S.; Opliger, P. S.; Laster, W. R.; Montgomery, J. A. *J. Med. Chem.* **1971**, *14*, 600.

0022-2623/80/1823-1095$01.00/0

Table I. Parameters Used for the Formulation of Equations 3-7



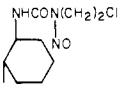
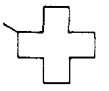
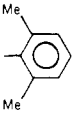
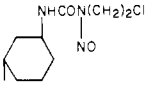
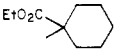
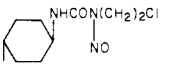
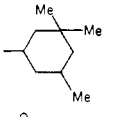
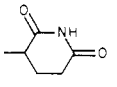
			3 log kill			LD_{10}						
no.	X	R	obsd[a]	calcd[b]	Δ log (1/C)	obsd	calcd[c]	Δ log (1/C)	log P	I_1	I_2	I_3
1	Cl	H	1.05 (3)	1.70	0.25	1.43	1.00	0.43	0.57[d]	0	0	0
2	Cl	-c-C_6H_{11}	1.25 (10)	1.31	0.06	0.613	0.69	0.08	2.83[d]	0	0	0
3	Cl		1.50 (3)	1.15	0.35	0.423	0.49	0.07	3.57	0	0	0
4	Cl		0.951 (3)	1.33	0.38	0.446	0.71	0.26	2.73[d]	0	0	0
5	Cl		1.16 (1)	1.13	0.03	0.664	0.48	0.18	3.62[e]	0	0	0
6	F	-c-C_6H_{11}	1.34 (5)	1.18	0.16	0.806	0.81	0.00	2.25	0	0	1
7	Cl		1.34 (4)	1.33	0.01	0.672	0.71	0.04	2.73[d]	0	0	0
8	Cl		1.21 (4)	1.28	0.07	0.591	0.65	0.06	2.98[d]	0	0	0
9	Cl	-c-C_5H_9	0.740[f] (2)	1.43	0.69	0.215[f]	0.82	0.60	2.19[d]	0	0	0
10	Cl		0.599 (4)	0.57	0.03	−0.573	−0.36	0.21	5.80	0	0	0
11	F	$(CH_2)_2F$	1.60 (3)	1.49	0.11	1.37	1.01	0.36	0.38	0	0	1
12	Cl		1.29 (1)	1.36	0.07	0.864	0.74	0.12	2.57[g]	0	0	0
13	Cl		1.09 (1)	1.19	0.10	0.315	0.55	0.24	3.37	0	0	0
14[b]	F(Cl)	$(CH_2)_2Cl(F)$[h]	0.673[f] (3)	1.41	0.73	0.490	0.98	0.49	0.95	0	0	1
15	Cl		1.41 (2)	1.33	0.08	0.146	0.71	0.56	2.74[e]	0	0	0
16	F		1.25 (2)	1.15	0.10	0.547	0.78	0.23	2.40	0	0	1
17	Cl		0.684 (1)	0.88	0.20	−0.212	0.13	0.34	4.66	0	0	0
18	Cl		1.55 (4)	1.21	0.34	0.771	0.57	0.20	3.30[d]	0	0	0
19	Cl		0.582 (2)	0.42	0.16	−0.214	−0.09	0.12	3.45	1	0	0
20	F		1.37 (3)	1.15	0.22	0.595	0.94	0.35	2.41[e]	0	0	1
21	Cl		0.294[f] (1)	1.03	0.74	−0.434[f]	0.34	0.78	4.06	0	0	0
22	Cl		1.80 (2)	1.73	0.07	1.16	1.01	0.15	0.37[d]	0	0	0
23	Cl	-C_6H_4-4-$CONMe_2$	1.87 (1)	1.63	0.24	1.27	0.97	0.30	1.07	0	0	0

Table I (*Continued*)

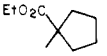
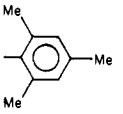
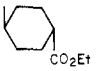
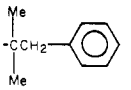
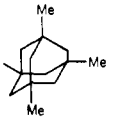
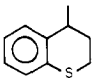
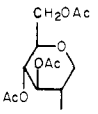
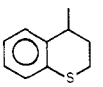
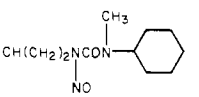
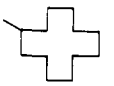
			3 log kill			LD_{10}						
no.	X	R	obsd[a]	calcd[b]	Δ log (1/C)	obsd	calcd[c]	Δ log (1/C)	log P	I_1	I_2	I_3
24	Cl		0.576 (1)	0.58	0.00	−0.030	0.10	0.13	2.70	1	0	0
25	Cl		0.562 (3)	0.54	0.02	0.072	0.06	0.01	2.88	1	0	0
26	Cl		1.41 (1)	1.43	0.02	0.786	0.82	0.03	2.20	0	0	0
27	Cl		1.58 (4)	1.25	0.34	0.675	0.62	0.06	3.11	0	0	0
28	Cl		2.10 (2)	1.79	0.31	1.25	0.84	0.41	2.07[d]	0	1	0
29	Cl		1.51 (1)	1.43	0.08	0.967	0.82	0.15	2.20	0	0	0
30	Cl		0.737 (1)	1.18	0.44	−0.422[f]	0.54	0.96	3.42[g]	0	0	0
31	Cl		1.93 (1)	1.75	0.18	1.22	1.01	0.21	0.19[d]	0	0	0
32	Cl		0.783 (1)	0.58	0.20	−0.384	−0.33	0.05	5.75[e]	0	0	0
33	F		1.92 (1)	1.59	0.33	0.922	1.01	0.09	−0.39	0	0	1
34	Cl		0.763 (1)	0.90	0.13	0.239	0.15	0.09	4.60	0	0	0
35	Cl		1.40 (3)	1.62	0.22	0.585	0.67	0.08	2.91	0	1	0
36	F		1.69 (1)	1.66	0.03	0.827	0.93	0.10	1.43[d]	0	1	1
37	Cl		1.79 (4)	1.49	0.30				1.87	0	0	0
38	Cl		1.55 (1)	1.53	0.02	0.813	0.90	0.09	1.66	0	0	0
39	Cl		1.82 (1)	1.70	0.12	1.13	1.00	0.13	0.57	0	0	0
40	F		1.40 (1)	1.35	0.05	0.884	0.94	0.06	1.33[e]	0	0	1
41	F		1.31 (1)	1.50	0.19	0.896	0.79	0.10	2.33	0	1	1
42	Cl		0.353 (1)	0.48	0.12	−0.207	−0.02	0.19	3.19	1	0	0
43	F		1.16 (1)	1.09	0.07	0.720	0.71	0.01	2.72	0	0	1
44	F					0.496	−0.10	0.59	5.22	0	0	1
45	Cl		1.81 (2)	1.78	0.03	1.07	0.84	0.23	2.08[d]	0	1	0
46	F		0.736 (1)	0.98	0.24	0.515	0.58	0.07	3.26	0	0	1

Table I (*Continued*)

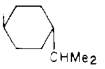
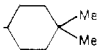
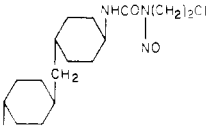
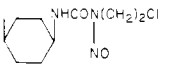
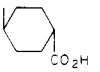
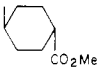
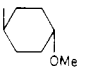
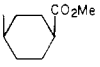
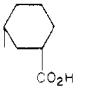
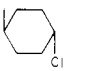
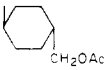
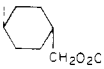
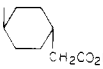
			3 log kill			LD_{10}						
no.	X	R	obsd[a]	calcd[b]	Δ log (1/C)	obsd	calcd[c]	Δ log (1/C)	log P	I_1	I_2	I_3
47	Cl		1.27 (1)	1.08	0.19	0.445	0.41	0.03	3.84	0	0	0
48	Cl		0.854 (1)	1.13	0.28	0.512	0.48	0.03	3.62	0	0	0
49	F		1.78 (1)	1.64	0.14	0.802	0.92	0.12	1.54[d]	0	1	1
50	F		1.05 (1)	0.88	0.17	0.388	0.46	0.08	3.67	0	0	1
51	F		0.797 (1)	1.02	0.23	0.592	0.64	0.04	3.04	0	0	1
52	F		1.09 (1)	1.00	0.09	0.674	0.61	0.06	3.13	0	0	1
53	Cl		1.16 (1)	1.38	0.22	0.848	0.77	0.08	2.45[d]	0	0	0
54	Cl					0.435	0.57	0.13	3.30	0	0	0
55	F		1.69 (1)	1.70	0.01	0.626	0.92	0.29	−1.48	0	0	1
56	Cl		1.90 (1)	1.88	0.02	0.842	0.98	0.14	−0.90[e]	0	0	0
57	Cl		1.36[f] (1)	0.49	0.87	0.283[f]	−0.49	0.77	6.07	0	0	0
58	Cl					0.660	0.49	0.17	3.57	0	0	0
59	Cl		1.63 (1)	1.53	0.10	1.03	0.90	0.13	1.68[e]	0	0	0
60	Cl		1.47 (2)	1.55	0.08	1.01	0.92	0.09	1.53[e]	0	0	0
61	Cl					1.05	0.87	0.18	1.89[g]	0	0	0
62	Cl		1.44 (1)	1.45	0.01	0.841	0.84	0.00	2.09	0	0	0
63	Cl		1.51 (1)	1.49	0.02	0.733	0.87	0.14	1.89	0	0	0
64	Cl		1.47 (1)	1.49	0.02	0.912	0.87	0.04	1.86	0	0	0
65	Cl		1.25 (1)	1.34	0.09	0.738	0.73	0.01	2.66	0	0	0
66	Cl	$-CH(CO_2H)(CH_2)_2-CO_2H$	0.870 (1)	0.94	0.07	0.037	0.38	0.34	0.60[e]	1	0	0
67	Cl		1.66 (1)	1.40	0.26	1.02	0.79	0.23	2.35[g]	0	0	0
68	Cl	$-CH(CH_3)CO_2H$	0.918 (1)	0.96	0.04	0.511	0.38	0.13	0.50	1	0	0
69	Cl	$-C[(CH_3)_2]CO_2H$	0.591 (1)	0.90	0.31	−0.049	0.36	0.41	0.91	1	0	0
70	Cl		0.605 (1)	0.75	0.14	0.308	0.26	0.05	1.82	1	0	0
71	Cl		1.62 (1)	1.50	0.12	1.12	0.87	0.25	1.85	0	0	0
72	Cl		1.65 (1)	1.44	0.21	0.863	0.82	0.04	2.16	0	0	0

Table I (*Continued*)

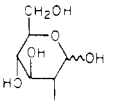
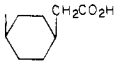
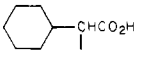
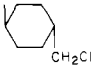
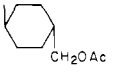
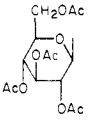
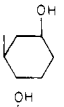
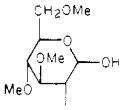
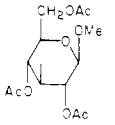
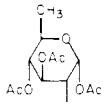
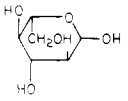
			3 log kill			LD_{10}						
no.	X	R	obsd[a]	calcd[b]	Δ log (1/C)	obsd	calcd[c]	Δ log (1/C)	log P	I_1	I_2	I_3
73	Cl	-CH(CO$_2$H)CH$_2$-C$_6$H$_5$	1.10 (1)	0.60	0.50	0.563	0.12	0.44	2.59	1	0	0
74	Cl		1.38 (1)	1.45	0.07	0.819	0.84	0.02	2.09	0	0	0
75	Cl		1.76 (2)	1.89	0.13	1.03	0.97	0.06	−1.02[e]	0	0	0
76	Cl		1.61 (1)	1.71	0.10	1.00	0.76	0.24	2.49	0	1	0
77	Cl		1.89 (2)	1.84	0.05	1.04	1.00	0.04	−0.53	0	0	0
78	Cl		1.35 (2)	1.43	0.08	0.701	0.81	0.11	2.22	0	0	0
79	Cl		1.24 (1)	1.48	0.24	0.776	0.86	0.08	1.96[g]	0	0	0
80	Cl		1.65 (1)	1.36	0.29	0.910	0.75	0.16	2.56[g]	0	0	0
81	Cl		1.89 (1)	1.84	0.05	0.801	1.00	0.20	−0.53	0	0	0
82	Cl		1.23 (1)	0.54	0.69	0.602	0.05	0.55	2.89	1	0	0
83	Cl	CH(CH$_3$)(CH$_2$)$_4$CH$_3$	0.853 (1)	1.15	0.30	0.642[f]	−0.12	0.76	3.55	0	0	0
84	Cl		1.27 (1)	1.27	0.00	0.658	0.64	0.02	3.01	0	0	0
85	Cl	-c-C$_6$H$_{10}$-Br-*p*	1.24 (1)	1.34	0.10	0.571	0.72	0.15	2.67	0	0	0
86	F		1.21 (1)	1.38	0.17				1.10	0	0	1
87	F		1.06 (1)	1.38	0.32				1.10	0	0	1
88	F		0.984 (1)	1.27	0.29	0.705	0.89	0.18	1.77	0	0	1
89	Cl		1.09 (1)	1.21	0.12	0.352	0.57	0.22	3.30	0	0	0
90	Cl					0.774	0.96	0.19	1.10[g]	0	0	0
91	Cl		1.36 (1)	1.62	0.26	0.735	0.96	0.23	1.11[g]	0	0	0
92	Cl		1.91 (1)	1.67	0.24	1.25	0.99	0.26	0.82[g]	0	0	0
93	Cl					1.41	0.97	0.44	1.04[g]	0	0	0
94	Cl		1.28 (1)	1.62	0.34	0.698	0.96	0.26	1.11	0	0	0
95	Cl		1.32 (1)	1.51	0.19	0.698	0.89	0.19	1.75[g]	0	0	0

Table I (*Continued*)

no.	X	R	3 log kill obsd[a]	3 log kill calcd[c]	3 log kill Δ log (1/C)	LD_{10} obsd	LD_{10} calcd[c]	LD_{10} Δ log (1/C)	log P	I_1	I_2	I_3
96	Cl		1.82 (2)	1.58	0.24	1.10	0.94	0.16	1.34[g]	0	0	0
97	Cl		1.44 (1)	1.64	0.20	0.795	0.96	0.17	1.00	0	0	0
98	Cl		1.48 (1)	1.67	0.19	0.682	0.99	0.31	0.80[g]	0	0	0
99	Cl					1.17	1.01	0.16	0.16[g]	0	0	0
100	Cl					1.06	1.00	0.06	0.67[g]	0	0	0
101	Cl		1.93 (1)	1.59	0.34	1.21	0.94	0.27	1.32[g]	0	0	0
102	Cl					1.12	1.00	0.12	0.69	0	0	0
103	Cl		1.46 (2)	1.86	0.40	0.928	1.00	0.07	−0.66[g]	0	0	0
104	Cl	$(CH_2)_2Cl$	1.59 (5)	1.55	0.04	0.718	0.92	0.20	1.53[d]	0	0	0

[a] Number of individual experiments used in the determination of 3 log kill given in parentheses. [b] Calculated using eq 3. [c] Calculated using eq 5. [d] Measured value; see ref 2a. [e] Measured value; see ref 4. [f] These points not used in the calculation of eq 3 or 5. [g] Measured value, this work; partitioning was done between octanol and distilled water. [h] The compound contains some of the isomeric product Cl, CH_2CH_2F.

Table II. Development of Equation 3

no.	intercept	log P	I_1	I_2	I_3	$(\log P)^2$	r	s	$F_{1,X}$
1	1.70	−0.18					0.598	0.325	49.1
2	1.77	−0.17	−0.72				0.805	0.241	71.9
3	1.74	−0.18	−0.69	0.30			0.830	0.228	11.2
4	1.80	−0.18	−0.74	0.36	−0.24		0.861	0.209	17.4
5	1.78	−0.13	−0.76	0.33	−0.24	−0.014	0.868	0.206	3.68

Table III. Squared Correlation Matrix for Variables Used in the Formulation of Equation 3

	log P	I_1	I_2	I_3
log P	1.00	0.00	0.00	0.01
I_1		1.00	0.01	0.03
I_2			1.00	0.03
I_3				1.00

R′ is aliphatic but not alicyclic and one instance where $N(c\text{-}C_6H_{11})CH_3$ is present. The negative coefficient with I_1 shows the detrimental effect of groups in this position regardless of their type. This would appear to be some kind of negative steric effect.

The variable I_2 is given the value of 1 for congeners that maintain an unoxidized sulfur atom (S but not SO_2). There are seven such examples in Table I. This structural feature increases activity by about twofold, other factors being constant. Sulfur is readily oxidized in vivo, yielding SO and SO_2. The oxidized forms would have much lower log P values and hence, in part, I is probably accounting for a latent hydrophilic property of these nitrosoureas. Congeners containing one or two SO_2 moieties are well fit without the use of indicator variables.

The variable I_3 takes the value of 1 where X of I is F and the value of 0 where X of I is Cl. Its small negative coefficient shows that F is inferior to Cl in I for antileukemia activity. In comparing eq 3 with eq 1, we note a large difference in intercepts; the reason for this is that C in eq 1 is in mol/kg, while in eq 3 and 4 it is in mmol/kg. Adding 3 to the intercept of eq 3 yields a value of 4.78, which agrees well with 4.53 of eq 1.

The QSAR for LD_{10} of I is given in eq 5–7. There is

$$\log 1/C = -0.041\ (\pm 0.007)\ (\log P)^2 - 0.62\ (\pm 0.15)\ I_1 + 1.01\ (\pm 0.06) \quad (5)$$

$$n = 96;\ r = 0.829;\ s = 0.221;\ \log P_0 = 0$$

$$\log 1/C = -0.038\ (\pm 0.008)\ (\log P)^2 - 0.53\ (\pm 0.17)\ I_1 + 0.98\ (\pm 0.08) \quad (6)$$

$$n = 101;\ r = 0.755;\ s = 0.276$$

$$\log 1/C = -0.040\ (\pm 0.009)\ (\log P)^2 + 0.95\ (\pm 0.08) \quad (7)$$

$$n = 96;\ r = 0.670;\ s = 0.291;\ \log P_0 = 0$$

less variance in the LD_{10} data and more noise so that the

correlation is poorer in terms of r and s. Only one indicator variable of the three used in eq 3 was found to be of value in eq 5 and 6; the coefficient of I_1 in eq 5 and 6 is essentially the same as that of eq 3, showing that this structural feature has the same effect on efficacy as on toxicity. The difference in the intercepts of eq 3 and 5 shows that, on the average, six times the concentration of I is needed to produce an LD_{10} as to cause a 1000-fold decrease in leukemia cells, assuming I_2 and I_3 to be constant, which is a measure of the therapeutic index of the nitrosoureas.

Although we cannot place confidence limits on log P_0 of eq 5 and 6 because of the method used in their calculation,[6] adding a term of log P to either equation does not produce a better correlation or yield an equation with a significant log P term; thus, there appears to be a real difference in log P_0 for antitumor activity and toxicity which parallels the result of eq 1 and 2. In looking for better antileukemia drugs, one should search among the more hydrophilic analogues of I.

Does this suggestion make sense when one looks at the most hydrophilic compounds in Table I? Actually, there are only seven analogues in Table I having negative log P values; these compounds are reasonably well fit by eq 3, which inclines one to think that studying nitrosoureas with log P_0 in the range of −1.50 to −2.50 would be worthwhile. In undertaking such a study, one's initial impulse might be to introduce moieties which would be charged at physiological pH (e.g., NH_2, COOH, SO_3H). This may not be a good idea; the derivatives in Table I which would be almost completely ionized at physiological pH are those containing the COOH function, but here we have found that log P for the neutral form of the molecule yields a better correlation. It has been our experience that log P or π for the ionized form of COOH does not give as good a correlation as that of the neutral form. This does not mean that such compounds are not ionized in body fluids; rather, it is fortuitous that neutral log P values work better. The higher activity which one might expect from COO^- because of its low log P may be offset by loss of drug binding to serum albumin (anions are strongly bound by albumin) or by increased difficulty of charged particles crossing biomembranes. Log P values for COO^- have been found using 0.1 N NaOH solution where the counterion is Na^+. What the most probable counterion in biological tissue is, is not known (Na^+ would be present in high concentration but its affinity might be much less than, say, NH_3^+ of an amino acid). It should be noted that it is well known from studies in physical organic chemistry that one usually cannot expect good correlation equations from data sets containing a mixture of charged and neutral substituents unless the ionic conditions under which the substituent constants were determined are the same as those for the system being correlated.[8]

The only suggestion that comes to mind for making nitrosoureas which, although un-ionized at pH 7.4, are still more hydrophilic than those in Table I would be to make R of I a disaccharide. Such molecules should have log P about 0.7 to 1.0 lower than the monosaccharides of Table I. About the same time that this thought occurred to us, a Japanese group reported[9] testing three compounds where R = β-maltosyl, β-lactosyl, and β-cellubiosyl. The respective ILS for these compounds acting on L-1210 leukemia are 403, 384, and 483 for doses of 20, 10, and 8 (mg/kg)/day. Not only are these highly potent, but the authors state that the maltosyl analogue appeared to be less toxic than the monosaccharide derivatives. Since experimental details differ, a direct comparison with these data cannot be made.

One of the advantages of formulating correlation equations is to uncover structural features which confer unusual biological activity. Those congeners whose activity is grossly underpredicted by eq 3 merit special consideration. There are only two such examples in Table I and, although these congeners (**57** and **82**) are rather badly underpredicted, neither is a very potent molecule with obvious characteristics for development.

There are three compounds (**9**, **14**, and **21**) that are much less active than expected, but these too afford us no special insight into the SAR of antitumor activity.

One must remember that eq 3 has been formulated for nitrosoureas producing 3 log kill of L-1210 leukemia cells. One would not expect and, in fact, does not find the same QSAR for nitrosoureas acting against other kinds of tumors. This is easily recognized from a study of the action of nitrosoureas against the solid Lewis lung tumor. An entirely different structure–activity pattern occurs in the Lewis lung tumor which, at the present time, does not seem to be amenable to QSAR treatment. The results with Lewis lung tumor incline one to think that the more lipophilic nitrosoureas are, in general, more effective.[10]

Finally, it should be mentioned that using Kubinyi's bilinear model[11] instead of log P + log P^2 in eq 3–6 yielded a slightly poorer correlation.

In summary, our analysis shows that the lipophilic character of the nitrosoureas is the most important parameter determining their antitumor activity. The one indicator variable in eq 3 with a positive coefficient shows that slight improvement in activity can be achieved by the inclusion of sulfur in R. A rather wide range of R groups has now been studied and no structural features of unusual significance have been uncovered. The only obvious conclusion to draw from eq 3 and 5 is that prospecting among more hydrophilic neutral congeners of I might yield analogues with a better threapeutic index.

(6) Hansch, C.; Steward, A. R.; Anderson, S. M.; Bentley, D. *J. Med. Chem.* **1968**, *11*, 1.
(7) Panthananickal, A.; Hansch, C.; Leo, A.; Quinn, F. R. *J. Med. Chem.* **1978**, *21*, 16.
(8) Hoefnagel, A. J.; Hoefnagel, M. A.; Wepster, B. M. *J. Org. Chem.* **1978**, *43*, 4720.
(9) Suami, T.; Machinami, T.; Hisamatsu, T. *J. Med. Chem.* **1979**, *22*, 247.
(10) Montgomery, J. A.; McCaleb, G. S.; Johnston, T. P.; Mayo, J. G.; Laster, W. R., Jr. *J. Med. Chem.* **1977**, *20*, 291.
(11) Kubinyi, H. *J. Med. Chem.* **1977**, *20*, 625.