

Visualizing properties of atomic and molecular systems

Ken Flurchick and Lee Bartolotti

North Carolina Supercomputing Center, Research Triangle Park, NC 27709, U.S.A.

In this article scientific visualization techniques that allow for the interactive computation and display of molecular structures and properties are presented. The electronic properties are computed and displayed using the application visualization system (AVS). By combining the visualization techniques of AVS with an orbital description of molecular systems, a better understanding of the electronic behavior of the systems can be achieved. Molecular properties such as the electrostatic potential and chemical reactivity index (the Fukui function) are computed and mapped onto isodensity displays of the electron density of the system.

Keywords: scientific visualization, density functional theory, AVS

INTRODUCTION

Scientific visualization is a powerful method to display the results of computer simulations and/or experimental measurements. This method allows for all the data to be displayed at one time, thereby showing relationships among the data that are difficult to determine otherwise. One visualization system that has been effective in displaying the results of electronic structure computations is the AVS, the application visualization system from Advanced Visual Systems, Inc. (Waltham, MA).¹ The AVS is an object-oriented approach to scientific visualization and the objects that AVS manipulates are the common graphic operations, namely, input, output, filters, and mappers. Input and output operations get data into and out of the AVS. Filters take data to data, either a subset of the original data set or data that are operated on. Mappers take data to geometric objects for display. The AVS is a flexible and extensible visualization system that can manipulate and display many types of systems and, in particular, molecular systems. The AVS

can display geometric structure data and various electronic properties that can be combined to extract specific chemical information about molecular systems. In this article we present a set of modules developed to read into the AVS the geometric structure and electron density data and, using the supported AVS modules, manipulate the data to compute and display physical properties.

Electronic and structural information can be obtained from either density functional theory or Schrödinger theory. Using various electronic structure programs, electron densities are computed as well as energy-optimized geometric structures. Input of the data into the AVS is accomplished using different input data readers as well as some specific input modules written to read data from some of the programs. Combining these modules with AVS built-in functions, additional molecular properties are determined and displayed. In addition, displaying the properties mapped onto the geometric structure shows how the electronic properties relate to chemical activity.

APPLICATION VISUALIZATION SYSTEM

The Network Editor, which is the primary AVS environment, is shown in Color Plate 1. The top section shows the list of modules. These modules are organized in columns of *data input*, *filters*, *mappers*, and *data output*. Below the module list is the work area where networks of modules are constructed. Networks like the one shown in Color Plate 1 were used to create the figures in this article, and are straightforward to construct. Modules are "dragged" from the module list into the network work area and, following a "data flow" paradigm, modules are connected by the users. The modules communicate with each other via different-colored input and output ports. Input ports are on top of the module and output ports are on the bottom. The different colors of the input and output ports represent different AVS data types. The fundamental data type for AVS is called a field and is depicted by blue ports. Colormap data are yellow, geometries (data to be rendered) are red, and integers are pink. The AVS data types and module descriptions are in the AVS manuals distributed by AVS.¹

To elucidate the connection between the colored ports on each module and the network, we discuss here the *isosurface* module. There are three input ports: (1) AVS field

Color plates for this article are on p. 50.

Address reprint requests to L. Bartolotti, North Carolina Supercomputing Center, MCNC 3021 Cornwallis Road, Research Triangle Park, Durham, NC 27709, U.S.A.

Received 23 November 1994; accepted 12 May 1994.

three-dimensional (3D) scalar for input data, (2) AVS field 3D scalar for optional colormap information, and (3) colormap to map field data to color values. In addition, a single output port, colored red to denote geometry, will output a geometry to be displayed.

The data for the visualization is processed in the following manner. The *display dmol* and *coord to geom* read input files from output of DMol. These two modules were developed by the authors to display the results of calculations using DMol from BIOSYM.² The *coord to geom* module processes the molecular structure information and generates an AVS geometry. The *display dmol* module reads the molecular orbital information and generates a three-dimensional AVS field. The field data are then passed to the *field to mesh* and *isosurface* modules to create the visualization. The modules included with the AVS can manipulate data in a variety of ways. There are many modules that can manipulate the AVS field data. For the work here, using the *field math* module, orbital data can be operated on; an example would be computing the square of the orbital.

MOLECULAR PROPERTIES

Many molecular properties can be determined from the electron density, which contains both quantitative and qualitative information about the system of interest. The electron density plays an important role in the visualization of atomic and molecular properties, such as molecular bonding, reactivity indices, electrostatic potentials (ESPs), and aromaticity. Electron densities can be computed using a number of available software packages. Some of the currently available software packages that compute electron density include DMol,² SPARTAN,³ MOPAC from QCPE⁴ and Gaussian92.⁵ These codes compute the electron density using either density functional theory (DFT) via the Kohn-Sham (KS)⁶ orbital density equations or wavefunction functional theory via the Schrödinger equation. For additional information about the different methodologies, see Szabo and Ostlund,⁷ Parr and Yang,⁸ and Lipkowitz and Boyd.⁹ MOPAC and DMol are used in the current work.

The visualization process may begin by reading the electron density information and other output data from the calculation and converting it to AVS field format. The AVS field format is an internal data format for the AVS. The electron density is then displayed as an isosurface, which is a three-dimensional surface of constant value. The isosurface may be overlaid on top of a figure of the geometry of the molecule and can be made transparent or displayed as a wireframe structure. The geometric display can be a stick, ball and stick, or CPK space-filling display. This is a standard type of display that can be done in the AVS using ChemViewer from MSI¹⁰ or CONVEXAVS from Convex.¹¹ Various molecular modeling packages, for example, SYBYL,¹² UNICHEM,¹³ QUANTA,¹⁰ INSIGHT,¹⁴ SPARTAN³ and others can also display this type of visualization. The AVS modules developed for this article are available from the anonymous ftp site for AVS.¹⁵

CHEMICAL REACTIVITY

In this work, the focus is on visualizing the electron density data to help elucidate reactive sites in a molecule. In addition,

other molecular properties are mapped onto the density to emphasize reactive sites.

The electrostatic potential $V(\mathbf{r}_1)^{16-18}$ is widely used to predict reactive sites in molecules. Regions of high negative charge represent sites where electrophilic attack is likely to occur, whereas sites of high positive charge reflect sites with a propensity of nucleophilic attack. The electron density can be obtained from the sum of the squares of the molecular orbitals, ϕ_i , namely,

$$\rho(\mathbf{r}_1) = \sum_{i=1}^N n_i |\phi_i(\mathbf{r}_1)|^2 \quad (1)$$

where N is the number of electrons in the molecule and n_i is the number of electrons in the i^{th} molecular orbital.

Frontier orbital theory⁸ is another method for predicting chemical reactivity. Density functional theory has provided additional insight into frontier orbital theory. In density functional theory, the Fukui reactivity index is defined as

$$f(\mathbf{r}_1) = \left[\frac{\partial \rho(\mathbf{r}_1)}{\partial N} \right]_{\nu(\mathbf{r}_1)} \quad (2)$$

where $\nu(\mathbf{r}_1)$ is the external potential (the nuclear-electron attraction potential in many instances). The quantity $f(\mathbf{r}_1)$ measures the spatial sensitivity of the system's chemical potential (electronegativity) to external perturbations. The derivative $(\partial \rho / \partial N)_{\nu}$ for some integral value of N is discontinuous. This discontinuity has an important implication. The left-hand derivative, $f_-(\mathbf{r}_1)$, measures the reactivity toward an electrophile, the right-hand derivative, $f_+(\mathbf{r}_1)$, measures reactivity toward a nucleophile, and their average, $\frac{1}{2}[f_-(\mathbf{r}_1) + f_+(\mathbf{r}_1)]$, measures reactivity toward radical attack. These derivatives can be written in terms of the molecular orbitals:

$$f_-(\mathbf{r}_1) = |\phi_{\text{HOMO}}(\mathbf{r}_1)|^2 + \sum_{i=1}^{N-1} n_i \frac{\partial}{\partial N} |\phi_i(\mathbf{r}_1)|^2 \quad (3)$$

$$f_+(\mathbf{r}_1) = |\phi_{\text{LUMO}}(\mathbf{r}_1)|^2 + \sum_{i=1}^N n_i \frac{\partial}{\partial N} |\phi_i(\mathbf{r}_1)|^2$$

In Eq. (3) HOMO refers to the highest occupied molecular orbital whereas LUMO refers to the lowest occupied molecular orbital. In many instances, the relaxation term in Eq. (3), $(\partial / \partial N) \phi_i(\mathbf{r}_1)^2$, can be ignored and the theory as first put forth by Fukui and others¹⁹⁻²² is recovered.

RESULTS

The distribution of electronic charge in a molecular system, displayed in conjunction with the molecular properties described above, provides insight into the chemical behavior of the molecular system. In Color Plates 2 and 3 the properties of interest are mapped onto an isosurface of the electron density, which totally encloses the van der Waals volume of the atoms in the molecules. The atom color scheme, for this and other structures, is white for hydrogen, green for carbon, red for oxygen, and blue for nitrogen. The AVS

tube module generated the tubular bonds. This mapping of properties may accentuate the reactive sites in the molecule.

The first molecule investigated is α -naphthalyl. The geometry and resulting orbitals were calculated using DMol, which solves the orbital density equations proposed by Kohn and Sham.⁶ The exchange-correlation energy density functional employed in this version of DMol is due to von Barth and Hedin.²⁶ DMol writes the optimized geometry, the selected MOs, the density $\rho(\mathbf{r}_1)$, and the electrostatic potential to separate files. The MOs, $\rho(\mathbf{r}_1)$, and the electrostatic potential are determined at points in a cube containing the molecule. AVS modules to read the geometry and the electronic data have been written to get the data into the AVS and to display structures. The optimized structure for α -naphthalyl is displayed in Color Plate 2. The second quantity displayed in Color Plate 2 is the electrostatic potential. The value of the ESP mapped onto the density isosurface is color coded from blue (most negative) to red (most positive). From this ESP picture, it is difficult to see that the position para to the NH_2 group is most susceptible to electrophilic attack.²⁵ The third object displayed in this figure in the HOMO density. The AVS *field math* module was used to generate the square of the HOMO. This quantity is the frontier orbital, as proposed by Fukui, which predicts the site(s) in the molecule susceptible to electrophilic attack. Here, red represents the most positive value of the frontier orbital and the most likely point of attack. Indeed, it does predict the para position as most reactive. The last object is the DFT Fukui function $f_-(\mathbf{r}_1)$, defined by Eq. (4). The left-hand derivative was calculated by finite difference, using the densities for the N and $N - 1$ electron systems,

$$f_-(\mathbf{r}_1) = \rho_N - \rho_{N-1} \quad (4)$$

The *field math* module calculated this difference. Again we see that $f_-(\mathbf{r}_1)$ (most positive value) correctly predicts the para position as being most susceptible to electrophilic attack.

The next molecule investigated is squalene, and the results are shown in Color Plate 3. This molecule is a precursor to lanosterol, the first intermediate in mammalian cells with the sterol nucleus.^{27,28} The calculations were performed with MOPAC using the PM3 Hamiltonian. The CONVEX AVS *read MOPAC* module was used to read in the graphical information. The coiled structure shown is the lowest energy configuration. Here we see that all three predict that electrophilic attack occurs at the double bonds located at the ends of the coil. However, both the ESP and the HOMO density seem to predict all double bonds as equally reactive to electrophilic attack.

CONCLUSION

We have taken advantage of the power and versatility of the AVS to display calculated properties of chemical systems. Because the AVS allows the data to be handled in many ways, we were able to select a visual display of some property that enhanced our understanding of that property. The current work used these visualization techniques to identify reactive sites in a series of molecules. Although we looked at only a few systems, this method for investigating chem-

ical reactivity should be widely applicable. This approach can be useful in educational as well as research environments.

REFERENCES

- 1 Upsom, C., Faulhaber, T., Jr., Kamins, D., Laidlaw, D., Schlegel, D., Vroom, J., Gurwitz, R., and van Dam, A. *The Application Visualization System: A Computational Environment for Scientific Visualization*, IEEE Computer Graphics and Applications, 9(#4), 30, 1989 AVS is available from Advanced Visual Systems, Inc., 300 Fifth Avenue, Waltham, MA 02154, U.S.A. (617)890-4300
- 2 Delly B. J. *Chem. Phys.* 1990, **92**, 508; DMol is available from BIOSYM Technologies, Inc., 9685 Scranton Road, San Diego, CA 92121, U.S.A.
- 3 SPARTAN version 3.0, Wavefunction, Inc., 18401 Von Karman #370, Irvine, CA 92715, U.S.A.
- 4 Stewart, J. MOPAC. *J. Computer-Aided Mol. Design* 1990, **4**; Quantum Chemistry Program Exchange #455
- 5 Frisch, M.J., Trucks, G.W., Head-Gordon, M., Gill, P.W., Wong, M.W., Foreman, J.B., Johnson, B.G., Schlegel, H.B., Robb, M.A., Replogle, E.S., Gomperts, R., Andes, J.L., Raghavachari, K., Binkley, G.S., Gonzalez, C., Martin, R.L., Fox, D.J., Duprees, D.J., Baker, J., Stewart, J.J.P., and Pople, J.A. *Gaussian92*, revision A. Gaussian, Inc., Pittsburgh, PA, U.S.A., 1992
- 6 Kohn, W. and Sham, L.J. *Phys. Rev.* 1965, **140**, A1133
- 7 Szabo, A. and Ostlund, N. *Modern Quantum Chemistry*. Macmillan, New York, 1982
- 8 Parr, R. and Yang, W. *Density-Functional Theory of Atoms and Molecules*. Oxford University Press, New York, 1989
- 9 Lipkowitz, Kenny B. and Boyd, Donald B. (Eds.). *Reviews in Computational Chemistry*. VCH, New York, 1992
- 10 ChemViewer, QUANTA. Molecular Simulations, Inc., 16 New England Executive Park, Burlington, MA 01803, U.S.A.
- 11 CONVEXAVS. M.A.R. Richardson, Convex Computer Corporation, 3000 Waterview Parkway, P.O. Box 833851, TX 75038, U.S.A.
- 12 SYBYL. Tripos Associates, 1699 S. Hanley Road, Suite 303, St. Louis, MO 63144, U.S.A.
- 13 UniChem is from Cray Research, Inc., 655E Lone Oak Drive Eagan, MN 55121, U.S.A.; UniChem is a registered trademark of Cray Research, Inc.
- 14 INSIGHT; available from BIOSYM Technologies, Inc., 9685 Scranton Road, San Diego, CA 92121, U.S.A.
- 15 AVS anonymous ftp site avs.ncsc.org.
- 16 Scrocco, E. and Tomasi, J. *Topics Curr. Chem.* 1973, **42**, 95
- 17 Tomasai, J. In: *Chemical Applications of Atomic and Molecular Electrostatic Potentials* (P. Politzer and D.G. Truhlar, Eds.), Plenum Press, New York, 1981, pp 257-294
- 18 Orazco, M. and Inque, F.J. *Comp. Chem.* 1993, **14**, 587

- 19 Fukui, K., Yonezawa, T., and Shingu, H. *J. Chem. Phys.* 1952, **320**, 722
- 20 Fukui, K., Yonezawa, T., Nagata, C., and Shingu, H. *J. Chem. Phys.* 1954, **22**, 1433
- 21 Fukui, K. *Theory of Orientation and Stereoselection*. Springer-Verlag, Berlin, 1975
- 22 Fukui, K. *Science* 1987, **218**, 747
- 23 Additional data were generated by Dr. Art Edwards, Department of Electrical Engineering, University of North Carolina at Charlotte, Charlotte, NC 28202, U.S.A.
- 24 Yang, W. and Parr, R. *Proc. Natl. Acad. Sci. U.S.A.* 1985, **82**, 6723
- 25 Flurry, R.L. *Molecular Orbital Theory of Bonding in Organic Molecules*. Marcel Dekker, New York, 1968, pp 79–81
- 26 von Barth, U. and Hedin, L. *J. Phys. Chem.* 1972; **5**, 1629
- 27 Robinson, R. *The Statistical Relations of Natural Products*. Clarendon Press, Oxford, 1955, pp 18–20
- 28 Light, R.L. *A Brief Introduction to Biochemistry*. W.A. Benjamin, New York, 1968, p 80