The Frontier Orbital Phase Angles: Novel QSAR Descriptors for Benzene Derivatives, Applied to Phenylalkylamine Hallucinogens

Brian W. Clare*

Division of Science, Murdoch University, Murdoch, W.A. 6150, Australia

Received March 9, 1998

A new empirical electronic descriptor, obtained from a molecular orbital calculation and applicable to benzene derivatives, is proposed. It is shown that this descriptor, the frontier orbital phase angle, correlates very strongly with the pharmacological activity in humans of a large series of hallucinogenic phenethylamines. In the largest QSAR study on such hallucinogens yet reported, it is demonstrated that the phase of mixing of degenerate frontier orbitals of benzene to form the frontier orbitals of the drug results in the best electronic descriptor yet found for hallucinogenic activity in phenylalkylamines.

Introduction

In the course of investigations of the quantitative structure−activity relationships (QSAR) of phenylalkylamine hallucinogen activity with data produced by quantum chemical calculations[1] on the drug molecules, schematic diagrams of the frontier orbitals of the molecules were made. Figure 1 shows a pair of such diagrams for two hallucinogens: mescaline, a drug of relatively low activity, and the much more potent drug known as DOM. A plot is made for both the highest occupied (HOMO) and lowest unoccupied (LUMO) molecular orbitals by running a molecular orbital calculation with the molecule oriented with the *Z* axis normal to the plane of the benzene ring. The molecule is then sketched in the *X*−*Y* plane, and circles are drawn on each atom, with a radius proportional to the coefficient of the atomic p_z orbital in the molecular orbital expansion. The circles are filled or not, depending on the sign of the coefficient.

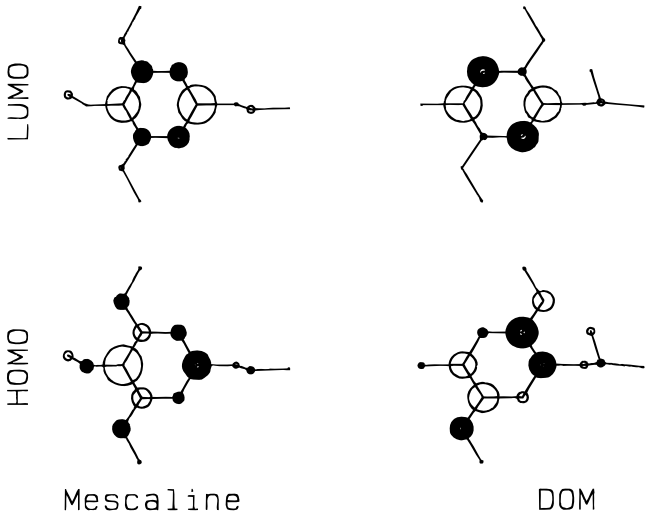

Figure 1. Frontier orbital plots for mescaline and DOM.

It can be seen that the plots look rather different, and it was observed[2] that this difference carried over to other drugs. To some degree, highly active drugs gave plots such as those for DOM, while drugs of low activity tended to give plots looking more like those for mescaline. This paper represents an attempt to quantify and analyze these differences in a much larger series. The results may well be applicable to other series of drugs based on benzene.

Hallucinogens of the phenethylamine, tryptamine, and ergoline classes are substances that produce characteristic changes in human consciousness, giving rise to a variety of abnormal phenomena.[3,4] While the three classes as a whole produce broadly similar effects, there are individual differences between members, and the effects can be studied satisfactorily only in humans. The drugs are of great importance for the light they throw on normal and abnormal mental function, and they are used in pharmacological preparations for mass screening of new drugs for antipsychotic properties. Unfortunately, because of perceptions of abuse, studies in humans have been totally discontinued, leaving only a relatively small amount of work done in animals using the drug discrimination paradigm.

It is obviously difficult to quantify the hallucinogenic effects of drugs from the subjective reports of individuals. Additionally, variation between individuals is significant, and it is more difficult to compare results for two different people than for two drugs in the same person. As these subjective effects are the primary concern, there is no way around this problem. Even if a quantifiable effect were found that correlated with the hallucinogenic effect in humans, in either humans or other species, it could only be used as long as the correlation held. Thus all of the data, however obtained, must ultimately be referred back to subjective reports, obtained from humans. Limitations on the quantifiability of the hallucinogen experience itself are acknowledged, but it is possible to estimate equivalent doses. Even at best, effective doses are not comparable between species.

A primary concern in this work is standardization. For the purposes of this study, it was decided to rely entirely on the very extensive human data collected by

* Also Department of Chemistry, The University of Western Australia, Nedlands, W.A. 6009, Australia.

S0022-2623(98)00144-7 CCC: $15.00
Published on Web 09/01/1998

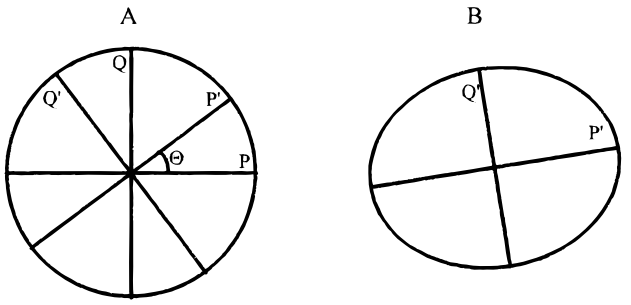

Figure 2. Representation of degenerate orbital pair showing phase angle Θ.

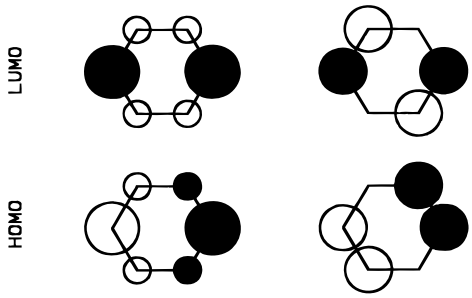

Figure 3. Degenerate HOMO and LUMO frontier orbital plots of benzene, P-like (left) and Q-like (right).

A. T. Shulgin and co-workers.[5,6] This includes data on nearly every known phenylalkylamine hallucinogen. As the drugs were tried and compared by the same small group of experienced observers, the data on different drugs are as comparable as it is possible for them to be. The data itself consists of effective dose: the mean of the threshold dose and dose required to produce the full effect of the drug. Limitations on the quality of both the data itself and its correlation with physical quantities may be assessed solely on the basis of the statistics of the data fitting.

Theory

A standard self-consistent field (SCF) calculation on benzene by any ab initio or semiempirical method gives six orbitals of π symmetry.[7] Two of these are degenerate and constitute the HOMO. Another pair form the degenerate LUMO. The HOMO and LUMO together comprise the frontier orbitals of the molecule. When the benzene molecule is oriented with its plane normal to the Z axis, the π orbitals take a particularly simple form. For all six π orbitals, the coefficients of the carbon $2p_x$ and $2p_y$ orbitals and the carbon 2s and hydrogen 1s orbitals are zero, and the orbitals may be represented by six-component eigenvectors of a matrix, giving the coefficients of the carbon $2p_z$ orbitals. They may be regarded as vectors in a six-dimensional space. Being orthogonal, they may be pictured as the mutually perpendicular axes of an ellipsoid, the lengths of which are proportional to the eigenvalues (i.e., the orbital energies). When two orbital energies are degenerate, one mutually perpendicular pair of equal axes defines a circular cross section, as in Figure 2A. Then any other mutually perpendicular pair of radii of this circle, such as $\mathbf{P}'$ and $\mathbf{Q}'$, that are related to $\mathbf{P}$ and $\mathbf{Q}$ by rotation through the angle Θ are then an equally acceptable pair of orbitals. For both the degenerate HOMO and LUMO orbitals of benzene, two pairs of orbitals take on a particularly symmetrical form. These will be designated as follows: $\mathbf{P_H}$ = (1/2$\sqrt{3}$, −1/2$\sqrt{3}$, −1/$\sqrt{3}$, −1/2$\sqrt{3}$, 1/2$\sqrt{3}$, 1/$\sqrt{3}$) and $\mathbf{Q_H}$ = (1/2, 1/2, 0, −1/2, −1/2, 0) as the components of the HOMO, and $\mathbf{P_L}$ = (−1/2$\sqrt{3}$, −1/2$\sqrt{3}$, 1/$\sqrt{3}$, −1/2$\sqrt{3}$, −1/2$\sqrt{3}$, 1/$\sqrt{3}$) and $\mathbf{Q_L}$ = (1/2, −1/2, 0, 1/2, −1/2, 0) as the components of the LUMO.

These orbitals are plotted in Figure 3 according to the scheme described above for mescaline and DOM. At first sight, it may appear that the carbon atoms of the benzene ring are not equivalent. It must be realized, however, as stated above, that because the orbitals are degenerate, any pair of orthogonal linear combinations of them is an equally satisfactory pair of orbitals. The orbitals themselves are precisely defined, even in the most approximate calculation.

Any linear combination of the HOMO pair, $C_p\mathbf{P_H} + C_q\mathbf{Q_H}$, is an equally valid HOMO, where C_p and C_q are any constants. It is convenient to normalize C_p and C_q so that $C_p^2 + C_q^2 = 1$. This leads to the introduction of the angular parameter Θ_H, so that the HOMO becomes $\cos\Theta_H\mathbf{P_H} + \sin\Theta_H\mathbf{Q_H}$, and the other orthogonal linear component is $-\sin\Theta_H\mathbf{P_H} + \cos\Theta_H\mathbf{Q_H}$. If Θ_H is zero, the first of these combinations is simply $\mathbf{P_H}$ and will be referred to as P-like, and if Θ_H is 90° it is $\mathbf{Q_H}$, and will be referred to as Q-like. The angular parameter Θ_H describes the phase of the degenerate HOMO orbitals as they mix. While the orbitals are degenerate, Θ_H has no physical significance. When the degeneracy is lifted, however, it describes the orientation, shape, and symmetry of the resulting orbital, such as that shown in Figure 2B. A similar argument may be applied to the LUMO. The orientations of $\mathbf{P}$ and $\mathbf{Q}$ for benzene given above, and hence the zeros of Θ_H and Θ_L were those obtained in a particular AM1 calculation for benzene and are arbitrary, and any other orientation could equally have been chosen as the reference. Where $\mathbf{P}$, $\mathbf{Q}$, and Θ are used without subscripts, they may be taken to refer to both HOMO and LUMO.

A plot of the first of these linear combinations for the HOMO and LUMO of benzene is shown in Figure 4, parts A and B, for various values of Θ_H and Θ_L. In all of these plots, the rightmost atom is atom 1 and the atoms are numbered counterclockwise. Collectively, Θ_H and Θ_L will be referred to as the frontier orbital phase angles.

From Figure 4 it can be seen that with increasing Θ the orbital diagram for both HOMO and LUMO goes from P-like at zero, through transitional forms at 15° to Q-like at 30°, through transitional forms back to P-like again at 60°, but with a 60° rotation. This then repeats with rotation every 60°. Thus angles of $\Theta + 60n°$ are equivalent to Θ except for orientation, for n = 0, 1, or 2. For $n = 3$, the plot returns to that for $n = 0$, with a change of sign. Because an eigenvector multiplied by any constant, including −1, is still an eigenvector, this change of sign has no physical significance. While Θ should not be thought of as an angle in the three-dimensional space of the molecule, its consequences in that space are apparent from the above.

For benzene itself, all of these combinations are completely equivalent. Assuming that the orbitals of a substituted benzene are perturbed only slightly from those of the parent compound, when benzene is substituted, the degeneracy of the HOMO and LUMO is

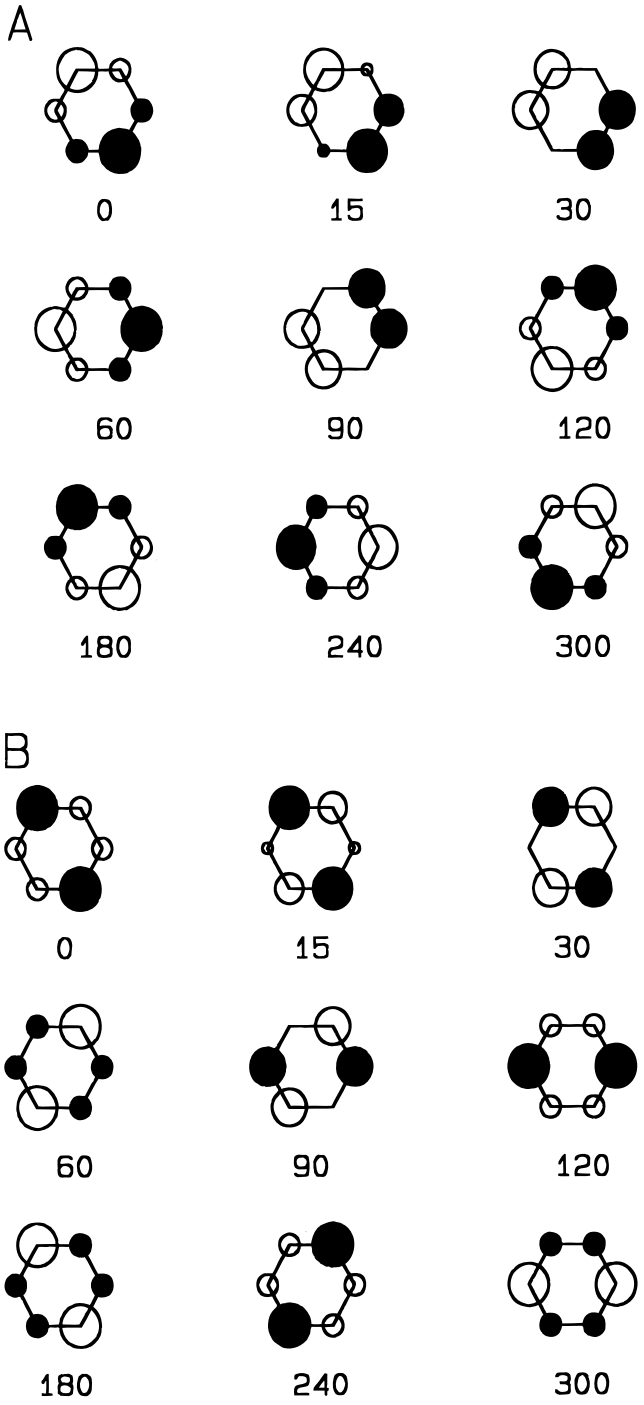

Figure 4. Degenerate frontier orbital plots of benzene, mixed according to P cos Θ + Q sin Θ, as a function of Θ: (A) HOMO, (B) LUMO.

broken, and the shapes formed by the orbitals shown diagrammatically in Figure 4 take on physical reality, and C_p and C_q, and so also Θ_H and Θ_L, are frozen to particular values, as shown in Figure 2B. As descriptors of the shape of the mixed orbitals, we may use sin 6Θ and cos 6Θ. This will allow a maximum of activity to occur at any value of Θ between 0° and 60°, 60° and 120°, or 120° and 180°. Their orientation may be described by ϑ, the nearest integer to Θ/60, if Θ is expressed in degrees. This procedure allows the shape and the orientation of the orbitals to be accounted for independently.

Even though Θ is not an angle in the three-dimensional space of the molecule, it does have effects in that space. The hexagons of Figure 4 may be thought of as representing the benzene rings of the drug molecules of this study, with the ethylamine residues bonded to the far right apexes. The angle Θ then describes both the shape (P-like or Q-like) and the orientation of the orbital.

The diagrams shown so far are schematic representations of the full three-dimensional orbitals that determine the reactivity of the molecule. The three-dimensional forms of the HOMO and LUMO orbitals of mescaline and DOM as displayed by the program Hyperchem[8] are shown in Figure 5, which may be compared with the schematic form in Figure 1.

Of course, in an arbitrary substituted benzene the symmetry of the diagrams seen in Figure 4 need not apply, and also there will be finite coefficients for p_z orbitals on the substituent atoms as well as the atoms of the ring. This is readily apparent in Figure 1. However, restricting ourselves to the common atoms of the series, those of the benzene ring, as is customary in QSAR, we can create plots similar to those of Figure 4, from the six coefficients of the carbon p_z orbitals, $C_1 - C_6$, in both the HOMO and LUMO. An empirical least squares Θ_H and Θ_L may be obtained by varying the parameter θ to minimize the function $\alpha = \Sigma(C_i - P_i \cos\theta - Q_i \sin\theta)^2$ for both the HOMO and LUMO. This least squares problem may be solved analytically, using the symbolic mathematics package Mathematica.[9]

To calculate Θ_H and Θ_L from the orbital coefficients C_1 to C_6, for the HOMO and LUMO orbitals, proceed as follows. For both HOMO and LUMO put $C = 1 + \Sigma C_i^2$ and, for the HOMO put $A = -C_1 - C_2 + C_4 + C_5$ and $B = -(C_1 - C_2 - 2C_3 - C_4 + C_5 + 2C_6)/\sqrt{3}$, and for the LUMO, put $A = -C_1 + C_2 - C_4 + C_5$ and $B = (C_1 + C_2 - 2C_3 + C_4 + C_5 - 2C_6)/\sqrt{3}$. The sum of squares to be minimized is then $C + A \sin\theta + B \cos\theta$. This expression can be converted by trigonometric identities to the form $C + L \sin(\theta+d)$, where $d = \tan^{-1}(B/A)$ and $L = A/\cos d = B/\sin d$. The value of θ that minimizes this function is obvious, and is denoted Θ_H for the HOMO, and Θ_L for the LUMO. The value of the objective function is given by $\alpha = C - A/\cos d = C - B/\sin d$, and varies from zero for a perfect fit to 2 for the poorest possible fit. As may be seen from Figure 4 angles of Θ greater than 180° result in duplicating the values from 0° to 180° with reversal of sign. For the reasons given above this change of sign has no physical significance and Θ may be regarded as defined in the range 0° to 180°. A PC program that implements this computational procedure is available by anonymous FTP.[10]

For both the HOMO and LUMO periodic functions sin 6Θ and cos 6Θ were constructed to give maxima or minima at the ring carbon atom positions, thus describing the orbital plot itself, and ϑ, defined as the nearest integer to Θ/60 (Θ in degrees) to allow for effects of orientation of the orbital plot with respect to the ethylamine side chain. The parameter ϑ takes integer values from 0 to 3, with 0 and 3 being equivalent.

As in the earlier study[1] an indicator variable was defined to allow for the effect of substitution in the α-position of the ethylamine side chain. I_{Me} has the value unity if the α-position of the ethylamine moiety

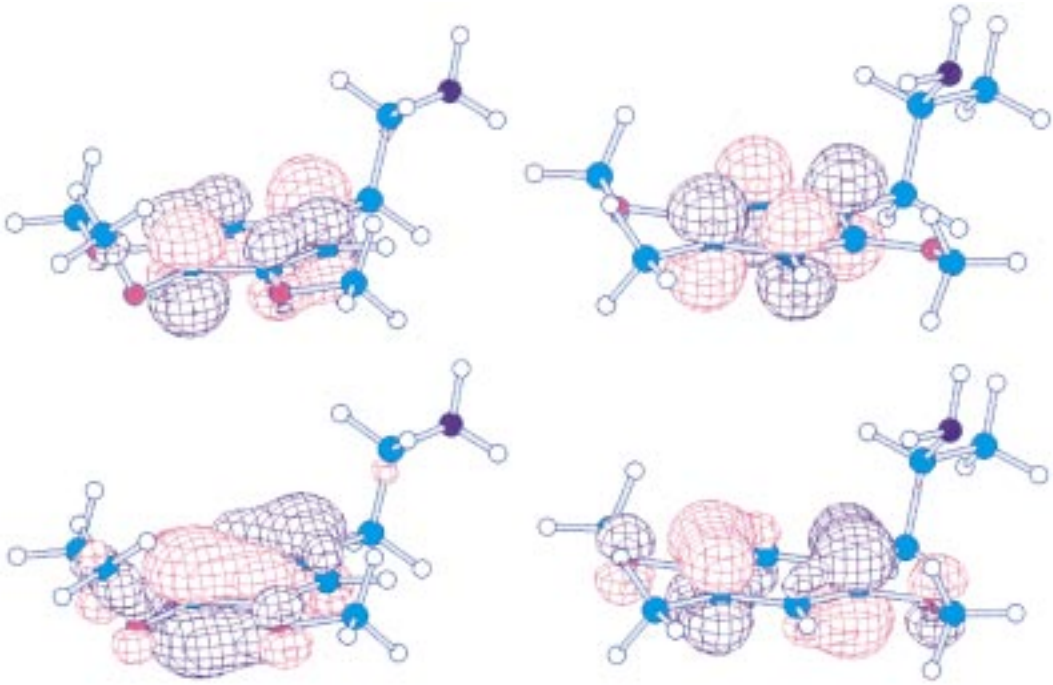

Figure 5. Three-dimensional form of the HOMO and LUMO orbitals of mescaline and DOM: (left) mescaline; (right) DOM; (top) LUMO; (bottom) HOMO. Isosurfaces include 95% of electron density, with red lines for positive phase, violet lines for negative.

is substituted with one or two alkyl groups and zero otherwise. Alkyl substitution on the α-position of a phenethylamine has little effect on the electronic structure of the benzene ring. Methyl substitution is known to potentiate the drugs, probably by protecting them from destruction by monoamine oxidase. Other alkyl groups are less effective. The only other substituent that occurred on the α-position in this group of drugs, methoxy, does not have this effect, and I_{Me} was set to zero in this case.

Calculations

All calculations were done on a Toshiba 4900 PC. The MOPAC programs were compiled with the Lahey F77L3-EM32 FORTRAN compiler.

The formulas and effective doses of the 92 hallucinogens listed in Table 1 were obtained from the literature.[6] Only those compounds that were studied by Shulgin's group were selected, and then only those for which a definite range of activities (a threshold and effective dose) were given. Even so, the group includes nearly every known phenylalkylamine hallucinogen. The mean of these two dosages, divided into the corresponding number for mescaline (taken as 300 mg), was the activity in mescaline units. These activities are given in Table 1. Mescaline units have long been a standard of activity in hallucinogenic drug work. The common logarithm of this value was used in the regressions. Several compound studied in the previous work[1] were dropped as being doubtful by the above criteria, and the activities of several other compounds were changed. Some 60 new compounds, not studied in the previous work, were added. The group of drugs now also includes a number of structural features, such as N-alkylation, N-hydroxylation, α,α-disubstitution, α-ethylation, and β-oxygen substitution, that were not present in the drugs of the original study.

The atomic coordinates of the molecules were calculated using the program PCMODEL,[11] and their geometries were optimized by the molecular mechanics routine MMX, which is part of PCMODEL. They were then further optimized with the program MOPAC 6[12] using the AM1 Hamiltonian.[13] They were defined with the benzene ring carbon atoms numbered anticlockwise from 1 to 6, and with the ethylamine side chain attached to atom 1. The origin was at the midpoint of atoms 1 and 4, with the X axis passing through atom 1 and with the midpoint of atoms 2 and 3 in the $X-Y$ plane. A single point calculation was then done with the program MOPAC 93,[14] using the keywords VECTORS to print out the eigenvectors and EPS=72.7 to take into account solvation in an aqueous medium by the COSMO[15] model. From the MOPAC 93 output the HOMO, SHOMO (second highest occupied molecular orbital), LUMO, and SLUMO (second lowest unoccupied molecular orbital) energies were recorded. The coefficients of the p_z atomic orbitals on the carbon atoms of the benzene ring of the phenethylamine moiety were used to calculate the phase angles Θ_H and Θ_L and the minimized objective function values α_H and α_L.

The log activity data was then regressed on the parameters η [$=(E_{LUMO} - E_{HOMO})/2$], I_{Me}, Δ_H [$=E_{HOMO} - E_{SHOMO}$], Δ_L [$= E_{SLUMO} - E_{LUMO}$], α_H, α_L, sin $6\Theta_H$, sin $6\Theta_L$, cos $6\Theta_H$, cos $6\Theta_L$, ϑ_H, and ϑ_L. The use of

Table 1. Identities of the Drugs and Their Activities *A* in Mescaline Units

no.	designation	chemical name	*A*
1	mescaline	3,4,5-trimethoxyphenethylamine	1
2	ME	3-ethoxy-4,5-dimethoxyphenethylamine	1
3	E	3,5-dimethoxy-4-ethoxyphenethylamine	6
4	P	3,5-dimethoxy-4-propoxyphenethylamine	7
5	ASB	3,4-diethoxy-5-methoxyphenethylamine	1.3
6	2C-E	2,5-dimethoxy-4-ethylphenethylamine	18
7	2C-D	2,5-dimethoxy-4-methylphenethylamine	8
8	3-TM	3,4-dimethoxy-5-methylthiophenethylamine	4
9	TM	3,5-dimethoxy-4-methylthiophenethylamine	10
10	3-TME	3,4-dimethoxy-5-ethylthiophenethylamine	4
11	4-TME	3-ethoxy-4-methylthio-5-methoxyphenethylamine	4
12	3-TE	3-methoxy-4-ethoxy-5-methylthiophenethylamine	4
13	4-TE	3,5-dimethoxy-4-ethylthiophenethylamine	12
14	3-TASB	3-ethylthio-4-ethoxy-5-methoxyphenethylamine	2
15	4-TASB	3-ethoxy-4-ethylthio-5-methoxyphenethylamine	4
16	5-TASB	3,4-diethoxy-5-methylthiophenethylamine	2
17	TP	3,5-dimethoxy-4-propylthiophenethylamine	16
18	TB	3,5-dimethoxy-4-butylthiophenethylamine	3
19	2C-I	2,5-dimethoxy-4-iodophenethylamine	17
20	4-MA	4-methoxyamphetamine	5
21	2,5-DMA	2,5-dimethoxyamphetamine	2.5
22	TMA	3,4,5-trimethoxyamphetamine	2
23	TMA-2	2,4,5-trimethoxyamphetamine	10
24	TMA-4	2,3,5-trimethoxyamphetamine	4
25	TMA-5	2,3,6-trimethoxyamphetamine	10
26	TMA-6	2,4,6-trimethoxyamphetamine	10
27	MEM	2,5-dimethoxy-4-ethoxyamphetamine	9
28	3C-BZ	3,5-dimethoxy-4-benzyloxyamphetamine	3
29	TA	2,3,4,5-tetramethoxyamphetamine	6
30	MDA	3,4-methylenedioxyamphetamine	3
31	MMDA	3-methoxy-4,5-methylenedioxyamphetamine	2
32	MMDA-2	2-methoxy-4,5-methylenedioxyamphetamine	8
33	MMDA-3A	2-methoxy-3,4-methylenedioxyamphetamine	6
34	DMMDA	2,5-dimethoxy-3,4-methylenedioxyamphetamine	6
35	DOM	2,5-dimethoxy-4-methylamphetamine	50
36	DOET	2,5-dimethoxy-4-ethylamphetamine	75
37	DOPR	2,5-dimethoxy-4-(n)-propylamphetamine	80
38	PARADOT	2,5-dimethoxy-4-methylthioamphetamine	40
39	ALEPH-4	2,5-dimethoxy-4-(i)-propylthioamphetamine	32
40	DOB	2,5-dimethoxy-4-bromoamphetamine	150
41	DOI	2,5-dimethoxy-4-iodoamphetamine	133
42	DOC	2,5-dimethoxy-4-chloroamphetamine	133
43	DON	2,5-dimethoxy-4-nitroamphetamine	80
44	2C-G	2,5-dimethoxy-3,4-dimethylphenethylamine	11
45	2C-G-3	5-(2-aminoethyl)-4,7-dimethoxyindane	14
46	2C-G-N	1,4-dimethoxynaphth-2-ylethylamine	10
47	2C-G-5	3,6-dimethoxy-4-(2-aminoethyl)benzonorbornane	22
48	2C-T-F	2,5-dimethoxy-4-(2-fluoroethylthio)phenethylamine	29
49	2C-T-13	2,5-dimethoxy-4-(2-methoxyethylthio)phenethylamine	9
50	AL	3,5-dimethoxy-4-allyloxyamphetamine	11
51	ALEPH-2	2,5-dimethoxy-4-ethylthioamphetamine	50
52	2C-B	2,5-dimethoxy-4-bromophenethylamine	16
53	2C-C	2,5-dimethoxy-4-chlorophenethylamine	10
54	CPM	3,5-dimethoxy-4-cyclopropylmethoxyphenethylamine	4
55	Ψ-DOM	2,6-dimethoxy-4-methylamphetamine	15
56	IP	3,5-dimethoxy-4-(i)-propoxyphenethylamine	5
57	BOD	4-methyl-2,5,*β*-trimethoxyphenethylamine	15
58	BOH	*β*-methoxy-3,4-methylenedioxyphenethylamine	3
59	4-BR-3,5-DMA	3,5-dimethoxy-4-bromoamphetamine	43
60	2C-P	2,5-dimethoxy-4-*n*-propylphenethylamine	37
61	2C-T	2,5-dimethoxy-4-methylthiophenethylamine	4
62	2C-T-2	2,5-dimethoxy-4-ethylthiophenethylamine	16
63	2C-T-4	2,5-dimethoxy-4-isopropylthiophenethylamine	21
64	2C-T-7	2,5-dimethoxy-4-*n*-propylthiophenethylamine	15
65	2C-T-8	2,5-dimethoxy-4-cyclopropylmethylthiophenethylamine	7
66	2C-T-9	2,5-dimethoxy-4-*tert*-butylthiophenethylamine	4
67	2C-T-17	2,5-dimethoxy-4-*sec*-butylthiophenethylamine	4
68	DMCPA	2-(2,5-dimethoxy-4-methylphenyl)cyclopropylamine	17
69	DOEF	2,5-dimethoxy-4-(2-fluoroethyl)amphetamine	110
70	FLEA	*N*-Hydroxy-*N*-methyl-3,4-methylenedioxyamphetamine	2.5
71	G-3	2,5-dimethoxy-3,4-trimethyleneamphetamine	20
72	G-5	3,6-dimethoxy-4-(2-aminopropyl)benzonorbornane	18
73	G	2,5-dimethoxy-3,4-dimethylamphetamine	12
74	HOT-2	2,5-dimethoxy-4-ethylthio-*N*-hydroxyphenethylamine	22
75	HOT-7	2,5-dimethoxy-*N*-hydroxy-4-*n*-propylthiophenethylamine	15
76	HOT-17	2,5-dimethoxy-4-*sec*-butylthio-*N*-hydroxyphenethylamine	3

Table 1 (Continued)

no.	designation	chemical name	*A*
77	J	2-amino-1-(3,4-methylenedioxyphenyl)butane	1.5
78	MDE	3,4-methylenedioxy-*N*-ethylamphetamine	2
79	MDOH	3,4-methylenedioxy-*N*-hydroxyamphetamine	2.3
80	MDPH	α,α-dimethyl-3,4-methylenedioxyphenethylamine	1.5
81	Methyl-J	2-methylamino-1-(3,4-methylenedioxyphenyl)butane	1.5
82	5-TOET	4-ethyl-2-methoxy-5-methylthioamphetamine	15
83	2-TOM	5-methoxy-4-methyl-2-methylthioamphetamine	4
84	5-TOM	2-methoxy-4-methyl-5-methylthioamphetamine	7
85	ALEPH-7	4-propylthio-2,5-dimethoxyamphetamine	55
86	BOB	2,5,*β*-trimethoxy-4-bromophenethylamine	20
87	3C-E	3,5-dimethoxy-4-ethoxyamphetamine	7
88	2C-T-13	2,5-dimethoxy-4-(2-methoxyethylthio)phenethylamine	9
89	2C-T-21	2,5-dimethoxy-4-(2-fluoroethylthio)phenethylamine	30
90	DESOXY	3,5-dimethoxy-4-methylphenethylamine	4
91	MAL	3,5-dimethoxy-4-methallyloxyphenethylamine	6
92	META-DOB	2,4-dimethoxy-5-bromoamphetamine	4

functions of 6Θ and the ϑ as variables allows separation of the effects of orientation and shape of the orbital combination. This will be referred to as the larger data set. Regression was also carried out using sin $2\Theta_H$, cos $2\Theta_H$, sin $2\Theta_L$, and cos $2\Theta_L$ in place of sin $6\Theta_H$, cos $6\Theta_H$, sin $6\Theta_L$, cos $6\Theta_L$, ϑ_H, and ϑ_L. This will be referred to as the smaller data set.

The regression was done using the all possible subsets module 9R of the BMDP statistical package.[16] The primary purpose of this study is to investigate the phase angles Θ_H and Θ_L. If these two are significantly involved, then obviously the degrees of splitting of the degeneracies, measured by Δ_H and Δ_L are also of interest. So also are the response surface minimum values, α_H and α_L which are a measure of lack of fit of the actual orbitals to the least squares linear combination, due either to asymmetry of the orbitals or to large coefficients on nonring atoms. A principal components calculation was done to determine whether the resulting model was stable to collinearity, and a randomization test[17,18] to determine whether chance correlation was likely.

Results

Initial analyses identified four serious outliers in the data. These were DON, **43**, and 2C-G-N, **46**, which had values of α_L much higher than the remainder of the compounds, and 4-BR-3,5-DMA, **59**, and especially META-DOB, **92**, which were identified by the diagnostics of the BMDP package as points of high leverage. DON, **43**, also had an isolated, very high value of Δ_L. In the regressions with the larger data set, the all possible subsets[19] results indicated that sin $6\Theta_L$ was the best single descriptor, explaining 16% of the variance in log *A*. It was followed by Δ_L (13%) and Δ_H (13%). The known important descriptors η and I_{Me} explained by themselves only 8% and 3%, respectively. When two variables were allowed into the equation, sin $6\Theta_L$ entered along with Δ_L, explaining together 29% of the variance in log *A*. When the smaller set was used, cos $2\Theta_L$ alone explained 22% of the variance, followed by Δ_L (13%) and Δ_H (13%), as in the larger regression. Again, η and I_{Me} explained by themselves 8% and 3%.

With the smaller data set, the best equation obtained was:

$$\log A = C_1 I_{Me} + C_2 \Delta_L + C_3 \cos 2\Theta_H + C_4 \sin 2\Theta_H + C_5 \sin 2\Theta_L + C_6 \quad (1)$$

	C_1	C_2	C_3	C_4	C_5	C_6
C	0.362	1.422	−0.250	−0.324	−0.363	0.011
se	0.078	0.256	0.084	0.066	0.099	0.125
α	0.00001	0.00000	0.00372	0.00000	0.00042	0.927

$$N = 92,\ R^2 = 0.584,\ Q^2 = 0.526,\ s = 0.34,\ F = 24.2,\ P = 4 \times 10^{-15},\ \Lambda = 1.28$$

Here *N* is the number of drugs in the regression, R^2 the squared multiple correlation coefficient, Q^2 the same based on the "leave one out" technique (i.e., on the predicted residuals), *s* the standard error of estimate, and *F* the Fisher variance ratio. *P* is the probability based on *F*; the *se* are the standard errors of estimate of the parameters and α the corresponding probabilities of obtaining the observed values if the true value were zero, as determined by a *t* test. Thus α is the statistical significance of the variable in the presence of the other variables: a value of α greater than 0.05 indicates that the variable does not contribute significantly to the regression.[20] The statistic Λ, defined as

$$\Lambda = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{\lambda_i}$$

where *n* is the number of variables and the λ_i are the eigenvalues of the correlation matrix of independent variables,[21] measures the amount of collinearity in the equation. Values of Λ exceeding 5.0 indicate unacceptable collinearity. Equation 1 leads to a maximum of activity at $\Theta_H = 116.2°$, $\Theta_L = 135°$, $I_{Me} = 1$ and large Δ_L, and to a minimum at $\Theta_H = 26.2°$, $\Theta_L = 45°$, $I_{Me} = 0$ and small Δ_L.

The best regression equation chosen based on Mallow's C_p[22] for the larger data set was the following:

$$\log A = C_1 I_{Me} + C_2 \Delta_L + C_3 \vartheta_H + C_4 \vartheta_L + C_5 \cos 6\Theta_H + C_6 \sin 6\Theta_H + C_7 \sin 6\Theta_L + C_8 \quad (2)$$

	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
C	0.368	1.661	0.432	0.376	0.189	0.226	0.478	−1.183
se	0.076	0.275	0.077	0.114	0.065	0.083	0.092	0.285
α	0.00001	0.00000	0.00000	0.00150	0.00473	0.00810	0.00000	0.00008

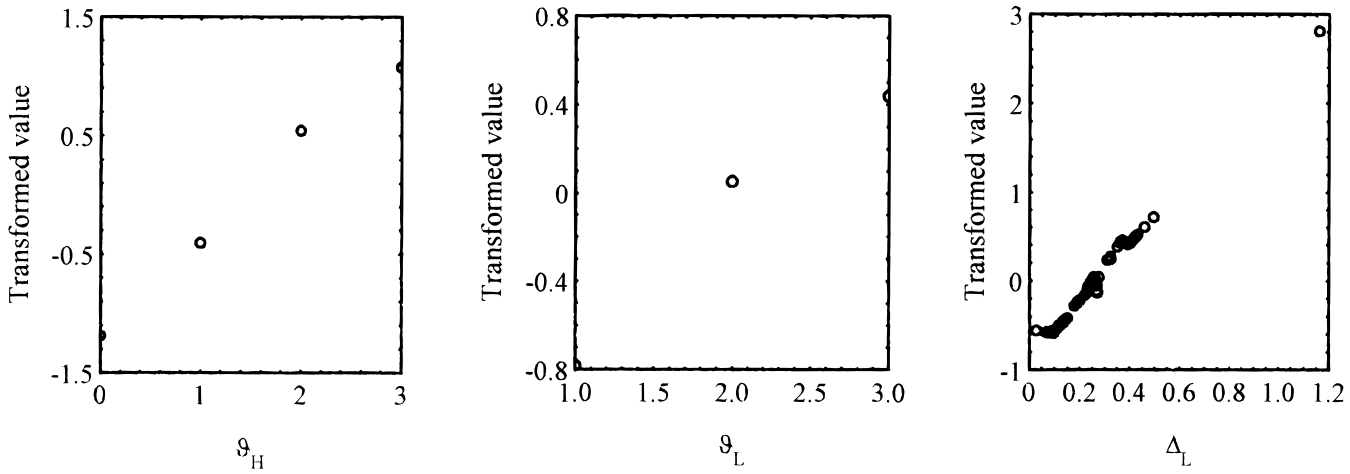

Figure 6. ACE transformation plots of ϑ_H, ϑ_L, and Δ_L for the variables of eq 1.

$$N = 92,\ R^2 = 0.613,\ Q^2 = 0.534,\ s = 0.28,$$
$$F = 19.0,\ P = 6 \times 10^{-15},\ \Lambda = 1.62$$

The same equation was obtained by stepwise regression be the Efroymson algorithm.[23]

Any procedure of selection of a subset of variables from a larger pool is subject to the inflation of statistical significance, because the standard F test assumes a single trial. For example, if a one-variable model is chosen from a pool of 100 variables, to obtain a true significance level α it is necessary to set a level $1 - (1 - \alpha')^{1/100}$, where α' is the value for a single model, to allow for the fact that there have been 100 trials, not just one. If one is selecting more than one variable in the model, the problem of what α value to use is too complex for analytical solution, but has been studied numerically by Topliss et al.[24,25] These authors recommended substitution of random numbers for the dependent variable and the carrying out of multiple subset selections on the resulting data.

Since the present results were obtained by selection of a subset of variables from a pool, it is necessary to confirm that the results are really statistically significant. The F test value of 6×10^{-15} is certainly too optimistic. The procedure of repeated applications of randomly reassigning the dependent variable, followed by stepwise regression was applied. In 10 000 such trials, a maximum R^2 of 0.308 was obtained. This does not approach the value of 0.613 in eq 2, and the estimated statistical significance from the reassignment procedure based on the distribution of R was 5×10^{-10}. This is a realistic estimate of the probability of obtaining such a correlation by chance, in the absence of a correlation between hallucinogenic activity and the variables in the equation. Thus, it is effectively certain that the nonzero correlation attained here for dependence of activity on the electronic descriptors did not arise by chance.

There is no a priori reason to believe that dependence of log A on these parameters should be linear. This is particularly true for the discrete variables ϑ_H and ϑ_L. To test this, the alternating conditional expectation (ACE) technique[26] of Breiman and Friedman was applied, using ordinal transformations for Δ_L, ϑ_H, and ϑ_L only. The resulting transformation plots are shown in Figure 6. It may be seen that for all three descriptors there is little deviation from linearity, and this is confirmed by the R^2 value obtained from the ACE calculation, 0.635, which is little better than that from the linear regression.

Equation 2 leads to maximum activity for $\Theta_L = 60n + 15°$ and $\Theta_H = 60n + 8.3°$ ($n = 0$, 1, or 2), and for $\vartheta_H = \vartheta_L = 3$. The activity predicted from $\vartheta_H = 0$ or $\vartheta_L = 0$ must be identical with that for $\vartheta_H = 3$ or $\vartheta_L = 3$, because reversal of sign has no physical significance. No compound has $\vartheta_L = 0$. Only two compound, DON, **43**, and 3-TASB, **14**, have $\vartheta_H = 0$, and thus the apparent violation of symmetry in the ACE plot for ϑ_H is an anomaly of the smoothing process, which does not enforce the necessary symmetry.

Using eq 2 the activities of each of the compounds was predicted using the data for the other compounds. The results are presented as residuals in the last column of Table 2[6]. Four compounds were poorly predicted, with residuals greater than 0.6 (i.e., more than a factor of 4 in activity). The worst two were 2,5-DMA, **21**, and meta-DOB, **92**, with residuals of −1.11 and 0.97, respectively. Meta-DOB, **92**, was initially identified as a point of high leverage. The other three compounds identified in the preliminary testing as anomalous, DON, **43**, 2C-G-N, **46**, and 4-bromo-3,5-DMA, **59**, were well predicted. The other two poorly predicted cases were borderline, being HOT-17, **76**, and TMA-5, **25**, with residuals of −0.67 and −0.64. The reason for the poor prediction in these two cases, and 2,5-DMA, **21**, is not clear.

Discussion

Much has been made of the presumed unreliability of data gathered from subjective accounts of hallucinogenic drug experiences, in the absence of traditional "double blind" approaches of classical pharmacology, which are not feasible in this area. Shulgin has employed highly trained volunteers, with wide experience of drugs of this kind, and validation is by agreement between individuals on the relative effects of different drugs. In view of the recognized variability of the effects of hallucinogens in different individuals, the placebo effect, and also the problem of tolerance, this approach may appear to be very unpromising. The results presented here should dispel any such perception. It is clear that random observer biases could not

Table 2. Calculated Phase Angles θ_H and θ_L (deg), Splitting Energies Δ_H and Δ_L (eV), Response Function Values α_H and α_L, η Values (eV), and Cross-Validated Residuals

drug	Θ_H (deg)	Θ_L (deg)	Δ_H (eV)	Δ_L (eV)	α_H	α_L	2η (eV)	d
1	53.3	120.8	0.046	0.243	0.041	0.005	9.190	−0.47
2	39.4	120.3	0.051	0.235	0.044	0.005	9.190	−0.13
3	55.6	120.8	0.070	0.238	0.042	0.005	9.172	0.29
4	55.8	120.9	0.065	0.240	0.042	0.005	9.175	0.34
5	46.5	120.4	0.071	0.238	0.044	0.005	9.164	−0.12
6	106.7	137.6	1.060	0.233	0.030	0.001	8.780	0.17
7	105.5	136.8	1.046	0.237	0.030	0.001	8.762	−0.19
8	113.5	85.4	0.671	0.264	0.214	0.007	8.375	−0.13
9	71.4	118.6	0.920	0.370	0.307	0.010	8.209	0.19
10	113.2	84.8	0.652	0.269	0.207	0.007	8.392	−0.16
11	59.7	119.4	0.919	0.362	0.200	0.011	8.209	−0.16
12	112.1	85.4	0.673	0.271	0.214	0.007	8.359	−0.09
13	59.8	118.7	0.886	0.383	0.197	0.013	8.231	0.34
14	7.8	154.8	0.642	0.271	0.207	0.007	8.385	−0.21
15	60.0	119.9	0.890	0.374	0.199	0.013	8.224	−0.22
16	112.2	85.4	0.655	0.270	0.214	0.007	8.362	−0.44
17	59.9	118.7	0.885	0.389	0.199	0.013	8.225	0.46
18	59.9	118.7	0.884	0.390	0.199	0.013	8.225	−0.31
19	111.7	131.0	0.982	0.389	0.040	0.006	8.670	−0.25
20	57.0	121.5	0.805	0.028	0.026	0.002	9.328	0.11
21	111.0	140.1	0.973	0.179	0.031	0.000	8.895	−1.11
22	49.4	120.2	0.060	0.235	0.041	0.005	9.182	−0.37
23	96.4	150.2	1.060	0.141	0.038	0.003	8.680	−0.14
24	135.2	103.6	0.501	0.131	0.040	0.005	9.013	−0.23
25	179.1	88.0	0.751	0.219	0.036	0.003	8.781	−0.64
26	65.4	120.3	0.399	0.067	0.058	0.009	9.086	0.26
27	96.1	150.5	1.065	0.139	0.039	0.003	8.674	−0.18
28	56.9	119.3	0.054	0.270	0.043	0.009	9.150	−0.43
29	104.6	138.8	0.443	0.136	0.032	0.003	8.887	−0.47
30	36.1	124.7	0.743	0.095	0.044	0.003	9.062	0.02
31	61.8	104.3	0.392	0.142	0.054	0.003	8.978	−0.06
32	96.6	132.7	1.019	0.123	0.042	0.003	8.712	−0.29
33	38.6	133.8	0.139	0.142	0.052	0.003	9.041	0.00
34	102.0	139.9	0.308	0.200	0.040	0.002	8.724	−0.52
35	104.3	136.4	0.954	0.232	0.032	0.001	8.760	0.30
36	106.2	137.5	0.990	0.232	0.030	0.001	8.780	0.45
37	106.0	137.2	0.985	0.235	0.031	0.001	8.774	0.48
38	85.7	132.0	0.993	0.379	0.123	0.004	8.150	0.24
39	85.6	131.9	0.891	0.432	0.122	0.007	8.167	0.04
40	109.2	131.1	0.874	0.391	0.041	0.007	8.644	0.45
41	111.0	131.8	0.877	0.386	0.041	0.006	8.670	0.33
42	106.2	132.9	0.902	0.320	0.037	0.004	8.674	0.56
43	2.0	114.5	0.857	1.162	0.063	0.141	7.952	0.20
44	104.4	129.8	0.506	0.080	0.029	0.001	9.036	0.33
45	99.7	133.9	0.407	0.088	0.028	0.002	9.067	0.41
46	115.7	64.2	1.070	0.323	0.172	0.200	8.027	0.15
47	100.1	137.6	0.454	0.111	0.027	0.003	9.097	0.59
48	87.3	131.7	0.973	0.394	0.114	0.005	8.168	0.50
49	87.2	131.8	0.975	0.391	0.115	0.005	8.168	−0.03
50	58.0	119.6	0.035	0.256	0.042	0.006	9.188	0.14
51	83.1	133.5	0.965	0.250	0.131	0.004	8.133	0.49
52	110.0	131.1	0.980	0.394	0.040	0.007	8.646	−0.24
53	107.1	132.9	1.005	0.324	0.036	0.003	8.675	−0.27
54	55.6	118.8	0.076	0.236	0.043	0.005	9.165	0.21
55	61.2	120.2	0.138	0.189	0.051	0.004	9.184	0.29
56	60.2	122.1	0.108	0.236	0.045	0.005	9.132	0.01
57	109.6	132.6	1.073	0.312	0.034	0.003	8.732	−0.13
58	34.6	124.8	0.762	0.150	0.053	0.003	9.071	0.30
59	150.8	120.3	0.099	0.351	0.043	0.013	9.127	0.02
60	106.6	137.4	1.056	0.239	0.031	0.001	8.769	0.49
61	83.5	133.6	0.965	0.393	0.135	0.004	8.134	−0.56
62	83.7	133.3	0.953	0.391	0.131	0.004	8.152	0.10
63	85.6	131.3	0.886	0.415	0.124	0.007	8.186	0.27
64	83.7	133.1	0.951	0.391	0.131	0.004	8.152	0.07
65	83.7	132.7	0.946	0.394	0.132	0.005	8.151	−0.28
66	89.8	128.6	0.747	0.496	0.128	0.020	8.193	−0.44
67	86.3	130.9	0.869	0.426	0.121	0.008	8.192	−0.48
68	99.0	133.4	0.895	0.277	0.042	0.002	8.607	0.17
69	107.0	134.3	0.967	0.277	0.034	0.001	8.734	0.52
70	36.6	124.0	0.682	0.089	0.045	0.003	9.058	−0.02
71	98.6	134.2	0.383	0.087	0.028	0.002	9.065	0.19
72	99.1	137.6	0.421	0.111	0.027	0.002	9.094	0.11
73	102.8	131.0	0.475	0.091	0.030	0.001	9.014	−0.04
74	93.3	133.6	0.969	0.402	0.136	0.004	8.131	0.03
75	83.6	133.5	0.967	0.404	0.131	0.004	8.129	0.04

Table 2 (Continued)

drug	Θ_H (deg)	Θ_L (deg)	Δ_H (eV)	Δ_L (eV)	α_H	α_L	2η (eV)	d
76	83.1	131.0	0.925	0.397	0.137	0.006	8.141	−0.67
77	36.0	124.6	0.753	0.096	0.045	0.003	9.059	−0.30
78	37.0	133.1	0.335	0.086	0.059	0.004	9.047	−0.42
79	35.8	129.5	0.838	0.091	0.043	0.003	9.072	−0.29
80	36.1	123.5	0.810	0.103	0.042	0.003	9.048	−0.26
81	36.9	132.4	0.411	0.085	0.052	0.004	9.053	−0.54
82	111.1	91.9	1.270	0.118	0.133	0.006	8.308	0.36
83	110.1	94.9	1.327	0.132	0.135	0.006	8.270	−0.09
84	111.8	92.1	1.216	0.128	0.129	0.007	8.312	−0.01
85	83.1	133.4	0.966	0.397	0.132	0.004	8.131	0.26
86	113.0	130.2	0.925	0.460	0.040	0.002	8.595	−0.33
87	58.4	122.3	0.087	0.227	0.040	0.006	9.171	−0.18
88	84.0	133.5	0.962	0.406	0.125	0.005	8.138	−0.18
89	85.1	131.5	0.917	0.425	0.124	0.008	8.146	0.40
90	57.1	120.4	0.077	0.191	0.041	0.003	9.238	0.16
91	52.1	120.4	0.075	0.260	0.046	0.007	9.144	0.39
92	87.2	56.3	0.752	0.136	0.056	0.019	8.908	0.97

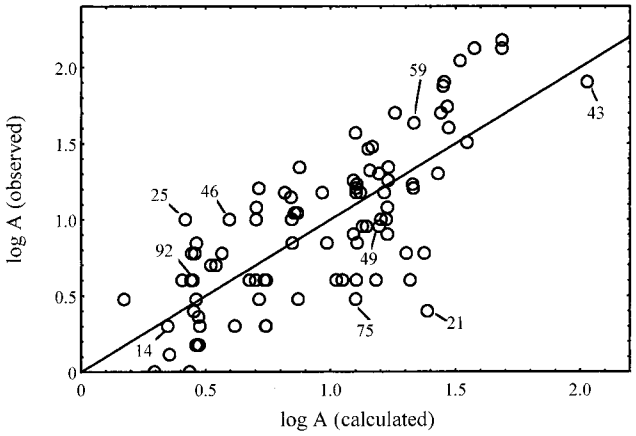

Figure 7. Plot of fitted versus observed log activity for eq 1 for the 92 drugs.

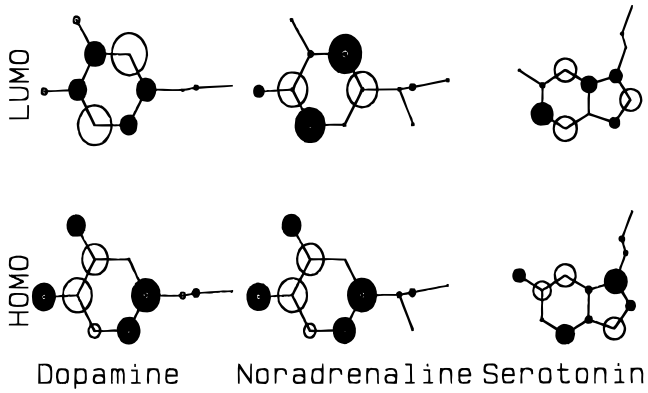

Figure 8. Frontier orbital plots for three neurotransmitters: dopamine, noradrenaline, and serotonin.

have led to drug activities that correlate so well with a theoretical index, which can only be obtained by an intricate calculation.

Hydrophobic and steric descriptors are known to be factors in explaining the activity of hallucinogens.[27−30] These properties have not been considered here because our object is not to obtain the best possible QSAR but to explore the significance of the electronic factors only. The work presented here explains up to 60% of the activity by electronic factors alone. This is in the largest group of hallucinogenic drugs ever subjected to a QSAR study, and the study yields an extremely high statistical significance. These results constitute a confirmation of both the electronic parameters as being significant, and of the validity of the data published by Shulgin's group. It also confirms the importance of frontier orbital involvement, which has long been claimed.[1,27,31]

Equation 2 explains the initial observation that led to this study. Equation 1, however, is simpler, and although not quite as good a fit, is equal to eq 2 as a predictor. Use of eq 1 shifts attention from the shape and orientation of the orbital plot to the actual value of the phase angles, and should probably be regarded as preferable. A plot of log activity, calculated by eq 1, against that observed, is given in Figure 7.

From the all-possible subsets results, there may be some correlation of activity with all of the descriptors studied. Correlation with I_{Me} is of course expected, and that with η, and also with E_H and E_L has been found in previous studies.[1,27,31] The significance of η is its relationship to the hardness of the molecule, in the Pearson sense.[32] The values of E_H and E_L can be related to the formation of charge-transfer complexes with electron donors and acceptors.[33,34] It is also noteworthy that the values of these descriptors for phenylalkylamines correlates poorly between semiempirical methods. Their difference 2η correlates much better.[35] Correlation with the Θ values is new to this work and seems to be of major importance. Given this, it is not unexpected that correlations be found with Δ_H and Δ_L, and both of these are positive (i.e., the more the degenerate levels are split, the more active the drug). Correlation also occurs with α_H, and this is negative, implying that the more the HOMO orbital departs from the symmetry of the original unperturbed benzene orbitals and the more the p_z contributions from the substituent atoms the less active the drug. Correlation with α_L is weaker, and it is positive. In general, the greater lack of fit, and hence the greater α values, occurred in the HOMO rather than the LUMO, and with S rather than O substitution. Particularly bad fit occurred with the nitro substituent of DON, **43**, and in 2C-G-N, **46**. Although the latter compound is well predicted, the applicability of the method to this, being a naphthalene derivative, is doubtful.

It is of interest to examine other substances which may bind to the same receptors as the hallucinogens. Orbital plots for three such substances are shown in Figure 8. It is particularly interesting to see that dopamine and noradrenaline give quite different LUMO

plots, even though the additional hydroxy group is not bonded to the aromatic ring. The Θ_L values for dopamine and noradrenaline are 91.7° and 102.3°, respectively, while the Θ_H values are almost identical at 37.4° and 36.3°. It may be noted that no hallucinogen listed in Table 2 comes close to matching both HOMO and LUMO phase angles of either dopamine or noradrenaline, though there are some matches for one of Θ_H and Θ_L, all for drugs of low activity. The plot for serotonin probably cannot be interpreted using the present technique, as it is difficult to relate the indole aromatic system to that of benzene, although the orbitals can perhaps be compared visually.

An earlier publication[1] has described more than 90% of the variance in the human activities of a subset of the drugs considered here in terms of electronic, hydrophobic, and steric factors. The principal equation of this study is reproduced here as eq 3:

$$\log A = C_1 I_{Me} + C_2 H_p + C_3 (E_L - E_H) + C_4 H_p^2 + C_5 H_m H_p + C_6 V_m V_p + C_7 V_p^2 + C_8 (Q_m - Q_o) + C_9 \quad (3)$$

	1	2	3	4	5	6	7	8	9
C	0.474	1.348	−10.89	−0.270	−1.885	2.11×10^{-4}	-2.78×10^{-4}	−0.676	6.29
se	0.077	0.112	1.759	0.056	0.256	5.8×10^{-5}	4.6×10^{-5}	0.129	0.96
α	0.00000	0.00000	0.00000	0.00002	0.00000	0.00072	0.00000	0.00001	0.00000

$N = 50$, $R^2 = 0.915$, $Q^2 = 0.880$, $s = 0.16$, $F = 54.9$, $P = 2 \times 10^{-19}$, $\Lambda = 10.4$

Here, H_p is the Hansch π hydrophobicity parameter[36] for the para (with respect to the ethylamine group) substituent, H_m is the sum of the π parameters for the two meta substituents, and V_p and V_m are the van der Waals volumes of the para and meta substituents (using in each case zero for an unsubstituted position). The charges Q_o and Q_m are respectively the sums of Mulliken charges of the carbon atoms 1 and 6, and 2 and 5 of the benzene ring, calculated by CNDO.[37] In this study, E_H and E_L were also calculated by CNDO.

The standard error of estimate in this earlier study implies an accuracy of ±45% in the accuracy of the measurement of activity. This is close to, if not beyond, what can be determined experimentally, leaving little scope for improvement with further parameters. The predictivity of eq 3 was improved somewhat by consideration of nonlinearity in a later publication.[18]

It is of interest to see how well the descriptors of eq 3 correlate with the phase angle descriptors of the present study. A cluster analysis correlation dendrogram[38] of the combined descriptors of the two studies is presented in Figure 9. This may be interpreted as follows: Each descriptor is represented by a horizontal line. Two horizontal lines are joined by a vertical line, the intersection of which with the horizontal scale at the top, with 1.0 on the left and 0.0 on the right, gives the magnitude of the correlation coefficient between the two descriptors, which are then combined to form a cluster. The correlation between clusters of descriptors is obtained by pooling the standardized descriptor values for each of the two groups and calculating the correlation between the two pooled groups, using the formulas given by Spearman.[39] The strongest correlation with log A is Δ_H, followed by sin $2\Theta_H$ and cos $2\Theta_H$. It may be seen that the strongest 1:1 correlation in the data is 0.934, between H_p and H_p^2. Because these are completely functionally related, this is not a problem for the regression analysis.

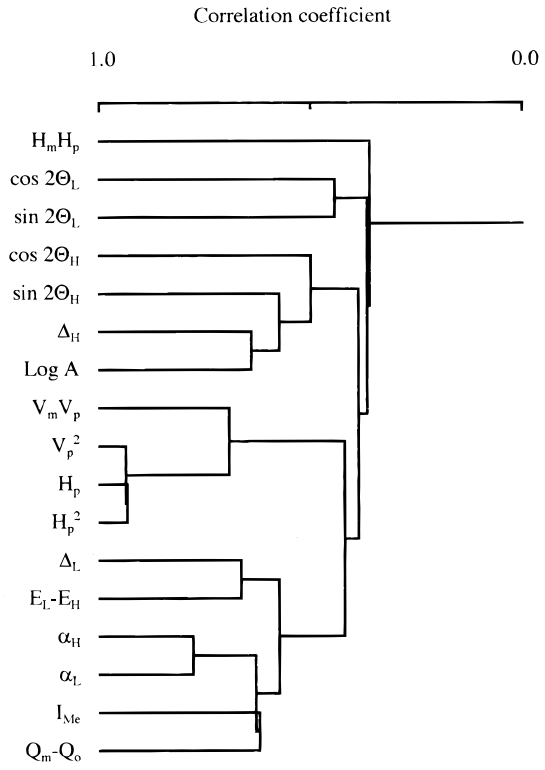

Figure 9. Correlation dendrogram of pooled variables, this paper, and eq 3.

The correlation of this cluster with V_p, however, is equally strong, and this correlation is the weakest feature of eq 3. The strongest correlation between the new frontier orbital descriptors is between α_H and α_L at 0.779. This is not very strong, and we have in any case not proposed any equations that include both. There are no other very strong correlations in the diagram.

This analysis does not exclude multicollinearities, involving three or more descriptors. To check for these, we may carry out a principal components analysis on the pooled descriptors of this paper and eq 3. The last four columns of the eigenvectors of the correlation matrix of these descriptors are given in Table 3. It is only the last two columns of the eigenvector matrix that take Λ beyond 5.0, so it is the large coefficients in these two columns that indicate the significant collinearities. Discounting those coefficients less than 0.1, it may be seen that there is a major multicollinearity involving V_p^2, V_mV_p, H_p^2, and to a lesser extent Δ_L and H_mH_p. A lesser multicollinearity involves H_p, H_p^2, and to a lesser extent V_p^2, V_mV_p, Δ_L, Δ_H, and α_L. These two multicollinearities involve the variables of Equation 3 much more than the descriptors presented in this paper.

When an all possible subsets regression is carried out on the pooled descriptors, using Mallow's C_p as criterion, eq 3 is the best eight-variable equation but is not represented in the 10 overall best regressions. All of its descriptors, however, are represented in all of them.

Table 3. Least Significant Four Eigenvectors of the Correlation Matrix of the Pooled Descriptors, This Paper, and Eq 3

	component			
	13	14	15	16
eigenvalue	0.091	0.052	0.025	0.013
α_H	−0.609	0.406	−0.034	0.017
α_L	0.306	−0.436	−0.115	0.026
cos $2\Theta_H$	0.074	0.040	−0.001	0.005
sin $2\theta_H$	−0.104	0.068	0.021	−0.095
cos $2\Theta_L$	−0.098	−0.077	−0.020	0.028
sin $2\Theta_L$	−0.032	−0.041	−0.043	0.021
Δ_L	−0.111	0.080	0.215	0.123
Δ_H	0.248	0.179	0.175	−0.033
I_{Me}	0.000	0.228	−0.015	−0.067
$E_L - E_H$	−0.205	0.304	0.091	0.088
H_p	−0.109	0.043	−0.807	−0.080
H_p^2	−0.422	−0.457	0.419	−0.292
$H_m H_p$	−0.096	−0.112	0.012	0.131
V_p^2	0.184	0.176	0.187	0.735
$V_m V_p$	0.400	0.434	0.176	−0.550
$Q_m - Q_o$	0.027	−0.072	−0.070	0.093

The best equation involves in addition α_H, cos $2\Theta_H$, and sin $2\Theta_L$. This equation leaves 20% less variance unexplained than does eq 3, and although in this equation the new coefficients are not statistically significant at the 0.05 level, the significance of the old terms is still in all but one case better than 0.0005. This indicates that the new descriptors in this equation are almost orthogonal with the old.

The hypothesis which led to this study results naturally in eq 2. Statistically, however, eq 1 is very nearly as satisfactory and may be preferred as being simpler. In either case there is an optimum value of Θ_H and Θ_L, leading to the greatest activity. Even if eq 2 is preferred, it appears that it is not pure Q-like orbitals, but intermediate forms, which maximize hallucinogenic activity. Thus, the original hypothesis that led to this study must be modified. The actual orientation of the nodes of the orbitals with respect to the ethylamine side chain is important, as also is the amount of splitting of the degenerate HOMO and LUMO orbitals of the benzene. It is likely that the physical basis of the correlation is the matching of the HOMO and LUMO orbitals of the benzene ring of the hallucinogen with complementary orbitals on some component of the receptor, such as an aromatic amino acid residue, in a form of π charge-transfer interaction.

Acknowledgment. I am indebted to Paula J. McLay of the Australian Government Analytical Laboratories for helpful suggestions after reading the manuscript.

References

(1) Clare, B. W. Structure-Activity Correlations for Psychotomimetics. 1. Phenylalkylamines: Electronic, Volume, and Hydrophobicity Parameters. *J. Med. Chem.* **1990**, *33*, 687−702.
(2) Clare, B. W. Unpublished observations.
(3) Cohen, S. The Chemical Transcendent State. In *The Psychopharmacology of Hallucinogens*, Stillman, R. C., Willette, R. E., Eds.; Pergamon: New York, 1978; pp 324−329.
(4) Cohen, S. The Psychotomimetic Agents. *Arzneim Forsch* **1970**, *15*, 68−102.
(5) Shulgin, A. T. Psychotomimetic Drugs: Structure-Activity Relationships. In *Handbook of Psychopharmacology;* Iversen, L. L., Iversen, S. D., Snyder S. H., Eds.; Plenum: New York, 1980; Vol. 11, pp 243−333.
(6) Shulgin, A.; Shulgin, A. *PIHKAL: A Chemical Love Story*; Transform Press: Box 13675, Berkeley, CA, 1991.
(7) Coulson, C. A. *Valence*; 2nd ed.; Oxford University Press: Oxford, 1961; p 254.
(8) *Hyperchem v. 5.1*, Available from Hypercube, Inc., 1115 NW 4th St., Gainesville, FL 32601-4256.
(9) *Mathematica v. 3.0*, available from Wolfram Research International, 100 Trade Center Dr., Champaign, IL 61820.
(10) File angle.zip at site ftp://central.murdoch.edu.au/pub/chem/clare.
(11) *PCMODEL v. 5.0*, Available from Serena Software, Box 3076, Bloomington, IN 47402-3076.
(12) Stewart, J. J. P. MOPAC: A general molecular orbital package, version 6.0. *QCPE Bull.* **1990**, *10*, 86−87.
(13) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AM1: A new general purpose quantum mechanical model. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(14) *MOPAC 93.00* (1993) Stewart, J. J. P., Fujitsu Ltd., Tokyo, Japan, also Stewart, J. J. P. MOPAC93 Release 2. *QCPE Bull.* **1995**, *15*, 13−14 (Copyright Fujitsu 1993, all rights reserved).
(15) Klamt, A.; Shüürman, G. COSMO: A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans 2* **1993**, 799−805.
(16) *BMDP Dynamic, Release 7.* Available from BMDP Statistical Software, Inc., 1440 Sepulveda Boulevard, Suite 316, Los Angeles, CA 90025.
(17) Clare, B. W. Alternating conditional expectations in QSAR. In *Methods and Principles in Medicinal Chemistry*; Mannhold R., Krogsgaard-Larsen P., Timmerman H., Eds.; VCH: Weinheim, 1994; Vol. 3. *Advanced Computer-Assisted Techniques in Drug Discovery*; van de Waterbeemd, H., Ed. pp 284−289.
(18) Clare, B. W. Structure-activity correlations for psychotomimetics. 2. Phenylalkylamines: a treatment of nonlinearity using the alternating conditional expectation technique *Chemom. Intell. Lab. Syst.* **1993**, *18*, 71−93.
(19) Furnival, G. M.; Wilson R. W. Regression by leaps and bounds. *Technometrics* **1974**, *16*, 499−511.
(20) Edwards, A. L. *Multiple Regression and the Analysis of Variance and Covariance*, 2nd ed.; Freeman: New York, 1985; pp 49−50.
(21) Chatterjee, S.; Price, B. *Regression Analysis by Example,* 1st ed.; Wiley: New York, 1977; pp 199−200.
(22) Daniel, C.; Wood, F. S. *Fitting Equations to Data*; Wiley: New York, 1971.
(23) Efroymson, M. A. Multiple Regression Analysis. In *Mathematical Methods for Digital Computers*; Ralston, A., Wilf, H. S., Eds.; Wiley: New York, 1962; p 191.
(24) Topliss, J. G.; Costello, R. J. Chance correlations in structure activity studies using multiple regression analysis. *J. Med. Chem.* **1972**, *15*, 1066−1068.
(25) Topliss, J. G.; Edwards, R. J. Chance Factors in Studies of Quantitative Structure-Activity Relationships. *J. Med. Chem.* **1979**, *22*, 1238−1244.
(26) Breiman, L.; Friedman, J. H. Estimating optimal transformations for multiple regression and correlation. *J. Am. Statist. Assoc.* **1985**, *80*, 580−619.
(27) Kang, S.; Green, J. P. Steric and electronic relationships among some hallucinogenic compounds. *Proc. Natl. Acad. Sci. U.S. A.* **1970**, *67*, 62−67.
(28) Nichols, D. E.; Kostuba, L. J. Steric effects of substituents on phenethylamine hallucinogens. 3,4-Methylenedioxyamphetamine analogues alkylated on the dioxole ring. *J. Med. Chem.* **1979**, *22*, 1264−1267.
(29) Nichols, D. E.; Shulgin, A. T. Directional lipophilic character in a series of psychotomimetic phenethylamine derivatives. *Life Sci.* **1977**, *21*, 569−576.
(30) Barfknecht, C. F.; Nichols, D. E. Correlations of psychotomimetic activity of phenethylamines and amphetamines with 1-octanol/water partition coefficients. *J. Med. Chem.* **1975**, *18*, 208−210.
(31) Snyder, S. H.; Merril, C. R. A relationship between the psychotomimetic activity of drugs and their electronic structure. *Proc. Natl. Acad. Sci. U.S.A.* **1965**, *54*, 258−266.
(32) Parr, R. G.; Pearson, R. G. Absolute hardness: companion property to absolute electronegativity. *J. Am. Chem. Soc.* **1983**, *105*, 7512−7516.
(33) Clare, B. W. The relationship of charge-transfer complexes to frontier orbital energies in QSAR. *J. Mol. Struct. (THEOCHEM)* **1995**, *331*, 63−78.
(34) Clare, B. W. Charge-transfer complexes and frontier orbital energies in QSAR: a congeneric series of electron acceptors. *J. Mol. Struct. (THEOCHEM)* **1995**, *337*, 139−150.
(35) Clare, B. W. Frontier orbital energies in quantitative structure-activity relationships: A comparison of quantum chemical methods. *Theor. Chim. Acta* **1994**, *87*, 415−430.
(36) Hansch, C.; Leo, A. *Substituent Constants for Correlation Analysis in Chemistry and Biology*; John Wiley & Sons: New York, 1979.

(37) Pople, J. A.; Beveridge, D. L. *Approximate Molecular Orbital Theory*; McGraw-Hill: New York, 1970.
(38) Sokal, R. R.; Sneath, P. H. *Principles of Numerical Taxonomy*; Freeman: San Fransisco, 1963; p 182.
(39) Spearman, C. Correlations of sums or differences. *Br. J. Psychol.* **1913**, *5*, 417−426.

JM980144C