

tion was refluxed. The reaction was irradiated with a 75-W flood-lamp for 2.5 hr. The solution was then cooled to room temperature and filtered. The  $\text{CCl}_4$  was removed under reduced pressure and the solid dissolved in diethyl ether. The compound was crystallized from  $\text{Et}_2\text{O}$ -hexane (yield 80%); mp 131–135°.

**Bis(dimethylsulphido)decahydrododecaborane** [ $\text{B}_{12}\text{H}_{10}[\text{S}(\text{CH}_3)_2]_2$ ]. The procedure for the synthesis of  $\text{B}_{12}\text{H}_{10}[\text{S}(\text{CH}_3)_2]_2$  was a modification of the procedure given to us by Kaczmarczyk.<sup>8</sup> In a typical reaction 3 g (15.9 mmol) of  $\text{Na}_2\text{B}_{12}\text{H}_{12}$  was dissolved in 20 ml of warm DMSO in a 100-ml round-bottom flask equipped with a reflux condenser and  $\text{N}_2$  inlet. Acetic anhydride (20 ml) was added slowly. When all of the acetic anhydride had been added the temperature of the reaction was raised to 90°. At this point an exothermic reaction took place and the temperature of the solution rose to 140–150°. After the exothermic reaction had stopped the flask was cooled to room temperature. The reaction mixture was then poured into 300 ml of  $\text{H}_2\text{O}$  containing 4 g of  $\text{NaCl}$ . If the product precipitated out as a solid, it was removed by filtration and crystallized from  $\text{EtOH-H}_2\text{O}$ . Many times the product precipitated as an oil. In this case the aqueous mixture was extracted with several portions of  $\text{CH}_2\text{Cl}_2$ . The  $\text{CH}_2\text{Cl}_2$  fractions were collected, the solvent was removed, and the product crystallized from  $\text{EtOH-H}_2\text{O}$  (mp 220–235°) in 60% yield.

**Tetramethylammonium Dimethylsulphidomethylsulphidodecahydrododecaborate** [ $(\text{CH}_3)_4\text{N}[(\text{CH}_3)_2\text{SB}_{12}\text{H}_{10}\text{SCH}_3]$ ].  $\text{B}_{12}\text{H}_{10}[\text{S}(\text{CH}_3)_2]_2$  (1 g, 3.78 mmol) was dissolved in 50 ml of dry DMF in a 100-ml three-neck flask equipped with a condenser and  $\text{N}_2$  inlet. Potassium phthalimide (1 g) was added and the temperature was raised to 100° and maintained there for 3 hr. The solution was then raised to 140° for 1 hr. The flask was cooled and the solution was filtered. After the solvent had been removed, the solid residue was extracted with three 500-ml portions of hot water. Addition of  $(\text{CH}_3)_4\text{NCl}$  to the aqueous solution resulted in the precipitation of  $(\text{CH}_3)_4\text{N}[(\text{CH}_3)_2\text{SB}_{12}\text{H}_{10}\text{SCH}_3]$ . The product was crystallized from hot water (90% yield).

**Conjugation of  $\text{B}_{10}\text{H}_8\text{S}(\text{CH}_3)_2\text{S}(\text{CH}_3)\text{CH}_2\text{C}_6\text{H}_4\text{NH}_2$  and Bovine Serum Albumin (BSA).** To 33 mg of  $\text{B}_{10}\text{H}_8\text{S}(\text{CH}_3)_2\text{S}(\text{CH}_3)\text{CH}_2\text{C}_6\text{H}_4\text{NH}_2$  (2e) was added 0.3 ml of a 1 M HCl solution in an ice bath.  $\text{NaNO}_2$  solution (1 mg/ml, 7 ml) was added and the solution was stirred for 1 hr. The mixture was filtered cold (4°) and diluted to 25 ml with 0.1 M HCl. This solution was added slowly to 65 mg of BSA in a phosphate buffer (pH 8). The solution tended to become acidic and was maintained at pH 8 with 0.1 M NaOH. After stirring for 24 hr at 4°, the mixture was dialyzed exhaustively against isotonic saline and the undialyzed product was analyzed for both boron and protein.

**Conjugation of  $\text{B}_{10}\text{H}_8\text{S}(\text{CH}_3)_2\text{S}(\text{CH}_3)\text{CH}_2\text{C}_6\text{H}_4\text{NCS}$  and BSA.**  $\text{B}_{10}\text{H}_8\text{S}(\text{CH}_3)_2\text{S}(\text{CH}_3)\text{CH}_2\text{C}_6\text{H}_4\text{NCS}$  (15 mg, 2b) was dissolved in 10 ml of dioxane. The solution was added slowly to 150

mg of BSA in 30 ml of a  $\text{CO}_3^{2-}/\text{HCO}_3^-$  buffer (pH 9). After the addition, the pH of the solution had dropped to 7 and was adjusted to 9 with  $\text{Na}_2\text{CO}_3$ . The solution was stirred at 0° overnight. The mixture was then filtered and dialyzed in a comparable manner to the above procedure.

**Coupling of RCOOH (2a and 2d) with BSA.** The desired boron hydride (27 mg) was dissolved in 20 ml of DMF along with 30 mg of BSA and 10 mg of dicyclohexylcarbodiimide (DDC) and stirred in an ice bath for 12 hr. The solution was then diluted with 30 ml of  $\text{H}_2\text{O}$  and dialyzed in a similar manner to the above procedure. The solution was then centrifuged to remove any suspended solids and analyzed for boron and protein content.

**Acknowledgment.** This work was supported in part by the National Institutes of Health through Grant P101 CA 12246 and the Ruth M. Israel Fund for Cancer Research. We are grateful to William N. Smolnycki for important technical assistance. We thank Mrs. Janette Messer of Massachusetts General Hospital for the protein and boron analyses and Dr. Alex Kaczmarczyk of Tufts University for supplying us with the method for preparing  $\text{B}_{12}\text{H}_{10}[\text{S}(\text{CH}_3)_2]_2$ .

## References

- (1) A. H. Soloway in "Progress in Boron Chemistry," Vol. 1, H. Steinberg and A. L. McCloskey, Ed., Pergamon Press, New York, N. Y., 1964, pp 203–234.
- (2) A. K. Asbury, R. G. Ojemann, S. L. Nielsen, and W. H. Sweet, *J. Neuropathol. Exp. Neurol.*, **31**, 279 (1972).
- (3) A. H. Soloway in "Radionuclides Application for Neurology and Neurosurgery," Y. Wang, Ed., Charles C Thomas, Springfield, Ill., 1970, pp 301–312.
- (4) A. H. Soloway, G. L. Brownell, R. G. Ojemann, and W. H. Sweet in "Preparation and Biomedical Applications of Labeled Molecules," Euraton, Brussels, 1964, pp 383–403.
- (5) M. F. Hawthorne and R. J. Wieksema, *J. Med. Chem.*, **15**, 449 (1972).
- (6) A. G. Mallinger, E. L. Jozwiak, Jr., and J. C. Carter, *Cancer Res.*, **32**, 1947 (1972).
- (7) W. H. Knott, W. R. Herther, and E. L. Muetterties, *Inorg. Chem.*, **4**, 280 (1965).
- (8) J. Wright and A. Kaczmarczyk, *Inorg. Chem.*, **6**, 1453 (1973).
- (9) A. H. Soloway and J. R. Messer, *Anal. Chem.*, **36**, 433 (1964).
- (10) E. Layne, *Methods Enzymol.*, **3**, 448 (1957).
- (11) W. Davis, J. J. Roberts, and W. C. H. Ross, *J. Chem. Soc.*, 890 (1955).

## Application of the Sequential Simplex Method in Designing Drug Analogs

Ferenc Darvas

EGYT Pharmacochemical Works, 1475, Budapest 10, Hungary. Received June 13, 1973

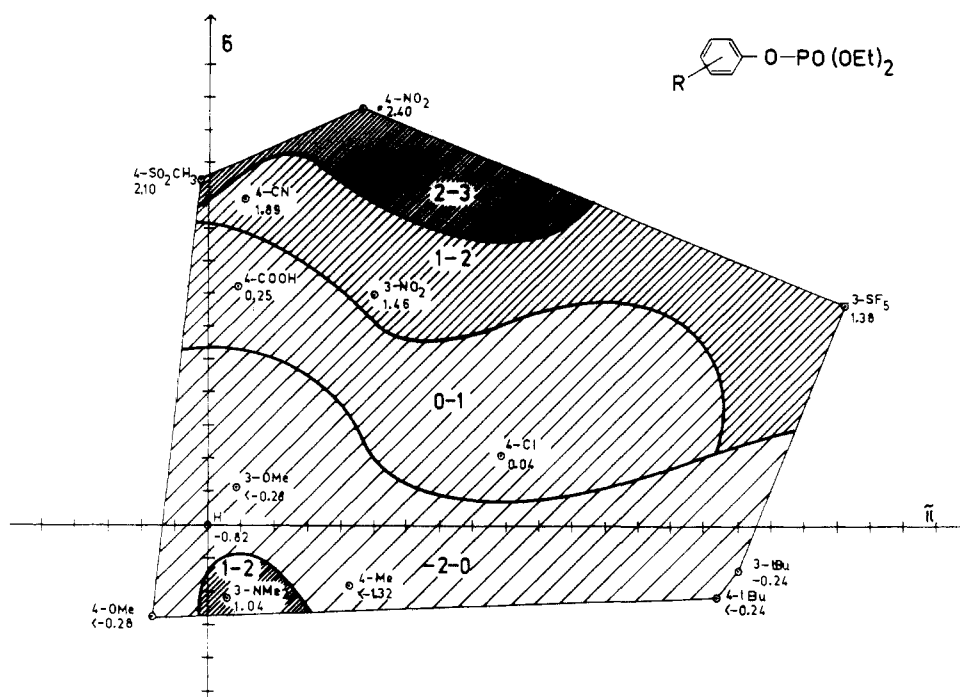
An EVOP sequential simplex optimization method is proposed for the design of drug analogs. The method is based on certain regularities in the structure of compounds having the same biological activity, and it does not require numerical calculation. The Hansch parameters and the biological activities were taken as coordinates in our work. The method has been illustrated by retrospective examples. According to the statistical analysis regarding a less favorable example, the method is superior to an unsystematic choice at a significance level of 0.01%.

In the past decade the Hansch method has become one of the most widespread methods in quantitative structure-activity relationship (QSAR) studies and in drug design. The method has been applied to more than 1500 members of about 150 different sets of compounds.<sup>1</sup>

First the Hammett constants and the partition coefficients were used to correlate the biological activity with the structure.<sup>2</sup> During the past 10 years the application of almost all the linear free energy related parameters has been used in correlations with biological activities of com-

pounds;<sup>1</sup> however, there have been few attempts to use methods other than regression analysis to utilize the correlations between these parameters and the biological activities. As far as we know, there have been three attempts to apply methods other than regression analysis in this field.<sup>3–5</sup>

It is well known that structure-activity correlations for series of compounds can be determined by Hammett-type constants and the Hansch  $\pi$  constant by some simple relationships.<sup>1</sup> So one may expect that biological effects are

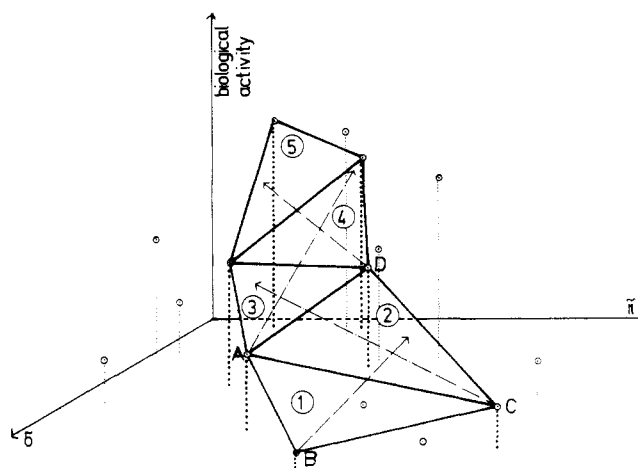


**Figure 1.** Positions of the groups of the biological activity of diethyl phenyl phosphates inactivating acetylcholinesterase in the  $\pi$ - $\sigma$  biological activity coordinate system. Biological activity = LD<sub>50</sub> toward houseflies in log 1/C units.

regularly distributed in a three-dimensional coordinate system determined by the above two types of variables and the biological activity.<sup>6</sup> The parameters calculated by regression analysis reflect this regularity only indirectly; due to their statistical nature they give no information on the actual location of biological activity. It was expected that the groups of different compounds having similar biological activities were located near one another in the coordinate system formed from the above-mentioned linear free energy related parameters. In our work, data of some papers<sup>7</sup> published earlier have been examined and definite proof of this assumption has been found.

As an example for these, Figure 1 shows the locations of 14 compounds in a coordinate system where the abscissa is the  $\sigma$  value measured for acetylcholinesterase inhibitor phenol derivatives<sup>8</sup> and the ordinate<sup>8</sup> is the  $\pi$  constant. In the diagrams throughout this paper, one division on the axes corresponds to 0.1  $\sigma$  and 0.1  $\pi$  units, respectively. The numbers beside the points of the compounds give the biological effects of the derivatives. The coefficient of the linear regression used for the approximation of the biological activity is good (0.987); the value of the twofold deviation is, however, high. The biological effects were distributed into four groups as shown in Figure 1 where the various groups are roughly separated by lines. The boundaries have been drawn by hand to bisect approximately the distances between the points belonging to different groups. The regular distribution of the groups with respect to the vertical axis is clearly perceptible, apart from the 3-N(CH<sub>3</sub>)<sub>2</sub> group in the  $+\pi - \sigma$  domain, whose activity value 1.04 represents a maximum not fitting into its environment at all.

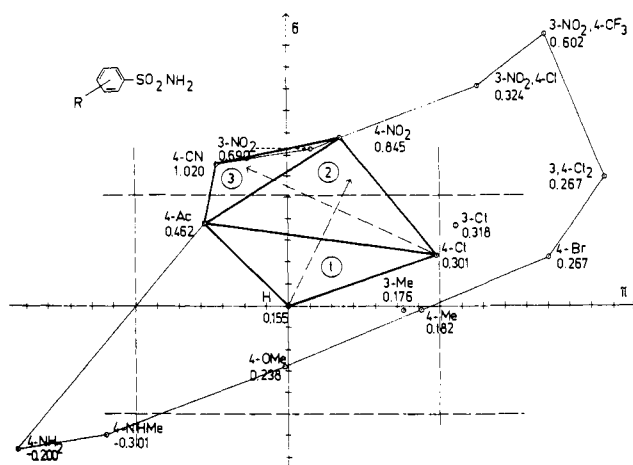
The more or less regular relative positions of the groups permit the application of function maximalization procedures. The method to be presented in this paper is an adaptation of the EVOP-type<sup>9</sup> "simplex self-directing" method to noncontinuous fields, where the regularities discussed above can be utilized, and the statistical disorder of the compounds can be eliminated without statistical calculations.



**Figure 2.** Three-dimensional coordinate system with a hypothetical set of compounds. A search is made for the derivative having maximal biological activity starting from compounds A, B, and C.

Since in the three-dimensional case this method requires neither a computer, nor numerical calculations, it provides a useful tool for medical chemists who prefer graphic representations and are reluctant to use mathematical statistics in designing biologically active analogs of a given compound.

The main principles of our procedure are presented in Figure 2, where a hypothetical set of compounds is plotted in a three-dimensional coordinate system. The two horizontal coordinates show the  $\pi$  and  $\sigma$  values of the substituents, respectively; the vertical coordinate reflects the biological activity of the compound carrying the substituent. To determine the compound exhibiting the highest biological activity, there is an easily intelligible, almost mechanical procedure. Three points are selected on the surface (in Figure 2, points A, B, and C), and the lowest lying point is determined by comparing their "altitudes." Spendley and his coworkers<sup>9</sup> proved by computerized



**Figure 3.** Simplex-type optimization of natriuretic sulfonamides. Biological activity = MIC with respect to carbonic anhydrase.

model experiments that the maximum of the surface can be found very quickly if the original three points are regarded as the apices of a triangle, and this triangle is projected through the side opposite to the lowest point. It is justified to assume that the new point (point D in Figure 2) will be located higher than the lowest point of the first triangle. The points are compared and, if the assumption proved to be correct, the lowest point is selected in the "mirrored" triangle, and the procedure is repeated until the maximum is found.

On the basis of the theoretical example, the method for designing optimal drug analogs consists of the following steps.

(1) A "parent compound" is chosen whose activity is to be improved, and the sites of change (substitution) are selected in the molecule.

(2) The parent compound is located in the appropriate  $\pi$ - $\sigma$  coordinate system or other coordinate systems used in the Hansch-type procedures.

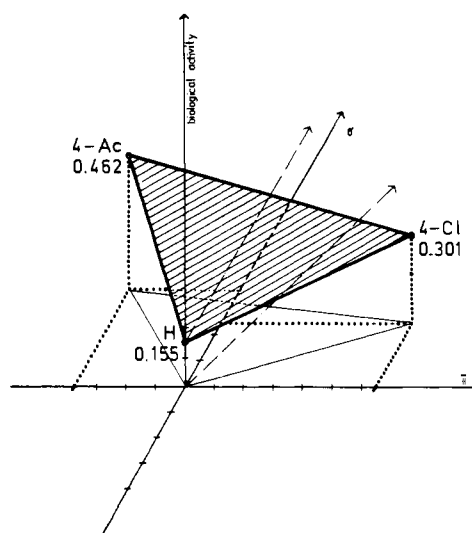
(3) Points corresponding to derivatives which presumably possess the desired effect are determined in the coordinate system. In compiling the list of substances, the special aspects of the synthetic work, biopharmacology, etc., can be taken into account.

(4) Two compounds are selected from the derivatives near the parent compound, prepared and tested. These are chosen following Topliss' suggestions,<sup>3</sup> that is, by looking for points where one of the parameters is systematically increased or decreased with respect to the original substance, the other remaining constant as far as possible.

(5) On the basis of the activities of these three compounds a decision is made about the next derivative to be prepared. The basis of the decision is always the simplex, *i.e.*, the triangle formed by the three substances. The directive derivative is determined practically by a plane-geometrical procedure; the point of the less effective derivative is connected with the midpoint of the opposite side of the triangle, and the new, supposedly more effective compound is searched for in this direction.

(6) The mechanical repetition of this procedure leads to the preparation of substances having outstanding activities as compared to their neighbors. This maximum is surrounded stepwise by simplexes, since the point with the highest effect in the triangle is always involved. If we wish to continue the optimization procedure, it is most preferable to include new substances in the optimization map around this maximum and to resume the optimization on a smaller scale.

It is to be emphasized that in certain cases other proce-



**Figure 4.** Spatial orientation of the starting simplex of natriuretic sulfonamides.

dures can be used equally well in searching the optimum; for instance, one may connect the point belonging to the most suitable derivative with the opposite midpoint in the triangle, and the new derivative is searched along the extension of this line beyond the best point. The searching strategy can be supplemented with heuristic features, such as with the conclusions of the conventional Hansch-type regression investigations concerning the presumably effective substances.

The procedure is illustrated in a retrospective manner on the example of natriuretic sulfonamides (Figure 3), studied by Kakeya, *et al.*<sup>10</sup> Let us assume that we wish to prepare compounds of possibly high natriuretic effect by aromatic substitution from benzenesulfonamide. As a first step the 15-20 substituents are chosen which may enhance the activity of the molecule or, in the absence of such type of information, which are easy to prepare. The aim of our method is to find the most effective member of the 16 compounds by preparing and testing the smallest number.

As the first step of the "optimization," two substituted compounds are selected, prepared, and tested for natriuretic effect. It is evident that one should prepare the 4-acetyl and 4-chloro derivatives, since the  $\sigma$  values of both compounds are higher than that of the parent molecule, while the  $\pi$  value is smaller in the first case and higher in the second than that of the unsubstituted compound.

If the biological activity is regarded as the third dimension, a triangle formed by the three compounds has the spatial position shown in Figure 4. If the sides of the simplex triangle are chosen to be long enough, it will be highly probable that the differences between the biological effects of the compounds are certainly not due to experimental errors in the determination of the biological activities or the  $\pi$  or  $\sigma$  values. The triangle covers about 50% of the total range of  $\pi$  and about 15% of that of  $\sigma$  (Figure 3). The activities of the three compounds are compared and the fourth, presumably more active compound, is looked for according to the above principle in the direction opposite to the less effective derivative in the triangle. It can be seen in Figure 4, and can be proved by the law of similar triangles, that the optimum direction can also be found by a plane-geometrical procedure; the point corresponding to the lowest value is connected to the midpoint of the opposite side of the triangle. In this direction the 4-nitro derivative can be found; this is now to be prepared and tested for natriuretic activity. Since

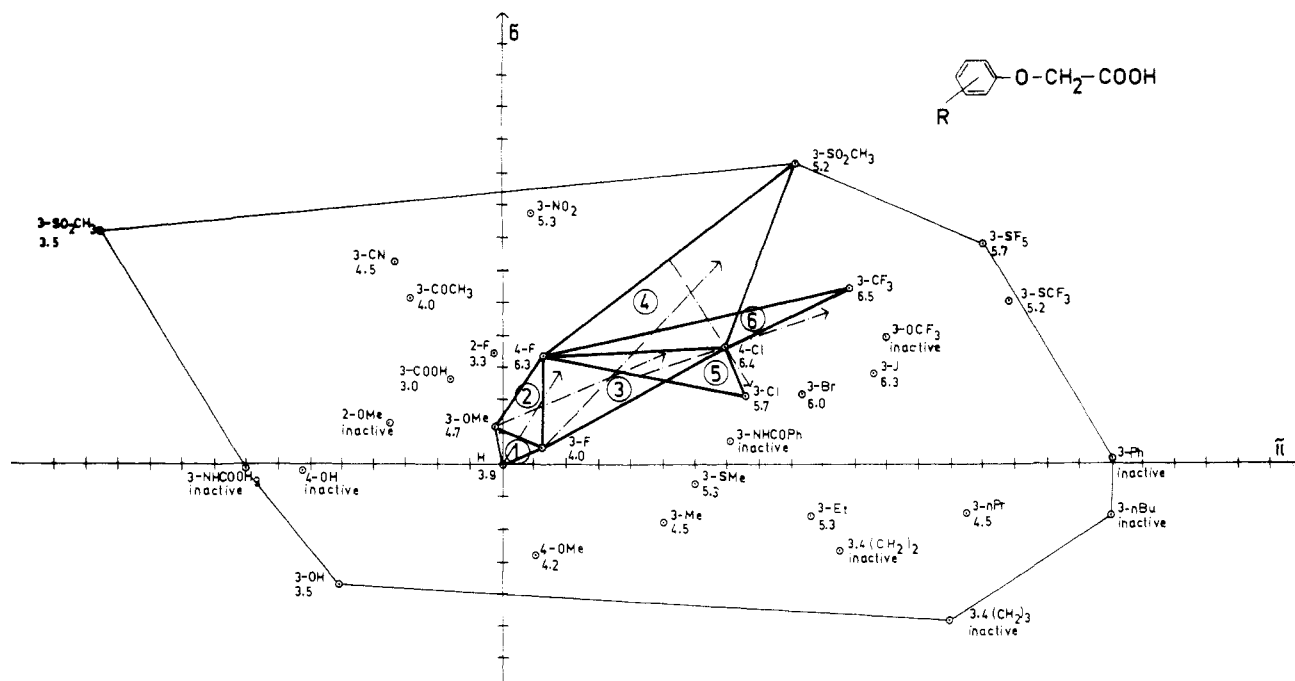


Figure 5. Simplex-type optimization of phenoxyacetic acids with auxin activity. Biological activity = Avena cell elongation.

its activity is substantially higher than that of the less active member of the simplex, *i.e.*, the unsubstituted derivative, a new simplex is constructed where this latter compound is replaced by the 4-nitrobenzenesulfonamide.

The least active compound of the new simplex is the 4-chloro derivative. In the direction opposite to this compound the most effective 4-cyanobenzenesulfonamide can be found. At this step the most effective derivative of the set, given by Kakeya, has been found.

The 3-nitro derivative is located between the outstandingly active 4-cyano and the very active 4-nitro derivatives. Since, presumably, this compound is also very effective, it is advisable to prepare and test its effect; the result (0.696) confirms the prominent character of the 4-cyano derivative.

Obviously, the 4-cyano derivative is surrounded by a number of other derivatives not given in the diagram, since they are not listed in the paper of Kakeya, *et al.* If further optimization is required, it is most preferable to include these derivatives around the compound in question into the optimization map and to accomplish the optimization procedure around the maximum on a smaller scale.

Similar results can be found if the optimization procedure is started from the triangles formed by the substituents hydrogen-4-acetyl-3-chloro, hydrogen-4-aminomethyl-4-acetyl, hydrogen-4-amino-4-acetyl, or hydrogen-4-methyl-4-bromo. Unfortunately, no substances are mentioned in the paper in the areas  $-0.7\pi$ ,  $-0.5\sigma$  and  $+0.7\pi$ ,  $0.5\sigma$ , respectively; thus, the result of the optimization starting from the triangles hydrogen-4-acetyl-4-methoxy and hydrogen-4-methoxy-4-methyl cannot be given. Of course, when applying this procedure in the practice, these areas are populated by further "compound points," too.

The next retrospective example pertains to the phenoxyacetic acids (Figure 5) investigated by Hansch, *et al.*<sup>2</sup> In the first step the 4-methoxy derivative, which has a higher value of  $\sigma$  and about the same  $\pi$  as the unsubstituted molecule, and the 3-fluoro derivative, which has higher values of  $\sigma$  and  $\pi$  than the unsubstituted phenoxyacetic acid, are prepared. The bisector, which indicates the slope

of the simplex, points unambiguously to the 4-fluoro derivative as the next compound to be prepared. In the new simplex, however, there is no unique decision about the choice of the next compound, since the 4-methoxy and 3-fluoro derivatives have about the same effect, but the two alternative directions are perpendicular to each other. If in the next step the 2-fluoro derivative is chosen, only the value of  $\sigma$  is increased; if the 4-chloro derivative is prepared, the values of both coordinates increase simultaneously. The simultaneous increase of the coordinates has already proved a successful method in the previous step, but now, in order to show the applicability of formal rules, let us "prepare" the 2-F derivative. This compound is inferior even to the less effective member of the simplex, so this does not lead to a new simplex, and therefore the preparation of the 4-Cl derivative should be attempted. From the fourth simplex obtained in the next step, however, simplex 3 is regained when applying the rules discussed. In such cases logical strategy is to look for the steepest ascent in the direction of the best point of the simplex. To obtain a new derivative, this point is connected with the midpoint and the line is extended beyond the simplex point. The same strategy should be applied to simplex 5; in this case the 3-CF<sub>3</sub> derivative of maximum activity is obtained.

These examples indicate a common feature of the simplex procedure; since the most effective member of the simplex is always retained when choosing a new simplex, derivatives with outstanding effects in their environment always remain in the simplex, and the corresponding point is, so to say, "walked around" by the simplex.

In the case of the antibacterial benzyl alcohols<sup>11</sup> the compounds are packed densely around the center of the optimization area (Figure 6). On the basis of the starting simplex (5-Br,2-OH-6-Br,3-OH-5-I,2-OH), the 5-C<sub>3</sub>H<sub>7</sub>,2-OH or the 3,5-Cl<sub>2</sub>,2-OH derivatives can be prepared. In both cases a new simplex is obtained which includes the prepared compound. In the model investigated, however, the simplex (6-Br,3-OH-5-I,2-OH-5-Pr,2-OH) does not lead to a new derivative because in the optimal direction no further compounds can be found. Now, going further from the simplex obtained after the preparation of the 3,5-

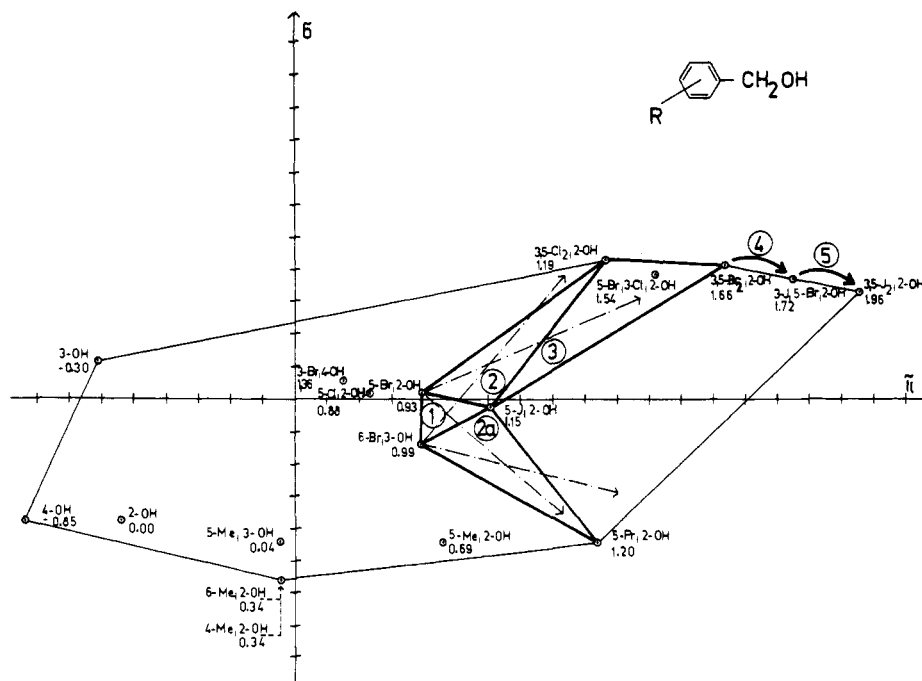


Figure 6. Simplex-type optimization of bactericidal benzyl alcohols. Biological activity = log relative bactericidal activity.

Br<sub>2</sub>,2-OH derivative in the direction of the new compound further substances with increasing activities can be prepared, namely the 3-I,5-Br,2-OH and the 3,5-I<sub>2</sub>,2-OH derivatives. In the case of the diethyl phenyl phosphates investigated by Hansch and Fujita<sup>8</sup> the starting simplex and then the 3-NO<sub>2</sub> derivative are chosen on the basis of the above considerations (Figure 7). The next choice is not unanimous, because no directive can be found opposite to the H-substituted product in the model. The maximum of the model can be found by going further in the direction of the most effective derivative in the simplex. If the 3-SF<sub>5</sub> derivative is prepared, the simplex (H, 4-Cl, 3-SF<sub>5</sub>) obtained leads to the 4-NO<sub>2</sub> derivative in the next step, which has maximum effect.

These examples reveal some of the characteristics of the optimization method.

(a) The performance of the method is not affected by compounds having extremely low biological effects (local minima) not fitting into the Hansch relationships. For instance, in the case of phenoxyacetic acids of auxin activity the 3-NHCOPh and the 3-COF<sub>3</sub> derivatives have extremely low activities. They are much less active than their environment. According to the characteristics of the optimization method, first these low values are replaced in the simplex; hence, they do not influence the final result, though they may lengthen the way toward the maximum.

(b) The presence of compounds with activities outstanding from their environment is a much greater problem. The point of the simplex with the highest values in a broad environment is never replaced. If there is no compound having higher activity than this maximum in the direction determined by the maximum and two environmental points, the optimization will terminate in this maximum, irrespective of the fact that this derivative is not the true maximum of the optimization area.

If this "pseudomaximum" is in the group of the most active compounds, and the compounds belonging to this group form a coherent area, this pseudomaximum usually causes no problem, because the real maximum is most probably found by going around. Such a pseudomaximum is, for instance, the 4-Cl derivative in the phenoxyacetic acid series. This derivative is surrounded by less active

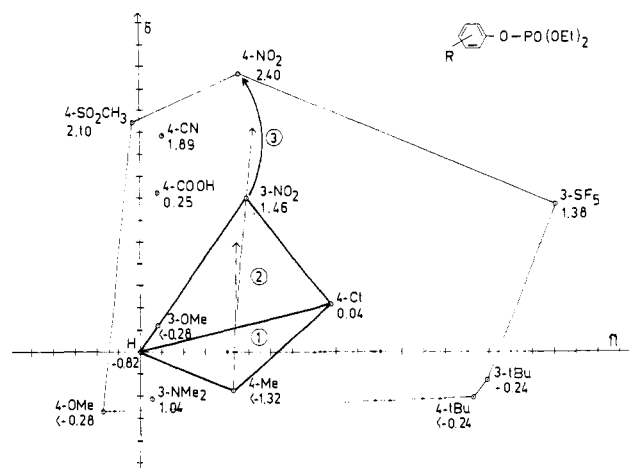


Figure 7. Simplex-type optimization of diethyl phenyl phosphates inactivating acetylcholinesterase. Biological activity = LD<sub>50</sub> toward houseflies in log 1/C units.

derivatives, but the simplexes containing the pseudomaximum and the neighboring derivatives always lead to the true maximum.

It may happen that in an environment of low activity, one derivative has an activity far higher than that of the environment. Such a compound is the diethyl 3-(dimethylamino)phenyl phosphate of cholinesterase inhibitor activity (Figure 7). In this case the whole system of structure-activity correlations fails, so it cannot be expected that the optimization method based on these correlations will succeed. When it is realized that the outstanding activity of the compound cannot be attributed to the coordinate values, incidental metabolic effects, or steric factors characteristic only of the given substituent, the optimization procedure must be repeated, neglecting this compound, or the whole system of correlations must be changed.

The most serious problem is caused by the local maxima included in groups with medium activities, for instance, by the 3-NO<sub>2</sub>,4-CF<sub>3</sub> derivative in the case of the natriuretic sulfonamides.

If the distance between the pseudomaximum and the

true maximum is not too large, the danger of ending up in the pseudomaximum can be avoided by choosing a basic net with evenly distributed rare points. If the two maxima are far apart, the surface formed by the biological activity values can obviously be approximated very poorly with the parabolic surface used in the Hansch method, but this is due to the inaccuracy of the basic relationships and not to the optimization procedure.

It should also be taken into account that if a "big jump" is made to the preparation of the next compound, maxima may be overlooked owing to the nonlinear behavior of the parameters. Because of the characteristics of the optimization method it is very probable that sooner or later the simplexes turn back into the direction of the maximum, but this roundabout approach decreases the efficiency of the procedure. The even distribution of the points, however, ensures that a single step does not lead very far from the original simplex.

To give a somewhat more quantitative picture about the performance of the simplex-type optimization, the natriuretic sulfonamide derivatives are discussed. In this model, which is less regular and has a "pseudomaximum," all the possible simplexes have been used as starting simplex, and it has been studied whether the derivative of maximum effect can be found in every case.

In order to construct the starting simplexes the optimization area was divided into different regions, shown by dotted lines in Figure 3. The starting simplexes were constructed in such a way that their points fell into three different regions, 0.1-1.0  $\pi - \sigma$  units apart and formed a triangle, and no further point fell inside the individual triangles. Under these conditions, 54 starting simplexes could be obtained in the optimization area. Where the two rules given above for finding the direction of more active derivatives were inapplicable, the new derivative was sought in the direction of medium activity or in the opposite direction. Derivatives with outstanding activities in comparison with that of the compounds already "prepared" during the optimization were regarded as maxima; whenever a maximum was found, the compounds located in its environment were prepared and the optimization process was terminated.

The maximum was hit in 49 cases, whereas in five cases (9.35%) the result was the "pseudomaximum" 3-NO<sub>2</sub>,4-CF<sub>3</sub> compound, which also belongs to the group of maximum activity. The maximum was produced as the fifth compound, on the average. This probability of hitting the maximum of 16 compounds by fivefold random choice is 31.2%, according to the formula of random sampling without replacement. The maximum can be hit with the certainty of 90.5% of the simplex method only if  $0.905 \times 16 = 14.4$  compounds of the series are prepared unsystematically. It can be found by subjecting the two numbers (14 and 5) to the  $\chi^2$  test that the sequential simplex method is significantly superior to the random choice even at the level of 0.01%.

The disadvantage of the widespread methods of drug design is that they need numerical data representing the biological activity, whereas the results of primary screening pharmacological tests are mostly given only in a semi-quantitative manner by the number of plus signs due to

the high experimental uncertainties. Several other features of the compounds are not expressible numerically, such as stability, preparation, and production facilities, etc. These are the aspects having major importance in practical drug research. In the methods based on regression analysis these aspects cannot be taken into account because of the lack of numerical data or the difficulties in converting the data into a common scale. In the procedure based on the simplex method, these aspects can be considered at the beginning of the optimization, *i.e.*, in the construction of the optimization net, as well as when the activity scale of the compound is fixed, because it is not necessary to construct this scale on the basis of a single, numerically expressed feature.

A promising but less spectacular application of the simplex method is the multidimensional optimization, where the given compound is substituted at several sites. The obvious advantage of this method is the elimination of numerous factorial-type experiments to investigate the relative importance of the individual substitution sites. In the multidimensional problems the error arising from the false evaluation of the magnitudes of changes is considerably lower, because the new data are compared to as many known data as the dimension of the problem is. According to the investigations of Spendley cited above, the efficiency of optimization increases with dimensionality: in four dimensions the efficiency is about 7.8 times and in ten dimensions about 10 times higher than that of the optimization method where the values assigned to two points are compared. As the dimensionality can be increased during the optimization, parameters held constant in the previous steps can also be taken into account. Likewise, the target function can also be augmented during the process.

**Acknowledgment.** The author is grateful to Dr. Laszlo Pallos for his interest and to Dr. Geza Nagy and Dr. Gyorgy Jalsovszky for a number of fruitful discussions and their help in preparing this manuscript.

## References

- (1) (a) C. Hansch "Drug Design," Vol. I, E. J. Ariens, Ed., Academic Press, New York and London, 1971, pp 271-405; (b) A. Verloop in ref 1a, Vol. III, 1972, Chapter 2, pp 133-138.
- (2) C. Hansch, R. M. Muir, T. Fujita, P. P. Maloney, F. Geiger, and M. Streich, *J. Amer. Chem. Soc.*, **85**, 2817 (1963).
- (3) J. G. Topliss, *J. Med. Chem.*, **15**, 1006 (1972).
- (4) M. L. Weiner and P. H. Weiner, *J. Med. Chem.*, **16**, 655 (1973).
- (5) C. Hansch, S. H. Unger, and A. B. Forsythe, *J. Med. Chem.*, **16**, 1217 (1973).
- (6) W. B. Neely and W. K. Whitney, *J. Agr. Food Chem.*, **16**, 571 (1968).
- (7) C. Hansch and R. Kerley, *J. Med. Chem.*, **13**, 957 (1970); J. K. Seydel, *Mol. Pharmacol.*, **2**, 259 (1966); C. Hansch and A. R. Steward, *J. Med. Chem.*, **7**, 691 (1964); R. W. Fuller, M. M. Marsh, and J. Mills, *J. Med. Chem.*, **11**, 397 (1968); C. Hansch and E. J. Lien, *J. Med. Chem.*, **14**, 653 (1971); *Biochem. Pharmacol.*, **17**, 709 (1968).
- (8) C. Hansch and T. Fujita, *J. Amer. Chem. Soc.*, **86**, 1616 (1964).
- (9) W. Spendley, G. R. Hext, and R. F. Himsworth, *Technometrics*, **4**, 441 (1962).
- (10) N. Kakeya, N. Yata, A. Kamada, and M. Aoki, *Chem. Pharm. Bull.*, **18**, 191 (1970).
- (11) C. Hansch and R. Kerley, *J. Med. Chem.*, **13**, 957 (1970).