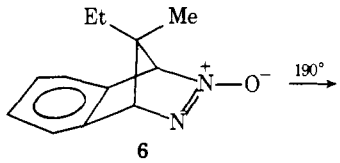
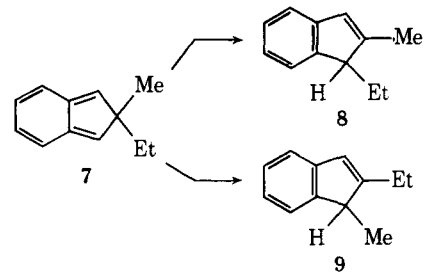


6



7 8 9

mixture of indenes **8** and **9**. A quantitative glpc analysis indicated the presence of less than 0.2% of crossover product **3**. Thus one can convincingly rule out the significant intervention of a free-radical-chain process in the rearrangement. Moreover, the ratio of **8**:**9** should be a reflection of the relative migratory aptitude of the ethyl vs. the methyl group in such suprafacial, concerted, thermal sigmatropic rearrangements. A measurement of relative migratory abilities in such processes has hitherto not been reported. Interestingly, the ethyl groups migrated at a rate seven times that of the methyl group. Insight into the factors involved in determining migratory aptitudes in such processes will no doubt be revealed through a systematic study, such as is presently underway.

Acknowledgment is made to the donors of the Petroleum Research Fund, administered by the American Chemical Society, and to the National Science Foundation for partial support of this research.

References and Notes

1. M. R. Willcott III, and I. M. Rathburn III, *J. Am. Chem. Soc.*, **96**, 938 (1974).
(2) M. A. M. Boersma, J. W. De Haan, H. Kloosterziel, and L. J. M. Van de Ven, *J. Chem. Soc. D*, 1168 (1970).
(3) K. Alder and M. Fremery, *Tetrahedron*, **14**, 190 (1961).
(4) J. P. Snyder, V. T. Bandurco, F. Darack, and H. Olsen, *J. Am. Chem. Soc.*, **96**, 5158 (1974).
(5) In contrast, the analogous azo compound would be expected to extrude N_2 at temperatures below 0°.[6]
(6) C. R. Flynn and J. Michl, *J. Am. Chem. Soc.*, **96**, 3280 (1974).
(7) J. Almy and D. J. Cram, *J. Am. Chem. Soc.*, **92**, 4316 (1970).
(8) L. L. Miller and R. F. Boyer, *J. Am. Chem. Soc.*, **93**, 650 (1971).
(9) With larger excesses of dienophile, essentially all of the sigmatropic process could be quenched.
(10) Nmr analysis indicates a threefold preference of the isomer shown.

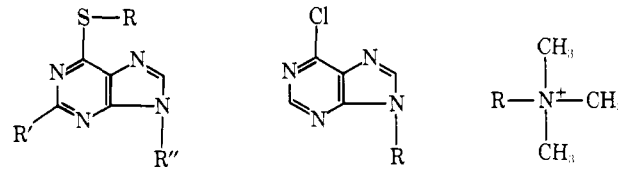
William R. Dolbier, Jr.,* Lenore McCullagh
Debra Rolison, Kent E. Anapolle
Department of Chemistry, University of Florida
Gainesville, Florida 32611
Received September 14, 1974

A Comment on Structure–Activity Correlations Obtained Using Pattern Recognition Methods

Sir:

Recently, pattern recognition techniques[1-3] have been applied to the screening of prospective anticancer drugs[4] which had previously been tested for activity in the Adenocarcinoma 755 (CA 755) screening system by the National Cancer Institute.[5] We wish to report that investigation of the structure–activity correlations resulting from this work has led us to conclude that there is a possibility of obtaining misleading relationships when using pattern recognition methods—particularly if there has been no interaction between a chemist and the data.

When we compared the 20 structure–activity correlations obtained for anticancer activity[6] with the structure–activity data of the compounds used to construct these correlations,[7] we observed a number of instances where the addition of one or more "+" correlation features had little effect on the activity and also several examples where the addition of a "+" feature or the replacement of a "+" feature by a different "+" feature caused significant reduction in activity. During this preliminary investigation, it became apparent that structure (1) (where R = H or organic side chain, R′ = H or NH_2, and R″ = H or a furanoside or pyranoside derivative) correlated strongly with anticancer activity. Analysis of the 233 nontoxic compounds in this test[8] revealed the following. (i) There were 80 compounds with structure **1**, 74 of

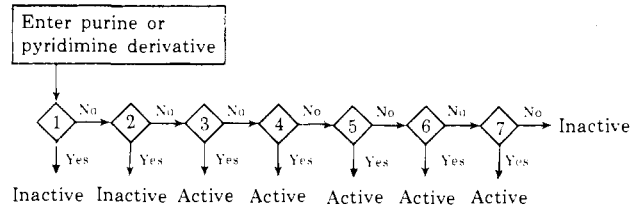


1 2 3

which were active. Five of the six inactive compounds of this structure contained either "fatty acid" or 2,3-*O*-isopropylideneribofuranosyl side chains. Of the 223 compounds, eight contained one or other of these two features and all were inactive. Only in structure **1** was sulfur active; all 13 compounds containing sulfur in other structures were inactive. (ii) Of the remaining 130 compounds which did not contain sulfur, 27 were active. Four of the 130 compounds contained fluorine (three of them active), 14 contained chlorine (seven active), 11 contained bromine (two active), and five contained iodine (all five active). Each of the seven active chlorine compounds and none of the inactive ones had structure 2 (R = ribofuranosyl or pyranosyl derivative). (iii) There were 10 active compounds that did not have structure **1** or contain a halogen. Five of these had structure **3** (R = substituted nucleoside). None of the other 218 compounds being investigated had this structure. (iv) There were no obvious structure–activity correlation features in the remaining five active compounds.

Using correlations based on the above analysis, we have constructed a flowchart (Scheme I) for the classification of the activity of prospective anticancer drugs. This scheme correctly classified 94 of the 101 "positive" (i.e., active) category and 119 of the 122 "negative" (i.e., inactive) cate-

Scheme I. Classification Flowchart



1. Presence of 2,3-*O*-isopropylideneribofuranosyl.
2. Presence of $-(CH_2)_{n-}$, where $n \geq 6$
3. Presence of structure 1, R = H or organic side chain
   R′ = H or NH
   R″ = H, furanoside, or pyranoside derivative
4. Presence of iodine
5. Presence of fluorine
6. Presence of structure 2, R = H, furanoside, or pyranoside derivative
7. Presence of structure 3, R = nucleoside derivative.

gory, giving 95.5% overall correct classification, which compared very well with the results obtained using the nearest neighbor method (93.5% correct[4]) using the leave-one-out procedure.[2,4] (Comparison of these performances is justified because leaving out any one of the 223 compounds would not have changed the conclusions drawn from the analysis or the resulting flowchart.)

The above analysis indicates that the previous classification scheme using pattern recognition methods[4] was able to produce satisfactory results using the 20 features[6] because of a favorable distribution of "inactive structural features" between the active and inactive classes, e.g., presence of sulfur (see point (i) above) and presence of phenyl. (Only ten compounds contained phenyl in structures other than structure **1**—all ten were inactive. There were 33 compounds that contained structure **1** and phenyl—31 of these were active.) This accounts for the misleading conclusions[4] that the presence of C–S bonds, S–H bonds, and phenyl or halogen substitution increase the tendency for a compound to be positive (i.e., active). Indeed, if the data set had contained more sulfur, phenyl, or bromine compounds with structures other than structure **1** (i.e., inactive sulfur, phenyl, and halogen compounds), the pattern recognition methods would not have made these correlations. This illustrates the advantage, with "difficult" data, of having an interactive system which enables the chemist to optimize the relevance of questions to be asked.

On the assumption that the incorporation of various biologically active features into one compound may enhance activity, the above analysis suggests different combinations of structures that may show enhanced activity. One such combination is structure **1** with the addition of either fluorine, iodine, or trimethylammonium ion. Several compounds combining structure **1** and fluorine have been tested[7]—none showed enhanced activity. Two recent reviews of biologically active nucleoside derivatives[9] make no mention of compounds incorporating iodine or trimethylammonium ion in structure **1**.

References and Notes

(1) T. L. Isenhour and P. Jurs, *Anal. Chem.*, **43**, 20A (1971), and references therein.
(2) B. R. Kowalski and C. F. Bender, *J. Am. Chem. Soc.*, **94**, 5632 (1972).
(3) J. B. Justice and T. L. Isenhour, *Anal. Chem.*, **46**, 223 (1974), and references therein.
(4) B. R. Kowalski and C. F. Bender, *J. Am. Chem. Soc.*, **96**, 918 (1974).
(5) A. Goldin, H. B. Wood, and R. R. Engle, *Cancer Chemother. Rep.*, **1**, 1 (1968).
(6) Data from Table I of ref 4.
(7) Data from Table 2 of ref 5.
(8) In this work we extracted the 259 compounds in ref 7 that had been tested for activity with the CA755 test and deleted the compounds that were toxic to test animals or had tumor weight inhibition (TWI) between 50 and 70%, as outlined.[4] This still left 101 compounds with TWI greater than 70% (i.e., active) and 122 compounds with TWI less than 50% (i.e., inactive), a total of 223, which we used. The previous investigators deleted a further 23 compounds and used only 200.[4]
(9) A. Bloch in "Medicinal Chemistry," Vol. 4, E. J. Ariens, Ed., Academic Press, New York, N.Y., 1973, Chapter 8; J. A. Montgomery, P. A. Johnston, and Y. F. Shealy, *ibid.*, Vol. 1, A. Burger, Ed., Wiley-Interscience, New York, N.Y., 1970, Chapter 28.

R. J. Mathews

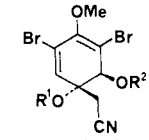
Australian Defence Scientific Service
Department of Defence, Materials Research Laboratories
Ascot Vale, Victoria, Australia

Received October 4, 1974

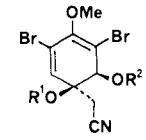
Synthesis of Aeroplysinin-1 and Related Compounds

Sir:

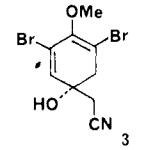
Marine sponges of the genus *Verongia* have yielded a number of interesting brominated metabolites,[1-3] some of which exhibit potentially useful biological activity. Aeroplysinin-1 (**1**), isolated as the (+)-enantiomer from *V. aerophoba*[2a] and as the (−)-enantiomer from *Ianthella ardis*,[4] displayed antibiotic activity against *S. Aureus* and antileukemia activity in the L-1210 screen. Aeroplysinin-1, the only naturally occurring 1-alkyl-1,2-dihydroxy-2*H*-arene, is of considerable biosynthetic interest since it is believed to be derived from an arene oxide *via* enzymatic hydration.[5] We wish to report the synthesis of aeroplysinin-1 (**1**) and the closely related compounds, isoaeroplysinin-1 (**2**) and 2-desoxyaeroplysinin-1 (**3**).



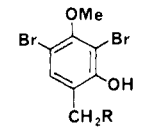
1 $R^1 = R^2 = H$
11 $R^1 = R^2 = Ac$



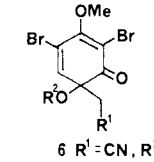
2 $R^1 = R^2 = H$
9 $R^1 = Ac$, $R^2 = H$
10 $R^1 = R^2 = Ac$



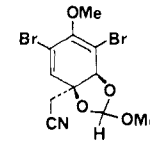
3



5 R = CN
7 R = H



6 $R^1 = CN$, $R^2 = Ac$
8 $R^1 = H$, $R^2 = Ac$
14 $R^1 = CN$, $R^2 = H$



12

3,5-Dibromo-2-hydroxy-4-methoxyphenylacetonitrile (**5**), mp 158°, obtained in 90% yield by the treatment of 2-hydroxy-4-methoxyphenylacetonitrile (**4**)[6] with 2 equiv of pyridinium hydrobromide perbromide in pyridine,[7] was oxidized with excess lead tetraacetate in acetic acid at 25° for 18 hr[8] to obtain the dienone **6**[9] in 35% yield. The presence of a nitrile function considerably lowered the yield of this reaction; using identical reaction conditions, the phenol **7** was oxidized to the corresponding dienone **8** in 75% yield. However, in a single step we had generated the correctly substituted diene system and introduced an oxygen atom at the required location.

Reduction of the dienone **6** with sodium borohydride in absolute ethanol at 0° gave three products: isoaeroplysinin-1 (**2**) (40%),[10] the corresponding monoacetate **9** (18%), and 2-desoxyaeroplysinin-1 (**3**) (22%).[11] On treatment with acetic anhydride in pyridine, the monoacetate **9** gave the diacetate **10**, which was not identical to the diacetate **11** obtained from authentic aeroplysinin-1 (**1**).[12] In order to show that the two hydroxyl functions in isoaeroplysinin-1 (**2**) were cis to one another, we converted **2** into a cyclic orthoformate **12**, using methyl orthoformate in refluxing benzene containing a catalytic amount of chloroacetic acid. Treatment of isoaeroplysinin-1 (**2**) with *p*-toluenesulfonic acid in refluxing benzene caused rapid dehydration to yield the phenol **5**. The structure of 2-desoxyaeroplysinin (**3**) was assigned on the basis of spectral data and confirmed by its dehydration under similar conditions to 3,5-dibromo-4-methoxyphenylacetonitrile (**13**).

The failure to obtain a trans diol was attributed to the influence of the acetoxy function. The dienone **6** was therefore converted into the corresponding keto alcohol **14** (45% conversion) by transesterification in methanol containing