

Display of quantum mechanical properties on van der Waals surfaces

P Quarendon, C B Naylor*, and W G Richards*

IBM UK Science Centre, Athelstan House, St Clement Street, Winchester, Hampshire, UK

*Physical Chemistry Laboratory, South Parks Road, Oxford, OX1 3QZ, UK

A technique is described for comparing graphically the electrostatic properties of molecules. Results from quantum mechanical calculations are mapped onto the van der Waals surface. The surface is defined by CSG (constructive solid geometry), a technique used in CAD (computer-aided design). Colour is used to represent the electrostatic properties and the results are displayed on a high resolution raster device as a 'realistic' model. This technique is expected to be of value in discriminating qualitatively between compounds of different biological activities.

Keywords: constructive solid geometry, van der Waals surface, raster display, colour mapping, quantum mechanical properties

received 21 November 1983, revised 19 December 1983

The differing reactivities and biological activities of a series of similar molecules depend in part on differences in the electronic properties of the series' members. These electronic properties should be susceptible to quantum mechanical calculation, but standard molecular orbital computer programs usually produce so great a numerical output that the results are not particularly helpful to a non-specialist. An obvious role for computer graphics is to display these data in a simplified way so that an experimental chemist may discriminate between the electronic features much more easily.

There have been other attempts to draw the results of molecular orbital calculations on a defined molecular surface¹, but these have generally used vector systems which produce the picture as a series of dots and require a molecular skeleton to be superimposed so that the user can identify particular molecules and interpret a fairly complex picture. The intention here is to use a raster system with colour shading and, most important, to present the molecular surface so that it looks like the CPK (Corey, Pauling, Koltun) space-filling model that is familiar to working chemists.

We want to be able to display a picture of a molecule which is recognizable to a chemist and to code the electronic information on the surface. The surface can be defined as in a CPK model by taking atoms as spheres of prescribed radius. The molecular properties we have chosen to display are the obvious and familiar electron density and electrostatic potential. The latter is the energy of interaction of a free proton with the molecule, an indication of electrophilic attack. Since the response to electrophiles will depend not only on the density of electrons at any point but also on how strongly they are held, we have also included calculations and pictures of 'super-delocalizabilities'. These are the electron densities divided by the orbital binding energies summed over all the occupied molecular orbitals.

As a model series known to exhibit variation in electronic properties we have chosen a series of substituted toluenes.

QUANTUM MECHANICAL CALCULATIONS

All the wave-functions were calculated with the standard GAUSSIAN'70 program² using the internal STO-3G basis set. Standard bond lengths and angles³ were employed throughout. A modified version of the supplementary DENPOT program⁴ was then used to compute the three electronic properties described below.

Electron density The electron density at any point P surrounding a molecule is defined as the square of the molecular wave-function at that point:

$$\rho(P) = \sum_i n_i |\phi_i(P)|^2$$

where n_i is the number of electrons in the molecular orbital ϕ_i .

Electrostatic potential The electrostatic potential at a

point P' surrounding a molecule is a measure of the electrostatic energy between that molecule and a unitary positive point charge placed at P' . It is the sum of two terms, an electron-point charge attraction and nucleus-point charge repulsion:

$$V_e(P') = - \frac{\rho(P')}{r_{pp'}} + \sum_{\alpha} \frac{Z_{\alpha}}{r_{p'\alpha}}$$

where Z_{α} is the charge of nucleus α , $r_{p'\alpha}$ is the nuclear-point charge distance, $\rho(P)$ is the electron density at a point P and $r_{pp'}$ is the distance P - P' . The electronic attraction term is integrated over all space, the nuclear term is summed over all nuclei.

Super-delocalizability The super-delocalizability⁵ at a point P is a measure of how tightly the electrons are held by the molecule at that point. It is the electron density at a point P due to orbital i , divided by the orbital energy, ϵ_i , summed over all orbitals:

$$SD(P) = - \sum_i \frac{n_i |\phi_i(P)|^2}{\epsilon_i}$$

DISPLAY OF PROPERTIES

The method used to construct the pictures is shown schematically in Figure 1. The major inputs are shown at the left of the diagram and the main processing components are shown in the centre. The data flow is indicated by arrows.

CSG model

The coordinates of the atoms in the molecule and the atomic radii to be assumed are fed to a small program which generates a description of the molecular geometry in CSG (constructive solid geometry). CSG is becoming familiar in CAD as a method of describing solid objects. For a review, see eg Requicha and Voelcker⁶. Briefly, an object is described by choosing a number of primitive objects, typically cylinders and rectangular blocks, and combining these with the set

operations union, intersection and difference. To construct a molecular model with a van der Waals surface, only spherical primitives are needed to represent the atoms, and the molecular shape is created by combining these with set-union operations. More complex surfaces, such as the solvent-accessible surface, require more primitives (in particular a torus) and a set-difference operation.

The CSG model of the molecule is stored as a file where it can be used to generate views of the molecule with standard colours to represent the different atom types in addition to the maps of electrical properties described here.

Surface generation

The file describing the surface geometry is read, checked for errors, and an internal data structure in the form of records linked by pointers is created to represent the molecular geometry. This data-structure is passed to the surface-generating program.

The surface is produced from its description by a method called 'spatial subdivision'. A cubic region of space is considered which contains the entire molecular model. The algorithm generates the view on one surface of the cube. To do this, the region is divided into eight smaller cubes, by bisecting each of the edges of the edges, and each of these new regions is inspected. Any which are empty are discarded and any region which is completely within an object is similarly ignored. Those which remain contain the surface of at least one object and are subdivided into eight further cubic regions. Again full and empty regions are discarded and subdivision is continued until the cubes each correspond to a single screen pixel. As all the remaining cubes contain part of the surface, the result is an approximation to the surface at pixel resolution.

In order to produce a picture it is not necessary to generate the entire surface, but only that part which is visible from the chosen view point. Therefore, cubic regions which are at the front of model space (those nearest the viewer) are inspected before those at the back. A map is kept, in the form of a binary tree, of areas of the screen which have been shaded. This allows any regions which are obscured by earlier shading to be ignored and so be eliminated from the image.

The basic algorithm is derived from work by Woodward and Quinlan⁷ and is described more fully by Quarendon⁸.

Computing property values

The value of the required property is computed at the centre of all the visible pixel volumes generated. A modified version of the DENPOT program is used. DENPOT takes as input the eigenvalues and coefficients generated by GAUSSIAN'70 to describe the electron wave-function of the molecule. The program was modified, first to add delocalizability, and then to allow each of the three electronic properties to be calculated at a single point in space using a function call. The Fortran subroutines were called from the main program which is written in Pascal.

For the size of molecule being considered (with more than 70 orbitals) calculating the electronic properties

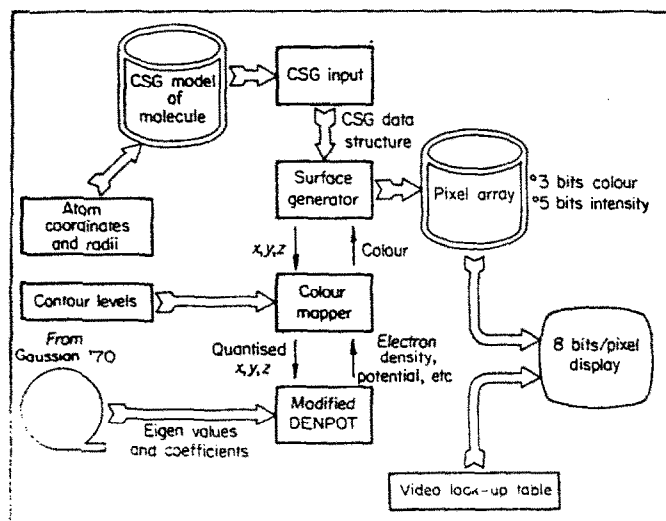


Figure 1. Schematic diagram of the method used to construct pictures

from the wave-functions is computationally quite expensive, varying from 0.1 to 1 second for each point. It is particularly expensive to calculate electrostatic potential. On the machine being used, an IBM 4341, this represents up to two million instructions and dominates the construction times for the pictures.

It would be prohibitively expensive to compute accurately the value at each pixel location. Therefore, the values are computed only on a regular 3D grid of points. Values on this grid are computed on demand. That is, when the value at a point is needed the eight nearest grid points are determined. These are at the corners of a cube. The values at these eight points are computed, using the modified DENPOT, and the results are saved in a 3D array corresponding to the grid. A subsequent evaluation requiring one or more of these grid values will use the stored value and will not require computation. The value at each point of interest is then found by linear interpolation from the eight neighbouring grid points.

The density of grid points was determined experimentally. Too low a density results in an inaccurate picture. Too high a density results in unnecessary storage and computing. The pictures shown were generated using a 0.2 Å grid. This involved evaluating the properties using DENPOT at 5000 points in a typical case. Doubling the grid size reduces the number of evaluations to one quarter but some of the detailed variations are not recorded.

Colour mapping

The pictures were designed to be shown on a display device with 8 bits/pixel. This permits 256 colour combinations to be shown simultaneously. To give a smooth gradation of intensity, these are divided into 32 intensity levels of eight colours. The computed values are therefore reduced to eight contour levels. Contour levels are defined by a table and in the examples are spaced equally between chosen minimum and maximum values.

The surface brightness on all pictures simulates that of a diffuse reflector. Specular reflection can also be simulated, but tends to obscure the colours in the region of the highlights and is not used in any of the examples. The intensity of each pixel is then simply the cosine of the angle between the surface normal and the direction assumed for the light source. In the examples the light is coming from a direction close to that of the eye.

The last control over the appearance which is available is through the video look-up table in the display device. With this, the colour used to represent each contour level can be chosen independently. Most experiments have used colours in spectral order so that the hues from the HSV colour cone⁶ are used in sequence to represent values.

RESULTS

Colour plate 1 (see page C1) shows a selection of results for the three computed properties on surfaces drawn for toluene; its hydroxy, chloro and methoxy derivatives. The colour code runs from red through to blue in eight divisions with in each case red

corresponding to the lowest or negative value and blue to the highest or positive. Thus in electron density terms the blue features denote areas of high density. The scaling chosen results in no very low electron density being indicated at these distances which are relatively close to the nuclei. For electrostatic potentials the full range of colour is observed with red regions being negative which corresponds to the attraction of a proton while blue denotes repulsion. Finally, the super-delocalizability similarly has regions from which electrons may be removed or combined with electrophiles denoted by colour at the blue end of the spectrum.

DISCUSSION

The series chosen does enable one to distinguish otherwise very similar molecules. The results should not be taken to indicate the rates of aromatic substitution since these are more properly a function of transition state stability rather than unperturbed electronic properties. However, the molecules are distinguished so that with development this scheme could be used for qualitative discrimination between active and inactive members of a series.

It must be emphasised that the choice of numerical range of any property and the corresponding assignment of colour is arbitrary. Some researchers prefer the opposite colour convention (red = positive, blue = negative) and look-up tables are available which permit the use of either convention. In any particular case some experimentation with the range is probably essential if one is to highlight just what is the best discriminator of activity as a function of electronic structure. The choice of molecular surface is also arbitrary and could lead to the neglect of aspects of the electrostatic potential at distances further from the molecule. The aim here is to produce a simple representation for comparative purposes.

We intend now to apply this system to a number of structure activity problems where conventional studies in terms of linear free energy relationships have proved inadequate.

ACKNOWLEDGEMENTS

C B Naylor is indebted to the Lord Dowding Fund for a research grant.

REFERENCES

1. Langridge R, Ferrin T E, Kuntz I D and Connolly M L *Science* No 211 (1981) pp 661-666
2. Hehre W J, Lantham W A, Ditchfield R, Newton M D and Pople J A 'Gaussian 70 Program' *QCPE* 236
3. Sutton L E *Tables of interatomic distances and configuration in molecules and ions* Special Publication, The Chemical Society, London No 11 (1958)
4. Peeters D, Sana M 'Denpot Program' *QCPE* 360
5. Fukui K, Yonezawa T, Nagata C *Bull. Chem. Soc. Jpn* No 27 (1954) p 423
6. Requicha A A G and Voelcker H B 'Solid modelling,

- a historical summary and contemporary assessment' *IEEE Comput. Graphics Appl.* Vol 2 No 2 (March 1982)
- 7 **Woodwark J R and Quinlan K M** 'The derivation of graphics from volume models by recursive subdivision of the object space' *Proc. Comput. Graphics 80 Conf.* (1980)
- 8 **Quarendon P** Paper submitted to J. Mol. Graphics
- 9 **Foley, J D and Van Dam, A** *Fundamentals of interactive computer graphics* p 613