

# The Molecular Mechanics of Quantized Valence Bonds

Vernon G. S. Box

Department of Chemistry, City College of C.U.N.Y., New York, NY 10031 (box@scisun.sci.cuny.cuny.edu)

Received: 15 July 1996 / Accepted: 18 February 1997 / Published: 12 March 1997

## Abstract

A new force field, Quantized Valence Bonds' Molecular Mechanics (QVBMM) has been included in the molecular modeling program STR3DI.EXE. The QVBMM force field successfully embraces and implements all of the pivotal concepts in VSEPR theory and uniquely integrates lone pairs into molecular mechanics. QVBMM facilitates a detailed analysis of the stereo-electronic effects that contribute to the structural and conformational preferences of organic molecules in their ground states, including those molecules that possess the common heteroatoms. The design, parameterization and application of the force field to a few representative molecules is discussed. The anomeric effect is also briefly examined.

**Keywords:** Molecular mechanics, force field, lone pair interactions, VSEPR, anomeric effect

## Introduction

### *Current molecular modeling methods*

Also a tremendous effort that has gone into theoretical chemistry and molecular modeling, there is still discord about the origins of many fundamental stereo-electronic effects and their manifestation. Unfortunately, these stereo-electronic effects form the theoretical basis upon which molecular modeling programs are constructed. Thus, molecular modeling is, and will continue to be, a controversial, very active and rapidly growing area of organic chemistry.

Currently, the four most widely used theoretical methods are the *ab initio* methods,[1] the Valence Shell Electron Pair Repulsion (VSEPR) theory,[2] the molecular mechanical (MM) methods,[3] and the semi-empirical molecular orbital methods.[4] While parts of VSEPR theory, like torsional potentials, are implicitly incorporated into most MM methods, the entire VSEPR theory is not totally implemented by computer programs and, hence, this model is still largely non-quantitative and approximate.

The methods based on molecular mechanics[3] are widely used and have gained considerable credibility because of their ability to simulate the structures of a very wide variety of organic molecules. Indeed, the molecular mechanical methods are used by many organic chemists for examining molecules that are too complex for study by the *ab initio* methods. There is still much room for new developments in the molecular mechanical methods, for example the treatment of lone pairs of electrons in d-orbitals, and new force field are still being developed and published.

VSEPR theory has continued to attract the attention of researchers, and there are efforts to rationalize its principles in terms of accepted quantum mechanical theory.[2] However, even the qualitative use of VSEPR theory is justifiable because of the remarkable, nearly unblemished, 35 years of successes this method has had in the prediction of the structures of covalent bonded molecules. The critically important concepts in this method are few and the method focuses attention on the "natural" geometrical preferences of clusters of atoms/domains about a central atom and the potential energies of isomeric molecules are easily, though qualitatively, compared. Indeed, this model has recently been illustrated

by tying air-filled balloons together, each balloon representing a domain - a bonding orbital, or lone pair, and its host atom(s).

VSEPR theory is also remarkable because the method is independent of the concept of orbitals. Molecular Orbital theory, while clearly an integral part of modern organic chemistry, often allows its users to generate interesting inconsistencies. For example, the restricted rotation about the ethane C-C bond has been ascribed to a "weak p-interaction" between the 2p orbitals on the carbons.[6] However, molecular orbital (bonding) theory has suggested that ethane should not possess atomic orbitals of any kind, and certainly not the p-orbitals needed for this p-interaction. Further, there is no experimentally documented instance of a p-bond, even a delocalized p-bond, that is longer than 1.43 Å, while the C-C bond in ethane is 1.53 Å long.[7] VSEPR theory rationalizes this restricted rotation quite simply as the repulsion between the eclipsed bonding pairs of electrons.

Given the success of VSEPR theory, it was obviously desirable to incorporate all of its principles, especially its highly visible dependence on the stereo-electronic roles of lone pairs of electrons, into a computational structure determination method. This would eventually give VSEPR theory a more quantitative profile. Molecular mechanics already used some of the principles of VSEPR theory and, like VSEPR theory, did not assume the existence of special types of orbitals. Molecular mechanics was therefore the most suitable vehicle to carry VSEPR theory forward. This work describes the *de novo* development of a force field guided by the desire to incorporate as little as is possible from existing force fields so that it would not simply be a "clone" of existing force fields. Further every attempt was made to fuse as many of the principles of VSEPR theory, as is possible, with the concepts (and consequences) of quantized bond types and incorporate these into the new molecular mechanics force field.

#### *Molecular modeling and stereo-electronic effects*

In principle, the homologous molecules, ethane, n-butane, n-hexane and n-pentadecane must have common stereo-electronic features, and the data that allows one to model n-butane must also enable the successful modeling of n-pentadecane. Therefore, if a molecular modeling program is parameterized to reliably reproduce the data of simple unstrained molecules, then it ought to serve as a tool for identifying those analogous molecules whose data are "unusual". Whenever an "unusual" molecule is encountered, the organic chemist can review the experimental data for that molecule and begin to make logical evaluations of these data in terms of what was predicted by a simple model and what was realized in practice. In this way stereo-electronic effects and their consequences can be brought to the forefront and examined.

This approach was used in the *de novo* development of the simply parameterized QVBMM force field. This approach has been successful and will enable us to identify, and probably answer, some of the remaining fundamental questions

about the roles of stereo-electronic effects in determining the structures of common organic molecules.

### **Molecular mechanics**

#### *The traditionally used energy potentials*

The energy potentials employed in molecular mechanics have been widely discussed.[3,4,5] The molecule's potential energy is usually calculated from the expression:

$$E_{\text{tot}} = E_{\text{str}} + E_{\text{ang}} + E_{\text{oop}} + E_{\text{ster}} + E_{\text{tor}} + E_{\text{elct}} + E_{\text{nb}} - E_{\text{solv}}$$

where  $E_{\text{str}}$  is the energy required to stretch/compress a bond,  $E_{\text{tor}}$  is the energy for rotation about a bond,  $E_{\text{ang}}$  is the energy required to compress or expand a bond angle,  $E_{\text{oop}}$  is the energy required to deform a p-systems from planarity,  $E_{\text{ster}}$  is the steric energy generated by steric repulsion between proximate groups/atoms,  $E_{\text{elct}}$  is the electrostatic potential between charges within the molecule,  $E_{\text{solv}}$  is the solvation energy, and  $E_{\text{nb}}$  encompasses all other intermolecular non-bonded interactions between non-polar molecules. The forms of these energy potentials have been widely discussed.[3,5]

The energy potential functions  $E_{\text{str}}$ ,  $E_{\text{ster}}$ ,  $E_{\text{oop}}$  and  $E_{\text{ang}}$  are usually estimated by a simple Hooke's Law derived function of the form:[3,5]

$$E = K \cdot (R - R_0)^2 \quad (1)$$

where  $R - R_0$  is the change in the length/width/out-of-plane-displacement of a bond/angle/atom.

The energy potential function  $E_{\text{tor}}$  relates the energy of four sequentially bonded atoms to their dihedral/torsion angle and is usually based on the observed preference for staggered over eclipsed conformations, especially whenever there is the possibility of "free" rotation about a bond. It is usually related to the torsion angle,  $w$ , by the expression:

$$E_{\text{tor}} = K \cdot [1 + (-1)^{J+1} \cdot \cos(J \cdot w)], \text{ for } J = 1 \text{ to } 3 \quad (2)$$

The forms of  $E_{\text{nb}}$  and  $E_{\text{elct}}$  are not as standardized as the other potentials listed above and vary from one implementation of molecular mechanics to the other.[3,5]

It is particularly important to remember that the potential energy functions used in molecular mechanics are empirically selected and their forms do not allude to the scientific underpinnings of molecular/atomic motion, or stereo-electronic effects.

#### *The effects of solvation*

Solvation is one of the remaining unsolved problems in molecular mechanics. Most molecular mechanics implementations bypass this difficulty by assuming that the molecular model is either in the gas phase, when all of the effects of

solvation are ignored, or in a homogeneous dipolar medium, when the often unsymmetrical topographical features of solvation are also ignored. The molecule is assumed to be pervaded the medium/solvent uniformly, like a tree in a dense fog. Even small globally symmetrical solvent molecules (like  $\text{CCl}_4$ ) can only interact with the external surface of a compact molecule and not with its interior, contrary to the tree/fog analogy. Therefore, solvation will not be a factor in many of the intramolecular interactions of compact molecules. In reality, an effective treatment of solvation must simultaneously embrace the topographical complexities of solvent - solute interactions and solvent - solvent interactions.

Most experimental determinations of structure or conformation are done in the solid phase or in solution. It is well known that the observed distribution of the conformational states of a molecule is dramatically affected by solvation, and that the nature of the solvent is of great importance.[8] In the solid phase, the occurrence of more than one conformation of a molecule in the unit cell, though not widespread, is well known and these different conformations of the same molecule usually do have slightly different bond lengths and angles.[8]

Some of the existing experimental data are inconsistent. In the parameterization process the use of this data must be approached with caution. For example, the cited experimentally determined conformational free energy of the hydroxyl group of cyclohexanol ranges from 0.31 to 1.5 kcal/mol, that of the methyl group of methylcyclohexane from 1.18 to 2.1 kcal/mol, that of the ethyl group of ethylcyclohexane from 1.67 to 2.27 kcal/mol, and that of the acetoxy group of cyclohexyl acetate from 0.36 to 1.6 kcal/mol.[9]

## The valence shell electron pair repulsion model (VSEPR)

### Torsional and steric interactions

The VSEPR model implicitly acknowledges the following features of covalent molecules:

a) The dominance of electron- electron repulsion.

These interactions are ranked in the hierarchy: lone pair - lone pair ( $n-n$ ) > lone pair - bond ( $n-\sigma$  and  $n-\pi$ ) > bond - bond ( $\sigma-\sigma$ ,  $\sigma-\pi$  and  $\pi-\pi$ ) repulsion.

This hierarchy reflects the increasing mollifying influence of the nuclei of the atoms involved on the size of the electronic interactions, since electron - nuclear interactions must be stabilizing while electron - electron interactions are destabilizing. The importance of electron pair - nuclear attractions is further realized in the examination of complex organic molecules that have intramolecular  $n$  - dipole interactions. The effective size (domain) of a lone pair orbital must be significant in order to logically support its importance in the hierarchy above. Valence shell lone pairs should also occupy  $sp^n$  hybridized orbitals, or  $p$ -orbitals, whose directionality will facilitate the orientation of these electrons away from other electron density and towards positively

charged sites, and minimize their interaction (orthogonality) with the other (core and bonding) electrons.

b) The orthogonality of orbitals borne by a given atom.

This concept establishes the geometrical requirements for minimal  $n-n$ ,  $n-\pi$ ,  $n-\sigma$ ,  $\sigma-\sigma$  and  $\pi-\sigma$  repulsions, and for minimal repulsion between these orbitals and core electrons. Orthogonality is a key feature of M. O. theory and Hückel M. O. theory, and modern M. O. theory does not require the delocalization of all proximal orbitals. Indeed, delocalization, even possibly leading to aromaticity, quite often does not occur in simple systems that seem ideally suited, and so is not an obligatory phenomenon.[7]

c) The role of the interactions of atomic/electronic domains in molecular geometry.

The determination of the structural features of a covalent molecule is due to the process of minimizing the energy of interaction between the various atoms/electrons, or domains, in that molecule. In the attainment of the minimum energy configuration, the net instability due to all repulsion and attractions discussed in a) and b) above are minimized.

## Quantized valence bonds, bond lengths, types and orders

During the development of the molecular graphics program STR3DI.EXE, an extensive review of the X-ray crystallographic data of several thousand organic molecules revealed very definitive relationships between bond lengths and bond types.[7] Further, these relationships were correlated with the results of VESCF-HMO bond order calculations.[7] These

**Table 1.** Bond type - bond length - bond order relationships

bond type	longest (pm)	cardinal (pm)	shortest (pm)	BOD	numeric type
C-C	164	153	142	< 1.5	1
C=C	142	133	125	1.5 to 2.5	2
C≡C	125	118	111	≥ 2.5	3
C-N	157	147	137	< 1.5	1
C=N	137	128	120	1.5 to 2.5	2
C≡N	120	116	107	≥ 2.5	3
C-O	153	143	133	< 1.5	1
C=O	133	124	117	1.5 to 2.5	2
C-S	201	188	175	< 1.5	1
C=S	175	164	154	1.5 to 2.5	2

(the cardinal bond type was that found in the smallest, structurally unperturbed molecule, e.g. the C-C bond of ethane and the C=C bond of ethene. BOD = bond order range - ex VESCF-HMO.)

bond length/type/order relationships have recently been refined and are shown below, (Table 1).

A recent review of these data suggested that a more realistic model should allow the observed range of lengths for any type of single bond to be  $\pm 7\%$  of the average bond length, that for a double bond to be  $+7\%$  to  $-6\%$  of the average bond length, and the observed range of lengths for a triple bond to be  $\pm 6\%$  of the average bond length. The new ranges do reflect the trend in the strengths of these bonds, since triple bonds are strongest. The bond length values given in the Table 1, above, have been rounded. The numeric value of a bond type (usually single, double, or triple) is obviously equal to the integral value of the VESCF-HMO calculated bond order obtained from MMX.[7]

The properties of the "components" of bonds (i.e. nuclei and electrons) are quantized and so one might argue that the properties of bonds, which are the "products" of nuclei and electrons, ought also to be quantized. The quantization of the properties of bonding molecular orbitals would further suggest that the lengths of bonds must also be quantized. This notion was supported by the data shown in Table 1, since the X-ray crystallographic data of several thousand molecules clearly showed that all  $\sigma$ - and  $\pi$ -bonds have lengths that fell within the ranges for the appropriate single or multiple bonds. Further the experimental data show that bond length ranges do not overlap. The X-ray data led to the conclusion that lengths and types of bonds were quantized.[7] Consequently, the bond lengths found in, and orders calculated for, delocalized  $\pi$ -systems definitively revealed the dominant  $\pi$ -bonding features in these systems. The application of these criteria to many simple conjugated  $\pi$ -systems has led to several important results concerning delocalization.[7]

### Molecular mechanics of quantized valence bonds

#### *Nature and mechanical features of covalent bonds*

The molecular mechanical treatment of bonds and bond angles as simple springs, is logical, but limited.[10] The fact that real springs are macroscopic and have bulk properties, while bonds and their angles do not, is ignored. When a real spring is stretched beyond its elastic limit it undergoes an irreversible plastic deformation and ceases to obey Hooke's Law. After plastic deformation, the spring requires less force to stretch, or compress, it than would be predicted by Hooke's Law, and the spring also loses its former equilibrium "unloaded" length. Covalent bonds and bond angles, lacking the macroscopic features of real springs, cannot undergo plastic deformations. A covalent bond that has been given sufficient energy to stretch it to, and beyond, its "elastic limit" ought to break abruptly, consistent with the quantized energy, and the limits of the length, of that bond. After the cleavage of the bond, any force that retards the movement of one fragment away from the other must be due to a coulombic attraction, rather than a "non-quantized, weak, covalent bonding"

interaction. The energy required to overcome these post-cleavage attractions should be very much less than that required to cleave the bond.

The concept of electronegativity[11] describes the strength of the attraction between the valence electrons of an atom and its nucleus. The nuclei of highly electronegative elements, like the halogens, interact very strongly with their valence electrons and so these elements can attract and retain "extra" electrons, so becoming negatively charged. These electronegative elements possess strongly positive nuclear electrostatic fields. On the other hand, the valence electrons of the alkali metals interact so weakly with their nuclei that they are easily lost, so converting their atoms into cations. These electropositive elements possess weakly positive nuclear electrostatic fields.

A simple covalent bond between two monovalent neutral atoms can be dissected into two electrons and two "nuclear bodies", each with effectively unit charge. The most important variable among covalent bonds is the electrostatic field generated by the two nuclei and their core electrons, in which the bonding electrons move. Thus, the electronegativities of the bonded atoms are the most important variables in the consideration of the properties of simple covalent bonds. Indeed, the lengths of covalent bonds seem to be dependent on the difference between the electronegativities of the bonded atoms. Indeed, the greater this difference, the shorter and stronger the resulting bond are found to be. Since all covalent bonds are similar in their "composition", then their "mechanical" properties must also be similar, even if their strengths, energies and lengths vary widely.

The degree of the polarization of a covalent bond has been associated with the difference between the electronegativities of two atoms. This very fundamental concept enables us to develop a qualitative appreciation for the way the bonds in organic molecules are polarized. Thus, the concept of electronegativity is of fundamental importance in organic chemistry and this concept was of critical importance in the development of QVBM.

#### *Bond cleavage, transition state geometry, bond energies and bond stretching force constants*

The assumption of the quantization of bond types and the existence of clearly defined bond length ranges, logically leads to the realization that stretching a double (p) bond beyond its 7% length range must result in its homolytic cleavage by the uncoupling of the p-bond, to generate a s-bond, so leading to the direct formation of a 1,2-diradical. The bond joining the atoms of this diradical must have a length within the range required for the corresponding single bond. Similarly, one must conclude that stretching a single (s) bond beyond its 7% length range must result in the homolytic cleavage of this bond. Internuclear distances in all transition states must be physically located at, or close to, these bond length range boundaries.[7,10] Heterolytic bond cleavages are obviously more complex processes, especially since the newly

separated charged fragments should attract each other very strongly, which will not be the case in homolyses. These factors normally cause heterolyses to require much more energy than homolyses.

#### Carbon - carbon bonds

At the point of homolytic cleavage of a simple carbon - carbon single bond (C1C) whose normal bond length ( $L_0$ ) was 1.532 Å, the potential energy ( $E_{str}$ ) would be the bond dissociation energy (about 85 kcal/mol) and the bond length ( $L$ ) should be approximately equal to  $L_0 \cdot 1.07$  Å, reflecting the 7% bond length limit. Then, by substitution into the traditionally used Hooke's Law expression (eq.1) we get:

$$85 = K_{C1C} \cdot (1.532 \cdot 0.07)^2 \quad (3)$$

Thus,  $K_{C1C}$ , the force constant  $K_{str}$  for the stretching of the C1C bond, was calculated to be 7391.03 kcal/mol/Å<sup>2</sup>. Similarly, the force constants for the stretching of C2C (double bond) and C3C (triple bond) can be calculated, by applying the respective bond length range to each bond type.

#### Other bonds

If BE represents bond energy and L the length of a covalent single bond (where  $L_{C1C} = 1.532$  Å and  $L_{C1Y}$  is the length of the C1Y bond between a carbon and any atom Y), then, from equation (3), at the breaking point of any single bond:

$$BE_{C1C} = K_{C1C} \cdot (L_{C1C} \cdot 0.07)^2 \quad \text{and}$$

$$BE_{C1Y} = K_{C1Y} \cdot (L_{C1Y} \cdot 0.07)^2$$

leading to

$$BE_{C1C} / BE_{C1Y} = (K_{C1C} / K_{C1Y}) \cdot (L_{C1C} \cdot 0.07 / L_{C1Y} \cdot 0.07)^2$$

hence

$$BE_{C1C} / BE_{C1Y} = (K_{C1C} / K_{C1Y}) \cdot (L_{C1C} / L_{C1Y})^2$$

thus,

$$K_{C1Y} = K_{C1C} \cdot (BE_{C1Y} / BE_{C1C}) \cdot (L_{C1C} / L_{C1Y})^2 \quad (4)$$

The bond length data available from x-ray crystallographic studies and the thermochemical homolytic bond energy data revealed another empirically derived expression linking these quantities:

$$(BE_{C1Y} / BE_{C1C}) = 1 + K_1 + L_{C1C} - L_{C1Y} \quad (5)$$

where:

$$\begin{aligned} K_1 &= -0.185 \text{ if } L_{C1C} > L_{C1Y}, \\ K_1 &= +0.185 \text{ if } L_{C1C} < L_{C1Y}, \\ K_1 &= 0 \text{ if } L_{C1C} = L_{C1Y} \end{aligned}$$

Thus, knowing the average bond energy of the C1C bond allows us to calculate the average bond energy of any single bond involving carbon. Further, all stretching force constants ( $K_{C1Y}$ ) for covalent single bonds involving carbon can be calculated using equation (6) below, which was derived from equations (4) and (5).

$$K_{C1Y} = K_{C1C} \cdot (1 + K_1 + L_{C1C} - L_{C1Y}) \cdot (L_{C1C} / L_{C1Y})^2 \quad (6)$$

Bond energies calculated using equation (5) are accurate to within 5% of the experimental values, provided that valid, mean bond lengths are used. The force constants that are calculated by using equation (6) also have a similar degree of accuracy. Equations (5) and (6) clearly indicate that short bonds will have high bond energies and high stretching force constants. These empirically derived equations (which were not developed or based on the concept of orbitals) are dependent only on the values used for the bond lengths involved, and so these equations seem to have a natural kinship with VSEPR theory and molecular mechanics.

A more general statement of equation (6) leads to equation (7), below, so extending these concepts to all single bonds by using the C1C bond as the reference bond.

$$K_{X1Y} = K_{C1C} \cdot (1 + K_1 + L_{C1C} - L_{X1Y}) \cdot (L_{C1C} / L_{X1Y})^2 \quad (7)$$

The equation (7) can be further extended to the formulation of equation (8), below, and applied to the calculation of stretching force constants of multiple bonds, where the integer n (1 to 3) represents bond type, and  $K_2 = K_3 = 0$ , while  $K_1$  has values stated above.

$$K_{XnY} = K_{CnC} \cdot (1 + K_n + L_{CnC} - L_{XnY}) \cdot (L_{CnC} / L_{XnY})^2 \quad (8)$$

These new ideas were implemented in QVBMM, and the undoubtedly successful use of the equation (8) in this force field can be seen as an indication of the (partial) validity of the concepts stated above. The use of equation (8) obviates the need for extensive tables of force constants and bond energies, since all of the force constants needed can be very rapidly calculated from the mean lengths of the bond in question and the data for the appropriate carbon - carbon bond.

#### Intramolecular non-bonding interactions and dielectric constants

The covalent radii and electronegativities of the participating atoms, and the dipoles associated with their polarized bonds, will be of primary importance in developing an understanding of intramolecular non-bonded interactions. The first "contact" between two molecules, or between two parts of the same molecule, will be the mutual repulsion of their

peripheral electron densities and the interaction of the electron density of each atom with the electrostatic fields of the approaching dipoles. These electrostatic repulsions and attractions are also partly responsible for torsional effects, the interactions of a given lone pair with other lone pairs, bonds, or atoms, coulombic interactions, London forces, and van der Waals repulsions.

Classical physics estimates the energy potential between two point charges by the equation:

$$E = (K \cdot Q_1 \cdot Q_2 / D^2) \cdot (D / \epsilon) \quad (9)$$

where  $K$  is a constant,  $\epsilon$  is the dielectric constant of the environment,  $Q_1$  and  $Q_2$  are the magnitudes of the point charges, and  $D$  is the distance between these charges.[12]

The numerical value of  $\epsilon$  for a perfect vacuum is 1, and the numerical values of  $\epsilon$  for the common hydrocarbon solvent media (like the alkanes, symmetrical alkenes and arenes) are usually less than 3. Since one could quite justifiably argue that the vacant interior of a molecule and intermolecular spaces must be quite close to being perfect vacuums, then the numerical size of the effective dielectric constant encountered in the intramolecular interactions within a simple hydrocarbon must be anywhere in the range 1 to 3. Long ranged intramolecular stereo-electronic interactions, in which the interacting atoms/bonds/dipoles are flanked or surrounded by other non-interacting molecular fragments, will be quite similar to intermolecular interactions, and the effective dielectric constant here should be close to 3 in magnitude. However, the short ranged intramolecular interactions should best be considered using an effective dielectric constant of unity.

Since this numerical range of the effective dielectric constant coincides with the numerical range of interatomic distances over which intramolecular stereo-electronic interactions occur, one intuitively realizes that the quantity  $D/\epsilon$  from equation (9) could be quite close to unity. Thus, the parameterization of QVBMM the force field for intramolecular interactions was first attempted by assuming that the quantity  $D/\epsilon$  was unity. As will be seen from the data below, the resounding success of this force field in the modeling of the structures of a wide variety of molecules does justify the use of this assumption and might even be considered to be strong evidence supporting the validity of the assumption that the quantity  $D/\epsilon$  is approximately unity for intramolecular interactions. Therefore, *in this work*, the  $D/\epsilon$  factors will be omitted from the equations, like equation (9), that describe and quantify intramolecular electrostatic interactions.

#### Bond angles and angle strain

Angle strain is defined as the energetic penalty incurred when a given bond angle is forced to depart from its equilibrium value, the value that is found in a simple, "unstrained" molecule. This concept implies that there are ideal values for bond angles, of 109.5° and 120° respectively, for non-p- and p- systems. Out-of-plane deformations do affect the

magnitudes of the bond angles in p-systems, but the rigidity with which the planarity of p-systems is enforced by nature reassures us that the use of punitive out-of-plane energy potentials is a justifiable method for handling this complication. That then leaves us to review the methods (and their validities) of assessing the potential energy stored in a given bond angle in three dimensional non-p systems and two dimensional p-systems.

The "normal" bond angles of 109.45°, at  $sp^3$  atoms, and 120°, at  $sp^2$  atoms, are only found in a very small group of highly symmetrical molecules like methane ( $sp^3$ ) and the carbonate dianion ( $sp^2$ ). These bond angles, especially those at  $sp^2$  atoms, are hardly ever found in other, even very simple, organic molecules. Attention has already been drawn to the wide-ranging acyclic bond angle sizes (from 104° to 115°) found at  $sp^3$  atoms in the X-ray crystallographic structural studies of monosaccharides.[13] Similarly, the variation in the sizes of the X-C-X angles of the halomethanes (ranging from 107° to 112°), in the alkenes  $X_2C=CH_2$  (ranging from 110° to 116°), and the halocarbonyls  $X_2C=O$  (ranging from 108° to 116°) have been reviewed.[2a] For the  $X_2C=Y$  molecules, the X-C-X angle sizes follow the sequence  $F < Cl < Br < CH_3 \sim H$ . Some of these data are summarized in Table 6 below. It is very clear that the experimentally observed bond angles are quite dependent on the sizes and electronegativities of the atoms involved, and on the presence of lone pairs on these atoms.

The wide variation in bond angle sizes, as shown above, clearly indicates that the traditionally assumed "normal" bond angle sizes might not be applicable to all systems, and it might be best to define "normal angles" based on the types and hybridizations of the atoms involved, rather than to use the "global" standards of 109.5° and 120°. On the other hand, if the "global" standards are retained, a more flexible approach to the assignment of force constants might be useful, so as to allow stereo-electronic effects to have a greater effect in the realization of the size of a "normal" bond angle. These observations prompted a careful study to identify the most efficacious potentials, and the best method for parameterizing QVBMM for angle strains. After much experimentation, a traditional Hooke's Law potential, like equation (1), was used to calculate angle strain, but great efforts were made to empirically select force constants that allowed torsional, steric and dipole - dipole interactions to influence the dynamics of the system so as to reproduce the observed "normal" angles observed in very simple molecules.

#### Energy potentials used in QVBMM

The development of QVBMM was guided by the ideas stated above, and the total enthalpy of a molecule was estimated by the equation (10).

$$E_{\text{tot}} = E_{\text{str}} + E_{\text{ang}} + E_{\text{oop}} + E_{\text{ster}} + E_{\text{rep}} + E_{\text{tor}} + E_{\text{lp}} + E_{\text{pln}} - E_{\text{solv}} \quad (10)$$

The traditional Hooke's Law potential were used to establish the terms  $E_{\text{str}}$  (bond stretch),  $E_{\text{ang}}$  (angle deformation),  $E_{\text{oop}}$  (out-of-plane deformation of p-bonds) and  $E_{\text{ster}}$  (steric). In general, the force constants for these potentials were determined empirically and were not simply adopted from the literature.

#### *Intramolecular non-bonded interactions*

Each C-H bond of the molecule ethane is weakly polarized (the electronegativities of C and H are 2.55 and 2.20 respectively), and we should anticipate that each hydrogen will bear an identical partial positive charge and each carbon bear an identical, but larger, partial negative charge. These charges on bonded atoms, calculated from the difference between the electronegativities of the bonded atoms, can then be regarded as "point charges" and their potentials can be measured by equation (9). The total charge on a multivalent atom is assumed to be the algebraic sum of the partial charges due to each of its bonds. Thus, every organic molecule, even hydrocarbons, must be seen to contain numerous dipoles (or atomic point charges) that can influence each other and thus influence the molecular energy and geometry.

Early studies in the development of QVBMM indicated that when only intramolecular interactions are considered, then the traditional coulombic and non-bonded potentials could be replaced by one new compound potential,  $E_{\text{rep}}$ , that measured the energy of the interaction of two atomic charges that were within the minimum distance for their interaction. Thus, by calculating the energies of all of the interactions of all of the atomic charges in a given molecule, we have essentially calculated the energy of all of the coulombic, dipole-dipole, and other electronic non-bonded interactions in that molecule.

The potential,  $E_{\text{rep}}$ , was estimated by applying equation (9) to the interaction of two charges (on atoms 1 and 2), C1 and C2, that were separated by distance D:

$$E_{\text{rep}} = K_{\text{rep}} \cdot C_1 \cdot C_2 / D^2 \quad (11)$$

The use of this potential suggests that the torsional interaction in the eclipsed conformation of ethane has three contributing factors:

- the dominant electronic bond pair - bond pair (torsional) repulsion, and
- the smaller repulsion of the C-H dipoles
- the still smaller steric repulsion of the eclipsed hydrogen atoms

Similarly, all of the situations in which atoms, or groups, approach each other closely enough to stimulate steric interactions will also stimulate dipolar interactions. These dipolar interactions are therefore very important in the estimation of the total energy of molecules. Energy potentials of this form are not widely used in molecular modeling force fields, but their validity in classical physics, combined with the demonstrated efficacy of QVBMM, support the uses of these assumptions.

#### *Lone pairs and their interactions*

It is well known that  $sp^2$  hybridized atoms are more electronegative, and smaller, than their corresponding  $sp^3$  states.[11] Thus, lone pairs of  $sp^2$  hybridized atoms should also be smaller, and less available for bonding and other interactions, than the lone pairs of the  $sp^3$  hybridized atoms. If we assign to the lone pairs of the saturated ( $sp^3$ ) hybridized atoms an "availability factor",  $F_{\text{lp}}$ , of one (1), then a value of 0.8688 has been assigned to the smaller availability factor for the lone pairs of p-atoms. This number, 0.8688, is the ratio of the covalent radii of all first row elements in their  $sp^2$  states to their covalent radii in their  $sp^3$  states, and was assumed to be the ratio of the sizes of lone pairs of these elements in their corresponding hybridization states.

The stereo-electronic interactions of lone pairs,  $E_{\text{lp}}$ , with other lone pairs or with atoms was also estimated by a Hooke's Law potential that embraces both repulsions and attractions.

$$E_{\text{lp}} = K_{\text{lp}} \cdot F_{\text{lp1}} \cdot F_{\text{lp2}} \cdot \text{CHRG} \cdot (D - D_0)^2 \quad (12)$$

Where  $F_{\text{lp1}}$  and  $F_{\text{lp2}}$  represent the availability of the lone pairs for interaction (if there is a lone pair-lone pair interaction, then  $\text{CHRG} = 1$ ),  $\text{CHRG}$  is the partial charge on an interacting atom (if there is a lone pair-atom interaction, then  $F_{\text{lp2}} = 1$ ), D is the distance between the interacting centers and  $D_0$  is the threshold distance for interaction to occur.

#### *Out-of-plane bending*

The potential  $E_{\text{pln}}$  is estimated by using the traditional Hooke's Law potential, equation (1), and is considered only when p-bonds are present in order to measure the enthalpy due to non-planarity of these p-bonds. The potential  $E_{\text{pln}}$  drives double bonds to be planar and is experimentally justifiable since only highly strained double bonds normally lack planarity.

#### *Torsional interactions*

The torsional potential used in QVBMM is based on the fundamental principle of the minimum energy of interaction of orthogonal orbitals and bonds. Torsional interactions that are due to repulsions between electrons in bonds or lone pairs must also be governed by this principle. The torsional energy is assumed to be zero when the dihedral angle between two bonds, or a bond and a lone pair, is  $90^\circ$ . The torsional potential  $E_{\text{tor}}$  used in QVBMM is quite different from the traditional form and can be represented as:

$$E_{\text{tor}} = K_{\text{tor}} \cdot K_a \cdot K_b \cdot [\cos(f)]^4 / D^2 \quad (13)$$

where  $K_a$  and  $K_b$  are determined by the types of orbitals that are interacting, f is the dihedral angle and D is the distance between the midpoints of the bonds. The factors  $K_a$  and  $K_b$  are 1 for s-bonds, 1.15 for p-bonds, and 1.25 for lone pairs, so representing the VSEPR hierarchy of interactions where

lone pair - lone pair > lone pair - bond pair > bond pair - bond pair. All p-bonds have more electron density than s-bonds and so must be more torsionally repulsive. Lone pairs are not counterbalanced by two nuclei, the way bond pairs are, and so generate the most torsional energy.

Note that the symmetries about the interacting atoms are no longer considered. There are many molecules that show a weak preference for an eclipsed conformation, e.g. acetaldehyde, rather than a staggered conformation, e.g. ethane. These preferences are determined by the total interactions of all of the stereo-electronic effects operating in the molecule ( $E_{\text{tor}}$  is only one of these stereo-electronic factors) and are not "ordained", as is suggested by the notion that ethane's dihedral angle of  $60^\circ$  is the "optimum" dihedral angle for all moieties that have  $sp^3$  -  $sp^3$  bonds. QVBMM was designed to allow all of the stereo-electronic effects present in a molecule to play their important roles in determining the geometry and energy of the molecule, knowing that the properly parameterized force field will show these preferences.

#### *Solvation and dielectric constant*

As explained above in the derivation of equation (9), the term  $D/e$  (distance/dielectric constant) was assumed to be unity in the parameterization of the force field for intramolecular interactions. One must therefore presume that the environment of the molecular model in QVBMM is either the gas phase, or a non-polarizable, non-polar, infinitely small (so that it can pervade the molecule), monatomic solvent. The force constants that will be affected by any subsequent consideration of the dielectric constant of the "medium" will be  $K_{\text{rep}}$ ,  $K_{\text{lp}}$  and  $K_{\text{tor}}$ .

### **Implementation of QVBMM**

#### *Energy minima search*

The final geometry of any molecule must be dependent solely on the dynamic interplay of the stereo-electronic effects within that molecule. QVBMM uses a pattern search[4] method to locate the lowest energy geometry for a molecule in any given conformation. In a variable amplitude pattern search (VAPS), each atom in the molecule is vibrated through varying amplitudes, in a spherical pattern and under the full influence of the stereo-electronic forces that affect it, until a more stable atomic position is found. The atom is then kept at this most stable site. This process is carried out on each atom in the molecular model, so moving each atom in the molecule to its currently most stable position in a given minimization cycle (iteration). The total enthalpy of the molecular model is calculated at the end of each iteration. If this enthalpy is lower than the enthalpy of the previous (starting) molecular model, then these new atomic coordinates are stored and the iterations continued. Otherwise, the amplitude of the atomic motion is altered and the iterations con-

tinued. Successive cycles of this algorithm are performed until the enthalpy of the molecular model cannot be lowered further and then the enthalpy of that conformation of the molecule is deemed to have been minimized.

The VAPS method, like all pattern search methods used in molecular mechanics, is limited by the fact that very small changes in molecular geometry can occur without incurring substantial energy penalties. Only about 0.008 kcal is required to stretch a C-C bond by 0.001 Å and 0.002 kcal to flex a bond angle by  $0.25^\circ$ . Thus, the process of defining the location of an atom whose bonds are weak ( $< 70$  kcal/mol) can be tedious since the low energied vibrations of these bonds will have relatively large amplitudes. Molecules with strong bonds will have more precisely defined geometries, but any flexible molecule will be able to rapidly fluctuate between several nearly iso-energetic similar geometries and the molecule's observed "minimum energy structure" will be a time-averaged structure embracing the iso-energetic conformations available to it. Frequently, the potential energy surface in the region in which the minimum energy structure of a molecule is located is much like a shallow, bumpy-surfaced, bowl. The location of the lowest (true local minimum) energy site on this kind of surface is very time consuming and requires a very careful exploration of the surface.

The VAPS method will rapidly relax most molecular models to within 0.1 kcal/mol of their nearest local energy minima. However, most molecular models normally emerge from the search when their enthalpies are less than 0.02 kcal from the nearest local energy minimum. A second minimization will get most molecular enthalpies to within 0.005 kcal of their local minima. We must also remember that a molecule that is at  $25^\circ\text{C}$  possesses a significant amount of vibrational energy, and so the observed structure of that molecule will be a time-averaged blend of the vibrational modes accessible to that conformation. The geometry of a small molecule, like a decalin, that is at its "static" local energy minimum will change infinitesimally if the molecule's energy is increased by 0.3 kcal/mol. Thus, the minimum energy structures obtained by use of the VAPS method will be almost identical to the structures of the molecules at their "static" local energy minima. The time and effort required to locate the "static" local energy minimum of a molecular model might be best spent on other activities.

QVBMM considers every atom and lone pair in the molecule. For example, methyl groups are not treated as "hard spheres". Thus, the minimization process will be sensitive to eclipsed/gauche arrangements between all vicinal bonds. The QVBMM module does not use algorithms like "ring flip flop" to broaden the search for lower energied conformations of cyclic molecules. Since iso-connected structures (diastereoisomeric, or isomeric and possessing the same connectivity, number and types of bonds) will normally have similar energies, and will be minimized with similar accuracies, then subtracting these energies to give their conformational enthalpy difference will also effectively cancel the minimum energy deviation errors.



### Parameterization of QVBMM

A given stereo-electronic effect must have the same origin in every molecule in which it is found, even though its observed magnitude will depend on the composition and structure of the host molecule. Thus, a properly designed molecular mechanics program, that has been parameterized by using simple molecules, should yield valid data for a very wide range of more complex, analogous systems. The simple molecules, ethane, propane, butane, the butynes, pentane, methyl acetate, acetaldehyde, acetone, acetamide, benzene, the butenes, 1,2-dimethoxyethane, the methyl- and ethylcyclohexanes, the cyclohexenes, the cyclohexylamines, and the halocyclohexanes were therefore used to parameterize QVBMM.

While this paper will concentrate on the design, parameterization and application of the force field to a few representative molecules, the undoubtedly successful performance of QVBMM in the minimization of the structure energies of a wide variety, and complexity, of molecules has clearly confirmed the validity of this approach. In all of these exercises, QVBMM produced structures that were either identical, or almost identical, with those generated by the widely used, commercially available molecular mechanics programs.

### Hydrocarbons

**Ethane:**  $K_{\text{str}}$  was determined, as above, for each bond type and the angular force constant,  $K_{\text{ang}}$ , was assumed to be 0.023 kcal/mol/degree<sup>2</sup>. The constants,  $K_{\text{ster}}$ ,  $K_{\text{rep}}$  and  $K_{\text{tor}}$ , were adjusted empirically until the enthalpy calculated for the gauche conformation of ethane was about 2.79 kcal/mol lower than that of the eclipsed conformation, that is consistent with the experimentally determined value.[5,9] At this point, the C1C bond length was 1.532 Ang., the C1H bond length was 1.09 Ang., and the force constants had been assigned the following values:  $K_{\text{ster}} = 25.000$ ,  $K_{\text{rep}} = 7.500$ , and  $K_{\text{tor}} = 3.750$ .

**Butane:** The possibility of eclipsed or staggered arrangements between each methyl group's hydrogens and the C2-C3 bond ensures that there are nine definitive conformations of n-butane. Conformational features involving the methyl hydrogens were designated as "S" - staggered, "E" -eclipsed, while those about the C2-C3 bond were designated as "A" -anti, and "G" -gauche. Proceeding sequentially from the arrangement between C1-C2 to C2-C3 to C3-C4 gives us the following nine conformations:

SAS	EAS	EAE	- anti carbon skeleton
SGS	EGS	EGE	- gauche carbon skeleton
SES	EES	EEE	- eclipsed carbon skeleton

Using the values of  $K_{\text{str}}$ ,  $K_{\text{ang}}$ ,  $K_{\text{ster}}$ ,  $K_{\text{rep}}$  and  $K_{\text{tor}}$  from the parameterization of ethane, above, QVBMM gave a value of 1.06 kcal/mol for the enthalpy difference between the SGS (gauche) and SAS (anti) butane conformers, in good agreement with the 0.97 kcal/mol value found in the gas phase.[5,9]

The dihedral angle of the carbon chain of the gauche conformation was found to be about 69.6°, somewhat larger than the often cited[5] theoretically calculated value of 65°. Interestingly, if the algorithm to calculate torsional strain was disabled, so that the gauche-butane dihedral angle was dependent only on steric effects, then the gauche dihedral angle was found to be 69.2°. Obviously, at the 69.2° dihedral angle, when steric strain is minimal, the torsional strain must also be very small and just manages to push the dihedral to the observed 69.6° value.

As expected, all of the eclipsed conformations of butane were much higher in energy than the most closely corresponding staggered conformations when only steric effects were invoked, and these energies were significantly greater when torsional effects were also considered.

### Methyl- and ethylcyclohexanes

The methyl- and ethylcyclohexanes and the 1,4-dimethylcyclohexanes are "benchmark" molecules for use in parameterizing a molecular mechanics force field because they have well-defined structures, with limited conformational opportunities. The free energy differences between their conformations and isomeric structures have been widely investigated experimentally. The two methyl groups of the 1,4-dimethylcyclohexanes are remote from each other, and so investigators have assumed the additivity of conformational factors will be valid, since distortions in the ring caused by one methyl group should not be felt significantly by the other.[5] The experimental enthalpy differences between the axial and equatorial isomers of methyl- and ethylcyclohexane have been determined by NMR studies (liquid phase) and estimated to be about 1.75 and 1.60 kcal/mol respectively, while the gas phase enthalpy difference between the cis and trans-isomers of the 1,4-dimethylcyclohexanes was found to be 1.89 kcal/mol.[5,9] As for butane, above, it is logical to expect the conformational energy of the methyl and ethyl groups of the respective cyclohexane to be higher in the gas phase.

Using the values of  $K_{\text{str}}$ ,  $K_{\text{ang}}$ ,  $K_{\text{ster}}$ ,  $K_{\text{rep}}$  and  $K_{\text{tor}}$  from the parameterization of ethane, above, QVBMM estimated the gas phase conformational enthalpy of the methyl group to be 1.84 kcal/mol in methylcyclohexane, 1.84 kcal/mol in 1,4-dimethylcyclohexane, and 1.56 kcal/mol in 1,1,4-trimethylcyclohexane. The conformational enthalpy of the ethyl group was calculated to be 1.62 kcal/mol in ethylcyclohexane. These estimates are consistent with the experimental data cited above.

### Dimethylcyclohexanes

The structure energies of the conformations of the dimethylcyclohexanes were minimized using QVBMM that had been parameterized as above. No other assumptions were made, nor was any further parameterization done. The relative energies of the stable conformations (di-equatorial and

**Table 2.** Conformational enthalpy differences for methylcyclohexanes calculated by the QVBMM force field

	enthalpy	isomer	enthalpy	DH (calc.)
<b>methylcyclohexane</b>				
1eq	5.807	1ax	7.646	1.84
<b>dimethylcyclohexanes</b>				
1eq, 4e	5.776	1eq, 4ax	7.621	1.84
		1ax, 4ax	9.231	3.45
1eq, 4ax	7.621	1ax, 4ax	9.231	1.61
1eq, 3eq	5.778	1eq, 3ax	7.513	1.73
		1ax, 3ax	12.546	6.76
1eq, 3ax	7.513	1ax, 3ax	12.546	5.03
1eq, 2eq	7.848	1eq, 2ax	9.329	1.58
		1ax, 2ax	9.529	1.68
1eq, 2ax	9.329	1ax, 2ax	9.529	0.20

eq - equatorial, ax - axial

axial, equatorial) of the dimethylcyclohexanes are shown in Table 2 (in kcal/mol).

The data in Table 2 for the 1,3- and 1,4-dimethylcyclohexanes, in which the methyl groups are non-adjacent and do not interact, are consistent with the experimental data.[5,9] These data are also internally consistent since the enthalpies for "similar" isomers, e.g. 1e,3a-dimethylcyclohexane and 1e,4a-dimethylcyclohexane, are similar. Thus, these sterically unhindered molecules do show the additivity of conformational effects.

The enthalpy of the 1a,2a-dimethylcyclohexane was calculated to be 0.29 kcal/mol larger than that of the 1a,4a-isomer. This slightly higher enthalpy could arise because the rings of 1a,4a-substituted cyclohexanes are much more flexible than their 1a,2a-analogs, so enabling the 1a,4a-isomer to distribute the strain among its bonds and angles, or relax, much more effectively than the 1a,2a-isomer. However, we must note, again, that when there are no gauche interactions between the substituents, the enthalpies of the isomers are quite similar.

The calculated enthalpy difference between the 1e,2e-isomer and the 1e,4e-isomer was 2.07 kcal/mol, and the calculated enthalpy difference between the 1e,2a-isomer and the 1e,4a-isomer was 1.70 kcal/mol. These enthalpy differences were larger than the 1.06 kcal/mol that would have been predicted for each of the "gauche-butane" interactions between the two methyl groups of these 1,2-disubstituted cyclohexanes. Thus, the QVBMM calculated enthalpies for the congested 1e,2e- and 1e,2a-isomers were about 1.01 and

0.64 kcal/mol respectively larger than would have been predicted.

The calculated enthalpy difference between the two more stable isomers (1e,2e- and 1e,2a-) of the 1,2-dimethylcyclohexanes, 1.58 kcal/mol, was similar to the liquid phase experimental value of 1.5 kcal/mol, but was slightly lower than the gas phase experimental value of 1.9 kcal/mol.[5,9] The calculated difference between the 1e,2e-isomer and the 1a,2a-isomer of 1.68 kcal/mol was much lower than the 2.7 kcal/mol that would have been expected.[5,9]

Since QVBMM provides good data for the uncongested 1,3-dimethylcyclohexanes and the 1,4-dimethylcyclohexanes, but not for the 1,2-dimethylcyclohexanes, the actual geometries of the cyclohexanes generated by QVBMM were compared with those predicted by MM2. The generated data for the cyclohexanes suggested that the dihedral angle between the methyl groups of the 1e,2a-dimethylcyclohexane was 55.55°, the dihedral angle of the 1e,2e-isomer was 61.44°, while that of gauche-butane was 69.6°. The higher enthalpy of these congested 1,2-dimethylcyclohexanes must therefore be due, in part, to the gauche interactions between their two methyl groups, that were much larger than that found in gauche-butane. The rigid 1e,2e-isomer also had fewer opportunities to disperse strain within the ring than the more flexible 1e,2a-isomer.

The MM2 calculated dihedral angle for the methyl groups of the 1e,2a-dimethylcyclohexane was 54.96°, the dihedral angle of the 1e,2e-isomer was 59.59°, while that for gauche-butane was 65°. The trend here is almost identical to that predicted by QVBMM, and so these dimethylcyclohexanes should also experience higher gauche-butane interactions. However, MM2 reported the "correct" enthalpy differences, presumably by making these gauche interactions either slightly attractive, or much less repulsive, and by other careful parameterization.

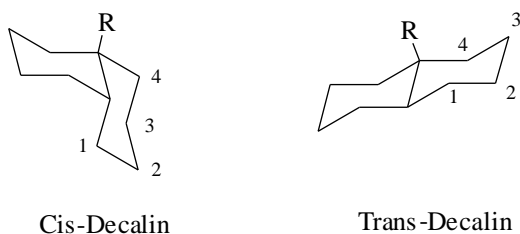
The additivity of conformational effects in congested molecules is known to be invalid, and it is not logical to expect steric and torsional interactions to decrease as the interacting entities are brought closer to each other.[5] Further, we must also remember that at small dihedral angles "gauche-butane" interactions must become intertwined with the "gearing" interactions of the methyl groups and these entropic might be difficult to estimate.[5] If we accept the accuracy of the experimentally determined enthalpies for these dimethylcyclohexanes, then we have obviously identified entropic effects, or stereo-electronic interactions, in the congested 1,2-dimethylcyclohexanes that cannot be logically explained from the data gathered from the isomeric, but uncongested, 1,3- and 1,4-dimethylcyclohexanes.

#### Decalins and hydrindanes

The conformational analysis of the hydrindanes is not as clear-cut as that for the decalins (which can be directly compared with the cyclohexanes). While the relative enthalpies of the decalins can be predicted from the experimental data obtained

for the dimethylcyclohexanes, the fusion of five and six-membered rings to form the hydrindane structures creates more subtle features. In general, the trans-decalins are usually more stable than their cis-isomers, while the cis-hydrindanes are usually more stable than their trans-isomers. The trans-hydrindanes and trans-decalins are conformationally quite rigid while their cis-isomers are much more flexible or conformationally mobile. This greater conformational mobility of the cis-isomers endows them with larger entropies and the ability to relax (disperse strain factors) much better than their trans-isomers. This trend is clearly seen in the data generated by QVBMM for these molecules.

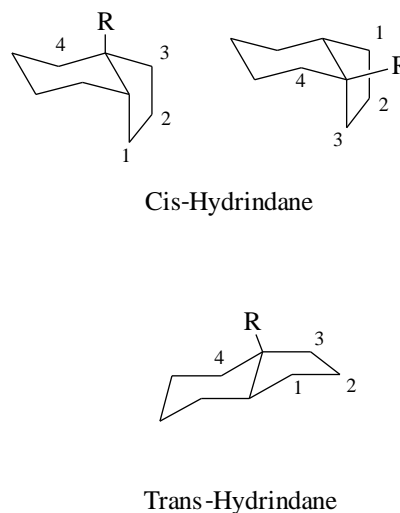
The QVBMM calculated enthalpy difference between trans-decalin and cis-decalin, 2.47 kcal/mol, was in good agreement with the experimental free energy data (liquid phase  $-2.7$  kcal/mol, gas phase  $-3.1$  kcal/mol).[5] The enthalpy differences between the substituted decalins also showed excellent agreement with the experimental data.



**Figure 1.**

The QVBMM calculated enthalpy for the trans-4a-methyldecalin was 0.33 kcal/mol higher than that of the less rigid cis-isomer, instead of the experimentally determined  $0.55 \pm 0.25$  kcal/mol lower.[5] However, since the entropy of cis-4a-methyldecalin must be greater than that of trans-4a-methyldecalin, then the predicted free energy difference will favour the trans-4a-methyldecalin. The enthalpy difference predicted for this pair of molecules, based on the gas phase conformational enthalpy of the methyl group, is about 0.95 kcal/mol.[5] The data for the 4-methyldecalins will be examined more closely to try to clarify this anomalous result.

The experimental free energy difference between the trans-3a-methyl- and the cis-3a-methylhydrindanes[5] is 2.3 kcal/mol and, since the entropy of the more flexible cis-isomers will be greater than that of the trans-isomer, the enthalpy difference should be even larger. QVBMM calculated the enthalpies of the two conformations of cis-3a-methylhydrindane (methyl group axial or equatorial to the cyclohexane ring) to be 21.13 and 22.37 kcal/mol respectively, while the enthalpy of the trans-3a-methylhydrindane was calculated to be 27.23 kcal/mol. The trans - cis enthalpy difference, of about 6 kcal/mol, is slightly larger than expected.



**Figure 2.**

QVBMM calculated the enthalpy of trans-hydrindane to be 1.49 kcal/mol higher than that of cis-hydrindane. However, the experimentally determined heats of combustion of these hydrindanes suggested that the enthalpy of trans-hydrindane was about 1.06 kcal/mol lower than that of cis-hydrindane.[5] This data will also be re-examined.

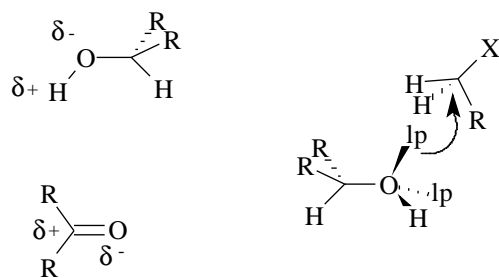
#### Other hydrocarbons

The calculations of the energies of many other cyclic and acyclic hydrocarbons have been performed and the data obtained were consistent with that cited in the literature. However, from the data above, it is apparent that QVBMM occasionally predicts elevated enthalpies (0.5 to 1. kcal/mol higher than found by experiment) for a few rigid and congested cyclic molecules. The fact that these deviations do not occur frequently, and only appear in a few of QVBMM's calculations for a given type of molecule, suggests that the force field might occasionally be unable to move these rigid structures out of metastable conformations to more stable minima. This matter will be explored.

#### Lone pairs in molecular modeling

##### Lone pair orbitals

Dipole - dipole interactions are usually discussed in terms of the partial negative charge on the heteroatom, but not involving the lone pair, as if the lone pair was hidden in an s orbital close to the heteroatom's core. Indeed, some of the most widely used molecular mechanics programs place lone pairs virtually in contact with their host nuclei.



**Figure 3.**

However, the nucleophilicity of an alcohol's oxygen and the basicity of an amine are discussed in terms of the availability of these atoms' lone pairs. These nucleophilic lone pairs are considered to be in  $p$ -, or  $sp^x$  hybridized, orbitals since these linearly extended orbitals should be more effective in nucleophilic processes than spherical  $s$ -orbitals. Further, the electrons in  $s$ -orbitals are closer to their nuclei and are more stable than those in  $p$ -type orbitals, and these features should ensure that transition states for reactions involving  $s$ -orbitals should require closer internuclear approaches, and be higher in energy, than those same transition states involving  $p$ -orbitals.

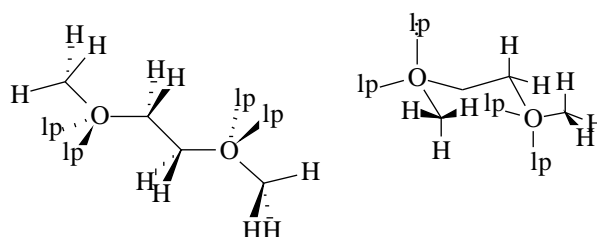
All the modern theoretical methods acknowledge the absolute importance of electron - electron repulsions, and the need to minimize these repulsions. VSEPR theory does not support the notion of a  $s$ -,  $p$ -pair of lone pair orbitals on oxygen, nor the presumed  $s$ - and  $p$ -orbital based lone pairs on the halogens, because these lone electron pairs would be much closer together (and repel each other more) than lone pairs in  $sp^2$  or  $sp^3$  orbitals. Thus QVBMM adds lone pairs on all heteroatoms that do bear lone pairs, and uses  $sp^2$  or  $sp^3$  orbital characteristics for these lone pairs.

VSEPR theory highlights the importance of the appreciable size of the domain of each lone pair. Further, we have traditionally accepted the notion that the sizes of the orbitals of an atom are in proportion to the radius of that atom. Thus, QVBMM assumes that the radial projection of the domain of a lone pair beyond the covalent radius of the host atom is about 1/3 of the covalent radius of the host atom. The lone pair domains for the first row elements are thereby smaller than a hydrogen atom.

#### *Dipole - dipole and lone pair - dipole interactions*

Most molecular mechanics programs treat lone pairs of heteroatoms as "virtual atoms" but otherwise do not consider their roles in the conformational stabilities of molecules. Most lone pair bearing atoms are treated as the negative ends of dipoles and the energies of their dipole - dipole interactions are included the overall potential energy analyses. However, these molecular mechanics programs (like MM2 and

MMX) erroneously predict that the anti-anti-anti-(AAA) conformation of 1,2-dimethoxyethane should be the most stable conformation and be about 2.04 kcal/mol more stable than the gauche-gauche-anti-(+G-GA) conformation.



**Figure 4.**

In fact, 1,2-Dimethoxyethane, above, has been shown to exist in the conformational populations shown in Table 3.

The gauche conformation +G-GA is the most stable. This molecule spends much more time in conformations in which the central  $CH_2$ - $CH_2$  bond is gauche (79%) than in conformations in which this bond is anti (21%).[14] The gauche conformers have larger dipole moments than the anti conformers and so the observed conformational preference cannot be explained in terms of dipole - dipole interactions.

Significantly, the most highly populated conformation, +G-GA, has a geometry that places a methyl group's hydrogen quite closely to an oxygen atom, as in atomic arrangement (A). The electronegativities of hydrogen, carbon and oxygen are 2.20, 2.55 and 3.44 respectively, and so the hydrogen in an R-O-C-H moiety must experience a cascading negative inductive effect, significantly greater than that experienced by a hydrogen in a hydrocarbon. Thus, the partial positive charge on this hydrogen must be comparable to, if not greater than, that on the adjacent carbon atom. The distance between the hydrogen and the oxygen is ideal for the interposition of a lone pair. This clearly indicates the presence of a strong attraction, between the oxygen's lone pair and the positive end of the -O-C-H dipole, that is very similar to the "hydrogen bond" found in alcohols[14b,14c,15a] and reminds us of other attractive interactions in organic chemistry.[15]

The strength of a hydrogen bond is known to depend on several factors, among them being: the temperature, the dielectric constant and polarity of the medium, the R-X-H...:Y-R' distances, and the apparent X-H...Y angle. The strengths of hydrogen bonds in alcohols have been measured and are estimated to be about 3 to 6 kcal/mol. These values are obviously the averages of the interaction energies of the energetically similar conformational arrangements of the interacting atoms. The maximum possible value to be found in the optimal conformations must then be larger than 6 kcal/mol.[15a]

The experimental fact of the greater stability of the +G-GA conformation of 1,2-dimethoxyethane over the AAA conformation indicates that the approach currently used by

**Table 3.** Conformational populations of 1,2-dimethoxyethane.

rotamer	population (%)	
AAA	13	
AAG	3	
GAG	5	subtotal: 21%
GGG	3	
AGA	23	
GGA	53	subtotal: 79%
		total: 100%

most molecular mechanics programs is slightly flawed. A new approach to the handling of lone pairs, dipoles and "gauche effects" in molecular mechanics was therefore needed. A viable alternative approach to the treatment of all of the "components" of a heteroatom is to replace its dipole - dipole interactions with a linear combination of all of the interactions of the neighbouring atomic point charges, and to include the n-n, n-s and n-p interactions generated by its lone pairs. This approach was adopted in the design of QVBMM and will be illustrated below.

#### Parameterization of QVBMM for lone pair - dipole and dipole - dipole interactions

A carbon atom that is attached to an electronegative atom, like oxygen, will be the positive end of that dipole and the bonding pairs of electrons in the other bonds to this carbon might reduce, or further increase, the size of the partial positive charge on the carbon, depending on the electronegativities of the other attached atoms. The overall interaction of a lone pair, or one point charge (of a dipole) with this carbon will be the sum of the interactions of the charge with the carbon's net charge and the bonding electrons that are arrayed about this carbon. This analysis should be applicable to all atoms in a molecule, regardless of their bonding circumstances.

For example, in a linear hydrogen bond,  $R_3X-H \cdots :O-R$ , in which the lone pair does not experience the full repulsion of the X-H and other adjacent bonding pairs, but experiences the full attraction of the positive end of the dipole, the overall interaction must be attractive and maximal. Any departure of the X - H - O angle from  $180^\circ$  must reduce the attractiveness of the interaction if there are no attractive interactions between R and  $R_3$ , or R and X. All of these interactions can be estimated by using the potentials described above, ultimately by equations (11), (12) and (13).

The +G-GA conformation of 1,2-dimethoxyethane has been calculated to be about 0.5 kcal/mol higher in enthalpy than the AAA conformation,[14b,14c] and this relationship was used to parameterize the force constant  $K_{lp}$  of equation (12), when  $K_{lp}$  was for lone pair - bond pair interactions. The

QVBMM calculated enthalpy difference was 0.42 kcal/mol at the optimal numerical value (without units) of  $K_{lp}$  of 32.500.

The structure energy minimized (QVBMM and *ab initio*), molecular model of the +G-GA conformation of 1,2-dimethoxyethane shows that the distance of closest approach of any pair of its non-geminal lone pairs should be greater than 3.0 Angstroms, at which the lone pairs should not repel/influence each other, and so n-n interactions were ignored. This minimum energy structure also confirmed that the "C-H.....O" hydrogen bond is indeed the dominant feature that confers stability on the +G-GA conformation. This "C-H" hydrogen bond will be seen to be of critical importance in understanding the anomeric effect shown by monosaccharides and related molecules.

#### Cyclohexanol and cyclohexylamine derivatives

The very wide range of conformational free energy values of cyclohexanol derivatives shown in Table 4 ought to be a source of concern and should stimulate their re-examination in order to separate the conformational enthalpy from the effects of solvation and aggregation. Similar criticisms[13] have been made of NMR studies of the conformations of monosaccharides. In any event, this fuzzy experimental data does place the following discussion in proper perspective.

The experimentally measured free energy differences between the conformers of the derivatives of cyclohexanol are almost constant and so too are those for the cyclohexylamines, Table 4, suggesting that the axial conformer is insensitive to the size of the group attached to the oxygen or the nitrogen. Thus, these substituent methyl and acetyl groups cannot lie above the ring in the most stable axial conformers.

The assessment of the conformational enthalpy of the hydroxyl and amino groups using QVBMM, that had been parameterized using 1,2-dimethoxyethane and the alkanes, was made more interesting by the discovery of two low energied chair conformers for each of the axial and equatorial isomers, and by the similarities of the enthalpies of these conformers. Both methoxycyclohexane and N-methylcyclohexylamine had one low energied chair conformer for each, axial and equatorial, conformer. These data are shown in Table 5.

If we consider the entropic contributions to the free energy, QVBMM calculated differences in enthalpy between the conformations of cyclohexanol, cyclohexylamine, and their methylated derivatives are in excellent agreement with the experimental values, Table 4, above. The QVBMM data also showed the importance of "C-H" hydrogen bonding, and n-s interactions in the molecular modeling of the cyclohexanols, cyclohexylamines, dimethoxyethane, and, as will eventually be seen, most other oxygenated molecules.

#### Halides

The parameterization and premises that enabled QVBMM to produce good results for nitrogen and oxygen also pro-

**Table 4.** Experimentally determined DG for some cyclohexane substituents.

Substituent	experimental (kcal/mol)[5,9] conformational free energy
-OH	0.3 to 1.5
-OCH <sub>3</sub>	0.4 to 0.74
-OAc	0.36 to 1.60
-NH <sub>2</sub>	1.23 to 1.7
-NHCH <sub>3</sub>	1.29

vided excellent data for fluorine. However, the modeling of molecules that contain chlorine, bromine and iodine offered some challenge, because the conformational enthalpies of these halogen atoms are smaller than would have been predicted from their covalent, or van der Waals, radii. On going from chlorine to iodine, the important factors that had to be considered and balanced included their decreasing electronegativities, increasing covalent radii (and a “topography” that becomes complicated by the symmetry of occupied d-orbitals), increasing C-X bond lengths, and their increasing polarizability (that must increasingly mollify the interactions of the halogen’s electron density with its environment). Indeed, the empirically conducted parameterization confirmed that the polarizability and the electronic symmetry, due to d-orbitals, of the halogens were very important factors.

The parameterization of QVBMM for iodine required additional empirical “tweaking”, having developed excellent parameters for chlorine and bromine, since the data for iodine do not fit into a logical extrapolation from those of the other halogens. Indeed, the experimental data reflect that iodine, the largest halogen, has a conformational free energy intermediate between those of fluorine and chlorine. It is clear that understanding the stereo-electronic requirements of iodine still presents a challenge. Also, it has been suggested that the conformational free energy of chlorine is higher than that of bromine due to entropic effects, and a simple extrapolation here would indicate that the conformational free energy of fluorine ought to be higher than that of chlorine.[5]

QVBMM was therefore parameterized for chlorine, bromine and iodine, but no parameterization was necessary for fluorine. The calculated conformational enthalpies of fluorine, chlorine, bromine and iodine in their halocyclohexane derivatives were 0.36, 0.40, 0.61 and 0.54 kcal/mol respectively. The experimental liquid phase conformational free energies of these atoms are 0.25 - 0.42, 0.53 - 0.64, 0.48 - 0.67 and 0.47 - 0.61, respectively.[5,9] Thus, the calculated conformational enthalpies of the halogens were in excellent agreement with their experimental values.

**Table 5.** Conformational enthalpy differences for some cyclohexanes calculated by the QVBMM force field.

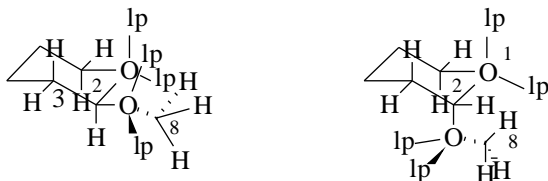
molecule	enthalpy (kcal/mol)	enthalpy difference ( $\Delta H$ )
eq-cyclohexanol	8.074	
	8.770	
ax-cyclohexanol	8.984	
	9.370	0.91
methoxycyclohexane		0.59
eq-cyclohexylamine	6.450	
	6.944	
ax-cyclohexylamine	7.508	
	8.389	1.05
N-methylcyclohexylamine		0.82

#### Anomeric effect [10,13,16]

*Parameterization:* The anomeric effect has bewildered carbohydrate chemists for decades and has been the subject of many reviews. Acetals are among the most widespread, simplest, and yet most important, organic molecules that can show significant n-n interactions and these molecules have been widely studied in order to quantify these n-n interactions. Currently, the anomeric effect is thought to be worth about 2.0 to 4.0 kcal/mol. One of the goals of the development of QVBMM was to address and quantify some aspects of the geometrical features of acetals and to see if these data would shed more light on the anomeric effects. It was very clear that one of the most surprising and powerful pieces of information encountered during the development and parameterization of QVBMM was the role of “C - H” hydrogen bonding in ethers, similar to those encountered in 1,2-dimethoxyethane. Logically, this “C - H” hydrogen bonding must also be present in acetals. However, it was not yet clear how this phenomenon would influence the overall picture of the anomeric effect that we hoped to develop during the study of the molecular modeling of acetals.

Initially QVBMM has been parameterized only using 1,2-dimethoxyethane and the alkanes, but could be used to model the isomers of 2-methoxy-6-methyltetrahydropyran in order to establish a framework for the later inclusion of a potential to assess the sizes of n-n interactions. This exercise, in which the lone pair n-n interaction were ignored, immediately highlighted the importance of the “C - H” hydrogen bonds because those conformers that possessed these hydrogen bonds were more stable than the others. These conformers are shown in figure 5, and are obviously the most stable conformers for

the anomers. Thus, the "C - H" hydrogen bonds dominated the situation and was in part responsible for the "exo-anomeric" conformers to be the most stable for the anomers.



**Figure 5.**

The "C - H" hydrogen bonding interactions between lone pairs of O-1 and the hydrogens of C-6, and between the lone pairs of O-7 and the methyl hydrogens (C-8) were common features of these anomers and so could be ignored. The a-anomer had significant "C - H" hydrogen bonds between O-1 and H-8, O-1 and H-2 (one for each lone pair), O-7 and H-2, and O-7 and axial H-6. The b-anomer had similar interactions between O-1 and H-2, O-7 and H-2, and O-1 and H-8. Thus, there were more stabilizing, attractive, "C - H" hydrogen bonding interactions in the a-anomer than in the b-anomer.

As the lone pair, n-n, interactions were introduced (the potential was activated in the computer program) and increased in size it became clear that the geometry of the a-anomer would not allow this anomer to show n-n interactions, since the lone pairs were too far apart. Even if the conformation was changed to place two lone pairs within interaction distance, the rotation of the C-2 - O-7 bond occurred to minimize/cancel this interaction. Similarly, when the starting geometry of the b-anomer allowed n-n interactions between the O-1 and the O-7 lone pairs, this geometry of the b-anomer changed, by rotation about the C2-O7 bond, to separate the axial lone pairs and so mollify the n-n interaction, while preserving the C-8 - H hydrogen bond with O-1. The geometries and energies of these systems were therefore significantly influenced by the n-n interactions, the "C - H" hydrogen bonds, and the conformational flexibility of the anomers.

We then considered using the alkoxides of the simple anomeric 2-hydroxytetrahydropyrans as suitable models for parameterizing the force field for n-n interactions, since these molecules would not have aglyconyl protons and so would not be subject to the "C - H" hydrogen bonding complications mentioned above. These alkoxides would also be isoelectronic with the 2-fluorotetrahydropyrans, whose data have not been reported to date, and so are particularly interesting. Unfortunately, our experimental investigations of the acylations of some glycosyl alkoxides showed that it might not be possible to establish a likely value for the conformational enthalpy difference between these anomeric alkoxide by kinetic studies and this option was terminated.[17] The

very interesting details of this investigation will be published elsewhere.

The experimentally measured[18] conformational equilibrium between the anomeric 2-methoxytetrahydropyrans favoured the axial a-anomeric isomer, by approximately a 4:1 ratio. Therefore, the n-n interactions in QVBMM were parameterized to make the b-anomeric 2-methoxytetrahydropyran 0.78 kcal/mol less stable than the isomeric a-anomeric acetal. Interestingly, halving the empirically determined value of  $K_{n-n}$  did not change the conformational enthalpy difference between the anomeric 2-methoxytetrahydropyrans significantly, showing that these anomers were able to mollify the destabilizing n-n interactions by conformational changes that adjusted the sizes of the gauche interactions and the stabilizing "C - H" hydrogen bonding. The force field that was applied to the corresponding anomeric alkoxides, that could not benefit from these aglyconyl "C - H" hydrogen bonds, and showed that the b-anomeric alkoxide was less stable than the a-anomeric alkoxide by 1.25 kcal/mol. This allowed us to generate a qualitative understanding of the relative importance of the n-n repulsions and the "C - H" hydrogen bonds.

#### *Enthalpy differences between some Anomeric conformations of tetrahydropyrans*

The importance of the anomeric effects in the glycosyl halides is very well known,[13,18] and these halides do not possess the complicating factor of "C - H" hydrogen bonds between O-5, the ring oxygen, and an aglyconyl group. The experimentally measured[18] conformational preference for the axial, a-anomeric, isomer of the 2-bromotetrahydropyrans and the 2-chlorotetrahydropyrans are about 96% in each case and this value is also found for the 2-halo-4-methyltetrahydropyrans. These axial/equatorial ratios indicate conformational free energy differences of about 1.82 kcal/mol for these anomers. Without any further parameterization QVBMM calculated the conformational enthalpy difference between the anomeric 2-chloro-6-methyltetrahydropyrans to be 1.48 kcal/mol, and that between the 2-bromo-6-methyltetrahydropyrans to be 1.59 kcal/mol, consistent with the experimental data.

Interestingly, the 6-methyl group seemed to exert a significant influence on the anomeric effect in the tetrahydropyrans. The QVBMM calculated conformational enthalpy difference between the anomeric 2-chlorotetrahydropyrans was 0.67 kcal/mol, the anomeric 2-chloro-4-methyltetrahydropyrans was 0.72 kcal/mol, the anomeric 2-bromotetrahydropyrans was 0.76 kcal/mol, and the anomeric 2-bromo-4-methyltetrahydropyrans was 0.78 kcal/mol. The experimentally measured conformational enthalpy differences between the 2-fluorotetrahydropyrans seems not to have been reported, however, QVBMM calculated the conformational enthalpy difference between the anomers of 2-fluoro-6-methyltetrahydropyran to be 1.59 kcal/mol, the anomers of 2-fluoro-4-methyltetrahydropyran to be 0.83 kcal/mol and the anomers of 2-fluoro-tetrahydropyran to be 0.75 kcal/mol.

Thus, the experimental, and QVBMM calculated, conformational enthalpy difference between the 2-halo-6-methyltetrahydropyrans were about 0.5 to 1 kcal/mol greater than the calculated data for the 2-halo-4-methyltetrahydropyrans and the 2-halotetrahydropyrans, the latter of which agreed quite closely. The conformational enthalpy difference between the anomeric 2-methoxy-6-methyltetrahydropyrans was calculated to be about 1.50 kcal/mol, while that of the anomeric 2-methoxyltetrahydropyrans was calculated to be about 0.78 kcal/mol, showing that the anomeric heteroatom was not the factor that influenced the calculations. Further, it cannot be coincidental that the two sets of tetrahydropyrans that are not substituted at C-6, and so have uncrowded acetal moieties, show the similar enthalpy characteristics, different from the 6-methyltetrahydropyrans.

The 1,3-disposition of the equatorial 6-methyl group with respect to the anomeric atom indicated that this effect and could not be a direct steric interaction. Thus, it seemed very likely that the 6-methyl group exerted a "buttr essing" effect[5] on the lone pairs of O-1, so preventing them from distorting away from the anomeric interactions, and so forcing these lone pairs to interact more strongly with the anomeric atom and its substituents. This buttressing effect was more significant in the b-anomers because these molecules have the equatorial anomeric atom, the ring oxygen and the equatorial 6-methyl group, in the same plane. The axial orientation of the anomeric atom of the a-anomer would effectively keep this atom away from the ring oxygen's lone pairs. Hence, we ought to be a bit more concerned about the roles of "simple" alkyl substituents in the evaluation of the anomeric effects.

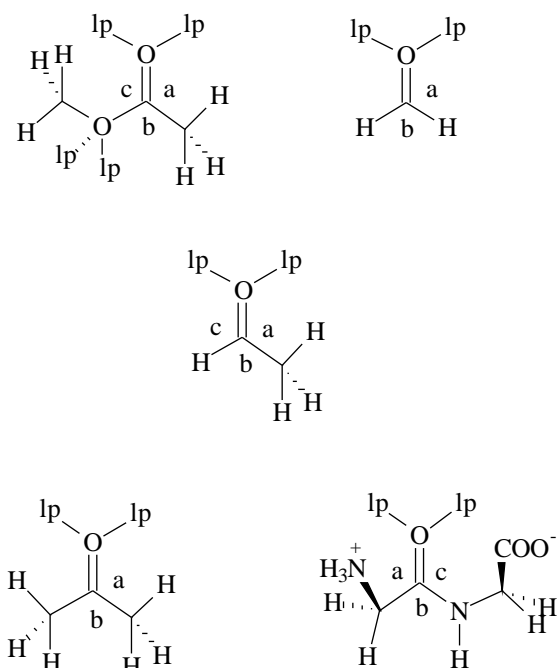
It is also well known that the nature of the solvent dramatically affects the percentage of the axial b-anomer of the 2-substituted tetrahydropyrans present at equilibrium.[18] In the case of the 2-methoxyltetrahydropyrans, the equilibrium distribution of the axial b-anomer ranged from a high of 82% in carbon tetrachloride to a low of 52% in water. Thus, either the stabilities of the b-anomers increased in polar solvents, or the stabilities of the a-anomers decreased in these polar solvents. The net effect, however, was the reduction of the sizes of the anomeric effects in polar solvents, and the enhancement of sizes of the anomeric effects in non-polar solvents. QVBMM does not take external (intermolecular) solvation into account directly, but rather factors the effects of solvation into the derivation of equation (9), above. In this context, therefore, the enthalpy values calculated by QVBMM for the simple 2-substituted tetrahydropyrans best match those expected from equilibria studied in solvents like chloroform and acetone, whereas the data calculated for the 6-methyltetrahydropyrans best match those observed in a wider range of relatively non-polar solvents.

The conformational enthalpy difference between the b-anomeric and the a-anomeric 2-hydroxy-6-methyltetrahydropyrans and the anomeric 2-amino-6-methyltetrahydropyrans were calculated to be about 1.02 kcal/mol and -0.28 kcal/mol respectively, showing the "reverse" anomeric effect (in which the b-anomer is predicted to be more stable than the a-

anomer) in the amino-sugar. Also, without any further parameterization, the "reverse" anomeric effects in the protonated anomers of 2-amino- and 2-hydroxy-6-methyltetrahydropyrans were excellently simulated by the QVBMM calculations.

#### Alkenes and carbonyl compounds

The bond angles of carbonyl compounds present a special challenge to molecular modelers because of the range they embrace, see Table 6 below. Esters are particularly interesting, and a survey of the bond angles found in the x-ray crystallographic data of monosaccharide acetates produced average bond angles shown below.[13] Similar angle sizes were also found in benzoates.[13] The bond angles surrounding the p-bond in simple alkenes correspond quite closely to those of simple aldehydes and ketones, but are quite different from those of molecules that contain lone pair bearing atoms.



**Figure 6.**

These variations in the sizes of the bond angles that surround the p-bond must be determined, at least in part, by the polarizability or "hardness" of the atoms attached to the p-bond, and the presence of lone pairs.

The parameterization of QVBMM to simulate the features of alkenes and carbonyl compounds was in fact the challenge of empirically selecting appropriate values for the



**Table 6.** Some representative bond angles subtended by  $sp^2$  carbon atoms [2a, 13].

compound	angle a [°]	angle b [°]	angle c [°]
H <sub>2</sub> C=O	121.7	116.5	121.7
(CH <sub>3</sub> ) <sub>2</sub> C=O	122.0	116.0	122.0
CH <sub>3</sub> -CO-H	123.9	117.5	118.6
R-O-CO-CH <sub>3</sub>	125.64	111.49	122.84
alanylglycine	119	115	126
F <sub>2</sub> C=O	126.2	107.7	
Cl <sub>2</sub> C=O	124.1	111.8	
Br <sub>2</sub> C=O	123.8	112.3	
F <sub>2</sub> C=CH <sub>2</sub>	124.7	110.6	
Cl <sub>2</sub> C=CH <sub>2</sub>	123.0	114.1	
(CH <sub>3</sub> ) <sub>2</sub> C=CH <sub>2</sub>	122.2	115.6	

for bond angle labels see figure 6.

“unstrained” bond angles and the force constants,  $K_{ang}$ , to be used in the Hooke’s law potentials that govern the bond angle vibrations. This exercise eventually provided the data shown in Table 7, and the force field then produced excellent results in the simulations of carbonyl compounds.

Note that these empirically determined force constants are all still quite small, except those for bond angles of esters, amides and angles involving lone pairs, again reinforcing the long-held view that bond angles are very easily deformed.

## Conclusion

The successful implementation of QVBMM in STR3DI.EXE represents a first step in the formal quantitation of the common lone pair interactions and the concepts embodied in VSEPR theory. This work validates the notion that molecular mechanics programs can be parameterized by using very simple molecules and the resulting programs should then be applicable to a wide range of simulations. The design of QVBMM involved relatively few assumptions and emphasized the general applicability of only a few concepts to all molecular types.

The molecular graphics capabilities of STR3DI.EXE were developed to enable the execution of this program on modestly capable MSDOS computers in order to stimulate the introduction of molecular modeling into the undergraduate curriculum. These capabilities have been exploited in a number of “computer laboratory” exercises for undergradu-

**Table 7.** Empirically determined “best values” for unstrained bond angles and their deformation force constants (for bond angle labels at  $sp^2$  atoms see figure 6)

saturated systems ( $sp^3$ apex atoms)		
atom type	unstrained bond angle	$K_{ang}$
first row $sp^3$ atoms, not including lone pairs	109.5	0.022
one first row $sp^3$ atom and 2 lone pairs	111	0.115
two first row $sp^3$ atoms and 1 lone pair	110	0.115
unsaturated systems ( $sp^2$ apex atoms)		
angle type	unstrained bond angle	$K_{ang}$
1-b and 5-b	109.5	0.022
2-b	116.5	0.022
3-a	119	0.022
3-b and 4-b	119	0.022
3-c	121.75	0.022
1-c and 5-c	120	0.115
1-a, 3-a, 4-a, 5-a	126	0.055
at delocalized p-atoms	120	0.055
angles involving lone pairs	120	0.115
others	120.5	0.022

ates[19] and in several research projects.[7,10,13] Thus, the implementation of the QVBMM module in STR3DI.EXE will now bring even more powerful molecular modeling capabilities within the reach of modestly equipped academic situations.

## References

- (a) Boese, R.; Haumann, T.; Stellberg, P. *Adv. Mol. Struct. Res.* **1995**, *1*, 201. (b) Hehre, W. J.; Radom, L.; Schleyer, P. V. R.; Pople, J. A. *Ab initio Molecular Orbital Theory*, Wiley, New York, 1986 and references cited therein.
- (a) Gillespie, R. J.; Hargittai, I. *The VSEPR Model of Molecular Geometry*, Allyn and Bacon, Boston, 1991 and references cited therein. (b) Gillespie, R. *J. Chem. Soc. Rev.* **1992**, 59 and references cited therein.

3. (a) Kollman, P. A. *Acc. Chem. Res.* **1996**, 29, 461. (b) Allinger, N. L. *Molecular Mechanics in Accurate Molecular Structures*, Domenicano A.; Hargittai, I. Eds.; Oxford University, New York, 1992; and references cited therein. (c) Rasmussen, K. *AIP Conf. Proc. (E.C.C.C. I Computational Chemistry)* **1995**, 330, 269. (d) Jaime, C. *Trends Org. Chem.* **1993**, 4, 469.
4. Dewar, M. J. S.; Jie, C.; Yu, J. *Tetrahedron* **1993**, 49, 5003 and references cited therein.
5. Eliel, E. L.; Wilen, S. H.; Mander, L. N. *Stereochemistry of Organic Compounds*; Wiley-Interscience, New York, 1994 and references cited therein.
6. Carey, F. A.; Sundberg, R. J. *Advanced Organic Chemistry*; Part A, 3rd Edn., Plenum, New York, 1990.
7. (a) Box, V. G. S. *Heterocycles* **1991**, 32, 2023. (b) Box, V. G. S. *Heterocycles* **1992**, 34, 1631.
8. (a) Lavallee, D. K.; Zeltmann, A. H. *J. Am. Chem. Soc.* **1974**, 96, 5552. (b) Jeffrey, G. A. *Conformational Studies in the Solid State: Extrapolation to Molecules in Solution*, in *Carbohydrates in Solution*; Isbell, H. S. Ed.; Advances in Chemistry Series, No. 117, American Chemical Society, Washington, D.C., 1973.
9. Hirsch, J. A. *Top. Stereochem.* **1967**, 1, 199.
10. Box, V. G. S. *Heterocycles* **1991**, 32, 795
11. March, J. A. *Advanced Organic Chemistry*, 4th Edn., Wiley-Interscience, New York, 1992.
12. Cutnell, J. D.; Johnson, K. W. *Physics*, 3rd Edn., John Wiley & sons, New York, 1995.
13. Box, V. G. S. *Heterocycles* **1990**, 31, 1157
14. (a) Astrup, E. E. *Acta chem. Scand.* **1979**, A33, 655. (b) Tsuzuki, S.; Uchimaru, T.; Tanabe, K.; Hirano, T. *J. Phys. Chem.* **1993**, 97, 1346. (c) Jaffe, R. L.; Smith, G. D.; Yoon, D. Y. *J. Phys. Chem.* **1993**, 97, 12745.
15. (a) Mascal, M. *Contemporary Org. Syn.* **1994**, 1, 31. (b) Juaristi, E. *J. Chem. Educ.* **1979**, 56, 438. (c) Beveridge, D. L.; Radna, R. *J. Am. Chem. Soc.* **1971**, 93, 3759. (d) Kingsbury, C. A. *J. Chem. Educ.* **1979**, 56, 431
16. (a) Box, V. G. S. *Heterocycles* **1982**, 19, 1939. (b) Box, V. G. S. *Heterocycles* **1984**, 22, 891.
17. Box, L. L.; Box, V. G. S.; Brown, D. A.; Das, K. K. unpublished work
18. (a) Tvaroska, I.; Bleha, T. *Adv. Carbohydr. Chem. Biochem.* **1989**, 47, 45. (b) Juaristi, E.; Cuevas, G. *Tetrahedron* **1992**, 48, 5019.
19. (a) Box, V. G. S. *J. Chem. Educ.* **1991**, 68, 662. (b) Box, V. G. S. *J. Chem. Educ.* **1994**, 70, 236.