FEATURE ARTICLE

Understanding Modern Molecular Dynamics: Techniques and Applications

Mark E. Tuckerman*

Department of Chemistry and Courant Institute of Mathematical Sciences, New York University, New York, New York 10003

Glenn J. Martyna

Department of Chemistry, Indiana University, Bloomington, Indiana 47405

Received: July 16, 1999; In Final Form: October 12, 1999

Downloaded via BROWN UNIV on June 26, 2025 at 03:53:27 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

Recent advances in molecular dynamics methodology have made it possible to study routinely the microscopic details of chemical processes in the condensed phase using high-speed computers. Thus, it is timely and useful to provide a pedagogical treatment of the theoretical and numerical aspects of modern molecular dynamics simulation techniques and to show several applications that illustrate the capability of these approaches. First, the standard Newtonian or Hamiltonian dynamics based method is presented followed by a discussion of theoretical advances related to non-Hamiltonian molecular dynamics. Examples of non-Hamiltonian molecular dynamics schemes capable of generating the canonical and isothermal−isobaric ensemble are analyzed. Next, the novel Liouville operator factorization approach to numerical integration is reviewed. The power and utility of this new technique are contrasted to more basic methods, particularly, in the development of multiple time scale and non-Hamiltonian integrators. Since the results of molecular dynamics simulations depend on the interparticle interactions employed in the calculations, modern empirical force fields and ab initio molecular dynamics approaches are discussed. An example calculation combining an empirical force field and novel molecular dynamics methods, the mutant T4 lysozyme M61 in water, will be presented. The combination of electronic structure with classical dynamics, the so called ab initio molecular dynamics method, will be described and an application to the structure of liquid ammonia discussed. Last, it will then be shown how the classical molecular dynamics methods can be adapted for quantum calculations using the Feynman path integral formulation of statistical mechanics. An application, employing both path integrals and ab initio molecular dynamics, to an excess proton in water will be presented.

1. Introduction

Molecular dynamics (MD) has had a long history[1−6] and has evolved into an important and widely used theoretical tool that allows researchers in chemistry, physics, and biology to model the detailed microscopic dynamical behavior of many different types of systems, including gases, liquids, solids, surfaces, and clusters. However, recent theoretical and numerical advances have increased the utility of the basic methodology and made a pedagogical overview for a nonspecialist audience timely and appropriate.

In a MD simulation, the classical equations of motion governing the microscopic time evolution of a many-body system are solved numerically subject to boundary conditions appropriate for the geometry or symmetry of the system. Thus, MD methodology is founded upon the basic principles of classical mechanics and can provide a window into the microscopic dynamical behavior of the individual atoms that make up a given system. From this information, the microscopic mechanisms of energy and mass transfer in chemical processes can be "observed" and dynamical properties such as absorption spectra, rate constants, and transport properties can be calculated. In addition to providing a microscopic dynamical picture, MD can also be employed as a means of sampling from a statistical mechanical ensemble and determining equilibrium properties. These properties include average thermodynamic quantities (pressure, temperature, volume, etc.), structure, and free energies along reaction paths.

In order to provide a picture of the microscopic behavior of a system from the laws of classical mechanics, MD requires, as an input, a description of the interparticle interactions. The quality of the results of an MD simulation depends on the accuracy of this description. One common approach involves the introduction of a model or *force field*. The relatively low computational overhead associated with a standard force-field has allowed large-scale calculations to be performed on proteins, membranes, and large biological assemblies. Recent applications employing common force fields include, among many other impressive examples that can be found in the literature, an exploration of protein folding pathways in solution,[7] structural and dynamical properties of ion channels,[8,9] a novel approach to conformational sampling applied to the mutant T4 lysozyme M6I (see section 7.2.1), and studies of the human immunodeficiency virus protease complexed with a new class of C_{60} fullerene-based inhibitor compounds.[10]

10.1021/jp992433y CCC: $19.00 © 2000 American Chemical Society
Published on Web 12/16/1999

The disadvantage of a model force-field is that a system is restricted to a single molecular connectivity. This prohibits force field models from describing chemical processes involving bond breaking and forming. An alternative approach is the combination of classical dynamics with electronic structure, which allows internuclear forces to be computed "on the fly" from an electronic structure calculation as a MD simulation proceeds.[11,12] This method, known as ab initio molecular dynamics, requires no input potential model and is capable of describing chemical events, although it has high computational overhead. Nonetheless, the ab initio MD approach has been successfully applied to study hydrogen-bonded liquids, such as water[13,14] and ammonia,[15] proton transfer in water,[16−19] ice,[20] and in a model ion channel,[21] Ziegler−Natta catalysis,[22] and the behavior of water on an alumina surface,[23] among many other systems.

2. Overview

In this article, the basic principles that underly modern MD methodology are described and a clear picture of why this approach is stable and accurate is developed for the nonspecialist. In order to develop a modern picture, the basic principles of Hamiltonian mechanics are reviewed and the microcanonical statistical thermodynamics generated by the dynamics discussed. Next, the use of non-Hamiltonian dynamical systems in molecular dynamics will be discussed and the theoretical statistical basis[24] underlying their use will be presented. Specific sets of non-Hamiltonian equations of motion that lead to the canonical and isothermal−isobaric ensembles will then presented in the context of this theoretical framework. The evolution operator formalism of classical mechanics is then used to derive numerical multiple time step integrators for both Hamiltonian (Newtonian) and non-Hamiltonian equations of motion whose long time stability can be clearly understood. However, the ability of MD calculations to describe physical systems depends on the accuracy of the interparticle interaction potentials. Both empirical force fields and ab initio techniques are discussed and their ability to treat large molecular systems and chemical reactions, respectively, evaluated. Example calculations based on state-of-the-art force fields and electronic structure techniques are presented. Finally, the problem of treating light nuclei in MD simulations will be discussed. In many examples involving light nuclei, the approximate validity of a classical description breaks down. Recently developed Feynman path integral molecular dynamics techniques, which allow this approximation to be relaxed, are discussed and example calculations presented. Although of considerable current interest, the topic of multiple electronic surfaces will be not be treated here.

3. The Basic Approach: Hamiltonian Mechanics

Molecular dynamics (MD) is employed to study the classical motion of a many-body system and extract from the dynamics the experimental observables. As MD calculations provide a window into the detailed motion of individual atoms in a system, the microscopic mechanisms of energy and mass transfer can be gleaned.

Consider a system consisting of N particles moving under the influence of the internal forces acting between them. The spatial positions of the particles as functions of time will be denoted by $\mathbf{r}_1(t)$, ..., $\mathbf{r}_N(t)$, and their velocities, $\mathbf{v}_1(t)$, ..., $\mathbf{v}_N(t)$. If the forces, $\mathbf{F}_1$, ..., $\mathbf{F}_N$, on the N particles are specified, then the classical motion of the system is determined by Newton's second law

$$m_i\ddot{\mathbf{r}}_i = \mathbf{F}_i \tag{3.1}$$

where m_1, ..., m_N are the masses of the N particles. Since the force on each particle is, in principle, a function of all of the N position variables, $\mathbf{F}_i = \mathbf{F}_i(\mathbf{r}_1, ..., \mathbf{r}_N)$, eqs 3.1 constitute a set of $3N$, or more generally, dN, where d is the number of spatial dimensions, coupled second-order differential equations. A unique solution to eqs 3.1 is obtained by choosing a set of initial conditions, $\{\mathbf{r}_1(0), ..., \mathbf{r}_N(0), \mathbf{v}_1(0), ..., \mathbf{v}_N(0)\}$. Newton's equations completely determine the full set of positions and velocities as functions of time and thus specify the *classical state* of the system at time t. Except in special cases, an analytical solution to the equations of motion, eqs 3.1, is not possible. An MD calculation, therefore, employs an iterative numerical procedure, called a *numerical integrator* or a *map*, to obtain an approximate solution.[6,25] The accuracy of the numerical solution is determined by the time discretization, Δt, referred to as the *time step*. In most cases, the forces, $\mathbf{F}_i(\mathbf{r}_1, ..., \mathbf{r}_N)$, are sufficiently nonlinear functions of position that, if the true solution could be obtained for a given choice of initial conditions, the numerical solution would bear little resemblance to it after enough iterations of the map. This is largely due to the fact that the initial conditions can only be specified to within a finite precision for numerical calculation. In a large system with highly nonlinear forces, small differences between two sets of initial conditions lead to a divergence between the trajectories that become exponentially large as time increases. However, the numerical solution is statistically equivalent to the true solution within a *bounded* error, and this is sufficient to ensure that the same physical observables are obtained on average. It is important to note that small systems with closed orbits possesses other such statistical equivalences.

In order to demonstrate the conditions required for the statistical equivalence of the numerical and true solutions to the equations of motion, it is first useful to recast eqs 3.1 in Hamiltonian form. The Hamiltonian for an N-particle system subject only to interparticle interactions is

$$H(\mathbf{p},\mathbf{r}) \equiv H(\mathbf{p}_1, ..., \mathbf{p}_N, \mathbf{r}_1, ..., \mathbf{r}_N) = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} + U(\mathbf{r}_1, ..., \mathbf{r}_N) \tag{3.2}$$

where $\mathbf{p}_1$, ..., $\mathbf{p}_N$ are the momenta of the particles defined by $\mathbf{p}_i = m_i\mathbf{v}_i$ and $U(\mathbf{r}_1, ..., \mathbf{r}_N)$ is the interparticle potential, in terms of which the forces are given by

$$\mathbf{F}_i = -\frac{\partial U}{\partial \mathbf{r}_i} \tag{3.3}$$

The equations of motion (3.1) can be derived from eq 3.2 according to Hamilton's equations,

$$\dot{\mathbf{r}}_i = \frac{\partial H}{\partial \mathbf{p}_i} = \frac{\mathbf{p}_i}{m_i}$$

$$\dot{\mathbf{p}}_i = -\frac{\partial H}{\partial \mathbf{r}_i} = -\frac{\partial U}{\partial \mathbf{r}_i} = \mathbf{F}_i(\mathbf{r}_1, ..., \mathbf{r}_N) \tag{3.4}$$

Taking the time derivative of both sides of the first of Hamilton's equations and substituting into the second, one easily arrives at eqs 3.1. Therefore, the classical state of a system at any instant in time can also be determined by specifying the complete set of particle positions and corresponding momenta. Alternatively, we may collect the full set of positions and momenta into a single vector $\mathbf{x} = (\mathbf{p}_1, ..., \mathbf{p}_N, \mathbf{r}_1, ..., \mathbf{r}_N)$ called the *phase space*

vector, which exists in a $2dN$-dimensional *phase space*. A classical state of the system corresponds to a single point in the phase space. The phase space is thus the union of *all possible* classical states of a system.

Two important properties of the equations of motion should be noted. One is that they are time reversible, i.e., they take the same form when the transformation $t \rightarrow -t$ is made. The consequence of *time reversal symmetry* is that the microscopic physics is independent of the direction of the flow of time. The second important property of the equations of motion is that they conserve the Hamiltonian eq 3.2. This can be easily seen by computing the time derivative of H and substituting eqs 3.4 for the time derivatives of position and momentum:

$$\frac{dH}{dt} = \sum_{i=1}^{N}\left[\frac{\partial H}{\partial \mathbf{r}_i}\dot{\mathbf{r}}_i + \frac{\partial H}{\partial \mathbf{p}_i}\dot{\mathbf{p}}_i\right] = \sum_{i=1}^{N}\left[\frac{\partial H}{\partial \mathbf{r}_i}\frac{\partial H}{\partial \mathbf{p}_i} - \frac{\partial H}{\partial \mathbf{p}_i}\frac{\partial H}{\partial \mathbf{r}_i}\right] = 0 \quad (3.5)$$

The conservation of the Hamiltonian is equivalent to the conservation of the total energy of the system and provides an important link between molecular dynamics and statistical mechanics. Recall that the latter connects the microscopic details of a system to physical observables such as equilibrium thermodynamic properties, transport coefficients, and spectra. Statistical mechanics is based on the *Gibbs' ensemble* concept. That is, many individual microscopic configurations of a very large system lead to the same macroscopic properties, implying that it is not necessary to know the precise detailed motion of every particle in a system in order to predict its properties. It is sufficient to simply average over a large number of identical systems, each in a different such microscopic configuration; i.e., the macroscopic observables of a system are formulated in terms of *ensemble averages*. Statistical ensembles are usually characterized by fixed values of thermodynamic variables such as energy, E; temperature, T; pressure, P; volume, V; particle number, N; or chemical potential, μ. One fundamental ensemble is called the *microcanonical* ensemble and is characterized by constant particle number, N; constant volume, V; and constant total energy, E, and is denoted as the *NVE* ensemble. Other examples include the canonical or *NVT* ensemble, the isothermal−isobaric or *NPT* ensemble, and the grand canonical or μVT ensemble. The thermodynamic variables that characterize an ensemble can be regarded as experimental control parameters that specify the conditions under which an experiment is performed.

Now consider a system of N particles occupying a container of volume V and evolving under Hamilton's equations of motion. According to eq 3.5, the Hamiltonian will be a constant, E, equal to the total energy of the system. In addition, the number of particles and the volume are assumed to be fixed. Therefore, a dynamical trajectory of this system will generate a series of classical states having constant N, V, and E, corresponding to a microcanonical ensemble. If the dynamics generates all possible states having a fixed N, V, and E, then an average over this trajectory will yield the same result as an average in a microcanonical ensemble. The energy conservation condition, $H(\mathbf{p}, \mathbf{r}) = E$, which imposes a restriction on the classical microscopic states accessible to the system, defines a hypersurface in the phase space called the *constant energy surface*. A system evolving according to Hamilton's equations of motion will remain on this surface. The assumption that a system, given an infinite amount of time, will cover the entire constant energy hypersurface is known as the *ergodic hypothesis*. Thus, under the ergodic hypothesis, averages over a trajectory of a system obeying Hamilton's equations are equivalent to averages over the microcanonical ensemble.

In mathematical terms, if $A(\mathbf{p}, \mathbf{r})$ is a function corresponding to a physical observable, then the microcanonical ensemble average of A is

$$\langle A\rangle = \frac{C_N}{h^{3N}\Omega(N,V,E)}\int d^N\mathbf{p}\int_{D(V)} d^N\mathbf{r}\, A(\mathbf{p},\mathbf{r})\, \delta(H(\mathbf{p},\mathbf{r}) - E) \quad (3.6)$$

where $\Omega(N,V,E)$ is the microcanonical partition function given by

$$\Omega(N,V,E) = \frac{C_N}{h^{3N}}\int d^N\mathbf{p}\int_{D(V)} d^N\mathbf{r}\,\delta(H(\mathbf{p},\mathbf{r}) - E) \quad (3.7)$$

Here, h is Planck's constant and C_N is a general combinatorial factor. As the prefactor C_N/h^{3N} does not affect the analyses presented herein, it will be omitted from expressions subsequently presented in this paper. Equation 3.7 is a device for "counting" the number of microscopic states of a system that obey the condition $H(\mathbf{p},\mathbf{r}) = E$ for a given number of particles N and container volume V. The integral over the N Cartesian positions is restricted by the spatial domain $D(V)$ defined by the walls of the container, while the momentum integral is unrestricted. The average of an observable, A, over a trajectory spanning a length of time, $\mathcal{T}$, is given by

$$\bar{A}_{\mathcal{T}} = \frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} dt\, A(\mathbf{p}(t),\mathbf{r}(t)) \quad (3.8)$$

for a trajectory starting at $t = 0$. The ergodic hypothesis is equivalent to the statement

$$\lim_{\mathcal{T}\rightarrow\infty} \bar{A}_{\mathcal{T}} = \langle A\rangle \quad (3.9)$$

(Note, the system need not be mixing or even chaotic in nature to obey eq 3.9. A one-dimensional harmonic oscillator is ergodic and samples all the phase space available to it!)

The meaning of the statistical equivalence between a numerical trajectory and the true trajectory of a system is now clear. Although a numerical trajectory may diverge in time from the true trajectory, as long as the numerical trajectory conserves the energy to within a given tolerance, ΔE, the numerical trajectory will also generate configurations belonging to the constant energy surface that are never in error by more than ΔE. (The existence of bounds on the error of numerical trajectories is discussed further in section 6.) Assuming ergodicity, a single numerical trajectory can also be used in eq 3.9 to compute the ensemble average of an observable. Note, this is equally true for a regular system with closed orbits and a chaotic or mixing system.

Finally, it should be noted that dynamical properties are also defined through ensemble averages. Time correlation functions are important because of their relation to transport coefficients and spectra via linear response theory.[26,27] Consider, for example, a time correlation function, $C_{AB}(t)$, between two observables, $A(\mathbf{p},\mathbf{r})$ and $B(\mathbf{p},\mathbf{r})$. In order to calculate $C_{AB}(t)$, one can use a set of trajectories generated by Hamilton's equations. Any trajectory is uniquely determined by its initial conditions. Suppose initial conditions for each trajectory in the set are sampled from an equilibrium phase space distribution function $f(\mathbf{p},\mathbf{r})$. The time correlation function is then defined to be

$$C_{AB}(t) = \frac{\int d^N\mathbf{p}\, d^N\mathbf{r}\, f(\mathbf{p},\mathbf{r})\, A(\mathbf{p},\mathbf{r})\, B(\mathbf{p}(t),\mathbf{r}(t))}{\int d^N\mathbf{p}\, d^N\mathbf{r}\, f(\mathbf{p},\mathbf{r})} \quad (3.10)$$

Thus, it can be seen that a time correlation function can be calculated by evolving a trajectory in time starting from each set of initial conditions and then averaging the product $A(\mathbf{p},\mathbf{r})B(\mathbf{p}(t),\mathbf{r}(t))$ over the set of trajectories at each instant in time. (In the microcanonical ensemble, $f(\mathbf{p},\mathbf{r}) = \delta(H(\mathbf{p},\mathbf{r}) - E)$.) In the thermodynamic limit, all equilibrium ensembles are equivalent, and thus, for very large systems, a single long trajectory can be used to generate a time correlation function, although the convergence of such an approach may be slow. For a detailed treatment of the properties of time correlation functions, the reader is referred to the review by Berne and Harp.[27]

Despite the utility of Hamiltonian molecular dynamics, its principle restriction is clear: although, given correct forces, the dynamics is exact in the classical limit; it can only generate equilibrium properties of the *NVE* ensemble. However, microcanonical conditions (*NVE*) are not consistent with the many experimental measurements under conditions of constant temperature and pressure or constant temperature and volume. In order to describe the thermodynamic properties of a system under these conditions, it is necessary to generate the corresponding ensemble. One of the more fruitful and interesting approaches to generating alternative ensemble averages is based on properties of non-Hamiltonian dynamical systems.

4. Principles of Non-Hamiltonian Statistical Mechanics

In the previous section, the concept of an ensemble was introduced, and the specific example of the microcanonical ensemble, as the ensemble of systems having $H(\mathbf{p},\mathbf{r}) = E$, was discussed. In general, an ensemble is defined by its phase space distribution function $f(\mathbf{p},\mathbf{r},t) \equiv f(\mathbf{x},t)$, which may possibly depend explicitly on time. The phase space distribution function must satisfy the *Liouville equation*, which for systems governed by Hamiltonian dynamics is

$$\frac{\mathrm{d}f(\mathbf{x},t)}{\mathrm{d}t} = \frac{\partial f(\mathbf{x},t)}{\partial t} + \dot{\mathbf{x}}\cdot\nabla f(\mathbf{x},t) = 0 \tag{4.1}$$

where ∇ is the $2dN$-dimensional gradient on the phase space. The Liouville equation results from the requirement that the rate of change of the number of ensemble members in an arbitrary phase space volume is equal to the flux of members through the boundary of the volume. It can be seen that the Liouville equation is a statement of the conservation of f. Since f is a probability distribution function, the existence of a conservation law for f implies the existence of a conserved phase space measure, $\mathrm{d}\mu$, For Hamiltonian systems the invariant measure is, $\mathrm{d}\mu = \mathrm{d}\mathbf{x} \equiv \mathrm{d}^N\mathbf{p}\,\mathrm{d}^N\mathbf{r}$. It will shortly be shown below why this is an invariant measure for Hamiltonian systems. In summary, given an ensemble distribution function satisfying eq 4.1, the average of any observable $A(\mathbf{x})$ can be defined by

$$\langle A\rangle_t = \frac{\int \mathrm{d}\mathbf{x}\, f(\mathbf{x},t)\, A(\mathbf{x})}{\int \mathrm{d}\mathbf{x}\, f(\mathbf{x},t)} \tag{4.2}$$

Non-Hamiltonian equations of motion are typically used to generate ensembles other than the microcanonical, for describing systems subject to nonholonomic constraints or for describing driven systems. The idea of generating ensembles dynamically began with the work of Andersen,[28] who showed that by extending the phase space beyond the $2dN$ dimensions of the physical system, a dynamical scheme could be constructed to generate an isobaric distribution of the physical subsystem. Isothermal extensions followed.[29,30] These original formulations of extended phase space dynamics were based on Hamiltonian systems, which possess certain undersirable features related to the definition of time. It was shown that these could be corrected by going over to a non-Hamiltonian formulation.[31] However, only recently has a consistent theoretical statistical framework underlying the use of general non-Hamiltonian systems been presented.[24] The theoretical underpinnings will be discussed below.

Consider the dynamical system

$$\dot{\mathbf{x}} = \boldsymbol{\xi}(\mathbf{x},t) \tag{4.3}$$

which is assumed to be non-Hamiltonian (i.e., expressible in the form of eqs 3.4). Here, $\boldsymbol{\xi}(\mathbf{x},t)$ is a generalized force, which may have an explicit time dependence. If the dynamical system is not Hamiltonian, then its *phase space compressibility*, defined to be

$$\kappa(\mathbf{x},t) = \nabla\cdot\mathbf{x} = \nabla\cdot\boldsymbol{\xi}(\mathbf{x},t) \tag{4.4}$$

which vanishes for a Hamiltonian system (the incompressibility property), will generally be nonzero and the phase space measure $\mathrm{d}\mathbf{x}$ is no longer invariant. In order to see this, one need only consider the Jacobian of the transformation from an initial phase space vector $\mathbf{x}_0$ a time-evolved vector $\mathbf{x}_t$ given by

$$J(\mathbf{x}_t;\mathbf{x}_0) = \frac{\partial(x_t^{\,1},\ldots,x_t^{\,n})}{\partial(x_0^{\,1},\ldots,x_0^{\,n})} \tag{4.5}$$

where n is the dimension of the phase space. It can be shown (see, e.g. ref 24) that $J(\mathbf{x}_t;\mathbf{x}_0)$ satisfies the following evolution equation:

$$\frac{\mathrm{d}}{\mathrm{d}t} J(\mathbf{x}_t,\mathbf{x}_0) = \kappa(\mathbf{x}_t,t)\, J(\mathbf{x}_t;\mathbf{x}_0) \tag{4.6}$$

with initial condition $J(\mathbf{x}_0;\mathbf{x}_0) = 1$. Equation 4.6 implies that J will only be 1 for all time if $\kappa = 0$, as it is for Hamiltonian systems. For non-Hamiltonian systems, the measure transforms according to

$$\mathrm{d}\mathbf{x}_t = J(\mathbf{x}_t;\mathbf{x}_0)\,\mathrm{d}\mathbf{x}_0 \tag{4.7}$$

which demonstrates that if $J \neq 1$, $\mathrm{d}\mathbf{x}_t \neq \mathrm{d}\mathbf{x}_0$.

A complete statistical theory of non-Hamiltonian systems has recently been presented in ref 24. The basic tenets of the theory are as follows:

(1) There is an invariant measure that takes the form

$$\mathrm{d}\mu = \sqrt{g(\mathbf{x},t)}\,\mathrm{d}\mathbf{x} \tag{4.8}$$

where the metric factor $\sqrt{g(\mathbf{x},t)}$ is given by

$$\sqrt{g(\mathbf{x},t)} = \mathrm{e}^{-w(\mathbf{x},t)} \tag{4.9}$$

and the function $w(\mathbf{x},t)$ is related to the compressibility $\kappa(x,t)$ by

$$\frac{\mathrm{d}w}{\mathrm{d}t} = \kappa \tag{4.10}$$

(2) There is a non-Hamiltonian generalization of the Liouville equation for a general ensemble distribution function $f(x,t)$, which takes the form

$$\frac{\partial}{\partial t}(\sqrt{g}f) + \nabla\cdot(\dot{\mathbf{x}}\sqrt{g}f) = 0 \tag{4.11}$$

(3) In the absence of external driving and explicitly time-dependent forces, a non-Hamiltonian microcanonical ensemble can be defined on the basis of the existence of the invariant measure. If the dynamical system in eq 4.3 has a set of M conserved quantities $K_\lambda(\mathbf{x})$, $\lambda = 1, ..., M$, satisfying $dK_\lambda/dt = 0$, then the microcanonical distribution function is given by

$$f(\mathbf{x}) = \prod_{\lambda=1}^{M} \delta(K_\lambda(\mathbf{x}) - \bar{K}_\lambda) \tag{4.12}$$

and the corresponding partition function is

$$\Omega(N,V,\bar{K}_1, ..., \bar{K}_M) = \int d\mathbf{x}\sqrt{g(\mathbf{x})} \prod_{\lambda=1}^{M} \delta(K_\lambda(\mathbf{x}) - \bar{K}_\lambda) \tag{4.13}$$

The microcanonical distribution, $f(\mathbf{x})$, together with the metric $\sqrt{g(\mathbf{x})}$ satisfies the time-independent form of eq 4.11. This is a necessary but not sufficient condition to guarantee that a given $f(\mathbf{x})$ is the correct equilibrium distribution function. For example, a distribution function $\bar{f}(\mathbf{x}) = \prod_{\lambda=1}^{M'}\delta(K_\lambda(\mathbf{x}) - \bar{K}_\lambda)$, where $M' < M$ (i.e., a distribution constructed from a subset of the conservation laws satisfied by a system), also satisfies the generalized Liouville equation. However, *all* conservation laws are *required* if the correct microcanonical distribution is to be constructed.

5. Designing Non-Hamiltonian Equations of Motion

It will now be shown how the basic principles of non-Hamiltonian statistical mechanics can be employed to design MD equations of motion that generate the canonical and isothermal−isobaric ensembles. In fact, the above principles lead to a procedure for designing general-purpose MD algorithms.[32]

The Nosé−Hoover chain (NHC) dynamics method[33] is a non-Hamiltonian MD scheme for generating the canonical ensemble. In this method, the ordinary phase space is extended to include a set of M thermostat variables $\eta_1, ..., \eta_M$ and their conjugate momenta $p_{\eta_1}, ..., p_{\eta_M}$, which act as a heat bath coupled to the system. The equations of motion take the form

$$\begin{aligned}
\dot{\mathbf{r}}_i &= \frac{\mathbf{p}_i}{m_i} \\
\dot{\mathbf{p}}_i &= \mathbf{F}_i - \frac{p_{\eta_1}}{Q_1}\mathbf{p}_i \\
\dot{\eta}_k &= \frac{p_{\eta_k}}{Q_k} \qquad k = 1, ..., M \\
\dot{p}_{\eta_k} &= G_k - \frac{p_{\eta_{k+1}}}{Q_{k+1}}p_{\eta_k} \qquad k = 1, ..., M-1 \\
\dot{p}_{\eta_M} &= G_M
\end{aligned} \tag{5.1}$$

where the thermostat forces are

$$\begin{aligned}
G_1 &= \sum_{i=1}^{N} \frac{\mathbf{p}_i^{\,2}}{m_i} - dNkT \\
G_k &= \frac{p_{\eta_{k-1}}^{\,2}}{Q_{k-1}} - kT \qquad k = 2, ..., M
\end{aligned} \tag{5.2}$$

The parameters Q_k, given by $Q_1 = dNkT\tau^2$ and $Q_k = kT\tau^2$, determine the time scale of the thermostat motion via the single time scale τ, which should be chosen corresponding to a characteristic time scale of the system, e.g., a vibrational period; In eqs 5.1, M thermostats, which successively thermostat each other, are coupled to the particles, thereby controlling/modulating the kinetic energy fluctuations of both particle and thermostat degrees of freedom.

Let us apply principles 1 and 2 above to the NHC equations. First, the compressibility and the invariant measure are computed. The compressibility can be seen to be

$$\begin{aligned}
\kappa &= \sum_{i=1}^{N}[\nabla_{\mathbf{r}_i}\cdot\dot{\mathbf{r}}_i + \nabla_{\mathbf{p}_i}\cdot\dot{\mathbf{p}}_i] + \sum_{k=1}^{M}\left[\frac{\partial\dot{\eta}_k}{\partial\eta_k} + \frac{\partial\dot{p}_{\eta_k}}{\partial p_{\eta_k}}\right] \\
&= -dN\frac{p_{\eta_1}}{Q_1} - \sum_{k=2}^{M}\frac{p_{\eta_k}}{Q_k} \\
&= -dN\dot{\eta}_1 - \sum_{k=2}^{M}\dot{\eta}_k
\end{aligned} \tag{5.3}$$

Since $\dot{w} = \kappa$, it is clear that $w = -dN\eta_1 - \sum_{k=2}^{M}\eta_k$, and the invariant measure according to eq 4.9 is

$$d\mu = \exp\left(\eta_1 + \sum_{k=2}^{M}\eta_M\right)d^N\mathbf{p}\, d^N\mathbf{r}\, d^M\eta\, d^M p_\eta \tag{5.4}$$

In order to determine the microcanonical partition function, the conserved quantities of eqs 5.1 are needed. One of the conserved quantities is the total energy of the extended system

$$H' = H(\mathbf{p},\mathbf{r}) + \sum_{k=1}^{M}\frac{p_{\eta_k}^{\,2}}{2Q_k} + kT\left[dN\eta_1 + \sum_{k=2}^{M}\eta_k\right] \tag{5.5}$$

In addition to the energy, there are d additional conservation laws if there are no external forces (i.e., when $\sum_{i=1}^{N}\mathbf{F}_i = 0$). These conservation laws take the form

$$\mathbf{K} = e^{\eta_1}\sum_{i=1}^{N}\mathbf{p}_i \equiv e^{\eta_1}\mathbf{P} \tag{5.6}$$

as can be seen by direct differentiation. Given the conservation laws, the microcanonical partition function can now be constructed according to eq 4.13. Suppose $d = 3$ and $M = 2$,

$$\begin{aligned}
\Omega(N,V,T,E,\mathbf{K}) = \int d^N\mathbf{p}\, d^N\mathbf{r}\, d\eta_1\, d\eta_2\, dp_{\eta_1}\, dp_{\eta_2} \exp(3N\eta_1 + \\
\eta_2)\,\delta\left(H(\mathbf{p},\mathbf{r}) + \frac{p_{\eta_1}^{\,2}}{2Q_1} + \frac{p_{\eta_2}^{\,2}}{2Q_2} + 3NkT\eta_1 + kT\eta_2 - E\right) \\
\delta(e^{\eta_1}P_x - K_x)\delta(e^{\eta_1}P_y - K_y)\delta(e^{\eta_1}P_z - K_z)
\end{aligned} \tag{5.7}$$

In order to show that eq 5.7 generates a canonical distribution in the system Hamiltonian, $H(\mathbf{p},\mathbf{r})$, the integrals over the thermostat variables need to be performed. The energy δ-function can be used to perform the integral over η_2, which requires that $\eta_2 = (E - H(\mathbf{p},\mathbf{r}) - p_{\eta_1}^{\,2}/2Q_1 - p_{\eta_2}^{\,2}/2Q_2 - 3NkT\eta_1)/kT$. One of the remaining δ-functions can be used to integrate over η_1, which will leave only ratios of components of $\mathbf{P}$ in the other two δ-functions. Finally, by, changing variables to center-of-

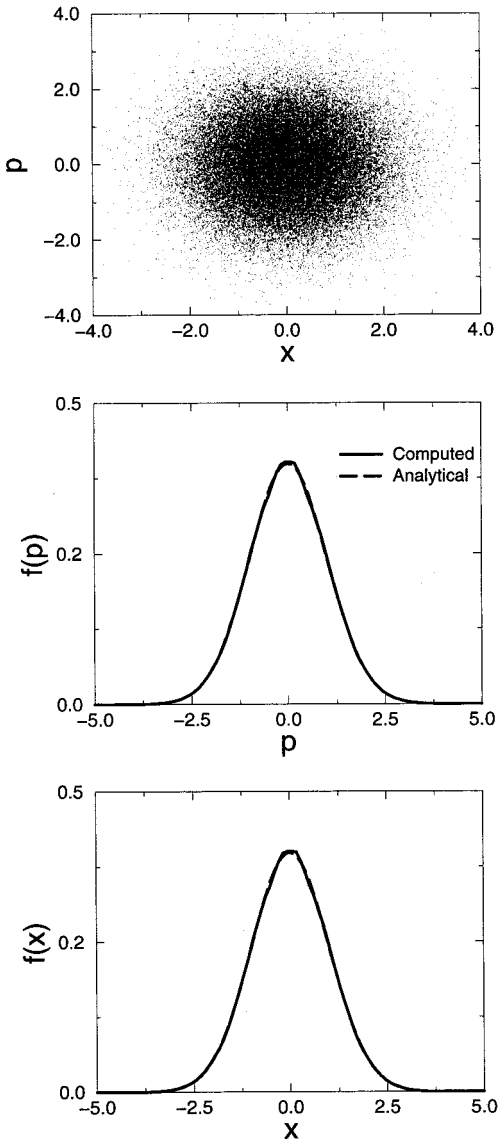

Figure 1. (a) The $x-p$ plane in the phase space of a one-dimensional harmonic oscillator coupled to a Nosé−Hoover chain (cf. eqs 5.1). The initial conditions are $x(0) = 0$, $p(0) = 1$, $\eta_k(0) = 0$, $p_{\eta_k}(0) = 1$. The mass and frequency of the oscillator are $m = 1$ and $\omega = 1$. The thermostat time scale is determined by $\tau = 1$ and $kT = 1$. The equations of motion are integrated for 10^7 steps with a time step of $\Delta t = 0.01$ using the methods in ref 40. (b) The distribution function of momentum for the oscillator in part a. The simulated distribution (solid line) is shown together with the analytical distribution (dashed line). (c) The distribution function of position for the oscillator in part a. The simulated distribution (solid line) is shown together with the analytical distribution (dashed line).

mass and normal mode momenta, ($\{\tilde{\mathbf{p}}\}$,$\mathbf{P}$), the expression can be simplified to yield

$$\Omega(N,V,E,K) \propto \int \mathrm{d}^{N-1}\tilde{\mathbf{p}}\, \mathrm{d}P\, \mathrm{d}^{N}\mathbf{r}\, P^2 \exp\left[-\beta\left(\frac{\tilde{\mathbf{p}}_i^{\,2}}{2\tilde{m}_i} + \frac{P^2}{2M} + U(\mathbf{r})\right)\right] \quad (5.8)$$

where $\{\tilde{m}, M\}$ are the normal mode and total masses, $\beta = 1/kT$, and the proportionality constant depends on E and $\mathbf{K}$. Equation 5.8 is the correct canonical partition function for the Hamiltonian $H(\mathbf{p},\mathbf{r})$. Hence, the NHC equations generate the canonical distribution function (within constants). Figure 1 shows the $x-p$ plane of phase space, and the position and momentum distribution functions for a one-dimensional harmonic oscillator generated by the NHC equations. The analytical distribution functions are also shown in the figure. As can be seen, the ensemble is correctly sampled by the NHC system. The significance of the harmonic oscillator will become clear later in the context of force fields (see section 7.2) and in path integral molecular dynamics (see section 8). Following the same procedure, it can be shown that the Nosé−Hoover thermostat method (corresponding to $M = 1$ in eqs 5.1) does not produce the correct canonical distribution when $\sum_{i=1}^{N} \mathbf{F}_i = 0$.[32] It is also interesting to note that if more than one thermostat chain is coupled to the system, then the momentum conservation laws are no longer present, and the proof simplifies. An extreme example, which has proved useful in a number of applications, is the coupling of a separate thermostat to each degree of freedom in the system. This scheme leads to very rapid equilibration of a system and plays an important role both in path integral molecular dynamics and simulations involving biological macromolecules.[34,35]

Next, the isothermal−isobaric or *NPT* ensemble is considered. In this ensemble, the volume, V, of the system must fluctuate such that the average internal, pressure of the system, $\langle P_{\mathrm{int}}\rangle$, is equal to an external applied pressure, P_{ext}. Thus, in designing a MD algorithm for the *NPT* ensemble, one should both incorporate the volume as a dynamical variable[28] and employ a thermostat to ensure that both instantaneous temperature and pressure fluctuations are generated properly. A useful non-Hamiltonian scheme for the *NPT* ensemble[36] is defined by the following equations of motion:

$$\begin{aligned}
\dot{\mathbf{r}}_i &= \frac{\mathbf{p}_i}{m_i} + \frac{p_\epsilon}{W}\mathbf{r}_i \\
\dot{\mathbf{p}}_i &= \mathbf{F}_i - \left(1 + \frac{1}{N}\right)\frac{p_\epsilon}{W}\mathbf{p}_i - \frac{p_\eta}{Q}\mathbf{p}_i \\
\dot{V} &= \frac{dVp_\epsilon}{W} \\
\dot{p}_\epsilon &= dV(P_{\mathrm{int}} - P_{\mathrm{ext}}) + \frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{p}_i^{\,2}}{m_i} - \frac{p_\eta}{Q}p_\epsilon \\
\dot{\eta} &= \frac{p_\eta}{Q} \\
\dot{p}_\eta &= \sum_{i=1}^{N}\frac{\mathbf{p}_i^{\,2}}{m_i} + \frac{p_\epsilon^{\,2}}{W} - (dN+1)kT
\end{aligned} \quad (5.9)$$

Here, p_ϵ is a momentum conjugate to the logarithm of the volume, W is its associated mass parameter, $\epsilon = \ln(V/V(0))$, P_{ext} is the external applied pressure, and P_{int} is the instantaneous internal pressure of the system given by

$$P_{\mathrm{int}} = \frac{1}{\mathrm{d}V}\left[\sum_{i=1}^{N}\frac{\mathbf{p}_i^{\,2}}{m_i} + \sum_{i=1}^{N}\mathbf{r}_i\mathbf{F}_i - (\mathrm{d}V)\frac{\partial U}{\partial V}\right] \quad (5.10)$$

Thus, the variable p_ϵ acts as a "barostat" which drives the system to the steady state $\langle P_{\mathrm{int}}\rangle = P_{\mathrm{ext}}$. For simplicity, eqs 5.9 are written with a single thermostat variable coupling both to the particles and to the barostat. Thus, the case $\sum_{i=1}^{N}\mathbf{F}_i \neq 0$ will be considered. However, it is clear that when $\sum_{i=1}^{N}\mathbf{F}_i \neq 0$, a thermostat chain should be used. In fact, the optimal algorithm employs separate thermostat chains on the particles and on the barostat.

The compressibility associated with eqs 5.9 is

$$\kappa = -(\mathrm{d}N + 1)\frac{p_\eta}{Q} = -(\mathrm{d}N + 1)\dot{\eta} \tag{5.11}$$

from which it can be seen that $w(\mathrm{x}) = -(\mathrm{d}N + 1)\eta$, and the invariant measure is

$$\mathrm{d}\mu = \mathrm{e}^{(\mathrm{d}N+1)\eta}\, \mathrm{d}^N\mathbf{p}\, \mathrm{d}^N\mathbf{r}\, \mathrm{d}V\, \mathrm{d}p_\epsilon\, \mathrm{d}\eta\, \mathrm{d}p_\eta \tag{5.12}$$

If the forces do not sum to zero, then the only conservation law that is always present is the energy,

$$H' = H(\mathbf{p},\mathbf{r}) + \frac{p_\epsilon^{\,2}}{2W} + \frac{p_\eta^{\,2}}{2Q} + (\mathrm{d}N + 1)kT\eta + P_{\mathrm{ext}}V \tag{5.13}$$

Using eq 4.13, the microcanonical partition function corresponding to eqs 5.9 can be constructed,

$$\Omega(N,P_{\mathrm{ext}},T,E) = \int \mathrm{d}V \int \mathrm{d}^N\mathbf{p} \int_{D(V)} \mathrm{d}^N\mathbf{r}\, \mathrm{d}p_\epsilon\, \mathrm{d}p_\eta\, \mathrm{d}\eta$$
$$\exp[(\mathrm{d}N + 1)\eta]\, \delta\left(H(\mathbf{p},\mathbf{r}) + \frac{p_\epsilon^{\,2}}{2W} + \frac{p_\eta^{\,2}}{2Q} + (\mathrm{d}N + 1)kT\eta + P_{\mathrm{ext}}V - E\right) \tag{5.14}$$

Integrating over η using the δ-function and then over p_η and p_ϵ yields

$$\Omega(N,P_{\mathrm{ext}},T,E) \propto \int \mathrm{d}V \exp(-\beta P_{\mathrm{ext}}V) \int \mathrm{d}\mathbf{p} \int_{D(V)} \mathrm{d}\mathbf{r}\, \mathrm{e}^{-\beta H(\mathbf{p},\mathbf{r})} \tag{5.15}$$

the correct partition function for the *NPT* ensemble.

As a simple demonstration of the *NPT* ensemble method, consider a single particle moving in a one-dimensional periodic potential,

$$U(x,V) = \frac{m\omega^2}{4\pi^2V^2}\left[1 - \cos\left(\frac{2\pi x}{V}\right)\right] \tag{5.16}$$

where V is the length of the box, i.e., the one-dimensional volume. Figure 2 shows the position and volume distribution functions, $P(x)$ and $P(L)$, corresponding to the choice $P_{\mathrm{ext}} = 1$, $W = 1$, $Q = 1$, and $kT = 1$. The distribution functions are compared with the analytical results. It can be seen that the proposed algorithm correctly generates the desired distribution functions. In this example, a separate NHC is coupled to the particle and to the barostat.

These examples show how the principles of classical non-Hamiltonian statistical mechanics can be applied in the design of MD algorithms that generate different statistical ensembles. Note, non-Hamiltonian dynamics are different from Hamiltonian dynamics. These non-Hamiltonian systems reduce to approximate Hamiltonian dynamics when the extended system coupling parameters (here, Q and W) are large. An important outstanding question concerns the possibility of employing non-Hamiltonian dynamical schemes to generate a grand canonical or μVT ensemble sampling method.

6. The Liouville Operator and Numerical Integration Methodology

It was noted in section 3 that a numerical Newtonian trajectory will diverge from the true Newtonian trajectory but that the numerical trajectory will be statistically equivalent to the true trajectory provided there is a well-defined energy conservation tolerance, ΔE, for all time, t. The error in a numerical trajectory decreases with decreasing time step, Δt. However, the smaller the time step, the more iterations of the numerical integrator/map that will be required to reach a given total time, $\mathcal{T}$. Thus, numerical integration is a balance between using the largest possible time step and maintaining an acceptable energy conservation tolerance, ΔE.

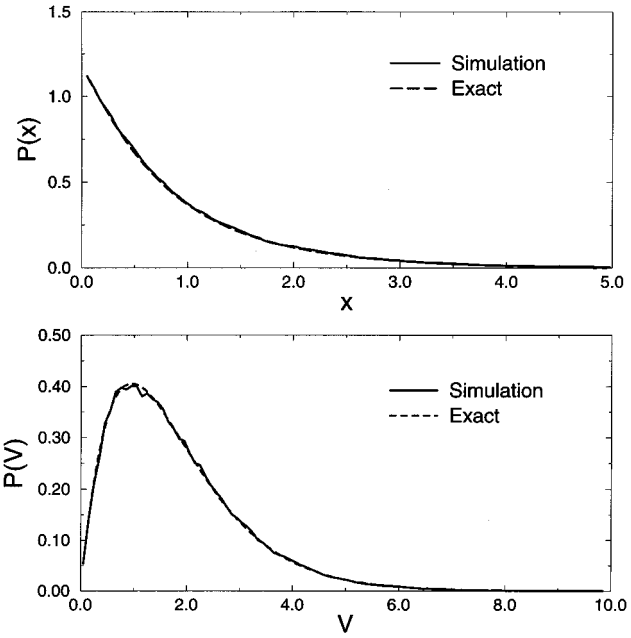

Figure 2. (a) The position distribution function for the model potential in eq 5.16. The simulation parameters are $m = 1$, $\omega = 1$, $kT = 1$, $P_{\mathrm{ext}} = 1$, and $W = 18$. A separate thermostat chain is used for the particle and for the barostat with $Q_{\mathrm{particle}} = 1$ and $Q_{\mathrm{barostat}} = 9$. The simulated distribution (solid line) is shown together with the analytical result (dashed line). (b) The volume distribution for the model potential described in part a. The simulated distribution (solid line) is shown together with the analytical result (dashed line).

The choice of an integration time step is determined by the nature of the forces acting on a system. In many-body systems, the forces arise from many different classes of interparticle interactions and generate motion on different, and often rather disparate, time scales. Consider, for example, a large protein in solution. There will be local intramolecular interactions which give rise to bond stretching, bending and torsional motion, long-range electrostatic and van der Waals interactions within the protein, and intermolecular interactions between the protein and solvent molecules which give rise to global rearrangements. Local intramolecular forces generate motion on a significantly shorter time scale than the long-range intermolecular interactions. However, the time step must be chosen such that the fastest motion in the system can be integrated stably and accurately. This fact leads to inefficient numerical procedures, because the computationally expensive slower forces are updated on time scales over which they do not change appreciably. This type of *multiple time scale problem* is almost always present in molecular dynamics simulations. Below, the general problem of numerical integration is discussed, and the proposed treatment is extended to the problem of multiple time scale motion in dynamical systems.

Several approaches can be employed to devise numerical integration procedures. One common technique develops low-order solutions using the time Taylor series expansion of the position and velocity, $\mathbf{r}_i(\Delta t)$, $\mathbf{v}_i(\Delta t)$, about $\Delta t = 0$, to obtain

useful maps. Careful consideration of the resulting functional form allows the time reversal symmetry of the equations of motion to be preserved.[6,28] The Taylor series based technique is appropriate for obtaining algorithms accurate to second order in time (local error) for Hamiltonian systems, but higher order schemes invariably require spatial derivatives of forces, which are difficult and computationally expensive to compute. When applied to non-Hamiltonian systems, such as eqs 5.1 or 5.9, the Taylor series approach generally yields schemes which do not preserve the invariant phase space measure.[37] Although time-reversal symmetry is formally satisfied, the methods often require iterative solutions which break the symmetry. Most textbook numerical integrators, such as Runge−Kutta or the Gear predictor−corrector, formally violate both conditions.

In this section, a third approach will be discussed, which will lead to time reversible methods that posses the correct metric. This new approach is based on an evolution operator formulation of classical mechanics.[38−40] Consider a Hamiltonian system described by eqs 3.4. These equations can be cast in the general form

$$\dot{\mathbf{x}} = iL\mathbf{x} \tag{6.1}$$

where $\mathbf{x}$ is the phase space vector and iL is the Liouville operator given by

$$iL = \{...,H\} \equiv \sum_{i=1}^{N}\left[\frac{\partial H}{\partial \mathbf{r}_i}\cdot\frac{\partial}{\partial \mathbf{r}_i} - \frac{\partial H}{\partial \mathbf{r}_i}\cdot\frac{\partial}{\partial \mathbf{p}_i}\right] = \sum_{i=1}^{N}\left[\frac{\mathbf{p}_i}{m_i}\cdot\frac{\partial}{\partial \mathbf{r}_i} + \mathbf{F}_i\cdot\frac{\partial}{\partial \mathbf{p}_i}\right] \tag{6.2}$$

Equations 6.1 have the formal solution

$$\mathbf{x}(t) = e^{iLt}\mathbf{x}(0) \tag{6.3}$$

Equation 6.3 is the starting point for the derivation of numerical integration procedures. The unitary operator, $\exp(iLt)$, is the *classical propagator*. Its action on $\mathbf{x}(0)$ cannot be determined analytically for any but a few simple cases. However, the formal solution to Hamilton's equations can be used to generate practical numerical integrators through introduction of an approximation to the classical propagator into eq 6.3. Suppose, for example, that the Liouville operator iL can be written as the sum of two parts, $iL = iL_1 + iL_2$, such that the action of the classical propagator on $\mathbf{x}(0)$ for each part can be evaluated analytically. The classical propagator can be rewritten using the Trotter theorem, which states

$$\exp(iLt) = \exp[(iL_1 + iL_2)t] = \lim_{P\to\infty}\left[\exp\left(\frac{iL_2t}{2P}\right)\exp\left(\frac{iL_1t}{P}\right)\exp\left(\frac{iL_2t}{2P}\right)\right]^P \tag{6.4}$$

Defining $t/P = \Delta t$ for finite P, the approximation

$$\exp(iL\Delta t) \approx \exp(iL_1\Delta t/2)\exp(iL_2\Delta t)\exp(iL_1\Delta t/2) + O(\Delta t^3)$$

$$\exp(iLP\Delta t) \approx \prod_{k=1}^{P}\exp(iL_1\Delta t/2)\exp(iL_2\Delta t)\exp(iL_1\Delta t/2) + O(t\Delta t^2) \tag{6.5}$$

can be made, which yields a numerical integration procedure that is accurate to the second order in the time step at long times. Consider the choice

$$iL_1 = \sum_{i=1}^{N}\frac{\mathbf{p}_i}{m_i}\cdot\frac{\partial}{\partial \mathbf{r}_i}$$

$$iL_2 = \sum_{i=1}^{N}\mathbf{F}_i\cdot\frac{\partial}{\partial \mathbf{p}_i} \tag{6.6}$$

Since the forces $\mathbf{F}_i$ are taken to depend only on the positions, the operator $\exp(iL_2\Delta t/2)$ becomes a translation operator on the momenta: $\mathbf{p}_i \rightarrow \mathbf{p}_i + (\Delta t/2)\mathbf{F}_i(\mathbf{r})$. Similarly, $\exp(iL_1\Delta t)$ is a translation operator on the positions: $\mathbf{r}_i \rightarrow \mathbf{r}_i + \Delta t(\mathbf{p}_i/m_i)$. Combining these two facts allows the action of the operator in eq 6.5 on the full set of positions and momenta to be evaluated analytically, yielding the approximate evolution:

$$\mathbf{r}_i(\Delta t) = \mathbf{r}_i(0) + \Delta t\mathbf{v}_i(0) + \frac{\Delta t^2}{2m_i}\mathbf{F}_i(0)$$

$$\mathbf{v}_i(\Delta t) = \mathbf{v}_i(0) + \frac{\Delta t}{2m_i}[\mathbf{F}_i(0) + \mathbf{F}_i(\Delta t)] \tag{6.7}$$

Equations 6.7 constitute the so-called *velocity Verlet*[41] integrator, derived here in a new and powerful way. The power of the operator technology lies in the fact that the same evolution can be obtained by viewing eq 6.5 as set of three sequential update steps without requiring the closed form expression of eqs 6.7. The three steps are simply (i) a velocity translation by an amount $(\Delta t/2m_i)\mathbf{F}_i$, (ii) a position translation by an amount $\Delta t\mathbf{v}_i$ using the velocity obtained in step i, and (iii) a velocity translation by an amount $(\Delta t/2m_i)\mathbf{F}_i$ using forces calculated at the position obtained in step ii. The scheme, therefore, appears as

for $i = 1$ to N

$\mathbf{v}_i \leftarrow \mathbf{v}_i + \frac{\Delta t}{2m_i}\mathbf{F}_i$

$\mathbf{r}_i \leftarrow \mathbf{r}_i + \Delta t\mathbf{v}_i$

endfor

get new forces

for $i = 1$ to N

$\mathbf{v}_i \leftarrow \mathbf{v}_i + \frac{\Delta t}{2m_i}\mathbf{F}_i$

endfor (6.8)

This procedure of translating each operator into an update step, which then can be turned into an instruction in computer code, is called the *direct translation technique*.[40] The direct translation technique, although seemingly trivial in this example, proves immensely powerful in complex, non-Hamiltonian systems of the type discussed in section 5. It should be noted that the Liouville operator approach has also been employed in the numerial propagation of quantum systems.[42,43]

An important property of the map generated by eq 6.5 is that it preserves the invariant phase space measure. This property is known as the *symplectic property*. For a Hamiltonian system, the symplectic property of a map is equivalent to the statement that its Jacobian is unity

$$J(\mathbf{x}_{\Delta t};\mathbf{x}_0) = \frac{\partial(x^1_{\Delta t}, ..., x^n_{\Delta t})}{\partial(x^1_0, ..., x^n_0)} \tag{6.9}$$

where $\mathbf{x}_{\Delta t}$ is the approximate evolution of $\mathbf{x}_0$ to $\mathbf{x}_{t=\Delta t}$ generated

by the numerical map. The significance of the symplectic property is that it ensures that the error of the map is bounded, i.e., there will be no secular growth in the energy conservation error which can affect the long-time statistical behavior of the dynamics, provided the time step is not too large. In order to see this, consider the example of a one-dimensional harmonic oscillator, $H = mv^2/2 + m\omega^2x^2/2$, for which the map in eq 6.7 exactly preserves the following time step-dependent Hamiltonian:[44]

$$\tilde{H}(\Delta t) = \frac{mv^2/2}{1 - \theta^2(\Delta t)} + \frac{m}{2}\omega^2 x^2 \tag{6.10}$$

where $\theta(\Delta t) = \omega\Delta t/2$. The difference between $\tilde{H}$ and the true Hamiltonian is

$$|\tilde{H}(\Delta t) - H| = \frac{mv^2}{2}\frac{\theta^2(\Delta t)}{1 - \theta^2(\Delta t)} \tag{6.11}$$

which shows that the energy conservation is rigorously bounded, since $\tilde{H}(\Delta t)$ is exactly conserved by the velocity Verlet map and the range of v is, thus, restricted. Notice, however, that $\tilde{H}(\Delta t)$ diverges at $\Delta t = 2/\omega$. At such a large time step, the orbits generated by the map undergo a transition from elliptical to hyperbolic, and the map breaks down. The power of the operator-based approach is that it allows symplectic integration methods to be constructed rather easily while the other approaches alluded to above do not. This is especially important in non-Hamiltonian systems, where it is necessary to ensure that the integrator preserves the more complicated phase space metric in eq 4.9.

The approach illustrated above can be easily extended to treat systems with multiple time scale motion. The method, which was first presented in ref 39, is based on the introduction of a reference system and a subdivision of the force, $\mathbf{F}_i$, into a contribution from the reference system and a deviation from the true force,

$$\mathbf{F}_i = \mathbf{F}_i^{(\mathrm{ref})} + \mathbf{F}_i^{(\mathrm{del})}$$
$$\mathbf{F}_i^{(\mathrm{del})} \equiv \mathbf{F}_i - \mathbf{F}_i^{(\mathrm{ref})} \tag{6.12}$$

If the reference system is chosen such that $\mathbf{F}_i^{(\mathrm{ref})}$ is a good approximation to the true force, then propagation of the system using the reference force with an occasional correction due to the difference force $\mathbf{F}_i^{(\mathrm{del})}$ could yield an efficient and accurate scheme. In particular, if $\mathbf{F}_i^{(\mathrm{ref})}$ is taken to be the fast local intramolecular forces, which are usually computationally inexpensive to evaluate, then an efficient multiple time step numerical procedure can be developed. By introducing a time step, δt, appropriate for the reference system, a Liouville operator, $iL^{(\mathrm{ref})}$, for the reference system and a deviation, $iL^{(\mathrm{del})}$, from the true Liouville operator according to

$$iL^{(\mathrm{ref})} = \sum_{i=1}^{N}\left[\frac{\mathbf{p}_i}{m_i}\cdot\frac{\partial}{\partial\mathbf{r}_i} + \mathbf{F}_i^{(\mathrm{ref})}\cdot\frac{\partial}{\partial\mathbf{p}_i}\right]$$
$$iL^{(\mathrm{del})} = \sum_{i=1}^{N}\mathbf{F}_i^{(\mathrm{del})}\cdot\frac{\partial}{\partial\mathbf{p}_i} \tag{6.13}$$

such that $iL = iL^{(\mathrm{ref})} + iL^{(\mathrm{del})}$, a propagator for the time step Δt can be constructed in which $iL^{(\mathrm{ref})}$ is used to propagate the system over n steps with a time step δt, and the correction $iL^{(\mathrm{del})}$ is applied in such a way that the overall scheme is time-reversible,

$$\exp(iL\Delta t) = \exp\left(iL^{(\mathrm{del})}\frac{\Delta t}{2}\right)\left[\exp\left(iL_2^{(\mathrm{ref})}\frac{\delta t}{2}\right)\exp(iL_1^{(\mathrm{ref})}\delta t)\exp\left(iL_2^{(\mathrm{ref})}\frac{\delta t}{2}\right)\right]^n\exp\left(iL^{(\mathrm{del})}\frac{\Delta t}{2}\right) \tag{6.14}$$

Here, $iL_1^{(\mathrm{ref})}$ and $iL_2^{(\mathrm{ref})}$ are defined analogously to eq 6.6 with $\mathbf{F}_i$ replaced by $\mathbf{F}_i^{(\mathrm{ref})}$. The mechanism of eq 6.14 can be gleaned by inspection. The operator in brackets corresponds to n steps of reference system propagation using the algorithm of eq 6.7 with a time step of δt. The correction $\exp(iL^{(\mathrm{del})}\Delta t/2)$ is applied both before and after the the reference system propagation, making the scheme time reversible. This numerical integration procedure is called the r-RESPA method,[39] which stands for *r*eversible *r*eference *s*ystem *p*ropagator *a*lgorithm. The advantage of this method is that if there is a wide separation in time scales, choosing the reference system force equal to the "fast" force will be a good approximation to the true force. This will allow n and Δt to be large. If, in addition, the fast forces are computationally inexpensive to evaluate, then there will be a large gain in cpu time, as the expensive slow forces will only need to be updated once every large time step, Δt. Various forms of this algorithm are now widely used.[45−49] However, under certain circumstances the time step can be limited by resonance phenomena.[50,51]

One can also imagine extending the r-RESPA algorithm to incorporate motion on more than two time scales. For example, local intramolecular forces might involve fast bond and bend forces and slower torsional forces. In addition, intermolecular forces often separate naturally into short and long range components, bringing in another time scale separation. In this case, short range forces generally consist mostly of strong, repulsive collisions, while long range forces tend to decay as $\pm 1/r^n$ where $n \leq 6$. When several time scales are present, eq 6.14 can be straightforwardly generalized to incorporate each time scale with its own time step.[39,40,46] For example, for a system with three characteristic time scales, a reference force $\mathbf{F}_i^{(\mathrm{ref})}$ and two corrections $\mathbf{F}_i^{(\mathrm{del})}$ and $\mathbf{F}_i^{(\mathrm{Del})}$ are chosen such that $\mathbf{F}_i = \mathbf{F}_i^{(\mathrm{ref})} + \mathbf{F}_i^{(\mathrm{del})} + \mathbf{F}_i^{(\mathrm{Del})}$. If the corresponding Liouville operators are $iL^{(\mathrm{ref})}$, $iL^{(\mathrm{del})}$, and $iL^{(\mathrm{Del})}$, defined analogously to eq 6.13, and the corresponding time steps are δt, Δt, and $\Delta\mathscr{T}$, respectively, then the three time step propagator is

$$\exp(iL\Delta\mathscr{T}) = \exp\left(iL^{(\mathrm{Del})}\frac{\Delta\mathscr{T}}{2}\right)\left\{\exp\left(iL^{(\mathrm{del})}\frac{\Delta t}{2}\right)\left[\exp(iL_2^{(\mathrm{ref})})\exp(iL_1^{(\mathrm{ref})}\delta t)\exp\left(iL_2^{(\mathrm{ref})}\frac{\delta t}{2}\right)\right]^n\exp\left(iL^{(\mathrm{del})}\frac{\Delta t}{2}\right)\right\}^m\exp\left(iL^{(\mathrm{Del})}\frac{\Delta\mathscr{T}}{2}\right) \tag{6.15}$$

Thus, the correction due to the slowest time scale is applied only every mn time steps, and the intermediate time scale correction is applied every n steps. Numerical procedures such as those of eqs 6.14 and 6.15 can lead to a considerable savings in the cpu time needed to perform a MD calculation. Examples will be discussed in the next section. (Unfortunately, cpu time saving can be limited by resonance behavior that occurs when the largest time step is close to the natural period of the fastest motion.[50,51])

Numerical integration of non-Hamiltonian systems, such as those given in eqs 5.1 or 5.9 present a series of challenges. First, velocity-dependent forces are present, which are generally difficult to treat. Second, in addition to ensuring that the conserved quantities are properly bounded by the numerical

map, one must make certain that the invariant measure $\sqrt{g(\mathbf{x})}\,d\mathbf{x}$ is preserved as well. It is in this aspect that the Taylor series approach fails for non-Hamiltonian systems, and this failure has been shown numerically to lead to nonsecular growth in the energy.[37] The operator approach, when applied correctly, solves these problems.[37]

A system of equations such as eqs 4.3 can be cast in the form of eq 6.1 by defining iL to be

$$iL = \sum_{i=1}^{n} \dot{x}_i \frac{\partial}{\partial x_i} = \sum_{i=1}^{n} \xi_i(x) \frac{\partial}{\partial x_i} \tag{6.16}$$

with a formal solution also given by eq 6.3. For the systems discussed in section 5, the approach that has been adopted[40] is based on a subdivision of the Liouville operator into a part that is purely Hamiltonian in the system variables $\{\mathbf{p},\mathbf{r}\}$ and the remainder, which contains all the non-Hamiltonian contributions. Thus,

$$iL = iL_{\mathrm{H}} + iL_{\mathrm{NH}} \tag{6.17}$$

where iL_{H} is given by eq 6.2, and $iL_{\mathrm{NH}} = iL - iL_{\mathrm{H}}$ is the non-Hamiltonian part of the operator. The operator iL_{H} is subdivided according to eq 6.6, and a propagator is constructed as follows:

$$\exp(iL\Delta t) = \exp\left(iL_{\mathrm{NH}}\frac{\Delta t}{2}\right)\exp\left(iL_{\mathrm{H}}^{(2)}\frac{\Delta t}{2}\right)\exp(iL_{\mathrm{H}}^{(1)}\Delta t)\exp\left(iL_{\mathrm{H}}^{(2)}\frac{\Delta t}{2}\right)\exp\left(iL_{\mathrm{NH}}\frac{\Delta t}{2}\right) \tag{6.18}$$

The form of eq 6.18 is justified by the fact that the Trotter theorem, eq 6.4, can be generalized to the exponential of a sum of more than two operators. The strategy of eq 6.18 is that the non-Hamiltonian part of the dynamics is applied, in essence, as a correction to a Hamiltonian reference system, in analogy to eq 6.14 for the multiple time scale problem. The operator iL_{NH} generally contains a large number of terms. Hence, it needs to be subdivided into simple contributions and a corresponding factorization of the operator $\exp(iL_{\mathrm{NH}}(\Delta t/2))$ made (see refs 52 and 40 for a more complete discussion).

Note, it is also possible to treat non-Hamiltonian systems with multiple time scales by combining eq 6.18 with eq 6.14 or 6.15. An example of such a system is a protein or peptide in solution under conditions of constant temperature and pressure (the *NPT* ensemble) or constant volume and temperature (the *NVT* ensemble). Results of a specific study are presented in section 7.2.1.

7. Interparticle Interaction Models in Molecular Dynamics

The preceding discussion of the MD technique has focused on general theoretical issues and specific numerical integration methodologies. It was assumed that the forces on the particles were known. However, developing an adequate description of the interparticle interactions is a difficult problem in developing good models for molecular systems. In this section, several approaches to this problem will be discussed. First, an empirical approach, based on the introduction of a mathematical model, or *force field* will be described, and an example study on a complex system, the mutant T4 lysozyme M61, under the CHARMM22 force field, will be presented. Next, a powerful alternative approach, which combines classical dynamics with electronic structure, referred to as ab initio molecular dynamics will be introduced and recent examples highlighting the capability of this method will be presented. The relative merits and disadvantages of each scheme will be noted, as well as prospects for combining the two approaches.

7.1. Boundary Conditions. Experimentally observable quantities are formally defined in the thermodynamic limit, where the number of particles, N, in a system and volume, V, are taken to infinity such that N/V remains constant. In actual computations, however, finite systems are used. This raises an important issue concerning the calculation of forces and energies in such systems. Particularly, given an N-particle system in a container of volume, V, how should the boundaries of the container be treated in order to minimize finite size effects?

One of the most common applications of the molecular dynamics technique is to bulk condensed phases such as liquids and solids, where it is natural to apply *periodic boundary conditions*. Under periodic boundary conditions, the system is replicated infinitely throughout all space. Periodic boundary conditions are a natural choice to use to treat a solid. However, in a liquid, they are also found to reduce the influence of the presence of boundaries on the properties of a system.

Although convenient, effective, and simple to apply, certain subtle problems arise when periodic boundary conditions are employed. In particular, long range forces, whose spatial range may extend beyond the boundaries of the container into surrounding images, present a challenge. Long range forces can only be correctly calculated by summing over all the periodic replicas of the original system. However, the associated computational effort is considerable. Fortunately, methods have been developed to treat this problem. Specifically, the Ewald summation technique, developed originally to treat Coulomb interactions and later extended to treat general interactions of the form $1/r^n$ for $n \leq 3$,[53] has proved enormously successful. The basic idea behind the technique is to divide the relevant part of the potential into a short range and a long range contribution. For the Coulomb potential, $1/r$, for example, this can be achieved via the identity

$$\frac{1}{r} = \frac{\mathrm{erf}(\alpha r)}{r} + \frac{\mathrm{erfc}(\alpha r)}{r} \tag{7.1}$$

where $\mathrm{erf}(x)$ and $\mathrm{erfc}(x)$ are the error function and complementary error function, respectively ($\mathrm{erf}(x) + \mathrm{erfc}(x) = 1$). The variable, α, is a convergence parameter, which can be optimized for each system studied. The short range term, $\mathrm{erfc}(\alpha r)/r$, is treated as an ordinary short range interaction, i.e., using a spherical cutoff to truncate the interaction at large spatial distances where the potential is small. The long range term, $\mathrm{erf}(\alpha r)/r$, is Fourier transformed into reciprocal space, where it takes the short-ranged form, $\exp(-g^2/4\alpha^2)$, and can be evaluated accurately by summing over only a small number of reciprocal space vectors of the simulation cell. Such reciprocal space sums can be evaluated with high a degree of efficiency ($N \log N$) using particle−mesh methods.[54−57] An efficient real-space alternative, the fast multipole method or FMM, has also been widely adopted.[58−60]

Long range forces in systems that are periodic in less than three dimensions, such as surfaces, wires, and clusters, are more difficult to treat using reciprocal space-based methods. However, a general reciprocal space-based technique has been developed for systems with periodicity less than three.[61,62] This new technique is a generalization of the Ewald summation method and requires only a small modification of the standard Ewald method to implement. It can also be easily incorporated into the particle−mesh Ewald scheme. In addition, the new method

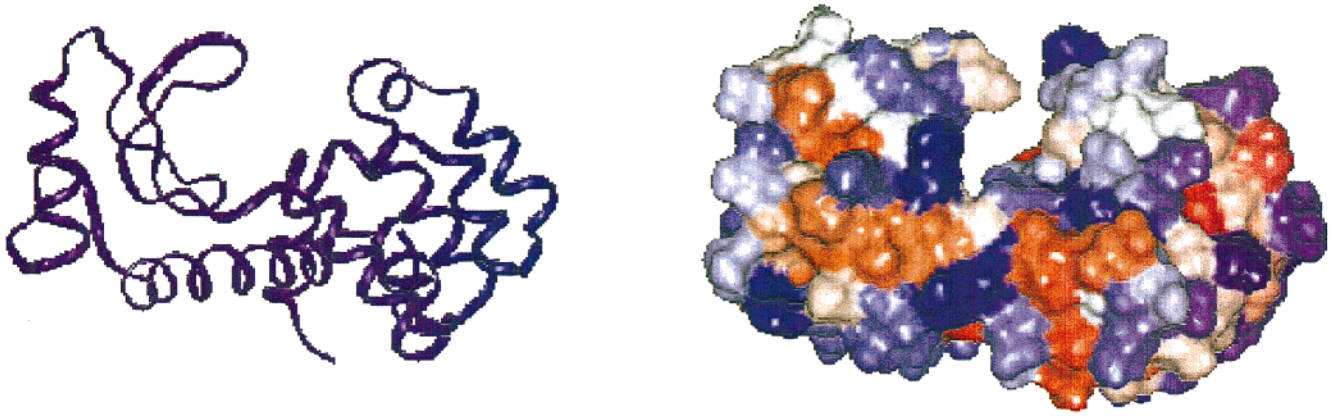

Figure 3. A conformer of the M6I T4 lysozyme in water solution at 300 K with the hinge-bending angle of $\theta = 95°$. Secondary structure elements shown on the left and the solvent-accessible Connolly surface on the right.

is applicable to both force-field based and ab initio molecular dynamics calculations.

7.2. Force Fields. One approach to modeling the interactions between particles in a MD calculation is to introduce an empirical force field model. Here, a specific mathematical forms for different interactions are postulated, and the parameters that characterize the individual functions are fit to experimental data and/or ab initio quantum chemical calculations. One example is the CHARMM22 force field.[63] The specific form for the interaction potential $U(\mathbf{r})$ in the CHARMM22 force field is[63]

$$U(\mathbf{r}) \equiv U(\mathbf{r}_1, ..., \mathbf{r}_N) = \sum_{\text{bonds}} K_b(b - b_0)^2 + \sum_{\text{angle}} K_\theta(\theta - \theta_0)^2 + \sum_{\text{UB}} K_{\text{UB}}(S - S_0)^2 + \sum_{\text{torsion}} K_\phi(1 + \cos(\eta\phi - \delta)) + \sum_{\text{imp}} K_{\text{imp}}(\phi_{\text{imp}} - \phi_{\text{imp}}^{(0)})^2 + \sum_{\text{nonbond}} 4\epsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] + \sum_{\text{nonbond}} \frac{q_i q_j}{r_{ij}} \quad (7.2)$$

In eq 7.2, the bond length, b; bend angle, θ; torsion angle, ϕ; 1−3 distance, S; improper torsion angle, ϕ_{imp}; and nonbonded distance, r_{ij}, are functions of the Cartesian coordinates of the system. The remaining parameters, the equilibrium bond lengths (b_0) and bond force constants (K_b), equilibrium bend angle (θ_0), and force constants (K_θ), etc., are the parameters that characterize the force field. A different force field, such as AMBER95,[64] generally has a similar mathematical form but is characterized by a different set of parameters. Note that eq 7.2 contains a sum over bonded pairs, bend and torsion angle terms, and nonbonded pairs. The bond and bend terms are harmonic oscillator functions. The torsion angle term is a cosine function, a multiple minimum function that allows for different conformations, e.g., trans and gauche. Nonbonded interactions include both Coulomb and Lennard−Jones terms. In addition, eq 7.2 contains Urey−Bradley (harmonic 1−3, interactions) and improper torsion angle terms. The former allows for a better fit to experimental data and the latter keeps planar species flat, when sp and sp^2 hybridized atoms are present. Note that the use of a force field requires imposition of a particular connectivity among the atoms and precludes the treatment of chemical events in which bonds are broken and formed.

In addition to the functional form of the potential, a force field is generally designed to work with a specific water model in solution phase. In the case of the CHARMM22, the particular water model is the TIP3P[65] model, a rigid water model with the OH distance and H−O−H angle fixed at 0.9572 Å and 104.52°, respectively, a charge of −0.834e on the oxygen and 0.417e on each hydrogen, and Lennard−Jones parameters $\epsilon =$ 76.54 K and $\sigma = 3.15$ Å.

Figure 4. A conformer of the M6I T4 lysozyme in water solution with a hinge-bending angle of $\theta = 95°$ aligned with a crystal conformer obtained from the Brookhaven Protein Databank (entry 150L, conformer A) (rms = 1.5 Å^2).

7.2.1. The mutant T4 Lysozyme M6I. The mutant phage T4 lysosyme M6I is a two-domain single macromolecular enzyme which belongs to the hinge-bending class of proteins. It is characterized by two large lobes connected by a narrow waist region. An active site cleft is formed between the lobes and a hinge-bend angle can be defined that measures the accessibility of the cleft (i.e., open or shut, see Figures 3 and 4).[66] The enzyme possesses two unique crystalline solid forms in which five different protein conformations, each with a different hinge-bend angle are present. These observed angles span a range of 30° (65°−95°) implying a high degree of mobility in the hinge-bend degree of freedom.[66] It is, therefore, of interest to study, theoretically, the hinge-bend angle equilibrium in both gas and solution phases and to examine the behavior of this highly mobile degree of freedom (the hinge-bend angle) as a function of environment. In addition, the studies will stringently test the effectiveness of modern force field models and methodology on a complex realistic system.

Recent theoretical work has been conducted on the mutant T4 lysozyme M6I in water solution to examine the mobility of the hinge-bend motion in this environment.[67−72] However, the utility of standard MD calculations is limited by the short time scales that can be examined and the comparatively long time

scale associated with the hinge-bend motion. As a result, simulation studies must be carried out to long times in order for the system to sample configuration space, adequately. Normal mode or essential dynamics techniques have been employed to perform postsimulation analysis of domain motion. No direct probes of equilibrium properties such as the hinge-bend probability distribution function have been undertaken due to the sampling problem.

Using methodology borrowed from path integrals molecular dynamics simulations (see section 8),[34] specifically order N noncanonical variable transformations, ergodic (for systems with small barriers) constant temperature molecular dynamics methods (cf. eqs 5.1), and the multiple time scale integration described in section 6, it is possible to enhance the sampling efficiency of Gaussian random coil calculations by a factor of over 200. Most of the increase in efficiency can be ascribed to the noncanonical variable transformations, which permit the long wavelength fluctuations of the coil to occur on a fast time scale. These transformations can be applied without loss of generality or any approximation to realistic all-atom models of proteins in order to enhance large scale domain motion. The new method performs well simply because the "essential" modes of the protein backbone have random coil-like character and properly exciting these modes drives large domain motion. In contrast to essential dynamics or normal mode techniques, input trajectories, atomic Hessians, and/or a matrix diagonalization are not required.

Here, random coil variable transformations are employed to improve configurational sampling of the hinge-bend motion during extended system molecular dynamics simulations (see section 8) on the T4 phage lysozyme. These are coupled to extended system multiple time step algorithms which alone increase simulation time by a factor of 6 over conventional methods. Finally, umbrella sampling techniques are also employed to ensure the effective sampling of the hinge-bend equilibrium. Using this novel suite of methodology, the hinge-bend probability distribution function in both gas and solution phases at $T = 300$K under the CHARMM22[63] force field was determined. The simulations were performed the PINY_MD simulation code.[73]

The efficiency of the random coil methodology is demonstrated in Figure 5. Without the transformations, the hinge-bend angle of the molecule samples only a fraction of its available configuration space during the 200 ps run. The full hinge-bend probability distribution function in the gas phase is given in Figure 6. The distribution is narrow because the lid of the cleft "collapses" to form hydrogen bond contacts (see Figure 7). In contrast, the solution phase distribution is broad and the conformers align well with those taken from crystal structures (rms ≤ 2 Å^2 for conformers with similar hinge-bend angles). This indicates that the methodology has not "damaged" the conformational equilibrium but rather sampled it, appropriately.

The results presented above indicate that the MD methodology is sufficiently accurate to test the force fields. Second, the force fields seem to yield a reasonable approximation to the biomolecular system. Thus, a synergy between new potential models and methods development has resulted in a leap forward for simulation.

7.3. Ab Initio Molecular Dynamics. An accurate force field is an important element in the MD method, as it permits large systems to be studied at relatively little computational cost. However, as noted previously, current force field technology is not capable of describing chemical events involving bond breaking and forming. Another deficiency of current force fields

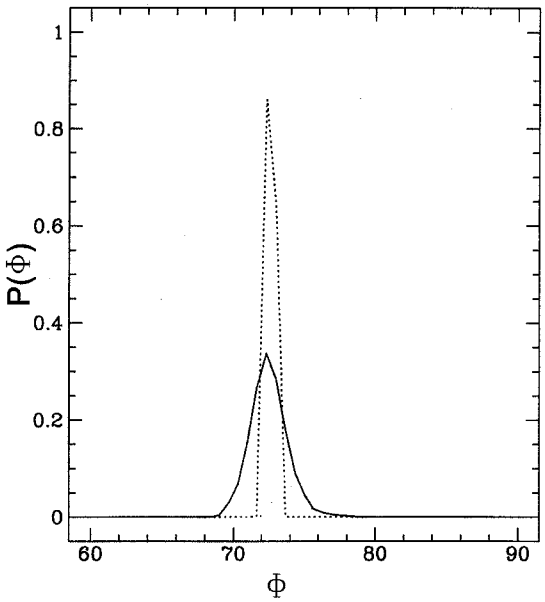

Figure 5. The biased hinge-bending angle probability distribution function of the M6I T4 lysozyme in vacuo with (solid line) and without (dashed line) random coil transformations.

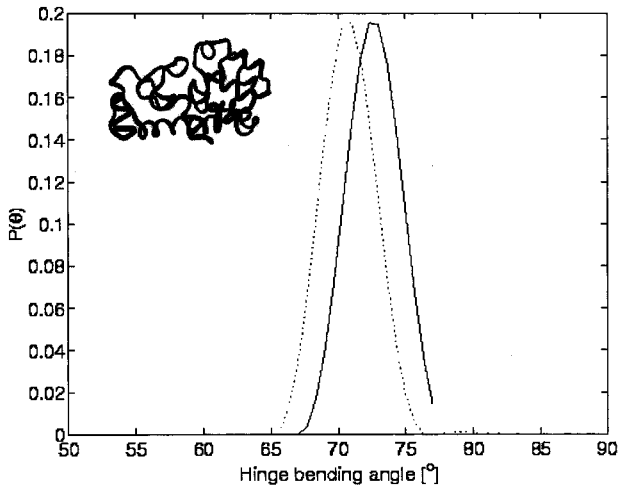

Figure 6. The true hinge-bending angle probability distribution function of the M6I T4 lysozyme in in vacuo at 300 K. The results of two independent runs of length 2.5 ns are shown.

is their failure to include polarization effects, although polarizable force fields have been introduced very recently.[74] The technique known as ab initio molecular dynamics (AIMD) solves these problems by combining "on the fly" electronic structure calculations with finite temperature dynamics. Not surprisingly, AIMD simulations are substantially more expensive than calculations based on empirical force fields. However, recent advances in electronic structure theory as well as readily available high-speed computers have begun to render the AIMD approach a viable one for studying chemical processes in the condensed phase.

The most important element in an AIMD calculation is the representation of the electronic structure. Clearly, calculation of the exact ground-state electronic wavefunction is intractable, and approximations must be used. The electronic structure theory employed should be reasonably accurate yet not too computationally demanding. One formulation of the electronic structure problem that satisfies these criteria is density functional theory (DFT).[75−77] DFT formulates the many-electron problem in terms of the electron density, $n(\mathbf{r})$, rather than the many-body

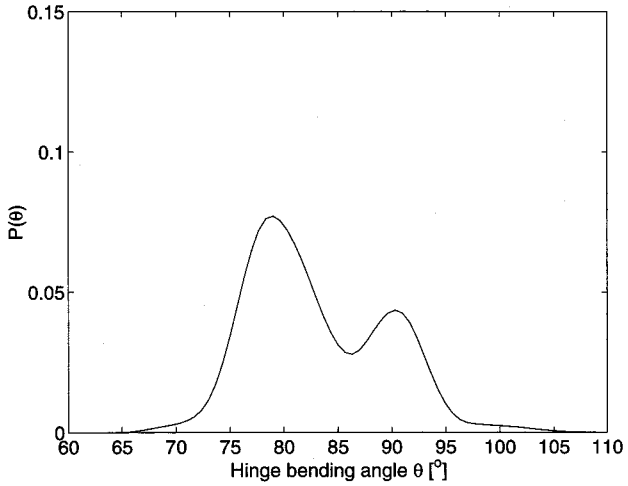

Figure 7. The true hinge-bending angle probability distribution function of the M6I T4 lysozyme in water solution at 300 K.

wavefunction. Thus, in principle, the central quantity is a function of just three rather than of 3*N* variables, a fact that renders calculations based on DFT computationally tractable. The basic tenet of DFT is that the energy of a quantum many-body system can be expressed as a unique functional of its density. By minimizing the density functional over all densities that give rise to a particular number of electrons, one obtains the ground state density and energy for a given system. Unfortunately, the explicit and unique form of this functional is not known. However, in the orbital-based formulation of DFT by Kohn and Sham, reasonable approximations to the density functional have been developed. In the Kohn−Sham formulation, the energy is expressed in terms of a set of *n* occupied single-particle orbitals $\psi_1(\mathbf{r})$, ..., $\psi_n(\mathbf{r})$ and the *N* nuclear positions, $\mathbf{R}_1$, ..., $\mathbf{R}_N$, and takes the form

$$E[\{\psi\},\{\mathbf{R}\}] = -\frac{1}{2}\sum_{i=1}^{n}\int \mathrm{d}\mathbf{r}\ \psi_i^*(\mathbf{r})\nabla^2\psi_i(\mathbf{r}) + \frac{1}{2}\int \mathrm{d}\mathbf{r}\ \mathrm{d}\mathbf{r}'\frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + E_{\mathrm{xc}}[n] + \int \mathrm{d}\mathbf{r}\ V_{\mathrm{ext}}(\mathbf{r},\mathbf{R}_1,...,\mathbf{R}_N)n(\mathbf{r}) \tag{7.3}$$

where the density $n(\mathbf{r})$ is related to the orbitals by

$$n(\mathbf{r}) = \sum_{i=1}^{n} |\psi_i(\mathbf{r})|^2 \tag{7.4}$$

and the orbitals are required to be mutually orthonormal,

$$\langle\psi_i|\psi_j\rangle = \delta_{ij} \tag{7.5}$$

In eq 7.3, V_{ext} represents the external potential due to the *N* nuclei and is given exactly by $V_{\mathrm{ext}} = -\sum_{I=1}^{N} = q_I/|\mathbf{r} - \mathbf{R}_I|$, where q_I is the charge on each nucleus. The first two terms in eq 7.3 are the electronic and Hartree energy terms, respectively. The functional $E_{\mathrm{xc}}[n]$ of the density, called the exchange-correlation functional, is unknown and must be approximated. For certain classes of systems, it has proved sufficient to approximate this functional by taking the exchange and correlation energies of. a homogeneous electron gas and substituting for the constant density, *n*, the inhomogeneous density $n(\mathbf{r})$. This approximation, known as the *local density approximation* (LDA), has proved useful and reasonably accurate in many problems of interest in metallic and semiconductor solids. However, it fails, for example, in hydrogen-bonded systems, where spatial variations in the electron density are too rapidly varying to be described adequately by LDA. The most common approach in such cases is to extend the dependence of-this functional to include the density $n(\mathbf{r})$ and its gradient $\nabla n(\mathbf{r})$: $E_{\mathrm{xc}} = E_{\mathrm{xc}}[n,\nabla n]$. This approximation is known as the *generalized gradient approximation* (GGA) and has become a popular approach due to the recent development of improved functionals.[78−83]

A possible strategy for combining electronic structure with molecular dynamics is the following: for a given set of initial nuclear positions $\mathbf{R}_1$, ..., $\mathbf{R}_N$, minimize the energy functional in eq 7.3 to obtain the ground state density $n_0(\mathbf{r})$ and corresponding orbitals $\psi_1^{(0)}(\mathbf{r})$, ..., $\psi_n^{(0)}(\mathbf{r})$. Given these quantities, the forces between the nuclei are given by the Hellman−Feynman theorem:

$$\mathbf{F}_I = -\frac{\partial}{\partial \mathbf{R}_I}E[\{\psi^{(0)}\},\{\mathbf{R}\}] \tag{7.6}$$

The forces are then fed into a numerical integration procedure together with a set of initial velocities for the nuclei, and a step of molecular dynamics is carried out, yielding a new set of positions and velocities. At the new nuclear positions, the energy functional is minimized again and a new set of forces is obtained and used to perform another step of MD propagation. This procedure is repeated until an entire trajectory has been generated. An elegant alternative formulation of this procedure was proposed by Car and Parrinello,[11] in which, rather than minimizing the functional at each new nuclear configuration, a fictitious dynamics for the electronic orbitals is introduced that allows them to follow the motion of the nuclei adiabatically. This dynamical procedure is constructed in such a way that if the orbitals are initially chosen corresponding to the ground state density at the initial nuclear configuration, they will remain approximately in the ground state as the nuclear configuration evolves in time. In the original formulation of the Car−Parrinello scheme, the orbitals are expanded in a plane wave basis,

$$\psi_i(\mathbf{r}) = \sum_{\mathbf{g}} c_{\mathbf{g}}^i \mathrm{e}^{i\mathbf{g}\cdot\mathbf{r}} \tag{7.7}$$

where $c_{\mathbf{g}}^i$ are the expansion coefficients. (This form of the plane wave expansion is actually a special case of a more general plane−wave expansion, in which the orbitals are assumed to be Bloch functions, $\psi_{i,\mathbf{k}}(\mathbf{r})$. Here, the choice $\mathbf{k} = (0,0,0)$, the so called *gamma point*, has been made.) The fictitious adiabatic dynamics is then formulated for the coefficients by introducing a set of velocities $v_{c_{\mathbf{g}}}^i = \dot{c}_{\mathbf{g}}^i$ and an associated mass parameter μ (having units of energy·(time)^2). In a Newtonian scheme, the equations of motion for the particles and coefficients then take the form

$$\mu\ddot{c}_{\mathbf{g}}^i = -\frac{\partial E}{\partial c^{*i}_{\mathbf{g}}} + \sum_j \Lambda_{ij}c_{\mathbf{g}}^j$$

$$M_I\ddot{\mathbf{R}}_I = -\frac{\partial E}{\partial \mathbf{R}_I} \tag{7.8}$$

where Λ_{ij} is a set of Lagrange multipliers for enforcing the orthonormality constraint. Various modifications of eqs 7.8, including versions for generating the *NVT*[52] and *NPT*[84,85]

ensembles and formulations which allow for nonorthogonal orbitals[86,87] or do not require explicit orthonormality constraints.[52,88]

Ab initio molecular dynamics has been used to predict, from first principles, the structural and dynamical properties of a number of neat molecular liquids and solids including water,[13,14,89] ice,[20,90,91] mono- and trihydrates of HCl,[92,93] nitromethane,[94] nitric acid trihydrate,[95] and ammonia.[15] It has also been used recently to study solutions, including trioxane in formaldehyde,[96,97] HCl,[98] HF,[99] and H_2SO_4,[100] water wires,[101,102] and the ions, H_3O^+ and OH^- [16−19,103] in water. In section 7.3.1, results of a recent study of the structure of neat liquid ammonia are presented.[15]

An important, developing area of MD methodology is the combination of the ab initio approach with empirical force fields. Such a combined scheme is expected to be of considerable utility in the treatment of large systems, in which chemical processes occur in a relatively localized region, e.g., at the active site of an enzyme or a chemical reaction in solution. In such systems, ab initio MD can be used to treat the chemically active region and a force field employed to describe the rest of the system. One of the difficult problems associated with a combined scheme is specifying how the electrons and nuclei in the ab initio region interact with the atoms in the force-field region. This is an especially challenging problem when it is necessary to "cut" bonds within a molecule, for example, in treating a reaction at the active site of an enzyme. The interested reader is referred to refs 104−108 for recent progress in this area.

7.3.1. *Ab Initio* Liquid Ammonia. Hydrogen-bonded liquids, such as water and ammonia, play an important role in solution chemistry. Liquid ammonia is employed as a solvent in various common organic reactions. In addition, metal−ammonia solutions (small amounts of metal dissolved in ammonia), exhibit many interesting properties, including a metal−insulator transition. At low concentration, these solutions are used to catalyze organic reactions. Predicting the structural and dynamical properties of hydrogen-bonded liquids has proved particularly challenging for the DFT, as the LDA approximation to the exchange-correlation functional fails. The situation improved dramatically with the advent of the GGA, and currently, structural and dynamical properties of water have been shown to be well reproduced[14,89] employing, for example, the B-LYP functional.[78,79] Although excellent empirical model potentials exist for ammonia (see, e.g., ref 109), it is important to demonstrate the applicability of DFT to this system if one desires to examine chemical reactions in this solvent. Here, we show that the B-LYP functional also provides a reasonably accurate description of the structure of bulk liquid ammonia.

In this study, a system of 32 ammonia molecules in a cubic, periodic box of length 11.229 Å was equilibrated at $T = 273$ K using the empirical liquid ammonia model of Impey and Klein.[109] Thereafter, Car−Parrinello ab initio molecular dynamics (cf. eqs 7.8) was performed using a plane wave basis set and the core electrons removed using the pseudopotentials of Bachelet−Hamann and Schlüter.[110] Tests comparing the use of these pseudopotentials against those of Troullier and Martins[111] at the chosen plane wave energy cutoff ($E_c = 70$ Rydbergs) yielded good agreement and, therefore, justifies their use. Equations 7.8 were integrated using a time step of 0.125 fs, generating a trajectory of 5 ps total length. These simulations were carried using the PINY_MD simulation code.[73]

In Figure 8, the NN, NH, and HH radial distribution functions generated from the ab initio trajectory are shown together with the experimental results of Soper et al.[112] It can be seen that

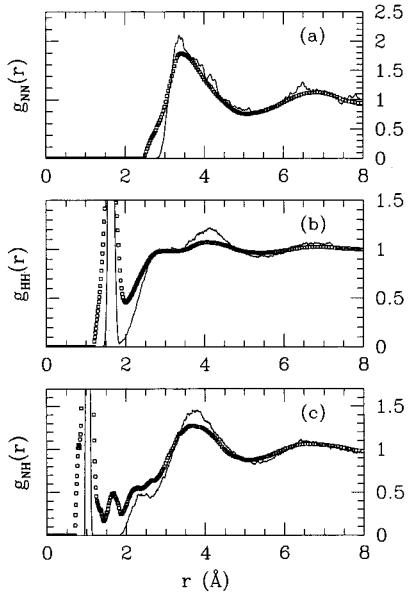

Figure 8. Experimental (square) and ab initio radial distribution functions of liquid ammonia at 273 K: (a) Nitrogen−nitrogen, (b) hydrogen−hydrogen, and (c) nitrogen−hydrogen.

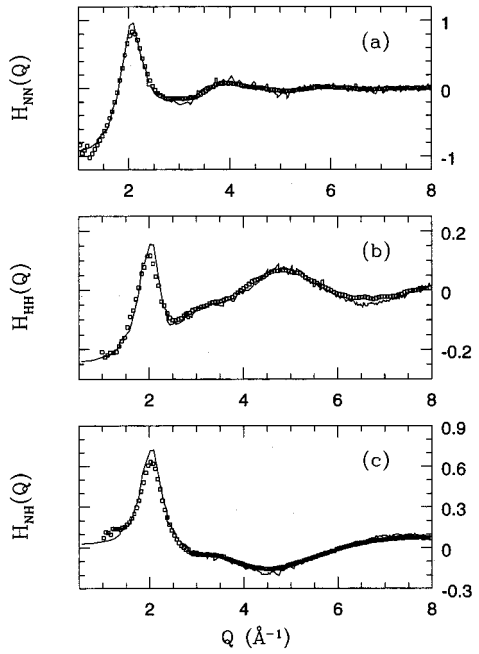

Figure 9. Experimental (square) and ab initio partial structure factors of liquid ammonia at 273 K: (a) Nitrogen−nitrogen, (b) hydrogen−hydrogen, and (c) nitrogen−hydrogen.

the DFT is capable of reproducing the experimental peak locations and approximate peak heights. In Figure 9, the corresponding structure factors are shown together with the experimental results. This quantity is a more stringent test of the results with experiment, and it can be seen that the DFT results are in good agreement with the experiment. In addition, the coordination numbers, computed from the area under the first peak of the radial distribution functions, yield 13.2 for $g_{NN}(r)$ and 41 for $g_{HH}(r)$. Both numbers are in good agreement with the experimental values of 14 and 42 for g_{NN} and g_{HH}, respectively. These results open up a number of interesting possibilities for ab initio MD studies involving chemical reactions in liquid ammonia.

A unique aspect of the AIMD method is the availability of electronic properties of the system, which can be used, for

example, to calculate the electric dipole moment[113] and, hence, IR intensities.[89] In ref 15, the electronic density of states and Kohn−Sham eigenvalue spectra for the ammonia monomer, dimer, and liquid at $T = 273$ K are discussed in detail.

8. Path Integral Molecular Dynamics. In the preceding discussion, it has been assumed that the N particles in the system under consideration could be treated as classical point particles. In many cases, this treatment is justifiable, however, there is a large class of systems for which such an approximation is not valid. In general, systems where hydrogen/proton motion is important, for example, proton transfer processes, often have significant nuclear quantum effects. The problem of treating nuclear quantum effects in a system at finite temperature requires the solution of a quantum statistical mechanical problem. One approach that has been applied with considerable success is based on the Feynman path integral formalism of statistical mechanics.[114,115]

Consider the quantum canonical partition function for a single particle in one spatial dimension. The partition function is given by the trace:

$$Q = \mathrm{Tr}(\mathrm{e}^{-\beta H}) = \int \mathrm{d}x \langle x|\mathrm{e}^{-\beta \mathrm{H}}|x\rangle \tag{8.1}$$

where the trace is carried out in the coordinate basis. Assuming $H = T + U$, where T is the kinetic energy operator, and U is the potential, the Trotter theorem, eq 6.4, allows $\exp(-\beta H)$ to be expressed as $[\exp(-\beta U/2P)\exp(-\beta T/P)\exp(-\beta U/2P)]^P$ in the limit $P \rightarrow \infty$. The Trotter theorem expression for $\exp(-\beta H)$ is then substituted into eq 8.1, and an identity operator in the form of $I = \int \mathrm{d}x|x\rangle\langle x|$ is inserted in between each factor of $\exp(-\beta U/2P)\exp(-\beta T/P)\exp(-\beta U/2P)$, yielding

$$Q = \lim_{P\rightarrow\infty} \int \mathrm{d}x_1 \cdots \mathrm{d}x_P \prod_{s=1}^{P} \langle x^{(s)}|\mathrm{e}^{-\beta U/2P}\mathrm{e}^{-\beta T/P}\mathrm{e}^{-\beta U/2P}|x^{(s+1)}\rangle|_{x^{(P+1)}=x^{(1)}} \tag{8.2}$$

Then, using the fact

$$\langle x^{(s)}|\mathrm{e}^{-\beta U/2P}\mathrm{e}^{-\beta T/2P}\mathrm{e}^{-\beta U/2P}|x^{(s+1)}\rangle = \left(\frac{mP}{2\pi\beta\hbar^2}\right)^{1/2} \exp\left[-\frac{\beta}{2P}(U(x^{(s)}) + U(x^{(s+1)}))\right]\exp\left[-\frac{mP}{2\beta\hbar^2}(x^{(s)} - x^{(s+1)})^2\right] \tag{8.3}$$

one obtains the final expression for Q as a function of P

$$Q = \lim_{P\rightarrow\infty}\left(\frac{mP}{2\pi\beta\hbar^2}\right)^{P/2}\int \mathrm{d}x^{(1)}\cdots\mathrm{d}x^{(P)}\mathrm{e}^{-\beta\Phi(x^{(1)},\ldots,x^{(P)})} = \lim_{P\rightarrow\infty} Q_p \tag{8.4}$$

where Φ is an effective potential given by

$$\Phi(x^{(1)}, \ldots, x^{(P)}) = \sum_{s=1}^{P}\left[\frac{1}{2}m\omega_P^2(x^{(s)} - x^{(s+1)})^2 + \frac{1}{P}U(x^{(s)})\right] \tag{8.5}$$

with $\omega_P = \sqrt{P}/\beta\hbar$.

Equation 8.4 is in the form of a configurational partition function for a P-particle system in one dimension subject to a potential, $\Phi(x_1, \ldots, x_P)$. The configurational partition function can also be expressed in a quasiphase space form by recognizing that the prefactor can be written as a product of P uncoupled Gaussian integrals

$$Q_P = N\int \mathrm{d}^P p\, \mathrm{d}^P x\, \mathrm{e}^{-\beta H} \tag{8.6}$$

where

$$H = \sum_{s=1}^{P}\frac{(p^{(s)})^2}{2m'} + \Phi(x^{(1)}, \ldots, x^{(P)}) \tag{8.7}$$

In eq 8.6, the constant N is an overall constant that ensures the equality of eqs 8.6 and 8.4. In addition, the mass m', being a fictitious mass, is arbitrary, a fact that can be exploited in devising an MD scheme for eq 8.6, as will be shown below. As was pointed out by Chandler and Wolynes,[116] eqs 8.6 and 8.7 together show that, for finite P, the path integral of a single quantum particle is isomorphic to a classical system of P particles subject with a Hamiltonian given by eq 8.7. Inspection of eq 8.5 shows that the P particles form a closed polymer chain with nearest neighbor harmonic coupling and are subject to a potential U. The classical isomorphism allows molecular dynamics to be used to simulate a finite-temperature quantum system. The extension of the path integral scheme to N particles in three dimensions is straightforward if it is assumed that the particles obey Boltzmann statistics, i.e., all spin statistics are neglected. In this case, the partition function is

$$Q_P = N\int \prod_{i=1}^{N} \mathrm{d}^P\mathbf{r}_i\, \mathrm{d}^P\mathbf{p}_i\, \mathrm{e}^{-\beta H} \tag{8.8}$$

where the classical Hamiltonian is given by

$$H = \sum_{s=1}^{P}\left\{\sum_{i=1}^{N}\left[\frac{(\mathbf{p}_i^{(2)})^2}{2m_i'} + \frac{1}{2}m_i\omega_P^2(\mathbf{r}_i^{(s)} - \mathbf{r}_i^{(s+1)})^2\right] + \frac{1}{P}U(\mathbf{r}_1^{(s)}, \ldots, \mathbf{r}_N^{(s)})\right\} \tag{8.9}$$

In principle, the equations of motion resulting from eq 8.9 could be implemented as a MD procedure, from which the quantum equilibrium properties of a system could be computed.[117] A number of well-known difficulties arise in a straightforward implementation of MD to the path integral. Primarily, since $\omega_P^2 \sim P$, the force constant of the harmonic coupling increases as P increases, giving rise to a stiff harmonic interaction and a time scale separation. As was shown by Hall and Berne,[118] this time scale separation gives rise to nonergodic trajectories that do not sample the available canonical phase space. A solution to this problem was first presented in ref 34. There, it was shown that several elements are needed to devise an efficient MD scheme for path integrals. First, a change of variables that diagonalizes the harmonic coupling is introduced. This has the effect of isolating the various time scales present in the Hamiltonian of eq 8.9. The change of variables is linear, having the general form

$$\mathbf{u}_i^{(s)} = \sum_{q=1}^{P}\mathbf{U}_{sq}\mathbf{r}_i^{(q)} \tag{8.10}$$

where the matrix $\mathbf{U}$ is a constant matrix of unit determinant. Two different choices of the matrix $\mathbf{U}$, discussed in refs 119 and 120, lead to the *staging* and *normal mode* transformations. The transformed coordinates $\mathbf{u}_i^{(s)}$ are known as staging or normal mode variables. If the change of variables is made in eq 8.8, then the corresponding classical Hamiltonian takes the form

$$H = \sum_{s=1}^{P}\left\{\sum_{i=1}^{N}\left[\frac{(\mathbf{p}_i^{(2)})^2}{2m_i^{(s')}} + \frac{1}{2}m_i^{(s)}\omega_P{}^2(\mathbf{u}_i^{(s)})^2\right] + \frac{1}{P}U(\mathbf{r}_1^{(s)}(\{\mathbf{u}_1\}), ..., \mathbf{r}_N^{(s)}(\{\mathbf{u}_N\}))\right\} \quad (8.11)$$

where the s-dependent masses $m_i^{(s)}$ result from the variable transformation. For a staging transformation, the masses are $m_i^{(1)} = 0$ and $m_i^{(s)} = m_i s/(s-1)$ for $s \geq 2$, while for the normal mode transformation, the masses are proportional to the normal mode eigenvalues. Thus, it is clear that the fictitious masses should be chosen according to $m_i^{(1')} = m_i$ and $m_i^{(s')} \propto m_i^{(s)}$. In this way, all modes will move on the same time scale, leading to maximally efficient exploration of the configuration space.

In addition to variable transformations, it is necessary to ensure that a canonical phase space is generated. This can be achieved via one of the non-Hamiltonian MD schemes for generating the *NVT* ensemble. It has been found that maximum efficiency is obtained if each Cartesian direction of each mode variable is coupled to its own thermostat, as was clearly demonstrated in ref 119, and multiple time scale integration techniques are employed.[34]

It is worth mentioning that the path integral MD scheme outlined here has been combined with ab initio MD to yield an ab initio path integral Car−Parrinello method.[119,121] This allows quantum effects on chemical processes to be studied. More recently, the ab initio path integral scheme has been extended to incorporate approximate quantum dynamical properties[122] via the so-called centroid dynamics method.[123,124] Finally, the path integral MD scheme has been modified to allow path integral simulations under conditions of constant temperature and pressure to be carried out.[120]

8.1. Ab Initio Path Integral Simulation of an Excess Proton in Water. Excess protons in liquid water have an anomalously high mobility. Explanations of this phenomenon began with the idea of structural diffusion proposed by de Grotthuss nearly 2 centuries ago.[125,126] In the Grotthuss picture, an overall motion of excess charge occurs via a chain of proton transfer reactions through the interconnected hydrogen bond network of water. Although the Grotthuss concept is generally believed to be the correct explanation of the high mobility of protons in water, the exact Grotthuss mechanism has remained a mystery for 200 years. Moreover, as a result of the high mobility, identifying the solvation structure of the hydrated proton has led to controversy. The classic freshman chemistry picture identifies the excess proton as attached to a water molecule, forming the hydronium ion, a picture that is generally agreed upon. However, the solvation structure of hydronium in water remains unresolved. Within the last 50 years, two predominant structural models emerged. Eigen[127,128] proposed the formation of a $H_9O_4^+$ complex in which the H_3O^+ core is strongly hydrogen bonded to three water molecules. If Eigen's model is correct, then the Grotthuss mechanism would be characterized by tunneling from one $H_9O_4^+$ complex state to another. Zundel,[129] on the other hand, based on certain broad features of spectroscopic data, proposed the existence of an $H_5O_2^+$ complex in which the proton is shared between two water molecules. Clearly, these two structural models lead to different interpretations of the Grotthuss picture of proton transport.

The combination of ab initio molecular dynamics and path integrals[119,121] provides a powerful tool for probing the controversial details of the solvation and transport of hydronium in water including nuclear quantum effects. Previous studies of D^+ in liquid D_2O using ab initio MD and classical nuclei revealed that $H_9O_4^+$ and $H_5O_2^+$ occur with roughly equal probability.[16−18] In addition, these studies revealed that proton transport is driven by coordination fluctuations of first solvation shell members of H_3O^+. When the coordination of a first shell water decreases from four to three as a result of the breaking of a hydrogen bond with a second solvation shell member, the undercoordinated water becomes a proton acceptor from H_3O^+. The $H_5O_2^+$ forms as an intermediate complex state during the proton transfer reaction. This mechanism was discovered independently in ref 130 and was shown to be consistent with available experimental data. A very rough estimate of the proton transfer rate from the ab initio MD trajectories gave a result of approximately 2 ps. Scaling this estimate by $\sqrt{M_H/M_D}$, therefore, yields an approximate proton transfer rate of 1.4 ps, in accordance with NMR measurements.[131]

In order to probe both thermal and quantum fluctuations of the hydrated proton, ab initio path integral simulations have been carried out on a system of 32 water molecules with one excess proton.[19] The imaginary time paths were discretized into $P = 8$ time slices, and trajectories consisting of 100 000 steps with a time step of 7.0 au were generated. Exchange and correlation were treated within the GGA (see discussion in section 7.3) using the B-LYP[78,79] functional, and the plane−wave basis was cut off at an energy of 70 Ry. Core electrons were treated using the Troullier−Martins pseudopotentials.[111] This particular DFT scheme has been shown to yield a good description of the water dimer,[132] liquid water,[14] and water dissociation.[103] Nosé−Hoover chain thermostats (see section 5)[33] were employed to maintain a temperature of 300 K. In addition to the path integral simulation, a corresponding ab initio MD simulation with classical nuclei ($P = 1$) was carried out. These simulations[19] were carried out using CPMD Version 3.0 (J. Hutter, P. Ballone, M. Bernasconi, P. Focher, E. Fois, S. Goedecker, D. Marx, M. Parrinello, M. Tuckerman).

The solvation structures of H_3O^+ in water can be analyzed by examining the two-dimensional probability distribution function $P(R_{OO}, \delta)$ of the shared proton in a hydrogen bond. The coordinate $\delta = R_{O_a}H - R_{O_b}H$ is the difference in distance between the proton and each of the two oxygens, i.e., proportional to the asymmetric stretch, and R_{OO} is the oxygen−oxygen separation. In order that the analysis be as unbiased as possible, this distribution function is analyzed in three stages. First, $P(R_{OO}, \delta)$ is computed for all hydrogen bonds, The distribution (not shown) is characterized by two high peaks at $(R_{OO}, \delta) = (2.8$ Å, 0.9 Å), which arise from hydrogen bonds between neutral water molecules. The distribution, however, also has a nonzero value around $|\delta| \approx 0$, which indicates the existence of centrosymmetric $H_5O_2^+$ complexes in which the excess proton is shared between two water molecules. Thus, a description solely in terms of H_3O^+ or H_9O^+ must be ruled out.

In the second stage, the analysis is refined by excluding all "irrelevant" hydrogen bonds. Irrelevant hydrogen bonds are defined to exist between neutral water molecules. In order to identify these irrelevant bonds, the defect site, H_3O^+, must be located. This can be accomplished uniquely by listing, for each hydrogen, the oxygen atom that is closest and then determining which oxygen atom appears three times in the list of oxygen atoms. Then, only the three hydrogen bonds involving this oxygen are considered. The corresponding distribution is shown in Figure 10a. The effect of excluding the irrelevant hydrogen bonds is to enhance the $|\delta| = 0$ corresponding to the centrosymmetric complex contribution and also to bring out two wings due to asymmetric hydrogen bonds.

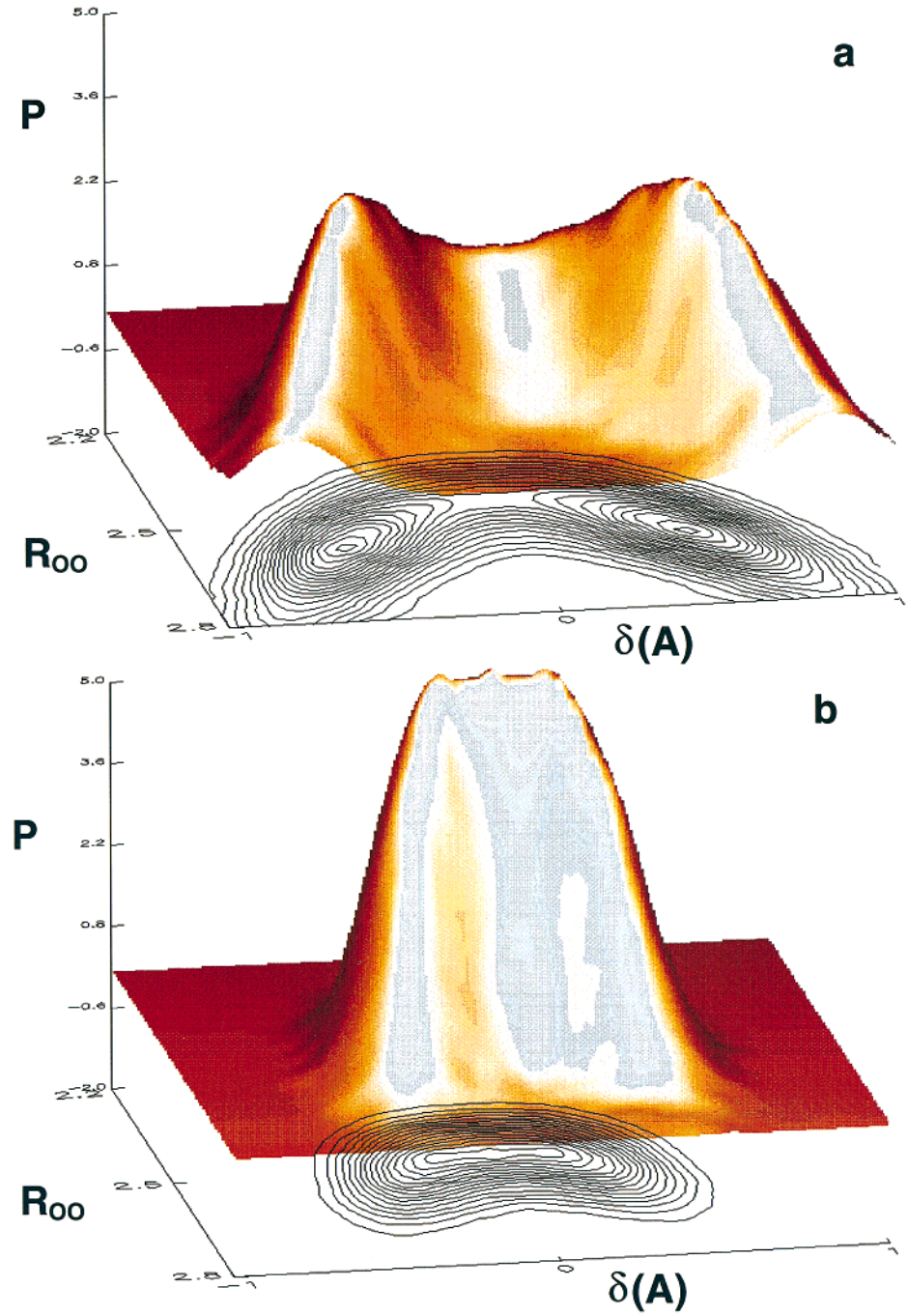

Figure 10. (a) The quantum mechanical two-dimensional probability distribution function $P(R_{OO}, \delta)$ for the three hydrogen bonds of the H_3O^+ complex in water. (b) The quanum mechanical two-dimensional probability distribution function $P(R_{OO}, \delta)$ for the "most active" of the three hydrogen bonds described in part a, as characterized by the smallest value of δ.

In the final refinement step, the distribution corresponding to the hydrogen bond with the smallest value of $|\delta|$ is computed. In this way, the "most active" hydrogen bond, i.e., the hydrogen bond through which proton transfer is most likely to occur, is isolated. The distribution, shown in Figure 10b, exhibits a broad, unstructured character. The unstructured nature of the distribution is a clear indication that an unambiguous identification of solely $H_9O_4^+$ or $H_5O_2^+$ complexes cannot be made. Rather, it is observed that for $|\delta|$ small, the complex corresponds to an equal sharing of the proton between two waters, in accordance with Zundel's view, and for $|\delta|$ large, the complex possesses the features associated with Eigen's $H_9O_4^+$ complex picture. Note, however, that between large and small $|\delta|$ values, the distribution function is characterized by a featureless, flat ridge, which indicates that essentially a continuum of other unclassified structures exists between these two "limiting" forms. In fact, the free energy profile, $F(\delta)$, obtained by integrating the distribution in Figure 10b over R_{OO} values and taking $-kT$ times the logarithm of the result, possesses a single, flat well characteristic of a fluxional complex (see Figure 11). This also implies that the $H_5O_2^+$ complex cannot be regarded as a typical transition state. Rather, the protonic defect complex is most accurately described as being of a fluxional nature, with Eigen's $H_9O_4^+$ and Zundel's $H_5O_2^+$ complexes as its limiting forms. Also shown in Figure 11 is the free energy profile computed from the classical trajectory. It can be seen that a small free energy barrier of ~0.56 kcal/mol exists (at 300 K, $kT \approx 0.59$ kcal/mol), which is washed out in the quantum case by zero-

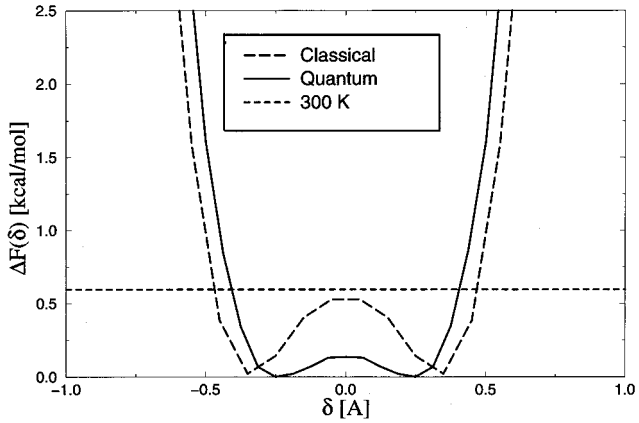

Figure 11. The free energy profiles along the proton transfer coordinate δ for the classical (dashed line) and quantum (solid line) simulations of an excess proton in water. The short dashed line corresponds to 300 K.

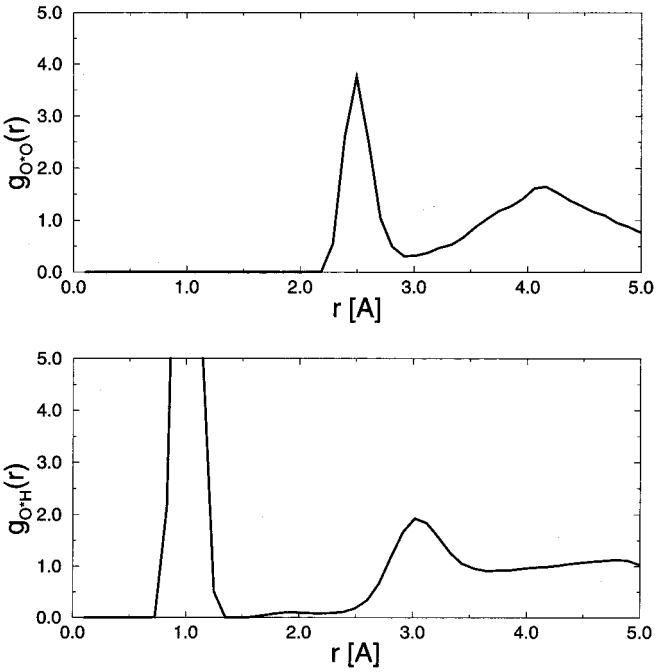

Figure 12. The O*−O (a) and O*−H (b) quantum radial distribution functions for an excess proton in water. Here, O* indicates the defect oxygen site.

point motion. Thus, an important manifestation of nuclear quantum effects is the destabilization of the $H_9O_4^+$ complex by removal of the free energy barrier.

The mechanism of proton transfer described in refs 17 and 18 was also investigated. Since dynamical properties are not available from these studies, an indirect approach was taken. The coordination number of the receiving water molecule in the hydrogen bond with the smallest $|\delta|$ was computed for different ranges of $|\delta|$. It was found that the coordination of this water, which is a first solvation shell water, decreased steadily from 4 at large $|\delta|$ to 3.5 at small $|\delta|$, indicating that the fluxional complex takes on the centrosymmetric $H_5O_2^+$ character when the receiving water is undercoordinated. This is in accordance with the mechanism obtained from our previous calculations and with that described in ref 130.

Finally, the quantum radial distribution functions $g_{O^*O}(r)$ and $g_{O^*H}(r)$ are shown in parts a and b of Figure 12, respectively, where *O** indicates the defect oxygen site. The first peak in the O*O distribution function at $r = 2.5$ Å shows clearly that the defect is characterized by strong, short hydrogen bonds. The first peak in the OO distribution function for water is at $r = 2.8$ Å. Note also the broad shoulder in the O*H distribution function between $r = 1.6$ Å and $r = 2.5$ Å. This, again, gives further evidence of the existence of a fluxional complex. This particular feature is also present in the classical O*H radial distribution function; however, it is somewhat less pronounced.

The analysis presented in the context of this ab initio path integral study underscores the complex behavior of the hydrated proton, indicating that both the Eigen and Zundel pictures are important but that the defect complex is actually of a fluxional character and cannot be understood entirely within either of these views. Typically, we tend to think of solution complexes in terms of well-defined solvation structures. The case of the hydrated proton serves to show that this concept can be misleading and suggests the possibility that many of the difficulties in the interpretation of experimental data might have arisen from this imposition of such a prejudice.

9. Conclusion

Recent developments in molecular dynamics methodology have been presented together with representative applications. The basis of MD in terms of Newtonian or Hamiltonian equations of motion has been reviewed. In addition, a theoretical statistical mechanical treatment has been developed which extends the Hamiltonian case to non-Hamiltonian dynamical systems. The latter can be used to generate statistical ensembles other than the microcanonical within the framework of a continuous dynamics. Thus, all the techniques for analyzing dynamical systems (e.g. conservation laws, stability analysis) can be applied. Dynamical systems capable of generating the canonical (*NVT*) and isothermal−isobaric (*NPT*) ensembles have been presented. Next, it was shown how the classical propagator expressed in terms of the Liouville operator can be used to derive stable, reversible numerical integration procedures for both Hamiltonian and non-Hamiltonian dynamical equations, including decompositions for systems with multiple time scale motion. In particular, it has been shown how multiple time step factorization schemes lead to more efficient procedures for dynamical systems characterized by a separation of two or more time scales.

The question of how interparticle interactions are computed in molecular dynamics has been addressed. The problem of the choice of boundary conditions has been discussed, and the existence of new methodology to treat zero-, one-, two- and three-dimensional periodicity within the same framework for systems with long range interactions was presented. The basic ideas behind modern force fields were given, and an application of force field model to a problem of biological interest, the mutant T4 lysozyme M61, has been presented, in order to demonstrate the use of force fields with the multiple time scale integration techniques discussed in section 6 and non-Hamiltonian evolution. Since the model force field approach breaks down in reacting chemical systems, the ab initio MD technique was reviewed. It was demonstrated how an efficient scheme for combining "on the fly" electronic structure with finite temperature dynamics could be achieved within the gradient-corrected Kohn−Sham local density functional theory and how the Car−Parrinello adiabatic dynamics approach leads to a stable scheme for carrying out ab initio MD calculations. The results of a recent application of ab initio MD to liquid ammonia were presented. It was seen that gradient-corrected DFT applied to liquid ammonia gives structural properties that are in good agreement with experiment.

Finally, it has been shown how quantum nuclear effects at finite temperature can be studied via Feynman's path integral formulation of quantum statistical mechanics. It was shown how the discrete Feynman path integral for the canonical partition function for an *N*-particle system could be mapped onto that of a classical *NP*-particle system consisting of interacting cyclic ring polymer chains, the latter being amenable to molecular dynamics methods. The classic problems surrounding MD evaluation of path integrals have been discussed and effective solutions to these problems presented. The combination of path integrals with ab initio MD was briefly discussed, and an application of the combined ab initio path integral scheme to the classic problem of the hydrated proton was presented. A number of critical conclusions were drawn from the analysis that could potentially resolve the 200 year-old questions of how does an excess proton in water solvate and what is the nature of the Grotthuss mechanism?

Molecular dynamics, based either on empirical force fields or "on the fly" ab initio-derived forces, is a powerful tool, both as a means of conformational sampling and for obtaining dynamical properties of systems. The flexibility of the method permits a good deal of creativity in the development of novel MD algorithms for different purposes. The authors are currently exploiting this flexibility to further extend multiple time scale integration methodology along the lines first introduced in ref 133 and to develop efficient MD-based conformational sampling techniques for biological macromolecules based on variable transformation techniques. Promising avenues for future development include the combining of multiple time scale methodology with the recently introduced transition path sampling method[134,135] for the calculation of rate constants and the development of multiple time scale hybrid Monte Carlo[136] algorithms for efficient conformational sampling. It is expected that, as the power of readily obtainable computational resources increases and new algorithms are devised, MD will start to play a greater role in the areas of drug design and screening and in investigations of the properties of novel materials. It is also expected that ab initio MD modules will soon appear as a feature in molecular modeling packages and, hence, will eventually become a standard tool in many experimental labs.

Acknowledgment. M.E.T. would like to acknowledge his collaborators, J. Hutter, D. Marx and M. Parrinello, on the proton in water project, the computational resources of the Max-Planck Institute in Stuttgart, and funding from PRF 33256-G and the Research Corporation RI0218. G.J.M. would like to acknowledge support by NSF CHE 95-5015 and PRF AC 32139.

References and Notes

(1) Fermi, E.; Pasta, J.; Ulam, S. *Los Alamos preprint* **1955**, *LA-1940*.
(2) Fermi, E. *Collected Papers* **1965**, *II*, 978.
(3) Alder, B. J.; Wainwright, T. E. *J. Chem. Phys.* **1957**, *26*, 1208.
(4) Alder, B. J.; Wainwright, T. E. *J. Chem. Phys.* **1959**, *31*, 459.
(5) Rahman, A. *Phys. Rev.* **1964**, *136A*, 405.
(6) Verlet, L. *Phys. Rev.* **1967**, *159*, 98.
(7) Duan, Y.; Kollman, P. A. *Science* **1998**, *282*, 740.
(8) Zhong, Q.; Moore, P. B.; Newns, D. M.; Klein, M. L. *FEBS Lett.* **1998**, *427*, 267.
(9) Zhong, Q.; Jiang, Q.; Moore, P. B.; Newns, D. M.; Klein, M. L. *Biophys. J.* **1998**, *74*, 3.
(10) Mi, H.; Tuckerman, M. E.; Schuster, D. I.; Wilson, S. R. *Electrochem. Soc. Proc.* **1999**, *99*, 256.
(11) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471−2474.
(12) Galli, G.; Parrinello, M. *Comput. Simul. Mater. Sci.* **1991**, *3*, 283.
(13) Laasonen, K.; Sprik, M.; Parrinello, M.; Car, R. *J. Chem. Phys.* **1993**, *99*, 9081.
(14) Sprik, M.; Hutter, J.; Parrinello, M. *J. Chem. Phys.* **1996**, *103*, 1142.
(15) Diraison, M.; Martyna, G. J.; Tuckerman, M. E. *J. Chem. Phys.* **1999**, *110*, 000.
(16) Tuckerman, M. E.; Laasonen, K.; Sprik, M.; Parrinello, M. *J. Phys. Condens. Matter* **1994**, *4*, A93.
(17) Tuckerman, M. E.; Laasonen, K.; Sprik, M.; Parrinello, M. *J. Phys. Chem.* **1995**, *99*, 5749.
(18) Tuckerman, M. E.; Laasonen, K.; Sprik, M.; Parrinello, M. *J. Chem. Phys.* **1995**, *103*, 150.
(19) Marx, D.; Tuckerman, M. E.; Hutter, J.; Parrinello, M. *Nature* **1999**, *397*, 601.
(20) Benoit, M.; Marx, D.; Parrinello, M. *Nature* **1998**, *392*, 258.
(21) Sagnella, D. E.; Laasonen, K.; Klein, M. L. *Biophys. J.* **1996**, *71*, 1172.
(22) M. Boero, M. P.; Terakura, K. *J. Am. Chem. Soc.* **1998**, *120*, 2746.
(23) K. C. Hass, W. F. S. A. C.; Andreoni, W. *Science* **1998**, *282*, 265.
(24) Tuckerman, M. E.; Mundy, C. J.; Martyna, G. J. *Europhys. Lett.* **1999**, *45*, 149.
(25) Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. *J. Chem. Phys.* **1982**, 76, 637.
(26) Kubo, R.; Toda, M.; Hashitsume, N. *Statistical Physics II*; Springer-Verlag: Berlin, 1978.
(27) Berne, B. J.; Harp, G. D. *Adv. Chem. Phys.* **1970**, *17*, 63.
(28) Andersen, H. C. *J. Chem. Phys.* **1980**, *72*, 2384.
(29) Nosé, S.; Klein, M. *Mol. Phys.* **1983**, *50*, 1055.
(30) Nosé, *S. J. Chem. Phys.* **1984**, *81*, 511.
(31) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695.
(32) Tuckerman, M. E.; Martyna, G. J.; Ciccotti, G. Manuscript in preparation.
(33) Martyna, G. J.; Tuckermana, M. E.; Klein, M. L. *J. Chem. Phys.* **1992**, *97*, 2635.
(34) Tuckerman, M. E.; Martyna, G. J.; Klein, M. L.; Berne, B. J. *J. Chem. Phys.* **1993**, *99*, 2796.
(35) Tobias, D. J.; Martyna, G. J.; Klein, M. L. *J. Phys. Chem.* **1993**, *97*, 12959.
(36) Martyna, G. J.; Tobias, D. J.; Klein, M. L. *J. Chem. Phys.* **1994**, *101*, 4177.
(37) Tuckerman, M. E.; Martyna, G. J. *J. Chem. Phys.* **1999**, *110*, 3623.
(38) Goldstein, H. *Classical Mechanics*, 2nd ed.; Addison-Wesley: Reading, MA, 1980.
(39) Tuckerman, M.; Martyna, G. J.; Berne, B. J. *J. Chem. Phys.* **1992**, *97*, 1990.
(40) Martyna, G. J.; Tuckerman, M.; Tobias, D.; Klein, M. *Mol. Phys.* **1996**, *87*, 1117.
(41) Andersen, H. C. *J. Comput. Phys.* **1983**, *52*, 24.
(42) Berne, B. J. *J. Stat. Phys.* **1986**, *43*, 911.
(43) Cao, J.; Voth, *G. A. J. Chem. Phys.* **1994**, *104*, 273.
(44) Friedman, A.; Auerbach, S. *J. Comp. Phys.* **1991**, *93*, 171.
(45) Watanabe, M.; Karplus, M. *J. Chem. Phys.* **1993**, *99*, 18063.
(46) Procacci, P.; Darden, T.; Marchi, M. *J. Phys. Chem.* **1996**, *100*, 10464.
(47) Mizan, T. I.; Savage, P. E.; Ziff, R. M. *J. Phys. Chem.* **1994**, *98*, 13067.
(48) Cui, S. T.; Cummings, P. T.; Cochran, H. D. *J. Chem. Phys.* **1996**, *104*, 255.
(49) Komeiji, Y.; Uebayasi, M.; Takata, R.; Shimizu, A.; Itsukashi, K.; Taiji, M. *J. Comput. Chem.* **1997**, *18*, 1546.
(50) Schlick, T.; Mandzuik, M. *Chem. Phys. Lett.* **1995**, *237*, 525.
(51) Schlick, T.; Mandzuik, M.; Skeel, R. D.; Srinivas, K. *J. Comput. Phys.* **1998**, *140*, 1.
(52) Tuckerman, M. E.; Parrinello, M. *J. Chem. Phys.* **1994**, *101*, 1302.
(53) Williams, D. E. *Acta Crystallogr. A* **1971**, *27*, 452.
(54) Hockney, R.; Eastwood, J. *Computer Simulation using Particles*; McGraw-Hill: New York, 1981.
(55) Pollock, E.; Gloshi, J. *Comput. Phys. Commun.* **1996**, *95*, 93.
(56) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T. A.; Lee, H. L.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
(57) Darden, T. A.; Toukmaji, A.; Pedersen, L. G. *J. Chim. Phys.* **1997**, *94*, 1346.
(58) Greengard, L.; Rokhlin, V. *Phys. Scripta* **1989**, *29A*, 139.
(59) White, C.; Head-Gordon, M. *J. Chem. Phys.* **1994**, *101*, 6593.
(60) Figueirido, F.; Zhou, R.; Levy, R.; Berne, B. *J. Chem. Phys.* **1997**, *106*, 9835.
(61) Martyna, G. J.; Tuckerman, M. E. *J. Chem. Phys.* **1999**, *110*, 2810.
(62) Martyna, G. J.; Tuckerman, M. E. Manuscript in preparation.
(63) MacKerell, A., Jr..; Bashford, D.; Bellott, M.; Dumbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kcuhnir, L.; Kuczera, K.; Lau, F.: Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; III, W. E. R.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.
(64) Cornell, W. D.; Cieplak, P.; Bayly, C, I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
(65) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

(66) Faber, H.; Matthews, B. *Nature* **1990**, *348*, 263.
(67) Hayward, S.; Berendsen, H. *Proteins: Struct., Funct., and Genet.* **1998**, *30*, 144−154.
(68) de Groot, B.; Hayward, S.; van Aalten, D.; Amadei, A.; Berendsen, H. *Proteins: Struct., Funct., Genet.* **1998**, *31*, 116−127.
(69) Hayward, S.; Kitao, A.; Berendsen, H. *Proteins: Struct., Funct., Genet.* **1997**, *27*, 425−437.
(70) Arnold, G.; Ornstein, R. *Biopolymers* **1997**, *41*, 533−544.
(71) Arnold, G.; Manchester, J.; Townsend, B.; Ornstein, R. *J. Biomol. Struct. Dynam.* **1994**, *12*, 457.
(72) Mchaourab, H.; Oh, K.; Fang, C.; Hubbell, W. *Biochemistry* **1997**, *36*, 307−316.
(73) Tuckerman, M. E.; Yarne, D. A.; Samuelson, S. D.; Hughes, A. L.; Martyna, G. J. *Comp. Phys. Comm.*, submitted.
(74) Banks, J. L.; Kaminski, G. A.; Zhou, R. H.; Mainz, D. T.; Berne, B. J.; Friesner, R. A. *J. Chem. Phys.* **1999**, *110*, 1999.
(75) Hohenberg, P.; Kohn, W. *Phys. Rev. B* **1964**, *136*, 864.
(76) Kohn, W.; Sham, L. J. *Phys. Rev. A* **1965**, *140*, 1133.
(77) Parr, R. G.; Yang, W. *Density Functional Theory of Atoms and Molecules*; Oxford University Press: New York, 1988.
(78) Becke, A. *Phys. Rev. A* **1988**, *38*, 3098.
(79) Lee, C.; Yang, W.; Parr, R. *Phys. Rev. B* **1988**, 37, 785.
(80) Perdew, J. P.; Wang, Y. *Phys. Rev. B.* **1991**, *45*, 13244.
(81) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.
(82) Proynov, E. I.; Salahub, D. R. *Phys. Rev. B* **1994**, *249*, 7874.
(83) Proynov, E. I.; Salahub, D. R. *Chem. Phys. Lett.* **1994**, *230*, 419.
(84) Focher, P.; Chiarotti, G. L.; Bernasconi, M.; Tosatti, E.; Parrinello, M. *Europhys. Lett.* **1994**, *26*, 345.
(85) Bernasconi, M.; Chiarotti, G. L.; Focher, P.; Scandolo, S.; Tosatti, E.; Parrinello, M. *J. Phys. Chem. Solids* **1995**, *56*, 501.
(86) Mauri, F.; Galli, G. *Phys. Rev. B* **1994**, *50*, 4316.
(87) Hutter, J.; Tuckerman, M. E.; Parrinello, M. *J. Chem. Phys.* **1995**, *102*, 859.
(88) Hutter, J.; Parrinello, M.; Vogel, S. *J. Chem. Phys.* **1994**, *101*, 3862.
(89) Silvestrelli, P. L.; Bernasconi, M.; Parrinello, M. *Chem. Phys. Lett.* **1997**, *277*, 478.
(90) Benoit, M.; Bernasconi, M.; Parrinello, M. *Phys. Rev. Lett.* **1996**, *76*, 2934.
(91) Bernasconi, M.; Silvestrelli, P. L.; Parrinello, M. *Phys. Rev. Lett.* **1998**, *81*, 1235.
(92) Tuckerman, M. E.; von Rosenvinge, T.; Klein, M. L. *Mater. Res. Soc. Symp. Proc.* **1996**, *408*, 477.
(93) von Rosenvinge, T.; Tuckerman, M. E.; Klein, M. L. *Farad. Discuss.* **1997**, *106*, 273.
(94) Tuckerman, M. E.; Klein, M. L. *Chem. Phys. Lett.* **1998**, *283*, 147.
(95) Mei, H. S.; Bagchi, K.; Tuckerman, M. E.; Klein, M. L. *J. Phys. Chem.* **1999**, *104*, *000*.
(96) Curioni, A.; Andreoni, W.; Hutter, J.; Schiffer, H.; Parrinello, M. *J. Am. Chem. Soc.* **1994**, *116*, 11251.
(97) Curioni, A.; Sprik, M.; Andreoni, W.; Schiffer, H.; Hutter, J.; Parrinello, M. *J. Am. Chem. Soc.* **1997**, *119*, 7218.
(98) Laasonen, K.; Klein, M. L. *J. Am. Chem. Soc.* **1994**, *116*, 11620.
(99) Laasonen, K.; Klein, M. L. *Mol. Phys.* **1996**, *88*, 135.
(100) Meijer, E. J.; Sprik, M. *J. Am. Chem. Soc.* **1998**, *120*, 6345.
(101) Mei, H. S.; Tuckerman, M. E.; Sagnella, D. E.; Klein, M. L. *J. Phys. Chem. B* **1998**, *102*, 10446.
(102) Pavese, M.; Berard, D. R.; Voth, G. A. *Chem. Phys. Lett.* **1999**, *300*, 93.
(103) Trout, B. L.; Parrinello, M. *Chem. Phys. Lett.* **1998**, *288*, 343.
(104) Warshel, A.; Weiss, R. M. *J. Am. Chem. Soc.* **1980**, *102*, 6218.
(105) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718.
(106) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700.
(107) Monard, G.; Loos, M.; Thery, V.; Baka, K.; Rivail, J.-L. *Int. J. Quant. Chem.* **1996**, *58*, 153.
(108) Gao, J.; Amara, P.; Alhambra, C.; Field, M. J. *J. Phys. Chem. A* **1998**, *102*, 4714.
(109) Impey, R. W.; Klein, M. L. *Chem. Phys. Lett.* **1984**, *104*, 579.
(110) Bachelet, G.; Hamann, D.; Schluter, M. *Phys. Rev. B* **1982**, *26*, 4199.
(111) Troullier, N.; Martins, J. *Phys. Rev. B* **1991**, *43*, 1993.
(112) Ricci, M. A.; Nardone, M.; Ricci, F. P.; Andreani, C.; Soper, A. K. *J. Chem. Phys.* **1995**, *102*, 7650.
(113) Marzari, N.; Vanderbilt, D. *Phys. Rev. B* **1997**, *56*, 12847.
(114) Feynman, R. P.; Hibbs, A. R. *Quantum Mechanics and Path Integrals*; McGraw-Hill: New York, 1965.
(115) Feynman, R. P. *Statistical Mechanics: A Set of Lectures*; Benjamin Cummings: Reading, 1972.
(116) Chandler, D.; Wolynes, *P.* G. *J. Chem. Phys.* **1981**, *74*, 4078.
(117) Parrinello, M.; Rahman, A. *J. Chem. Phys.* **1984**, *80*, 860.
(118) Hall, R. W.; Berne, B. J. *J. Chem. Phys.* **1984**, *860*, 3641.
(119) Tuckerman, M. E.; Marx, D.; Klein, M. L.; Parrinello, M. *J. Chem. Phys.* **1997**, *104*, 5579.
(120) Martyna, G. J.; Hughes, A.; Tuckerman, M. E. *J. Chem. Phys.* **1999**, *110*, 3275.
(121) Marx, D.; Parrinello, M. *J. Chem. Phys.* **1997**, *104*, 4077.
(122) Marx, D.; Tuckerman, M. E.; Martyna, G. J. *Comput. Phys. Commun.* **1999**, *118*, 000.
(123) Cao, J.; Voth, G. A. *J. Chem. Phys.* **1993**, *99*, 10070.
(124) Cao, J.; Voth, G. A. *J. Chem. Phys.* **1994**, *101*, 6157.
(125) de Grotthuss, C. J. T. *Ann. Chim.* **1806**, *LVIII*, 54.
(126) Atkins, P. W. *Physical Chemistry*, 6th ed.; Oxford University Press: Oxford, 1998.
(127) Wicke, E.; Eigen, M.; Ackerman, T. *Z. Phys. Chem.* **1954**, *1*, 340.
(128) Eigen, M. *Angew. Chem. Int. Edn. Engl.* **1964**, *3*, 1.
(129) Zundel, G.; Metzger, H. *Z. Physik. Chem. (N. F.)* **1968**, *58*, 225.
(130) Agmon, N. *Chem. Phys. Lett.* **1995**, *244*, 456.
(131) Meiboom, S. *J. Chem. Phys.* **1961**, *34*, 375.
(132) Mok, D. K. W.; Handy, N. C.; Amos, R. D. *Mol. Phys.* **1997**, *92*, 667.
(133) Martyna, G. J.; Tuckerman, M. E. *J. Chem. Phys.* **1995**, *102*, 8071.
(134) Dellago, C.; Bolhuis, P. G.; Csajka, F. S.; Chandler, D. *J. Chem. Phys.* **1988**, *108*, 1964.
(135) Dellago, C.; Bolhuis, P. G.; Chandler, D. *J. Chem. Phys.* **1998**, *110*, 6617.
(136) Duane, S.; Kennedy, A. D.; Pendleton, B. J.; Roweth, D. *Phys. Lett. B* **1987**, *195*, 216.