Journal of Medicinal Chemistry

VOLUME 11, NUMBER 2 FEBRUARY 26, 1968

Application of Regression Analyses to Antitumor Activities of Various Acetylenic Carbamates[1]

WILLIAM P. PURCELL AND JOHN M. CLAYTON

Department of Medicinal Chemistry, University of Tennessee College of Pharmacy, Memphis, Tennessee 38103

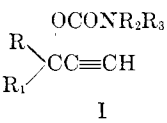
Received October 25, 1967

Regression analyses by the Free and Wilson method[2] were applied to the tumor inhibition (per cent inhibition/mg of drug per kg of test animal) and cure potency (per cent prolongation of life/mg of drug per kg of test animal) of 69 substituted acetylenic carbamate analogs.[8] The original biological response parameters yielded more meaningful results than logarithms of these data. For regressions considered of predictive value, correlations were significant at greater than the 90% level. It was found that statistical tests alone are not always reliable means of judging the predictive utility of regressions of this type. Compounds predicted to be most active against these tumors were among those not tested; some of the more promising compounds would contain 2-naphthyl, 4-fluorophenyl, or phenyl groups or combinations of them at the nitrogen with cyclohexyl, cycloheptyl, or cyclopentyl groups or combinations at the 1,1-(2-propynyl) positions. In compounds with the highest calculated activities, substituents on the nitrogen appear to contribute more to the total activity than do substituents at the 1,1-(2-propynyl) positions.

In the search for a method of accurate prediction of therapeutically active molecules for specific pharmacological actions, the application of regression analyses, of both the mathematical[2,3] and linear free-energy[4,5] models, continues to hold much promise and interest. In consideration of the labor of synthesis and testing associated with drug development, any mechanism suggesting molecules having a high probability of success would be invaluable. An apparently good correlation was found, for example, in the application of Free and Wilson's method[2] to an analogous series of cholinesterase inhibitors.[6] More recently, application of this technique to hypoglycemic activities of several piperidinesulfamyl semicarbazides has also given interesting results.[7]

For meaningful application of the Free and Wilson model,[2] the biological data should meet three basic prerequisites: (1) molecules in the series should be closely similar (to increase the probability of a constant mechanism of action), (2) biological activity selected should be accurate, quantitative, and measured under uniform conditions for the series, and (3) the group contributions (to the chosen activity parameters) must be intrinsically additive. Also, it is desirable for the data to have a high number of degrees of freedom since the greater the ratio of number of observations to number of unknowns, the more significant are the results.

Dillard, *et al.*,[8] have recently reported experimental results[9] of the antitumor activities of 85 acetylenic carbamates, most of which are well suited for application of the regression analysis. Following the procedures of Johnson, *et al.*,[10] the antitumor activities of various analogs of the substituted acetylenic carbamates (I) were tested against subcutaneously implanted tumors in mice.[8] The tumors used were X5563, a plasma-cell tumor, and C1498, an atypical myelogenous leukemia.[8] Dillard, *et al.*,[8] reported the per cent inhibition of the tumor X5563 and the per cent prolongation of life for those animals with tumor C1498. We analyzed these data (1) to rank the antitumor activities of the substituent groups and note possible structure–activity relationships, and (2) to predict the compounds of the series not tested, and possibly not synthesized, which would have the greatest potential as tumor inhibitors.



I

Calculations.—By assuming the activity contributions of the substituent groups on the parent structure to be constant and additive to the total activity of the

(1) This research is being supported by the U. S. Army Medical Research and Development Command (DA-49-193-MD-2779) and the National Science Foundation (GB-4453). This paper is Contribution No. 283 from the Army Research Program on Malaria. Computer facilities were provided through Grant HE-09495 from the National Institutes of Health.

(2) S. M. Free, Jr., and J. W. Wilson, *J. Med. Chem.*, **7**, 395 (1964).

(3) J. Kopecký, K. Boček, and D. Vlachová, *Nature*, **207**, 981 (1965).

(4) C. Hansch and T. Fujita, *J. Am. Chem. Soc.*, **86**, 1616 (1964).

(5) C. Hansch, E. W. Deutsch, and R. N. Smith, *ibid.*, **87**, 2738 (1965).

(6) W. P. Purcell, *Biochim. Biophys. Acta*, **105**, 201 (1965).

(7) W. R. Smithfield and W. P. Purcell, *J. Pharm. Sci.*, **56**, 577 (1967).

(8) R. D. Dillard, G. Poore, D. R. Cassady, and N. R. Easton, *J. Med. Chem.*, **10**, 40 (1967).

(9) It should be recognized that the authors[8] noted that the activities reported "are the results of a specific dose–response test for each compound and should be considered in a qualitative manner in comparing relative potencies." This limitation should not be overlooked when analyzing the results of the regression.

(10) I. S. Johnson, H. F. Wright, G. H. Svoboda, and J. Vlantis, *Cancer Res.*, **20**, 1016 (1960).

molecule, the data of Dillard, *et al.*,[8] were analyzed according to the method of Free and Wilson.[2] For this model, one generates a linear equation for each observation (*e.g.*, eq 1) where a, b, c, and d represent

$$[C_6H_5]_a + [C_6H_5]_b + [H]_c + [CH_3]_d + \mu = 7.533 \quad (1)$$

positions R, R_1, R_2, and R_3, respectively, on the parent structure while the brackets indicate the activity contribution of the substituent group at that position to the total activity; μ represents the activity contribution of the parent structure. Thus, $[C_6H_5]_a$ is the activity contribution of a phenyl group at position R; $[C_6H_5]_b$ represents the activity contribution of a phenyl group at position R_1; $[H]_c$ is the activity contribution of a hydrogen at position R_2; and $[CH_3]_d$ is the activity contribution of a methyl group at position R_3. The sum of these individual group contributions plus μ in each equation is equal to the biological response parameter (eq 2). In searching for

$$\text{Biological Response} = \mu + \Sigma \text{ group contributions} \quad (2)$$

a biological response parameter which could be used comparatively, the antitumor activity percentages reported by Dillard, *et al.*,[8] divided by the dosages (mg/kg) administered to the test animals were selected. Thus, the activity contributions in these regressions were set equal to the per cent inhibition of the tumor/mg of drug per kg of the test animal in system X5563 and the per cent prolongation of life/mg of drug per kg of the animals tested with tumor C1498 (*e.g.*, biological response in eq 1 is 7.533% prolongation of life/mg per kg).

From the assumption that position R is equivalent to position R_1, and R_2 is equivalent to R_3, the linear equations can be simplified by reducing the number of unknowns. For example, eq 1 reduces to eq 3. One

$$2[C_6H_5]_a + [H]_c + [CH_3]_c + \mu = 7.533 \quad (3)$$

can then generate 65 simultaneous equations with 36 unknowns for system X5563 and 69 equations with 37 unknowns for tumor C1498. In addition, following the method of Free and Wilson,[2] two restrictions (symmetry equations) were applied, namely, summations to zero of the group contributions at positions R,R_1 and R_2,R_3. The 67 simultaneous equations for system X5563 and the 71 simultaneous equations for tumor C1498 were then solved independently by the method of least squares using the IBM 1620 computer.

Results and Discussion

Least-squares solution of the linear equations yielded the calculated activity contribution of each substituent group as well as that of the parent structure. The calculated total activity of each molecule was found by summation of these group contributions and μ (eq 2). These calculated activities were then plotted against the observed biological activities; most of the points were very near the 45° line of the graph, indicating a relatively good correlation between the two sets of values. It was observed, however, that certain groups (cyclohexyl, propynyl, and tetramethylene) substituted at positions R_2,R_3 probably possessed nonadditive properties because there were large deviations between their corresponding points (plot of tumor X5563) and the ideal 45° line. Of these, only the cyclohexyl group, which had the highest calculated activity, exhibited an activity contribution (5.040% inhibition/mg per kg) sufficient to be considered as active. Similarly, the regression analysis of system C1498 indicated that three substituent groups (cyclohexyl, tetramethylene, and ethyl) exhibited nonadditive activity contribution properties. Here, too, the magnitude of the activity contribution of only the cyclohexyl group (6.679% prolongation of life/mg per kg) warrants its consideration in potential antitumor agents of this type.

These observations prompted the deletion of data obtained from the "nonadditive" groups, and the simultaneous equations were once again solved.

From the results of these second analyses, the ranks of antitumor activity contributions by the substituted groups in each tumor system were found to be quite similar (Table I). It was observed that 5 of the 18 groups substituted at positions R,R_1 and 6 of the 14 groups substituted at positions R_2,R_3 contributed constructively (positive sign, Table I) to the antitumor activity of the compounds tested against tumor X5563. On the other hand, only 3 out of the

TABLE I

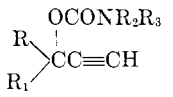
SUBSTITUENT GROUP CONTRIBUTION TO ANTITUMOR ACTIVITY

Tumor X5563		Rank	Tumor C1498	
Group	Activity[a]		Group	Activity[b]
Substituents at Position R,R_1				
2-$C_{10}H_7$	2.604	1	2-$C_{10}H_7$	2.047
4-FC_6H_4	1.162	2	4-FC_6H_4	1.499
1-$C_{10}H_7$	0.675	3	C_6H_5	1.240
$ClCH_2$	0.570	4	4-ClC_6H_4	−0.381
C_6H_5	0.570	5	1-$C_{10}H_7$	−1.297
4-ClC_6H_4	−0.025	6	4-IC_6H_4	−1.341
4-IC_6H_4	−0.435	7	4-$C_6H_5C_6H_4$	−1.845
4-BrC_6H_4	−1.012	8	4-BrC_6H_4	−2.176
2-C_5H_4N	−2.051	9	3,4-$Cl_2C_6H_3$	−3.416
2-ClC_6H_4	−2.246	10	2,4-$Cl_2C_6H_3$	−3.416
3-ClC_6H_4	−2.322	11	3-ClC_6H_4	−3.443
2,4-$Cl_2C_6H_3$	−2.360	12	3-BrC_6H_4	−3.665
3,4-$Cl_2C_6H_3$	−2.382	13	CH_3	−3.697
CH_3	−2.657	14	H	−3.697
H	−2.657	15	4-$CF_3C_6H_4$	−3.697
4-$O_2NC_6H_4$	−2.657	16	4-$CH_3C_6H_4$	−3.697
4-$CH_3C_6H_4$	−2.657	17	2-C_4H_3S	−3.697
2-C_4H_3S	−2.657	18	2-ClC_6H_4	−3.970
		19	2-C_5H_4N	−4.087
Substituents at Position R_2,R_3				
Cycloheptyl	3.407	1	Cyclopentyl	5.555
$CH_2CH{=}CH_2$	1.265	2	Cycloheptyl	4.285
Cyclopentyl	1.207	3	$CH_2CH{=}CH_2$	1.141
$(CH_2)_3N(CH_3)_2$	0.873	4	$CH_2C{\equiv}CH$	0.589
CH_3	0.372	5	CH_3	0.412
Cyclooctyl	0.073	6	H	0.021
C_2H_5	−0.027	7	$(CH_2)_3N(CH_3)_2$	−1.915
H	−0.033	8	Cyclopropyl	−2.315
$CH_2CH_2C_6H_5$	−2.477	9	Cyclooctyl	−2.348
4-ClC_6H_4	−3.060	10	NH_2	−4.248
Cyclopropyl	−3.260	11	$CH_2CH_2C_6H_5$	−4.582
CH_2CH_2OH	−3.260	12	4-ClC_6H_5	−4.675
$CH_2C_6H_5$	−3.260	13	CH_2CH_2OH	−4.915
NH_2	−3.260	14	$CH_2C_6H_5$	−4.915

[a] Activity is given as per cent inhibition of the tumor/mg of drug per kg of the test animal. [b] Activity is given as per cent prolongation of life/mg of drug per kg of the test animal.

TABLE II

OBSERVED AND CALCULATED ANTITUMOR ACTIVITY OF VARIOUS SUBSTITUTED ACETYLENIC CARBAMATES AGAINST TUMOR X5563

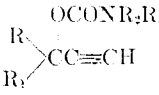


	Substituents at position R,R1																					Substituents at position R2,R3																Antitumor activity[a]	
No.	ClCH2	CH3	C6H5	H	4-ClC6H4	3-ClC6H4	2-ClC6H4	3,4-Cl2C6H3	2,4-Cl2C6H3	4-BrC6H4	4-FC6H4	4-IC6H4	4-O2NC6H4	4-CH3C6H4	2-C10H7	1-C10H7	2-C5H4N	2-C4H3S	H	CH3	C2H5	CH2CH=CH2	Cyclopropyl	CH2CH2OH	CH2C6H5	CH2CH2C6H5	Cyclopentyl	Cycloheptyl	Cyclooctyl	(CH3)2N((CH2)2	4-ClC6H4	NH2	Obsd	Calcd[b]
1	1	1																	2														0.000	0.000
2		1	1																2														0.000	0.000
3			2																2														2.213	3.227
4			2																1	1													4.267	3.632
5			2																1		1												3.233	3.233
6			2																1			1											5.555	4.525
7			2																1				1										0.000	0.000
8			2																1					1									0.000	0.000
9			2																1						1								0.000	0.000
10			2																1							1							0.783	0.783
11			2																1								1						4.467	4.467
12			2																1									1					6.667	6.667
13			2																1										1				3.333	3.333
14			2																1											1			4.133	4.133
15			2																1												1		0.200	0.200
16			2																1													1	0.000	0.000
17			2																	2													4.333	4.037
18			1	1															2														0.000	0.000
19			1		1														2														2.500	2.632
20			1		1														1	1													3.200	3.038
21			1		1														1			1											2.900	3.930
22			1		1															2													3.333	3.442
23			1			1													2														0.000	0.335
24			1			1													1	1													1.111	0.741
25			1			1														2													1.111	1.146
26			1				1												2														0.911	0.411
27			1				1													2													0.722	1.221
28			1					1											2														0.160	0.275
29			1					1												2													0.493	1.084
30			1						1										2														0.000	0.297
31			1							1									2														1.583	1.644
32			1							1									1	1													1.667	2.050
33			1							1										2													2.900	2.455
34			1								1								2														2.000	3.819
35			1								1									2													6.667	4.629
36			1									1							2														2.222	2.222
37			1										1						2														0.000	0.000
38			1											1					2														0.000	0.000
39			1												1				2														8.000	5.261
40			1												1					2													3.333	6.071
41			1													1			2														3.333	3.333
42			1														1		2														0.667	0.606
43			1														1		1	1													1.533	1.011
44			1														1			2													0.833	1.416
45			1															1	2														0.000	0.000
46					2														2														1.550	2.038
47					2														1	1													4.167	2.444
48					2															2													1.667	2.848
49					1			1											2														0.387	−0.319
50					1				1										2														0.000	−0.297
51											2								2														2.857	4.411
52											2									2													6.667	5.222

[a] Activity is given as per cent inhibition of tumor/mg of drug per kg of test animal. [b] Calculated using eq 2 where $\mu = 2.153\%$ inhibition of tumor/mg of drug per kg of test animal.

TABLE III

OBSERVED AND CALCULATED ANTITUMOR ACTIVITY OF VARIOUS SUBSTITUTED ACETYLENIC CARBAMATES AGAINST TUMOR C1498



	Substituents at position R,R1																					Substituents at position R2,R3																					Antitumor activity[a]	
No.	C_6H_5	CH_3	H	4-ClC_6H_4	3-ClC_6H_4	2-ClC_6H_4	3,4-$Cl_2C_6H_3$	2,4-$Cl_2C_6H_3$	4-BrC_6H_4	3-BrC_6H_4	4-FC_6H_4	4-IC_6H_4	4-$CF_3C_6H_4$	4-$CH_3C_6H_4$	4-$C_6H_5C_6H_4$	2-$C_{10}H_7$	1-$C_{10}H_7$	2-C_5H_4N	2-C_4H_3S	H	CH_3	$CH_3CH_2CH_2$	$CH_3CH(CH_3)$	Cyclopropyl	CH_2CH_2OH	$CH_2C_6H_5$	$CH_2CH_2C_6H_5$	Cyclopentyl	Cyclohexyl	Cycloheptyl	$(CH_3)_2N(CH_2)_2$	4-ClC_6H_4	NH_2	Obsd	Calcd[b]
1	2																			1	1													7.533	5.327
2	2																			1		1												6.056	6.056
3	2																			1			1											5.600	5.504
4	2																			1				1										2.600	2.600
5	2																			1					1									0.000	0.000
6	2																			1						1								0.000	0.000
7	2																			1							1							0.333	0.333
8	2																			1								1						10.467	10.467
9	2																			1									1					9.200	9.200
10	2																			1										1				2.567	2.567
11	2																			1											1			3.000	3.000
12	2																			1												1		0.240	0.240
13	2																			1													1	0.667	0.667
14	2																				2													4.533	5.717
15	1	1																		2														0.000	0.000
16	1		1																	2														0.000	0.000
17	1			1																2														2.525	3.345
18	1			1																1	1													2.300	3.706
19	1			1																	2													2.067	4.096
20	1				1															2														0.000	0.254
21	1				1															1	1													1.089	0.644
22	1				1																2													0.844	1.035
23	1					1														2														0.000	−0.274
24	1					1															2													0.233	0.507
25	1						1													2														0.000	0.280
26	1						1														2													0.000	1.060
27	1							1												2														0.000	0.280
28	1							1													2													0.000	1.060
29	1								1											2														2.000	1.521
30	1								1											1	1													1.367	1.911
31	1								1												2													2.367	2.302
32	1									1										2														0.370	0.031
33	1									1											2													0.473	0.812
34	1										1									2														2.140	5.195
35	1										1									1			1											5.667	5.762
36	1										1										2													13.800	5.976
37	1											1								2														2.356	2.356
38	1												1							2														0.000	0.000
39	1													1						2														0.000	0.000
40	1														1					2														3.133	1.851
41	1														1						2													1.350	2.632
42	1															1				2														6.800	5.715
43	1															1					2													5.467	6.524
44	1																1			2														2.400	2.400
45	1																	1		2														0.000	−0.390
46	1																	1		1	1													0.000	0.000
47	1																				2													0.000	0.390
48	1																		1	2														0.000	0.000
49				2																2														1.000	1.604
50				2																1	1													4.625	2.084
51				2																	2													1.400	2.474
52				1			1													2														0.000	−1.341
53				1				1												2														0.000	−1.341
54											2									2														2.686	5.451
55											2										2														6.667	6.235

[a] Activity is given in per cent prolongation of life/mg of drug per kg of test animal. [b] Calculated using eq 2 where $\mu = 2.414$% prolongation of life/mg of drug per kg of test animal.

19 groups substituted at positions R,R_1 and 6 groups of 14 substituted at R_2,R_3 were not deleterious to the activity of those compounds inhibiting tumor C1498. Also, these rankings were noticeably similar to those of the original regressions in that the number of active groups was the same in each and many of them were in the same relative order. Too, the significance of the correlation increased markedly in the second regressions indicating that statistically they are considerably better: the correlation coefficient, level of significance of F ratio,[11] and Ψ[12] for system X5563 changed from 0.816, 0.940, and 0.578 to 0.915, 0.995, and 0.403, respectively, while the corresponding values for system C1498 went from 0.800, 0.900–0.950, and 0.600 to 0.882, 0.975–0.995, and 0.471, respectively. Tables II and III give the calculated and observed activities for the second regression analyses. Included in the total calculated activities is the calculated value of μ for each system; μ = 2.153% inhibition of tumor/mg of drug per kg of test animal for tumor X5563 and μ = 2.414% prolongation of life/mg of drug per kg of test animal in system C1498. Deviations between the calculated and observed activities for most observations are quite small; of course, there is necessarily no activity deviation for those compounds with substituent groups observed only once.

In other attempts to find a more significant regression, the linear equations were solved using the logarithms of the biological responses as the activity parameters since logarithms of biological activity data are often considered free-energy related, and therefore may be additive. These calculations were based on all substituent groups analyzed in the first calculations (*i.e.*, no groups of the original data were deleted in these regressions). From calculation of the F ratios,[11] Ψ,[12] and correlation coefficients, it was found that these regressions of logarithms were statistically better than the original calculations. In order to make a valid comparison of the degree of fit between the linear data and logarithmic data, calculated total activity was plotted against the observed total activity for each observation using the antilogarithms of the results of the logarithmic regression. Table IV summarizes the statistical results and makes it clear that the preferred choice of biological response parameter for tumor C1498 is the original linear data and not their logarithms.

It is important to emphasize that one could be misled from the statistics of the regression analysis using the logarithms of the original linear data (correlation coefficient = 0.927, level of F ratio[11] = 0.995, Ψ[12] = 0.373) which are better than those for the original linear data (Table IV). Statistical calculations (correlation coefficient, F ratio,[11] and Ψ[12]) alone are not suitable as a means of judging the predictive utility of regression analyses of this type.

Perhaps the most interesting point in this study is the fact that several molecules which were not tested *in vivo* have calculated antitumor activities greater than those tested. It would appear that the best compounds would contain 2-naphthyl, 4-fluorophenyl, or phenyl groups or combinations of them substituted at positions R and R_1 with cyclohexyl, cycloheptyl, or cyclopentyl groups or combinations at R_2 and R_3. For example, the calculated activity of the molecule, 1,1-(2,2′-dinaphthyl)-2-propynyl N,N-dicycloheptylcarbamate, with 2-naphthyl groups substituted at positions R and R_1 and cycloheptyl groups substituted at R_2 and R_3 is 14.175% inhibition/mg per kg against tumor X5563 while the most potent molecule tested, 1-(2-naphthyl)-1-phenyl-2-propynyl carbamate, had an observed activity of only 8.000% inhibition/mg per kg. On the other hand, when 2-naphthyl groups are substituted at positions R and R_1 along with cyclopentyl groups at R_2 and R_3, the molecule, 1,1-(2,2′-dinaphthyl)-2-propynyl N,N-dicyclopentylcarbamate, exhibits a calculated activity of 17.618% prolongation of life/mg per kg against tumor C1498 compared with the 13.800% prolongation of life/mg per kg of 1-(4-fluorophenyl)-1-phenyl-2-propynyl N,N-dimethylcarbamate which is the highest observed activity. One would conclude that these and similar promising compounds are worthy of synthesis and testing.

In the more active molecules of this series, substituents at positions R_2 and R_3 contribute more to the calculated total activity of the molecule than do substituents at R and R_1. An example of this is 1,1-diphenyl-2-propynyl N-cyclopentylcarbamate (**8**, Table III), the compound tested with the highest calculated total activity. Of the total calculated activity of 10.467% prolongation of life/mg per kg, an activity of only 2.480% prolongation of life/mg per kg was contributed by positions R and R_1 compared with 5.576% prolongation of life/mg per kg contributed by positions R_2 and R_3. This observation holds true uniquely well for those molecules predicted as potent antitumor agents and is also illustrated in Table I; the calculated activities of the more potent substituent groups at R and R_1 (2.047 and 1.499% prolongation of life/mg per kg) are only half these of the more active substituent groups at R_2 and R_3 (5.555 and 4.285% prolongation of life/mg per kg).

TABLE IV

STATISTICAL RESULTS OF REGRESSION ANALYSES OF LINEAR AND LOGARITHMIC DATA (SYSTEM C1498)

Results of regression analysis of	Cor coef	Signif of F ratio[a]	Ψ[b]	Σ (obsd − calcd)^2[c]
Original linear data	0.800	0.90–0.95	0.600	527.6
Logarithms of original linear data (results converted to antilogarithms)	0.636	<0.75	0.793	920.8

[a] Reference 11. [b] Reference 12. [c] Sum of squares of deviations between the observed and calculated activities.

Acknowledgment.—The authors wish to thank Mr. M. M. Marsh, Dr. N. R. Easton, and Mr. R. D. Dillard of the Lilly Research Laboratories for their helpful comments on the manuscript. They also wish to thank Mr. K. Sundaram and Miss J. A. Singer for reading and commenting on the manuscript.

(11) G. W. Snedecor, "Statistical Methods," 5th ed, The Iowa State College Press, Ames, Iowa, 1956, pp 417–420, 276–279.

(12) O. Exner, *Collection Czech. Chem. Commun.*, **31**, 3222 (1966).