Chapter 3

Molecular Mechanics

> *We don't give a damn where the electrons are.*
> Words to the author, from the president of a well-known chemical company, emphasizing his firm's position on basic research.

3.1 PERSPECTIVE

Molecular mechanics (MM) [1] is based on a mathematical model of a molecule as a collection of balls (corresponding to the atoms) held together by springs (corresponding to the bonds) (Fig. 3.1). Within the framework of this model, the energy of the molecule changes with geometry because the springs resist being stretched or bent away from some "natural" length or angle, and the balls resist being pushed too closely together. The *mathematical* model is thus conceptually very close to the intuitive feel for molecular energetics that one obtains when manipulating molecular models of plastic or metal: the model resists distortions (it may break!) from the "natural" geometry that corresponds to the bond lengths and angles imposed by the manufacturer, and in the case of space-filling models, atoms cannot be forced too closely together. The MM model clearly ignores electrons.

The principle behind MM is to express the energy of a molecule as a function of its resistance toward bond stretching, bond bending, and atom crowding, and to use this energy equation to find the bond lengths, angles, and dihedrals corresponding to the minimum-energy geometry – or more precisely, to the various possible potential energy surface minima (chapter 2). In other words, MM uses a conceptually *mechanical* model of a molecule to find its minimum-energy geometry (for flexible molecules, the geometries of the various conformers). The form of the mathematical expression for the energy, and the parameters in it, constitute a *forcefield*, and MM methods are sometimes called forcefield methods. The term arises because the negative of the first derivative of the potential energy of a particle with respect to displacement along some direction is the force on the particle; a "forcefield" $E(x, y, z$ coordinates of atoms) can be differentiated to give the force on each atom.

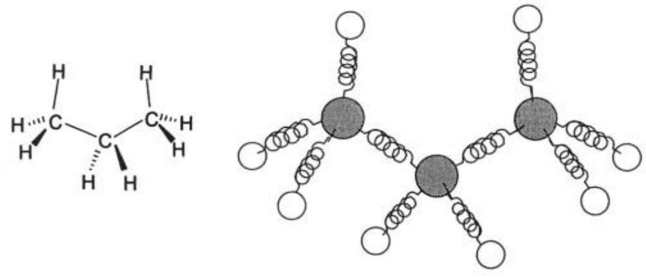

Figure 3.1. Molecular mechanics (the forcefield method) considers a molecule to be a collection of balls (the atoms) held together by springs (the bonds).

The method makes no reference to electrons, and so cannot (except by some kind of empirical algorithm) throw light on *electronic* properties like charge distributions or nucleophilic and electrophilic behaviour. Note that MM implicitly uses the Born–Oppenheimer approximation, for only if the nuclei experience what amounts to a static attractive force, whether from electrons or springs, does a molecule have a distinct geometry (section 2.3).

An important point, which students sometimes have a problem with, is that the concept of a *bond* is central to MM, but not essential – although often useful – in *electronic* structure calculations. In MM a molecule is defined by the atoms and the bonds, which latter are regarded almost literally as springs holding the atoms together. Usually, bonds are placed where the rules for drawing structural formulas require them, and to do a MM calculation you must specify each bond as single, double, etc., since this tells the program how strong a bond to use (sections 3.2.1 and 3.2.2). In an electronic structure calculation – *ab initio* (chapter 5), semiempirical (SE) (chapter 6), and density functional theory (chapter 7) – a molecule is defined by the relative positions of its atomic nuclei, the charge, and the "multiplicity" (which follows easily from the number of unpaired electrons). An oxygen nucleus and two protons with the right *x, y, z* coordinates, no charge, and multiplicity one (no unpaired electrons) is a water molecule. There is no need to mention bonds here, although the chemist might wish to somehow extract this useful concept from this picture of nuclei and electrons. This can be done by calculating the electron density and associating a bond with, for example, a path along which electron density is concentrated, but there is no unique definition of a bond in electronic structure theory. It is worth noting, too, that in some graphical interfaces used in computational chemistry bonds are specified by the user, while in others they are shown by the program depending on the separation of pairs of atoms. The novice may find it disconcerting to see a specified bond still displayed even when a change in geometry has moved a pair of atoms far apart, or to see a bond vanish when a pair has moved beyond the distance recognized by some fudge factor.

Historically [2], MM seems to have begun as an attempt to obtain quantitative information about chemical reactions at a time when the possibility of doing quantitative quantum mechanical (chapter 4) calculations on anything much bigger than the hydrogen molecule seemed remote. Specifically, the principles of MM, as a potentially general method for studying the variation of the energy of molecular systems with their geometry, were formulated in 1946 by Westheimer[1] and Meyer [3a], and by Hill [3b], In this same year Dostrovsky, Hughes[2] and Ingold[3] independently applied MM concepts of to the quantitative analysis of the S_N2 reaction, but they do not seem to have recognized the potentially wide applicability of this approach [3c]. In 1947 Westheimer [3d] published detailed calculations in which MM was used to estimate the activation energy for the racemization of biphenyls.

Major contributors to the development of MM have been Schleyer[4] [2b,c] and Allinger[5] [1c,d]; one of Allinger's publications on MM [1d] is, according to the Citation Index, one of the most frequently cited chemistry papers. The Allinger group has, since the 1960s, been responsible for the development of the "MM-series" of programs, commencing with MM1 and continuing with the currently widely-used MM2 and MM3, and the recent MM4 [4]. MM programs [5] like Sybyl and UFF will handle molecules involving much of the periodic table, albeit with some loss of accuracy that one might expect for trading breadth for depth, and MM is the most widely-used method for computing the geometries and energies of large biological molecules like proteins and nucleic acids (although recently SE (chapter 6) and even *ab initio* (chapter 5) methods have begun to be applied to these large molecules).

3.2 THE BASIC PRINCIPLES OF MM

3.2.1 Developing a forcefield

The potential energy of a molecule can be written

$$E = \sum_{\text{bonds}} E_{\text{stretch}} + \sum_{\text{angles}} E_{\text{bend}} + \sum_{\text{dihedrals}} E_{\text{torsion}} + \sum_{\text{pairs}} E_{\text{nonbond}} \qquad (3.1)$$

where E_{stretch} etc. are energy contributions from bond stretching, angle bending, torsional motion (rotation) around single bonds, and interactions between atoms or groups which are nonbonded (not directly bonded together). The sums are over all the bonds, all the angles defined by three atoms A–B–C, all the dihedral angles defined by four atoms A–B–C–D, and all pairs of significant nonbonded interactions. The mathematical form of these terms and the parameters in them constitute a particular forcefield. We can make this clear by being more specific; let us consider each of these four terms.

The bond stretching term. The increase in the energy of a spring (remember that we are modelling the molecule as a collection of balls held together by springs) when it is

---

[1] Frank H. Westheimer, born Baltimore, Maryland, 1912. Ph.D. Harvard 1935. Professor University of Chicago, Harvard.

[2] Edward D. Hughes, born Wales, 1906. Ph.D. University of Wales, D.Sc. University of London. Professor, London. Died 1963.

[3] Christopher K. Ingold, born London 1893. D.Sc. London 1921. Professor Leeds, London. Knighted 1958. Died London 1970.

[4] Paul von R. Schleyer, born Cleveland, Ohio, 1930. Ph.D. Harvard 1957. Professor Princeton; institute codirector and professor University of Erlangen-Nürnberg, 1976–1998. Professor University of Georgia.

[5] Norman L. Allinger, born Rochester New York, 1930. Ph.D. University of California at Los Angeles, 1954. Professor Wayne State University, University of Georgia.

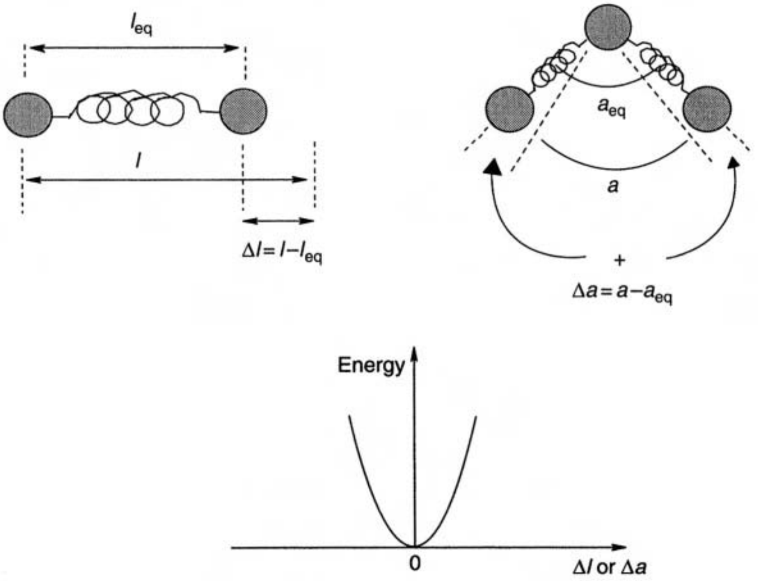

Figure 3.2. Changes in bond lengths or in bond angles result in changes in the energy of a molecule. Such changes are handled by the E_{stretch} and E_{bend} terms in the MM forcefield. The energy is approximately a quadratic function of the change in bond length or angle.

stretched (Fig. 3.2) is approximately proportional to the square of the extension:

$$\Delta E_{\text{stretch}} = k_{\text{stretch}}(l - l_{\text{eq}})^2$$

where k_{stretch} is the proportionality constant (actually one-half the *force constant* of the spring or bond [6]; but note the warning about identifying MM force constants with the traditional force constant from, say, spectroscopy – see section 3.3); the bigger k_{stretch}, the stiffer the bond/spring – the more it resists being stretched; l is the length of the bond when stretched; and l_{eq} is the equilibrium length of the bond, its "natural" length.

If we take the energy corresponding to the equilibrium length l_{eq} as the zero of energy, we can replace $\Delta E_{\text{stretch}}$ by E_{stretch}:

$$E_{\text{stretch}} = k_{\text{stretch}}(l - l_{\text{eq}})^2 \tag{3.2}$$

The angle bending term. The increase in energy of system ball–spring–ball–spring–ball, corresponding to the triatomic unit A–B–C (the increase in "angle energy") is approximately proportional to the square of the increase in the angle (Fig. 3.2); analogously to Eq. (3.2):

$$E_{\text{bend}} = k_{\text{bend}}(a - a_{\text{eq}})^2 \tag{3.3}$$

where k_{bend} is a proportionality constant (one-half the angle bending force constant [6]; note the warning about identifying MM force constants with the traditional force constant from, say, spectroscopy –see section 3.3); a is the size of the angle when distorted; and a_{eq} is the equilibrium size of the angle, its "natural" value.

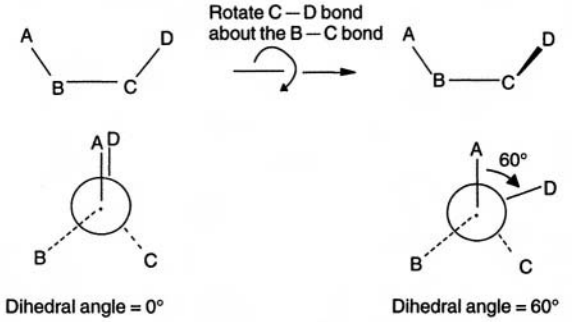

Figure 3.3. Dihedral angles (torsional angles) affect molecular geometries and energies. The energy is a periodic (cosine or combination of cosine functions) function of the dihedral angle; see, for example, Figs. 3.4 and 3.5.

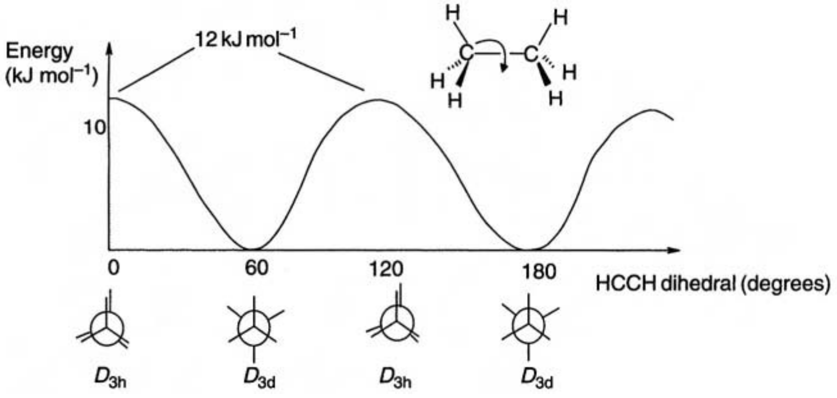

Figure 3.4. Variation of the energy of ethane with dihedral angle. The curve can be represented as a cosine function.

The torsional term. Consider four atoms sequentially bonded: A–B–C–D (Fig. 3.3). The dihedral angle or torsional angle of the system is the angle between the A–B bond and the C–D bond as viewed along the B–C bond. Conventionally this angle is considered positive if regarded as arising from clockwise rotation (starting with A–B covering or eclipsing C–D) of the back bond (C–D) with respect to the front bond (A–B). Thus in Fig. 3.3 the dihedral angle A–B–C–D is 60° (it could also be considered as being –300°). Since the geometry repeats itself every 360°, the energy varies with the dihedral angle in a sine or cosine pattern, as shown in Fig. 3.4 for the simple case of ethane. For systems A–B–C–D of lower symmetry, like butane (Fig. 3.5), the torsional potential energy curve is more complicated, but a *combination* of sine or cosine functions will reproduce the curve:

$$E_{\text{torsion}} = k_0 + \sum_{r=1}^{n} k_r[1 + \cos(r\theta)] \tag{3.4}$$

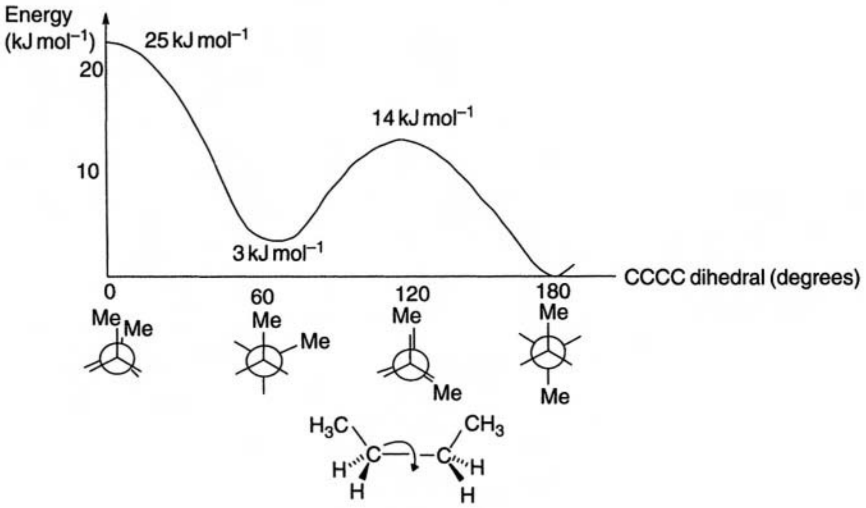

Figure 3.5. Variation of the energy of butane with dihedral angle. The curve can be represented by a sum of cosine functions.

The nonbonded interactions term. This represents the change in potential energy with distance apart of atoms A and B that are not directly bonded (as in A–B) and are not bonded to a common atom (as in A–X–B); these atoms, separated by at least two atoms (A–X–Y–B) or even in different molecules, are said to be nonbonded (with respect to each other). Note that the A–B case is accounted for by the bond stretching term E_{stretch}, and the A–X–B term by the angle bending term E_{bend}, but the nonbonded term E_{nonbond} is, for the A–X–Y–B case, superimposed upon the torsional term E_{torsion}: we can think of E_{torsion} as representing some factor inherent to resistance to rotation about a (usually single) bond X–Y (MM does not attempt to explain the theoretical, electronic basis of this or any other effect), while for certain atoms attached to X and Y there may also be nonbonded interactions.

The potential energy curve for two nonpolar nonbonded atoms has the general form shown in Fig. 3.6. A simple way to approximate this is by the so-called Lennard–Jones 12-6 potential [7]:

$$E_{\text{nonbond}} = k_{\text{nb}}\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{3.5}$$

where r is the distance between the centers of the nonbonded atoms or groups.

The function reproduces the small attractive dip in the curve (represented by the negative term) as the atoms or groups approach one another, then the very steep rise in potential energy (represented by the raising the positive, repulsive term raised to a large power) as they are pushed together closer than their van der Waals radii. Setting $dE/dr = 0$, we find that for the energy minimum in the curve the corresponding value of r is $r_{\text{min}} = 2^{1/6}\sigma$, i.e.

$$\sigma = 2^{-1/6} r_{\text{min}} \tag{3.6}$$

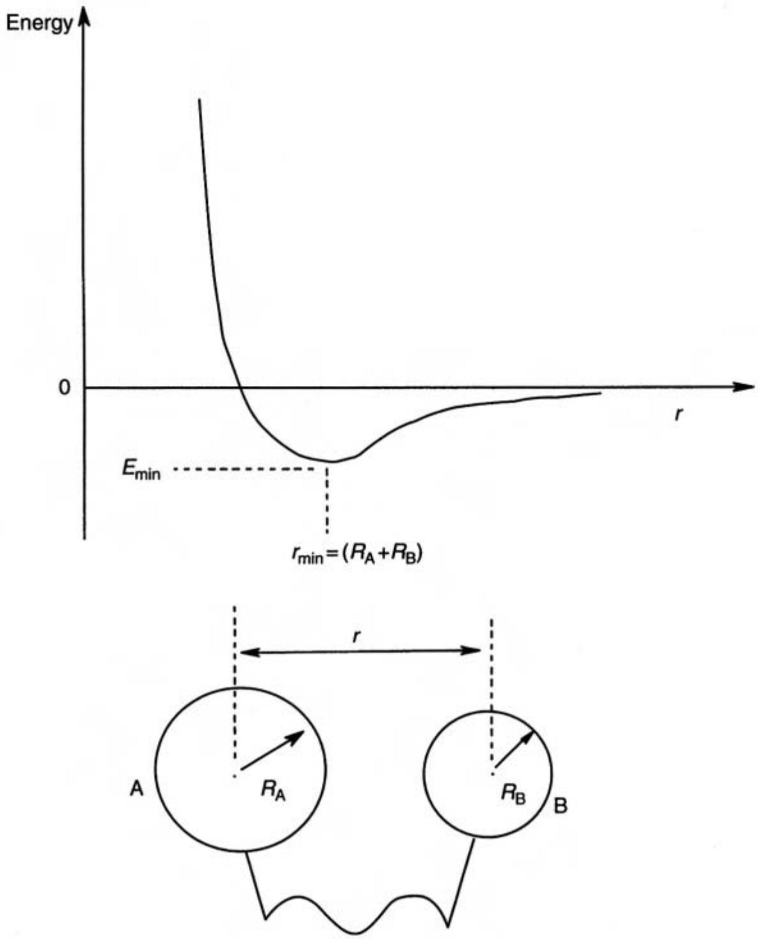

Figure 3.6. Variation of the energy of a molecule with separation of nonbonded atoms or groups. Atoms/groups A and B may be in the same molecule (as indicated here) or the interaction may be intermolecular. The minimum energy occurs at van der Waals contact. For small nonpolar atoms or groups the minimum energy point represents a drop of a few kJ mol^{-1} ($E_{min} = -1.2$ kJ mol^{-1} for CH_4/CH_4), but short distances can make nonbonded interactions destabilize a molecule by many kJ mol^{-1}.

If we assume that this minimum corresponds to van der Waals contact of the nonbonded groups, then $r_{min} = (R_A + R_B)$, the sum of the van der Waals radii of the groups A and B. So

$$2^{1/6}\sigma = (R_A + R_B)$$

and so

$$\sigma = 2^{-1/6}(R_A + R_B) = 0.89(R_A + R_B) \tag{3.7}$$

Thus σ can be calculated from r_{min} or estimated from the van der Waals radii.

Setting $E = 0$, we find that for this point on the curve $r = \sigma$, i.e.

$$\sigma = r(E = 0) \tag{3.8}$$

If we set $r = r_{\text{min}} = 2^{1/6}\sigma$ (from Eq. (3.6)) in Eq. (3.5), we find

$$E(r = r_{\text{min}}) = (-1/4)k_{\text{nb}}$$

i.e.

$$k_{\text{nb}} = -4E(r = r_{\text{min}}) \tag{3.9}$$

So k_{nb} can be calculated from the depth of the energy minimum.

In deciding to use equations of the form (3.2)–(3.5) we have decided on a particular MM forcefield. There are many alternative forcefields. For example, we might have chosen to approximate E_{stretch} by the sum of a quadratic and a cubic term:

$$E_{\text{stretch}} = k_{\text{stretch}}(l - l_{\text{eq}})^2 + k(l - l_{\text{eq}})^3$$

This gives a somewhat more accurate representation of the variation of energy with length. Again, we might have represented the nonbonded interaction energy by a more complicated expression than the simple 12-6 potential of Eq. (3.5) (which is by no means the best form for nonbonded repulsions). Such changes would represent changes in the forcefield.

3.2.2 Parameterizing a forcefield

We can now consider putting actual numbers, k_{stretch}, l_{eq}, k_{bend}, etc., into Eqs (3.2)–(3.5), to give expressions that we can actually use. The process of finding these numbers is called *parameterizing* (or parametrizing) the forcefield. The set of molecules used for parameterization, perhaps 100 for a good forcefield, is called the *training set.* In the purely illustrative example below we use just ethane, methane and butane.

Parameterizing the bond stretching term. A forcefield can be parameterized by reference to experiment (empirical parameterization) or by getting the numbers from high-level *ab initio* or density functional calculations, or by a combination of both approaches. For the bond stretching term of Eq. (3.2) we need k_{stretch} and l_{eq}. Experimentally, k_{stretch} could be obtained from IR spectra, as the stretching frequency of a bond depends on the force constant (and the masses of the atoms involved) [8], and l_{eq} could be derived from X-ray diffraction, electron diffraction, or microwave spectroscopy [9].

Let us find k_{stretch} for the C/C bond of ethane by *ab initio* (chapter 5) calculations. Normally high-level *ab initio* calculations would be used to parameterize a forcefield, but for illustrative purposes we can use the low-level but fast STO-3G method [10]. Eq. (3.2) shows that a plot of E_{stretch} against $(l - l_{\text{eq}})^2$ should be linear with a slope of k_{stretch}. Table 3.1 and Fig. 3.7 show the variation of the energy of ethane with stretching of the C/C bond, as calculated by the *ab initio* STO-3G method. The equilibrium bond length has been taken as the STO-3G length:

$$l_{\text{eq}}(\text{C–C}) = 1.538\,\text{Å} \tag{3.10}$$

Table 3.1. Change in energy as the C–C bond in CH_3–CH_3 is stretched away from its equilibrium length

C–C length, l	$l - l_{eq}$	$(l - l_{eq})^2$	$E_{stretch}$ (kJ mol^{-1})
1.538	0	0	0
1.550	0.012	0.00014	0.29
1.560	0.022	0.00048	0.89
1.570	0.032	0.00102	1.86
1.580	0.042	0.00176	3.15
1.590	0.052	0.00270	4.75
1.600	0.062	0.00384	6.67

The calculations are *ab initio* (STO-3G; chapter 5). Bond lengths are in Å.

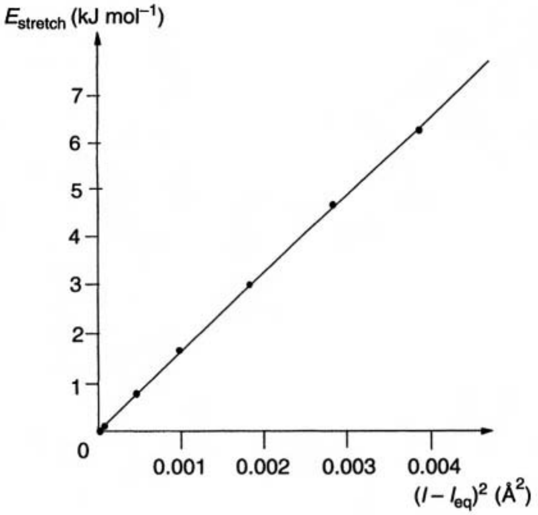

Figure 3.7. Energy vs. the square of the extension of the C–C bond in CH_3–CH_3. The data in Table 3.1 were used.

The slope of the graph is

$$k_{stretch}(\text{C–C}) = 1735\,\text{kJ mol}^{-1}\,\text{Å}^{-2} \tag{3.11}$$

Similarly, the CH bond of methane was stretched using *ab initio* STO-3G calculations; the results are

$$l_{eq}(\text{C–H}) = 1.083\,\text{Å} \tag{3.12}$$

$$k_{stretch}(\text{C–H}) = 1934\,\text{kJ mol}^{-1}\,\text{Å}^{-2} \tag{3.13}$$

Parameterizing the angle bending term. From Eq. (3.3), a plot of E_{bend} against $(a - a_{eq})^2$ should be linear with a slope of k_{bend}. From STO-3G calculations on bending the H–C–C angle in ethane we get (cf. Table 3.1 and Fig. 3.7)

$$a_{eq}(\text{HCC}) = 110.7^\circ \tag{3.14}$$

$$k_{\text{bend}}(\text{HCC}) = 0.093\,\text{kJ mol}^{-1}\,\text{degree}^{-2} \tag{3.15}$$

Calculations on staggered butane gave for the C–C–C angle

$$a_{eq}(\text{CCC}) = 112.5^\circ \tag{3.16}$$

$$k_{\text{bend}}(\text{CCC}) = 0.110\,\text{kJ mol}^{-1}\,\text{degree}^{-2} \tag{3.17}$$

Parameterizing the torsional term. For the ethane case (Fig. 3.4), the equation for energy as a function of dihedral angle can be deduced fairly simply by adjusting the basic equation $E = \cos\theta$ to give $E = 1/2E_{\text{max}}[1+\cos 3(\theta+60)]$. For butane (Fig. 3.5), using Eq. (3.4) and experimenting with a curve-fitting program shows that a reasonably accurate torsional potential energy function can be created with five parameters, k_0 and $k_1 - k_4$:

$$E_{\text{torsion}}(\text{CH}_3\text{CH}_2\text{–CH}_2\text{CH}_3) = k_0 + \sum_{r=1}^{4} k_r[1 + \cos(r\theta)] \tag{3.18}$$

The values of the parameters $k_0 - k_5$ are given in Table 3.2. The calculated curve can be made to match the experimental one as closely as desired by using more terms (Fourier analysis).

Parameterizing the nonbonded interactions term. To parameterize Eq. (3.5) we might perform *ab initio* calculations in which the separation of two atoms or groups in different

Table 3.2. The experimental potential energy values for rotation about the central C–C bond of $\text{CH}_3\text{CH}_2\text{–CH}_2\text{CH}_3$ can be approximated by

$$E_{\text{torsion}}(\text{CH}_3\text{CH}_2\text{–CH}_2\text{CH}_3) = k_0 + \sum_{r=1}^{4} k_r[1 + \cos(r\theta)]$$

with $k_0 = 20.1$, $k_1 = -4.7$, $k_2 = 1.91$, $k_3 = -7.75$, $k_4 = 0.58$. Experimental energy values at 30°, 90°, and 150° were interpolated from those at 0°, 60°, 120°, and 180°; energies are in kJ mol^{-1}

θ (degrees)	E (calculated)	E (experimental)
0	0.15	0
30	6.7	7.0
60	14	14
90	8.8	9.0
120	3.5	3.3
150	15	15
180	25	25

molecules (to avoid the complication of concomitant changes in bond lengths and angles) is varied, and fit Eq. (3.5) to the energy vs. distance results. For nonpolar groups this would require quite high-level calculations (chapter 5), as van der Waals or dispersion forces are involved. We shall approximate the nonbonded interactions of methyl groups by the interactions of methane molecules, using experimental values of k_{nb} and σ, derived from studies of the viscosity or the compressibility of methane. The two methods give slightly different values [7b], but we can use the values

$$k_{nb} = 4.7\,\mathrm{kJ\,mol^{-1}} \tag{3.19}$$

and

$$\sigma = 3.85\,\text{Å} \tag{3.20}$$

Summary of the parameterization of the forcefield terms. The four terms of Eq. (3.1) were parameterized to give:

$$E_{\text{stretch}}(\text{C–C}) = 1735(l - 1.538)^2 \tag{3.21}$$

$$E_{\text{stretch}}(\text{C–H}) = 1934(l - 1.083)^2 \tag{3.22}$$

$$E_{\text{bend}}(\text{HCH}) = 0.093(a - 110.7)^2 \tag{3.23}$$

$$E_{\text{bend}}(\text{CCC}) = 0.110(a - 112.5)^2 \tag{3.24}$$

$$E_{\text{torsion}}(\text{CH}_3\text{CCCH}_3) = k_0 + \sum_{r=1}^{4} k_r[1 + \cos(r\theta)] \tag{3.25}$$

The parameters k of Eq. (3.25) are given in Table 3.2.

$$E_{\text{nonbond}}(\text{CH}_3/\text{CH}_3) = 4.7\left[\left(\frac{3.85}{r}\right)^{12} - \left(\frac{3.85}{r}\right)^{6}\right] \tag{3.26}$$

Note that this parameterization is only illustrative of the principles involved; any really viable forcefield would actually be much more sophisticated. The kind we have developed here might at the very best give crude estimates of the energies of alkanes. An accurate, practical forcefield would be parameterized as a best fit to many experimental and/or calculational results, and would have different parameters for different kinds of bonds, e.g. C–C for acyclic alkanes, for cyclobutane and for cyclopropane. A forcefield able to handle not only hydrocarbons would obviously need parameters involving elements other than hydrogen and carbon. Practical forcefields also have different parameters for various *atom types,* like sp^3 carbon vs. sp^2 carbon, or amine nitrogen vs. amide nitrogen. In other words, a different value would be used for, say, stretching involving an sp^3/sp^3 C–C bond than for an sp^2/sp^2 C–C bond. This is clearly necessary since the force constant of a bond depends on the hybridization of the atoms involved; the IR stretch frequency for the $sp^3\text{C}/sp^3\text{C}$ bond comes at roughly $1200\,\text{cm}^{-1}$, while that for the $sp^2\text{C}/sp^2\text{C}$ bond is about $1650\,\text{cm}^{-1}$ [8]. Since the vibrational frequency of a bond is proportional to the square root of the force constant, the force constants are in the ratio of about $(1650/1200)^2 = 1.9$; for corresponding atoms, force constants are in fact generally roughly proportional to bond order (double bonds and triple bonds are about two and three times as stiff, respectively, as the corresponding single bonds). Some forcefields account for the variation of bond order with conformation (twisting *p* orbitals out of alignment reduces their overlap) by performing a simple PPP molecular orbital calculation (chapter 6) to obtain the bond order.

A sophisticated forcefield might also consider H/H nonbonded interactions explicitly, rather than simply subsuming them into methyl/methyl interactions (combining atoms into groups is the feature of a *united atom* forcefield). Furthermore, nonbonding interactions between *polar* groups need to be accounted for in a field not limited to hydrocarbons. These are usually handled by the well-known potential energy/electrostatic charge relationship

$$E = k(q_1 q_2/r)$$

which has also been used to model hydrogen bonding [11].

A subtler problem with the naive forcefield developed here is that stretching, bending, torsional, and nonbonded terms are not completely independent. For example, the butane torsional potential energy curve (Fig. 3.5) does not apply precisely to all CH_3–C–C–CH_3 systems, because the barrier heights will vary with the length of the central C–C bond, obviously decreasing (other things being equal) as the bond is lengthened, since there will be a decrease in the interactions (whatever causes them) between the CH_3's and H's on one of the carbons of the central C–C and those on the other carbon. This could be accounted for by making the *k*'s of Eq. (3.25) a function of the X–Y length. Actually, partitioning the energy of a molecule into stretching, bending, etc. terms is somewhat formal; e.g. the torsional barrier in butane can be considered to be partly due to nonbonded interactions between the methyl groups. It should be realized that there is no one, right functional form for an MM forcefield (see, e.g. [1b]); accuracy, versatility, and speed of computation are the deciding factors in devising a forcefield.

3.2.3 A calculation using our forcefield

Let us apply the naive forcefield developed here to comparing the energies of two 2,2,3,3-tetramethylbutane ($(CH_3)_3CC(CH_3)_3$, i.e. *t*-Bu-Bu-*t*) geometries. We compare the energy of structure **1** (Fig. 3.8) with all the bond lengths and angles at our "natural" or standard values (i.e. at the STO-3G values we took as the equilibrium bond lengths and angles in section 3.2.2) with that of structure **2**, where the central C–C bond has been stretched from 1.538 to 1.600 Å, but all other bond lengths, as well as the bond angles and dihedral angles, are unchanged. Figure 3.8 shows the nonbonded distances we need, which would be calculated by the program from bond lengths, angles and dihedrals. Using Eq. (3.1):

$$\left(E = \sum_{\text{bonds}} E_{\text{stretch}} + \sum_{\text{angles}} E_{\text{bend}} + \sum_{\text{dihedrals}} E_{\text{torsion}} + \sum_{\text{pairs}} E_{\text{nonbond}}\right)$$

Figure 3.8. Structures for a simple MM "by hand" calculation on the effect of changing the central C–C length of $(CH_3)_3C–C(CH_3)_3$ from 1.538 to 1.600 Å.

For structure **1**

$$\sum_{\text{bonds}} E_{\text{stretch}}(\text{C–C}) = 7 \times 1735(1.538 - 1.538)^2 = 0$$

Bond stretch contribution cf. structure with $l_{\text{eq}} = 1.538$.

$$\sum E_{\text{stretch}}(\text{C–H}) = 18 \times 1934(1.083 - 1.083)^2 = 0$$

Bond stretch contribution cf. structure with $l_{\text{eq}} = 1.083$.

$$\sum_{\text{angles}} E_{\text{bend}}(\text{HCH}) = 18 \times 0.093(110.7 - 110.7)^2 = 0$$

Bond bend contribution cf. structure with $a_{\text{eq}} = 110.7°$.

$$\sum_{\text{angles}} E_{\text{bend}}(\text{CCC}) = 12 \times 0.110(112.5 - 112.5)^2 = 0$$

Bond bend contribution cf. structure with $a_{\text{eq}} = 112.5°$

$$\sum_{\text{dihedrals}} E_{\text{torsion}}(\text{CH}_3\text{CCCH}_3) = 6 \times 3.5 = 21.0$$

Torsional contribution cf. structure with no gauche–butane interactions

$$\sum_{\text{nonbond}} E_{\text{nonbond}}(\text{anti–CH}_3/\text{CH}_3) + \sum_{\text{nonbond}} E_{\text{nonbond}}(\text{gauche–CH}_3/\text{CH}_3)$$

$$= 3 \times 4.7\left[\left(\frac{3.85}{3.931}\right)^{12} - \left(\frac{3.85}{3.931}\right)^{6}\right] + 6 \times 4.7\left[\left(\frac{3.85}{3.065}\right)^{12} - \left(\frac{3.85}{3.065}\right)^{6}\right]$$

$$= 3 \times (-0.487) + 6 \times (54.05) = -1.463 + 324.3$$

$$= 323\,\text{kJ mol}^{-1}$$

nonbonding contribution cf. structure with noninteracting CH_3s.

Actually, nonbonding interactions are already included in the torsional term (as *gauche*–butane interactions); we might have used an ethane-type torsional function and accounted for CH_3/CH_3 interactions entirely with nonbonded terms. However, in comparing calculated *relative* energies the torsional term will cancel out.

$$E_{\text{total}} = E_{\text{stretch}} + E_{\text{bend}} + E_{\text{torsion}} = 0 + 0 + 21.0 + 323\,\text{kJ mol}^{-1} = 344\,\text{kJ mol}^{-1}$$

For structure **2**

$$\sum_{\text{bonds}} E_{\text{stretch}}(\text{C–C}) = 6 \times 1735(1.538 - 1.538)^2 + 1 \times 1735(1.600 - 1.538)^2$$

$$= 0 + 6.67 = 6.67\,\text{kJ mol}^{-1}$$

Bond stretch contribution cf. structure with $l_{\text{eq}} = 1.538$.

$$\sum_{\text{bonds}} E_{\text{stretch}}(\text{C–H}) = 18 \times 1934(1.083 - 1.083)^2 = 0$$

Bond stretch contribution cf. structure with $l_{\text{eq}} = 1.083$.

$$\sum_{\text{angles}} E_{\text{bend}}(\text{HCH}) = 18 \times 0.093(110.7 - 110.7)^2 = 0$$

Bond bend contribution cf. structure with $a_{\text{eq}} = 110.7°$

$$\sum_{\text{angles}} E_{\text{bend}}(\text{CCC}) = 12 \times 0.110(112.5 - 112.5)^2 = 0$$

Bond bend contribution cf. structure with $a_{\text{eq}} = 112.5°$

$$\sum_{\text{dihedrals}} E_{\text{torsion}}(\text{CH}_3\text{CCCH}_3) = 6 \times 3.5 = 21.0$$

Torsional contribution cf. structure with no gauche–butane interactions.

The stretching and bending terms for structure **2** are the same as for structure **1**, except for the contribution of the central C–C bond; strictly speaking, the torsional term should be smaller, since the opposing $\text{C(CH}_3)$ groups have been moved apart.

$$\sum_{\text{nonbond}} E_{\text{nonbond}}(\text{anti–CH}_3/\text{CH}_3) + \sum_{\text{nonbond}} E_{\text{nonbond}}(\text{gauche–CH}_3/\text{CH}_3)$$

$$= 3 \times 4.7\left[\left(\frac{3.85}{3.974}\right)^{12} - \left(\frac{3.85}{3.974}\right)^{6}\right] + 6 \times 4.7\left[\left(\frac{3.85}{3.120}\right)^{12} - \left(\frac{3.85}{3.120}\right)^{6}\right]$$

$$= 3 \times (-0.673) + 6 \times (41.97) = -2.019 + 251.8$$

$$= 250\,\text{kJ mol}^{-1}$$

nonbonding contribution cf. structure with noninteracting CH_3s.

$$E_{\text{total}} = E_{\text{stretch}} + E_{\text{bend}} + E_{\text{torsion}} = 6.67 + 0 + 21.0 + 250 = 277\,\text{kJ mol}^{-1}$$

So the relative energies are calculated to be

$$E(\text{structure } \mathbf{2}) - E(\text{structure } \mathbf{1}) = 277 - 344 = -67\,\text{kJ mol}^{-1}$$

This crude method predicts that stretching the central C/C bond of 2,2,3,3-tetramethylbutane from the approximately normal sp^3-C-sp^3-C length of 1.583 Å (structure **1**) to the quite “unnatural” length of 1.600 Å (structure **2**) will lower the potential energy by $67\ \mathrm{kJ\ mol^{-1}}$, and indicates that the drop in energy is due very largely to the relief of nonbonded interactions. A calculation using the accurate forcefield MM3 [12] gave an energy difference of $54\ \mathrm{kJ\ mol^{-1}}$ between a “standard” geometry approximately like structure **1**, and a *fully optimized* geometry, which had a central C/C bond length of 1.576 Å. The surprisingly good agreement is largely the result of a fortuitous cancellation of errors, but this does not gainsay the fact that we have used our forcefield to calculate something of chemical interest, namely the relative energy of two molecular geometries. In principle, we could have found the minimum-energy geometry according to this forcefield, i.e. we could have optimized the geometry (chapter 2). Geometry optimization is in fact the main use of MM, and modern programs employ analytical first and second derivatives of the energy with respect to the geometric coordinates for this (chapter 2).

3.3 EXAMPLES OF THE USE OF MM

If we consider the applications of MM from the viewpoint of the goals of those who use it, then the *main* applications have been:

(1) to calculate the geometries (and perhaps energies) of small to medium-sized (i.e. nonpolymeric) molecules, very often in order to a reasonable starting geometry for another type (e.g. *ab initio*) of calculation;

(2) to calculate the geometries and energies of polymers (mainly proteins and nucleic acids);

(3) to calculate the geometries and energies of transition states (infrequent);

(4) as an aid to organic synthesis;

(5) to generate the potential energy function under which molecules move, for molecular dynamics calculations.

These applications are not all independent. For example, a chemist planning a synthesis might use MM to obtain a plausible geometry for an intermediate involved in the synthesis (the use of MM in synthesis is now so common it is likely that this is often not reported in the literature), and a protein or nucleic acid could be studied with molecular dynamics. Examples of these five facets of the use of MM will be given.

3.3.1 Geometries and energies of small- to medium-sized molecules

Molecular mechanics is used mainly to calculate geometries and energies for small- (roughly C_1 to about C_{10}) and medium-sized (roughly C_{11} to C_{100}) organic molecules. It is by no means limited to organic molecules, as forcefields like SYBYL and UFF [5] have been parameterized for most of the periodic table, but the great majority of MM calculations have been done on organics, probably largely because MM was the creation of organic chemists (this is probably because the concept of geometric structure has long been central in organic chemistry). The most frequent use of MM is undoubtedly to obtain reasonable starting structures for *ab initio*, SE, or DFT (chapters 5–7) calculations. Nowadays this is usually done by building the molecule with an interactive builder in a graphical user interface, then optimizing it with MM with the click of a mouse. The resulting structure is then subjected to an *ab initio,* etc. calculation. MM calculations are usually done only for equilibrium structures (i.e. relative minima on the PES), but by constraining geometric parameters one can approximate roughly transition states (below).

The two salient features of MM calculations on small to medium-sized molecules is that they are *fast* and they *can* be *very accurate.* Times required for a geometry optimization of unbranched $C_{20}H_{42}$, of C_{2h} symmetry, with the Merck Molecular Force Field (MMFF), the SE AM1 (chapter 6) and the *ab initio* HF/3-21G (chapter 5) methods, as implemented with the program SPARTAN [13], were 1.2, 16 s, and 57 min, respectively (on an obsolescent machine a few years ago; these times would now by shorter by a factor of at least 2). Clearly as far as speed goes there is no contest between the methods, and the edge in favor of MM increases with the size of the molecule. In fact, MM was till recently the only practical method for calculations on molecules with more than about 100 heavy atoms (in computational chemistry a heavy atom is any atom heavier than helium). Even programs not designed specifically for macromolecules will handle molecules with 1000 or more atoms on machines of modest power (e.g. a good PC).

Molecular mechanics energies can be very accurate *for families of compounds for which the forcefield has been parameterized.* Appropriate parameterization permits calculation of ΔH_f° (heat of formation, enthalpy of formation) in addition to strain energy [1f]. For the MM2 program (see below), for standard hydrocarbons ΔH_f° errors are usually only $0–4\,\mathrm{kJ\,mol^{-1}}$, which is comparable to experimental error, and for oxygen containing organics the errors are only $0–8\,\mathrm{kJ\,mol^{-1}}$ [14]; the errors in MM conformational energies are often only about $2\,\mathrm{kJ\,mol^{-1}}$ [15]. MM geometries are usually reasonably good for small to medium-sized molecules [4,9a,16]; for the MM3 program (see below) the RMS error in bond lengths for cholesteryl acetate was only about 0.007 Å [4]. "Bond length" is, if unqualified, somewhat imprecise, since different methods of measurement give somewhat different values [4,9a] (section 5.5.1). MM geometries are routinely used as input structures for quantum-mechanical calculations, but in fact the MM geometry and energy are in some cases as good or better than those from a "higher-level calculation" [17]. The benchmark MM programs for small to medium-sized molecules are probably MM2 and MM3, which will presumably be gradually supplanted by MM4 [4]; the MMFF [18] is likely to become very popular too, not least because of its implementation in SPARTAN [13].

3.3.2 Geometries and energies of polymers

Next to generating geometries and energies of small to medium-sized molecules, the main use of MM is to model polymers, mainly biopolymers (proteins, nucleic acids, polysaccharides). Forcefields have been developed specifically for this; two of the most widely-used of these are CHARMM (Chemistry at Harvard using Molecular Mechanics) [19] (the academic version; the commercial version is CHARMm) and the forcefields in the computational package AMBER (Assisted Model Building with Energy Refinement) [20]. CHARMM was designed to deal with biopolymers, mainly proteins, but has been extended to handle a range of small molecules. AMBER is perhaps the most widely used set of programs for biological polymers, being able to model proteins, nucleic acids, and carbohydrates. Programs like AMBER and CHARMM that model large molecules have been augmented with quantum mechanical methods (SE [21] and even *ab initio* [22]) to investigate small regions where treatment of electronic processes like transition state formation may be critical.

An extremely important aspect of the modelling (which is done largely with MM) of biomolecules is designing pharmacologically active molecules that can fit into active sites (the pharmacophores) of biomolecules and serve as useful drugs. For example, a molecule might be designed to bind to the active site of an enzyme and block the undesired reaction of the enzyme with some other molecule. Pharmaceutical chemists computationally craft a molecule that is sterically and electrostatically complementary to the active site, and try to *dock* the potential drug into the active site. The binding energy of various candidates can be compared and the most promising ones can then be synthesized, as the second step on the long road to a possible new drug. The computationally assisted design of new drugs and the study of the relationship of structure to activity (quantitative structure–activity relationships, QSAR) is one of the most active areas of computational chemistry [23].

3.3.3 Geometries and energies of transition states

By far the main use of MM is to find reasonable geometries for the ground states of molecules, but it has also been used to investigate transition states. The calculation of transition states involved in conformational changes is a fairly straightforward application of MM, since "reactions" like the interconversion of butane or cyclohexane conformers do not in involve the deep electronic reorganization that we call bond-making or bond-breaking. The changes in torsional and nonbonded interactions that accompany them are the kinds of processes that MM was designed to model, and so good transition state geometries and energies can be expected for this particular kind of process; transition state *geometries* cannot be (readily) measured, but the MM energies for conformational changes agree well with experiment: indeed, one of the two very first applications of MM [3a,d] was to the rotational barrier in biphenyls (the other was to the S_N2 reaction [3c]). Since MM programs are usually not able to optimize an input geometry toward a saddle point (see below), one normally optimizes to a minimum subject to the symmetry constraint expected for the transition state. Thus for ethane, optimization to a minimum within D_{3h} symmetry (i.e. by constraining the HCCH dihedral to be $0°$, or by starting with a structure of exactly D_{3h} symmetry) will give the transition state, while optimization with D_{3d} symmetry gives the ground-state conformer (Fig. 3.9). Optimizing an input C_{2v} cyclohexane structure (Fig. 3.10) gives the stationary point nearest this input structure, which is the transition state for interconversion of enantiomeric twist cyclohexane conformers.

There are several examples of the application of MM to actual chemical reactions, as distinct from conformational changes; the ones mentioned here are taken from the review by Eksterowicz and Houk [24]. The simplest way to apply MM to transition states is to approximate the transition state by a ground-state molecule. This can sometimes give surprisingly good results. The rates of solvolysis of compounds RX to the

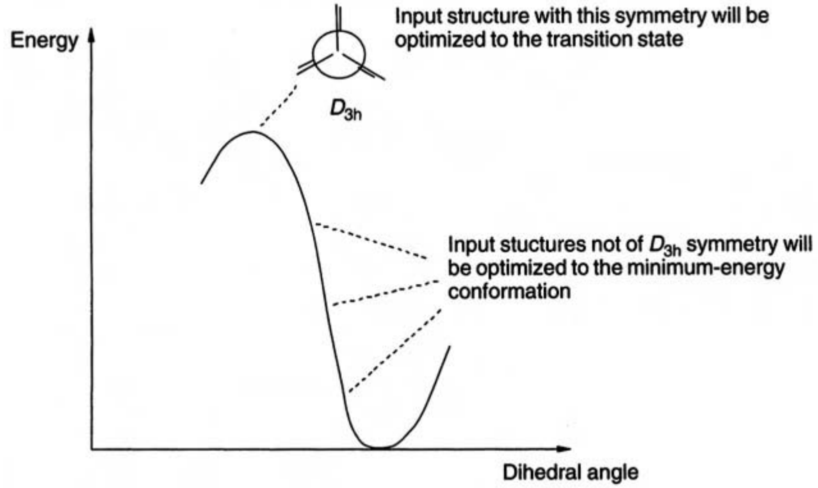

Figure 3.9. Optimizing ethane within D_{3h} symmetry (i.e. by constraining the HCCH dihedral to be 0°, or by inputting a structure with exact D_{3h} symmetry) will give the transition state, while optimization without requiring D_{3h} symmetry gives the ground-state conformer.

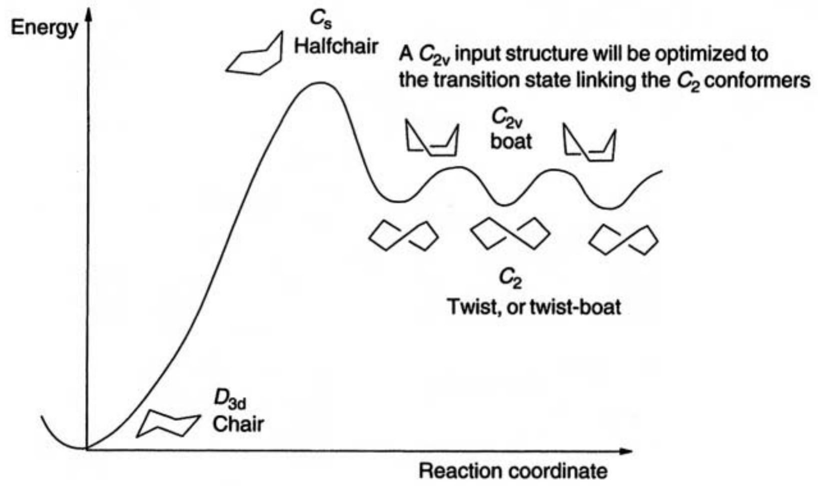

Figure 3.10. Optimizing cyclohexane within C_{2v} symmetry gives a transition state, not one of the minima.

cation correlated well with the energy difference between the hydrocarbon RH, which approximates RX, and the cation R^+, which approximates the transition state leading to this cation. This is not entirely unexpected, as the Hammond postulate [25] suggests that the transition state should resemble the cation. In a similar vein, the activation energy for solvolysis has been approximated as the energy difference between a "methylalkane", with CH_3 corresponding to X in RX, and a ketone, the sp^2 carbon of which corresponds to the incipient cationic carbon of the transition state.

One may wish a more precise approximation to the transition state geometry than is represented by an intermediate or a compound somewhat resembling the transition state. This can sometimes be achieved by optimizing to a minimum subject to the constraint that the bonds being made and broken have lengths believed (e.g. from quantum mechanical calculations on simple systems, or from chemical intuition) to approximate those in the transition state, and with appropriate angles and dihedrals also constrained. With luck this will take the input structure to a point on the potential energy surface near the saddle point. For example, an approximation to the geometry of the transition state for formation of cyclohexene in the Diels–Alder reaction of butadiene with ethene can be achieved (Fig. 3.11) by essentially building a boat conformation of cyclohexene, constraining the two forming C/C bonds to about 2.1 Å, and optimizing, using the CH_2 bridge (later removed) to avoid twisting and to maintain C_s symmetry; optimization with a dihedral constraint removes steric conflict between two hydrogens and gives a reasonable starting structure for, say, an *ab initio* optimization.

The most sophisticated approach to locating a transition state with MM is to use an algorithm that optimizes the input structure to a true saddle point, that is to a geometry characterized by a Hessian with one and only one negative eigenvalue (chapter 2). To do this the MM program must be able not only to calculate second derivatives, but must also be parameterized for the partial bonds in transition states, which is a feature lacking in standard MM forcefields.

MM has been used to study the transition states involved in S_N2 reactions, hydroborations, cycloadditions (mainly the Diels-Alder reaction), the Cope and Claisen rearrangements, hydrogen transfer, esterification, nucleophilic addition to carbonyl groups and electrophilic C/C bonds, radical addition to alkenes, aldol condensations, and various intramolecular reactions [24].

3.3.4 MM in organic synthesis

In the past 15 years or so MM has become widely used by synthetic chemists, thanks to the availability of inexpensive computers (personal computers will easily run MM programs) and user-friendly and relatively inexpensive programs [5]. Since MM can calculate the energies and geometries of ground state molecules and (within the limitations alluded to above) transition states, it can clearly be of great help in planning syntheses. To see which of two or more putative reaction paths should be favored, one might (1) use MM like a hand-held model: examine the substrate molecule for factors like steric hindrance or proximity of reacting groups, or (2) approximate the transition states for alternative reactions using an intermediate or some other plausible proxy (cf. the treatment of solvolysis in the discussion of transition states above), or (3) attempt to calculate the energies of competing transition states (cf. the above discussion of transition state calculations).

The examples given here of the use of MM in synthesis are taken from the review by Lipkowitz and Peterson [26]. In attempts to simulate the metal-binding ability of biological acyclic polyethers, the tricyclic **1** (Fig. 3.12) and a tetracyclic analogue were synthesized, using as a guide the indication from MM that these molecules resemble the cyclic polyether 18-crown-6, which binds the potassium ion; the acyclic compounds were found to be indeed comparable to the crown ether in metal-binding ability.

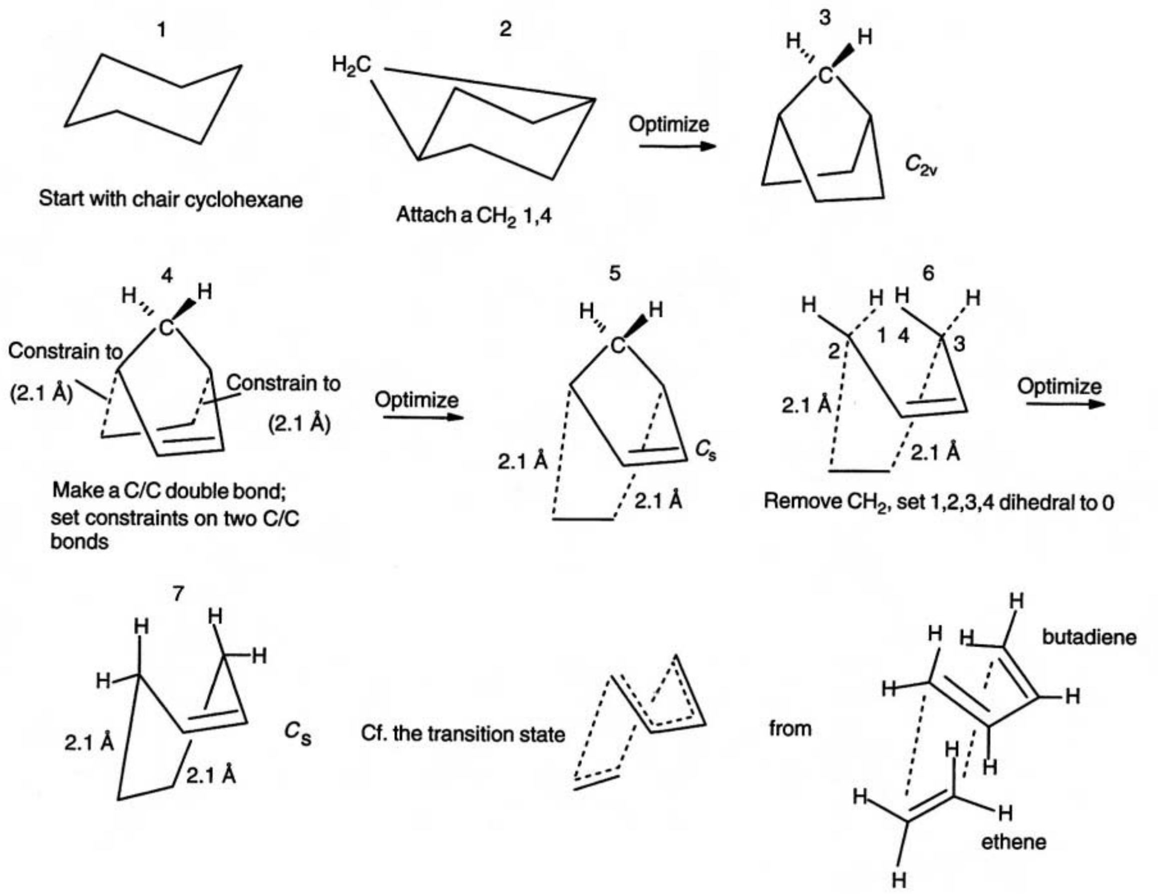

Figure 3.11. Using MM to get the (approximate) transition state for the Diels–Alder reaction of butadiene with ethene.

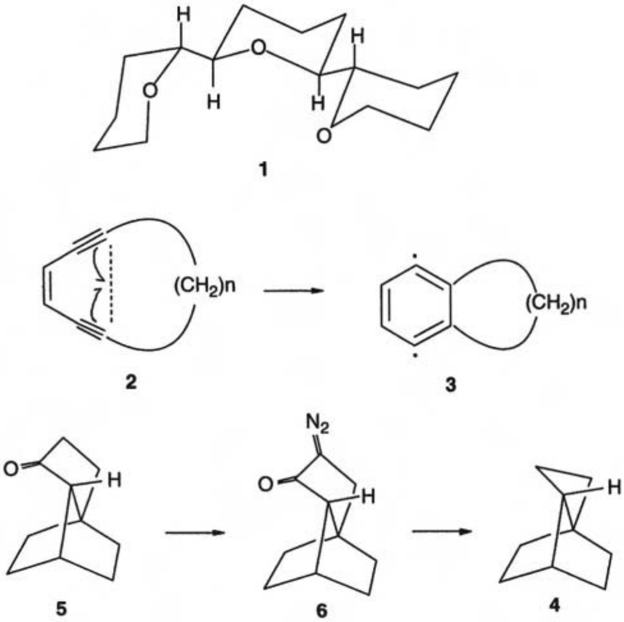

Figure 3.12. Some molecules (1, 2, 4) which have been synthesized with the aid of MM.

Enediynes like **2** (Fig. 3.12) are able to undergo cyclization to a phenyl-type diradical **3**, which *in vivo* can attack DNA; in molecules with an appropriate triggering mechanism this forms the basis of promising anticancer activity. The effect of the length of the constraining chain (i.e. of n in **2**) on the activation energy was studied by MM, aiding the design of compounds (potential drugs) that were found to be more active against tumors than are naturally-occurring enediyne antibiotics.

To synthesize the very strained tricyclic system of **4** (Fig. 3.12), a photochemical Wolff rearrangement was chosen when MM predicted that the skeleton of **4** should be about $109\,\mathrm{kJ\,mol^{-1}}$ less stable than that of the available **5**. Photolysis of the diazoketone **6** gave a high-energy carbene which lay above the carbon skeleton of **4** and so was able to undergo Wolff rearrangement ring contraction to the ketene precursor of **4**.

A remarkable (and apparently still unconfirmed) prediction of MM is the claim that the perhydrofullerene $C_{60}H_{60}$ should be stabler with some hydrogens *inside* the cage [27].

3.3.5 Molecular dynamics and Monte Carlo simulations

Programs like those in AMBER are used not only for calculating geometries and energies, but also for simulating molecular motion, i.e. for molecular dynamics [28], and for calculating the relative populations of various conformations or other geometric arrangements (e.g. solvent molecule distribution around a macromolecule) in Monte Carlo simulations [29]. In molecular dynamics Newton's laws of motion are applied to molecules moving in a MM forcefield, although relatively small parts of the system (system: with biological molecules in particular modelling is often done not on an isolated molecule but on a molecule and its environment of solvent and ions) may be simulated with quantum mechanical methods [21,22]. In Monte Carlo methods random numbers decide how atoms or molecules are moved to generate new conformations or geometric arrangements (states) which are then accepted or rejected according to some filter. Tens of thousands (or more) of states are generated, and the energy of each is calculated by MM, generating a Boltzmann distribution.

3.4 GEOMETRIES CALCULATED BY MM

Figure 3.13 compares geometries calculated with the MMFF with those from a reasonably high-level *ab initio* calculation (MP2(FC)/6-31G*; chapter 5) and from experiment. The MMFF is a popular force field, applicable to a wide variety of molecules. Popular prejudice holds that the *ab initio* method is "higher" than MM and so should give superior geometries. The set of 20 molecules in Fig. 3.13 is also used in chapters 5, 6, and 7, to illustrate the accuracy of *ab initio,* SE, and density functional calculations in obtaining molecular geometries. The data in Fig. 3.13 are analyzed in Table 3.3. Table 3.4 compares dihedral angles for eight molecules, which are also used in chapters 5–7.

This survey suggests that for common organic molecules the MMFF is nearly as good as the *ab initio* MP2(FC)/6-31G* method for calculating geometries. Both methods give good geometries, but while these MM calculations all take effectively about one second, MP2 geometry optimizations on these molecules require typically a few minutes (CH_3COCH_3, 2.4 min; CH_3Cl, 1.4 min; $(CH_3)_2SO$, 3.7 min; 1.5GHz Pentium 4). For larger molecules where MP2 would need hours, MM calculations might still take only seconds. Note, however, that *ab initio* methods provide information that MM cannot, and are far more reliable for molecules outside those of the kind used in the MM training set (section 3.2.2). The worst MMFF bond length deviation from experiment among the 20 molecules is 0.021 Å (the C=C bond of propene; the MP2 deviation is 0.020 Å); most of the other errors are ca. 0.01 Å or less. The worst bond angle error is 13.6°, for HOF, and for HOCl the deviation is 7.9°, the second worst angle error in the set. This suggests a problem for the MMFF with X–O–Halogen angles, but while for CH_3OF deviation from the MP2 angle (which is likely to be close to experiment) is $\text{MMFF}-\text{MP2} = 110.7° - 102.8° = 7.9°$, for CH_3Cl the deviation is only $112.0° - 109.0° = 3.0°$.

MMFF dihedral angles are remarkably good, considering that torsional barriers are believed to arise from subtle quantum mechanical effects. The worst dihedral angle error is 10°, for HOOH, and the second worst, –5.0°, is for the analogous HSSH. The popular *ab initio* HF/3-21G (chapter 5) and SE PM3 (chapter 6) methods also have trouble with HOOH, predicting a dihedral angle of 180°. For those dihedrals not involving OO or SS bonds, (an admittedly small selection), the MMFF errors are only ca. 1°–2°, cf. ca. 2°–6° for MP2.

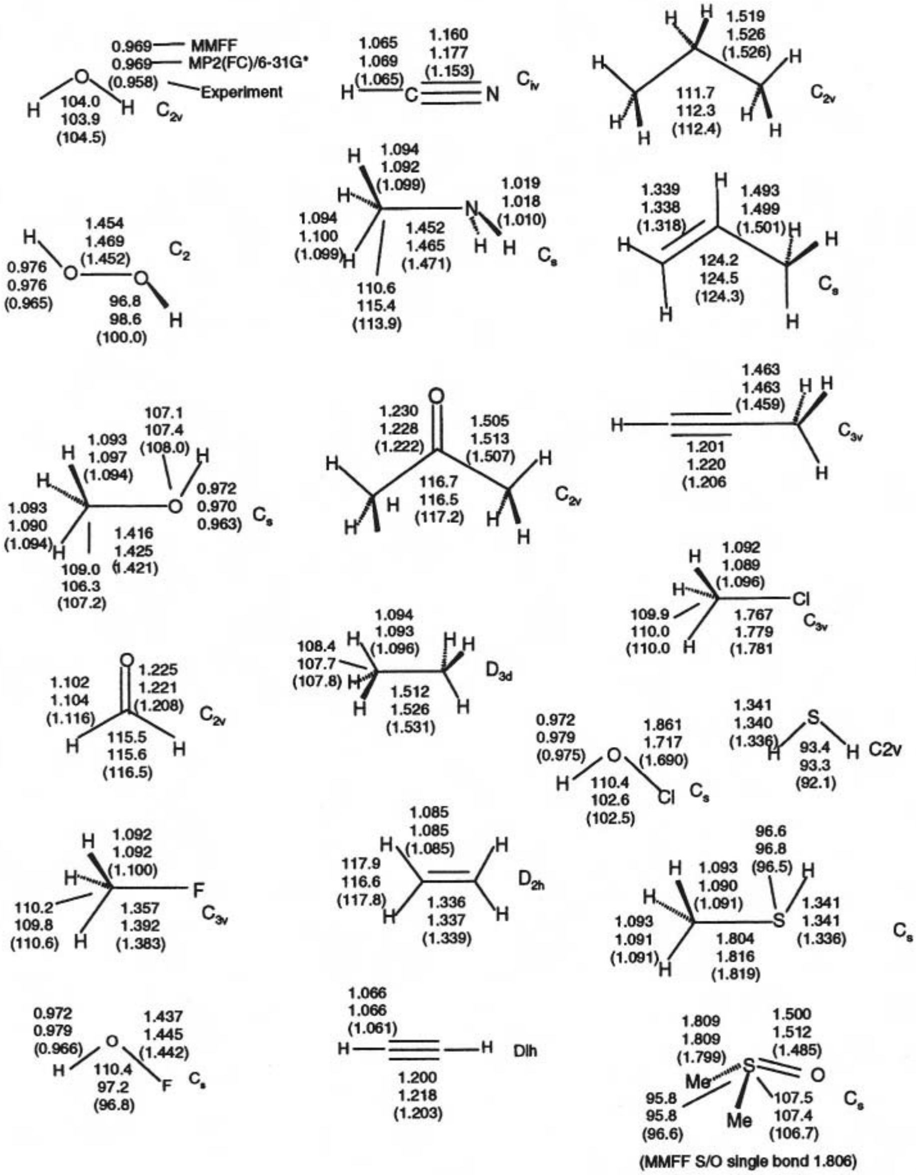

Figure 3.13. A comparison of some MMFF, MP2(FC)/6-31G* and experimental geometries. Calculations are by the author and experimental geometries are from Ref. [30]. Note that all CH bonds are ca. 1 Å, all other bonds range from ca. 1.2–1.8 Å, and all bond angles (except for linear molecules) are ca. 90–120°.

Table 3.3. Error in MMFF MM and MP2(FC)/6-31G* bond lengths and angles, from Fig. 3.13

Bond length errors, $r - r_{exp}$ (Å)				Bond angle errors, $a - a_{exp}$
C–H	O–H, N–H, S–H	C–C	C–O, N, F, Cl, S	Angles
MeOH −0.001/−0.004 −0.001/0.003	H_2O 0.011/0.011	Me_2CO −0.002/0.006	MeOH −0.005/0.007	H_2O (HOH) −0.5/−0.6
HCHO −0.014/−0.012	H_2O_2 0.011/0.011	CH_3CH_3 −0.019/−0.005	HCHO 0.017/0.013	H_2O_2 (HOO) −3.2/−1.4
MeF −0.008/−0.008	MeOH 0.009/0.007	CH_2CH_2 −0.013/−0.017/−0.002	MeF −0.005/−0.007	MeOH (HCO)
HCN 0.000/0.004	HOF 0.006/0.013	HCCH −0.003/0.015	HCN 0.007/0.024	HCHO (HCH) −1.0/−0.9
$MeNH_2$ −0.005/0.001 −0.005/−0.007	$MeNH_2$ 0.009/0.008	$CH_3CH_2CH_3$ −0.007/0.000	$MeNH_2$ −0.019/−0.006	MeF (HCH) −0.4/−0.8
CH_3CH_3 −0.002/−0.003	HOCl −0.003/0.004	CH_2CHCH_3 −0.008/−0.002 0.021/0.020	Me_2CO 0.008/0.006	HOF (HOF) 13.6/0.4
CH_2CH_2 0.000/0.000	H_2S 0.005/0.004	$HCCCH_3$ 0.004/0.004 −0.005/0.014	MeCl −0.014/−0.002	$MeNH_2$ (HCN) −3.3/1.5

CHCH	MeSH		MeSH	Me_2CO (CCC)
0.005/0.005	0.005/0.005		−0.015/−0.003	−0.5/−0.8
MeCl			Me_2SO	CH_3CH_3 (HCH)
−0.004/−0.007			0.010/0.010	0.6/−0.1
MeSH				CH_2CH_2 (HCH)
0.002/0.000				0.1/−1.2
0.002/−0.001				
				$CH_3CH_2CH_3$ (CCC)
				−0.7/−0.1
				CH_2CHCH_3 (CCC)
				−0.1/0.2
				MeCl (HCH)
				−0.1/0.0
				H_2S (HSH)
				1.3/1.2
				MeSH (CSH)
				0.1/0.3
				Me_2SO (CSC)
				−0.8/−0.8
				(CSO)
				0.8/0.7
3+, 8−, two 0	7+, 1−, none 0	2+, 7−, none 0	4+, 5−, none 0	7+, 11−, none 0
4+, 7−, two 0	8+, 0−, none 0	5+, 3−, one 0	6+, 3−, none 0	6+, 11−, one 0
mean of 13:	mean of 8:	mean of 9:	mean of 9:	mean of 18:
0.004/0.004	0.007/0.008	0.009/0.008	0.011/0.009	1.7/0.7

Errors are given as MMFF/MP2. In some cases (e.g. MeOH) errors for two bonds are given, on one line and on the line below. A minus sign means that the calculated value is less than the experimental. The numbers of positive, negative, and zero deviations from experiment are summarized at the bottom of each column. The averages at the bottom of each column are arithmetic means of the absolute values of the errors.

Table 3.4. MMFF, MP2(FC)/6-31G* and experimental dihedral angles (degrees).

	Dihedral Angles			
Molecule	MMFF	MP2/6-31G*	Exp.	Errors
HOOH	129.4	121.3	119.1[a]	10/2.2
FOOF	90.7	85.8	87.5[b]	3.2/−1.7
FCH_2CH_2F (FCCF)	72.1	69.0	73[b]	−1.0/−4
FCH_2CH_2OH				
(FCCO)	65.9	60.1	64.0[c]	1.9/−3.9
(HOCC)	53.5	54.1	54.6[c]	−1.1/−0.5
$ClCH_2CH_2OH$				
(ClCCO)	65.7	65.0	63.2[b]	2.5/1.8
(HOCC)	56.8	64.3	58.4[b]	−1.6/5.9
$ClCH_2CH_2F$				
(ClCCF)	69.8	65.9	68[b]	1.8/−2.1
HSSH	84.2	90.4	90.6[a]	−6.4/−0.2
FSSF	82.9	88.9	87.9[b]	−5.0/1.0
				Deviations: 5+, 5−/4+, 6− mean of 10: 3.5/2.3;

Errors are given in the *Errors* column as MMFF/MP2/6-31G*. A minus sign means that the calculated value is less than the experimental. The numbers of positive and negative deviations from experiment and the average errors (arithmetic means of the absolute values of the errors) are summarized at the bottom of the *Errors* column. Calculations are by the author; references to experimental measurements are given for each measurement. The AM1 and PM3 dihedrals vary by a fraction of a degree depending on the input dihedral. Some molecules have calculated minima at other dihedrals in addition to those given here, e.g. FCH_2CH_2F at FCCF 180°.

[a] Hehre [30], pp. 151, 152.

[b] M. D. Harmony, V. W. Laurie, R. L. Kuczkowski, R. H. Schwenderman, D. A. Ramsay, F. J. Lovas, W. H. Lafferty, A. G. Makai, "Molecular Structures of Gas-Phase Polyatomic Molecules Determined by Spectroscopic Methods," J. Phys. Chem. Ref. Data, 1979, *8*, 619–721.

[c] J. Huang and K. Hedberg, J. Am. Chem. Soc., 1989, *111*, 6909.

3.5 FREQUENCIES CALCULATED BY MM

Any method that can calculate the energy of a molecular geometry can in principle calculate vibrational frequencies, since these can be obtained from the second derivatives of energy with respect to molecular geometry (section 2.5), and the masses of the vibrating atoms. Some commercially available MM programs, for example the MMFF as implemented in SPARTAN [13], can calculate frequencies. Frequencies are useful (section 2.5) (1) for characterizing a species as a minimum (no imaginary frequencies) or a transition state or higher-order saddle point (one or more imaginary frequencies), (2) for obtaining zero-point energies to correct frozen-nuclei energies (section 2.2), and (3) for interpreting or predicting infrared spectra.

(1) *Characterizing a species.* This is not often done with MM, because MM is used mostly to create input structures for other kinds of calculations, and to study known (often biological) molecules. Nevertheless MM can yield information on the curvature of the potential energy surface (as calculated by that particular forcefield, anyway) at the point in question. For example, the MMFF-optimized geometries of D_{3d} (staggered) and D_{3h} (eclipsed) ethane (Fig. 3.3) show, respectively, no imaginary frequencies and one imaginary frequency, the latter corresponding to rotation about the C/C bond. Thus the MMFF (correctly) predicts the staggered conformation to be a minimum, and the eclipsed to be a transition state connecting successive minima along the torsional reaction coordinate. Again, calculations on cyclohexane conformations with the MMFF correctly give the boat an imaginary frequency corresponding to a twisting motion leading to the twist conformation, which latter has no imaginary frequencies (Fig. 3.10). Although helpful for characterizing conformations, particularly hydrocarbon conformations, MM is less appropriate for species in which bonds are being formed and broken. For example, the symmetrical (D_{3h}) species in the $F^- + CH_3 - F$ S_N2 reaction, with equivalent C/F partial bonds, is incorrectly characterized by the MMFF as a minimum rather than a transition state, and the C/C bonds are calculated to be 1.289 Å long, cf. the value of ca. 1.8 Å from methods known to be trustworthy for transition states.

(2) *Obtaining zero-point energies (ZPEs).* ZPEs are essentially the sum of the energies of each normal-mode vibration. They are added to the raw energies (the frozen-nuclei energies, corresponding to the stationary points on a Born-Oppenheimer surface; section 2.3) in accurate calculations of relative energies using ab initio (chapter 5) or DFT (chapter 7) methods. However, the ZPEs used for those corrections are usually obtained from an *ab initio* or DFT calculation.

(3) *Infrared spectra.* The ability to calculate the energies (cm^{-1}) and relative intensities of molecular vibrations amounts to being able to calculate infrared spectra. MM as such cannot calculate the *intensities* of vibrational modes, since these involve changes in dipole moments (section 5.5.3), and dipole moment is related to electron distribution, a concept that lies outside MM. However, approximate intensities can be calculated by assigning dipole moments to bonds or charges to atoms, and such methods have been implemented in MM programs [31], although MM programs that calculate intensities are not yet widely used. Figures 3.14–3.17 compare the IR spectra of acetone, benzene, dichloromethane and methanol, calculated with the MM3 program[6] with the experimental spectra, and with spectra calculated by the *ab initio* MP2(FC)/6-31G* method; the data for Figs 3.14–3.17 are in Tables 3.5–3.8 (in chapters 5–7 spectra for these four molecules, calculated by *ab initio,* SE, and density functional methods, are also given). The MP2 spectra generally match experiment better than the MM3, although the latter method furnishes a rapid way of obtaining approximate IR spectra. For a series of related compounds, MM3 might be a reasonable way to quickly investigate trends in frequencies and intensities. Extensive surveys of MMFF and MM4 frequencies showed that MMFF root-mean-square errors are ca. $60\,cm^{-1}$, and MM4 errors $25–52\,cm^{-1}$ [5b].

---

[6] The MM3 frequencies and intensities were kindly provided by Dr. J. -H. Lii of the Department of Chemistry of the University of Georgia, Athens, Georgia, USA.

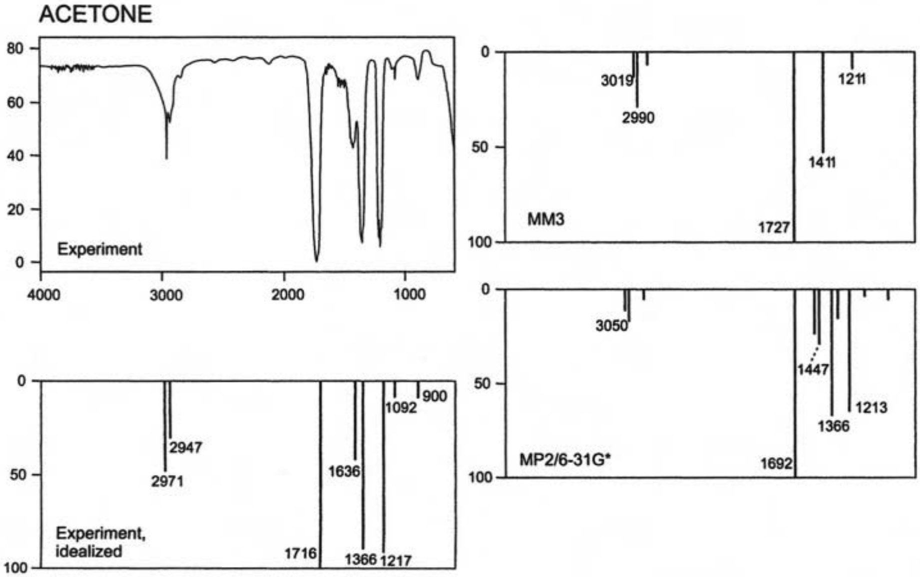

Figure 3.14. Experimental (gas phase) and MM (MM3) and MP2(FC)/6-31G* calculated infra red spectra of acetone. The MM3 spectrum is based on the data in Table 3.5; the MP2 spectrum is that shown in Fig. 5.33.

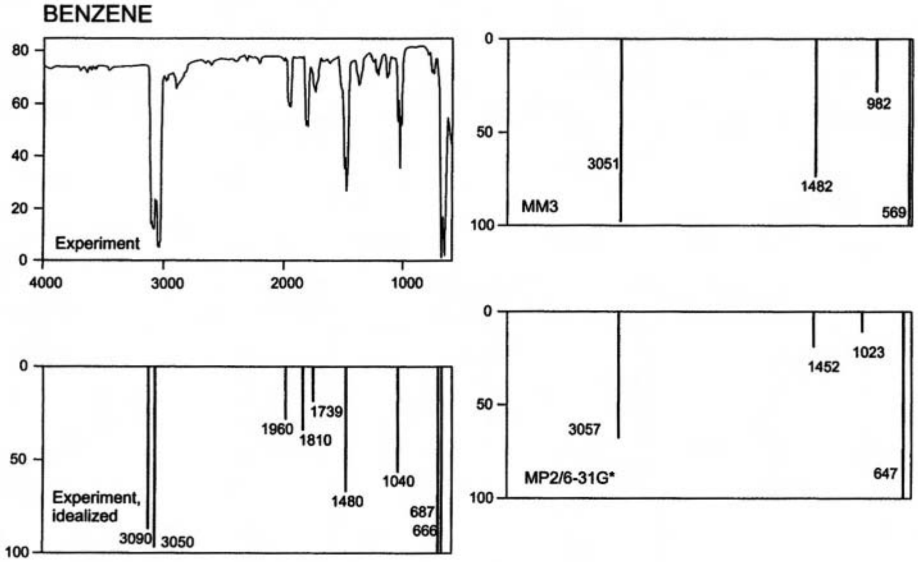

Figure 3.15. Experimental (gas phase) and MM (MM3) and MP2(FC)/6-31G* calculated infra red spectra of benzene. The MM3 spectrum is based on the data in Table 3.6; the MP2 spectrum is that shown in Fig. 5.34.

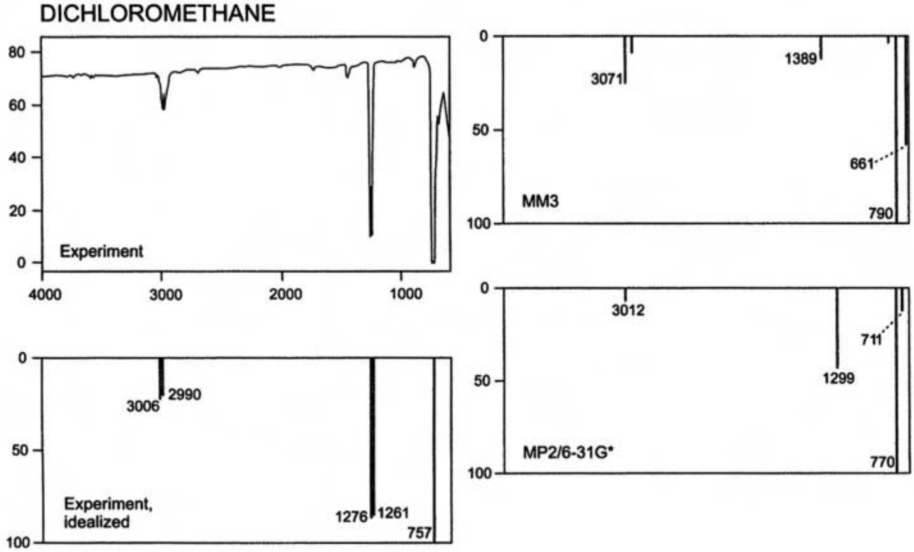

Figure 3.16. Experimental (gas phase) and MM (MM3) and MP2(FC)/6-31G* calculated infra red spectra of dichloromethane. The MM3 spectrum is based on the data in Table 3.7; the MP2 spectrum is that shown in Fig. 5.35.

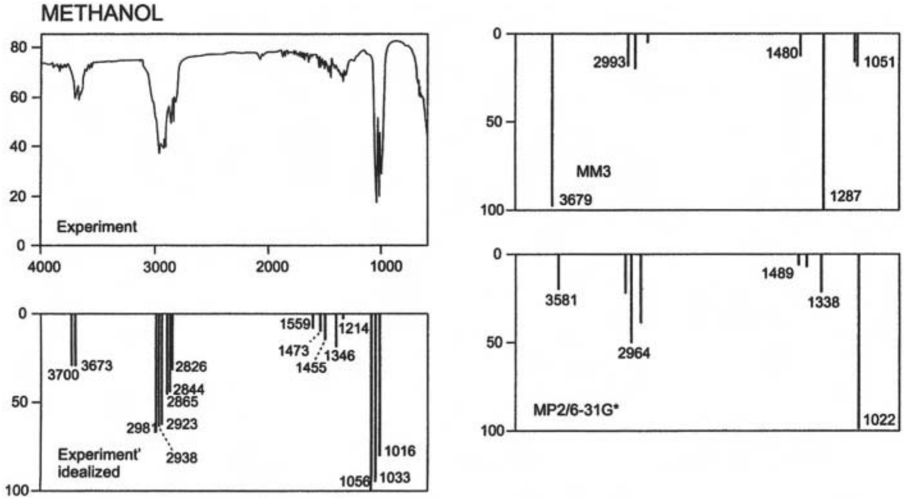

Figure 3.17. Experimental (gas phase) and MM (MM3) and MP2(FC)/6-31G* calculated infra red spectra of methanol. The MM3 spectrum is based on the data in Table 3.8; the MP2 spectrum is that shown in Fig. 5.36.

Table 3.5. Calculated (MM3) IR spectrum of acetone

Frequency (cm^{-1})	Relative intensity
3018	17
3016	11
2990	27
2902	9
1727	100
1411	49
1211	7

Intensities have been normalized, and those of less than 2 percent relative intensity have been ignored.

Table 3.6. Calculated (MM3) IR spectrum of benzene

Frequencies	Intensities
3051	96
1482	75
982	29
569	100

The intensities of degenerate frequencies have been combined, intensities have been normalized, and those of less than 2 percent relative intensity have been ignored.

Table 3.7. Calculated (MM3) IR spectrum of dichloromethane

Frequencies	Intensities
3071	25
3003	8
1389	13
813	3
790	100
661	59

Intensities have been normalized, and those of less than 2 percent relative intensity have been ignored.

Table 3.8. Calculated (MM3) IR spectrum of methanol

Frequencies	Intensities
3679	98
2993	17
2930	19
2847	3
1480	12
1287	100
1084	15
1051	19

Intensities have been normalized, and those of less than 2 percent relative intensity have been ignored.

3.6 STRENGTHS AND WEAKNESSES OF MM

3.6.1 Strengths

Molecular mechanics is *fast,* as shown by the times for optimization of $C_{20}H_{42}$ in section 3.3. The speed of MM is not always at the expense of *accuracy*: for the kinds of molecules for which it has been parameterized, it can rival or surpass experiment in the reliability of its results (sections 3.3 and 3.4). MM is *undemanding* in its hardware requirements: except perhaps for work on large biopolymers, MM calculations on moderately well-equipped personal computers are quite practical. The characteristics of speed, (frequent) accuracy and modest computer requirements have given MM a place in many modelling programs.

Because of its speed and the availability of parameters for almost all the elements (section 3.3), MM – even when it does not provide very accurate geometries – can supply reasonably good input geometries for SE, *ab initio* or density functional calculations, and this is one of its main applications. The fairly recent ability of MM programs to calculate IR spectra with some accuracy [16,32] may presage an important application, since frequency calculation by quantum mechanical methods usually requires considerably more time than geometry optimization). Note that MM frequencies should be calculated using the MM geometry – unfortunately, MM cannot be used as a shortcut to obtaining frequencies for a species optimized by a quantum mechanical calculation (*ab initio,* density functional or SE), since frequencies must be calculated using the same method used for the geometry optimization (section 2.5).

3.6.2 Weaknesses

The possible pitfalls in using MM are discussed by Lipkowitz [33]. The weaknesses of MM stem from the fact that it ignores electrons. The philosophy behind MM is to think of a molecule as a collection of atoms subject to forces and to use any practical mathematical treatment of these forces to express the energy in terms of the geometric parameters. By parameterization MM can “calculate” electronic properties; for example, using bond dipoles it can find a dipole moment for a molecule, and using values that have been calculated for various atom types by quantum mechanics it can assign charges to atoms. However, such results are obtained purely by analogy, and their reliability can be negated by unexpected electronic factors to which MM is oblivious. MM cannot provide information about the shapes and energies of molecular orbitals nor about related phenomena such as electronic spectra.

Because of the severely empirical nature of MM, interpreting MM parameters in terms of traditional physical concepts is dangerous; for example, the bond-stretching and angle-bending parameters cannot rigorously be identified with spectroscopic force constants [33]; Lipkowitz suggests that the MM proportionality constants (section 3.2.1) be called *potential constants.* Other dangers in using MM are the following:

(1) *Using an inappropriate forcefield.* A field parameterized for one class of compounds is not likely to perform well for other classes.

(2) *Transferring parameters form one forcefield to another.* This is usually not valid.

(3) *Optimizing to a stationary point that may not really be a minimum* (it could be a “maximum”, a transition state), and certainly may not be a *global* minimum (chapter 2). If there is reason to be concerned that a structure is not a minimum, alter it slightly by bond rotation and reoptimize; a transition state should slide down toward a nearby minimum (e.g. eclipsed ethane altered slightly from the D_{3h} geometry and optimized goes to the staggered conformer (Fig. 3.9).

(4) *Being taken in by vendor hype.* MM programs, more so than SE ones and unlike *ab initio* or DFT programs, are ruled by empirical factors (the form of the forcefield and the parameters used in it), and vendors do not usually caution buyers about potential deficiencies.

(5) *Ignoring solvent and nearby ions.* For polar molecules using the *in vacuo* structure can lead to quite wrong geometries and energies. This is particularly important for biomolecules. One way to mitigate this problem is to explicitly add solvent molecules or ions to the system, which can considerably increase the time fora calculation. Another might be to subject various plausible *in vacuo*-optimized conformations to single-point

(no geometry optimization) calculations that simulate the effect of solvent and take the resulting energies as being more reliable than the *in vacua* ones.

(6) *Lack of caution about comparing energies calculated with MM.* The method calculates the energy of a molecule relative to a hypothetical strainless idealization of the molecule. Using MM to calculate the relative energy of two isomers by comparing their strain energies (the normal MM energies) is dangerous because the two strain energies are not necessarily relative to the same hypothetical unstrained species (strain energies are not an unambiguous observable [34]). This is particularly true for functional group isomers, like $(CH_3)_2O/CH_3CH_2OH$ and $CH_3COCH_3/H_2C{=}C(OH)CH_3$, which have quite different atom types. For isomers consisting of the same kinds of atoms (alkanes cf. alkanes, say), and especially for conformational isomers and *E/Z* isomers (geometric isomers), a good MM forcefield should give strain energies which reasonably represent relative enthalpies. For example, the MMFF gives for $CH_3COCH_3/H_2C{=}C(OH)CH_3$ strain energies of $6.9/{-}6.6\,\mathrm{kJ\,mol^{-1}}$, i.e. relative energies of $0/{-}0.3\,\mathrm{kJ\,mol^{-1}}$, but the experimental value is ca. $0/44\,\mathrm{kJ\,mol^{-1}}$, i.e. $H_2C{=}C(OH)CH_3$ is much the higher-energy molecule. On the other hand, the MMFF yields for *gauche*-butane/*anti*-butane strain energies of $-21.3/-18.0\,\mathrm{kJ\,mol^{-1}}$, i.e. relative energies of $0/3.3\,\mathrm{kJ\,mol^{-1}}$, reasonably close to the experimental value of $0/2.8\,\mathrm{kJ\,mol^{-1}}$. For chair ($D_{2d}$), twist ($D_2$), and boat ($C_{2v}$) cyclohexane, the MMFF strain energies are –14.9, 9.9, and $13.0\,\mathrm{kJ\,mol^{-1}}$, i.e. relative energies of 0, 24.8 and $27.9\,\mathrm{kJ\,mol^{-1}}$, cf. the experimental the estimates of 0,24 and $29\,\mathrm{kJ\,mol^{-1}}$. MM programs can be parameterized to give, not just strain energy, but enthalpies of formation [1f], and the use of these enthalpies should make possible energy comparisons between isomers of disparate structural kinds.

Although chemists often compare stabilities of isomers using enthalpies, we should remember that equilibria are actually determined by free energies. The lowest-enthalpy isomer is not *necessarily* the one of lowest free energy: a higher-enthalpy molecule may have more vibrational and torsional motion (it may be springier and floppier) and thus possess more entropy and hence have a lower free energy. Free energy has an enthalpy and an entropy component, and to calculate the latter, one needs the vibrational frequencies. Programs that calculate frequencies will usually also provide entropies, and with parameterization for enthalpy this can permit the calculation of free energies. Note that the species of lowest free energy is not always the major one present: one low-energy conformation could be outnumbered by one hundred of higher energy, each demanding its share of the Boltzmann pie.

(7) *Assuming that the major conformation determines the product.* In fact, in a mobile equilibrium the product ratio depends on the relative reactivities, not relative amounts, of the conformers (the Curtin–Hammett principle [35]).

(8) *Failure to exercise judgement: small energy differences* (say up to $10\text{–}20\,\mathrm{kJ\,mol^{-1}}$) mean nothing in many cases. The excellent energy results referred to in section 3.3 can be expected only for families of molecules (usually small to medium-sized) for which the forcefield has been parameterized.

Many of the above dangers can be avoided simply by performing test calculations on systems for which the results are known (experimentally, or "known" from high-level quantum mechanical calculations). Such a reality check can have salutary effects on the reliability of one's results, and not only with reference to MM.

3.7 SUMMARY OF CHAPTER 3

This chapter explains the basic principles of MM, which rests on a view of molecules as balls held together by springs. MM began in the 1940s with attempts to analyze the rates of racemization of biphenyls and S_N2 reactions.

The potential energy of a molecule can be written as the sum of terms involving bond stretching, angle bending, dihedral angles and nonbonded interactions. Giving these terms explicit mathematical forms constitutes devising a forcefield, and giving actual numbers to the constants in the forcefield constitutes parameterizing the field. An example is given of the devising and parameterization of an MM forcefield.

MM is used mainly to calculate geometries and energies for small to medium-sized molecules. Such calculations are fast and can be very accurate, provided that the forcefield has been carefully parameterized for the types of molecules under study. Calculations on biomolecules is a very important application of MM; the pharmaceutical industry designs new drugs with the aid of MM: for example, examining how various candidate drugs fit into the active sites of biomolecules (docking) and the related aspect of QSAR are of major importance. MM is of some limited use in calculating the geometries and energies of transition states. Organic synthesis now makes considerable use of MM, which enables chemists to estimate which products are likely to be favored and to devise more realistic routes to a target molecule than was hitherto possible. In molecular dynamics MM is used to generate the forces acting on molecules and hence to calculate their motions, and in Monte Carlo simulations MM is used to calculate the energies of the many randomly generated states.

MM is fast, it can be accurate, it is undemanding of computer power, and it provides reasonable starting geometries for quantum mechanical calculations. MM ignores electrons, and so can provide parameters like dipole moment only by analogy. One must be cautious about the applicability of MM parameters to the problem at hand. Stationary points from MM, even when they are relative minima, may not be global minima. Ignoring solvent effects can give erroneous results for polar molecules. MM gives strain energies, the difference of which for structurally similar isomers represent enthalpy differences; parameterization to give enthalpies of formation is possible. Strictly speaking, relative amounts of isomers depend on free energy differences. The major conformation (even when correctly identified) is not necessarily the reactive one.

REFERENCES

[1] General references to MM: (a) A. K. Rappe, C. L. Casewit, “Molecular mechanics across chemistry,” University Science Books, Sausalito, CA, 1997, website http://www.chm.colostate.edu/mmac; (b) A. R. Leach, “Molecular Modelling, Principles and Applications,” Addison Wesley Longman, Essex (UK), 1996, chapter 3; (c) U. Burkert, N. L. Allinger, “Molecular Mechanics”, ACS Monograph 177, American Chemical Society, Washington, DC, 1982; (d) N. L. Allinger, Calculation of Molecular Structures and Energy by Force Methods”, in *Advances in Physical Organic Chemistry, 13,* V. Gold and D. Bethell, Eds., Academic Press, New York, 1976; (e) T. Clark, “A Handbook of Computational Chemistry” Wiley, New York, 1985; (f) I. N. Levine, “Quantum Chemistry,” 4th edn, Prentice Hall, Engelwood Cliffs, New Jersey, 1991, pp. 583–587; (g) Issue

No. 7 of Chem. Rev., 1993, *93*; (h) Conformational energies: I. Pettersson, T. Liljefors, in *Reviews in Computational Chemistry,* 1996, *9*; (i) Inorganic and organometallic compounds: C. R. Landis, D. M. Root, T. Cleveland, in *Reviews in Computational Chemistry,* 1995, *6;* (j) Parameterization: J. P. Bowen, N. L. Allinger, *Reviews in Computational Chemistry,* 1991, *2.*

[2] MM history: (a) Ref. [1]; (b) E. M. Engler, J. D. Andose, P. von R. Schleyer, J. Am. Chem. Soc., 1973, *95,* 8005 and references therein; (c) MM up to the end of 1967 is reviewed in detail in: J. E. Williams, P. J. Stang, P. von R. Schleyer, Annu. Rev. Phys. Chem., 1968, *19*, 531.

[3] (a) F. H. Westheimer, J. E. Mayer, J. Chem. Phys., 1946, *14,* 733; (b) T. L. Hill, J. Chem. Phys., 1946, *14,* 465; (c) I. Dostrovsky, E. D. Hughes, C. K. Ingold, J. Chem. Soc., 1946, 173; (d) F. H. Westheimer, J. Chem. Phys., 1947, *15*, 252.

[4] B. Ma, J-H Lii, K. Chen, N. L. Allinger, J. Am. Chem. Soc., 1997, *119*, 2570 and references therein.

[5] (a) Information on and references to MM programs may be found in Ref. [1] and the website of Ref. [1a]; (b) For papers on the popular MMFF and the MM4 forcefield (and information on some others) see the issue of J. Comp. Chem., 1996, *17.*

[6] The force constant is defined as the proportionality constant in the equation force $= k\times$ extension (of length or angle), so integrating force with respect to extension to get the energy (= force × extension) needed to stretch the bond gives $E = (k/2)(\text{extension})^2$, i.e. $k = \text{force constant} = 2k_{\text{stretch}}$ (or $2k_{\text{bend}}$).

[7] (a) A brief discussion and some parameters: P. W. Atkins, "Physical Chemistry," 5th edn, Freeman New York, 1994, pp.772, 773; it is pointed out here that $e^{-r/\sigma}$ is actually a much better representation of the compressive potential than is r^{-12}; (b) W. J. Moore, "Physical Chemistry," 4th edn, Prentice-Hall, New Jersey, 1972, p. 158 (from J. O. Hirschfelder, C. F. Curtis and R. B. Bird, "Molecular Theory of Gases and Liquids," Wiley, New York, 1954). Note that our k_{nb} is called 4ϵ here and must be multiplied by 8.31/1000 to convert it to our units of kJ mol^{-1}.

[8] Infra red spectroscopy: R. M. Silverstein, G. C. Bassler and T. C. Morrill, "Spectrometric Identification of Organic Compounds," 4th edn, Wiley, New York, 1981, chapter 3.

[9] Different methods of structure determination give somewhat different results; this is discussed in Ref. [4] and in: (a) B. Ma, J-H Lii, H. F. Schaefer, N. L. Allinger, J. Phys. Chem., 1996, *100*, 8763; (b) A. Domenicano, I. Hargittai, Eds., "Accurate Molecular Structures," Oxford Science Publications, New York, 1992.

[10] To properly parameterize a MM force field only high-level *ab initio* calculations (or density functional) calculations would actually be used, but this does not affect the principle being demonstrated.

[11] Ref. [1b, pp. 148–181].

[12] MM3: N. L. Allinger, Y. H. Yuh, J-H Lii, J. Am. Chem Soc., 1989, *111*, 8551. The calculation was performed with MM3 as implemented in Spartan (Ref. [13]).

[13] Spartan SGI Version 4.0.4; Spartan is a comprehensive computational chemistry program with MM, *ab initio,* density functional and SE capability, combined with powerful graphical input and output.

[14] Ref. [1c, pp. 182–184].

[15] K. Gundertofte, T. Liljefors, P.-O. Norby, I. Pettersson, J. Comp. Chem., 1996, *17*, 429.

[16] Comparison of various force fields for geometry (and vibrational frequencies): T. A. Halgren, J. Comp. Chem., 1996, *17*, 553.

[17] The Merck force field (Ref. [18]) often gives geometries that are satisfactory for energy calculations (i.e. for single-point energies) with quantum mechanical methods; this could be very useful for large molecules: W. J. Hehre, J. Yu, P. E. Klunzinger, "A Guide to Molecular Mechanics and Molecular Orbital calculations in Spartan," Wavefunction Inc., Irvine CA, 1997, chapter 4.

[18] T. A. Halgren, J. Comp. Chem., 1996, *17,* 490.

[19] M. C. Nicklaus, J. Comp. Chem., 1997, *18*, 1056; the difference between CHARMM and CHARMm is explained here.

[20] (a) www.amber.ucsf.edu/amber/amber.html; (b) W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M Ferguson, D. C Sellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, J. Am. Chem. Soc., 1995, *117*, 5179; (c) V. Barone, G. Capecchi, Y. Brunei, M.-L. D. Andries, R. Subra, J. Comp. Chem., 1997, *18*, 1720.

[21] (a) M. J. Field, P. A. Bash, M. Karplus, J. Comp. Chem., 1990, *11*, 700; (b) P. A. Bash. M. J. Field, M. Karplus, J. Am. Chem. Soc., 1987, *109*, 8092.

[22] U. C. Singh, P. Kollman, J. Comp Chem., 1986, *7,* 718.

[23] (a) Reference [1b, chapter 10]; (b) H.-D. Höltje, G. Folkers "Molecular Modelling, Applications in Medicinal Chemistry," VCH, Weinheim, Germany, 1996; (c) "Computer-Assisted Lead Finding and Optimization," H. van de Waterbeemd, B. Testa, G. Folkers, Eds., VCH, Weinheim, Germany, 1997.

[24] J. E. Eksterowicz and K. N. Houk, Chem. Rev., 1993, *93,* 2439.

[25] M. B. Smith and J. March, "March's Advanced Organic Chemistry," Wiley, New York, 2000, pp. 284–285.

[26] K. B. Lipkowitz, M. A. Peterson, Chem Rev., 1993, *93*, 2463.

[27] M. Saunders, Science, 1991, *253*, 330.

[28] (a) Ref. [1b, chapter 6]; (b) M. Karplus, G. A. Putsch, Nature, 1990, *347,* 631; (c) C. L. Brooks III, D. A. Case, Chem. Rev., 1993, *93,* 2487; (d) M. Eichinger, H. Grubmüller, H. Heller, P. Tavan, J. Comp. Chem., 1997, *18,* 1729; (e) G. E. Marlow, J. S. Perkyns, B. M. Pettitt, Chem. Rev., 1993, *93,* 2503; (f) J. Aqvist, A. Warshel, Chem. Rev., 1993, *93,* 2523.

[29] Ref. [1b, chapter 7].

[30] W. J. Hehre, L. Radom, p. v. R. Schleyer, J. A. Pople, "*Ab initio* Molecular Orbital Theory," Wiley, New York, 1986.

[31] J.-H. Lii and N. L. Allinger, J. Comp. Chem., 1992, *13*, 1138. The MM3 program with the ability to calculate IR intensities can be bought by academic users from QCPE (the Quantum Chemistry Exchange Program of the University of Indiana) and by commercial users from Tripos Associates of St. Louis, Missouri.

[32] N. Nevins, N. L. Allinger, J. Comp. Chem., 1996, *17*, 730 and references therein.

[33] K. B. Lipkowitz, J. Chem. Ed., 1995, *72,* 1070.

[34] (a) K. Wiberg, Angew. Chem., Int. Ed. Engl., 1986, *25*, 312; (b) Issue No. 5 of Chem. Rev., 1989, *89*; (c) S. Inagaki, Y. Ishitani, T. Kakefu, J. Am. Chem. Soc., 1994, *116*, 5954; (d) S. Nagase, Acc. Chem. Res., 1995, *28*, 469; (e) S. Gronert, J. M. Lee, J. Org. Chem., 1995, *60,* 6731; (f) A. Sella, H. Basch, S. Hoz, J. Am. Chem. Soc., 1996, *118*, 416; (g)

S. Grime, J. Am. Chem. Soc., 1996, *118,* 1529; (h) V. Balaji, J. Michl, Pure Appl. Chem., 1988, *60*, 189; (i) K. B. Wiberg, J. W. Ochterski, J. Comp. Chem., 1997, *18*, 108.

[35] J. I. Seeman, Chem. Rev., 1983, *83*, 83.

EASIER QUESTIONS

1. What is the basic idea behind MM?
2. What is a forcefield?
3. What are the two basic approaches to parameterizing a forcefield?
4. Why does parameterizing a forcefield for transition states present special problems?
5. What is the main advantage of MM, generally speaking, over the other methods of calculating molecular geometries and relative energies?
6. Why is it not valid in all cases to obtain the relative energies of isomers by comparing their MM strain energies?
7. What class of problems cannot be dealt with by MM?
8. Give four applications for MM. Which is the most widely used?
9. MM can calculate the values (cm^{-1}) of vibrational frequencies, but without "outside assistance" it can't calculate their intensities. Explain.
10. Why is it not valid to calculate a geometry by some slower (e.g. *ab initio)* method, then use that geometry for a fast MM frequency calculation?

HARDER QUESTIONS

1. One big advantage of MM over other methods of calculating geometries and relative energies is speed. Does it seem likely that continued increases in computer speed could make MM obsolete?
2. Do you think it is possible (in practical terms? In principle?) to develop a forcefield that would accurately calculate the geometry of any kind of molecule?
3. What advantages or disadvantages are there to parameterizing a forcefield with the results of "high-level" calculations rather than the results of experiments?
4. Would you dispute the suggestion that no matter how accurate a set of MM results might be, they cannot provide insight into the factors affecting a chemical problem, because the "ball and springs" model is unphysical?
5. Would you agree that hydrogen bonds (e.g. the attraction between two water molecules) might be modelled in MM as weak covalent bonds, as strong van der Waals or dispersion forces, or as electrostatic attractions? Is any one of these three approaches to be preferred in principle?
6. Replacing small groups by "pseudoatoms" in a forcefield (e.g. CH_3 by an "atom" about as big) obviously speeds up calculations. What disadvantages might accompany this simplification?
7. Why might the development of an accurate and versatile forcefield for inorganic molecules be more of a challenge than for organic molecules?

8. What factor(s) might cause an electronic structure calculation (e.g. *ab initio* or DFT) to give geometries or relative energies very different from those obtained from MM?
9. Compile a list of molecular characteristics/properties that cannot be calculated purely by MM.
10. How many parameters do you think a reasonable forcefield would need to minimize the geometry of 1,2-dichloroethane?