The Use of Molecular Modeling and VSEPR Theory in the Undergraduate Curriculum to Predict the Three-Dimensional Structure of Molecules

Teaching with Technology ─────────────────────────────────

The Use of Molecular Modeling and VSEPR Theory in the Undergraduate Curriculum to Predict the Three-Dimensional Structure of Molecules

Brian W. Pfennig* and Richard L. Frock
Department of Chemistry, Vassar College, Poughkeepsie, NY 12604; **brpfennig@vassar.edu*

The purpose of this paper is to illustrate how valence shell electron pair repulsion (VSEPR) theory and molecular modeling can be used in a complementary fashion to predict the three-dimensional structure of molecules. Students in an advanced general chemistry course or an upper-level physical inorganic course are given a series of small molecules to investigate. After sketching the Lewis structures for these molecules, the students use VSEPR theory to predict their three-dimensional structures, including site preferences and bond length differences for ligands in the trigonal bipyramidal geometry, deviations from the "ideal" bond angles in molecules with mixed ligands, and the relative magnitudes of the bond dipoles in these molecules. The students then use the Editor mode of the CAChe molecular modeling program[1] to draw the molecule's skeletal structure, instruct the computer to perform a molecular mechanics calculation to minimize the steric energy of the molecule, perform a semiempirical quantum mechanical calculation using MOPAC to further optimize this structure, and finally visualize the molecule in three dimensions by rotating it on the computer screen. The CAChe program is also used to probe the resulting bond lengths, bond angles, partial charges, and dipole moments; and these values are compared to those qualitatively predicted by VSEPR theory. The resulting molecular information is compared with structural data from the chemical literature to verify the efficacy of the valence bond and molecular orbital models in predicting the three-dimensional structures of molecules. At Vassar College, students worked in groups of 3–4 over two laboratory periods on this project. Student responses on end-of-the-semester evaluations ranked this exercise highly among the seven projects performed during the semester. They commented specifically on how the software allowed them to visualize the three-dimensional aspects of molecules more effectively than simple ball-and-stick models and on how the molecular modeling reinforced their understanding of the more subtle features of VSEPR theory, such as bond angle deviations, site preferences, and bond dipoles.

Using the VSEPR Model to Predict the 3-D Structures of Molecules

Since its original formulation by Gillespie and Nyholm (*1*) in the late 1950s, valence shell electron pair repulsion (or VSEPR) theory has become one of the most fundamental models introduced in the chemistry curriculum. VSEPR theory owes its immense popularity to the fact that it is a relatively simple model and can be used to predict the three-

dimensional geometry of molecules from their Lewis structures. The principal assumption of the model is that the valence electron pairs in molecules occupy certain regions of space called *domains* and that these electron pair domains will tend to repel each other electrostatically. An electron pair domain is a region in space around an atom in a molecule in which there is a high probability of finding one or more electron pairs. The sulfur atom in thionyl chloride ($SOCl_2$), for instance, has four electron pair domains around it: two S–Cl single bond domains, a lone pair domain localized on the S atom, and a S=O double bond domain. The S=O double bond in $SOCl_2$ is counted as a single domain because each of the four electrons in the double bond is forced to be localized in the region of space between the S and O nuclei. In general, multiple bond domains will take up more space around a central metal atom than single bond domains because more electrons are forced into the region of space between two atoms in such domains. A more complete explanation of multiple bonding and the VSEPR model can be found in a recent *Journal* article by Gillespie (*2*). Likewise, a lone pair domain, which is localized on a particular atom, will occupy more space around that atom than will a single bonding pair domain, in which the electron pair must be shared with a second atom. For bonds between atoms of differing electronegativities, the electrons in these bonding pair domains will lie closer to the more electronegative atom. Thus, different kinds of electron pair domains compete unequally for the available space around the central atom(s) in molecules.

Once the number of domains around the central atom has been ascertained, the basic electron geometry of the molecule can be determined by recognizing that electron pair domains will repel each other in order to achieve a configuration around the central atom in which they are as far apart from each other as is physically possible. The basic electron geometries for 2–6 bonding electron pair domains and their bond angles are already familiar to most students of chemistry and are listed in Table 1.[2] With the exception of the trigonal bipyramid, all ligand sites in each of these geometries are symmetrically equivalent and there are no specific site preferences for different ligands. The nonequivalence of equatorial and axial sites in a trigonal bipyramidal geometry has important implications with regard to the site preferences of the different kinds of electron pair domains, as reviewed in an excellent article by McKenna and McKenna (*3*). The larger domains will favor equatorial sites, at which there is less steric interaction than at the axial sites.

The presence of one or more lone pair electron domains

Table 1. Electron Geometry of Molecules with Differing Numbers of Electron Pair Domains, with Experimental and Calculated Bond Angles for Representative Molecules

No. of Domains	Electron Geometry	3-D Shape	Molecular Example	Bond Angles (calcd)[a]	Bond Angles (exptl)[b]
2	linear		O=C=O	180°	180°
3	trigonal planar		Cl_3B	120°	120°
4	tetrahedral		$SiCl_4$	109.5°	109.5°
5	trigonal bipyramidal		PF_5	in plane:120° $F_{ax}-P-F_{eq}$: 90°	in plane:120° $F_{ax}-P-F_{eq}$: 90°
6	octahedral		SF_6	in plane: 90° $F_{ax}-S-F_{ax}$: 180°	in plane: 90° $F_{ax}-S-F_{ax}$: 180°

[a]Calculated by CAChe. [b]Reference 9.

on the central atom in a molecule has several important consequences in determining the molecule's three-dimensional shape. The most obvious of these effects is that because electrons are so much smaller than atomic nuclei, ordinary methods of structural characterization cannot detect the presence of the lone pair electron domains. Therefore, the observed molecular geometry will appear to be different from that of the basic electron geometry and the resulting structure will be named differently, as shown in Table 2.

The second consequence of a lone pair domain on the central atom is a distortion of the bond angles from their idealized geometry, since lone pair domains occupy more space around the central atom than do single bond domains. Thus, the equatorial F–S–F bond angle in the seesaw-shaped SF_4 molecule, which has a trigonal bipyramidal electron geometry with a lone pair occupying the more spacious equatorial site, is only 101.6°, significantly less than the expected 120° bond angle. The F_{ax}–S–F_{eq} bond angle in this molecule is also smaller than the predicted 90° bond angle at 86.6°.

The larger effective size of lone pair domains around the central atom can affect bond lengths as well as bond angles. For instance, the axial S–F bond length in SF_4 is 1.65 Å, while the equatorial S–F bond lengths are 1.55 Å. The VSEPR interpretation of these experimental results is that the lone pair in the equatorial plane repels the two axial fluorine atom bonding pairs, which are closer to it, to a greater extent than it repels the two equatorial fluorine atom bonding pairs. It should be noted, however, that other factors, such as the size of the ligands themselves, electronegativity differences, and hybridization also affect these bond lengths by influencing the degree of orbital overlap. These bond length and bond angle deviations can be qualitatively predicted by the students and then compared to the results of a molecular modeling calculation and actual structural data.

Multiple bond domains are also larger than single bond domains and will have similar effects on the bond lengths and bond angles in molecules as did the lone pair domains. For instance, these will also prefer the equatorial sites in a trigonal bipyramidal electron geometry, where there is more

room for them to expand. An interesting question that might arise is, "In a molecule with both a lone pair and a double bond, which will be the more dominant domain?" In the square pyramidal molecule $XeOF_4$, the lone pair and Xe=O double bond domains are opposite each other, as one might expect, to minimize the electron–electron repulsions between these two larger domains. The O=Xe–F bond angle in this molecule is 91.6°, slightly larger than the predicted 90° value. This result indicates that the Xe=O double bond domain defeats the lone pair domain in dictating the geometry of this compound. In contrast, the O=S=O bond angle in SO_2 is 119.3°, less than the predicted bond angle for a trigonal planar electron geometry, indicating that the lone pair dictates the geometry in this example. The answer to this question is therefore a complex one, which depends on a number of factors, including the bond polarity. The results of molecular modeling calculations are therefore especially useful in these more ambiguous cases in predicting the correct molecular structure.

Even in molecules in which there are no multiple or lone pair domains, deviations from ideal bond angles result from electronegativity differences between the different atoms. Examples of the observed and calculated bond angles for several pairs of molecules which differ only in the electronegativities of their substituents are listed in Table 3. A comparison of the bond angles in H_2O (104.5°) and H_2S (92.1°) is particularly illustrative. The Pauling electronegativities of H, O, and S are 2.1, 3.5, and 2.5, respectively. Thus, the bonding electrons in H_2O are held much closer to the central atom than in H_2S, owing to the larger electronegativity of O as compared to S. As the two lone pair domains in both of these molecules repel each other, causing the ideal bond angle for this tetrahedral electron geometry to decrease from 109.5°, the bonding electron pairs in H_2O will repel each other sooner than those in H_2S because they are held closer to the more electronegative O atom. Students can use the other examples in this table to make similar comparisons.

Using Molecular Modeling to Predict the 3-D Structure of Molecules

A second goal of this paper is for students to be able to use molecular modeling software to determine and visualize the optimum three-dimensional structure for a molecule. The CAChe molecular modeling package consists of a number of programs that perform the various tasks involved in this process. The details of how to use this software (4) will not be discussed here. Instead, an overview of the basic procedure will be presented.

The first step in determining the three-dimensional structure of a molecule is to construct its Lewis structure. This is done using the Editor program of CAChe, which allows one to build and modify the molecule and to define certain input parameters that are to be used in performing the

Table 2. Electron and Molecular Geometry of Molecules with Differing Numbers of Bonding Pair and Lone Pair Domains, with Experimental and Calculated Bond Angles for Representative Molecules

No. of Domains	Electron Geometry	No. of Lone Pair Domains	Molecular Geometry	Molecular Example	Bond Angles (calcd)[a]	Bond Angles (exptl)[b]
3	trigonal planar	1	bent	Cl–Ge–Cl	103.3°	100.4°
4	tetrahedral	1	trigonal pyramidal	PF_3	98.0°	97.8°
4	tetrahedral	2	bent	H_2O	103.5°	104.5°
5	trigonal bipyramidal	1	seesaw	SF_4	in plane: 104.2° F_{ax}–S–F_{eq}: 86.9°	in plane: 101.6° F_{ax}–S–F_{eq}: 86.6°
5	trigonal bipyramidal	2	T-shaped	BrF_3	81.5°	85°
5	trigonal bipyramidal	3	linear	XeF_2	180°	180°
6	octahedral	1	square pyramidal	BrF_5	in plane: 90° F_{ax}–Br–F_{eq}: 85.4°	in plane: 90° F_{ax}–Br–F_{eq}: 85.1°
6	octahedral	2	square planar	$[ICl_4]^-$	90°	90°

[a]Calculated by CACHe. [b]Reference 9.

quantum mechanical calculations on the molecule. Students are asked to select each atom from CACHe's built-in periodic table, defining the valence configuration, charge, and display color. Atoms are then bonded by clicking and dragging from one atom to the other on the computer screen. Double and triple bonds are constructed by doing this two or three times in succession, respectively. Once the configuration of each atom and the bonds have been defined, the molecular representation can be tidied up using the Beautify command. The molecule file is then saved in an appropriate folder on the hard drive.

The second step in obtaining the three-dimensional structure of the molecule is to perform a series of calculations to optimize the molecular geometry. Students first run the Mechanics program on their molecular file. This program employs classical theory to predict a reasonable lowest-energy configuration of the molecule based on minimizing its steric interactions. Several good references have appeared in this *Journal* describing the principles of molecular mechanics (*5, 6*). The force field that a molecule feels includes empirically derived equations from classical physics to describe bond stretching, bond angle bending, bond twisting (or torsional strain), nonbonded interactions (van der Waals forces), and hydrogen bonding. The Mechanics program makes small changes in each of these parameters and calculates the resulting steric energy of the molecule (*4*). The steric energy is the difference in potential energy between the current configuration of the molecule and the form in which all these parameters are at their ideal values.

After obtaining a reasonable structure of the molecule using molecular mechanics, students perform a semiempirical quantum mechanical calculation using MOPAC (*7*). In our course, students use this program to calculate the electronic structure of the molecule by solving the Schrödinger equation using the AM1 semiempirical Hamiltonians developed by Dewar (*8*). These parameters are especially effective at considering the effects of the lone pairs as MOPAC further optimizes the molecular structure. The resulting structure can then be visualized in three-dimensional space and manipulated on screen using the Visualizer+ program of CACHe. Both space-filling and ball-and-stick representations of the molecule can be displayed. The Visualizer+ program also allows students to select certain bond lengths or bond angles to be evaluated and these data can be recorded in their notebooks. In our class, these (calculated) data are compared with the literature (experimental) values obtained by structural methods, as shown in Tables 1–3 (*9*). Reasonably good agreement was found between the calculated and experimental bond lengths and angles for these molecules, reinforcing the predictive capabilities of both the VSEPR and molecular modeling approaches in determining molecular structure.

Table 3. Electron and Molecular Geometry of Pairs of Molecules with Differing Bond Polarities, with Their Experimental and Calculated Bond Angles

No. of Domains	Electron Geometry	No. of Lone Pair Domains	Molecular Geometry	Molecular Example	Bond Angle (calcd)[a]	Bond Angle (exptl)[b]
4	tetrahedral	1	trigonal pyramidal		105.6°	100.3°
4	tetrahedral	1	trigonal pyramidal		98.0°	97.8°
4	tetrahedral	2	bent		102.5°	103.1°
4	tetrahedral	2	bent		103.5°	104.5°
4	tetrahedral	2	bent		95.4°	92.1°

[a]Calculated by CAChe. [b]Reference 9.

Partial Charges and Dipole Moments

The concept of partial charge is usually invoked in the introductory chemistry curriculum when trying to determine which of a number of Lewis structures is the best resonance form. Typically, formal charges are used for this purpose. However, this formulation assumes that the bonding is perfectly covalent, that is, that the bonding electrons are equally shared by both atoms in the bond. This is rigorously correct, of course, only when the two atoms have identical electronegativities, as is the case for homonuclear diatomics. Most compounds, however, have some degree of polar covalent bonding, in which the bonding electrons are held closer to one atom in the bond than to the other atom, leading to partial negative and positive charges, respectively. In order to predict a more accurate partial charge in polar covalent bonds, Allen has developed a partial charge formula, eq 1, which accounts for these electronegativity differences and which has recently been incorporated into general chemistry textbooks (*10, 11*).

$$(Q)_A = (VE)_A - (LPE)_A - 2 \sum_{\text{all bonds}} \frac{\chi_A}{\chi_A + \chi_B} \qquad (1)$$

In this formula, the partial charge (Q) of an atom (A) is equal to the number of valence electrons that the atom would possess if it were a free atom (VE) minus the number of lone pair electrons the atom has in the molecule (LPE) minus twice the sum (over all bonds) of the electronegativity of that atom (χ_A) divided by the sum of the electronegativities of each atom in the bond ($\chi_A + \chi_B$). Thus, for example, while the formal charges on each atom in formaldehyde, CH_2O, are all zero, yielding no information about the bond polarities or molecular dipole moment, the partial charges calculated by eq 1 are nonzero and are shown below in eqs 2–4. The Pauling electronegativities of C, H, and O are 2.5, 2.1, and 3.5, respectively.

$$(Q)_H = 1 - 0 - 2\left(\frac{2.1}{2.1 + 2.5}\right) = +0.09 \qquad (2)$$

$$(Q)_O = 6 - 4 - 2\left(\frac{3.5}{3.5 + 2.5} + \frac{3.5}{3.5 + 2.5}\right) = -0.33 \qquad (3)$$

$$(Q)_C = 4 - 0 - 2\left(\frac{2.5}{2.5 + 3.5} + \frac{2.5}{2.5 + 3.5} + \frac{2.5}{2.5 + 2.1} + \frac{2.5}{2.5 + 2.1}\right) = +0 \qquad (4)$$

These partial charges are far more useful than formal charges in predicting that the bonds are polar and that formaldehyde will have a molecular dipole moment.

Students are asked to calculate the partial charges on each atom for the molecules in Table 4 using eq 1, predict from the molecule's geometry the relative magnitude of its molecular dipole moment, and then compare their predictions to the partial charges and dipole moments obtained using the CAChe molecular modeling program. This is accomplished by first optimizing the molecular structure using the procedure mentioned above and then using the Project Leader program to obtain the partial charges on each atom. The calculation that Project Leader performs is an AM1 semiempirical quantum mechanical calculation. The results for formaldehyde and a number of other small molecules are shown in Table 4 and are compared with the partial charges obtained using eq 1. Although the agreement is not perfect in every case because of the many simplifying assumptions made in eq 1, the results for this set of molecules, at least, demonstrate yet another way that molecular modeling can be employed in the chemistry curriculum in the estimation of the three-dimensional structure and physical properties of molecules.

The net dipole moment for a molecule results from the vector addition of each bond dipole. Thus, in highly symmetric molecules such as PF_5, whose polar P–F bonds precisely cancel each other out, the molecular dipole moment is zero. In asymmetric compounds, however, the net dipole moment is often nonzero. For these molecules, the dipole moment (μ) can be defined as the product of the magnitude of the vectorally added partial charges (q) and the distance (d) between the vector sum of the positive and negative partial charges. Table 4 lists the dipole moments calculated by Project Leader for a number of small molecules, as well as the experimentally measured values (*12, 13*). These values can be compared to those qualitatively predicted by a consideration of the sizes of the partial charges on each atom, the molecular geometry determined by VSEPR theory, and bond lengths.

Summary

VSEPR theory and molecular modeling using the CAChe software package was performed by small groups of students in a two-week laboratory experiment to demonstrate the complimentary nature of these models in predicting the three-dimensional geometry of molecules and to calculate partial charges and dipole moments for these molecules. The

Table 4. Molecular Geometry of Some Small Molecules, with Their Dipole Moments and Partial Charges

Molecular Structure	Molecular Geometry	Dipole Moment[a]		Partial Charge (calcd)	
		Calcd	Exptl	CAChe	Eq 1
O=C=O	linear	0.00	0.00	C +0.41 O −0.21	C +0.67 O −0.33
(bent structure)	bent	1.86	1.85	O −0.38 H +0.19	O −0.50 H +0.25
(trigonal planar structure)	trigonal planar	2.32	2.34	C +0.14 H +0.07 O −0.28	C +0.16 H +0.09 O −0.33
(trigonal pyramidal structure)	trigonal pyramidal	1.85	1.49	N −0.40 H +0.13	N −0.53 H +0.18
(tetrahedral structure)	tetrahedral	0.00	0.00	Si +1.10 Cl −0.28	Si +1.00 Cl −0.25
(tetrahedral structure)	tetrahedral	1.50	1.58	C −0.10 H +0.13 Cl −0.08	C +0.01 H +0.09 Cl −0.09

[a]References *12* and *13*.

laboratory project was timed to coincide with the coverage of the VSEPR, valence bond, and molecular orbital bonding models in the lecture section. While VSEPR theory is incredibly useful in its simplicity, a major strength of the molecular modeling approach is that it allows students to quantitatively explore the more subtle implications of VSEPR theory using an alternative model. For instance, the molecular modeling approach is particularly useful for demonstrating the bond length and bond angle deviations from ideal geometry that accompany the presence of lone pair or multiple bond domains or result from differences in the electronegativity of the bonded atoms. These features can also be qualitatively predicted using the VSEPR model, but cannot easily be demonstrated using conventional ball-and-stick models. The calculated values for bond angles, bond lengths, partial charges, and dipole moments were all found to be in reasonable agreement with literature values, adding a quantitative aspect to the project. In today's computer-based society in which powerful molecular modeling programs such as CAChe have become increasingly affordable, this two-pronged approach to the treatment of molecular geometry in the undergraduate curriculum can be highly effective.

Notes

1. CAChe[TM] is a registered trademark of the Oxford Molecular Group Inc.
2. It is difficult to represent this three-dimensional structure on a two-dimensional sheet of paper. Therefore, certain conventional representations must first be defined. Any solid line represents atoms that lie in the plane of the paper. Those atoms that lie in front of this plane are indicated with a solid wedge, whereas those that lie behind this plane are indicated with a dashed wedge.

Literature Cited

1. Gillespie, R. J.; Nyholm, R. S. *Q. Rev. Chem. Soc.* **1957**, *11*, 339.
2. Gillespie, R. J. *J. Chem. Educ.* **1992**, *69*, 116.
3. McKenna, A. G.; McKenna, J. F. *J. Chem. Educ.* **1984**, *61*, 771.
4. *CAChe[TM] Reference Version 3.9*; Oxford Molecular Group Inc.: Campbell, CA, 1996.
5. Boyd, D. B.; Lipkowitz, K. B. *J. Chem. Educ.* **1982**, *59*, 269.
6. Cox, P. J. *J. Chem. Educ.* **1982**, *59*, 275.
7. *MOPAC Guide Version 3.8*; Oxford Molecular Group Inc.: Campbell, CA, 1995.
8. Dewar, M. J. S.; Zoebish, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
9. Gillespie, R. J.; Hargittai, I. *The VSEPR Model of Molecular Geometry*; Allyn and Bacon: Needham Heights, MA, 1991.
10. Bodner, G. M.; Rickard, L. H.; Spencer, J. N. *Chemistry: Structure and Dynamics*; Wiley: New York, 1996.
11. Moog, R. S.; Farrell, J. J. *Chemistry: A Guided Inquiry*; Wiley: New York, 1996.
12. Moore, W. J. *Physical Chemistry*, 4th ed.; Prentice-Hall: Englewood Cliffs, NJ, 1972; p 704.
13. Carey, F. A.; Sundberg, R. J. *Advanced Organic Chemistry, Part A: Structure and Mechanisms*, 2nd ed.; Plenum: New York, 1984; p 593.