Computer Recognition of Activity Class from Molecular Transforms

Sir:

We wish to report the first successful application of computer pattern recognition to molecular transforms derived from x-ray diffraction data. Hitherto, pattern recognition studies of compound activity (as opposed to compound identity) have been based on molecular fragment coding as the source of structure description.[1,2] Although successful activity recognition from multidimensional structure descriptions suggests the possibility of extracting subtle structure–activity relationships, work based on the use of fragment coding as the source of structural descriptors has drawn criticism.[3] The difficulty is that, by describing a molecule in terms of a predetermined list of molecular attributes, such as molecular weight, ring size, presence of carbonyl, and so forth, one may prejudice the possible results and thus draw misleading conclusions regarding which structural moieties are significant in producing activity.

Molecular transforms[4] are computed directly from three-dimensional atomic coordinates and thus partake of none of the arbitrariness of fragment coding. For class recognition where molecular shape is important, molecular transforms provide a direct source of data embodying that shape. We have used a Wierl-type expression to calculate molecular transforms:

$$I(\mathrm{s}) = \sum_{i=2}^{N} \sum_{j=1}^{i-1} Z_i Z_j \frac{\sin(sr_{ij})}{sr_{ij}}$$

Here, Z_i is the atomic number of the ith atom in an N-atom molecule, and r_{ij} is the distance between atom i and atom j. The transforms were computed at 100 equi-spaced points in the range $1\ \text{Å}^{-1} \leq s \leq 31\ \text{Å}^{-1}$ and reduced to 100-dimensional binary patterns by tabulating "1" 's in intervals where a zero-crossing (i.e., $I(\mathrm{s}) = 0$) occurred and "0" 's elsewhere. A typical pattern contained about 20 zero-crossings in the s-range studied. The use of zero-crossings to characterize the molecular transforms in binary form is analogous to peak/no-peak coding of NMR or mass spectra.[5] For the relatively broad peaks encountered in a molecular transform, determination of the peak positions is more difficult than for NMR or mass spectral peaks, whereas the zero-crossings of the molecular transform are easily identified.

Our preliminary results are reported for a set of 89 compounds, of which 42 are known to be sedatives and 47 are known to be tranquillizers.[6] These are a subset of the 219 compounds used by Stuper and Jurs[1] and include phenothiazines, benzodiazepines, barbiturates, carbamates, carbinols, and aliphatic amides. The molecular transforms for 13 of these compounds were computed from three-dimensional coordinates obtained from the Cambridge Crystallographic Data Centre data file[7] (Summer 1975 update) via the National Institutes of Health (Division of Computer Research and Technology) Chemical Information System. For those compounds not present in the file, coordinates were generated from those of similar molecules with the use of a modification of the program ATCOOR[8] and standard bond lengths, angles, and torsion angles. The molecular transforms were reduced to binary patterns, as described above, and these patterns were used to train binary pattern classifiers by the error-correction-feedback method which has been described elsewhere.[9,10] Although a data base of 89 patterns does not justify the use of 100 features, feature elimination by the weight-sign method[9,10] reduced the dimensionality to 25, which is an acceptable value for this size data base.[11]

Using molecular transforms in this manner, we have produced: (1) a 25-component weight vector, trained on all 89 compounds, capable of correctly classifying all 42 sedatives and all 46 tranquillizers; (2) a 21-component weight vector, trained on 79 compounds, capable of correctly classifying 9 out of 10 tranquillizers not present in the training set; (3) a 25-component weight vector, trained on 79 compounds, capable of correctly classifying 10 out of 10 sedatives not present in the training set; (4) a 21-component weight vector, trained on 79 compounds, capable of correctly classifying 5 tranquillizers and 4 sedatives out of 5 tranquillizers and 5 sedatives not present in the training set. These findings demonstrate that molecular transforms, prepared and sampled in the manner we have described, contain sufficient information to permit machine recognition of these two activity classes.

It should be pointed out that, since the molecular transform as we have computed it depends only on interatomic distances and not on absolute atomic positions, it may be possible to use x-ray structure-factor data directly to compute the molecular transform. The F^2_{obsd} would be used to compute a Patterson function,[12] and the height and position of the Patterson peaks would supply the Z_iZ_j and r_{ij} information needed for the molecular transform computation. The potential difficulty would be the possible inclusion of some intermolecular distances, but the prospect of achieving activity classification of a substance directly from diffraction data prior to structure solution seems worth pursuing.

Although "cranking out structures" is generally frowned upon by crystallographers, our findings argue for the collection of extensive files of three-dimensional atomic coordinates for major compound classes to be used in computer-aided investigations of structure–activity correlations. The current state of the art in x-ray structure determination and the existence of the Cambridge Crystallographic Data Centre project make this a realistic goal.

Acknowledgment. This research was supported by the National Science Foundation under Grant MPS-74-01249.

References and Notes

(1) A. J. Stuper and P. C. Jurs, *J. Am. Chem. Soc.*, **97,** 182 (1975).
(2) B. R. Kowalski and C. F. Bender, *J. Am. Chem. Soc.*, **96,** 916 (1974).
(3) R. J. Mathews, *J. Am. Chem. Soc.*, **97,** 935 (1975).
(4) H. Lipson and C. A. Taylor, "Fourier Transforms and X-Ray Diffraction", G. Bell and Sons, Ltd., London, 1958, p 9.
(5) H. B. Woodruff, G. L. Ritter, S. R. Lowry, and T. L. Isenhour, *Appl. Spectrosc.*, **30,** 213 (1976).
(6) E. Usdin and D. H. Efron, "Psychotropic Drugs and Related Compounds", 2d ed, DHEW Publication No. (HSM)72-9074, 1972.
(7) O. Kennard, D. Watson, F. Allen, W. Motherwell, W. Town, and J. Rodgers, *Chem. Brit.*, **11,** 213 (1975).
(8) J. E. Nordlander, Quantum Chemistry Program Exchange, Program No. 250.
(9) T. L. Isenhour and P. C. Jurs, *Anal. Chem.*, **43**, 20A (1971).
(10) C. L. Wilkins, R. C. Williams, T. R. Brunner, and P. J. McCombie, *J. Am. Chem. Soc.*, **96,** 4182 (1974).
(11) R. O. Duda and P. E. Hart, "Pattern Classification and Scene Analysis", Wiley, New York, N.Y., 1973, pp 69–70.
(12) See, for example, G. H. Stout and L. H. Jensen, "X-Ray Structure Determination", Macmillan, New York, N.Y., 1968, pp 260, 270. The Patterson function is a density function $\rho(xyz)$ which has maxima at points corresponding to interatomic vectors r_{ij}; the amplitudes of the Patterson maxima are proportional to the product of the atomic numbers Z_iZ_j of the corresponding atoms.
(13) On sabbatical leave from Simmons College, Boston, Mass.

Leonard J. Soltzberg,[13] Charles L. Wilkins*
Department of Chemistry, University of Nebraska-Lincoln
Lincoln, Nebraska 68588

Received January 26, 1976

Electron Spin Resonance Spectrum of the Perfluorocyclobutane Radical Anion[1]

Sir:

Perfluorocycloalkanes are known to form negative ions by electron attachment in the gas phase.[2] Also, chemical studies have shown that these molecules compete very effectively against nitrous oxide for the electrons released in the radiolysis

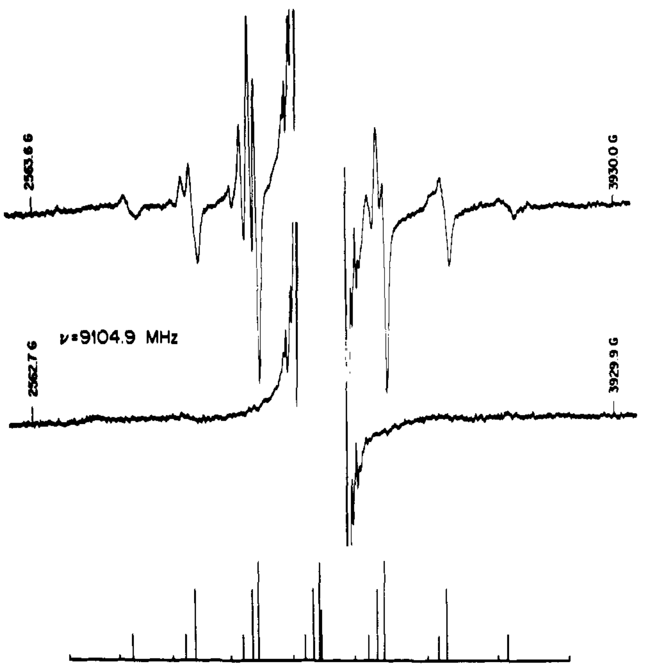

Figure 1. First-derivative ESR spectra of a γ-irradiated solid solution of perfluorocyclobutane in neopentane at 130 K recorded before (upper spectrum) and after (lower spectrum) exposure to unfiltered light from a tungsten lamp. The center portions of the spectra containing intense lines from the neopentyl radical ($\cdot CH_2CMe_3$) have been deleted for the sake of clarity. The line diagram indicates the location of the expected second-order components for the ESR spectrum of the perfluorocyclobutane radical anion possessing eight equivalent fluorines.

of liquid cyclohexane.[3] Taken together with the recent discovery of the hypervalent F_3NO^- radical anion[4] which is isoelectronic with CF_4^-, these facts suggested that radical anions of *saturated* perfluorocarbons might be directly observable in the condensed phase at low temperatures. We now wish to report the ESR spectrum of the perfluorocyclobutane radical anion.

The title radical was prepared by γ irradiation at 77 K of a solid solution containing 5 mol % of perfluorocyclobutane in neopentane. It has been known for some time[5-7] that isotropic ESR spectra are observed for small and highly symmetrical radical species in the plastically crystalline neopentane matrix.[8] More recently, the isotropic ESR parameters for the inorganic radical anions SiF_4^-, SnH_4^-, and PF_5^- were obtained from γ irradiation studies using this matrix.[9,10] Neopentane undergoes a solid–solid phase transition at 130 K[11] with an entropy change of 4.39 cal deg^{-1} and above this temperature the solid clearly possesses the properties of a rotator phase suitable for isotropic ESR studies of trapped radicals.

In the upper spectrum of Figure 1 recorded at 130 K, the hyperfine pattern lying outside the central region is characteristic of a radical possessing eight equivalent ^{19}F nuclei ($I^F = 1/2$) with $a_F = 147.7 \pm 0.6$ G and $g = 2.0021 \pm 0.0017$. This is clearly demonstrated by the detailed second-order structure in the spectrum which is completely resolved for the four $M_I = +1$ components, this latter group being present with the intensity ratios (1:7:20:28) and second-order splittings ($4a_F^2/H$, $3a_F^2/H$, $2a_F^2/H$) predicted by theory.[12] Although the lines are selectively broadened according to the (M_I, I) values to the extent that the weak outermost (± 4, 4) components are undetectable in the present spectrum, the ESR parameters calculated independently from the (± 1, 1), (± 2, 2), and (± 3, 3) components were found to be in excellent agreement. The origin of this type of broadening is well understood and involves the modulation of the anisotropic ^{19}F hyperfine coupling tensor by the tumbling motion of the radical.[13] Not surprisingly, it appears that the rotational correlation time is sufficiently long to permit a substantial dipolar contribution to the line width.[14]

Additional evidence for the integrity of the spectrum and the anionic nature of the species was provided by photobleaching experiments. As shown in Figure 1, all the line components assigned to the perfluorocyclobutane radical anion were removed cleanly by exposure of the sample to unfiltered light from a tungsten lamp. A plausible mechanism for the photobleaching is that of electron detachment, a process which is commonly undergone by trapped anionic intermediates.[15]

The equivalence of the eight fluorines in the ESR spectrum does not necessarily require the perfluorocyclobutane radical anion to assume the planar and highly symmetrical D_{4h} structure. A puckered D_{2d} structure similar to that of the neutral molecule[16] would also be consistent with the results if the "axial" and "equatorial" fluorines were rapidly interchanged on the ESR timescale by ring inversion. Moreover, such an internal motion would greatly contribute to the reduction of anisotropy associated with the fluorine hyperfine tensors.

It is interesting that the a_F value of 148 G for the perfluorocyclobutane radical anion is of the same order as that found for hypervalent inorganic radical anions such as SF_6^-,[17] PF_5^-,[10] SiF_4^-,[9] and F_3NO^- [4] where the unpaired electron occupies an a_1 orbital composed of an antibonding combination of the ligand σ orbitals and the central atom orbitals of the appropriate symmetry.[18] Similarly in the perfluorocyclobutane radical anion, the half-occupied molecular orbital can be regarded as derived from a suitable combination of localized C–F σ^* orbitals. This description also fits in very well with the recent study of the perfluorobenzene radical anion[19,20] for which it was strongly argued[19] that the magnitude of the fluorine coupling (137 G) is more consistent with the occupancy by the unpaired electron of the lowest σ^* orbital rather than of the next available π^* orbital. Further studies on the radical anions of perfluorocycloalkanes and perfluorocycloalkenes are in progress and will be reported in due course.

References and Notes

(1) This work was supported by the U.S. Energy Research and Development Administration (Document No. ORO-2968-100).

(2) For recent papers, see (a) K. M. Bansal and R. W. Fessenden, *J. Chem. Phys.*, **59**, 1760 (1973); (b) F. J. Davis, R. N. Compton, and D. R. Nelson, *J. Chem. Phys.*, **59**, 2324 (1973).

(3) (a) L. A. Rajbenbach, *J. Am. Chem. Soc.*, **88**, 4275 (1966); (b) L. A. Rajbenbach and U. Kaldor, *J. Chem. Phys.*, **47**, 242 (1967); (c) N. H. Sagert, *Can. J. Chem.*, **46**, 95 (1968).

(4) K. Nishikida and F. Williams, *J. Am. Chem. Soc.*, **97**, 7166 (1975).

(5) R. W. Fessenden and R. H. Schuler, *J. Chem. Phys.*, **39**, 2147 (1963).

(6) J. Lin and F. Williams, *J. Phys. Chem.*, **72**, 3707 (1968).

(7) A. J. Colussi, J. R. Morton, and K. F. Preston, *J. Chem. Phys.*, **62**, 2004 (1975).

(8) E. F. Westrum, Jr., *J. Chem. Educ.*, **39**, 443 (1962).

(9) J. R. Morton and K. F. Preston, *Mol. Phys.*, **30**, 1213 (1975).

(10) A. R. Boate, A. J. Colussi, J. R. Morton, and K. F. Preston, *Chem. Phys. Lett.*, **37**, 135 (1976).

(11) (a) J. G. Aston and G. H. Messerly, *J. Am. Chem. Soc.*, **58**, 2354 (1936); (b) A. H. Mones and B. Post, *J. Chem. Phys.*, **20**, 755 (1952).

(12) R. W. Fessenden, *J. Chem. Phys.*, **37**, 747 (1962).

(13) (a) P. Meakin and P. J. Krusic, *J. Am. Chem. Soc.*, **95**, 8185 (1973); (b) K. S. Chen, P. J. Krusic, P. Meakin, and J. K. Kochi, *J. Phys. Chem.*, **78**, 2014 (1974).

(14) Our studies reveal that even for the highly symmetrical SF_6^- radical, there is a small but significant dipolar contribution to the line widths of the ESR spectrum in neopentane. A much more pronounced anisotropic contribution to the line shapes is observed in the spectrum of SiF_4^- in neopentane at 130 K.

(15) See, for example, J. B. Gallivan and W. H. Hamill, *J. Chem. Phys.*, **44**, 1279 (1966).

(16) C. H. Chang, R. F. Porter, and S. H. Bauer, *J. Mol. Struct.*, **7**, 89 (1971).

(17) (a) R. W. Fessenden and R. H. Schuler, *J. Chem. Phys.*, **45**, 1845 (1966); (b) J. R. Morton, K. F. Preston, and J. C. Tait, *ibid.*, **62**, 2029 (1975).

(18) For a discussion of the electronic structure of the hypervalent phosphoranyl radicals, see (a) T. Gillbro and F. Williams, *J. Am. Chem. Soc.*, **96**, 5032 (1974); (b) A. J. Colussi, J. R. Morton, and K. F. Preston, *J. Phys. Chem.*, **79**, 651, 1855 (1975).

(19) L. F. Williams, M. B. Yim, and D. E. Wood, *J. Am. Chem. Soc.*, **95**, 6475 (1973).

(20) In the course of our studies, we have found that the perfluorobenzene radical anion is generated by γ irradiation of a solid solution of C_6F_6 in neopentane at 77 K. The ESR spectrum of $C_6F_6^-$ in neopentane at 130 K shows residual anisotropy as manifested by well-defined line shapes which are charac-

teristic of those observed in powder spectra for axially symmetric hyperfine tensors. From the clearly separated parallel and perpendicular features corresponding to the outermost (±3, 3) components, corrected values of $A_\parallel = 143.3$ G and $A_\perp = 130.3$ G were calculated. The derived isotropic coupling, $a_F = 135$ G, is in satisfactory agreement with the value of 137 G obtained directly from the completely isotropic spectrum of $C_6F_6^-$ in the adamantane matrix at 218 K.[19]

Masaru Shiotani, Ffrancon Williams*
Department of Chemistry, University of Tennessee
Knoxville, Tennessee 37916

Received March 29, 1976

The Role of the Excited Singlet State of Dyes in Dye Sensitized Photooxygenation Reactions

Sir:

The role of the triplet state of dyes in dye sensitized photooxygenation reactions has been the subject of much discussion.[1] There is now no doubt that in many oxygenation reactions the triplet dye acts as a source of energy for the formation of the highly reactive species, singlet oxygen ($^1\Delta_gO_2$). In other oxygenation reactions, the triplet dye has been shown to react with a substrate to generate radicals which subsequently react with oxygen to give oxidation products derived from the substrate. In very few instances has the role of the excited singlet state of the dye been discussed apart from its relevance to the process of intersystem crossing.

Examination of the excited singlet states of common sensitizing dyes such as rose bengal and methylene blue is facilitated by the fact that these dyes fluoresce. Thus, substrate-singlet dye interactions can be quantitatively assessed by measuring the effectiveness of quenching the dye fluorescence by the substrate. We have previously shown that triethylamine quenches the excited singlet state of rose bengal.[2] As can be seen from Table I, many other amines quench the fluorescence of the dyes. In accordance with the view that the quenching occurs via a charge-transfer process, i.e.,

$$\mathrm{Dye}_{S_1} + \mathrm{Am} \rightarrow \mathrm{Dye}\cdot^- + \mathrm{Am}\cdot^+$$

where Am = amine, the lower the ionization potential of the amine, the more effective a quencher it becomes. Another feature of the quenching process is that its efficiency is solvent dependent (see Table II). However, until singlet lifetimes for rose bengal in a variety of solvents have been determined, the solvent dependency of the K_{SV} values cannot be analyzed. However, the K_{SV} values do show how the part played by the excited singlet state of the dyes in the sensitized oxidation of amines is very dependent upon the solvent which is employed.

The K_{SV} values determined for aliphatic amines such as DABCO show that the quenching of the excited singlet state of the dye is important when high concentrations of amines are used ($>5 \times 10^{-2}$ M). It is therefore most important that this fact be borne in mind when amines are being used to test for the participation of singlet oxygen in oxygenation reactions. In order to be certain that the retardation of an oxygenation reaction is not due to any quenching of the excited singlet state of the dye, the amine concentration should be kept below 5×10^{-3} M. Another hazard of using high amine concentrations is that the triplet state of the dye may be appreciably quenched.

Aromatic amines such as *N,N*-dimethylaniline are remarkably efficient quenchers of the fluorescence of the dyes. It is therefore not surprising that indoles are also very efficient quenchers (see Table I). Thus any photooxidation of indoles carried out at $>10^{-3}$ M substrate concentration will involve the excited singlet state of the dye and makes attempts to relate observed products to particular reactive intermediates a precarious business.[4] A further consequence of the quenching process is that the rate of photooxidation of indoles does not markedly increase with increase in indole concentration. In those cases where the dyes become adsorbed to a substrate (e.g., a protein) the quenching process will become very important if a tryptophyl residue is near the site of adsorption.

Table I. Stern–Volmer Constants[a] for Quenching of the Fluorescence of Rose Bengal and Methylene Blue

Substrate	K_{SV} (M) Rose bengal[b]	Methylene blue[b]
N,N-Dimethylaniline	32.5[c]	20.6[c]
Tribenzylamine		5.0[c]
1,4-Diazabicyclo[2.2.2]octane	6[c]	2.6[c]
Triethylamine	1.1[c]	1.0[c]
3-Methylindole	23.0[c]	24.0[c]
Tryptamine	23.0[c]	17.0[c]
Sodium iodide	2.0[d]	14.6,[d] 15.4[e]
Sodium bromide	0.1[d]	0.8[e]
Sodium chloride		0.2[e]
Sodium azide	0.6[d]	3.6[e]
Sodium thiosulfate		3.1[e]
2,5-Diphenylisobenzofuran		57[c]
β-Carotene		57[f]

[a] Calculated from $\phi_0/\phi = 1 + K_{SV}[Q]$ where ϕ_0 = quantum yield of fluorescence in absence of quencher, ϕ = quantum yield of fluorescence when quencher present at concentration [Q], and $K_{SV} = k_q\tau$ where k_q = bimolecular quenching rate constant and τ = fluorescence lifetime of the dye. In all cases aerated solutions were used. Maximum error in K_{SV} is estimated at ±10%. [b] Optical density of dyes = 0.1 at λ_{excit}. [c] Methanol. [d] Methanol/water (50:50 V/V). [e] Methanol/water (20:80 V/V). [f] Methanol–benzene (20:80 V/V), corrections applied for absorption by *β*-carotene at λ_{excit} and λ_{emis} of dye.

Table II. Stern–Volmer Constants for the Quenching of Rose Bengal Fluorescence by Triethylamine in Various Solvents

Solvent	K_{SV} (M)	Solvent	K_{SV} (M)
Methanol	1.1	Acetone	0.4
Methanol-water	1.3	Acetonitrile	~1.5[b]
Water	4–5[a,b]	Ethanol	1.5
Pyridine	0.25		

[a] Minimum amount of methanol added to aid solubility (~1%). [b] Some curvature was observed in the Stern–Volmer plot.

The quenching of the excited singlet states of dyes by anions[5] is particularly relevant to the report[6] by Fenical, Kearns, and Radlick, that addition of azides to photooxygenation reactions of olefins leads to azido–hydroperoxide formation. The azide ion is a powerful quencher of the excited singlet state of methylene blue and, at an azide concentration of 1.0 M, quenching is ~78% efficient. Electron transfer from the anion to the excited dye will produce azide radicals which can subsequently react with the olefin in the presence of oxygen to give the observed products. Gollnick has previously demonstrated the feasibility of the radical anion.[7] The azide ion has also been shown[8] to be a quencher of singlet oxygen. If therefore it is to be used for testing for the intermediacy of singlet oxygen by adding varying amounts of the anion to a reaction, its concentration must be kept relatively low.

Azide ions, like halogen anions quench the fluorescence of rose bengal far less efficiently than methylene blue fluorescence. This lower quenching efficiency can be attributed to the fact that the dye molecule is an anion and consequently the quenching process requires the close approach of two anions, i.e., the development of an unfavorable Coulombic interaction.

The excited singlet state of methylene blue is efficiently quenched by *β*-carotene and 2,5-diphenylisobenzofuran. *β*-Carotene, like DABCO, has been proposed as being a suitable quencher for use in a diagnostic test for the participation of