Biophys. Acta, **212**, 116 (1970).
(12) M. Friedkin, E. J. Crawford, and D. K. Misra, *Fed. Proc., Fed. Am. Soc. Exp. Biol.*, **21**, 176 (1962).
(13) L. T. Plante, E. J. Crawford, and M. Friedkin, *J. Biol. Chem.*, **242**, 1466 (1967).

(14) J. R. Bertino, J. P. Perkins, and D. G. Johns, *Biochemistry*, **4**, 839 (1965).
(15) J. K. Coward, K. N. Parameswaran, A. R. Cashmore, and J. R. Bertino, *Biochemistry*, **13**, 3899 (1974).
(16) A. C. Farthing, *J. Chem. Soc.*, 3213 (1950).

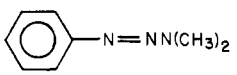
Use of Cluster Analysis in the Development of Structure–Activity Relations for Antitumor Triazenes

W. J. Dunn, III,* M. J. Greenberg, and Soledad S. Callejas

Department of Medicinal Chemistry, College of Pharmacy, University of Illinois at the Medical Center, Chicago, Illinois 60680. *Received January 12, 1976*

A series of antitumor triazenes in which the members of the series are physicochemically distinct was designed using the cluster analysis approach as proposed by Hansch and his co-workers. The series that resulted was tested against Sarcoma 180 in the mouse and the antitumor activities were analyzed using regression techniques. The structure–activity relations that resulted are discussed in terms of proposed mechanisms of action.

1-Phenyl-3,3-dimethyltriazene (I) was reported in 1955 to be active against Sarcoma 180 in the mouse.[1] More recent studies by Lin et al.[2] showed that the 2-, 3-, and 4-$CONH_2$ and 2-, 3-, and 4-CO_2CH_3 analogues of I were active against L1210 mouse leukemia. Also, a number of analogues of I have been reported by Shealy et al.[3] to be active against L1210.



I

While a number of reports of the antitumor activity of analogues of I have appeared, there have been no reports of the systematic design of a series of I on which quantitative structure–activity studies have been done.

We recently undertook such a study which led to the preparation and testing of seven analogues (compounds 1–4, 6, 7, and 14, Table I). Attempts to derive quantitative relationships from these data yielded ambiguous results due to considerable interrelationships between the physicochemically derived independent variables considered. At this time, we became aware of the work of Hansch et al.[4] in which they proposed that cluster analysis could be used in the design of series such as ours. The result is a series in which the physicochemical properties of each analogue are unique and the chance of collinearity between these variables is minimized. This report deals with our results of the use of cluster analysis in the design of antitumor derivatives of the lead substance I.

Cluster Analysis and Its Application. Hansch et al.[4] selected 90 substituent groups and applied to their physicochemical constants, heirarchial clustering to factor the groups according to their properties. These properties are in terms of π, the Hansch constant, F and R, the field and resonance constants, respectively, molar refractivity (MR), and molecular weight (MW). Four sets (sets 1–4 in their report) were treated in this way with their set 1 factored according to the parameters π^2, π, F, R, MR, and MW, set 2 according to the parameters π^2, π, F, and R, set 3 according to π, F, R, and MR, and set 4 according to π, F, and R. Within each set four levels of factorization were carried out. In set 2, for example, the lowest level produced from the 90 substituents five clusters, the next highest gave 10 clusters, the next 20, and the last 60 clusters. The higher the level of factorization the less forcing occurs on clustering so the more distinction is obtained.

For this study we have selected for investigation the dependence of activity on π^2, π, F, R, MR, and MW. This requires set 1 of Hansch's clusters. We selected the 10 level of this set and from this 14 analogues resulted. We considered only monosubstituted 3- and 4-substituted triazenes and assumed the 3-substituted and 4-substituted analogues to be a single series. This assumption is based on the data reported by Lin et al.[2] These workers found that the placing of the same substituent in the 3 or 4 position had little, if any, effect on the level of antitumor activity.

The analogues designed and prepared are given in Table I. Also included in this table are the biological data and physicochemical substituent constants. Table II summarizes the physical data of the triazenes synthesized in this study and Table III is a correlation matrix of the groups and their variables.

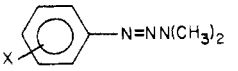
Experimental Section

Synthesis. The majority of the triazenes in this study were prepared by coupling the appropriate aryldiazonium cation with dimethylamine according to the procedure of Lin et al.[2] This procedure could not be used for the preparation of *p*-(3,3-dimethyl-1-triazeno)cinnamic acid. In this case the aryldiazonium cation was isolated as the stable fluoborate salt and this species was coupled with dimethylamine according to the procedure of Kolar.[5] The results are given in Table III and the melting points reported are uncorrected. Structures and purity were verified by thin-layer chromatography, mass spectra, and infrared and ^1H NMR spectroscopy.

Biological Testing. The triazenes were tested for activity against Sarcoma 180 ascitic tumor in the mouse using a modification of the procedure of Sartorelli.[6] Donor mice were sacrificed by asphyxiation with chloroform, followed by removal of the top layer of skin of the intraperitoneal cavity, and fluid was removed. A threefold dilution of this fluid was made with sterile saline and 0.1 ml (approximately 2×10^6 cells) was injected into 15–18-g female albino Swiss mice. Each test group was composed of six mice and a control of six mice was determined each time a test was made.

The triazenes were suspended in sterile saline with Tween 20 and administered on alternate sides of the intraperitoneal cavity 24 h after implantation. This was repeated daily for 3 days. Initial dose levels were set at 0.25, 0.38, and 0.50 of the LD_{50} for that particular drug. Other dosage levels were set depending on initial activity and toxicity as determined by weight loss. The test and control groups were observed daily and the day of death was recorded. A dose–response curve was generated from which the log dose (mg/kg) required to give T/C (%) = 130 was obtained. The error in such a determination can be as high as 22% as

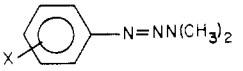
Table I. Biological Data and Physicochemical Constants



Compd	X	pC_{130}[a]	pC_{130} calcd by eq 2	pLD_{50}	π[b]	MR[b]	σ[b]	MW[b]	F[b]	R[b]
1	H	3.42	3.41	2.84	0.00	1.03	0.00	1.0	0.00	0.00
2	3-CH_3	3.37	3.46	2.91	0.56	5.75	−0.07	15.0	−0.04	−0.13
3	4-CN	2.91	2.96	2.56	−0.57	6.33	0.66	26.0	0.51	0.19
4	3-Cl	3.16	3.16	2.68	0.71	6.03	0.37	35.4	0.41	−0.15
5	3-CF_3	3.18	3.11	2.80	0.88	5.02	0.43	69.0	0.38	0.19
6	3-CO_2H	3.01	3.16	2.55	−0.32	6.93	0.37	45.0	0.33	0.15
7	3-SCH_3	3.33	3.31	2.76	0.61	13.82	0.15	47.1	0.20	−0.18
8	4-F	3.24	3.37	2.99	0.14	0.92	0.06	19.0	0.43	−0.34
9	4-*n*-C_3H_7	3.56	3.50	2.91	1.55	14.96	−0.13	43.1	−0.06	−0.08
10	3-$NHCOCH_3$	3.45	3.27	2.70	−0.97	14.93	0.21	58.1	0.28	−0.26
11	4-COC_6H_5	3.16	3.11	2.82	1.05	30.33	0.43	105.1	0.30	0.16
12	4-C_6H_5	3.43	3.52	2.66	1.96	25.36	−0.01	77.1	0.08	−0.08
13	4-$CH=CHCO_2H$	2.79	2.79	2.57	0.00	17.91	0.90	71.1	−0.15	1.04
14	4-NO_2		2.87	3.00	−0.28	7.36	0.78	46.0	0.67	0.16

[a] C = mol/kg. [b] See ref 10.

Table II. Synthesis and Physical Properties



Compd	Bp (mm) or mp, °C[a]	Yield, %[b]	Formula[c]
1	89.3 (2.3) [113-114 (12[d])]	45	$C_8H_{11}N_3$
2	88.2 (0.75) [79 (0.50[e])]	28	$C_9H_{13}N_3$
3	112-113 [112[e]]	62	$C_9H_{10}N_4$
4	105 (1.0) [97-99 (0.4[f])]	33	$C_8H_{10}N_3Cl$
5	73.5 (0.4) [76-78 (1.3[f])]	45	$C_9H_{10}N_3F_3$
6	113 [115[g]]	14	$C_9H_{11}N_3O_2$
7	107 (0.1)	40	$C_9H_{13}N_3S$
8	37-38 [37.5[e]]	38	$C_8H_{10}N_3F$
9	112 (0.2)	37	$C_{11}H_{17}N_3$
10	132-134	46	$C_{10}H_{14}N_4O$
11	57-58	30	$C_{15}H_{15}N_3O$
12	72-73 [70-71[h]]	57	$C_{14}H_{15}N_3$
13	177-178	39	$C_{11}H_{13}N_3O_2$
14	97 [98[f]]	89	$C_8H_{10}N_4O_2$

[a] Crystallization solvent was methanol. Literature values are in brackets. [b] Based on starting arylamine. [c] C, H, and N analyses were obtained for all compounds; results were within ±0.4% of the theoretical values. [d] Reference 11. [e] Reference 12. [f] Reference 13. [g] Reference 5. [h] Reference 14.

Table III. Correlation Coefficient Matrix

	Act.	σ	π	F	R	MW	MR
Act.	1.00	0.92	0.40	−0.28	−0.71	−0.29	0.05
σ	0.92	1.00	−0.45	0.25	0.79	0.46	0.06
π	0.40	−0.45	1.00	−0.37	−0.14	0.23	0.46
F	−0.28	0.25	−0.15	1.00	−0.38	0.11	−0.28
R	−0.71	0.77	−0.07	−0.38	1.00	0.39	0.23
MW	−0.29	0.46	0.27	0.11	0.39	1.00	0.78
MR	0.05	0.06	0.46	−0.28	0.23	0.78	1.00

determined from our data and error theory.

LD_{50} values were determined by the procedure of Turner[7] using suspensions prepared as previously described. The LD_{50} values reported are the moles per kilogram of animal which are lethal to 50% of a group of five mice in 1 week. Each LD_{50} was determined from a dose–response curve derived from at least four test groups. The activities are reported in Table I.

Statistical Analysis. Regression studies of the biological activity were carried out using the Statistical Analysis System and the IBM 370/155 computer of the Research Resources Laboratory of the University of Illinois at the Medical Center.

Results

The correlation matrix for the set of triazene substituents which resulted from cluster analysis is given in Table III. The difficulty in our original seven compounds was due to the fact that two were in cluster 1 (H, CH_3) and the remainder from cluster 3. Our final series of 14 analogues shows little if any covariance among variables. This emphasizes an advantage of using this approach to analogue design.

This approach should lead to a series in which quantitative structure–activity relationships may be derived. From the standpoint of mechanism of action studies, this can be significant. In an attempt to determine such relationships, the antitumor and toxicity data were submitted to regression analysis. The relationship represented between antitumor activities and the physicochemical parameters of these analogues is given in eq 1 and 2. It can be seen from these equations that hydrophobicity is not a significant variable in determining the level of activity while the electronic nature of the substituent as measured by its Hammett σ constant explains 85% of the variance in the antitumor activity. A similar result is found for the dependence of toxicity on physicochemical properties (eq 3). While this correlation

$$pC_{130} = 4.01\ (\pm 0.88) + 0.12\ (\pm 0.10)\ \pi^2 - 0.65\ (\pm 0.61)\ \pi \quad (1)$$
$$n = 13;\ r^2 = 0.22;\ s = 6.22$$

$$pC_{130} = 3.41\ (\pm 0.03) - 0.69\ (\pm 0.09)\ \sigma \quad (2)$$
$$n = 13;\ r^2 = 0.85;\ s = 0.09$$

$$pLD_{50} = 2.84\ (\pm 0.04) - 0.33\ (\pm 0.10)\ \sigma \quad (3)$$
$$n = 14;\ r^2 = 0.48;\ s = 0.11$$

with σ is not as high as that for antitumor activity, the equation is significant at >95% ($F_{1,13} = 11.11$; $F_{1,13;\alpha=0.05} = 4.67$). The low correlation is probably due to the narrow range in toxicities found in the series (mean = 2.74, standard deviation = 0.15) which results in a very small variance about the mean. No significant dependence of toxicity on π was observed.

Since σ is a composite of both resonance and field components, an attempt to dissect the σ term in eq 2 and 3 to the Swain–Lupton F and R constants was made. In order to do this the 3- and 4-aryl analogues were considered

to be two different sets but were treated in the same equation. With antitumor activity as the dependent variable, this treatment yielded eq 4, which shows that the

$$pC_{130} = 3.42\ (\pm 0.05) - 0.77\ (\pm 0.18)\ F_{para} - 0.68\ (\pm 0.11)\ R_{para} - 0.63\ (\pm 0.20)\ F_{meta} \quad (4)$$

$$n = 13;\ r^2 = 0.83;\ s = 0.11;\ F_{3,13} = 14.27;\ F_{3,13;\alpha=0.05} = 3.41$$

electronic effect on activity is due to both field and resonance effects. The field effect dominates for substituents in the 3 position. A similar result was observed for toxicity as is seen in eq 5. A comparison of eq 4 and

$$pLD_{50} = 2.87\ (\pm 0.04) - 0.55\ (\pm 0.13)\ F_{para} - 0.22\ (\pm 0.10)\ R_{para} - 0.41\ (\pm 0.18)\ F_{meta} \quad (5)$$

$$n = 14;\ r^2 = 0.67;\ s = 0.10;\ F_{3,14} = 6.90;\ F_{3,14;\alpha=0.05} = 3.34$$

5 suggests that the separation of toxic and antitumor activities would be difficult.

The mechanism of action of the triazenes is not known. Studies indicate that these agents may exert their effect by one of the two following modes: (1) metabolic conversion of the triazene to an active substance(s); or (2) decomposition of the triazene to an active substance(s).

The rationale for the first proposal is, in part, due to the observation that 3,3-disubstituted triazenes, in order to be active, must have as a 3-substituent the methyl group. It has been shown that such triazenes are readily demethylated in vitro and it is thought that some resulting intermediate may be responsible for activity.[8]

In the second case the active substance has been proposed to be the diazonium cation formed by hydrolysis or photolysis of the triazene. Photolysis should be of little significance of vivo.

The proposal that triazenes may owe their activity to a diazonium cation formed from hydrolysis is generally discounted due to the fact that, in general, the diazonium cation resulting from such a reaction is inactive.[8]

To reject the idea that the diazonium intermediate is not involved in the mechanism of antitumor activity due to the fact that the administered diazonium cation is inactive does not consider that the triazene may be transferred into the cell and there undergo hydrolysis. To discount such a mechanism may be premature at this time, especially since an administered diazonium cation would have little chance of entering an intact cell.

Preussmann and Kolar[9] have determined the half-life for hydrolysis of 14 3- and 4-substituted 1-phenyl-3,3-dimethyltriazenes. Some of the triazenes in their study are the subject of this report. From their data the following regression equation of log k vs. σ was obtained. In this equation k is the pseudo-first-order rate constant for

$$\log k = -2.70\ (\pm 0.16) - 4.27\ (\pm 0.38)\ \sigma \quad (6)$$

$$n = 14;\ r^2 = 0.91;\ s = 0.48$$

hydrolysis at pH of 7.0 and 37 °C. The equation is significant at >99% and shows that hydrolysis and antitumor activity are dependent upon σ in the same way. Log k and pC_{130} were similarly related to the F and R constants as seen by eq 7 which is significant at >99%.

$$\log k = -2.92\ (\pm 0.22) - 4.98\ (\pm 0.65)\ R_{para} - 3.11\ (\pm 0.61)\ F_{para} - 4.27\ (\pm 0.62)\ F_{meta} \quad (7)$$

$$n = 14;\ r^2 = 0.93;\ s = 0.46$$

The correlation between antitumor activity and hydrolysis is remarkable; they are dependent on electronic effects in the same way which suggests that hydrolysis of the triazene to the diazonium cation may be involved in the cytotoxic mechanism. Further work on this is in progress as are studies to determine if other chemical pathways may be operating which are compatible with these results.

Because of the complexity of studies of drug design and action, systematic approaches, such as the use of cluster analysis, are becoming attractive. This report illustrates the application of the technique and its potential for studying mechanisms of drug action.

Acknowledgment. The investigators acknowledge the support of this work by the American Cancer Society Institutional Grant No. 1N-9N.

References and Notes

(1) D. A. Clarke, R. K. Barclay, C. C. Stock, and C. S. Rondestvedt, *Proc. Soc. Exp. Biol. Med.*, **90**, 484 (1955).
(2) Y. T. Lin, T. L. Loo, S. Vadlamudi, and A. Goldin, *J. Med. Chem.*, **15**, 201 (1972).
(3) Y. F. Shealy, C. A. O'Dell, J. D. Clayton, and C. A. Krauth, *J. Pharm. Sci.*, **60**, 1426 (1971).
(4) C. Hansch, S. H. Unger, and A. B. Forsythe, *J. Med. Chem.*, **16**, 1217 (1973).
(5) G. F. Kolar, *Z. Naturforsch., B*, **27**, 1183 (1972).
(6) A. C. Sartorelli, B. A. Booth, and K. C. Agrawal, *J. Med. Chem.*, **11**, 700 (1968).
(7) R. A. Turner, "Screening Methods in Pharmacology", Academic Press, New York, N.Y., 1965.
(8) R. C. S. Audette, T. A. Connors, H. G. Mandel, K. Merai, and W. C. J. Ross, *Biochem. Pharmacol.*, **22**, 1855 (1973).
(9) G. F. Kolar and R. Preussmann, *Z. Naturforsch., B*, **26**, 950 (1971).
(10) C. Hansch, A. Leo, S. H. Unger, K. H. Kim, D. Nikaitani, and E. J. Lien, *J. Med. Chem.*, **16**, 1207 (1973).
(11) J. Elks and D. H. Hey, *J. Chem. Soc.*, 441 (1943).
(12) V. Zverina, J. Divis, J. Marhold, and M. Matrka, *Cesk. Farm.*, **18**, 33 (1969); *Chem. Abstr.*, **71**, 70216q (1969).
(13) C. S. Rondestvedt and S. J. Davis, *J. Org. Chem.*, **22**, 200 (1957).
(14) R. J. W. LeFevre and T. H. Liddicoet, *J. Chem. Soc.*, 2743 (1951).