Theor Chim Acta (1994) 87: 415–430

Theoretica Chimica Acta

Frontier orbital energies in quantitative structure–activity relationships: A comparison of quantum chemical methods

Brian W. Clare

School of Mathematical and Physical Sciences, Murdoch University, Murdoch, Western Australia 6150, Australia

Received March 12, 1993/Accepted June 24, 1993

Summary. The energies of the highest occupied and lowest unoccupied molecular orbitals (HOMO and LUMO) have long been used as descriptors in QSAR (Quantitative Structure–Activity Relationships). It is shown that different quantum chemical methods of calculating these energies yield results which sometimes correlate poorly with each other. This could seriously affect physical interpretation of QSAR equations. A comparison is made between HOMO and LUMO energies and their differences and sums (hardness and electronegativity) calculated by some of the best known *ab initio* and semi-empirical methods for two series of simple organic molecules. The difference between the HOMO and LUMO energies correlates better between methods than does either alone, and their sum correlates relatively poorly. MINDO/3 (Modified Intermediate Neglect of Differential Overlap, version 3) is the poorest method in terms of correlation with the more extended basis set *ab initio* methods, followed by CNDO (Complete Neglect of Differential Overlap) and INDO (Intermediate Neglect of Differential Overlap). The best semi-empirical methods, in terms of correlation with experiment and the more extended basis set *ab initio* calculations, are MNDO (Modified Neglect of Differential Overlap), AM1 (Austin Model 1) and PM3 (Parametric Method 3). The simplest *ab initio* method, STO-3G, does not agree as well with the extended basis set calculations or with experimental results as the more advanced semi-empirical methods.

Key words: QSAR – HOMO – LUMO – Electron affinity – Ionization potential – Semi-empirical – *Ab initio* – frontier orbitals

1 Introduction

From early in the development of quantitative structure–activity relationships (QSAR) descriptors derived from quantum chemistry have been correlated with biological activity. The descriptors used have included orbital energies, orbital coefficients, Mulliken charges on atoms and superdelocalizabilities [1]. The negative of the highest occupied molecular orbital (HOMO) energy is used as an estimate of the ionization potential (IP) and that of the lowest unoccupied molecular orbital (LUMO) energy is used as a measure of the electron affinity (EA). More

recently, the identification [2, 3] that half the sum of the IP and EA is the electronegativity of the molecule, and half their difference is its hardness has made these two quantities attractive QSAR descriptors.

The computational methods used to determine these quantities have ranged from Huckel theory through CNDO (Complete Neglect of Differential Overlap) to the more rigorously developed semi-empirical methods of Dewar and the *ab initio* methods developed by Pople and others. The CNDO method, and the closely related INDO (Intermediate Neglect of Differential Overlap) method were parameterized to agree with the early *ab initio* methods, while Dewar's strategy was to use experimental information to secure agreement with measured quantities, which included IP's but not EA's. All of these methods have the advantage that they can be applied in advance of the synthesis of the drug, and that they describe properties which are not experimentally determinable.

There are a number of QSAR studies in the literature relating the activity of drugs to frontier orbital energies, calculated by CNDO [4, 5] and more recent [6, 7] quantum chemical methods. If the object of such a study is the prediction of the activities of unknown compounds, it is of little consequence if the different methods yield numerical values which do not correlate with real properties of the molecules, provided that good correlations with biological activities are obtained. If however the object is physical interpretation of the QSAR equation, it is crucial that the parameter in question correlates well with the physical variable which it describes.

In early studies, Kang and Green [8] and Snyder and Merril [9] found, using Huckel theory and CNDO, that psychotomimetic activity in phenylalkylamines and indolylalkylamines was related to HOMO energies. In a later study, Clare [10] using the CNDO method and a much larger group of phenylalkylamines found that a better relationship could be obtained with LUMO energy, and better still with the HOMO-LUMO energy difference. This negates the conclusion of the earlier workers that psychotomimetics act as electron donors in the formation of charge-transfer complexes. Further unpublished work [11] showed that there was little correlation between HOMO energies calculated by CNDO and by PM3.

This of course creates a dilemma for the interpretation of the QSAR's: Which, if either, method should be used? If the QSAR is to be interpreted in physical terms, it is necessary to know what significance may be attached to the calculated energies, and which computational method gives results more related to physical reality. Otherwise, the interpretation of the results may depend strongly on which quantum chemical method is used in the study, and may well be completely misleading. Thus a study of the relationship between the values of the frontier orbital energies calculated by different methods, and measured IP's and EA's, is timely. The object of this paper is to determine how well HOMO and LUMO energies generated by these methods compare, both with each other and with equivalent quantities derived from experiment.

The negative of the HOMO and LUMO energies are usually identified with the IP and EA respectively. The justification for this is Koopmans' Theorem, the derivation of which assumes that the geometry of the compound is unchanged on the gain or loss of an electron, and that the orbitals of the molecule and resulting ions are identical, and also ignores the change in correlation energy. These errors tend to cancel for IP's, but not for EA's [12]. Also, in the calculation of an SCF (Self Consistent Field) wave function, the occupied but not the unoccupied orbitals are optimized [13]. Thus it may be expected that the HOMO energy will be a much better measure of the IP than the LUMO energy is of the EA.

An improved way of estimating the EA of a molecule may be by way of the HOMO energy of the negative ion. This is because for a given nuclear configuration, the HOMO of the negative ion is by definition equal to the LUMO of the corresponding molecule, and the HOMO of the negative ion should be more reliable for the reasons given above. A direct measure of the IP and the EA would be given by the difference between the total energy of the positive or negative ion respectively, and that of the molecule. Problems of accuracy are to be expected in this computation, as the result is obtained as a small difference between two large numbers, so this method is likely to work only with extended basis sets, and an effective treatment of correlation.

Even when interpretation of the QSAR equation is the goal, correlation of the descriptor with the underlying physical quantity and not numerical agreement is all that is required. Correlation implies a linear relationship, and this relationship will be incorporated into the resulting QSAR equation, and the statistics of this equation will be unaffected, as will predicted values. In using such an equation, it is of course necessary to calculate the descriptor in the same way as was done in its derivation. Because this work is primarily concerned with correlation, the results are reported in the form of dendrograms of the correlation matrix.

Since the objective of this work is understanding the role of the frontier orbitals in QSAR, an ideal data set would be several groups of congeneric series, for which experimental IP and EA data is available. There are relatively few simple neutral organic molecules for which both experimental IP's and experimental EA's are known, so an exhaustive systematic study of the reliability of the calculated HOMO and LUMO energies, and their sum and difference is at this time not feasible. Also, it is difficult to imagine a congeneric series the members of which are small enough for an *ab initio* study. Four sets of data will be treated.

Firstly, a set of phenylalkylamines with psychotomimetic properties, for which a QSAR has been previously reported [10] will be examined. These molecules are typical of those used in QSAR studies in that they consist of a parent molecule with varying substituents. It will be determined whether there is agreement between the HOMO and LUMO energies, determined by the two different methods, CNDO and PM3 (Parametric Method 3). There is little experimental data on IP's and EA's for these compounds, so they cannot be used as a check on the actual values of the calculated energies. There is also no possibility of an extended basis set *ab initio* calculation for these compounds, because of their size.

Secondly, a set of 22 small organic molecules for which IP's are available will be subjected to a fairly exhaustive treatment by both semi-empirical and *ab initio* methods. These compounds are amenable to *ab initio* calculation, but are not typical of series encountered in QSAR, as they are of diverse structure. Most of these compounds, like the majority of simple organic compounds, have positive LUMO energies, and so their EA's would be very difficult to measure, and are not known.

The third set comprises 19 somewhat larger organic molecules for which experimental EA's are known. These are mostly too large to permit any but the most rudimentary *ab initio* treatment. Finally, to bridge the gap between the group of small but diverse molecules and the congeneric series of phenylalkylamines, 33 psychotomimetics of more diverse structures will be added to the latter set. These include tryptamines, ergolines, β-carbolines, ibogaine and yohimbine [11]. This set will serve to test the hypothesis that part of the problems found with the first data set are due to the narrow range of variation in that data set.

2 Calculations

Most calculations were done on a Toshiba 5200 Personal Computer using a Lahey F77L EM32 FORTRAN compiler. Some of the larger *ab initio* calculations were done on a Hewlett Packard Series 9000 Model 720 computer. The hierarchical cluster analysis program was written by the author. The geometry of each of the molecules in Table 1 was set up using the molecular modelling program DTMM [14], and fully optimized using the PM3 Hamiltonian of the semi-empirical program MOPAC, Version 6 [15]. PM3 is the most recent, and probably the most reliable of the semi-empirical methods.

An earlier study [10] related psychotomimetic activity in a group of phenylalkylamines to calculated properties, including frontier orbital energies. These energies were calculated by the CNDO method [16]. The HOMO and LUMO energies of the 50 active members of this set of compounds were recalculated with PM3, and these energies were also calculated for the 33 additional compounds by both CNDO and PM3.

A plot of the HOMO energy calculated by PM3 against that calculated by CNDO for the 50 phenylalkylamines is given in Fig. 1A. It can be seen that there is little correlation between the two. Figure 1B shows a similar plot for the LUMO energy, indicating that there is only slightly more correlation than with the HOMO energies. However, the difference between the two energies is very strongly correlated between the two methods, as may be seen in Fig. 1C. The sums of the frontier

Table 1. Test Compounds used in correlation analysis

List 1 Compounds of known ionization potential	List 2 Compounds of known electron affinity
Acetaldehyde	Anthracene
Acetone	Anthraquinone
Acetonitrile	Azulene
Acetylene	Benzene
Allene	Benzonitrile
Benzene	Benzophenone
Cyclopropane	Benzoquinone
Dimethyl ether	*o*-Dicyanobenzene
Ethane	*t*-Dicyanoethylene
Ethanol	Fluoranil
Ethylene	4-Fluorobenzophenone
Ethylene oxide	Hexafluorobenzene
Fluoroethane	Maleic anhydride
Formaldehyde	1-Naphthaldehyde
Formic acid	Naphthalene
Furan	Nitrobenzene
Hydrogen isocyanate	Nitromethane
Ketene	Phenylbenzoquinone
Methyl isocyanide	Tetracyanoquinodimethane
Methylamine	
Nitromethane	
Pyrrole	

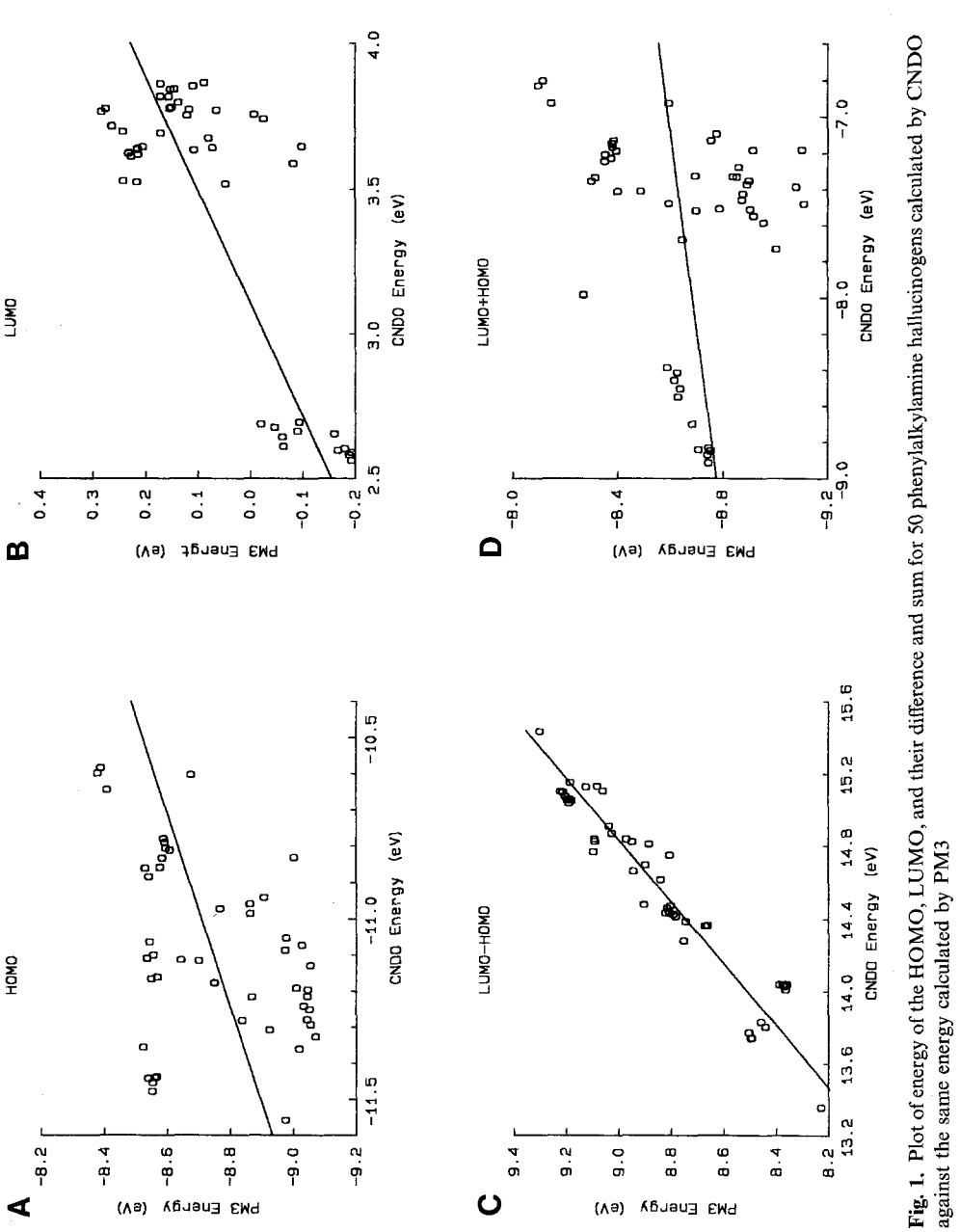
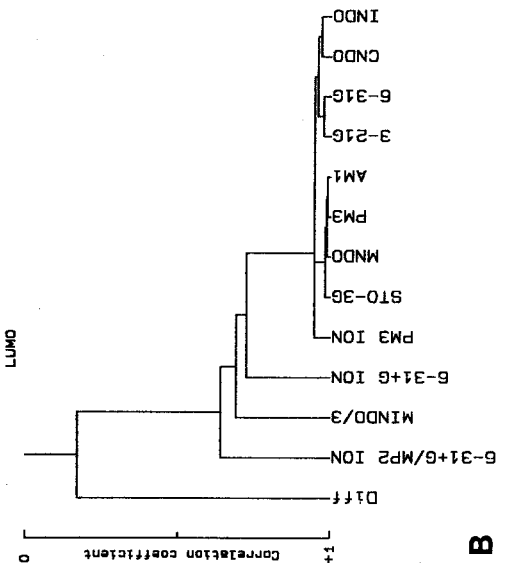
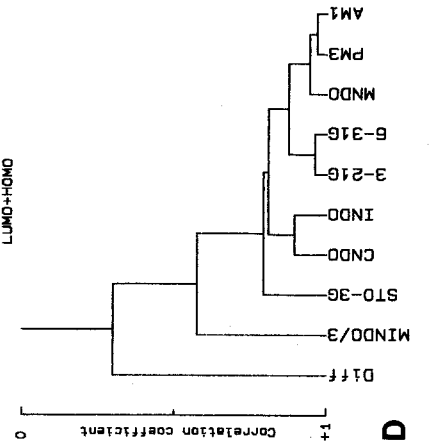
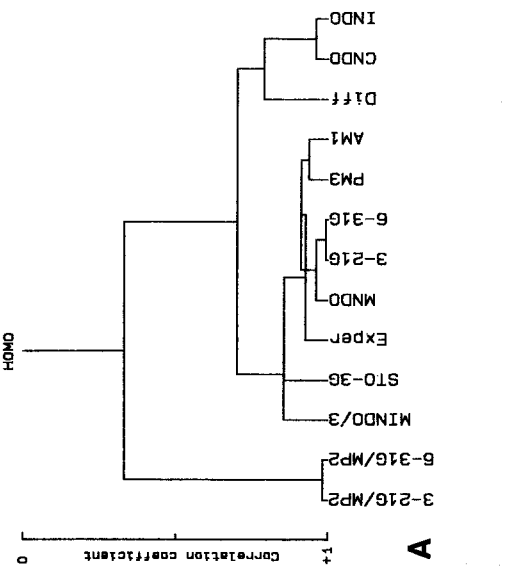
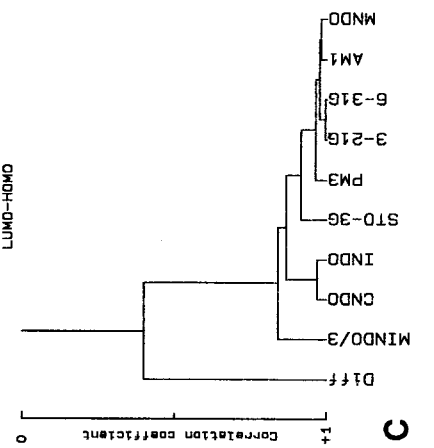


Fig. 1. Plot of energy of the HOMO, LUMO, and their difference and sum for 50 phenylalkylamine hallucinogens calculated by CNDO against the same energy calculated by PM3


INDO
CNDO
6-31G
3-21G
AM1
PM3
MNDO
STO-3G
PM3 ION
6-31+G ION
MINDO/3
6-31+G/MP2 ION
Diff
LUMO
Correlation coefficient
0
+1
B


AM1
PM3
MNDO
6-31G
3-21G
INDO
CNDO
STO-3G
MINDO/3
Diff
LUMO+HOMO
Correlation coefficient
0
+1
D


INDO
CNDO
Diff
AM1
PM3
6-31G
3-21G
MNDO
Exper
STO-3G
MINDO/3
6-31G/MP2
3-21G/MP2
HOMO
Correlation coefficient
0
+1
A


MNDO
AM1
6-31G
3-21G
PM3
STO-3G
INDO
CNDO
MINDO/3
Diff
LUMO-HOMO
Correlation coefficient
0
+1
C

orbital energies calculated by the same two methods, are not statistically related, even at the 90% significance level as shown in Fig. 1D.

The frontier orbital energies of the molecules in Table 1, list 1 were calculated by the semi-empirical methods CNDO and INDO using the program CNINDO [16], by MINDO/3 (Modified Intermediate Neglect of Differential Overlap, version 3), MNDO (Modified Neglect of Differential Overlap), AM1 (Austin Model 1) and PM3 using the program MOPAC [15], and at three levels of *ab initio* theory (STO-3G, 3-21G and 6-31G) by the program GAMESS [17, 18]. As well as the frontier orbital energies, the IP and EA were also calculated from the difference between the total energies of the neutral molecules and the appropriate ion at the 6-31 + G level [19]. The geometries of the ions were not optimized, so the calculated IP's and EA's are vertical rather than adiabatic, and so should correlate with the frontier energies rather than the experimental values. The calculations on the molecules were done with the restricted closed-shell Hartree-Fock method. Those on the ions were done with the unrestricted Hartree-Fock method, and in all cases a doublet was assumed to be the ground state. The frontier orbital energies for each of the molecules in Table 1, list 2 were calculated by the semi-empirical methods only. The experimental IP's [20] and EA's [21] were taken from the literature.

The sets of frontier orbital energies, IP's and EA's, of the compounds in Table 1, list 1 as well as their sum and difference, were individually treated by hierarchical clustering, using Spearman's formulae [22] for combining correlations of sums. The results of these cluster analyses are given in Fig. 2. The vertical axis of these plots is the correlation coefficient, ranging from 1.0 on the bottom to 0.0 on the top. The horizontal scale identifies the calculation method used. Thus two vertical lines, connected by a horizontal line, means that the energies calculated by the methods identified by the vertical lines correlate at the level indicated by the intercept of the horizontal line on the vertical axis. When two or more methods combine in this way, their correlation with other methods or groups of methods is obtained by summing the standardized values of the energies for the methods in the groups in question.

Also plotted in Fig. 2A are the HOMO energies resulting from the second-order Møller-Plesset corrections to the 3-21G and 6-31G results, and plotted in Fig. 2B are the energies of the HOMO of the negative ions, calculated with PM3, and at the 6-31+G level with and without the second-order Møller-Plesset correction. Correlations between methods for sums and differences of frontier orbital energies calculated by those methods common to the HOMO and LUMO sets of Fig. 2A and 2B are shown in Fig. 2C and 2D. The method labelled Diff in Fig. 2A to 2D is that involving the difference between the total energies of the

Fig. 2. Dendrograms of the correlations between calculation methods of the energy of the HOMO, LUMO, their difference and sum for the compounds of Table 1 list 1. Plot **A** also includes measured ionization potential and ionization potential calculated by difference between the total energy of the cation and molecule, at the 6-31+G/MP2 level, and **B** includes electron affinities calculated by the difference between the total energy of the anion and molecule, also at the 6-31+G/MP2 level, and the HOMO energy of the anion at the same level, and the HOMO of the anion by PM3. The correlation between methods or groups of methods is given by the height of the horizontal line connecting them. Thus in **A** the 3-21G/MP2 and 6-31G/MP2 methods are very closely correlated with each other (correlation coefficient 0.98), but are poorly correlated with all other methods, including the experimental values

molecule and the positive or negative ion, calculated at the 6-31+G level. The methods in Fig. 2B labelled PM3 ION, 6-31+G ION and 6-31+G/MP2 ION are those using the HOMO energy of the negative ion.

3 Discussion

The group of compounds in Table 1 list 1 is not typical of those normally encountered in QSAR, in that they do not constitute a congeneric series. This is unfortunate, since such series are of primary interest to medicinal chemists. It is however unavoidable, as the molecules encountered in congeneric series of drugs are much larger, and cannot be studied by *ab initio* methods because of the excessive demands of the latter on computer time and memory resources. The compounds such as those used in producing Fig. 1 are much more homogeneous in type than the 22 compounds of Table 1 list 1, which were used to generate Fig. 2. Thus the variability of the data in list 1 is greater than that in the usual QSAR study, such as that shown in Fig. 1, the HOMO and LUMO energies in Fig. 1A and 1B ranging over 0.9 and 0.6 eV respectively, while those of the compounds in list 1 range over 3.6 and 4.0 eV.

To illustrate the effect of homogeneity of the compound set on the correlation between frontier energies calculated by different methods, the plots of Fig. 3A to D were constructed. These correspond to Fig. 1A to D, but with the additional 33 compounds. These plots show an appreciably better correlation between the two methods than those in Fig. 1A to D, with the more homogeneous group of compounds. The HOMO and LUMO energies in these plots range over 1.2 and 0.9 eV respectively, which is somewhat more than those of Fig. 1, but still much less than those calculated for the test data of Table 1 list 1. This suggests that in series of closely related compounds covering a narrow range of frontier energies there is an indeterminacy in the zero of the energy scale. Thus while the HOMO and LUMO energies of the phenylalkylamines calculated by the two methods each show poor correlation, their difference correlates very well between methods, and their sum very poorly. Increasing the chemical diversity of the compounds studied raises the range of both HOMO and LUMO energies, and reduces the apparent effect of the variation of the energy zero.

The energies of the HOMO's in Fig. 2A seem to converge well as the computational method becomes more sophisticated, but they do not converge to the experimental IP's. This may be expected from the approximations involved in Koopman's Theorem, which relies on the cancellation of the energy change in redistributing the orbitals with the change in correlation energy. Thus when the MP2 correction for the 3-21G and 6-31G wavefunctions are applied, the correlation between methods actually becomes much poorer, as the cancellation of errors is removed. The method based on the calculated difference between the total energy of the molecule and that of the positive ion also fares rather badly, correlating better with CNDO and INDO than with the improved semi-empirical and the extended basis set *ab initio* HOMO energies, or the experimental IP's.

It has been stated that LUMO energies are not expected to mirror EA's as well as HOMO energies do IP's [12, 13]. It can be seen from Fig. 2B that in spite of this, they generally correlate much better between calculation methods than do HOMO energies. The exceptions are the MINDO/3 method LUMO energy, the total energy difference at the 6-31+G level and the HOMO of the negative ion at this level with and without the MP2 correction. Even the HOMO of the negative ion

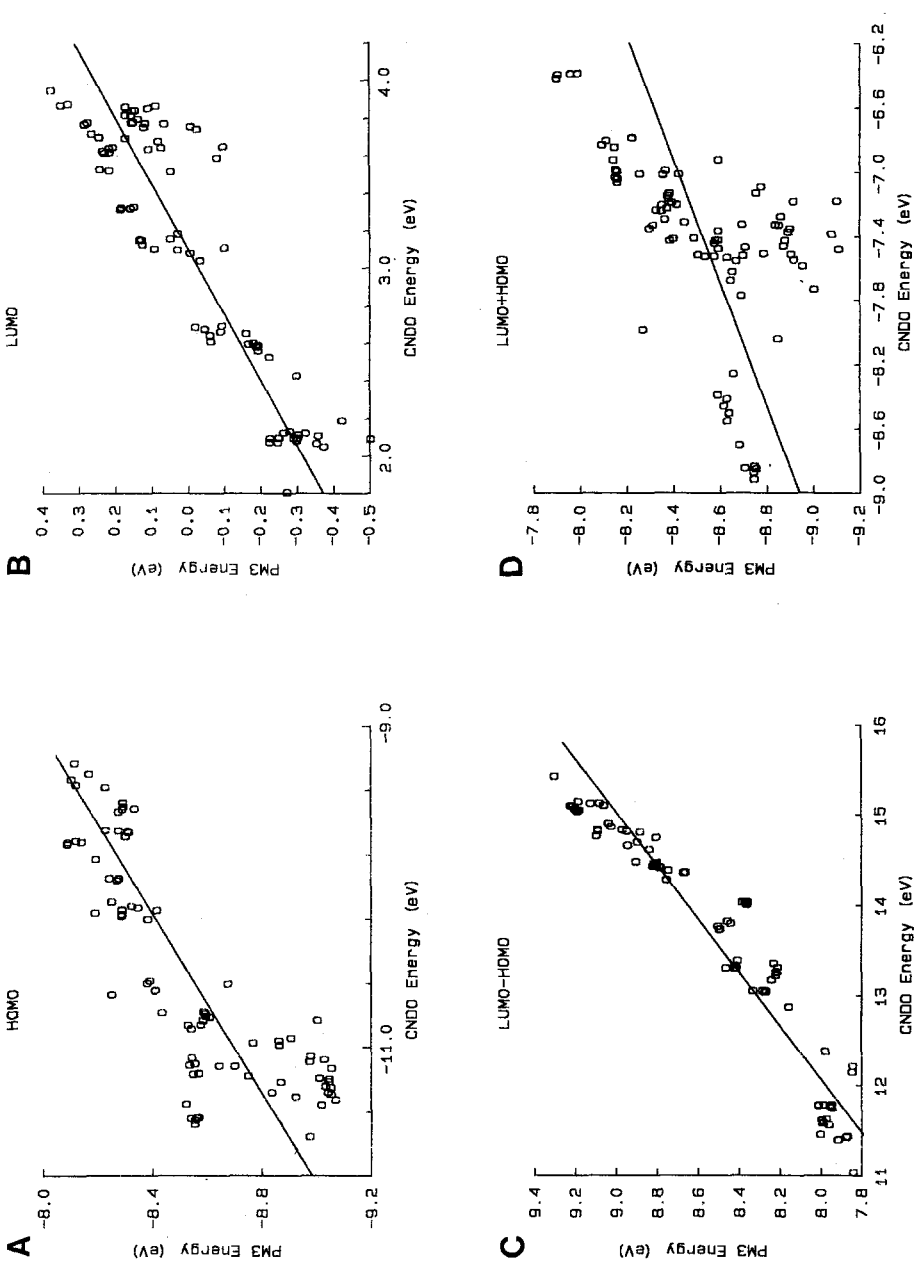


Fig. 3. Plot of the energy of the HOMO, LUMO and their difference and sum for the compounds of Fig. 1, plus 33 more tryptamines, ergolines and miscellaneous hallucinogens calculated by CNDO against the same energy calculated by PM3

Table 2. Correlation coefficients between HOMO energies of compounds of List 1 (Table 1), calculated by different methods, corresponding to Fig. 2A

Exper	1.000											
CNDO	0.743	1.000										
INDO	0.746	0.961	1.000									
PM3	0.902	0.720	0.726	1.000								
AM1	0.923	0.728	0.686	0.968	1.000							
MNDO	0.908	0.712	0.672	0.926	0.945	1.000						
MINDO	0.802	0.693	0.711	0.821	0.793	0.833	1.000					
STO-3G	0.814	0.819	0.785	0.789	0.808	0.898	0.792	1.000				
321-G	0.903	0.691	0.638	0.891	0.923	0.966	0.875	0.889	1.000			
3-21G/MP2	0.365	0.199	− 0.005	0.268	0.380	0.339	0.313	0.265	0.477	1.000		
6-31G	0.883	0.623	0.566	0.872	0.910	0.953	0.864	0.835	0.993	0.504	1.000	
6-31G/MP2	0.394	0.085	− 0.025	0.319	0.422	0.366	0.359	0.228	0.498	0.984	0.539	1.000
Diff	0.626	0.750	0.815	0.541	0.529	0.412	0.488	0.462	0.406	0.038	0.354	0.041
	Exper	CNDO	INDO	PM3	AM1	MNDO	MINDO	STO3G	3-21G	3-21G/MP2	6-31G	6-31G/MP2

Table 3. Correlation coefficients between LUMO energies of compounds of List 1 (Table 1), calculated by different methods, corresponding to Fig. 2B

CNDO	1.000											
INDO	0.969	1.000										
PM3	0.943	0.893	1.000									
AM1	0.955	0.910	0.990	1.000								
MNDO	0.974	0.922	0.982	0.988	1.000							
MINDO	0.679	0.784	0.647	0.628	0.627	1.000						
STO3G	0.924	0.851	0.979	0.983	0.973	0.562	1.000					
3-21G	0.962	0.936	0.956	0.967	0.966	0.645	0.928	1.000				
6-31G	0.941	0.947	0.906	0.918	0.922	0.691	0.855	0.977	1.000			
Diff	0.088	0.165	0.128	0.124	0.125	0.081	0.081	0.153	0.234	1.000		
A	0.731	0.838	0.641	0.665	0.686	0.655	0.578	0.746	0.824	0.522	1.000	
B	0.602	0.650	0.568	0.582	0.595	0.519	0.518	0.652	0.677	0.035	0.637	1.000
C	0.940	0.898	0.955	0.940	0.937	0.642	0.924	0.921	0.875	0.113	0.629	0.538
	CNDO	INDO	PM3	AM1	MNDO	MINDO	STO3G	3-21G	6-31G	Diff	A	B

A. 6-31G anion
B. 6-31G/MP2 anion
C. PM3 anion

calculated by PM3 correlates well with all of the LUMO's except MINDO/3. The total energy difference between the molecule and anion correlates extremely poorly with all of the other energies, and does not reach statistical significance, even at the 90% level. This is also reflected in a poor correlation of this method for the sum and difference of the HOMO and LUMO energies, as may be seen in Fig. 2C and 2D.

The correlations in Fig. 2C indicate that the difference between the HOMO and LUMO energies correlates better between calculation methods than does the HOMO energy. There is relatively good agreement between the larger basis set *ab initio* methods and the MNDO, AM1 and PM3 methods, but as before, the total energy difference ("Diff") correlates poorly with all the other energies. Figure 2D shows that the sum of the HOMO and LUMO energies correlates relatively poorly between methods, although the correlations within the groups MNDO, AM1 and PM3, and 3-21G and 6-31G are still quite good. The total energy difference, MINDO/3 method HOMO, STO-3G method HOMO, and to a lesser extent the CNDO and INDO method HOMO energies show the smallest correlation with the larger basis set *ab initio* frontier orbital energies, and those derived from the more recent semi-empirical methods. Plots derived from CNDO and PM3 calculations on the test compounds of Table 1, list 1, are shown in Fig. 4. These may be used to aid interpretation of the correlations in Fig. 2, and may be compared with the corresponding plots in Fig. 1 and Fig. 3.

Figures 2D and 4D show that in the more varied test data set of list 1, the sum of the HOMO and LUMO energies correlates better between CNDO and PM3 than it does with the phenylalkylamines in Fig. 1D or the mixed psychotomimetic data in Fig. 3D.

A cluster analysis of the frontier orbital energies of the compounds in the second list of Table 1, using only the semi-empirical methods, but including experimental EA's, is shown in Fig. 5. This indicates that the HOMO energy of the negative ion correlates relatively poorly with the experimental EA, followed by the MINDO/3 LUMO energy. The older CNDO and INDO method LUMO energies, and the more recent MNDO, AM1 and PM3 method LUMO energies correlate equally well with experimental values, and very well within their own groups. Comparison of Fig. 5 with Fig. 2B shows that the correlation structure is very similar between the two groups of compounds. The differences between the two figures are that MINDO/3 method LUMO energy correlates better, and the

Table 4. Correlation coefficients between HOMO-LUMO energy difference of compounds of List 1 (Table 1), calculated by different methods, corresponding to Fig. 2C

CNDO	1.000								
INDO	0.969	1.000							
PM3	0.853	0.828	1.000						
AM1	0.887	0.839	0.981	1.000					
MNDO	0.875	0.811	0.959	0.988	1.000				
MINDO	0.782	0.804	0.847	0.842	0.810	1.000			
STO3G	0.867	0.808	0.892	0.922	0.946	0.768	1.000		
3-21G	0.867	0.812	0.948	0.975	0.978	0.815	0.910	1.000	
6-31G	0.853	0.803	0.944	0.969	0.966	0.812	0.872	0.995	1.000
Diff	0.490	0.586	0.372	0.363	0.288	0.475	0.157	0.310	0.347
	CNDO	INDO	PM3	AM1	MNDO	MINDO	STO3G	3-21G	6-31G

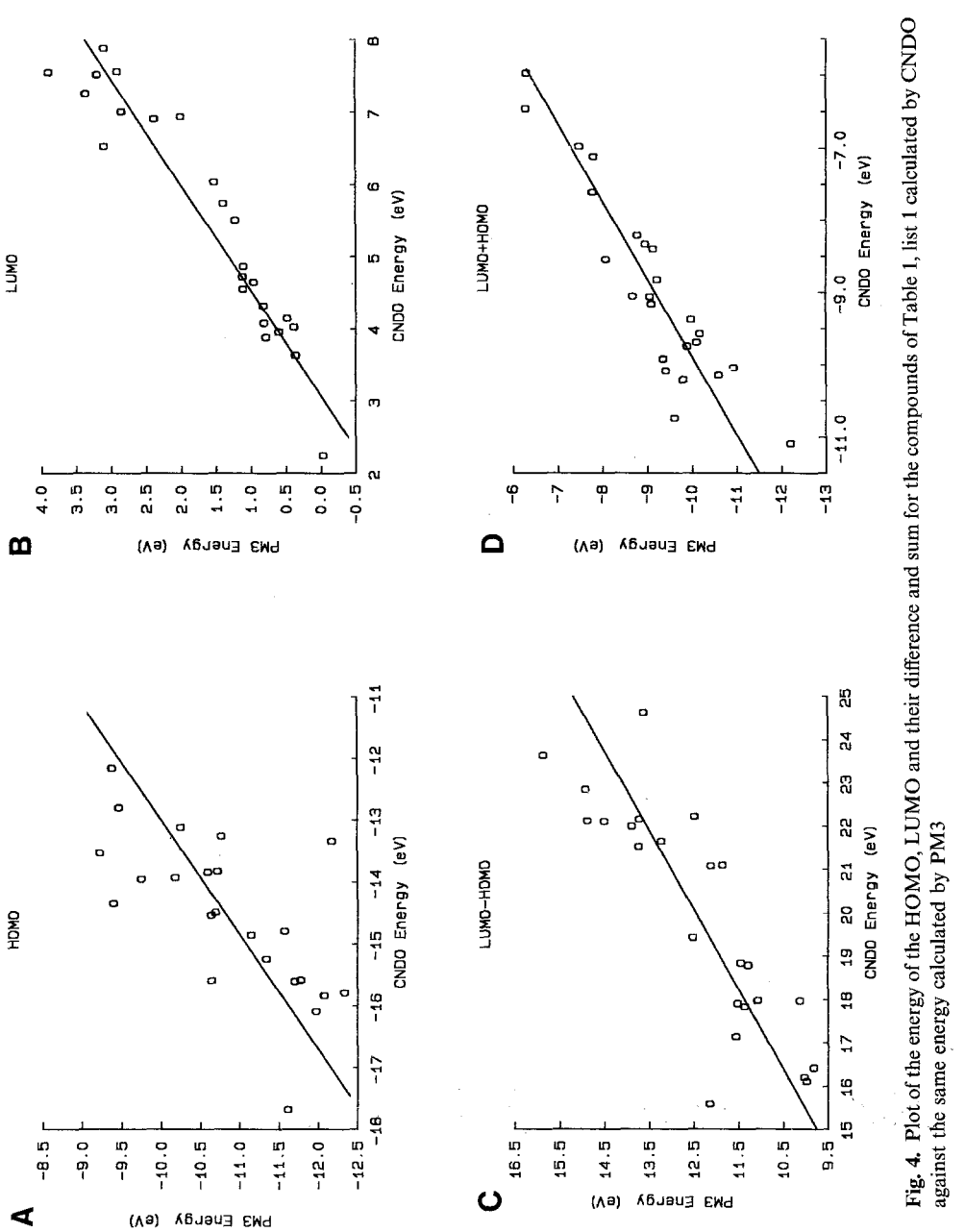


Fig. 4. Plot of the energy of the HOMO, LUMO and their difference and sum for the compounds of Table 1, list 1 calculated by CNDO against the same energy calculated by PM3

Table 5. Correlation coefficients between HOMO-LUMO energy sum for compounds of List1 (Table 1), calculated by different methods, corresponding to Fig. 2D

CNDO	1.000								
INDO	0.895	1.000							
PM3	0.897	0.755	1.000						
AM1	0.886	0.698	0.978	1.000					
MNDO	0.840	0.720	0.960	0.943	1.000				
MINDO	0.465	0.642	0.520	0.443	0.576	1.000			
STO3G	0.812	0.519	0.845	0.877	0.772	0.143	1.000		
3-21G	0.815	0.693	0.885	0.895	0.942	0.664	0.740	1.000	
6-31G	0.670	0.610	0.780	0.789	0.883	0.762	0.554	0.961	1.000
Diff	0.507	0.650	0.260	0.201	0.187	0.238	0.135	0.143	0.061
	CNDO	INDO	PM3	AM1	MNDO	MINDO	STO3G	3-21G	6-31G

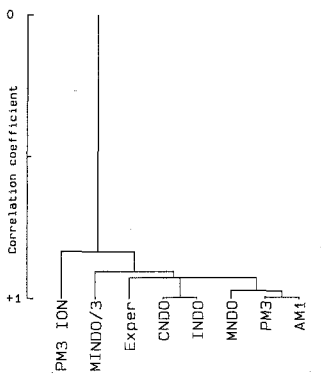


Fig. 5. Dendrogram of experimental electron affinities and LUMO energies of compounds of Table 1 list 2, by semi-empirical methods

Table 6. Correlation coefficients between LUMO energies of compounds of List 2 (Table 1), calculated by different methods, corresponding to Fig. 5

Exper	1.000						
PM3	0.915	1.000					
AM1	0.929	0.994	1.000				
MNDO	0.874	0.965	0.974	1.000			
MINDO	0.831	0.865	0.891	0.928	1.000		
CNDO	0.899	0.927	0.945	0.883	0.867	1.000	
INDO	0.908	0.918	0.943	0.887	0.884	0.996	1.000
A	0.869	0.869	0.858	0.740	0.586	0.865	0.842
	Exper	PM3	AM1	MNDO	MINDO	CNDO	INDO

A: HOMO energy of cation

PM3 method HOMO energy of the negative ion correlates worse with the remaining methods for the compounds of list 2 than would have been expected by comparison with Fig. 1B.

It may be concluded from this work that in terms of correlation with experimental IP's and EA's, the HOMO energy calculated by CNDO or INDO is not a satisfactory descriptor for use in QSAR. The LUMO energy calculated by any of

the methods other than MINDO/3 is better than the HOMO, as is the difference between the HOMO and LUMO energies, calculated preferably by MNDO, AM1 or PM3. The uniformly poorest method in terms of correlation with experimental and larger basis set *ab initio* values is MINDO/3. The poorest descriptor is the sum of the HOMO and LUMO energies, in the case of the PM3 and CNDO analysis of the psychotomimetic data, as illustrated in Figs. 1D and 3D. With the data of list 1 however, the sum gives a reasonable correlation between the two methods, although less so than the LUMO energy, as seen in Fig. 4D.

The *ab initio* methods are too demanding of computer resources to be routinely used in QSAR studies. Perhaps surprisingly, the very time-consuming *ab initio* methods involving the difference in energy between the molecule and ion, and also the estimation of the EA as the HOMO energy of the negative ion correlate very poorly with both experimental IP's and EA's, and with the Koopmans' Theorem values. This may possibly be correctable by the use of a more extended basis set.

The disconcerting observation of the lack of correlation between the HOMO and LUMO energies calculated by different methods, evident in Fig. 1A and 1B, is not reproduced by those calculated for the compounds of list 1, shown in Fig. 2A and 2B, and in Fig. 4A and 4B. Consequently, the spectacular improvement in correlation between methods obtained by using the HOMO-LUMO difference, as in Fig. 1C, is not seen in Fig. 4C. It is suggested that this is because of the close similarity of the psychotomimetic compounds, and hence the small spread of orbital energies. That homogeneity of compound type leads to poorer correlations is a very surprising observation, with far-reaching implications for QSAR studies.

Acknowledgements. I am indebted to the Research Centre for Advanced Materials and Mineral Processing for making available the Hewlett Packard computing facility, and to Paula J. McLay of the School of Mathematical and Physical Sciences, Murdoch University for a criticism of the manuscript.

References

1. Franke R (1984) Theoretical drug design methods. Elsevier, NY, pp 115–123
2. Parr RG, Pearson RG (1983) J Am Chem Soc 105:7512
3. Pearson RG, Palke WE (1992) J Phys Chem 96:3283
4. Bushelev SN, Stepanov NF (1989) Z Naturforsch 44:212
5. Kawakami Y, Hopfinger AJ (1990) Chem Res Toxicol 3:244
6. Tuppurainen K, Lotjonen S, Laatikainen R, Vartiainen T, Maran U, Strandberg M, Tamm T (1991) Mutation Research 247:97
7. Ramos MN, Neto B de B (1990) J Comput Chem 11:569
8. Kang S, Green JP (1970) Nature 226:645
9. Snyder SH, Merril CR (1965) Proc Natl Acad Sci 54:258
10. Clare BW (1990) J Med Chem 33:687
11. Clare BW (1992) Unpublished data
12. Szabo A, Ostlund NS (1989) Modern Quantum Chemistry, Revised First Ed., McGraw Hill, New York, pp 123–128
13. Sadlej J (1985) Semi-Empirical Methods of Quantum Chemistry, Ellis Horwood, Chichester, p 282
14. Crabbe MJC, Appleyard JR (1991) Desktop Molecular Modeller, Version 2.0, Oxford University Press, England
15. Stewart JJP (1990) QCPE Bull 10:86
16. Dobosh P (1975) QCPE 10:281
17. Schmidt MW, Baldridge KK, Boatz JA, Jensen JH, Koseki S, Gordon MS, Nguyen KA, Windus TL, Elbert ST (1990) QCPE Bull 10:52

18. Dupis M, Spangler D, Wendoloski JJ (1980) National Resource for Computations in Chemistry Software Catalogue, University of California: Berkely, CA, USA, Program QG01
19. Gutsev GL, Boldyrev AI (1985) Adv Chem Phys 61:169
20. Stewart JJP (1990) J Comput-Aided Mol Des 4:1
21. Chen ECM, Wentworth WE (1974) J Chem Phys 63:3183
22. Spearman C (1913) Brit J Psychol 5:417