**Method II.** (a) A solution of the sodium salt of the compound with a hydroxy group (0.1 mM) in $H_2O$ (50 μg/mL) was extracted with an organic solvent one-half volume × 2) containing TOMAC (25 mM), and the organic solvent layer was refluxed. The reaction mixture was reextracted with 30 mM NaI/$H_2O$. The purification procedure was similar to that of method I (Table VIII).

(b) To a solution of **6** (144 mg) in DMF (12 mL) was added 5% TOMAC/$CH_2Cl_2$ (500 mL), and the mixture was refluxed for 22 h. The reaction mixture was extracted with 6% NaI/$H_2O$ (170 mL), and the aqueous layer was treated similarly to give **17** (40.6 mg) as a freeze–dried powder.

The physicochemical properties of the derivatives are shown in Table II.

**Determination of in Vitro and in Vivo Antibacterial Activity.** The MIC was determined by the agar dilution method.[12] The protective effect in Slc:ICR mice was determined as described previously.[12] The 50% effective dose ($ED_{50}$) was calculated by the method of Reed and Muench[13] from the survival rate recorded 5 days after infection.

**Determination of β-Lactamase Inhibitory Activity and Antibacterial Synergy Test.** The β-lactamase inhibitory activity was determined as described previously[5] and expressed in terms of $I_{50}$, the concentration required to inhibit β-lactamase activity by 50%. The potentiation of the antibacterial activity of ampicillin and cefotiam by carbapenem antibiotics was examined by the 2-fold dilution method with Mueller–Hinton agar (Difco) as described previously.[6]

**Determination of the Stability to Mouse Renal Enzyme(s).** The carbapenem antibiotic (50 μg/mL) was incubated in a 10% mouse kidney homogenate at 30 °C. At intervals, we determined the amount of the residual carbapenem antibiotic by assaying the activity to inhibit the β-lactamase.

**Acknowledgment.** We are grateful to Drs. M. Yoneda, T. Kishi, and A. Imada in our Central Research Division for their encouragement throughout this work. Thanks are also due to the members of the fermentation and physical analysis sections for their support and to Y. Nohara for his skillful assistance.

**Registry No.** 1, 76025-74-6; **2**, 76035-86-4; **3**, 83510-01-4; **4**, 57459-82-2; **5**, 12795-21-0; **6**, 68510-62-3; **7**, 68421-49-8; **8**, 83916-36-3; **9**, 80994-11-2; **10**, 83916-37-4; 11, 83916-38-5; **12**, 80994-12-3; **13**, 83916-39-6; **14**, 83916-40-9; **15**, 83916-41-0; **16**, 83916-42-1; **17**, 75443-31-1; **18**, 75443-29-7; **19**, 83916-43-2; **20**, 83916-44-3; **21**, 83916-45-4; **22**, 83916-46-5; tri-*n*-octylmethylammonium chloride, 5137-55-3.

(12) K. Tsuchiya, M. Kida, M. Kondo, H. Ono, M. Takeuchi, and T. Nishi, *Antimicrob. Agents Chemother.*, 14, 557 (1978).
(13) L. J. Reed and H. Muench, *Am. J. Hyg.*, 27, 493 (1938).

# *Notes*

## Selection of Test Series by a Modified Multidimensional Mapping Method

Raymond Wootton

*Department of Biophysics, Wellcome Research Laboratories, Langley Court, Beckenham, Kent, England.* *Received May 3, 1982*

Maintaining a realistic minimum distance between compounds in a defined multidimensional parameter space ensures well-spread sets of parameter values. It has been suggested, however, that the use of this multidimensional mapping method may lead to series of compounds with high multicollinearities of parameter values. An alternative method, multidimensional mapping by distance and determinant, is discussed here. This method maximizes the determinant of the interparameter correlation matrix as well as maintaining the minimum distance criterion. Its performance is compared with other methods, and it is shown that collinearities may be overcome or maintained at low levels when this method is used.

Multiple linear regression analysis is the most commonly used technique for identifying quantitative relationships between biological activity and the physicochemical parameter values of series of compounds. The chances of establishing such relationships, which will be of use for the prime objective of predicting the activity of further members of the series, depend critically upon there being sufficient variation in the parameter values and no seriously high interparameter correlations (collinearities) for the compounds used to derive the relationships. Multidimensional mapping (MM) is a technique that has been used to help achieve these goals in the design of series of compounds.[1,2] Good variation is ensured by maintaining a set minimum distance between compounds in the scaled multidimensional parameter space, and low collinearities are achieved by choosing compounds close to the center of gravity in space of the compounds already selected. This method has been criticized in two recent papers[3,4] on the grounds that insufficient weight is given to ensuring lack of correlation between parameters. Martin and Panas[3] examined the physicochemical properties of a series of methoxychlor analogues[2] that were designed partly with the help of this method and concluded that its use resulted in a computer-designed series of compounds with too high a degree of multicollinearity. In fact, in this practical application of QSAR methodology, three different methods of computer-aided compound choice were used, and the analysis of Martin and Panas is therefore oversimplified. If the properties of the compounds are examined in greater detail, it becomes clear that the MM method does not, of itself, lead to high multicollinearities and, indeed, goes some way toward decreasing collinearities already present in a data set.

However, if, as was the case for the methoxychlor analogues, there are very high collinearities for the compounds already made, then it may be necessary to adopt an alternative strategy to aid compound choice. Franke et al.[4]

(1) R. Wootton, R. Cranfield, G. C. Sheppey, and P. J. Goodford, *J. Med. Chem.*, 18, 607 (1975).
(2) P. J. Goodford, A. T. Hudson, G. C. Sheppey, R. Wootton, M. H. Black, G. J. Sutherland, and J. C. Wickham, *J. Med. Chem.*, 19, 1239 (1976).
(3) Y. C. Martin and H. N. Panas, *J. Med. Chem.*, 22, 784 (1979).

0022-2623/83/1826-0275$01.50/0 

Table I. Test Series of Substituents Selected by MMDD[a]

| no. | substituents selected[b] | | | | | | | | | | $D_s$[c] | $V_s$[c] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 3 | 14 | 19 | 21 | 22 | 25 | 30 | 32 | 35 | 0.843 | 1.036 |
| 2 | 1 | 6 | 14 | 17 | 19 | 20 | 27 | 28 | 31 | 32 | 0.454 | 1.353 |
| 3 | 2 | 8 | 9 | 10 | 14 | 17 | 22 | 23 | 27 | 35 | 0.517 | 1.236 |
| 4 | 1 | 3 | 8 | 9 | 14 | 17 | 18 | 20 | 32 | 35 | 0.453 | 1.176 |
| 5 | 2 | 7 | 14 | 16 | 20 | 21 | 25 | 30 | 32 | 35 | 0.598 | 1.118 |
| 6 | 1 | 3 | 8 | 14 | 19 | 21 | 30 | 31 | 32 | 35 | 0.731 | 1.158 |
| 7 | 1 | 3 | 12 | 17 | 19 | 20 | 26 | 30 | 31 | 32 | 0.570 | 1.294 |
| 8 | 1 | 3 | 8 | 10 | 12 | 17 | 18 | 20 | 22 | 34 | 0.400 | 1.236 |
| 9 | 1 | 4 | 10 | 13 | 17 | 20 | 26 | 27 | 31 | 34 | 0.548 | 1.202 |
| 10 | 2 | 7 | 14 | 19 | 21 | 22 | 25 | 30 | 32 | 35 | 0.717 | 1.123 |

[a] MMDD = multidimensional mapping by distance and determinant. Minimum distance between compounds in the scaled $\pi$, $F$, $R$, and MR parameter space = 0.45. [b] 1 = H; 2 = Me; 3 = Et; 4 = *n*-Pr; 5 = *i*-Pr; 6 = *n*-Bu; 7 = *t*-Bu; 8 = Ph; 9 = $CF_3$; 10 = OH; 11 = OMe; 12 = OEt; 13 = O-*n*-Pr; 14 = O-*i*-Pr; 15 = O-*n*-Bu; 16 = O-*n*-Am; 17 = OPh; 18 = OAc; 19 = $NH_2$; 20 = $NMe_2$; 21 = NHAc; 22 = $NO_2$; 23 = CHO; 24 = Ac; 25 = $CO_2Me$; 26 = $CO_2Et$; 27 = $CONH_2$; 28 = SMe; 29 = $SO_2Me$; 30 = $SO_2NH_2$; 31 = CN; 32 = F; 33 = Cl; 34 = Br; 35 = I. [c] $D_s$ is the determinant of the interparameter correlation matrix for the selected substituents, and $V_s$ is the variance coefficient (see text).

have presented such an alternative by combining the techniques of multidimensional mapping and principal component analysis. This method (PCMM) gives lower correlations generally, but these are obtained at the expense of a loss of variance. An alternative, much simpler modification of the basic MM method is discussed here. This method emphasizes correlation by selection of compounds to maximize the determinant of the interparameter correlation matrix and may be called multidimensional mapping by distance and determinant (MMDD). In fact, the MMDD method was also used in the design of the methoxychlor series and was successful in reducing collinearities generally, within the constraints of this real series design. The present paper compares the MMDD method with other methods of choosing substituents for an ideal hypothetical series and examines the performance of the MM and MMDD methods in the design of methoxychlor analogues.

## Method

Franke et al.[4] compared various methods of selecting substituents to occupy a single para position on an aromatic ring by selecting 10 sets of 10 substituents using each method. The same approach was adopted here. A standard parameter set of $\pi$, $F$, $R$, and MR was used, and sets of 10 substituents were chosen from the 35 in the Wellcome data bank.[5] Parameter values were scaled to lie in equivalent ranges (±0.5) as described previously.[1] The first substituent was chosen at random. The second was chosen as the nearest to the first but greater than a set minimum distance from it. Subsequent substituents were then chosen to satisfy two criteria: (1) those substituents closer than the preset minimum distance from previously selected substituents were rejected; (2) the parameters for each remaining eligible substituent were included in turn with those for the selected substituents, and the one giving the highest value for the determinant of the resulting interparameter correlation matrix was chosen.

## Results and Discussion

The results of 10 such selections of 10 substituents by the MMDD method with a minimum distance of 0.45 unit between substituents are given in Table I. Also shown in this table are $D_s$, the determinant of the interparameter correlation matrix, and $V_s$, a measure of the variance in the parameters, as used by Franke et al.[4] and defined as:

$$V_s = \frac{1}{m} \sum_{i=1}^{m} \frac{S_{is}^2}{S_{ip}^2}$$

Where $m$ is the number of parameters, $S_{is}^2$ is the variance of the $i$th parameter for the 10 selected substituents and $S_{ip}^2$ is the variance of this parameter for the 35 data bank substituents. In general, the higher the values of $D_s$ and $V_s$ the better. $D_s$ approaches 1 as interparameter correlations approach 0, but an upper limit on $V_s$ is not well-defined. However, $V_s$ should ideally be greater than 1.

Table II. Comparison of Average $D_s$ and $V_s$ Values from Different Selection Methods

| | $\bar{D}_s$[a] | $s_{\bar{D}_s}$ | $\bar{V}_s$[a] | $s_{\bar{V}_s}$ |
|---|---|---|---|---|
| total substituents | 0.57 | | 1 00 | |
| MM[b] method | 0.41 | 0.14 | 1.18 | 0.08 |
| PCMM[b] method | 0.56 | 0.11 | 1.13 | 0.15 |
| MMDD[b] method | 0.58 | 0.14 | 1.19 | 0.09 |

[a] Mean determinant, $\bar{D}_s$, and variance coefficient, $\bar{V}_s$, for 10 selections of 10 substituents for the various selection methods. $s_{\bar{D}_s}$ and $s_{\bar{V}_s}$ are the corresponding standard deviations. [b] MM = multidimensional mapping according to Wootton et al.[1] PCMM = principal component analysis with multidimensional mapping according to Franke et al.[4] MMDD = multidimensional mapping by distance and determinant as used here.

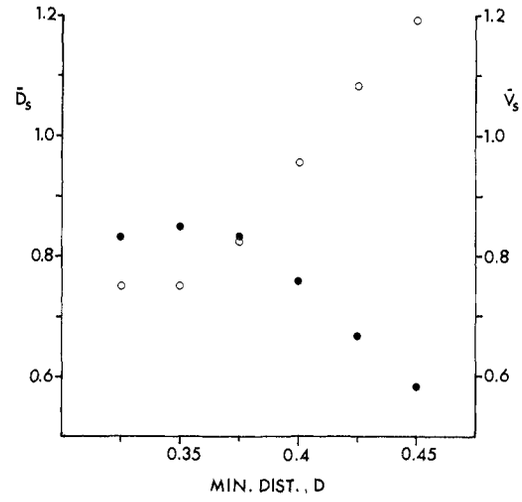


**Figure 1.** The effect of the minimum distance between substituents on the mean determinant of the interparameter correlation matrix, $\bar{D}_s$ (●), and the mean variance coefficient, $\bar{V}_s$ (○), for 10 selections each of 10 substituents.

Table II gives the mean values of $D_s$ and $V_s$ for the 10 series selected by the MMDD method and the corresponding figures for the whole data bank of 35 substituents and the 10 sets of 10 substituents selected by Franke et al.[4] using the simple MM method and the PCMM method. The MMDD method used here gives the highest values for $\bar{D}_s$ and $\bar{V}_s$ overall, although there is considerable variation in the values for individual selections for all three methods as shown by the large standard deviations. The MMDD method would therefore appear to be the method of choice if interparameter correlations are a real problem, since it is simpler to use than the PCMM method and requires no subjective decisions.

In fact, the reduction of interparameter correlations can easily be given further emphasis in the MMDD method by lowering the minimum distance between substituents. However, the resulting lower correlations are obtained at the expense of a loss of variance, since lowering the min-

(4) W. J. Streich, S. Dove, and R. Franke, *J. Med. Chem.*, **23**, 1452 (1980).
(5) F. E. Norrington, R. M. Hyde, S. G. Williams, and R. Wootton, *J. Med. Chem.*, 18, 604 (1975).

Table III. Interparameter Correlation Matrices and Squared Multiple Correlation Coefficients, $R^2$, at Various Stages in the Design of Methoxychlor Analogues

| | $\pi$ | $F$ | $R$ | $MR_o$ | $MR_{mp}$ | $R^2$ |
|---|---|---|---|---|---|---|
| | | | A. $n = 5$ | | | |
| $\pi$ | 1.00 | | | | | |
| $F$ | −0.76 | 1.00 | | | | |
| $R$ | 0.32 | −0.78 | 1.00 | | | |
| $MR_o$ | 0.78 | −0.85 | 0.67 | 1.00 | | |
| $MR_{mp}$ | 0.72 | −0.19 | −0.28 | 0.13 | 1.00 | |
| | | | B. $n = 12$ | | | |
| $\pi$ | 1.00 | | | | | 0.95 |
| $F$ | −0.58 | 1.00 | | | | 0.69 |
| $R$ | −0.37 | −0.10 | 1.00 | | | 0.54 |
| $MR_o$ | 0.34 | −0.10 | −0.03 | 1.00 | | 0.84 |
| $MR_{mp}$ | 0.72 | −0.34 | −0.22 | −0.29 | 1.00 | 0.91 |
| | | | C. $n = 16$ | | | |
| $\pi$ | 1.00 | | | | | 0.73 |
| $F$ | −0.37 | 1.00 | | | | 0.24 |
| $R$ | −0.18 | −0.24 | 1.00 | | | 0.15 |
| $MR_o$ | 0.31 | −0.18 | 0.17 | 1.00 | | 0.52 |
| $MR_{mp}$ | 0.67 | −0.18 | −0.22 | −0.27 | 1.00 | 0.71 |

imum distance criterion will naturally lead to poorer variance in parameter values. This is shown in Figure 1, which was generated by selecting 10 sets of 10 substituents for various values of the minimum distance between substituents. As the minimum distance $D$ is lowered, the values of $\bar{D}_s$ increase, but the values of $\bar{V}_s$ fall dramatically.

The selections presented here were performed automatically with an initial random choice of first substituent. In the practical design of a series of compounds, however, it is necessary to consider synthetic feasibility as an important criterion in compound choice and also to include existing compounds within the strategy. Such constraints were evident during the design of methoxychlor analogues[2] discussed earlier, and it is instructive to examine retrospectively the performance of both the simple MM method and the MMDD method in this real series design. Problems with parameter collinearities were experienced throughout the design of this series, which were mainly due to the prolific use of alkyl and alkoxy substituents and a poor method of selection of the first set of compounds.

The first five compounds (see Table I of Wootton et al.[2]) were selected by constructing a regular polyhedron in the five-dimensional parameter space of $\pi$, $F$, $R$, $MR_o$ (molar refraction at the ortho position), and $MR_{mp}$ (molar refraction summed over the meta and para positions). Synthetically feasible compounds were then chosen to lie as close as possible to the apexes of this polyhedron in the hope of obtaining effectively orthogonal parameter sets. In fact, considerable difficulties were encountered in the practical application of this method, and extensive manipulation of the size, position, and orientation of the polyhedron was required before synthetically feasible compounds could be found. Table III (A) shows the resulting interparameter correlation matrix for the five compounds eventually made; far from being orthogonal, the parameters contain high collinearities. It was therefore clear that this method was not achieving the desired results, although the ranges of parameter values were good for the five compounds. An alternative method was therefore sought.

The next seven compounds were selected by the simple MM method (and synthetic feasibility criteria), and the resulting interparameter correlation matrix after the synthesis of 12 compounds is shown in Table III (B). Collinearities have been reduced generally, but there are still two high correlations that are significant at the 5% level. There are also high multicollinearities present in the data set, as noted by Martin and Panas.[3] This is shown by the high values for $R^2$ [Table III (B)] that represents the proportion of the variance in each parameter, which may be explained by a linear combination of the other four.

At ths stage in the design process, it was apparent that the MM method was not producing sufficient reduction in collinearities. Therefore, the decision was taken to emphasize collinearity by switching to the MMDD method and also lowering the permitted minimum distance between compounds (cf. Figure 1). An additional four compounds were selected by this new method (compounds 19–22 in Table I of Wootton et al.[2]), and the resulting interparameter correlation matrix for the final set of 16 compounds chosen with the aid of a computer is shown in Table III (C). Collinearities have been reduced to acceptable levels, except for the one remaining correlation between $\pi$ and $MR_{mp}$ which is still significant at the 5% level. The $R^2$ values are also much improved compared with the situation after 12 compounds had been made and again indicate that the only significant correlation remaining is between $\pi$ and $MR_{mp}$. The use of the MMDD method has therefore achieved the desired effect as far as was possible in this real series design. Breaking the correlation between $\pi$ and $MR_{mp}$ would have required introducing large hydrophilic substituents, and such substituents were not represented in the data bank used, which was deliberately restricted as part of the drug design exercise.

The present paper has shown that interparameter correlations can be minimized with the MMDD method of compound selection. In fact, the originally proposed MM method[1] does not lead to serious collinearities in the ideal case; of the 10 selections reported by Franke et al.[4] by the MM method, only one contains a correlation that is significant at the 5% level. However, as with any selection method, when it is compromised by allowing other factors to dominate compound choice, then it is possible that considerable deviations from the ideal may occur. The experience with the use of both methods in the design of methoxychlor analogues shows that the MMDD method may well be more robust toward such compromise.