Let us review again the sense of Eqs. (2.4) and (2.9). In both instances, the minimum value for the energy is zero (assuming positive force constants and sensible behavior for odd power terms). An energy of zero is obtained when the bond length or angle adopts its equilibrium value. Thus, a 'strain-free' molecule is one in which every coordinate adopts its equilibrium value. Although we accepted a negative torsional term in our fluoromethanol example above, because it provided some chemical insight, by proper choice of phase angles in Eq. (2.10) we could also require this energy to have zero as a minimum (although not necessarily for the dihedral angle $\omega = \pi$). So, neglecting non-bonded terms for the moment, we see that the raw force-field energy can be called the 'strain energy', since it represents the positive deviation from a hypothetical strain-free system.

The key point that must be noted here is that strain energies for two different molecules *cannot be meaningfully compared unless the zero of energy is identical.* This is probably best illustrated with a chemical example. Consider a comparison of the molecules ethanol and dimethyl ether using the MM2(91) force field. Both have the chemical formula C_2H_6O. However, while ethanol is defined by the force field to be composed of two sp^3 carbon atoms, one sp^3 oxygen atom, five carbon-bound hydrogen atoms, and one alcohol hydrogen atom, dimethyl ether differs in that all six of its hydrogen atoms are of the carbon-bound type. Each strain energy will thus be computed relative to a different hypothetical reference system, and there is no *a priori* reason that the two hypothetical systems should be thermodynamically equivalent.

What is necessary to compute a heat of formation, then, is to define the heat of formation of each hypothetical, unstrained atom type. The molecular heat of formation can then be computed as the sum of the heats of formation of all of the atom types plus the strain energy. Assigning atom-type heats of formation can be accomplished using additivity methods originally developed for organic functional groups (Cohen and Benson 1993). The process is typically iterative in conjunction with parameter determination.

Since the assignment of the atomic heats of formation is really just an aspect of parameterization, it should be clear that the possibility of a negative force-field energy, which could derive from addition of net negative non-bonded interaction energies to small non-negative strain energies, is not a complication. Thus, a typical force-field energy calculation will report any or all of (i) a strain energy, which is the energetic consequence of the deviation of the internal molecular coordinates from their equilibrium values, (ii) a force-field energy, which is the sum of the strain energy and the non-bonded interaction energies, and (iii) a heat of formation, which is the sum of the force-field energy and the reference heats of formation for the constituent atom types (Figure 2.8).

For some atom types, thermodynamic data may be lacking to assign a reference heat of formation. When a molecule contains one or more of these atom types, the force field cannot compute a molecular heat of formation, *and energetic comparisons are necessarily limited to conformers, or other isomers that can be formed without any change in atom types.*

2.4 Geometry Optimization

One of the key motivations in early force-field design was the development of an energy functional that would permit facile optimization of molecular geometries. While the energy

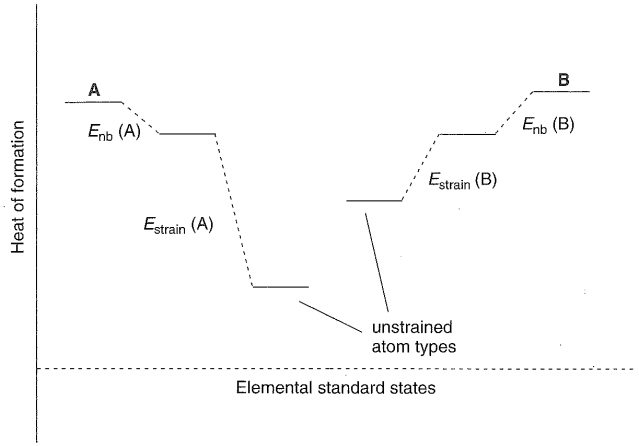

Figure 2.8 Molecules **A** and **B** are chemical isomers but are composed of different atomic types (atomic typomers?). Thus, the sums of the heats of formation of their respective unstrained atom types, which serve as their zeroes of force-field energy, are different. To each zero, strain energy and non-bonded energy (the sum of which are force-field energy) are added to determine heat of formation. In this example, note that **A** is predicted to have a lower heat of formation than **B** even though it has a substantially larger strain energy (and force-field energy); this difference is more than offset by the difference in the reference zeroes

of an *arbitrary* structure can be interesting, real molecules vibrate thermally about their equilibrium structures, so finding minimum energy structures is key to describing equilibrium constants, comparing to experiment, etc. Thus, as emphasized above, one priority in force-field development is to adopt reasonably simple functional forms so as to facilitate geometry optimization. We now examine the optimization process in order to see how the functional forms enter into the problem.

2.4.1 Optimization Algorithms

Note that, in principle, geometry optimization could be a separate chapter of this text. In its essence, geometry optimization is a problem in applied mathematics. How does one find a minimum in an arbitrary function of many variables? [Indeed, we have already discussed that problem once, in the context of parameter optimization. In the case of parameter optimization, however, it is not necessarily obvious how the penalty function being minimized *depends* on any given variable, and moreover the problem is highly underdetermined. In the case of geometry optimization, we are working with far fewer variables (the geometric degrees of freedom) and have, at least with a force field, analytic expressions for how the energy depends on the variables. The mathematical approach can thus be quite different.] As the problem is general, so, too, many of the details presented below will be general to any energy

functional. However, certain special considerations associated with force-field calculations merit discussion, and so we will proceed first with an overview of geometry optimization, and then examine force-field specific aspects.

Because this text is designed primarily to illuminate the conceptual aspects of computational chemistry, and not to provide detailed descriptions of algorithms, we will examine only the most basic procedures. Much more detailed treatises of more sophisticated algorithms are available (see, for instance, Jensen 1999).

For pedagogical purposes, let us begin by considering a case where we do not know how our energy depends on the geometric coordinates of our molecule. To optimize the geometry, all we can do is keep trying different geometries until we are reasonably sure that we have found the one with the lowest possible energy (while this situation is atypical with force fields, there are still many sophisticated electronic structure methods for which it is indeed the only way to optimize the structure). How can one most efficiently survey different geometries?

It is easiest to proceed by considering a one-dimensional case, i.e., a diatomic with only the bond length as a geometric degree of freedom. One selects a bond length, and computes the energy. One then changes the bond length, let us say by shortening it 0.2 Å, and again computes the energy. If the energy goes down, we want to continue moving the bond length in that direction, and we should take another step (which need not necessarily be of the same length). If the energy goes up, on the other hand, we are moving in the wrong direction, and we should take a step in the opposite direction. Ultimately, the process will provide three adjacent points where the one in the center is lower in energy than the other two. Three non-collinear points uniquely define a parabola, and in this case the parabola must have a minimum (since the central point was lower in energy than the other two). We next calculate the energy for the bond length corresponding to the parabolic minimum (the degree to which the computed energy agrees with that from the parabolic equation will be an indication of how nearly harmonic the local bond stretching coordinate is). We again step left and right on the bond stretching coordinate, this time with smaller steps (perhaps an order of magnitude smaller) and repeat the parabolic fitting process. This procedure can be repeated until we are satisfied that our step size falls below some arbitrary threshold we have established as defining convergence of the geometry. Note that one can certainly envision variations on this theme. One could use more than three points in order to fit to higher order polynomial equations, step sizes could be adjusted based on knowledge of previous points, etc.

In the multi-dimensional case, the simplest generalization of this procedure is to carry out the process iteratively. Thus, for LiOH, for example, we might first find a parabolic minimum for the OH bond, then for the LiO bond, then for the LiOH bond angle (in each case holding the other two degrees of freedom fixed), and then repeat the process to convergence. Of course, if there is strong coupling between the various degrees of freedom, this process will converge rather slowly.

What we really want to do at any given point in the multi-dimensional case is move not in the direction of a *single* coordinate, but rather in the direction of the greatest downward slope in the energy with respect to *all* coordinates. This direction is the opposite of the

gradient vector, $\mathbf{g}$, which is defined as

$$\mathbf{g}(\mathbf{q}) = \begin{bmatrix} \dfrac{\partial U}{\partial q_1} \\ \dfrac{\partial U}{\partial q_2} \\ \dfrac{\partial U}{\partial q_3} \\ \vdots \\ \dfrac{\partial U}{\partial q_n} \end{bmatrix} \tag{2.32}$$

where $\mathbf{q}$ is an n-dimensional coordinate vector ($n = 3N - 6$ where N is the number of atoms if we are working in internal coordinates, $n = 3N$ if we are working in Cartesian coordinates, etc.) If we cannot compute the partial derivatives that make up $\mathbf{g}$ analytically, we can do so numerically. However, that numerical evaluation requires at least one additional energy calculation for each degree of freedom. Thus, we would increase (or decrease) every degree of freedom by some step size, compute the slope of the resulting line derived from the energies of our initial structure and the perturbed structure, and use this slope as an estimate for the partial derivative. Such a 'forward difference' estimation is typically not very accurate, and it would be better to take an additional point in the opposite direction for each degree of freedom, and then compute the 'central difference' slope from the corresponding parabola. It should be obvious that, as the number of degrees of freedom increases, it can be particularly valuable to have an energy function for which the first derivative is known *analytically*.

Let us examine this point a bit more closely for the force-field case. For this example, we will work in Cartesian coordinates, in which case $\mathbf{q} = \mathbf{X}$ of Eq. (1.4). To compute, say, the partial derivative of the energy with respect to the x coordinate of atom A, we will need to evaluate the changes in energy for the various terms contributing to the full force-field energy as a function of moving atom A in the x direction. For simplicity, let us consider only the bond stretching terms. Clearly, only the energy of those bonds that have A at one terminus will be affected by A's movement. We may then use the chain rule to write

$$\frac{\partial U}{\partial x_A} = \sum_{\substack{i \\ \text{bonded} \\ \text{to A}}} \frac{\partial U}{\partial r_{\mathrm{A}i}} \frac{\partial r_{\mathrm{A}i}}{\partial x_A} \tag{2.33}$$

Differentiation of E with respect to $r_{\mathrm{A}i}$ for Eq. (2.4) gives

$$\frac{\partial U}{\partial r_{\mathrm{A}i}} = \tfrac{1}{2}[2k_{\mathrm{A}i} + 3k_{\mathrm{A}i}^{(3)}(r_{\mathrm{A}i} - r_{\mathrm{A}i,\mathrm{eq}}) + 4k_{\mathrm{A}i}^{(4)}(r_{\mathrm{A}i} - r_{\mathrm{A}i,\mathrm{eq}})^2](r_{\mathrm{A}i} - r_{\mathrm{A}i,\mathrm{eq}}) \tag{2.34}$$

The bond length r_{Ai} was defined in Eq. 2.15, and its partial derivative with respect to x_A is

$$\frac{\partial r_{Ai}}{\partial x_A} = \frac{(x_A - x_i)}{\sqrt{(x_A - x_i)^2 + (y_A - y_i)^2 + (z_A - z_i)^2}} \tag{2.35}$$

Thus, we may quickly assemble the bond stretching contributions to this particular component of the gradient. Contributions from the other terms in the force field can be somewhat more tedious to derive, but are nevertheless available analytically. This makes force fields highly efficient for the optimization of geometries of very large systems.

With $\mathbf{g}$ in hand, we can proceed in a fashion analogous to the one-dimensional case outlined above. We step along the direction defined by $-\mathbf{g}$ until we locate a minimum in the energy for this process; since we are taking points in a linear fashion, this movement is called a 'line search' (even though we may identify our minimum by fitting our points to a polynomial curve). Then, we recompute $\mathbf{g}$ at the located minimum and repeat the process. Our new search direction is necessarily orthogonal to our last one, since we minimized E in the last direction. This particular feature of a steepest descent curve can lead to *very* slow convergence in unfavorable cases.

A more robust method is the Newton–Raphson procedure. In Eq. (2.26), we expressed the full force-field energy as a multidimensional Taylor expansion in arbitrary coordinates. If we rewrite this expression in matrix notation, and truncate at second order, we have

$$U(\mathbf{q}^{(k+1)}) = U(\mathbf{q}^{(k)}) + (\mathbf{q}^{(k+1)} - \mathbf{q}^{(k)})\mathbf{g}^{(k)} + \tfrac{1}{2}(\mathbf{q}^{(k+1)} - \mathbf{q}^{(k)})^{\dagger}\mathbf{H}^{(k)}(\mathbf{q}^{(k+1)} - \mathbf{q}^{(k)}) \tag{2.36}$$

where the reference point is $\mathbf{q}^{(k)}$, $\mathbf{g}^{(k)}$ is the gradient vector for the reference point as defined by Eq. (2.32), and $\mathbf{H}^{(k)}$ is the 'Hessian' matrix for the reference point, whose elements are defined by

$$H_{ij}^{(k)} = \left.\frac{\partial^2 U}{\partial q_i \partial q_j}\right|_{\mathbf{q}=\mathbf{q}^{(k)}} \tag{2.37}$$

If we differentiate Eq. (2.36) term by term with respect to the ith coordinate of $\mathbf{q}^{(k+1)}$, noting that no term associated with point k has any dependence on a coordinate of point $k+1$ (and hence the relevant partial derivative will be 0), we obtain

$$\frac{\partial U(\mathbf{q}^{(k+1)})}{\partial q_i^{k+1}} = \frac{\partial \mathbf{q}^{(k+1)}}{\partial q_i^{k+1}}\mathbf{g}^{(k)} + \frac{1}{2}\frac{\partial \mathbf{q}^{(k+1)^{\dagger}}}{\partial q_i^{k+1}}\mathbf{H}^{(k)}(\mathbf{q}^{(k+1)} - \mathbf{q}^{(k)}) + \frac{1}{2}(\mathbf{q}^{(k+1)} - \mathbf{q}^{(k)})^{\dagger}\mathbf{H}^{(k)}\frac{\partial \mathbf{q}^{(k+1)}}{\partial q_i^{k+1}} \tag{2.38}$$

The l.h.s. of Eq. (2.38) is the ith element of the vector $\mathbf{g}^{(k+1)}$. On the r.h.s. of Eq. (2.38), since the partial derivative of $\mathbf{q}$ with respect to its ith coordinate is simply the unit vector in the ith coordinate direction, the various matrix multiplications simply produce the ith element of the multiplied vectors. Because mixed partial derivative values are independent of the order of differentiation, the Hessian matrix is Hermitian, and we may simplify

Eq. (2.38) as

$$g_i^{(k+1)} = g_i^{(k)} + [\mathbf{H}^{(k)}(\mathbf{q}^{(k+1)} - \mathbf{q}^{(k)})]_i \tag{2.39}$$

where the notation $[]_i$ indicates the ith element of the product column matrix. The condition for a stationary point is that the l.h.s. of Eq. (2.39) be 0 for *all* coordinates, or

$$\mathbf{0} = \mathbf{g}^{(k)} + \mathbf{H}^{(k)}(\mathbf{q}^{(k+1)} - \mathbf{q}^{(k)}) \tag{2.40}$$

which may be rearranged to

$$\mathbf{q}^{(k+1)} = \mathbf{q}^{(k)} - (\mathbf{H}^{(k)})^{-1}\mathbf{g}^{(k)} \tag{2.41}$$

This equation provides a prescription for the location of stationary points. In principle, starting from an arbitrary structure having coordinates $\mathbf{q}^{(k)}$, one would compute its gradient vector $\mathbf{g}$ and its Hessian matrix $\mathbf{H}$, and then select a new geometry $\mathbf{q}^{(k+1)}$ according to Eq. (2.41). Equation (2.40) shows that the gradient vector for this new structure will be the $\mathbf{0}$ vector, so we will have a stationary point.

Recall, however, that our derivation involved a truncation of the full Taylor expansion at second order. Thus, Eq. (2.40) is only approximate, and $\mathbf{g}^{(k+1)}$ will not necessarily be $\mathbf{0}$. However, it will probably be smaller than $\mathbf{g}^{(k)}$, so we can repeat the whole process to pick a point $k+2$. After a sufficient number of iterations, the gradient will hopefully become so small that structures $k+n$ and $k+n+1$ differ by a chemically insignificant amount, and we declare our geometry to be converged.

There are a few points with respect to this procedure that merit discussion. First, there is the Hessian matrix. With n^2 elements, where n is the number of coordinates in the molecular geometry vector, it can grow somewhat expensive to construct this matrix at every step even for functions, like those used in most force fields, that have fairly simple analytical expressions for their second derivatives. Moreover, the matrix must be *inverted* at every step, and matrix inversion formally scales as n^3, where n is the dimensionality of the matrix. Thus, for purposes of efficiency (or in cases where analytic second derivatives are simply not available) approximate Hessian matrices are often used in the optimization process – after all, the truncation of the Taylor expansion renders the Newton–Raphson method *intrinsically* approximate. As an optimization progresses, second derivatives can be estimated reasonably well from finite differences in the analytic first derivatives over the last few steps. For the first step, however, this is not an option, and one typically either accepts the cost of computing an initial Hessian analytically for the level of theory in use, or one employs a Hessian obtained at a less expensive level of theory, when such levels are available (which is typically *not* the case for force fields). To speed up slowly convergent optimizations, it is often helpful to compute an analytic Hessian every few steps and replace the approximate one in use up to that point. For *really* tricky cases (e.g., where the PES is fairly flat in many directions) one is occasionally forced to compute an analytic Hessian for *every* step.

Another key issue to note is that Eq. (2.41) provides a prescription to get to what is usually the *nearest* stationary point, but there is no guarantee that that point will be a

minimum. The condition for a minimum is that all coordinate second derivatives (i.e., all diagonal elements of the Hessian matrix) be positive, but Eq. (2.41) places no constraints on the second derivatives. Thus, if one starts with a geometry that is very near a transition state (TS) structure, the Newton–Raphson procedure is likely to converge to that structure. This can be a pleasant feature, if one is looking for the TS in question, or an annoying one, if one is not. To verify the nature of a located stationary point, it is necessary to compute an accurate Hessian matrix and inspect its eigenvalues, as discussed in more detail in Chapter 9. With force fields, it is often cheaper and equally effective simply to 'kick' the structure, which is to say, by hand one moves one or a few atoms to reasonably distorted locations and then reoptimizes to verify that the original structure is again found as the lowest energy structure nearby.

Because of the importance of TS structures, a large number of more sophisticated methods exist to locate them. Many of these methods require that two minima be specified that the TS structure should 'connect', i.e., the TS structure intervenes in some reaction path that connects them. Within a given choice of coordinates, intermediate structures are evaluated and, hopefully, the relevant stationary point is located. Other methods allow the specification of a particular coordinate with respect to which the energy is to be maximized while minimizing it with respect to all other coordinates. When this coordinate is one of the normal modes of the molecule, this defines a TS structure. The bottom line for all TS structure location methods is that they work best when the chemist can provide a reasonably good initial guess for the structure, and they tend to be considerably more sensitive to the availability of a good Hessian matrix, since finding the TS essentially amounts to distinguishing between different local curvatures on the PES.

Most modern computational chemistry software packages provide some discussion of the relative merits of the various optimizers that they make available, at least on the level of providing practical advice (particularly where the user can set certain variables in the optimization algorithm with respect to step size between structures, tolerances, use of redundant internal coordinates, etc.), so we will not try to cover all possible tricks and tweaks here. We will simply note that it is usually a good idea to visualize the structures in an optimization as it progresses, as every algorithm can sometimes take a pathologically bad step, and it is usually better to restart the calculation with an improved guess than it is to wait and hope that the optimization ultimately returns to normalcy.

A final point to be made is that most optimizers are rather good at getting you to the *nearest* minimum, but an individual researcher may be interested in finding the *global* minimum (i.e., the minimum having the lowest energy of all minima). Again, this is a problem in applied mathematics for which no one solution is optimal (see, for instance, Leach 1991). Most methods involve a systematic or random sampling of alternative conformations, and this subject will be discussed further in the next chapter.

2.4.2 Optimization Aspects Specific to Force Fields

Because of their utility for very large systems, where their relative speed proves advantageous, force fields present several specific issues with respect to practical geometry optimization that merit discussion. Most of these issues revolve around the scaling behavior

that the speed of a force-field calculation exhibits with respect to increasing system size. Although we raise the issues here in the context of geometry optimization, they are equally important in force-field simulations, which are discussed in more detail in the next chapter.

If we look at the scaling behavior of the various terms in a typical force field, we see that the internal coordinates have very favorable scaling – the number of internal coordinates is $3N - 6$, which is linear in N. The non-bonded terms, on the other hand, are computed from pairwise interactions, and therefore scale as N^2. However, this scaling assumes the evaluation of *all* pairwise terms. If we consider the Lennard–Jones potential, its long-range behavior decays proportional to r^{-6}. The total number of interactions should grow at most as r^2 (i.e., proportional to the surface area of a surrounding sphere), so the net energetic contribution should decay with an r^{-4} dependence. This quickly becomes negligible (particularly from a gradient standpoint) so force fields usually employ a 'cut-off' range for the evaluation of van der Waals energies – a typical choice is 10 Å. Thus, part of the calculation involves the periodic updating of a 'pair list', which is a list of all atoms for which the Lennard–Jones interaction needs to be calculated (Petrella *et al.* 2003). The update usually occurs only once every several steps, since, of course, evaluation of interatomic distances *also* formally scales as N^2.

In practice, even though the use of a cut-off introduces only small disparities in the energy, the discontinuity of these disparities can cause problems for optimizers. A more stable approach is to use a 'switching function' which multiplies the van der Waals interaction and causes it (and possibly its first and second derivatives) to go smoothly to zero at some cut-off distance. This function must, of course, be equal to 1 at short distances.

The electrostatic interaction is more problematic. For point charges, the interaction energy decays as r^{-1}. As already noted, the number of interactions increases by up to r^2, so the total energy in an infinite system might be expected to diverge! Such formal divergence is avoided in most real cases, however, because in systems that are electrically neutral there are as many positive interactions as negative, and thus there are large cancellation effects. If we imagine a system composed entirely of neutral groups (e.g., functional groups of a single molecule or individual molecules of a condensed phase), the long-range interaction between groups is a dipole–dipole interaction, which decays as r^{-3}, and the total energy contribution should decay as r^{-1}. Again, the actual situation is more favorable because of positive and negative cancellation effects, but the much slower decay of the electrostatic interaction makes it significantly harder to deal with. Cut-off distances (again, ideally implemented with smooth switching functions) must be quite large to avoid structural artifacts (e.g., atoms having large partial charges of like sign anomalously segregating at interatomic distances just in excess of the cut-off).

In infinite periodic systems, an attractive alternative to the use of a cut-off distance is the Ewald sum technique, first described for chemical systems by York, Darden and Pedersen (1993). By using a reciprocal-space technique to evaluate long-range contributions, the total electrostatic interaction can be calculated to a pre-selected level of accuracy (i.e., the Ewald sum limit is exact) with a scaling that, in the most favorable case (called 'Particle-mesh Ewald', or PME), is $N\log N$. Prior to the introduction of Ewald sums, the modeling of polyelectrolytes (e.g., DNA) was rarely successful because of the instabilities introduced

by cut-offs in systems having such a high degree of localized charges (see, for instance, Beveridge and McConnell 2000).

In aperiodic systems, another important contribution has been the development of the so-called 'Fast Multipole Moment' (FMM) method (Greengard and Rokhlin 1987). In essence, this approach takes advantage of the significant cancellations in charge–charge interactions between widely separated regions in space, and the increasing degree to which those interactions can be approximated by highly truncated multipole–multipole interactions. In the most favorable case, FMM methods scale linearly with system size.

It should be remembered, of course, that scaling behavior is informative of the relative time one system takes compared to another of different size, and says *nothing* about the *absolute* time required for the calculation. Thus, FMM methods scale linearly, but the initial overhead can be quite large, so that it requires a very large system before it outperforms PME for the same level of accuracy. Nevertheless, the availability of the FMM method renders conceivable the molecular modeling of extraordinarily large systems, and refinements of the method, for example the use of multiple grids (Skeel, Tezcan, and Hardy 2002), are likely to continue to be forthcoming.

An interesting question that arises with respect to force fields is the degree to which they can be used to study reactive processes, i.e., processes whereby one minimum-energy compound is converted into another with the intermediacy of some transition state. As noted at the beginning of this chapter, one of the first applications of force-field methodology was to study the racemization of substituted biphenyls. And, for such 'conformational reactions', there seems to be no reason to believe force fields would not be perfectly appropriate modeling tools. Unless the conformational change in question were to involve an enormous amount of strain in the TS structure, there is little reason to believe that any of the internal coordinates would be so significantly displaced from their equilibrium values that the force-field functional forms would no longer be accurate.

However, when it comes to reactions where bonds are being made and/or broken, it is clear that, at least for the vast majority of force fields that use polynomial expressions for the bond stretching energy, the 'normal' model is inapplicable. Nevertheless, substantial application of molecular mechanics to such TS structures has been reported, with essentially three different approaches having been adopted.

One approach, when sufficient data are available, is to define new atom types and associated parameters for those atoms involved in the bond-making/bond-breaking coordinate(s). This is rather tricky since, while there may be solid experimental data for activation energies, there are unlikely to be any TS structural data. Instead, one might choose to use structures computed from some QM level of theory for one or more members of the molecular data set. Then, if one assumes the reaction coordinate is highly transferable from one molecule to the next (i.e., this methodology is necessarily restricted to the study of a single reaction amongst a reasonably closely related set of compounds), one can define a force field where TS structures are treated as 'minima' – minima in quotes because the equilibrium distances and force constants for the reactive coordinate(s) have values characteristic of the transition state.

This methodology has two chief drawbacks. A philosophical drawback is that movement along the reaction coordinate raises the force-field energy instead of lowering it, which

is opposite to the real chemical system. A practical drawback is that it tends to be data limited – one may need to define a fairly large number of parameters using only a rather limited number of activation energies and perhaps some QM data. As noted in Section 2.2.7, this creates a tension between chemical intuition and statistical rigor. Two papers applying this technique to model the acid-catalyzed lactonization of organic hydroxy-acids illustrate the competing extremes to which such optimizations may be taken (Dorigo and Houk 1987; Menger 1990).

An alternative approach is one that is valence-bond like in its formulation. A possible TS structure is one whose molecular geometry is computed to have the same energy *irrespective of whether the atomic connectivity is that of the reactant or that of the product* (Olsen and Jensen 2003). Consider the example in Figure 2.9 for a hypothetical hydride transfer from an alkoxide carbon to a carbonyl. When the C–H bond is stretched from the reactant structure, the energy of the reactant-bonded structure goes up, while the energy of the product-bonded structure goes down because that structure's C–H bond is coming closer to its equilibrium value (from which it is initially very highly displaced). The simplest way to view this process is to envision two PESs, one defined for the reactant and one for the product. These two surfaces will intersect along a 'seam', and this seam is where the energy is independent of which connectivity is employed. The TS structure is then defined as the *minimum* on the seam. This approach is only valid when the reactant and product energies are computed

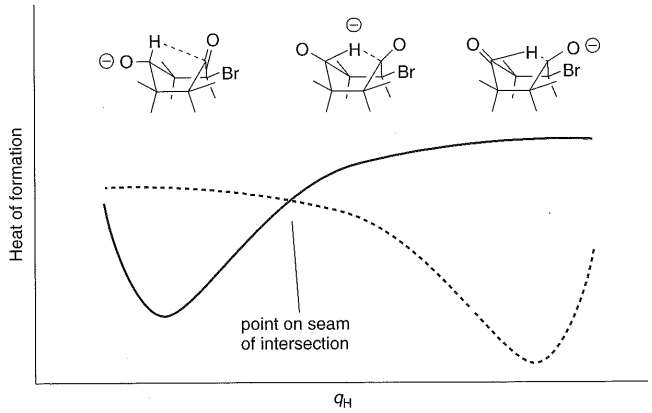

Figure 2.9 Slice through two intersecting enthalpy 'surfaces' along an arbitrary coordinate describing the location of a transferring H atom. The solid curve corresponds to bond stretching of the solid bond from carbon to the H atom being transferred. The dashed curve corresponds analogously to the dashed bond. At the point of intersection, the structure has the same energy irrespective of which bonding scheme is chosen. [For chemical clarity, the negative charge is shown shifting from one oxygen to the other, but for the method to be valid the two oxygen atom types could not change along either reaction coordinate. Note also that the bromine atom lifts the symmetry that would otherwise be present in this reaction.]

relative to a common zero (e.g., heats of formation are used; see Section 2.3), but one of its chief advantages is that it should properly reflect movement of the TS structure as a function of reaction thermicity. Because the seam of intersection involves structures having highly stretched bonds, care must be taken to use bond stretching functional forms that are accurate over larger ranges than are otherwise typical. When the VB formalism goes beyond the seam approach, and is adopted in full, a new ground-state potential energy surface can be generated about a true TS structure; such an approach is sometimes referred to as multiconfiguration molecular mechanics (MCMM) and is described in detail in Section 13.4.

The third approach to finding TS structures involves either adopting bond making/breaking functional forms that are accurate at all distances (making evaluation of bond energies a rather unpleasant N^2 process), or mixing the force-field representation of the bulk of the molecule with a QM representation of the reacting region. Mixed QM/MM models are described in detail in Chapter 13.

2.5 Menagerie of Modern Force Fields

2.5.1 Available Force Fields

Table 2.1 contains an alphabetic listing of force fields which for the most part continue to be in use today. Nomenclature of force fields can be rather puzzling because developers rarely change the name of the force field as development progresses. This is not necessarily a major issue when new development extends a force field to functionality that had not previously been addressed, but can be singularly confusing if pre-existing parameters or functional forms are changed from one version to the next without an accompanying name change. Many developers have tried to solve this problem by adding to the force field name the last two digits of the year of the most recent change to the force field. Thus, one can have MM3(92) and MM3(96), which are characterized by, *inter alia*, different hydrogen bonding parameters. Similarly, one has consistent force field (CFF) and Merck molecular force field (MMFF) versions identified by trailing year numbers. Regrettably, the year appearing in a version number does not necessarily correspond to the year in which the modifications were published in the open literature. Moreover, even when the developers themselves exercise adequate care, there is a tendency for the user community to be rather sloppy in referring to the force field, so that the literature is replete with calculations inadequately described to ensure reproducibility.

Further confusing the situation, certain existing force fields have been used as starting points for development by new teams of researchers, and the name of the resulting product has not necessarily been well distinguished from the original (which may itself be in ongoing development by its original designers!). Thus, for instance, one has the MM2* and MM3* force fields that appear in the commercial program MACROMODEL and that are based on early versions of the unstarred force fields of the same name (the * indicates the use of point charges to evaluate the electrostatics instead of bond dipoles, the use of a non-directional 10–12 potential for hydrogen bonding in place of an MM3 Buckingham potential, and a different formalism for handling conjugated systems). The commercial program Chem3D

Table 2.1 Force fields

Name (if any)	Range	Comments	Refs	Σ(error)[a]
–	Biomolecules (2nd generation includes organics)	Sometimes referred to as AMBER force fields; new versions are first coded in software of that name. All-atom (AA) and united-atom (UA) versions exist.	Original: Weiner, S. J., Kollman, P. A., Nguyen, D. T., and Case, D. A. 1986. *J. Comput. Chem.*, **7**, 230. Latest generation: Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G. M., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J. M., and Kollman, P. A. 2003. *J. Comput. Chem.*, **24**, 1999.; Ryjacek, F., Kubar, T., and Hobza, P. 2003. *J. Comput. Chem.*, **24**, 1891. See also amber.scripps.edu	
–	Organics and biomolecules	The program MACROMODEL contains many modified versions of other force fields, e.g., AMBER*, MM2*, MM3*, OPLSA*.	Mohamadi, F., Richards, N. J. G., Guida, W. C., Liskamp, R., Lipton, M., Caufield, C., Chang, G., Hendrickson, T., and Still, W. C. 1990. *J. Comput. Chem.* **11**, 440. Recent extension: Senderowitz, H. and Still, W. C. 1997. *J. Org. Chem.*, **62**, 1427. See also www.schrodinger.com	7 (AMBER*) 4 (MM2*) 5 (MM3*)
BMS	Nucleic Acids		Langley, D. R. 1998. *J. Biomol. Struct. Dyn.*, **16**, 487.	

(*continued overleaf*)

Table 2.1 (*continued*)

Name (if any)	Range	Comments	Refs	Σ(error)[a]
CHARMM	Biomolecules	Many versions of force field parameters exist, distinguished by ordinal number. All-atom and united-atom versions exist.	Original: Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. 1983. *J. Comput. Chem.*, **4**, 187; Nilsson, L. and Karplus, M. 1986. *J. Comput. Chem.*, **7**, 591. Latest generation: MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, T. F. K., Mattos, C., Michnick, S., Nago, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiórkievicz-Kuczera, J., Yin, D., and Karplus, M. 1998. *J. Phys. Chem. B*, **102**, 3586; MacKerell, A. D. and Banavali, N. 2000. *J. Comput. Chem.*, **21**, 105; Patel, S. and Brooks, C. L. 2004. *J. Comput. Chem.*, **25**, 1. See also yuri.harvard.edu	
CHARMm	Biomolecules and organics	Version of CHARMM somewhat extended and made available in Accelrys software products.	Momany, F. A. and Rone, R. 1992. *J. Comput. Chem.*, **13**, 888. See also www.accelrys.com	

Chem-X	Organics	Available in Chemical Design Ltd. software.	Davies, E. K. and Murrall, N. W. 1989. *J. Comput. Chem.*, **13**, 149.	12
CFF/CVFF	Organics and biomolecules	CVFF is the original; CFF versions are identified by trailing year digits. Bond stretching can be modeled with a Morse potential. Primarily available in Accelrys software.	CVFF: Lifson, S., Hagler, A. T., and Stockfisch, J. P. 1979. *J. Am. Chem. Soc.*, **101**, 5111, 5122, 5131. CFF: Hwang, M.-J., Stockfisch, T. P., and Hagler, A. T. 1994. *J. Am. Chem. Soc.*, **116**, 2515; Maple, J. R., Hwang, M.-J., Stockfisch, T. P., Dinur, U., Waldman, M., Ewig, C. S., and Hagler, A. T. 1994. *J. Comput. Chem.*, **15**, 162; Maple, J. R., Hwang, M.-J., Jalkanen, K. J., Stockfisch, T. P., and Hagler, A. T. 1998. *J. Comput. Chem.*, **19**, 430; Ewig, C. S., Berry, R., Dinur, U., Hill, J.-R., Hwang, M.-J., Li, C., Maple, J., Peng, Z., Stockfisch, T. P., Thacher, T. S., Yan, L., Ni, X., and Hagler, A. T. 2001. *J. Comput. Chem.*, **22**, 1782. See also www.accelrys.com	13 (CVFF) 7 (CFF91)
DREIDING	Main-group organics and inorganics	Bond stretching can be modeled with a Morse potential.	Mayo, S. L., Olafson, B. D., and Goddard, W. A., III, 1990. *J. Phys. Chem.* **94**, 8897.	10

(*continued overleaf*)

Table 2.1 (*continued*)

Name (if any)	Range	Comments	Refs	Σ(error)[a]
ECEPP	Proteins	Computes only non-bonded interactions for fixed structures. Versions identified by /(ordinal number) after name.	Original: Némethy, G., Pottle, M. S., and Scheraga, H. A. 1983. *J. Phys. Chem.*, **87**, 1883. Latest generation: Kang, Y. K., No, K. T., and Scheraga, H. A. 1996. *J. Phys. Chem.*, **100**, 15588.	
ESFF	General	Bond stretching is modeled with a Morse potential. Partial atomic charges from electronegativity equalization.	Original: Barlow, S., Rohl, A. L., Shi, S., Freeman, C. M., and O'Hare, D. 1996. *J. Am. Chem. Soc.*, **118**, 7578. Latest generation: Shi, S., Yan, L., Yang, Y., Fisher-Shaulsky, J., and Thacher, T. 2003. *J. Comput. Chem.*, **24**, 1059.	
GROMOS	Biomolecules	Coded primarily in the software having the same name.	Daura, X., Mark, A. E., and van Gunsteren, W. F. 1998. *J. Comput. Chem.*, **19**, 535.; Schuler, L. D., Daura, X., and van Gunsteren, W. F. 2001. *J. Comput. Chem.*, **22**, 1205. See also igc.ethz.ch/gromos	

MM2	Organics	Superseded by MM3 but still widely available in many modified forms.	Comprehensive: Burkert, U. and Allinger, N. L. 1982. *Molecular Mechanics*, ACS Monograph 177, American Chemical Society: Washington, DC.	5 (MM2(85), MM2(91), Chem-3D)
MM3	Organics and biomolecules	Widely available in many modified forms.	Original: Allinger, N. L., Yuh, Y. H., and Lii, J.-H. 1989. *J. Am. Chem. Soc.*, **111**, 8551. MM3(94): Allinger, N. L., Zhou, X., and Bergsma, J. 1994. *J. Mol. Struct. (Theochem)*, **312**, 69. Recent extension: Stewart, E. L., Nevins, N., Allinger, N. L., and Bowen, J. P. 1999. *J. Org. Chem.* **64**, 5350.	5 (MM3(92))
MM4	Hydrocarbons, alcohols, ethers, and carbohydrates		Allinger, N. L., Chen, K. S., and Lii, J. H. 1996. *J. Comput. Chem.*, **17**, 642; Nevins, N., Chen, K. S., and Allinger, N. L. 1996. *J. Comput. Chem.*, **17**, 669; Nevins, N., Lii, J. H., and Allinger, N. L. 1996. *J. Comput. Chem.*, **17**, 695; Nevins, N. and Allinger, N. L. 1996. *J. Comput. Chem.*, **17**, 730. Recent extension: Lii, J. H., Chen, K. H., and Allinger, N. L. 2004. *J. Phys. Chem A*, **108**, 3006.	

(*continued overleaf*)

Table 2.1 (*continued*)

Name (if any)	Range	Comments	Refs	Σ(error)[a]
MMFF	Organics and biomolecules	Widely available in relatively stable form.	Halgren, T. A. 1996. *J. Comput. Chem.*, **17**, 490, 520, 553, 616; Halgren, T. A., and Nachbar, R. B. 1996. *J. Comput. Chem.*, **17**, 587. See also www.schrodinger.com	4 (MMFF93)
MMX	Organics, biomolecules, and inorganics	Based on MM2.	See www.serenasoft.com	5
MOMEC	Transition metal compounds		Original: Bernhardt, P. V. and Comba, P. 1992. *Inorg. Chem.*, **31**, 2638. Latest generation: Comba, P. and Gyr, T. 1999. *Eur. J. Inorg. Chem.*, 1787 See also www.uni-heidelberg.de/institute/fak12/AC/comba/molmod_momec.html	
OPLS	Biomolecules, some organics	Organic parameters are primarily for solvents. All-atom and united-atom versions exist.	Proteins: Jorgensen, W. L., and Tirado-Rives, J. 1988. *J. Am. Chem. Soc.*, **110**, 1657; Kaminski, G. A., Friesner, R. A., Tirado-Rives, J., and Jorgensen, W. L. 2001. *J. Phys. Chem. B*, **105**, 6474.	

			Nucleic acids: Pranata, J., Wierschke, S. G., and Jorgensen, W. L. 1991. *J. Phys. Chem. B*, **113**, 2810. Sugars: Damm, W., Frontera, A., Tirado-Rives, J., and Jorgensen, W. L. 1997. *J. Comput. Chem.*, **18**, 1955. Recent extensions: Rizzo, R. C., Jorgensen, W. L. 1999. *J. Am. Chem. Soc.*, **121**, 4827. Carbohydrates: Kony, D., Damm, W., Stoll, S., and van Gunsteren, W. F. 2002. *J. Comput. Chem.*, **2**, 1416.	
PEF95SAC	Carbohydrates	Based on CFF form.	Fabricius, J., Engelsen, S. B., and Rasmussen, K. 1997. *J. Carbohydr. Chem.*, **16**, 751.	
PFF	Proteins	Polarizable electrostatics	Kaminski, G. A., Stern, H. A., Berne, B. J., Friesner, R. A., Cao, Y. X., Murphy, R. B., Zhou, R., and Halgren, T. A. 2002. *J. Comput. Chem.*, **23**, 1515.	

(*continued overleaf*)

Table 2.1 (*continued*)

Name (if any)	Range	Comments	Refs	Σ(error)[a]
SHAPES	Transition metal compounds		Allured, V. S., Kelly, C., and Landis, C. R. 1991. *J. Am. Chem. Soc.*, **113**, 1.	
SYBYL/Tripos	Organics and proteins	Available in Tripos and some other software.	Clark, M., Cramer, R. D., III, and van Opdenbosch, N. 1989. *J. Comput. Chem.*, **10**, 982. See also www.tripos.com and www.scivision.com	8–12
TraPPE	Organic	Primarily for computing liquid/vapor/supercritical fluid phase equilibria	Original: Martin, M. G. and Siepmann, J. I. 1998. *J. Phys. Chem. B*, **102**, 2569. Latest Generation: Chen, B., Potoff, J. J., and Siepmann, J. I. 2001. *J. Phys. Chem. B*, **105**, 3093.	
UFF	General	Bond stretching can be modeled with a Morse potential.	Rappé, A. K., Casewit, C. J., Colwell, K. S., Goddard, W. A., III, and Skiff, W. M. 1992. *J. Am. Chem. Soc.*, **114**, 10024, 10035, 10046.	21
VALBOND	Transition metal compounds	Atomic-orbital-dependent energy expressions.	Root, D. M., Landis, C. R., and Cleveland, T. 1993. *J. Am. Chem. Soc.*, **115**, 4201.	

[a] Kcal mol^{-1}. From Gundertofte *et al.* (1991, 1996); see text.

also has force fields based on MM2 and MM3, and makes no modification to the names of the originals.

As a final point of ambiguity, some force fields have not been given names, *per se*, but have come to be called by the names of the software packages in which they first became widely available. Thus, the force fields developed by the Kollman group (see Table 2.1) have tended to be referred to generically as AMBER force fields, because this software package is where they were originally coded. Kollman preferred that they be referred to by the names of the authors on the relevant paper describing their development, e.g., 'the force field of Cornell *et al.*' This is certainly more informative, since at this point the AMBER program includes within it *many* different force fields, so reference to the 'AMBER force field' conveys no information.

Because of the above ambiguities, and because it is scientifically unacceptable to publish data without an adequate description of how independent researchers might reproduce those data, many respected journals in the chemistry field now have requirements that papers reporting force-field calculations include as supplementary material a complete listing of all force field parameters (and functional forms, if they too cannot be adequately described otherwise) required to carry out the calculations described. This also facilitates the dissemination of information to those researchers wishing to develop their own codes for specific purposes.

Table 2.1 also includes a general description of the chemical space over which the force field has been designed to be effective; in cases where multiple subspaces are addressed, the order roughly reflects the priority given to these spaces during development. Force fields which have undergone many years worth of refinements tend to have generated a rather large number of publications, and the table does not try to be exhaustive, but effort is made to provide key references. The table also includes comments deemed to be particularly pertinent with respect to software implementing the force fields. For an exhaustive listing, by force field, of individual papers in which parameters for specific functional groups, metals, etc., were developed, readers are referred to Jalaie and Lipkowitz (2000).

2.5.2 Validation

The vast majority of potential users of molecular mechanics have two primary, related questions: 'How do I pick the best force field for my problem?' and, 'How will I know whether I can trust the results?' The process of testing the utility of a force field for molecules other than those over which it was parameterized is known as 'validation'.

The answer to the first question is obvious, if not necessarily trivial: one should pick the force field that has previously been shown to be most effective for the most closely related problem one can find. That demonstration of effectiveness may have taken place within the process of parameterization (i.e., if one is interested in conformational properties of proteins, one is more likely to be successful with a force field specifically parameterized to model proteins than with one which has not been) or by post-development validation. Periodically in the literature, papers appear comparing a wide variety of force fields for some well-defined problem, and the results can be quite useful in guiding the choices of subsequent