Computer-Generated Space-Filling Molecular Models

GRAHAM M. SMITH* and PETER GUND

Merck Sharp & Dohme Research Laboratories, Rahway, New Jersey 07065

Received May 15, 1978

INTRODUCTION

Many chemical and physical properties of a molecule can be related to its three-dimensional structure. However the chemists' two-dimensional structural diagrams do not convey very accurate geometrical information. Mechanical molecular models, such as Dreiding models, also have limitations. For example, strained molecules are difficult to construct; flexible molecules often have several possible geometries; and it is difficult to compare molecules of different but related structures. The Merck Molecular Modeling System (MMMS)[1] has been created to assist medicinal chemists in perceiving structural relationships in drug molecules.

MMMS is an integrated set of computer programs, intended to be used directly by the chemists, for creating and manipulating three-dimensional molecular structues. Molecules may be constructed from internal coordinates (bond lengths and angles, dihedral angles) or fractional coordinates when an x-ray structure is available. Once created, a structure may be manipulated and viewed via interactive computer graphics, it may be subjected to classical strain energy minimization, and quantum mechanical calculations (CNDO, INDO) may be performed.

MMMS graphics programs initially all represented molecules as "stick" models with stereoscopic views to impart three-dimensional information. While stick models are simple to program and easily recognized by the chemist, they fail to convey the space-filling characteristics of atoms. Several workers[2,3] have reported programs which give such a space-filling representation. By removing hidden lines, three-dimensional information can be obtained from a single two-dimensional image.

In order to add this capability to MMMS, we obtained a copy of the program written by Professor P. K. Warme[2] and implemented it on our computer. The results from this program elicited a great deal of interest from chemists; however, the program has a major limitation. In order to simplify the equations, the program requires all atoms to have the same radius. The shapes of molecules generated by the program are thus usually unrealistic.

To overcome this limitation we had to redesign and reimplement Warme's program. The resulting improved program, named SPACEFIL, and some applications are described in this paper.

PROGRAM DESCRIPTION

The basic program organization, shown in the flow chart (Figure 1), is similar to that in Warme's program.[2] The space-filling representation consists of two picture elements: the great circle of each sphere drawn perpendicularly to the viewing axis, and the circle of intersection for all overlapping atoms. Each of the picture elements is derived from a standard circle of 24 line segments by scaling, rotation, and translation. A more detailed description of the main routines and a listing of the program appear in the Appendix.

All major routines of Warme's program had to be rewritten to allow scaling of the spheres to different radii. The changes are centered in the routines which find the center of a circle of intersection, its radius, and two orientating angles. The derivation of the equations is given in the Appendix.

SPACEFIL automatically selects standard[5] van der Waals radii for each atom type; a stereoscopic view of 5-methyl-5-ethyl-5, 6-dihydro-2(1*H*)-pyridone[4] is shown (Figure 2a). An experimental version of SPACFIL can mark spheres with great circles rotated around the x axis and around the y axis to give curved cross hatching (Figure 2b) or with contour circles (Figure 2c). Each method of marking can be used to differentiate atom types or groups, as in Figure 2d, where the cross hatching indicates a nitrogen and the contour circles indicate an oxygen.

The SPACFIL program has now been incorporated into MMMS and is being used by Merck chemists at laboratories in Rahway, N.J., West Point, Penn., and Montreal, Canada.

The original version of this program used a Tektronix 4010-1 Display, with the PLOT10 graphics package.[6] The program was run on our IBM 370/168 under the TSO Operating System, with a 1200 baud telephone link. Since this system used a storage tube display, very large structures (up to about 1300 atoms) could be drawn. The only limit to structure size was the clutter from drawing too many lines on the screen.

Once developed, the program was modified by our MIS Research Systems and Programming Group for our new Digital Equipment Corp. GT42-GT43 Display terminals, using the GT40L graphics package.[7] These displays have the advantage of a larger and brighter screen; however, they have the disadvantage that more than about 70 atoms causes flicker. This problem can be partly overcome by photographing the display with a Polaroid camera. Despite flicker, stereopairs of large molecules can be photographed and viewed with stereoscopic viewers. More recently MIS Systems and Programming has incorporated the ability to reproduce the displayed image on a Calcomp Plotter.

PROGRAM APPLICATIONS

The program is being used for an increasing number of studies at Merck, some of which are summarized here.

Holography[8] is perhaps the best method available for the representation of three-dimensional molecular structure, because it requires no special viewing glasses or motion to impart depth to the image. Creating an "Integral Hologram" of Merck's new analgesic, Dolobid® (Figure 3), was the first successful application of SPACFIL and was accomplished as follows.

The Dolobid structure was created from standard bond lengths and angles, and CNDO/2 calculations were used to determine the preferred twist angle between the rings. The SPACFIL program was modified by our systems and programming group so that it would accept a constant rotation angle for the y axis and display a new rotated image on command. Each of these rotated images was photographed with a modified Bolex movie camera, one frame at a time. The resulting film showed one complete (360°) rotation of the molecule in 1080 frames. This film was then processed holographically[9] to yield a cylindrical transparency. Illumination of this transparency by a point source of light from inside and below causes a three-dimensional image of the structure to appear suspended in the center of the cylinder. This hologram was exhibited at a Clinical Congress[10] in London in October 1977.

0095-2338/78/1618-0207$01.00/0

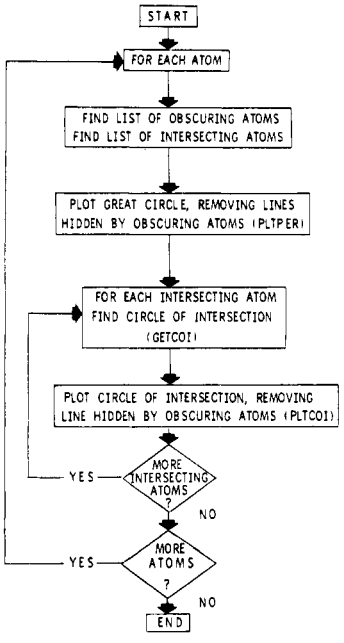

Figure 1. Flow chart of basic program organization.

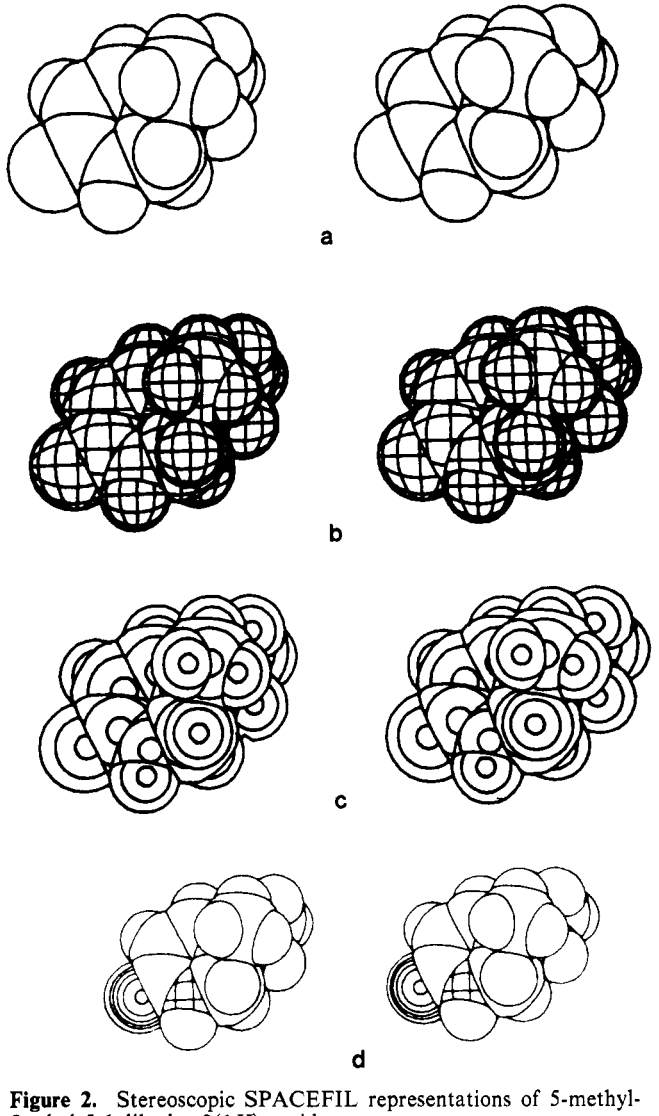

Figure 2. Stereoscopic SPACEFIL representations of 5-methyl-5-ethyl-5,6-dihydro-2(1*H*)-pyridone.

The program is also useful for visualizing the steric bulk of a molecule. For example, in the stick figure representation of the crystal structure of Valinomycin complexed with potassium ion[11] (Figure 4a), it is difficult to perceive to what degree the ion is enveloped within the cavity and isolated from the environment.

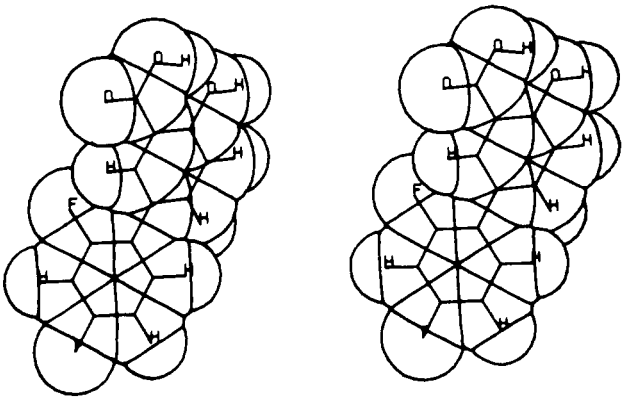

Figure 3. Stereoscopic view of Dolobid.

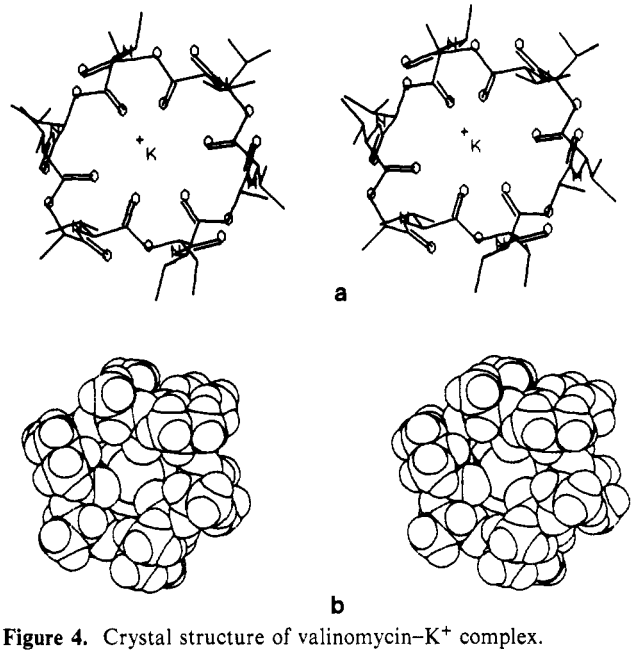

Figure 4. Crystal structure of valinomycin–K^+ complex.

However, in the SPACFIL representation (Figure 4b) it is clear that the potassium ion is deeply imbedded in the cavity. Also visible are three of the six carbonyl oxygens which coordinate to the ion in the cavity. The stick figure (Figure 5a) of the uncomplexed Valinomycin crystal structure[12] shows that, when an ion is not present, the molecule flattens. In addition to this, it can be seen from the space-filling image (Figure 5b) that the cavity becomes completely closed.

Similarly, a space-filling representation of the crystallographically determined dihydrofolate reductase[13] active site gives an idea of the shape of the cavity.

A slightly modified version of the program takes advantage of the variable radius capability to represent types of data other than atomic radii. For example, quantum chemically derived net electron densities or frontier orbital coefficients may be represented. Thus, Figure 6 shows this CNDO/2 derived data for nicotinic acid (cross-hatching indicates a negative value). The radii have been automatically scaled to eliminate overlapping circles.

Merck is routinely solving the crystal structures in-house of interesting organic compounds, such as the new antibiotic Cefoxitin.[14] As these structures are solved, they are normally placed on the MMMS so that interested chemists may examine and manipulate them. In this context the SPACFIL

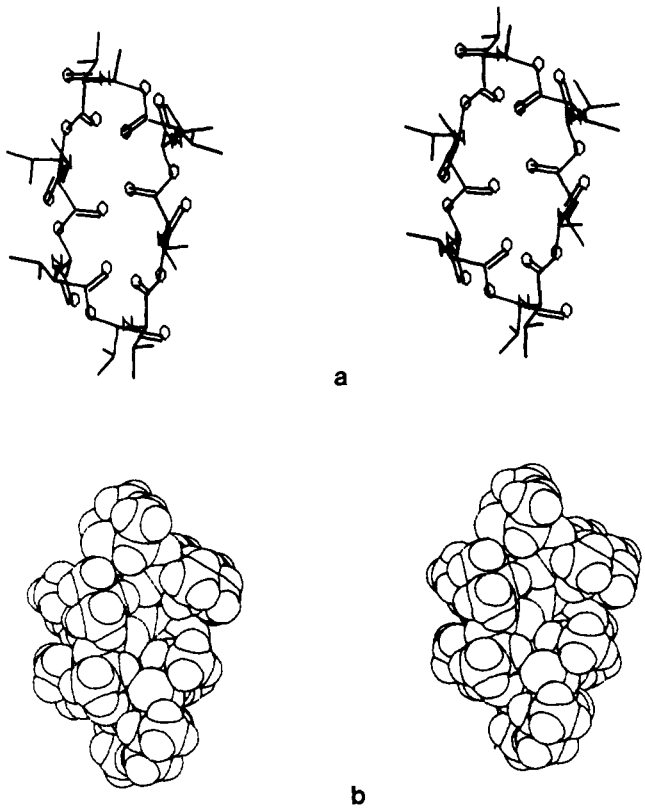

Figure 5. Crystal structure of uncomplexed valinomycin.

program is being used extensively to view new structures in different orientations.

We are confident that the increased perception of three-dimensional relationships afforded by SPACFIL and other MMMS programs will aid the chemists in efficiently designing new, specifically acting drugs.

ACKNOWLEDGMENT

The authors thank Dr. J. D. Andose for modifying the program for display on the GT-42/43 graphics terminals and for programming the "movie" option, Mr. Jack Kath for filming Dolobid®, and MSDRL chemists for the applications and many helpful suggestions.

APPENDIX

This program uses two major algorithms, one for hidden-line removal and the second for locating the circle of intersection of two spheres.

The hidden-line routine, which is different from Warme's, decides if a line segment is visible, hidden, or partially visible. Hidden-line removal takes place in two parts since a line may be obscured in two ways, i.e., inside a sphere (Figure 7a) or behind a sphere (Figure 7b). In both cases the new end point may be found straightforwardly by the analytical solution of the equations of a line and a sphere in three dimensions.

Line

$$Z = NX + C \tag{1}$$

$$Y = MX + B \tag{2}$$

Sphere

$$R^2 = X^2 + Y^2 + Z^2 \tag{3}$$

$$R^2 = X^2 + (MX + B)^2 + (NX + C)^2 \tag{4}$$

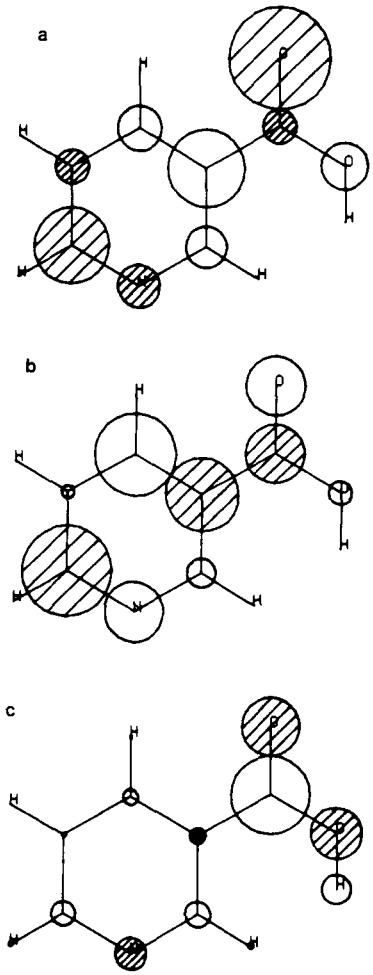

Figure 6. (a) Net charge, (b) HOMO, and (c) LUMO coefficients for nicotinic acid.

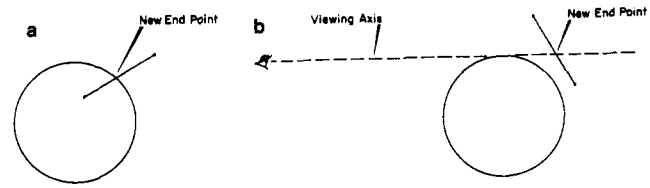

Figure 7. Two ways a sphere may obscure a line.

Expand the equation and collect X^2, X, and constant terms to get:

$$0 = DX^2 + EX + F \tag{5}$$

where

$$D = 1 + M^2 + N^2 \tag{6}$$

$$E = 2MB + 2NC \tag{7}$$

$$F = B^2 + C^2 - R^2 \tag{8}$$

Using the quadratic formula to solve for the X values, gives

$$X = \frac{-E \pm \sqrt{E^2 - 4DF}}{2D} \tag{9}$$

two values of X, because a line can cross a sphere at two points. The values degenerate if the line is tangent, and are imaginary if the line does not pass through the sphere. Given the X values, Y and Z can be found from the equations of the line. Now a decision must be made as to which X value to choose. The new end point for the line segment is the point of intersection closest to that end point which is visible.

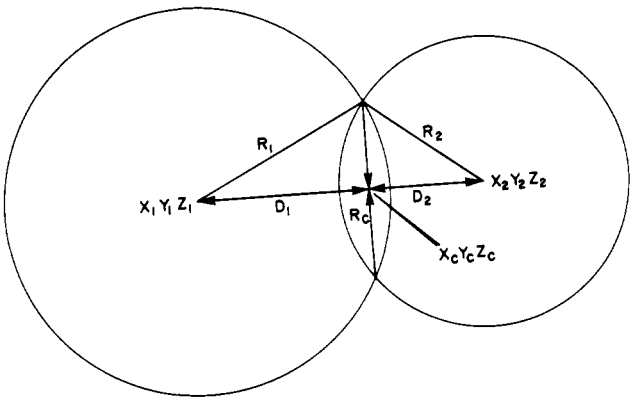

Figure 8. Finding the circle of intersection of two spheres.

One limitation of this hidden line algorithm should be noted. A line is judged to be partly hidden if one of its end points is inside a sphere. However, if a line segment were long enough to pass through a sphere with both its ends outside, the algorithm then would fail to detect the hidden part of the line. However, in practice this never occurs, provided that the longest line segment is kept substantially shorter than the length of the radius of the smallest sphere.

The second algorithm locates the circle of intersection of two contiguous spheres. Three pieces of information are needed to locate the circle of intersection in space:

1. The radius, R_c (Figure 8)
2. The coordinates of the center, X_c, Y_c, Z_c (Figure 8)
3. Two angles, PHIX and PHIZ (Figure 9)

The radius R_c is found first by solving the following three equations for D_1.

$$R_1^2 = D_1^2 + R_c^2 \tag{11}$$

$$R_2^2 = D_2^2 + R_c^2 \tag{12}$$

$$D_{12} = D_1 + D_2 \tag{13}$$

R_1, R_2, and D_{12} are known from the structure.

$$D_1 = (R_1^2 - R_2^2 + D_{12}^2)/(2 * D_{12}) \tag{14}$$

This value can then be used to evaluate eq 11 to give the radius, R_c.

The coordinates of the center (X_c, Y_c,Z_c) can be found with the following equations.

$$X_c = ((D_1/D_{12})*(X_2 - X_1)) + X_1 \tag{15}$$

$$Y_c = ((D_1/D_{12})*(Y_2 - Y_1)) + X_1 \tag{16}$$

$$Z_c = ((D_1/D_{12})*(Z_2 - Z_1)) + Z_1 \tag{17}$$

Last, the two orientation angles can be found using Figure 9 and the following equations:

$$D_{xy} = \sqrt{D_x^2 + D_y^2} \tag{18}$$

$$\text{PHIX} = -\text{ARCTAN}\ (D_{xy}/D_z) \quad (\text{unless } D_z = 0, \text{ then PHIX} = -90°) \tag{19}$$

$$\text{PHIZ} = -\text{ARCTAN}\ (D_x/D_y) \quad (\text{unless } D_y = 0, \text{then PHIZ} = -90°) \tag{20}$$

NOTE: If $D_y = 0$ then sign information (which indicates the direction of rotation around the X axis) is lost, so the sign of D_x must be given to D_{xy}.

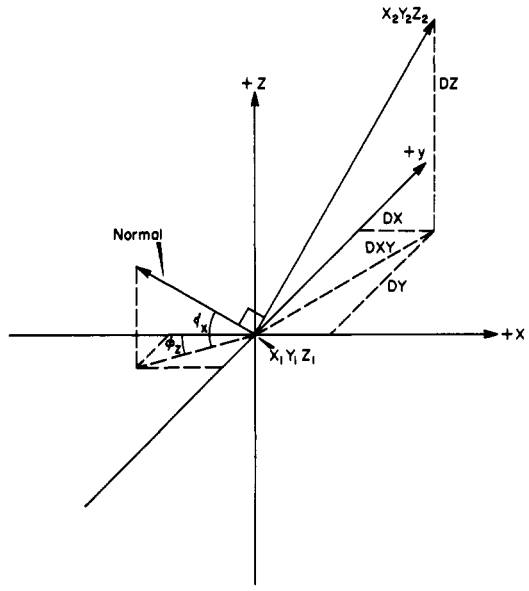

Figure 9. Orientation angles for intersecting spheres.

The negative inverse of the slope is used because the circle of intersection is perpendicular to the line connecting the centers of the two spheres.

The standard circle is initially in the XY plane having a unit radius. It is rotated around the X axis, around the Z axis, scaled with R_c, and translated to be centered at coordinates X_c, Y_c, Z_c.

Supplementary Material Available: A program listing (IBM Fortran IV, Tektronix PLOT10 graphics calls) for a Tektronix 4010 display, dimensioned for 250 atoms (22 pages). Supplementary material, which will appear in the microfilm edition of the journal, can be ordered from Business Operations, Books and Journals Division, American Chemical Society, 1155 16th St., N.W., Washington, D.C., or call (202) 872-4600. Orders must state photocopy or microfiche. Full bibliographic citation including names of all authors and prepayment are required: $3.00 for microfiche or $5.50 for photocopy (prices subject to change).

REFERENCES AND NOTES

(1) J. D. Andose, P. H. Gund, and J. B. Rhodes, 11th MARM, April 23, 1977, Abstract CO10.
(2) P. K. Warme, *Comput. Biomed. Res.*, **10**, 75 (1977).
(3) Among others, K. Knowlton and L. Cherry, *Comput. Chem.* **1**, 161 (1977); J. Staudhammer, *Comp. Graphics*, **12** (3), 167 (1978); T. K. Porter, *ibid.*, **12** 282 (1978); S. Motherwell, "Cambridge Crystallographic Files Users Manual", PLUTO 1, 1976; C. K. Johnson, Oak Ridge Thermal Ellipsoid Program (ORTEP), 1965.
(4) Structure modeled using MMMS.
(5) L. Pauling, "The Nature of the Chemical Bond", 3rd ed., Cornell University Press, Ithaca, N.Y., 1960.
(6) Tektronix, Inc., Beverton, Ore. 97005.
(7) Digital Electronics Corp., Maynard, Mass. 01754.
(8) D. L. Vickers, G. S. Smith, and S. R. Levine, *Trans. Am. Crystallogr. Assoc.*, **12**, 85 (1976).
(9) Holographic Film Co., 37 E. 4th St., New York, N.Y.
(10) Symposium on Pain and Pain Relief, Wembley Conference Centre, London, Oct. 12, 1977.
(11) K. Neupert-Laves and M. Dobler, *Helv. Chim. Acta*, **58**, 432 (1975).
(12) I. L. Karle, *J. Am. Chem. Soc.*, **97**, 4379 (1975).
(13) D. A. Matthews, R. A. Alden, J. T. Bolin, S. T. Freer, R. Hamlin, N. Xuong, J. Kraut, M. Poe, M. Williams, and K. Hoogsteen, *Science*, **197**, 452 (1977).
(14) H. B. Woodruff, E. O. Stapley, H. Wallick, H. R. Onishi, S. B. Zimmerman, A. K. Miller, and D. Hendlin, *Microb. Drug Resist.*, 409 (1975).