Aust. J. Chem., 1995, **48**, 1385–1400

Structure–Activity Correlations for Psychotomimetics. III* Tryptamines

Brian W. Clare

School of Mathematical and Physical Sciences, Murdoch University, Murdoch, W.A. 6150.

Abstract

Based on electronic properties calculated by the semiempirical PM3 method, a quantitative structure–activity relationship has been derived for a group of psychotomimetic tryptamines. Data analysis by partial least squares reveals two significant components. The major property affecting potency was the difference in energy between the highest occupied and lowest unoccupied molecular orbital. Less significant was the effect of molecular size, shape and polarizability, in particular, the dimension in the direction of the ethylamine side chain, involving the substituents on the amine nitrogen. The results are consistent with earlier studies on phenylalkylamines.

Introduction

In Parts I and II[1] of this series, quantitative structure–activity relationships were derived for a series of phenylalkylamines. In the present paper, a similar study is conducted on the tryptamines. It must be noted at the outset that the number of known psychotomimetic tryptamines is much smaller than that of the phenylalkylamines, probably because of the ease of synthesis of the latter. The experimental data are also less homogeneous than in that series, which was largely attributable to a single research group, that of A. T. Shulgin and his collaborators.

While the tryptamines constitute an important subclass of the psychotomimetics that resemble LSD and mescaline in their effects, there have been few attempts to develop a quantitative structure–activity relationship for them. In their review, Gupta, Singh and Bindal[2] noted a number of studies on receptor ligand binding and other *in vitro* tests, involving quantum chemical indices such as frontier orbital densities and charges on atoms of the indole ring system, total orbital energy and hydrophobicity of the 7-substituent, but none on psychotomimetic potency. Snyder and Merril[3] mentioned two tryptamines (psilocin and 6-hydroxy-*N*,*N*-diethyltryptamine) in a list of psychotomimetics of several classes, showing a general increase of psychotomimetic activity with HOMO (highest occupied

* Part II, *Chemom. Intell. Lab. Syst.*, 1993, **18**, 71.

[1] Clare, B. W., *Chemom. Intell. Lab. Syst.*, 1993, **18**, 71.
[2] Gupta, S. P., Singh, P., and Bindal, M. C., *Chem. Rev.* 1983, **83**, 633.
[3] Snyder, S. H., and Merril, C. R., *Proc. Natl Acad. Sci. U.S.A.*, 1965, **54**, 258.

Manuscript received 2 February 1995

molecular orbital) energy, and Kang and Green[4] made similar observations on the same two compounds. The 6-hydroxy compound is probably not psychotomimetic in humans. Kline *et al.*[5] found some correlation between the first two ionization potentials (measured by photoelectron spectroscopy) and behavioural activity in rats in a series of five tryptamines. Migliaccio *et al.*[6] have shown by n.m.r. measurements that an interaction occurs between the hydroxy and amine groups in psilocin, and suggest that this facilitates entry into the central nervous system. Glennon,[7] in a study of oxygen-substituted *N*,*N*-dimethyltryptamines by the Pariser Parr Pople technique, concluded that several quantum chemical properties of the tryptamines did not correlate with activity. This series of tryptamines included 6- and 7-hydroxy and methoxy derivatives that were indicated by early animal experiments to be psychotomimetic.[8] These indications have not been confirmed in humans, and so will not be considered in this study.

Sixteen tryptamines with known hallucinogenic properties in humans were found in the literature,[9,10] and their molecular electronic properties were calculated with the PM3 Hamiltonian of the quantum chemical package MOPAC, Version 6.[11] Only descriptors that can be computed or estimated without needing access to the drug were employed. This enables the activity of new substances to be calculated without the need for their synthesis. This sample represents nearly all of the known psychotomimetic tryptamines. Bufotenin was not included, as there are some doubts as to the nature of its effects.[12] Likewise, psilocybin was not included as it is believed to be rapidly hydrolysed to psilocin *in vivo*, and psilocin is the active form.

A difficulty in interpretation arises because the route of administration of the drug is not constant throughout the series. Some of the drugs, e.g. *N*,*N*-dimethyltryptamine, are inactive when taken orally. Others, however, were only administered by this route. Since there seems to be little prospect of collecting further systematic data on these compounds, this problem is unavoidable. It is not obvious how reliably to correct for this deficiency in the data, so all activities were treated as comparable, and the mode of administration should be considered parenteral. Since in no case were the drugs given intravenously, the discrepancies introduced should not be large. In this connection, it should be noted that, in one key study involving psilocybin, similar doses were used orally and parenterally, and no difference in effects was reported.[13]

It is desirable to study as many and varied a group of drugs as possible. This enables the separation of contributions to drug activity from different molecular

[4] Kang, S., and Green, J. P., *Proc. Natl Acad. Sci. U.S.A.*, 1970, **67**, 62.
[5] Kline, T. B., Benington, F., Morin, R. D., Beaton, J. M., Glennon, R. A., Domelsmith, L. N., Houk, K. N., and Rozeboom, M. D., *J. Med. Chem.*, 1982, **5**, 1381.
[6] Migliaccio, G. P., Shieh, T. N., Byrn, S. R., Hathaway, B. A., and Nichols, D. E., *J. Med. Chem.*, 1981, **24**, 206.
[7] Glennon, R. A., *Res. Commun. Chem. Path. Pharmacol.*, 1974, **9**, 185.
[8] Gessner, P. K., Godse, D. D., Krull, A. H., and McMullan, J. M., *Life Sci.*, 1968, **7**, 267.
[9] Shulgin, A. T., in 'Psychotropic Agents, Part III' (Eds F. Hoffmeister and G. Stille) pp. 3–29 (Springer: New York 1982).
[10] Repke, D. B., Grotjahn, D. B., and Shulgin, A. T., *J. Med. Chem.*, 1985, **28**, 892.
[11] Stewart, J. J. P., *Quantum Chem. Program Exch. Bull.*, 1990, **10**, 86.
[12] Glennon, R. A., and Rosecrans, J. A., *Neurosci. Biobehav. Rev.*, 1982, **6**, 489.
[13] Hollister, L. E., Prusmack, J. J., Paulsen, J. A., and Rosenquist, N., *J. Nerv. Ment. Dis.*, 1960, **131**, 428.

properties. However, the use of animal data has been avoided because of the risk of contamination of the results. This risk is present even with the best animal models, such as the discrimination paradigm. Although activities in rats determined by this method correlate with activity in humans, the correlation in the case of phenylalkylamines is not as good as the quantitative structure–activity relationship correlations that can be achieved,[14] and the animal data should be used to settle more qualitative questions, such as the discrimination between hallucinogenic, psychomotor stimulant and entactogenic activity. Thus only human data have been used in the present study, and there is little possibility of extending the series in the future, because of legal and ethical strictures.

The technique used by Shulgin and coworkers to establish the potency of hallucinogens is referred to as 'double conscious'.[15] The standard 'double blind' technique of medical experimentation is inapplicable in this area, firstly because an active drug is easily recognized as such by the subject, and secondly because the active cooperation of the subject is required in a series of experiments, to elucidate the nature of the effects. Briefly, after establishment of safe levels in several animal species, the drug is given to human subjects, starting at submilligram levels and increasing, at 50% increments, at intervals of several days until activity is observed. The effective dose is considered to be the mean of the minimum dose to produce a recognizable effect and that necessary to produce the maximum effect. This is then expressed as mescaline units (MU), the ratio of the effective dose of mescaline (3·75 mg/kg) to that of the substance in question.

Recent research has indicated that psychotomimetic activity has more than one dimension.[16] Thus some drugs, such as LSD, are classified as hallucinogens and others, such as methylenedioxymethamphetamine, as entactogens. Both of these groups overlap with psychomotor stimulants of the amphetamine type. To the present time, few of the compounds dealt with in this paper have been decisively classified into one or another of these groups, and so it is not possible to treat this distinction at this time. Thus it must be remembered that the psychotomimetic activity predicted by equations derived in this work is not homogeneous, but contains a measure of at least three components. This does not invalidate the analysis, but obviously a more useful quantitative structure–activity relationship could be developed if this information were available.

The identities of the drugs, and their activities, taken from the literature, are given in Table 1.

The general strategy adopted in this study was to calculate a large number of trial descriptors that have been shown to be important in other cases, or are suggested in theories of drug–receptor binding, and, as far as possible, eliminate those which were redundant. Because the number of compounds is small and that of the descriptors is large, the relatively new partial least-squares (PLS) method[17] was used in preference to stepwise regression in the analysis of the results.

[14] Clare, B. W., *J. Med. Chem.*, 1990, **33**, 687.
[15] Shulgin, A. T., Sargent, T., and Naranjo, C., *Nature (London)*, 1969, **221**, 537.
[16] Nichols, D. E., *J. Psychoactive Drugs*, 1986, **18**, 305.
[17] Sjostrom, M., Wold, S., Lindberg, W., Persson, J. A., and Martens, H., *Anal. Chim. Acta*, 1983, **150**, 61.

The number of significant components is determined by cross-validation.[18] Thus groups of points are left out of the construction of the equation, and are then predicted, based on the remaining points. A sum of squares of residuals of the excluded points (PRESS) is accumulated until each point has been excluded once and only once. Then a cross-validated R^2 is constructed:

$$R_{cv}{}^2 = (S_t - S_p)/S_t$$

where S_t is the sum of squares of deviations from the mean, and S_p (= PRESS) is the sum of squares of the predicted residuals. The number of components selected is that which maximizes the $R_{cv}{}^2$.

Table 1. Identities and activities of drugs considered in this study

Drug No.	Tryptamine	Activity (MU)
(1)	*N*,*N*-dimethyl	5
(2)	*N*,*N*-diethyl	5
(3)	5-methoxy-*N*,*N*-dimethyl	50
(4)	5-methoxy-*N*,*N*-diisopropyl	50
(5)	α-methyl	15
(6)	5-methoxy-α-methyl	100
(7)	*N*,*N*-diisopropyl	8
(8)	*N*,*N*-dipropyl	5
(9)	*N*,*N*-diallyl	5
(10)	4-hydroxy-*N*,*N*-dimethyl (psilocin)	40
(11)	4-hydroxy-*N*,*N*-diethyl	40
(12)	4-hydroxy-*N*-isopropyl-*N*-methyl	40
(13)	4-hydroxy-*N*,*N*-diisopropyl	20
(14)	5-methoxy-*N*-isopropyl-*N*-methyl	70
(15)	*N*-isopropyl-*N*-methyl	14
(16)	4-methoxy-*N*-methyl-*N*-isopropyl	13

The quality of a PLS model can be degraded by the presence of irrelevant variables. A procedure (GOLPE) which avoids the problem of chance correlation[19,20] has recently been described for selection of variables in PLS.

Partial least squares used in this way has considerable advantages over multiple regression. It is not affected by collinearity, as orthogonal components are extracted. The methods of selecting the number of components, the subset of variables and the determination of errors in regression coefficients are all non-parametric, and hence not dependent on any assumptions concerning the distribution of the variables. The GOLPE technique maximizes the predictive ability of the model by eliminating an optimum number of redundant variables.

Calculations

The GOLPE calculations were done on a Hewlett Packard series 9000 model 735 computer. All other calculations were done on a Toshiba T5200 personal computer

[18] Cramer, R. D., III, Bunce, J. D., and Patterson, D. E., *Quant. Struct.-Act. Relat.*, 1988, **7**, 18.
[19] Cruciani, G., Baroni, M., Clementi, S., Costantino, G., and Skagerberg, B., *J. Chemom.*, 1992, **6**, 335.
[20] Baroni, M., Clementi, S., Cruciani, G., Costantino, G., Riganelli, D., and Oberrauch, E., *J. Chemom.*, 1992, **6**, 347.

with a Lahey F77L3/EM32 FORTRAN compiler. The stepwise regression, PLS, GOLPE and cluster analysis programs were written by the author. The multiple regression was done with the BMDP package.[21] The Cartesian coordinates of the atoms were calculated from standard bond lengths and angles with the program EUCLID,[22] and the geometries of the neutral molecules were then fully optimized with the PM3 method. It is recognized that the resulting conformations are not those of the molecule on the receptor site. The actual optimized conformations were similar for all substances, except for the *N*,*N*-dipropyl (8) and the *N*,*N*-diallyl (9) derivatives, and may be seen for *N*,*N*-dimethyltryptamine in Fig. 1. The descriptors chosen are not very sensitive to conformation. The electronic descriptors are mainly properties of the indole nucleus. For all of the drugs, these properties would be influenced in the same way by protonation of the amine nitrogen. Thus although it is uncertain whether the drugs are active in their unprotonated or protonated form, the influence of the protonation should tend to cancel.

Since lysergic acid diethylamide (LSD) contains tryptamine as a substructure, and since LSD is the most powerful psychotomimetic known and its effects are similar to those of the tryptamines, it is plausible that the conformation of the tryptamine moiety in LSD is especially favourable for activity. To test this hypothesis, the structure of LSD was optimized with PM3, as for the tryptamines. By using the molecular modelling program DTMM,[23] the indole moieties of the tryptamines were superimposed on that of LSD and the dihedral angles of the ethylamine side chains of the tryptamines were adjusted so as to minimize the distance between the amine nitrogen of the tryptamine and that of the LSD. Then, by using the molecular mechanics program MM2,[24] the dihedrals of the ethylamine moiety were constrained to these values, and all other geometric parameters were optimized. This process, of adjusting the dihedrals to best match LSD and optimizing the rest of the structure, was iterated to convergence. The iteration did not converge in all cases. The difference in energy between the strained tryptamine so generated and the corresponding geometry-optimized tryptamine was calculated by MM2. This energy was found not to correlate significantly with the activity of the tryptamine, and was not considered further. The values of the energy difference, for those cases which were tried, and where the iteration converged, are shown in Table 2.

The GOLPE method was applied to the tryptamine series as described in the Introduction;[20,21] 5 groups of points were used in cross-validation, and the PRESS statistic was averaged over 30 random selections of points in the groupings. This was done in both the PLS and GOLPE calculations. The $R_{cv}{}^2$ and σ_p (standard deviation of prediction errors, $\sigma_p = \sqrt{S_p/n}$) parameters were calculated from this averaged PRESS. The GOLPE analysis was carried out by using a fractional factorial design of resolution IV, with 128 runs,[25] augmenting the design matrix

[21] BMDP Dynamic, Release 7, available from BMDP Statistical Software, Inc., 1440 Sepulveda Boulevard, Suite 316, Los Angeles, CA 90025, U.S.A.

[22] Essen, H., *Quantum Chem. Program Exch. Bull.*, 1983, **3**, 13.

[23] Crabbe, M. J. C., and Appleyard, J. R., Desktop Molecular Modeller, Version 2.0, 1991, available from Oxford University Press, England.

[24] Allinger, N. L., MM2(87), available from Quantum Chem. Program Exch., and also from Molecular Design Ltd, 1122 B Street, Hayward, CA 94541, U.S.A.

[25] Box, G. E. P., Hunter, W. G., and Hunter, J. S., 'Statistics for Experimenters' Ch. 12 (John Wiley: New York 1978).

with σ_p, and carrying out PLS on the augmented matrix. The variables were ranked in order of increasing values of their loadings in the first PLS component. That number of contiguous variables which gave a minimium σ_p was selected.

The descriptors that were calculated and the symbols that will represent them are given in Table 3. Because it is usually important in drug action, the hydrophobicity of the whole molecule was included as a descriptor. It was calculated by the method of Broto.[26] This method is readily implemented on a computer,[27] and unlike the better known methods of Hansch[28] and Rekker[29] is entirely automatic, and requires no chemical judgment. Hydrophobicities calculated by Broto's method correlate well with both experimental data and those calculated by the alternative methods. Broto gives an estimate of prediction error for substances not in the training set of 0·4.[26] The corresponding value for Rekker's and Hansch's methods, given by Rekker for a varied group of drugs is similar[30] (0·40 and 0·52 respectively). Meylan and Howard[31] give comparable values. The sum of electrophilic superdelocalizabilities over all atoms was calculated, as another indicator of hydrophobicity effects.[32]

A variety of electronic descriptors were calculated from the PM3 results.[33–36] These included the frontier orbital energies, and the magnitudes of the charges of the most negatively charged and most positively charged atoms in the molecule. The former charge is related to hydrogen bond basicity.[37] In every case the same two atoms were involved. The most negative atom was C2 of the indole, and the most positive was the indole nitrogen. Other charge-related descriptors were the mean of the absolute Mulliken charges on all atoms, and the local dipole index. The latter is the mean absolute value of the charge difference between the members of each bonded pair of atoms.

Other global descriptors were the three linear dimensions derived from the inertial ellipsoid by the formula $L_x = 2\sqrt{5(I_y + I_z - I_x)/2M}$, and corresponding formulae for L_y and L_z. These represent the lengths (in Å) of the axes of an ellipsoid of uniform density 1 u/Å^3 (1·66 g/ml) with the same moments of inertia I_x, I_y and I_z as the molecule. The molecular weight of the drug is M. These dimensions are measures of gross steric effects. The van der Waals volume and area, calculated by the program MOLSV,[38] also served as steric descriptors.

[26] Broto, P., Moreau, G., and Vandycke, C., *Eur. J. Med. Chem.*, 1984, **19**, 71.
[27] Croizet, F., Dubost, J. P., Langlois, M. H., and Audry, E., *Quant. Struct.–Act. Relat.*, 1991, **10**, 211.
[28] Hansch, C., and Leo, A., 'Substituent Constants for Correlation Analysis in Chemistry and Biology' (John Wiley: New York 1979).
[29] Rekker, R. F., and Mannhold, R., 'Calculation of Drug Lipophilicity, the Hydrophobic Fragmental Constant Approach' (VCH: Weinheim 1992).
[30] Rekker, R. F., ter Laak, A. M., Mannhold, R., *Quant. Struct.–Act. Relat.*, 1993, **12**, 152.
[31] Meylan, W. M., and Howard, P. H., *J. Pharm. Sci.*, 1995, **84**, 83.
[32] Cammarata, A., in 'Molecular Orbital Studies in Chemical Pharmacology' (Ed. L. B. Kier) pp. 156–190 (Springer: New York 1970).
[33] Kikuchi, O., *Quant. Struct.–Act. Relat.*, 1987, **6**, 179.
[34] Lukovits, I., *J. Med. Chem.*, 1983, **26**, 1104.
[35] Cartier, A., and Rivail, J. L., *Chemom. Intell. Lab. Syst.*, 1987, **1**, 335.
[36] Brown, R. E., and Simas, A. M., *Theor. Chim. Acta*, 1982, **62**, 1.
[37] Famini, G. R., and Penski, C. A., *J. Phys. Org. Chem.*, 1992, **5**, 395.
[38] Smith, G. M., QCMP053 (converted by K. J. Tupper), *Quantum Chem. Program Exch. Bull.*, 1988, **8**, 150.

The components of the dipole moment referred to the inertial axes, and also the magnitude of the resultant dipole moment, were included as indicators of electrostatic effects. The three diagonal components of the polarizability tensor served as indicators of van der Waals interactions. As a possibly better indicator of these latter interactions, the ratio of the average polarizability to the van der Waals volume[39] was also included. The directions of the inertial axes relative to the indole nucleus were very similar for all of the molecules except (8) and to a lesser degree (9), and are illustrated for (1) (*N*,*N*-dimethyltryptamine) in Fig. 1. The molecule was in each case oriented so that the amine nitrogen atom had all Cartesian coordinates positive with respect to the origin, which was the centre of gravity of the molecule.

Because only two substituents occur at the 4-position in the indole (hydroxy or methoxy), and one at the 5-position (methoxy), volume and hydrophobicity descriptors would both be uninterpretable, and an indicator variable was used for each of these oxygen substitution patterns. A small range of substituents occurs on the amine nitrogen, and these were represented by their van der Waals volumes and Hansch π values. The values used for the volume and hydrophobicity terms are the same as those used in Part I,[14] and are given in Table 4.

Results

The correlation structure of the data matrix is shown in the dendrogram in Fig. 2. Two features of this graph are striking. First, there is a strong correlation of $\log A$ with E_δ, and to a lesser extent with ΔH_f and P_{yy}. Second, there is a group of five descriptors very highly correlated with the molecular weight, involving area, volume, polarizability and the sum of electrophilic superdelocalizabilities. These are so strongly correlated that they form in effect a single variable. Less strongly correlated with this group are the total energy E_t, the hydrophobicity, the polarizability component P_{xx} and length component L_x along the ethylamine side chain, and the volume V_N and hydrophobicity π_N of the substituents on the amine nitrogen.

PLS *Results*

From the data set consisting of all of the descriptors in Table 3, two components were found to be significant by cross-validation, giving $R^2 = 0{\cdot}844$ and $R_{cv}{}^2 = 0{\cdot}749$. Reducing the matrix with the GOLPE technique, as described in the Calculations section left 24 descriptors, and slightly improved the predictive ability of the equation, with an R^2 of $0{\cdot}825$ and an $R_{cv}{}^2$ of $0{\cdot}773$. A plot of the calculated against the observed logarithm of activity is shown in Fig. 3. The exhibited strong trend, for only two components, inspires confidence in the validity of the analysis.

A loadings plot for the two significant components in this regression is shown in Fig. 4*a*, and a scores plot in Fig. 4*b*. Interpretation of the significance of the components is facilitated by the latter plot. Component **1** separates the drugs into three groups. Those with negative values of component **1** are without substitution on the indole nucleus. Those with positive values less than 2 have

[39] Famini, G. R., Ashman, W. P., Mickiewicz, A. P., and Wilson, L. Y., *Quant. Struct.-Act. Relat.*, 1992, **11**, 162.

Table 4. Volumes and hydrophobicities of the substituents on the amine nitrogen

Substituent	Volume (Å^3)	Hansch π	Substituent	Volume (Å^3)	Hansch π
Methyl	18·6	0·56	propyl	56·0	1·56
Ethyl	37·3	1·06	isopropyl	56·0	1·36
			allyl	52·0	1·26

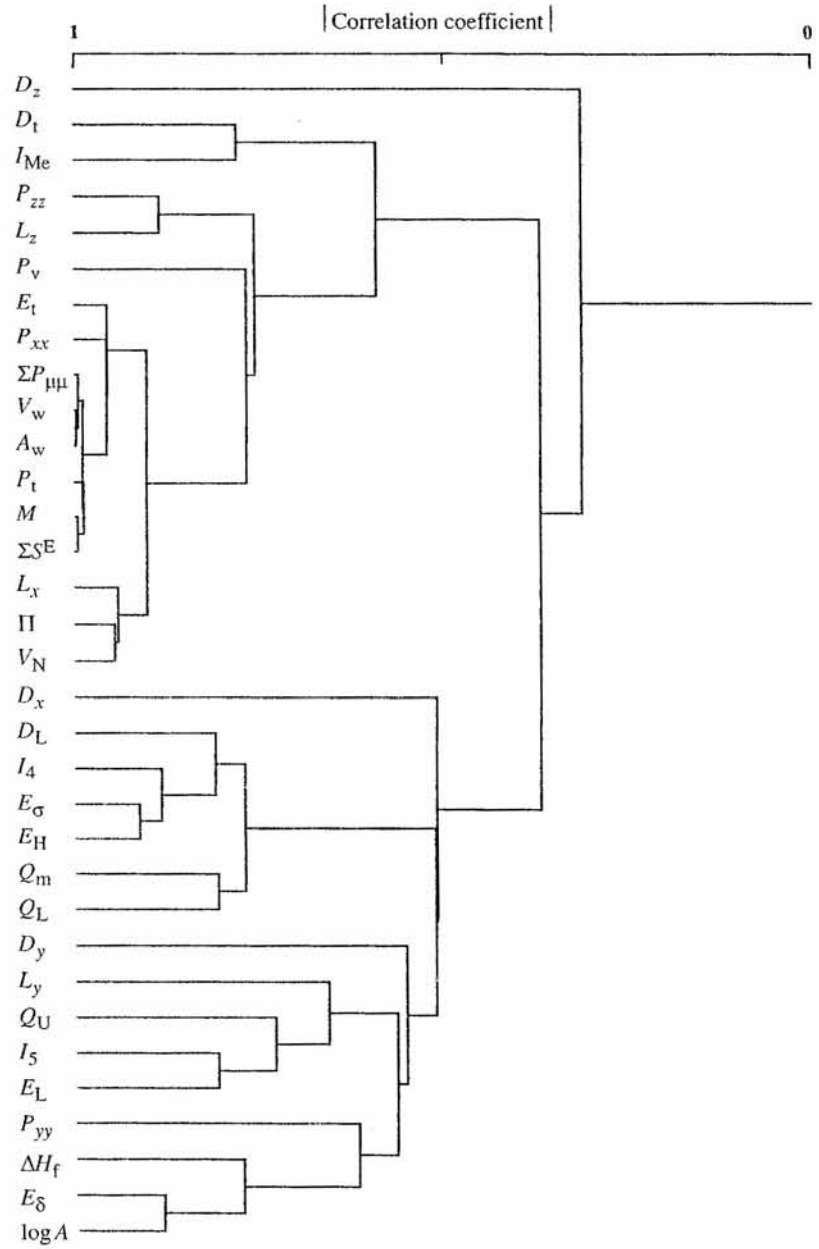

Fig. 2. Dendrogram of the correlation matrix of all descriptors used in this study (log A is the logarithm of the activity).

4-hydroxy or 4-methoxy substitution, while those with positive values greater than 2 have 5-methoxy substitution. While the order of the unsubstituted, 4-hydroxy and 5-methoxy compounds along this component can be explained by E_δ alone, the magnitude of the effect requires either non-linearity or some other contributing variable. Within each of these groups, the interpretation of component **2** is also straightforward. Increasing values of component **2** are associated with increasing size and polarizability (particularly in the X direction) and hydrophobicity. The substituents on the amine nitrogen were the main contributor here, in the order H < methyl < ethyl < allyl ≈ propyl ≈ isopropyl. The value of component **1** decreases slightly along this same series; thus the lines separating the three groups slope upwards and slightly to the left. Component **1** has a much greater influence on psychotomimetic activity than component **2**.

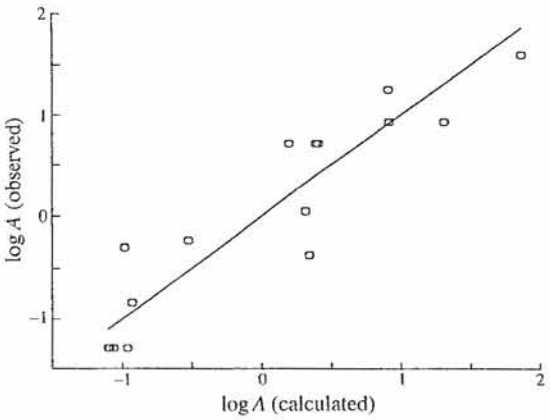

Fig. 3. Plot of observed log activity against that calculated by PLS.

Examination of the loadings of the PLS model showed that the first component is dominated by the difference between the energies of the HOMO and the lowest unoccupied molecular orbital (LUMO). Lesser contributors are the calculated heat of formation, the dipole moment, the indicator variable I_5 and the Y component of the dipole moment. The second component includes contributions from the molecular weight, volume, surface area, average polarizability, sum of self-atom polarizabilities, and the component of the polarizability tensor and the linear dimension of the molecule along the direction of the ethylamine side chain, and the hydrophobicity. Many of these variables may be seen from Fig. 2 to be highly intercorrelated. The second component is thus very much an indication of size and polarizability, both total and in the direction of the ethylamine side chain, and to a lesser extent hydrophobicity.

It is apparent from Fig. 4*a* that there is strong clustering of the descriptors, that is, groups of descriptors compete to explain the same variation. For example, the cluster containing descriptors 1, 4, 6, 15, 16, 17, 21 and 26 (M, P_{xx}, P_{zz}, Π, A_w, V_w, P_t and L_x) contains variables that describe the size of the molecule, particularly along the direction of the side chain, and also its hydrophobicity, which in this data set is highly correlated with size. These descriptors, the most prominent contributors to the second component, are also seen to be highly correlated in Fig. 2. The first component is dominated by $E_L - E_H$ in a negative sense, and dipole moment, especially in the Y direction, and ΔH_f and I_5 in a positive sense. The terms I_5 and D_y, although obviously related, have only a

weak linear correlation. The values of the most important descriptors in this model, except for the indicator variables, are given in Table 5.

Multiple Regression Results

To test further the contributions from the different descriptors, multiple regression was carried out on the main contributing descriptors. Topliss and Edwards[40] have shown that, in stepwise selection of variables in multiple regression, a serious risk of chance correlation arises. As a control on this, they suggested repeating the stepwise regression with random numbers substituted for the dependent variable. This has been adapted to the present study, by repeated stepwise regression

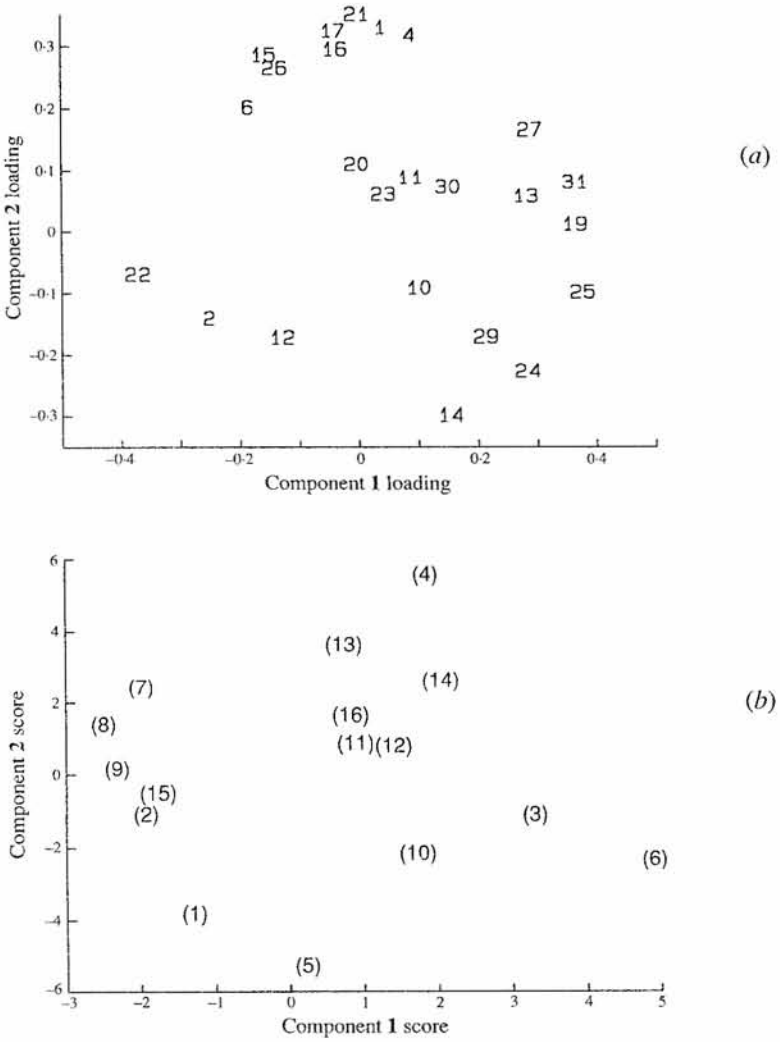

Fig. 4. (*a*) PLS loadings for the tryptamines of Table 1 and the reduced set of descriptors. The numbering is that of Table 3. (*b*) PLS scores plot for the tryptamines of Table 1 and the reduced set of descriptors. The numbering is that of Table 1.

[40] Topliss, J. G., and Edwards, R. J., *J. Med. Chem.*, 1979, **22**, 1238.

Table 5. Values of the most important descriptors for each drug

The symbols are those of Table 2

Drug	M	P_{xx} (Å^3)	P_{yy} (Å^3)	P_{zz} (Å^3)	E_L (eV)	E_δ (eV)
(1)	188·3	139·8	137·1	54·7	0·134	8·419
(2)	216·3	155·5	145·4	75·4	0·128	8·414
(3)	218·3	159·8	158·6	61·5	0·029	8·217
(4)	274·4	186·5	188·2	93·3	−0·006	8·243
(5)	174·3	131·2	128·5	44·9	0·027	8·408
(6)	204·3	160·7	140·5	52·3	−0·100	8·222
(7)	244·4	170·6	156·5	93·1	0·049	8·463
(8)	244·4	169·6	110·0	140·2	0·134	8·420
(9)	240·4	170·9	134·6	124·0	0·126	8·412
(10)	204·3	145·6	146·7	54·8	0·185	8·271
(11)	232·3	160·9	154·7	76·2	0·182	8·270
(12)	232·3	165·2	159·9	66·5	0·157	8·275
(13)	260·4	176·6	167·0	92·0	0·148	8·287
(14)	246·4	174·0	174·5	73·8	−0·005	8·261
(15)	216·3	158·5	151·2	66·1	0·105	8·444
(16)	246·3	173·6	171·8	73·3	0·168	8·295

Drug	D_L (e)	D_x (D)	D_y (D)	D_t (D)	E_t (eV)	P_v
(1)	0·1386	0·4448	−0·3394	0·6087	−1992·7	0·6959
(2)	0·1354	0·3997	−0·4043	0·5886	−2291·6	0·6763
(3)	0·1480	0·6890	0·4595	0·9519	−2435·5	0·7080
(4)	0·1510	0·6449	−0·3249	0·8604	−3033·5	0·6704
(5)	0·1454	0·5192	−0·8867	1·0914	−1843·5	0·7057
(6)	0·1532	0·4751	1·0462	1·1511	−2286·4	0·7166
(7)	0·1405	−0·2674	−0·3005	0·5025	−2590·7	0·6592
(8)	0·1307	0·4211	0·3850	0·5964	−2590·7	0·6584
(9)	0·1452	0·4549	−0·4045	0·6173	−2526·8	0·6947
(10)	0·1794	0·6635	−0·0366	0·7212	−2286·5	0·6997
(11)	0·1707	0·6431	−0·1029	0·6999	−2585·4	0·6800
(12)	0·1745	0·6533	0·0428	0·6656	−2585·5	0·6780
(13)	0·1731	0·6349	−0·0141	0·6607	−2884·5	0·6626
(14)	0·1476	0·7346	−0·3297	0·8059	−2734·6	0·6836
(15)	0·1387	−0·3992	−0·2678	0·4829	−2291·7	0·6731
(16)	0·1543	0·5339	0·1337	0·5598	−2734·5	0·6752

Drug	ΔH_f (kJ/mol)	Q_U (e)	Π	L_x (Å)	L_y (Å)	L_z (Å)
(1)	131·5	0·2914	1·66	15·46	7·19	3·61
(2)	97·1	0·2904	2·20	17·23	7·07	5·36
(3)	−25·1	0·2688	1·79	15·98	9·61	3·38
(4)	−98·8	0·2486	3·42	18·69	9·42	5·42
(5)	118·1	0·2965	0·85	13·30	7·15	3·30
(6)	−38·5	0·2583	0·98	14·76	8·76	3·11
(7)	55·3	0·2486	3·29	17·99	7·44	5·53
(8)	53·6	0·2921	3·12	18·45	8·84	6·16
(9)	299·4	0·2924	2·63	18·26	8·15	5·91
(10)	−50·7	0·2968	1·27	14·95	7·75	3·58
(11)	−85·0	0·2972	1·82	16·71	7·69	5·12
(12)	−93·8	0·2832	1·81	16·80	7·85	3·77
(13)	−125·2	0·2739	2·90	17·65	7·91	5·25
(14)	−66·6	0·2459	2·33	17·91	9·60	3·69
(15)	88·8	0·2739	2·20	17·37	7·39	3·87
(16)	−68·6	0·2742	2·33	16·33	8·93	3·69

following random reassignment of the original dependent variable ($\log A$) to the drugs. From these results a statistical significance can be calculated.[41] Since there are 16 drugs, only those stepwise regressions with four or fewer variables were accepted. In 500 regressions with reassigned dependent variable, the maximum R^2 obtained was 0·876 with four variables selected (significance 4×10^{-4}), and 0·841 with three variables selected (significance $1{\cdot}4\times10^{-4}$). This establishes a risk of chance correlation within acceptable limits.

The regressions which follow were carried out with BMDP-2R, the 'stepwise regression' module.

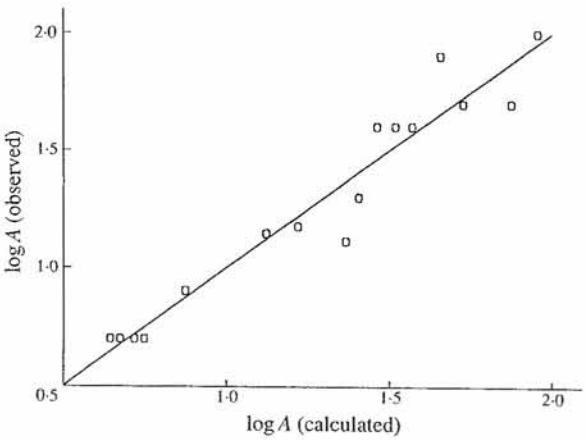

Fig. 5. Conventional regression plot of observed log activity against that calculated from the two descriptors E_δ and D_x.

The orbital energy difference alone gave a very significant regression:

$$\log A = -4{\cdot}56(\pm0{\cdot}66)E_\delta + 39{\cdot}2(\pm5{\cdot}4) \qquad (1)$$

$$N = 16, R^2 = 0{\cdot}773, S = 0{\cdot}224, F = 48{\cdot}0, R_{cv}{}^2 = 0{\cdot}700$$

Here N is the number of points, R^2 the square of the multiple correlation coefficient, S the standard error of estimate, F the Fisher variance ratio and $R_{cv}{}^2$ the cross-validated R^2. The numbers in parentheses are standard errors of estimate. The best two-descriptor regression was:

$$\log A = -5{\cdot}93(\pm0{\cdot}81)E_\delta + 0{\cdot}52(\pm0{\cdot}22)D_x + 51{\cdot}0(\pm6{\cdot}8) \qquad (2)$$

$$N = 16, R^2 = 0{\cdot}842, S = 0{\cdot}194, F = 34{\cdot}9, R_{cv}{}^2 = 0{\cdot}777$$

A plot of observed against calculated log activity based on this equation is shown in Fig. 5. The best four-descriptor regression was:

$$\log A = -7{\cdot}10(\pm0{\cdot}92)E_\delta + 0{\cdot}286(\pm0{\cdot}081)D_t - 0{\cdot}86(\pm0{\cdot}19)D_x - 0{\cdot}34(\pm0{\cdot}13)D_y + 60{\cdot}3(\pm7{\cdot}7) \qquad (3)$$

$$N = 16, R^2 = 0{\cdot}932, S = 0{\cdot}14, F = 38{\cdot}0, R_{cv}{}^2 = 0{\cdot}834$$

[41] Clare, B. W., in 'Methods and Principles in Medicinal Chemistry' (Eds R. Mannhold, P. Krogsgaard-Larsen and H. Timmerman) Vol. 3 'Advanced Computer-Assisted Techniques in Drug Discovery' (Ed. H. van de Waterbeemd) pp. 284, 289 (VCH: Weinheim 1994).

A result similar to equation *(3)* is obtained with I_{Me} replacing D_t. The descriptor I_{Me} is an indicator variable for tertiary versus primary amine, as well as for unsubstituted versus α-methyl-substituted alkylamine chain. Thus the second PLS component can represent such effects as steric accommodation of the molecule, its polarizability and hydrophobicity, pK_a of the amine, and the known blocking of destruction of the drug by monoamine oxidase. These ambiguities can be removed only by a study of more varied structures. The predicted activity of each drug by the PLS model and by equation *(3)* is given in Table 6.

Table 6. Comparison of observed activity for each drug with activity calculated by PLS and multiple regression equation *(3)*

Drug No.	Observed activity	Calculated activity PLS	Calculated activity Equation *(3)*	Drug No.	Observed activity	Calculated activity PLS	Calculated activity Equation *(3)*
(1)	5	6	4	(9)	5	5	5
(2)	5	6	6	(10)	40	48	32
(3)	50	43	53	(11)	40	44	37
(4)	50	42	75	(12)	40	47	30
(5)	15	14	17	(13)	20	44	25
(6)	100	95	90	(14)	80	34	45
(7)	8	7	7	(15)	14	6	13
(8)	5	4	5	(16)	13	32	23

There was no satisfactory three-descriptor regression. Although all the terms of equations *(1)–(3)* are statistically significant at the 95% level, equation *(3)* contains four descriptors for 16 points, and thus must be considered borderline. Its standard error of estimate corresponds to 37% accuracy in determining the effective dose, a value which approaches the estimated accuracy of determination of activity. The amount of collinearity in this equation is small, as may be seen from Fig. 2. As a predictor, however, the PLS result is probably safer because of the larger range of properties that it covers.

No significant regression was obtained with the hydrophobicity term, or with the HOMO and LUMO energies separately. These energies together however gave highly significant regressions, with nearly equal coefficients, of opposite sign.

Discussion

By far the strongest trend in these data is the dependence of activity on the frontier orbital energy difference. This contrasts with the phenylalkylamines,[1,14] where this descriptor was less important (although still very highly significant), and hydrophobicity of the *para* substituent and the *meta–para* interaction were dominant. These last parameters cannot be tested in the present data set because of the paucity of substituents on the indole ring system. It would perhaps be unsafe to conclude on the basis of the present data set alone that the electronic effect can be attributed to the HOMO–LUMO difference, as it includes only three substitution patterns: unsubstituted, 4-oxygen-substituted, and 5-oxygen-substituted. However, this descriptor has been identified as important in the phenylalkylamines, in which a much wider range of substituents is present. It is also the most prominent among many calculated electronic descriptors in the case of the tryptamines. This strongly suggests that, in the case of the

tryptamines, the same identification can be made, as the two groups of drugs very probably have the same mode of action.

The prominence of the E_δ term suggests the importance of charge transfer in the formation of the drug–receptor complex. The negative sign of the coefficient of this term means that the activity of the drug increases with increasing softness of the drug, in the 'hard and soft acid and base' sense. Half of E_δ has been identified as 'hardness'.[42] The drugs considered here all have a hardness of approximately 4·1 eV, which makes them very soft in the Pearson sense, and implies that part of the receptor is also very soft, and may be an electron-deficient aromatic system, which acts as a π electron acceptor.

The theory of charge-transfer interactions suggests that, if the drug is acting as an electron donor, the energy of the HOMO should be the relevant parameter, and, if as an acceptor, the energy of the LUMO. Empirically, it is found that the difference gives a better correlation than either, and also that the difference correlates better between quantum theoretic methods than do either of the frontier orbital energies.[43] More work is needed to clarify this situation.

The importance of P_{xx} in the second PLS component, and D_t and D_y in the first PLS component and in equations *(2)* and *(3)*, suggests that electrostatic effects may also be important, and may imply that the receptor is charged, or at least strongly dipolar. A similar effect was noted with the phenylalkylamines.[1] The strongest indications are that PLS component **2** is an indication of the size, polarizability or hydrophobicity of the molecule. Since there is only a small range of substituents on the amine nitrogen (alkyl or alkenyl), it is difficult to narrow the interpretation to any single property of these substituents. Thus while it seems highly probable that factors other than electronic are operative, and while it is likely that these include either the size, polarizability or hydrophobicity of the substituent on the amine nitrogen (from Fig. 4*b*), it is not safe to decide between these properties.

By the use of multivariate statistics, it has been possible to uncover relationships that are not apparent from inspection of the data. Thus while one could readily see that the oxygen substitution pattern on the indole is important, it is far from obvious that the substituent on the amine nitrogen are also important. Thus while the activities of *N*,*N*-dimethyltryptamine (1), *N*,*N*-diethyltryptamine and *N*,*N*-dipropyltryptamine (8) are equal, *N*,*N*-diisopropyltryptamine (7) is somewhat more active, and *N*-isopropyl-*N*-methyltryptamine (15) is considerably more active. Without the statistical treatment, it would be impossible to decide whether the effect of the *N*-alkyl groups was significant. The fact that the second component is significant by cross-validation strongly suggests that it is.

It should also be noted that the substitution on the nitrogen is not without effect on the energy term E_δ, as may be seen from Table 5. This would further complicate naive attempts to discover quantitative structure–activity relationships by inspection, and is probably responsible for the slope of the lines separating the groups in Fig. 4*b*.

A quantitative structure–activity relationship for a drug determined in whole-organism experiments reflects effects other than interaction of the drug with its

[42] Parr, R. G., and Pearson, R. G., *J. Am. Chem. Soc.*, 1983, **105**, 7512.
[43] Clare, B. W., *Theor. Chim. Acta*, 1994, **87**, 415.

primary target. Thus the volume or hydrophobicity effect could be a result of effects other than hydrophobic interactions at the receptor site. Such effects may include an influence on transportation of the drug to its receptor, or on its removal from the system by monoamine oxidase. In this connection, it is relevant to note that *N*,*N*-dimethyltryptamine (1) and *N*,*N*-diethyltryptamine (2) are effective only on parenteral administration, and are of brief duration of action, about 1 h. This is not true of the α-methyl compounds and some of the *N*-isopropyl derivatives, which are orally active and last up to 12 h. It is plausible that the rapid destruction of the shorter-acting drugs does not permit the attainment of an effective concentration at the receptor when given orally. When injected, however, they are distributed within a few minutes, before significant destruction can occur. Longer-acting compounds are not so rapidly removed, and reach comparable local concentrations, whether injected or taken orally.

Andrews and Lloyd[44] have proposed a common pharmacophore for central nervous system-active drugs, based on a study of conformationally rigid substances, including LSD. This involves hydrophobic bonding of an aromatic ring in the drug to a region of the receptor, and hydrogen bonding of the amine nitrogen. The receptors for neurotransmitters have evolved over time from a common precursor, diverging to the variety of neurotransmitters known at present, and to the variety of subtypes for each neurotransmitter. The common pharmacophore has a nitrogen 0·3 Å out of the plane of the aromatic ring. This compares with 0·17 Å for LSD, and 1·5 Å for *N*,*N*-dimethyltryptamine. Given the conformational flexibility of the latter, agreement is not to be expected, and, as has been shown, *N*,*N*-dimethyltryptamine can adopt the conformation of LSD with relatively little cost in energy (Table 2). In addition to the hydrophobic bonding proposed by Andrews and Lloyd, the present and previous[14] studies suggest the involvement of an electronic factor, such as charge transfer or acid–base hardness, in the binding of the aromatic nucleus to the receptor.

Acknowledgments

I am indebted to the Research Centre for Advanced Materials and Mineral Processing for making available the Hewlett Packard computing facility, to Paula J. McLay of the Australian Government Analytical Laboratories, Cottesloe, Western Australia, for a criticism of the manuscript, and to Jennifer M. Hopwood of the Department of Mathematics, University of Western Australia, for helpful discussions on PLS.

[44] Andrews, P. R., and Lloyd, E. J., *Prog. Med. Chem.*, 1986, **23**, 91.