A Two-Component Approach to Predicting Antitumor Activity from Chemical Structure in Large-Scale Screening

Louis Hodes

National Cancer Institute, Bethesda, Maryland 20892. *Received March 28, 1986*

A new way of combining physicochemical parameters such as the octanol/water partition coefficient (or log *P*) with molecular structure features has been devised to predict antitumor activity on a large, diverse set of compounds. This is done by adding the log *P* parameter as a separate component to an earlier method based on structure. The two-component approach is motivated by the dual concepts of accessibility and specificity. Extensive testing shows improvement in performance for the two-component method over the use of structure alone. All the compounds with definitive biological results in the in vivo NCI prescreen form the training set. The method separates the training set into disjoint subsets depending on the range of log *P*. Therefore, structure fragments receive activity weights that vary with the log *P* range. This change in weights accounts for the improved performance rather than any difference in the structural characteristics of the compounds in the different log *P* ranges.

This is a report on a method for combining the Hansch octanol/water partition coefficient[1] with molecular structure fragments as predictors of anticancer activity for a large, diverse set of compounds. The concepts of accessibility and specificity of compounds will be used to motivate a two-component approach. Then this approach will be applied to the National Cancer Institute data.

The original method[2] was first described in 1977 and has been used by the NCI Developmental Therapeutics Program since 1980 as an aid in selecting compounds for screening. Early studies[3] to validate the method include a comparison of its performance with that of a chemist familiar with the data.

The method was designed to handle large numbers of diverse compounds that cannot readily be classified. The use of thousands of structure fragments as variables precluded the use of regression and other quantitative structure–activity relationship (QSAR) methods normally employed[4] to predict activity within one or more classes.

Under these conditions, one cannot merely account for the partition coefficient as another variable in addition to the large number of structure fragments. This would imply a single optimum partition coefficient for all the diverse structures. What we require is a method that allows the optimum partition coefficient to vary with structure. Or failing that, we allow structure fragment activity weights to depend on the partition coefficient.

Review of the Original Method. All the compounds with definitive biological results in the in vivo NCI prescreen[5] form the training set. These compounds are subjected to exhaustive generation of structure fragments of certain sizes and types.[6] This process yields about 10000 distinct fragments from the current training set of over 100000 compounds.

The incidence of a given fragment in the entire training set yields its expected incidence in the active portion under the asumption of homogeneity, that is, irrelevance of the fragment to activity. The actual incidence of the fragment in the active portion of the training set minus this expected incidence, expressed in standard deviations, is taken as the activity weight for the fragment.

A new compound is evaluated by summing the weights of the fragments it contains. A routine has been established under which a 20% subset of the training set is run through to check the performance and establish percentile scores for further use. This will be amplified further when the new method is described.

For example, the entire training set is represented in the center of the diagram in Figure 1. Here it is assumed that the active compounds comprise 10% of the training set, at the top. To the left is represented the subset of compounds containing some fragment A. About 7% of these compounds are active. Fragment A will therefore receive a negative weight for activity. Similarly the right-hand box represents the portion of compounds containing some other fragment B. This fragment would get a positive weight for activity because about 15% of the compounds are active. Of course, all three boxes would overlap, but they are separated for illustration only.

An Anomaly and Its Resolution. Now suppose fragment A were a simple isolated (unfused) benzene ring and fragment B were a sugar ring. The above are actually the kinds of statistics these two specific fragments yield. However they are not generally considered to be involved in anticancer activity. For a long time it was felt that these readings represented statistical artifacts in the data set.[7]

The anomaly can be resolved if one imposes a hypothetical accessibility dimension onto the data as shown in Figure 2. Thus one can separate the inaccessible compounds which will have no chance to be active. Let us assume and show in Figure 2 that in the entire file 40% of the compounds, for one reason or another, wound up as inaccessible. We can then derive the proportions of compounds containing fragments A and B so that, for the accessible compounds, exactly the same fraction will be active as in the entire file.

In other words, as is shown in Figure 2, the negative activity weight of fragment A and the positive activity weight of fragment B are a result of accessibility properties rather that specific drug–receptor interaction. Benzene rings tend to render their compounds inaccessible. It is probably not possible in practice to separate the effects of the two components. However, such a hypothetical separation can be used in interpreting the results and modifying the method.

The large negative activity weight of the benzene ring was the main reason for an early modification. The me-

(1) Hansch, C.; Fujita, T. *J. Am. Chem. Soc.* **1964**, *86*, 1610.
(2) Hodes, L.; Hazard, G. F.; Geran, R. I.; Richman, S. *J. Med. Chem.* **1977**, *20*, 469.
(3) Hodes, L. *J. Chem. Inf. Comput. Sci.* **1981**, *21*, 128.
(4) See, for example: Nakayama, A.; Inamura, H.; Fujita, T. *J Med. Chem.* 1984, *27*, 1493.
(5) For a description of the P388 mouse leukemia prescreen, see: *In Vivo Cancer Models*, NIH Publication No. 84-2635; U.S. Government Printing Office: Washington, DC, 1984. The training set consisted of compounds with conclusive P388 results, as of September 1983, from which certain well-known classes of compounds have been removed. For the A, C, and N activity criteria, see ref 7.
(6) Hodes, L. *J. Chem. Inf. Comput. Sci.* **1981**, *21*, 132.
(7) Hodes, L. In *Computer-Assisted Drug Design*; Olson, E. C., Cristoffersen, R. E., eds.; ACS Symposium Series 112; American Chemical Society: Washington, DC, 1979; pp 592–593.

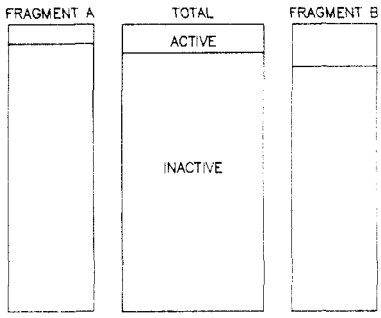

Figure 1. Hypothetical training set of compounds. Ten percent of the total are active. Subset containing fragment A extracted to the left in which 7% of the compounds are active. Subset containing fragment B extracted to the right in which 15% of the compounds are active.

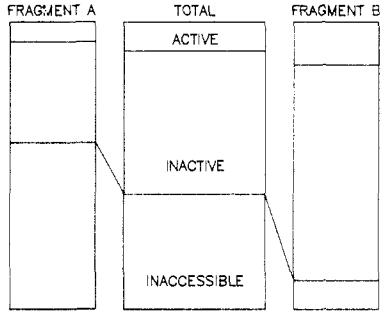

Figure 2. Compounds of Figure 1 where the difference of activity between fragments A and B is explained by accessibility.

thod was modified to diminish the effect of high incidence fragments, such as the benzene ring, which seemed to have extreme weights. For every fragment with incidence greater than 10 compounds, the weight was divided by the logarithm of the incidence. Thus, if the benzene ring appeared in 50 000 compounds and had an activity weight of −15 standard deviations, the weight would be effectively −15/4.7 or −3.2. The weight of the sugar ring, having less incidence, would not be reduced so dramatically.

Accessibility and Specificity. As implied in the previous section, we can separate the action of a drug into two components even though this distinction may not always be useful in practice. The first component will be called accessibility. Following Cramer,[8] this refers to noncovalent interactions that can be derived from physicochemical parameters, that is, those due to averages of tumbling molecules rather than specific shapes or structure. Generally, accessibility implies those passive interactions such as solubilities that allow a drug to reach its receptor.

The second component is what Cramer calls specificity. This encompasses the covalent or shape-related interactions between drug and receptor, which depend on precise structure.

Most work in QSAR consists of series optimization where the distinction between accessibility and specificity is not often important. One or the other can become the dominant form of expression. Moreover, the variables are low in number and constrained so that almost any systematic method can produce equivalent, if not satisfactory, results.

For a truly diverse set of compounds, however, the distinction between accessibility and specificity becomes more important. The use of nonspecific accessibility variables causes Cramer to limit the range of his work to areas like anethesia where the passive accessibility variables account for most of the reaction. Other works that explicitly combine the two types of components either restrict the compounds so they are not truly diverse or else do not allow the full interplay between the accessibility and specificity variables. The work reported here attempts to allow this interplay in a more adequate manner.

log *P* as a Separate Component. The earlier analysis of the seemingly anomalous weights has shown how activity weights for structural fragments do incorporate both accessibility and specificity aspects. However, if these aspects can be separated somehow, then the activity weights will be more precise. For example, a sugar ring in a hydrophobic molecule, by moderating the hydrophobicity, should have a better biological effect than a sugar ring in an already hydrophilic molecule.

The two-component method will be illustrated with use of the octanol/water partition coefficient or log *P* to provide the accessibility component. We chose log *P* because of the large amount of available data and its empirical relation to accessibility phenomena. The method involves dividing the training set into two or more disjoint subsets according to ranges of log *P*. Thus, the activity weight of a fragment will depend on the estimated range of log *P* of the compound in which it occurs.

The available methods for calculating log *P* did not apply very readily to our diverse and large set of compounds. However, there was not need for actual estimates of log *P* but only an estimate of to which of several ranges of log *P* each compound belongs. This was done by using measured log *P* values for 4013 compounds[9] with structures obtained from their CAS registry numbers.

The 4013 compounds with measured log *P* values were sorted in increasing order of log *P*. Most of the compounds were on the positive, hydrophobic side. The lowest 1000 went from −5 to +0.57. These were taken as a low log *P* training set. Opposed to these were the top 2000 compounds, taken as the high log *P* training set, with values from 1.7 to 8. In this manner, the same method ordinarily used for anticancer data or toxicity data was used to create a model for log *P*. That is, fragment weights were computed based on the difference of incidence of the fragment in the low log *P* set vs. the high log *P* set. A compound could then be evaluated for low log *P* by summing the weights of its fragments. This sum provides a measure that the compound belongs to the low log *P* range. A test run of a 20% subset showed a good separation of the log *P* compounds.

First Version. The entire P388 training set[5] was then passed through the log *P* model, yielding a continuous estimate of relative log *P* over all the P388 compounds. Then an arbitrary score, in this case zero, was chosen to separate the low from the high log *P*.

This procedure yielded a two-component training set where the first component, log *P*, was simply two-valued, high or low. The separating score of zero placed a majority of the P388 compounds in the high log *P* subset. It was notable that the active compounds tended to be less hydrophobic than the inactive compounds in our collection. The ratio was about 2:1 for the highly active compounds, labeled A; 3:1 for the moderately active compounds, labeled C; and 4:1 for the inactive compounds, labeled N.

The activity weights for the two-component model were then generated. As expected, the sugar-related fragment

(8) R. D. Cramer, III *Quant. Struct.-Act. Relat. Pharmacol., Chem. Biol.* **1983**, 2, 7.

(9) Pomona College Medicinal Chemistry Project Data Base, Issue 23, July 1983.

Table I. Comparison of Performance of Two-Component Method vs. Original Method[a]

	cumulative % actives			
	original		two component	
percentile	A	C	A	C
99	27	10	27	10
98	39	16	43	17
95	71	30	72	33
90	88	44	91	47
80	92	59	95	63
70	97	70	99	72
50	100	82	100	85
30		92		94
10		99		99
0		100		100

[a] Test sets are identical 20% subset of training set. Cumulative percentages of active compounds are shown at selected percentile levels. The columns labeled A refer to the compounds from the highly active portion of the training set. The columns labeled C refer to the moderately active compounds, which still meet the criterion for passing the screen. Note the concentration of active compounds above the 80th percentile of the ranking.

Table II. log *P* Distribution of Compounds with CAS Registry Numbers and Measured log *P*

log *P* range	no. of compounds
−5 to −1	270
−1 to 1	1094
1 to 3	1990
3 to 8	659
total	4013

O R– C R– C – O (O–C–C–O where the first two bonds are ring bonds) received higher activity weight in the high log *P* range. This fragment appeared in 340 actives out of 608 compounds in the high log *P* range for a weight of 20.9 and 402 actives out of 1310 total in the low log *P* range for a weight of 13.9.

Overall performance did show improvement for the two-component model when run on the same 20% subset of the training set. Performance can be measured by the relative number of actives scoring highly. The cumulative percentages of actives at selected percentile levels are compared to those for the original unsplit model in Table I.

The distributions in Table I were tested for statistical significance by the Smirnov test.[10] This yielded a *p* value of about 0.22 in the one-tailed form. A more powerful test, the Wilcoxon rank-sum test, yielded a difference of about 1.5 standard deviations or a *p* value of 0.06, one-tailed.

Further experimentation showed that the gain in performance was due to a somewhat larger gain in performance on the high log *P* subset. Perhaps splitting the compounds into more ranges of log *P*, especially at the lower end, would produce better results. Also, all the log *P* data could be used by no longer eliminating the roughly 1000 midrange compounds from the log *P* data.

Second Version. Now the 4013 compounds with measured log *P* values were separated into four ranges of log *P* as shown in Table II. Each subset of the log *P* data was taken in turn as the active set to produce four separate fragment-weight tables. In each table the weights now signified the likelihood that a compound with the fragment belonged to the given range of log *P*.

(10) All statistical tests are taken from Lehmann, E. L. *Nonparametrics: Statistical Methods Based on Ranks*; Holden-Day: San Francisco, 1975.

Table III. Derived log *P* Distribution of P388 Training Set[a]

	number of compounds		
log *P* range	A	C	N
−5 to −1	199	387	6806
−1 to 1	73	522	11665
1 to 3	146	867	27427
3 to 8	334	1576	47065
total	752	3352	92963

[a] Note the proportions of the highly active A compounds, the moderately active C compounds, and the remaining inactive N compounds.

Table IV. Sugar Ring Fragment Statistics in P388 Training Set[a]

	incidence		
log *P* range	active (A + C)	total (A + C + N)	activity weight
−5 to −1	202	679	10.54
−1 to 1	130	475	8.43
1 to 3	225	442	16.86
3 to 8	184	324	17.15

[a] The incidence of the sugar ring is given for the compounds as enumerated in Table III.

Table V. Comparison of Performance of the Second Version of the Two-Component Method vs. the Original Method[a]

	cumulative percent actives			
	original		two component	
percentile	A	C	A	C
99	27	10	37	11
98	39	16	48	20
95	71	30	76	33
90	88	44	86	47
80	92	59	97	63
70	97	70	99	73
50	100	82	99	88
30		92	100	94
10		99		99
0		100		100

[a] Test sets are identical 20% subset of training set. Cumulative percentages of active compounds are shown at selected percentile levels. These are the same compounds subjected to the first version presented in Table I. The two-component method now has the training set stratified as shown in Table III.

At this stage the P388 training set was run through all four log *P* models using the tables just created. Each compound was assigned to that range of log *P* for which it received the highest score. The distribution of log *P* determined in this manner for the highly active compounds (A), the moderately active compounds (C), and the inactive compounds (N) is shown in Table III. In this manner, a new two-component model was created. For example, the sugar fragment O R– C R– C – O now had four different weights depending on the range of log *P* as shown in Table IV. The lack of differentiation between the values at the low end suggests that it would not be worthwhile to split the log *P* ranges further. For comparison, the same 20% subset used earlier was run through the new version. Each compound was evaluated by summing the fragment weights in its assigned log *P* range. The cumulative percentages of actives at selected percentile levels are shown in Table V.

These results were somewhat better than those from the first version. Notice the 37% of the A's fell in the top percentile vs. 27% in the first version and also in the original unsplit version. The best difference for the C's occurs at the 54th percentile, skipped in Table V, where 88% of the C's are included but only 80% of the C's in the

Table VI. Derived log P Distribution of Disjoint Test Set[a]

log P range	number of compounds A	C	N
−5 to −1	5	19	322
−1 to 1	9	27	492
1 to 3	15	49	1215
3 to 8	28	60	1987
total	57	155	4016

[a] Compare the distribution to that of Table III.

Table VII. Comparison of Performance of the Second Version of the Two-Component Method vs. the Original Method[a]

percentile	cumulative percent actives: original A	original C	two component A	two component C
99	11	5	14	6
98	14	8	23	15
95	35	21	44	24
90	56	39	60	37
80	72	52	70	55
70	84	62	82	62
50	93	77	91	74
30	100	86	95	86
10		98	100	99
0		100		100

[a] Test set is that shown in Table VI disjoint from the training set. Cumulative percentages of active compounds are shown at selected percentile levels. Note that the concentration of the active compounds above the 80th percentile level of the ranking is not as pronounced as that for the subset test set of Tables I and V.

original version. The Smirnov test is a measure of the significance of this maximum difference. In this case we get a significance level of 0.044. The Wilcoxon rank-sum test, again performed roughly by considering the data at the same level in Table V to be at tied rank, gives a significance level of 0.03, one-tailed.

Although this experiment shows a definite enhancement in performance through the use of log P, some of the improvement must be due to an increased subset effect. When the test set is a subset of the training set, performance tends to be somewhat better than for a disjoint test set. For example, several compounds will get high scores if they have fragments that occur only once, but these fragments are ignored in the disjoint case.

Since both runs were done with exactly the same subset test set, the comparison seems fair. However, the addition of the log P parameter effectively multiplies the number of fragments, increasing the likelihood that a compound will have a single two-component fragment. Would a disjoint test provide a better comparison?

A Disjoint Test Set. A test set disjoint from the training set was obtained by using the compounds in the March 1984 update of the training set that were not in the September 1983 training set. These additional data consisted of 57 compounds in set A, 155 compounds in set C, and 4016 compounds in set N.

Prior to running the new test set against the two-component model just described as the second version, each compound in the test set was rated as to range of log P with use of the same four log P training sets that were used to stratify the two-component model. The split of the new test set into ranges of log P is shown in Table VI. In this small set we do not see a large proportion of hydrophilic active compounds as Table III shows in the earlier training set.

This test set was evaluated first with use of the original September 1983 training set without the log P enhancement and again with use of the same training set in its two-component version. The two outcomes are presented in Table VII at selected percentile levels. The results show a large improvement in performance for the two-component model at the 98th percentile level, which deteriorates as the percentiles decrease. To compare the performance on this test set, we were prepared to use a more refined

Table VIII. Preparation of Active Compounds for Wilcoxon Signed-Rank Test

no.	percentile: two component	original	difference	difference midrank	no.	percentile: two component	original	difference	difference midrank
1	98	98	0	n/a	30	79	97	−18	47
2	98	100	−2	23	31	40	32	8	41.5
3	94	94	0	n/a	32	100	100	0	n/a
4	62	64	−2	23	33	100	98	2	23
5	66	47	19	48	34	93	91	2	23
6	87	87	0	n/a	35	99	97	2	23
7	85	63	22	49.5	36	76	86	−10	43
8	65	87	−22	49.5	37	100	100	0	n/a
9	97	93	4	33.5	38	82	76	6	38
10	100	100	0	n/a	39	99	98	1	15
11	92	84	8	41.5	40	100	100	0	n/a
12	99	98	1	15	41	29	52	−23	51
13	96	68	28	53.5	42	76	75	1	15
14	88	92	−4	33.5	43	94	91	3	29.5
15	18	45	−27	52	44	98	81	17	46
16	72	85	−13	44.5	45	98	96	2	23
17	41	77	−36	55	46	98	95	3	29.5
18	63	70	−7	40	47	83	96	−13	44.5
19	68	74	−6	38	48	96	95	1	15
20	97	93	4	33.5	49	76	82	−6	38
21	75	31	44	56	50	99	96	3	29.5
22	100	100	0	n/a	51	96	95	1	15
23	98	96	2	23	52	93	95	−2	23
24	98	97	1	15	53	93	90	3	29.5
25	100	100	0	n/a	54	94	92	2	23
26	100	100	0	n/a	55	99	98	1	15
27	94	89	5	36	56	24	73	−49	57
28	84	56	28	53.5	57	75	79	−4	33.5
29	91	91	0	n/a					

Table IX. Comparison of Performance When Test Compounds Are Scrambled by Assignment to Arbitrary log *P* Ranges

	cumulative percent actives			
	scrambled		two component[a]	
percentile	A	C	A	C
99	14	3	14	6
98	23	11	23	15
95	33	20	44	24
90	42	28	60	37
80	61	43	70	55
70	67	55	82	62
50	77	68	91	74
30	89	81	95	86
10	100	97	100	99
0		100		100

[a] The two-component performance is copied from Table VII.

statistical test, the Wilcoxon signed-rank test. For each active compound, the difference in percentile levels is taken with a positive sign if the two-component result was higher and a negative sign if the original result was higher. Compounds at the same percentile level are considered tied. Table VIII shows the percentile levels of the 57 A compounds and the ranking of their differences that is used in the Wilcoxon signed-rank test.

The Wilcoxon signed-rank test on the data of Table VIII shows an improvement of 0.8 standard deviation for the two-component model, whereas the C compounds show a smaller improvement of 0.34 standard deviation. Failure to produce a larger improvement may be due to another important effect that was masked by the subset effect described earlier, the effect of a smaller training set. That is, the two-component method has the effect of splitting the training set into disjoint sets, one for each predetermined range of log *P*. Thus, we are effectively dealing with four smaller training sets. All of our experience has shown that performance improves with the size of the training set. A larger size helps counteract diversity. In this case, it is possible that a given fragment may occur in the training set at one range of log *P* and the test set at different ranges of log *P*. The fragment would be evaluated to zero, regardless of its activity weight.

The following experiment was performed to get a measure of the small training set effect. The two-component model was run by assigning test set compounds to log *P* ranges regardless of their actual estimated log *P* range, merely requiring that the distribution was the same so that each training set range received the same number of compounds as in the earlier two-component run. That is, following Table VI, for the A's the first five compounds were assigned to log *P* below −1, the next nine to log *P* between −1 and 1, and so forth.

The cumulative percent actives at the selected percentile levels are shown in Table IX. As expected, the performance is a great deal worse than the correctly assigned two-component run. The A's were 1.1 standard deviations lower and the C's 2.27 standard deviations lower by the Wilcoxon signed-rank test.

This experiment exaggerates the small set effect since the training set had been systematically separated by log *P* and the test set had not, causing a possible mismatch. At this point there seemed to be another way to eliminate the small set effect and even improve the performance of the two-component model. The new idea was to restore the effect of the complete training set by filling gaps in the two-component fragment-weight table. A test compound would no longer be restricted to a single log *P* range, but would be allowed to use the fragment weight of the closest log *P* range. That is, if a section does not contain a given fragment, obtain the fragment and its weight from the closest adjacent section.

Table X. Number of Distinct Fragments in Training Set at Various Ranges of log *P*[a]

log *P* range	no. of fragments
−5 to −1	3016
−1 to 1	4476
1 to 3	6038
3 to 8	7129
total	9478

[a] These are the fragments generaged for all the compounds tabulated in Table III.

Table XI. Comparison of Performance of the Second Version of the Two-Component Method vs. the Same Method with Augmented Fragment-Weight Table[a]

	cumulative percent actives			
	augmented		two component	
percentile	A	C	A	C
99	14	6	14	6
98	23	15	23	15
95	44	24	44	24
90	60	37	60	37
80	70	55	70	55
70	82	63	82	62
50	91	74	91	74
30	95	86	95	86
10	100	99	100	99
0		100		100

[a] Test set is disjoint from the training set. Cumulative percentages of active compounds are shown at selected percentile levels. The two-component performance is again copied from Table VII.

A Distributed Fragment-Weight Table. The number of fragments in each log *P* range is shown in Table X. There were a total of 9748 distinct fragments among the four ranges, which was equal to the number of fragments in the intact September 1983 P388 training set. Each range was augmented to 9748 fragments by means of the following procedure.

The fragments that were missing from any log *P* range or section were supplied with a log *P* value taken from the closest adjacent section. If the fragment was missing from one of the two middle sections and present on both immediately adjacent sections, the average of the weights from the two adjacent sections was used. Although the number of fragments was greatly increased, the results, shown in Table XI, were almost exactly the same as those from the earlier two-component run. It seemed as though each log *P* range had automatically selected fragments pertinent to that range, so that test compounds that were log *P* rated rarely had fragments outside their assigned range.

But the enlarged distributed fragment-weight table should improve performance on the previous run where the log *P* range of the test compounds was ignored and scrambled. This last experiment was performed with the expectation that the mismatched test compounds would now have their fragments matched with something like their true fragment weights since these weights would be shifted to all the ranges. Thus there should be a dramatic improvement in performance. The experiment just described was performed, and there was surprisingly little change from the earlier results of Table IX. See Table XII.

Now it became clear that the large number of fragments that were distributed across the log *P* ranges were, like the tail of a comet, irrelevant to the performance. They are mostly contained in small numbers of usually inactive

Table XII. Comparison of Performance When Test Compounds Are Scrambled by Assignment to Arbitrary log P Ranges[a]

	cumulative percent actives			
	scrambled		augmented	
percentile	A	C	A	C
99	14	3	14	3
98	23	11	23	12
95	33	20	33	19
90	42	28	47	30
80	61	43	61	43
70	67	55	67	55
50	77	68	79	71
30	89	81	91	81
10	100	97	100	97
0		100		100

[a] Original from Table IX vs. augmented fragment-weight table.

compounds. The true determinants of performance in the two-component model are distinctions among the differentiated weights of fragments already common to all the log P ranges.

Remarks and Conclusions

In constructing a two-component approach to apply to a diverse set of compounds, it was necessary to radically depart from some of the concepts used in standard Hansch analyses. First, there cannot be a single optimum value of log P. Second, the use of indicator variables for fragments allowed only one weight upon the presence of a fragment. Here, the weight is dependent on the range of log P.

Other stratifications of the training set besides log P can be tried. Along these lines, an earlier experiment in separating large and small compounds was not very satisfactory.

In summary, the experiments on the disjoint test set showed a significant loss in performance when the compounds were randomized over the log P ranges. When the fragment-weight table was augmented, the results were not greatly changed. This shows that the difference in preformance was due to a difference in weights for varying log P of fragments in the original table.

The large amount of testing of the second, more definitive version of the two-component model indicates the amount of improvement in performance that can be expected when ranges of log P are introduced into the earlier model that was based on structure alone. The improvement shown in the two main tests, Tables V and VII, appears mostly in the upper two percentiles of the score. Thus, the two-component model would be especially useful for automated literature surveillance where only the top few percent of compounds are examined.

Some of the compounds appearing in Table VIII that had poor ranking under the two-component model were examined. They would have ranked much higher with a different log P assignment. Perhaps with a more discriminating log P model they would have been classified into a more appropriate log P range. This may be achieved if there will be a lot more measured log P data. That points to the weakness of this approach. Data on 4000 compounds were used to classify 100000 more diverse compounds. The log P data were especially lacking toward the low log P end where performance was worst.

Acknowledgment. All of the programs and much of the programming were performed for NCI as part of a contract by Chemical Abstracts Service. Arthur Levitt, in charge of this work at CAS, contributed a great deal, including the essential idea to use the original method for the log P model.

Antibacterial Activity of Phosphono Dipeptides Related to Alafosfalin

Barbara Lejczak, Paweł Kafarski, Helena Sztajer, and Przemysław Mastalerz*

Institute of Organic and Physical Chemistry, Technical University of Wrocław, 50-370 Wrocław, Poland.
Received December 23, 1985

A series of dipeptides containing N-terminal alanine or leucine and a wide range of P-terminal racemic 1-aminoalkanephosphonates were prepared and tested in vitro for their ability to inhibit the growth of various bacterial species. The results demonstrate that peptides containing 4-amino-4-phosphonobutyric acid and 1-amino-1-methylethanephosphonic acid exhibit antibacterial activity comparable with that observed in the case of peptides containing P-terminal racemic 1-aminoethanephosphonic acid (analogue of alanine) used as a positive control.

For a substance to be an effective antimicrobial agent, it must be able to interfere with an essential function of the microbial cell. Target sites within the cell are often susceptible to inhibitors when tested in cell-free systems, but the intact microbe is often not susceptible to the same agents. This difference in inhibitory activity between intact and cell-free systems is commonly attributed to cell permeability, whereby elements of the cell membrane restrict the access of external molecules from the environment.

In recent years a variety of naturally occurring, as well as synthetic, antibiotics have been recognized that are analogues of small peptides and that function by entering susceptible microorganisms via peptide permeases and attacking intracellular targets. The inhibitory agent may be an intact peptide or a moiety released from it by intracellular hydrolysis.[1-3]

The most extensively studied antibiotics have been analogues of small peptides in which the C-terminal amino acid is replaced by the mimetics of alanine.[4-10] These

(1) Alper, M. D.; Ames, B. N. *J. Bacteriol.* 1978, *133*, 149.
(2) Ringrose, P. S. In *Microorganisms and Nitrogen Source*; Payne, J. W., Ed.; Wiley: Chichester, 1980; pp 641-692 and 805-807.
(3) Ringrose, P. S. *Biochem. Soc. Trans.* 1983, *11*, 804.
(4) Allen, J. G.; Atherton, F. R.; Hall, M. J.; Hassall, C. H.; Holmes, S. W.; Lambert, R. W.; Nisbet, L. J.; Ringrose, P. S. *Nature (London)* 1978, *272*, 56.
(5) Atherton, F. R.; Hall, M. J.; Hassall, C. H.; Lambert, R. W.; Ringrose, P. S. *Antimicrob. Agents Chemother.* 1979, *15*, 677.

0022-2623/86/1829-2212$01.50/0