Torsional Path Integral Monte Carlo Method for the Quantum Simulation of Molecular Systems

Thomas Francis Miller

University College,
University of London

M.Phil. thesis submitted for the

Department of Chemistry

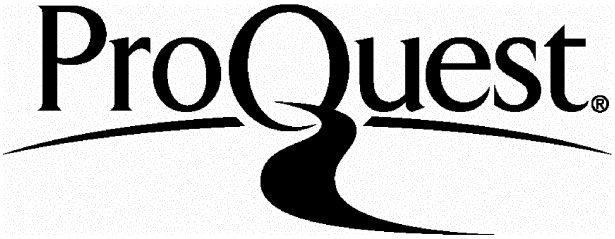
in November 2002

ProQuest Number: U641971

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.



ProQuest U641971

Published by ProQuest LLC(2015). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code.
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

A molecular application is introduced for calculating quantum statistical mechanical expectation values of large molecules at non-zero temperatures. The torsional path integral Monte Carlo (PIMC) technique applies an uncoupled winding number formalism to the torsional degrees of freedom in molecular systems. The internal energy of the molecules ethane, n-butane, n-octane, and enkephalin are calculated at standard temperature using the torsional PIMC technique and compared to the expectation values obtained using the harmonic oscillator approximation and a variational technique. All studied molecules exhibited significant quantum mechanical contributions to their internal energy expectation values according to the torsional PIMC technique. The harmonic oscillator approximation approach to calculating the internal energy performs well for the molecules presented in this study but is limited by its neglect of both anharmonicity effects and the potential coupling of intramolecular torsions.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

The realisation that theoretical methods yield results useful for the understanding of biomolecular activity has caused computational studies of proteins and other large biomolecules to become increasingly common in modern chemistry.[27,55,99,137,151,160,177,198] Theoretical analysis enables the calculation of structural properties, thermodynamic properties, and the electronic structure of a wide range of biochemical systems. Studies regarding the folding and unfolding of proteins in solvent, protein docking at a functional site, and the modelling of ion transfer channels are now published with regularity, and the range of biomolecular reaction processes accessible to theoretical analysis continues to expand in breadth and variety. Also, computational techniques perform a critical role in facilitating the interpretation of incomplete or ambiguous experimental data. Common examples of theoretical contribution to experimental analysis are found in the fields of crystallographic and NMR spectroscopy.[83,84]

A particularly promising direction of research in theoretical biochemistry is the calculation of kinetic and thermodynamic properties of large systems. Thermodynamic quantities, and free energy in particular, describe the tendency of molecular systems to associate and react, whereas kinetic quantities determine the rate at which they do so.[12] Reliable and efficient dynamical simulations would enable the prediction of molecular solvation, ligand binding to proteins and nucleic acids, sequence-dependent stabilities of proteins and nucleic acids, and environmental effects on reactions in solutions.

Most biochemical simulations assume the validity of the Born-Oppenheimer (BO) approximation, in which electrons and nuclei move in different time scales. It is therefore assumed that electrons can be modelled in the environment of fixed nuclei, and nuclei can be modelled in the mean field of the electrons. Application of the BO approximation allows treatment of the electronic part of a system to be separated from treatment of the nuclear motion.[159]

Upon application of the Born-Oppenheimer approximation, biochemical simulation may be broken into two parts: (1) the method for simulating the electronic degrees of freedom or electronic potential energy surface (PES) of the system and (2) the method for simulating the nuclear degrees of freedom or dynamics of the system. Practically all techniques of biochemical simulation involve the application of a potential energy surface and a dynamical technique. The quality of these two factors determines the quality of the calculated results.

Determination of an electronic potential energy surface is the domain of electronic structure theory. For applications on the biological scale, the highest-quality calculations performed in modern quantum chemistry generally utilise density functional theory methods and *ab initio* wavefunction methods based upon the Hartree-

Fock mean-electron-field approximation.[104,135,164,197] Fully quantum treatments that scale linearly with the system size have been developed and applied to biochemical systems.[53,105,108,109,133,158,167,182,183,200] Although impressive in accuracy, these techniques generally require a substantial investment of computational resources to obtain the energy of just a single structure. More affordable electronic structure techniques are based on the eigenfunctions of an experimentally parametrized Hamiltonian operator. These semi-empirical quantum chemistry methods include the AM1, PM3, and MNDO techniques in which orbital interactions are chosen to fit reliable data.[48,49,157]

The most affordable and straightforward potential energy surfaces are generated from molecular mechanics force field calculations. This approach is entirely empirical and does away with solution of the Schrodinger equation. The force field potential is defined as an analytical function of the nuclear coordinates and adjusted to reproduce experimental observations or higher-level calculations. In their simplest implementations, force fields potentials are a sum of uncoupled energy terms that correspond to bond lengths, bond angles, torsional angles, Van der Waals interactions, and electrostatic interaction.[20,40,87,103,144,159] Typically, a harmonic approximation is used for the bond length and bond angle terms and a sinusoidal potential is used for the torsional angles. A Lennard-Jones 6-12 potential is used to describe the Van der Waals interaction and Coulomb's law is used for the electrostatic term. The parameters in the energy terms are individually selected for each possible atom combination, resulting in a great many (usually between 10 and 100) variables that are chosen to perform as well as possible. Prominent examples of molecular mechanics force fields include MM2, MM3, AMBER, CHARMM, YETI, GROMOS, OPLS, and TIP3P.[7–9,25,42,95,110,111,114,115,187,192,194,195] Because each force field is

parametrized in a unique fashion to a unique source of reference data, it is critical for users to exercise great care in selecting a potential for calculations intended to produce numerical accuracy.

It is worthwhile to note that a class of hybrid electronic structure theory techniques is enjoying increasing application to protein systems. These QM/MM techniques treat the majority of the biochemical system with a molecular mechanics force field and use a higher level of theory for the chemically interesting portion.[1,61,62,65,67–69,71,74,76,132,155,156,165,193] These techniques are particularly appropriate for reaction dynamics calculations in which accurate description of the transition state is essential.[75]

In additional to the potential energy surface, a biomolecular simulation requires the utilisation of a dynamical technique. That is, a description of the nuclear motion must be established. Historically, this nuclear component has been treated primarily with classical mechanics.[6,26,101] For most calculations of real-time molecular dynamics in biochemical systems, the nuclear degrees of freedom are propagated via numerical integration of the classical equations of motion. Considerable care must be taken in this process to ensure time reversibility and the conservation of energy in long calculations.[29,81,176] Techniques for the evaluation of classical thermodynamic quantities in biochemical systems include molecular dynamics and Monte Carlo approaches.[101] These methods enable the calculation of ensemble averages by sampling a sufficient portion of the total geometric configuration space. Classical determination of important biological quantities such as the relative free energy of solvation and protein-ligand binding have been performed.[15,16,92,127,199] Other impressive early studies examined biological catalysis and the impact of amino acid sequence on the function and stability of proteins and enzymes.[63,98,143,190,191] Cal-

culations of this sort generally employ free energy perturbation, thermodynamic integration, or slow growth approximations in which free energy differences are repeatedly evaluated for configurations that differ only very slightly along the simulation pathway.[93,94,101,102,136,141,142,146,184–186]

The appearance of quantum mechanics into large-scale molecular simulations began with improvement of the potential energy surface. Primarily for application to reaction dynamics, techniques that coupled semi-empirical or crude quantum mechanical descriptions of the electronic wavefunction with a classical mechanical model of the nuclear motion were first developed.[43,44,61,66,72,73,76] These techniques generally require specification of a reaction coordinate and the evaluation of a potential of mean force (PMF) that is related to the free energy of activation of the ensemble of possible transition states.[35,45–47,70,172–174] The PMF is obtained by averaging over all coordinates along the reaction pathway. Umbrella sampling and free energy perturbation techniques are usually employed in the evaluation of the PMF.[181,201]

Transition state theory techniques based on evaluation of the potential of mean force have been developed in many ways to include the quantum behaviour of nuclear coordinates. In particular, a quantum harmonic oscillator approximation to nuclear vibrations has been employed.[2–4,23,77] Improved calculations of the PMF using the centroid path-integral dynamics method have also been attempted for enzymes.[31,78,88,90,91,113,128,166,189] The centroid path-integral formalism has also been utilised in the quantum-classical path (QCP) method in which a classical free energy calculation is first performed and then adjusted to include quantum mechanics.[57,88,89,188] Finally, new techniques for the extension of quantum description to the entire reaction coordinate have been reported.[3,171] Many of these applications

underline the important role of quantum mechanics in reaction dynamics.[2–5,23,56,77]

Although transition state theory has clearly proven to be an important tool for the quantum analysis of chemical reactivity, it is not a particularly general technique. Defining the reaction coordinate is a necessary and nontrivial inconvenience unique to each application. Molecular dynamics techniques in which the system is propagated in real time is a more general approach to analysing chemical behaviour. As was previously mentioned, real-time molecular dynamics calculations originally employed a molecular mechanics force-field and a classical description of the nuclear motion.[175–177]

Introduction of quantum mechanics to the electronic portion of a real-time molecular dynamics simulation can be achieved with Car-Parrinello molecular dynamics (CPMD).[32,64,145] CPMD combines a plane-wave density functional description of the electronic structure with a classical description of the nuclear motion. Originally developed for solid-state applications, CPMD has yielded impressive results in applications to bulk molecular systems.[13,52] To a limited number of systems, real time molecular dynamics in which a quantum description is provided for the nuclear motion has been performed using centroid path integrals.[31,78,88,90,91,113,128,189] This technique, however, tends to fail for asymmetric potentials at low temperatures.[116,129,131,149]

Finally, several modern techniques have been utilised to calculate equilibrium quantities of general molecular systems using a quantum description of the nuclear motion. For non-zero temperature applications, these techniques employ a path integral formulation of statistical mechanics that establishes an isomorphism between the thermodynamics of a quantum particle and the thermodynamics of a related classical polymer.[38,59] Within the isomorphism, most simulations have been

performed using classical molecular dynamics techniques.[82,118,168] However, path integral molecular dynamics (PIMD) calculations of this sort are generally slowed by the fact that they require force evaluations, and an alternative path integral Monte Carlo approach is available. Nonetheless, the PIMD technique has recently been extended to include a Car-Parrinello description of the electronic portion of the system.[19,121–124,178] These *ab initio* path integral molecular dynamics calculations have led to new insights to the role of quantum mechanics in the nuclear motion of condensed small molecule systems.[125,179,180]

In the aforementioned work, transition state theory studies have shown that inclusion of quantum mechanical effects in the nuclear motion at the transition state is important for large molecules.[2–5,23,56,77] Also, it was mentioned that the *ab initio* PIMD technique has led to new understanding of bulk small molecule dynamics via the inclusion of a quantum description of nuclear motion.[125,179,180] However, few studies have proposed and applied general techniques for the quantum description of nuclear motion to large molecules, so the importance of this factor in thermodynamic quantities remains unclear.[154] This research will focus on the contributions to the thermodynamic internal energy arising from the quantum mechanical behaviour of nuclei.

1.2 Selection of a Dynamical Technique

Because this thesis focuses on obtaining molecular internal energies, it is critical that the theoretical techniques employed are suited to the analysis of equilibrium thermodynamics. For a large molecular system, myriad excited states of the nuclear motion are thermally accessible at room temperature. Explicit calculation of

each excited stated for the extraordinarily complicated potential energy surface of a large molecule would be an impossible task. Thankfully, with the help of statistical mechanics it is also an unnecessary task. The strategy of statistical mechanics is to calculate equilibrium expectation values for a system from the known probability that the system can be found in a particular state. It is assumed, at thermal equilibrium, that a chemical system can be found in a particular state of energy E with probability $\sim e^{-\frac{E}{kT}}$, where T is temperature and k is a constant. It is also assumed that the system will move throughout this state space fast enough that any expectation value for the system will simply be a sum weighted according to these probabilities.

The path integral formulation of quantum mechanics developed by Feynman offers a singularly convenient approach to the calculation of quantum statistical mechanical expectation values.[58–60] The general strategy exploits an observation by Dirac about the relationship between the infinitesimal quantum mechanical transition amplitude and the value of the classical action.[51,153] This path integral relationship between classical and quantum mechanics enables the formulation of exact quantum mechanical expectation values in terms of the classical expectation values of a slightly modified physical system.[38,59]

The general idea behind incorporating quantum mechanics via path integrals is shown schematically in Figure (1.1). Given a particular physical particle, say an electron or atom, classical mechanics requires that both the position and momentum are simultaneously defined. That is, the classical expectation values are calculated while representing the particle in position space as a single well-defined point. However, to calculate the quantum statistical mechanical expectation values of the physical particle within path integral theory, the physical particle is replaced

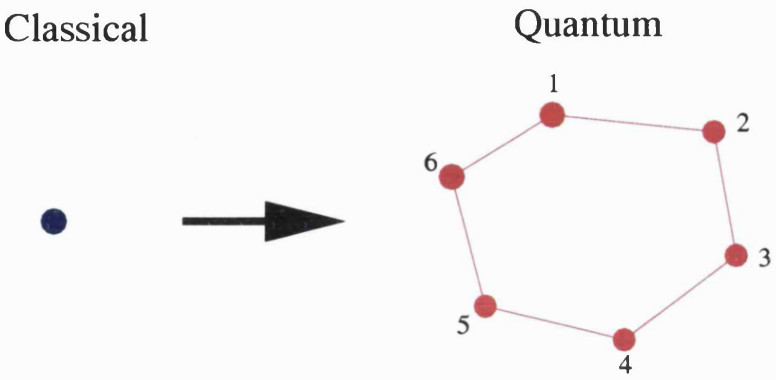


Figure 1.1: Path integral theory incorporates the effects of quantum mechanics by replacing each classical particle in the system by a ring of Trotter beads.

by a ring of classical points, called Trotter beads, and the classical expectation value is again calculated. This change in representation of the physical system, called the *classical isomorphism*, was presented by Chandler and Wolynes.[38] It enables the simulation of the quantum particle with a ring of classical beads. In the limit of many Trotter beads, the *classical isomorphism* becomes a mathematically exact model for the quantum behaviour of the original particle. To obtain a desired level of accuracy, highly quantum mechanical particles must be replaced by a ring with more beads than is required for a less quantum mechanical system.

Qualitatively, the inclusion of quantum effects via such a change of representation makes sense. Rudimentary quantum mechanics immediately dispels the simplistic notion of a particle with finite momentum represented as a point in position space. The Heisenberg Uncertainty Principle states that the certainty with which the momentum and position of a particle can simultaneously be known is bounded from above.[12] That is, for a particle of finite mass and velocity, any reliable quantum mechanical model must incorporate the broadening of the distribution function of the particle in position space. By replacing the quantum particle with a

ring of classical beads, path integral theory introduces this broadening effect in an especially tangible fashion.

The diffusion Monte Carlo (DMC) technique has long been regarded as an alternative to path integrals for the calculation of quantum behaviour in large molecular systems.[10,11,24,28,33,41,54,79,161] This technique has been used to calculate structural properties and the zero-point energy in the quantum ground state for small proteins.[41] However, despite some progress in applying the technique to excited states, DMC is primarily applicable only to the analysis of ground states.[24,41,161] In molecular systems held at biochemically interesting temperatures, many excited states of the nuclear wave function are occupied and potentially important in the evaluation of chemical properties. The inability of the DMC technique to incorporate these excited states curtails its utility for analysing statistical thermodynamic quantities.

Other common approaches to the evaluation of statistical mechanical properties in large molecules are founded on the techniques of molecular dynamics.[175,177] These approaches assume that a sufficiently long trajectory will explore the state space in the same fashion as Monte Carlo sampling. Unfortunately, these techniques tend to be less efficient than path integral Monte Carlo approaches because of the need to calculate forces or potential gradients in order to perform the dynamics.[175]

For the specific objectives of this research, it is expected that the path integral Monte Carlo (PIMC) technique using the quantum statistical mechanics formulation of path integrals described above will be the most practical method to employ. The research in this thesis will focus on the development of the PIMC approach and its applications to biomolecules.

1.3 Selection of a Potential Energy Surface

Having decided upon using the path integral Monte Carlo technique for the calculation of quantum statistical expectation values of molecular systems, the question remains as to how the actual molecules will be represented. We need to select a model for describing the geometry and electronic structure of the molecules.

It was discussed previously that wavefunction methods are accurate and increasingly computationally accessible, but the fact remains that each evaluation of the molecular potential energy for a given set of nuclear positions is a costly endeavour. For statistical sampling techniques, such as the path integral Monte Carlo method we intend to employ in this thesis, thousands and millions of energy evaluations will need to be performed. Therefore, it will be necessary to select a molecular model for which energy evaluations can be executed more quickly.

At the other end of the spectrum of molecular models are those for which large portions of the molecular structure are represented by single units whose interactions are parametrized either to experiment or to more sophisticated calculations. For example, in large biochemical calculations, the protein structure will be defined in terms of its amino acid sequence or in terms of connected units of prescribed secondary structure. Some path integral Monte Carlo calculations on biochemical systems preceding the work in this thesis used this sort of simplified structural model.[50,106,107]

Lee and Berne employ path integrals in the formulation of a quantum thermal annealing technique for locating the global minimum of proteins,[106,107] and Dewey derives a statistical mechanical treatment of large biomolecules utilising the sequence information as an internal coordinate.[50] However, techniques based on such dramatically simplified structural models risk introducing non-physical manifestations while

neglecting subtle, yet important, aspects of reality.

The benefit, of course, of using such simple structural models is that the number of degrees of freedom in the calculations is reduced, and so is the required computational time. However, a more appealing strategy for reducing the number of degrees of freedom in biochemical calculations has been proposed.[41] Upon noting how successful previous calculations had been upon separating the low-frequency and high-frequency motions in quantum mechanical simulations, Clary reduced the coordinate space of an all-atom structural model only to the torsional degrees of freedom.[17,18,28,79,80,112]

In this approach, a standard ball-and-stick molecular mechanics model using a generic parametrized all-atom potential energy surface is employed, but the intra-molecular bond stretches and angle distortions are forbidden. The molecule is held rigid except for twisting motions about the torsional degrees of freedom. These twisting modes are known from basic spectroscopy theory to be the lowest-frequency internal degrees of freedom in the molecule.[86] Furthermore, the torsional modes are expected to be the most important degrees of freedom for biomolecular simulation because their low frequencies facilitate substantial variation in molecular configuration at typical temperatures.[41] Therefore, the inclusion of an all-atom structural model with a reduced torsional configuration space is an appropriate compromise between chemical accuracy and computational feasibility. It will be the structural model employed in this research.

1.4 This Research

The objective of this thesis is to apply the path integral Monte Carlo technique to torsional degrees of freedom of large molecular systems for the purpose of calculating quantum mechanical internal energies at non-zero temperatures. All path integral Monte Carlo calculations, variational calculations, and harmonic oscillator calculations reported in this study are performed with code written by the author in the c programming language. The only codes utilised throughout the entire study that were not written by the author are for the single-point energy calculations and the coordinate transformations between internal- and xyz-coordinates, both of which were obtained from the TINKER molecular mechanics program.[139]

A theoretical obstacle emerges upon the application of the path integral technique to systems, such as torsions, that exhibit periodic boundary conditions. Although PIMC has never, to our knowledge, been applied to the torsions of molecular bonds, numerous path integral Monte Carlo studies have been performed on the free rotation of small molecules and the rotation of small molecules on surfaces.[30,119,120,147] It emerges from these studies that an additional factor, termed the winding number, must be included if an accurate path integral Monte Carlo result is to be obtained.[119,152] Another primary objective of this research is to investigate the contribution of the PIMC winding number term to molecular torsions and evaluate its importance.

Chapter 2

Translational Path Integral Theory

2.1 Introduction

The path integral Monte Carlo (PIMC) method combines the distinct techniques of path integral theory and Monte Carlo theory.[14,21,22,37,38] In this chapter, the ideas regarding path integrals previously discussed in conjunction with Figure (1.1) are formalised and developed. The contribution of Monte Carlo theory to the PIMC method is largely numerical, and discussion of Monte Carlo techniques are put off until the following chapter.

Much like the Schrodinger equation or the Heisenberg equation, Feynman's path integrals are an independent and complete formulation of quantum mechanics.[58,60] However, with an eye to our objectives, the formalism is presented within the framework of statistical mechanics. The chapter begins with a postulated definition of the density matrix operator and knowledge of Boltzmann statistics. It concludes with a computationally useful expression for the canonical partition function and a

strategy for the calculation of quantum mechanical ensemble averages.

In this chapter, we consider only a single quantum mechanical particle not constrained by periodic boundary conditions. Although application to large molecular systems will require a many-particle torsional formalism, discussion of this extension is delayed until Chapter 4. For the time being, the physical scenario will be simplified as much as possible.

2.2 Theoretical Basics

Any quantum system is completely described by a density matrix operator ρ such that

$$\rho = \sum_i w_i \left|\phi_i\right\rangle \left\langle\phi_i\right| \tag{2.1}$$

and the following four conditions are obeyed.[59]

1. The set $|\phi_i\rangle$ is a complete orthonormal set of vectors.
2. $w_i \geq 0$.
3. $\sum_i w_i = 1$.
4. The expectation value of an operator A is $\langle A\rangle = \text{Tr } \rho A$.

Upon observing that

$$\begin{aligned} \langle A\rangle &= \text{Tr } \rho A = \sum_i \left\langle\phi_i\right| \rho A \left|\phi_i\right\rangle \\ &= \sum_i \sum_j w_j \left\langle\phi_i|\phi_j\right\rangle \left\langle\phi_j\right| A \left|\phi_i\right\rangle = \sum_i w_i \left\langle\phi_i\right| A \left|\phi_i\right\rangle , \end{aligned} \tag{2.2}$$

it follows from conditions (2) and (3) that the w_i can be interpreted as the probability that the system is in state ϕ_i.

However, the probability that a system with Hamiltonian H can be found in the state $|\phi_i\rangle$ corresponding to eigenvalue E_i is

$$\frac{1}{Q}e^{-\beta E_i} \tag{2.3}$$

where $Q = \sum_i e^{-\beta E_i}$ and $\beta = \frac{1}{kT}$ is the partition function of the system. Therefore, for such a system, the density matrix operator is

$$\rho = \frac{1}{Q}\sum_i e^{-\beta E_i} |\phi_i\rangle \langle\phi_i| . \tag{2.4}$$

Since $H|\phi_i\rangle = E_i|\phi_i\rangle$, it follows that

$$\rho = \frac{1}{Q}\sum_i e^{-\beta H} |\phi_i\rangle \langle\phi_i| = \frac{e^{-\beta H}}{Q} \tag{2.5}$$

and

$$Q = \sum_i e^{-\beta E_i} = \sum_i e^{-\beta E_i} \langle\phi_i|\phi_i\rangle = \sum_i \langle\phi_i| e^{-\beta H} |\phi_i\rangle = \text{Tr } e^{-\beta H}. \tag{2.6}$$

That is, the partition function Q is written as the trace of the unnormalised density matrix operator. For systems described in x-coordinate space, the position trace of the unnormalised density matrix operator assumes the form

$$Q = \text{Tr } e^{-\beta H} = \int dx \langle x| e^{-\beta H} |x\rangle = \int dx \rho(x, x, \beta) . \tag{2.7}$$

The position matrix element $\rho(x, x', \beta)$ of the unnormalised density matrix operator will henceforth be referred to simply as the density matrix element. We can now proceed to develop the single-particle path integral formalism.

2.2.1 Property of Density Matrix Elements

Before moving on, we pause briefly to explore a very useful property of density matrix elements. Because the Hamiltonian trivially commutes with itself, it follows

that

$$\rho(x, x'; \beta) = \langle x| e^{-\beta H/2} e^{-\beta H/2} |x'\rangle . \tag{2.8}$$

Finally, using the completeness property $\int dx''|x'' >< x''| = 1$, we obtain the relationship

$$\rho(x, x'; \beta) = \int dx'' \langle x| e^{-\beta H/2} |x''\rangle \langle x''| e^{-\beta H/2} |x'\rangle . \tag{2.9}$$

This crucial property illustrates that a density matrix element can be written in terms of a linked pair of density matrix elements corresponding to twice the temperature. It will be very useful in our derivation of path integrals.

2.2.2 Free Particle Density Matrix Element

It will be illustrative at this point and useful shortly to explicitly calculate the density matrix element for a particle in one dimension experiencing zero potential. For this case, the Hamiltonian H is simply equivalent to the kinetic energy operator

$$H = \frac{p^2}{2m} = \frac{-\hbar^2}{2m}\frac{\partial^2}{\partial x^2} \tag{2.10}$$

where m is the mass of the particle. We assume initially that the particle is confined within a one dimensional box over the interval $[-L/2, L/2]$. For finite L, the system exhibits discrete states

$$\langle x|\phi_n\rangle = \frac{1}{\sqrt{L}} e^{ik_n x}, \quad k_n = \frac{2\pi n}{L} \tag{2.11}$$

which correspond to energies

$$E_n = \frac{\hbar^2}{2m} k_n^2. \tag{2.12}$$

We can then write the following expression for the free particle density matrix element:

$$\begin{aligned}\rho_0(x,x',\beta) &= \langle x|\, e^{-\beta H}\, |x'\rangle \\ &= \sum_n \sum_{n'} \langle x|\phi_n\rangle \left\langle \phi_n \left| e^{-\beta H} \right| \phi_{n'} \right\rangle \langle \phi_{n'}|x'\rangle \\ &= \sum_n \frac{1}{L} e^{ik_n(x-x')} e^{-\frac{\beta\hbar^2 k_n^2}{2m}}. \end{aligned} \tag{2.13}$$

Now, allowing the length L to extend so that $L \to \infty$, the discrete wave numbers k_n approach the continuous variable k, and

$$dk \approx \Delta k = k_{n+1} - k_n = \frac{2\pi}{L}. \tag{2.14}$$

By rewriting Equation (2.13), we then arrive at the analytical expression for the density matrix element of the free particle

$$\begin{aligned}\rho_0(x,x',\beta) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\; e^{-ik(x-x')} e^{-\frac{\beta\hbar^2 k^2}{2m}} \\ &= \left[\frac{m}{2\pi\beta\hbar^2}\right]^{\frac{1}{2}} e^{-\frac{m(x-x')^2}{2\beta\hbar^2}}. \end{aligned} \tag{2.15}$$

2.3 Single-Particle Translational Path Integrals

As will be seen in the subsequent derivation, the path integral technique is remarkable for its suitability to problems of quantum statistical mechanics.[59] Our derivation of path integrals proceeds largely by analogy to classical statistical mechanics. We begin with the expression for the canonical partition function from Equation(2.7),

$$Q(\beta) = \int dx\; \langle x|\, e^{-\beta H}\, |x\rangle = \int dx \rho(x,x,\beta) \tag{2.16}$$

where $\beta = 1/kT$. Throughout this entire chapter, all integrals run from $-\infty$ to ∞. Equation (2.16) utilises the density matrix element $\rho(x, x', \beta)$ introduced in Equation (2.7). It is clear that $Q(\beta)$ is the position trace of the Boltzmann-like operator $e^{-\beta H}$ and is a function only of temperature via β. This simple functional dependence of the partition function on temperature is a feature that is preserved throughout the development of path integrals and exemplifies their applicability to quantum statistical mechanics.

The following property was proven in Section (2.2.1):

$$\rho(x, x'; \beta) = \int dx'' \rho(x, x''; \beta/2)\, \rho(x'', x'; \beta/2)\,. \tag{2.17}$$

By inductively applying Equation (2.17) to Equation (2.16), we arrive at a discretized form of the canonical partition function

$$Q(\beta) = \prod_{m=1}^{P} [\int dx_m] \prod_{t=1}^{P} \rho(x_t, x_{t+1}; \beta/P) \tag{2.18}$$

where $x_1 = x_{P+1}$. Thus far, all manipulations have preserved mathematical exactness. However, the physical interpretation of the new form of Q is worth considering. By comparing the integrands of Equations (2.16) and (2.18), it is clear that the original density matrix element has been split into P new density matrix elements whose starting and ending points are consecutively linked. These new density matrix elements correspond to the links between Trotter beads depicted in Figure (1.1), so P is equal to the number of Trotter beads introduced through the *classical isomorphism*. Each of the new density matrix elements corresponds to a fractionally smaller value of β or, equivalently, larger value of the temperature. The strategy of path integral theory is to accurately approximate these high temperature path segments to evaluate the quantum partition function.

2.3.1 Derivation of the Path Integral Approximation

The Trotter product formula applied to the unnormalised density matrix operator $\hat{\rho}(\beta/P)$ yields, assuming $H = T + V$:[162,163,169]

$$\hat{\rho}(\beta/P) = e^{-\beta V/2P} e^{-\beta T/P} e^{-\beta V/2P} + \mathcal{O}\left[(\beta/P)^3\right]. \tag{2.19}$$

Moreover, the Trotter theorem states that[100]

$$\hat{\rho}(\beta/P) = \lim_{P\to\infty} \left[e^{-\beta V/2P} e^{-\beta T/P} e^{-\beta V/2P}\right]^P. \tag{2.20}$$

Taking the position matrix element of $\hat{\rho}$ in Equation (2.20) and substituting it into Equation (2.18) leads to the result

$$Q(\beta) = \lim_{P\to\infty} \prod_{m=1}^{P} \left[\int dx_m\right] \prod_{n=1}^{P} \tilde{\rho}(x_n, x_{n+1}, \beta/P) \tag{2.21}$$

where $x_1 = x_{P+1}$ and

$$\begin{aligned} \tilde{\rho}(x_n, x_{n+1}, \beta/P) &= \langle x_n | e^{-\beta V/2P} e^{-\beta T/P} e^{-\beta V/2P} | x_{n+1}\rangle \\ &= e^{-\beta V(x_n)/2P} \langle x_n | e^{-\beta T/P} | x_{n+1}\rangle e^{-\beta V(x_{n+1})/2P} \\ &= \rho_0(x_n, x_{n+1}, \beta/P)\, e^{-\beta(V(x_n)+V(x_{n+1}))/2P}. \end{aligned} \tag{2.22}$$

Note from the error term in Equation (2.19) that approximation of the matrix element $\rho(x_n, x_{n+1}, \beta/P)$ with $\tilde{\rho}(x_n, x_{n+1}, \beta/P)$ improves with higher temperature or larger numbers of Trotter beads. The function ρ_0 in Equation (2.22) is the free particle density matrix element and was found in Section (2.2.2) to have the exact form[59,150]

$$\rho_0(x, x', \beta) = \langle x_n | e^{-\beta T} | x_{n+1}\rangle = \left[\frac{m}{2\pi\beta\hbar^2}\right] e^{-m(x-x')^2/2\beta\hbar^2}. \tag{2.23}$$

Substitution of Equations (2.23) and (2.22) into Equation (2.21) finally yields the completed path integral representation of the canonical partition function

$$Q(\beta) = \lim_{P\to\infty} \left(\frac{mP}{2\pi\beta\hbar^2}\right)^{P/2} \prod_{m=1}^{P} \left[\int dx_m\right] e^{-\beta(V_{int}+V_{ext})} \tag{2.24}$$

where $x_1 = x_{P+1}$ and with

$$V_{int} = \frac{mP}{2\beta^2\hbar^2}\sum_{n=1}^{P}(x_n - x_{n+1})^2 \text{ and } V_{ext} = \frac{1}{P}\sum_{n=1}^{P}V(x_n)\,. \tag{2.25}$$

The path integral approximation arises from the practical selection of some finite number of Trotter beads P. Chandler and Wolynes pointed out that the integrand in Equation (2.24) is analogous to the classical Boltzmann factor for a particular ring polymer with P monomers linked by a quadratic interaction (V_{int}) and subject to the influence of an external potential (V_{ext}).[38]

Note that throughout this development, the entire impact of temperature is conveniently contained in the parameter β. Also note that the capacity to describe quantum behaviour can literally be modulated with an adjustment of the number of Trotter beads.

Chapter 3

Monte Carlo Theory

3.1 Introduction

Equations (2.24) and (2.25) arrived at in the previous chapter comprise an appealing expression for the canonical partition function Q. However, it has yet to be explained how this intimidating multi-dimensional integral can actually be computed or how useful physical properties will be extracted. These feats require the employment of the useful Monte Carlo method.[34,96,97,196]

Monte Carlo techniques comprise a varied and important class of numerical methods for solving statistical problems in chemistry and physics.[34,134] Frequently, as in the case of path integrals, a chemical problem can be formulated with relative ease, but the complicated nature of the expression prohibits exact calculation.[27] It is in this venue that Monte Carlo methods find their greatest utility. Although the following discussion corresponds generally to the excellent introduction to Monte Carlo theory by Newman and Barkema, it has been adapted in this presentation to

the specific application of PIMC techniques.[134]

Consider, momentarily, the expectation of some quantity A of a system of discrete states in thermal equilibrium,

$$\langle A \rangle = \frac{1}{Q} \sum_{\mu} A_{\mu} e^{-\beta E_{\mu}}. \tag{3.1}$$

The practical difficulty in calculating $\langle A \rangle$ with this simple expression is that the number of states available to the system may exceed that which can be readily computed. However, the form of the summand in Equation (3.1) suggests a more tractable approach. States high in energy will make vanishingly small contributions to the total expectation value because of the Boltzmann weighting term $e^{-\beta E_{\mu}}$ and may thus be excluded from the summation. For most physically relevant systems, relatively few states are low enough in energy to contribute significantly to the expectation value. Therefore, it seems likely that statistical properties of such systems at thermal equilibrium can be accurately obtained by sampling only the lowest-energy states.

The strategy of Monte Carlo (MC) methods is just that: to obtain reliable statistical calculations of systems in thermal equilibrium by sampling relatively few of the total number of states conceivably occupied by the system.[134] Although an astonishing variety of MC techniques have been derived and implemented in virtually every discipline that utilises numerical analysis, each method evolves from the same general idea. The equilibrium process is simulated for a model system. The system is allowed to move through its various states in such a fashion that each state appears with its appropriate probability. Quantities of interest will be sampled during this simulated dynamical process to yield the desired expectation values. For the model system, the probability distribution of the states will be known *a priori*, and the rules for transferring between states during the simulation will be subject

to the constraint that this probability distribution is reproduced.

3.2 Path Integral Probability Distribution

The *a priori* distribution for systems in thermal equilibrium at temperature T is generally the Boltzmann distribution[134]

$$\rho_\mu = \frac{1}{Q} e^{-\beta E_\mu} \tag{3.2}$$

where Q is the partition function. However, we will consider the version of this distribution subjected to path integral discretization[38]

$$\rho^{PI}(\vec{x}) = \frac{1}{Q} e^{-\beta(V_{int}+V_{ext})} \tag{3.3}$$

from Equation(2.24). Note that the path integral probability distribution is a function of the coordinate vector of the P Trotter beads. The terms V_{int} and V_{ext} are also functions of $\vec{x}$ and are defined in Equation (2.25).

3.3 Path Integral Energy Estimator

Some care must be taken with regard to the calculation of observable quantities in light of the fact that the path integral probability distribution varies from the exact Boltzmann distribution. For example, we evaluate the internal energy expectation value $\langle E \rangle$ using the exact Boltzmann distribution simply according to[134]

$$\langle E \rangle = \frac{1}{Q} \sum_\mu E_\mu e^{-\beta E_\mu}. \tag{3.4}$$

However, it is incorrect to assume that substituting $V_{int}+V_{ext}$ for E_μ in the preceding equation will yield the same internal energy expectation value $\langle E \rangle$. An internal

energy estimator E^{PI} must be derived such that[85]

$$\langle E \rangle = \frac{1}{Q} \prod_{m=1}^{P} \left[\int dx_m \right] E^{PI}(\vec{x})\, e^{-\beta(V_{int}+V_{ext})}. \tag{3.5}$$

Using the path integral representation of the canonical partition function in Equation (2.24) and the fact that

$$\langle E \rangle = -\frac{\partial \log Q}{\partial \beta}, \tag{3.6}$$

we obtain the following form for the path integral energy estimator[14]

$$E^{PI}(\vec{x}) = \frac{P}{2\beta} - \frac{mP}{2\beta^2\hbar^2} \sum_{n=1}^{P} (x_n - x_{n+1})^2 + \frac{1}{P} \sum_{n=1}^{P} V(x_n) \tag{3.7}$$

or

$$E^{PI}(\vec{x}) = \frac{P}{2\beta} - V_{int} + V_{ext}. \tag{3.8}$$

The path integral energy estimator presented here is often referred to as the Kinetic Energy Estimator.[14] Another popular path integral energy estimator is the Virial Energy Estimator which is often found to give better numerical convergence for path integral simulations requiring a large number of Trotter beads.[85] However, it is not found that the coupled molecular torsion problem requires large numbers of Trotter beads, so since the Virial Energy Estimator requires the computationally expensive task of determining energy gradients, we discuss only the Kinetic Energy Estimator in this study.

3.3.1 Relationship of Path Integrals to Classical Mechanics

As a brief (but interesting and important) side note, we display in this section the simple manner in which the path integral formalism may be utilised to perform a classical simulation. The classical energy estimator is first derived in the same fashion as the path integral energy estimator.

Again utilising the relationship,

$$\langle E\rangle = -\frac{\partial \log Q}{\partial \beta}, \tag{3.9}$$

we insert the classical partition function, yielding

$$\langle E\rangle = -\frac{C}{Q}\frac{\partial}{\partial\beta}\int dx \int dp\; e^{-\beta(T(p)+V(x))}, \tag{3.10}$$

where C is a normalisation constant for the partition function.[37] It is assumed that the kinetic energy $T(p) = \frac{p^2}{2m}$ is a function only of the momentum and the potential energy $V(x)$ is a function only of the position. Standard manipulations then yield

$$\begin{aligned}\langle E\rangle &= \frac{\int dx \int dp\; \left(\frac{p^2}{2m}+V(x)\right) e^{-\beta\left(\frac{p^2}{2m}+V(x)\right)}}{\int dx \int dp\; e^{-\beta\left(\frac{p^2}{2m}+V(x)\right)}} \\ &= \frac{\int dp\; \frac{p^2}{2m} e^{-\beta\frac{p^2}{2m}} \int dx\; e^{-\beta V(x)}}{\int dp\; e^{-\beta\frac{p^2}{2m}} \int dx\; e^{-\beta V(x)}} + \frac{\int dx\; V(x)\, e^{-\beta V(x)} \int dp\; e^{-\beta\frac{p^2}{2m}}}{\int dx\; e^{-\beta V(x)} \int dp\; e^{-\beta\frac{p^2}{2m}}}.\end{aligned} \tag{3.11}$$

We then cancel like terms to obtain

$$\langle E\rangle = \frac{\int dp\; \frac{p^2}{2m} e^{-\beta\frac{p^2}{2m}}}{\int dp\; e^{-\beta\frac{p^2}{2m}}} + \frac{\int dx\; V(x)\, e^{-\beta V(x)}}{\int dx\; e^{-\beta V(x)}}. \tag{3.12}$$

The first term on the right-hand side of Equation (3.12) is immediately identified as $\left\langle \frac{p^2}{2m} \right\rangle$, the average kinetic energy for a free classical particle. It can be evaluated analytically to yield $\left\langle \frac{p^2}{2m} \right\rangle = \frac{1}{2\beta}$. Finally,

$$\langle E\rangle = \frac{\int dx\; \left(\frac{1}{2\beta}+V(x)\right) e^{-\beta V(x)}}{\int dx\; e^{-\beta V(x)}}, \tag{3.13}$$

so the classical energy estimator is

$$E^{classical}(x) = \frac{1}{2\beta} + V(x), \tag{3.14}$$

and the probability distribution utilised in the MC simulation is

$$\rho^{classical} = e^{-\beta V(x)}. \tag{3.15}$$

Note that classical Equations (3.14) and (3.15) recover the path integral Equations (3.8) and (3.3) when the number of Trotter beads P is set to 1 in Equation (2.25). That is, a classical simulation may be performed within the path integral formalism by simply setting the number of Trotter beads P to 1.

3.4 Importance Sampling

The discretized form of Equation (3.5) may be written as

$$\langle E \rangle = \frac{\prod_{m=1}^{P} \left[\sum_{t_m}\right] E^{PI}\left(\vec{x}_t\right) e^{-\beta(V_{int}+V_{ext})}}{\prod_{m=1}^{P} \left[\sum_{t_m}\right] e^{-\beta(V_{int}+V_{ext})}}. \tag{3.16}$$

The best possible estimate of $\langle E \rangle$ upon truncation of Equation (3.16) is then

$$\langle E \rangle = \frac{\prod_{m=1}^{P} \left[\sum_{t_m=1}^{M_m}\right] E^{PI}\left(\vec{x}_t\right) \rho\left(\vec{x}_t\right)^{-1} e^{-\beta(V_{int}+V_{ext})}}{\prod_{m=1}^{P} \left[\sum_{t_m=1}^{M_m}\right] \rho\left(\vec{x}_t\right)^{-1} e^{-\beta(V_{int}+V_{ext})}} \tag{3.17}$$

where $\rho\left(\vec{x}_t\right)$ is an arbitrary distribution function that may be chosen to have any form.[134] Judicious selection of this *a priori* distribution function may dramatically reduce the necessary size of the elements in $\vec{M}$ by avoiding the inclusion of states too high in energy to contribute to the expectation value. Importance sampling is the process of including states according to the probability of a particular distribution function.[134]

Suppose we choose $\rho\left(\vec{x}_t\right)$ to take the form of the Boltzmann-like path integral distribution function

$$\rho^{PI}\left(\vec{x}\right) = \frac{e^{-\beta(V_{int}+V_{ext})}}{\prod_{m=1}^{P} \left[\sum_{t_m=1}^{M_m}\right] e^{-\beta(V_{int}+V_{ext})}}. \tag{3.18}$$

Inserting Equation (3.18) into Equation (3.17) we arrive at a much simplified form

for the expectation value:

$$\langle E \rangle = \frac{\prod_{m=1}^{P} \left[\sum_{t_m=1}^{M_m} \right] E^{PI}(\vec{x}_t)}{\prod_{m=1}^{P} [M_m]}. \tag{3.19}$$

A Markov process is employed to ensure that the Monte Carlo simulation generates states according to the selected *a priori* distribution probability.

3.5 Markov Processes

Monte Carlo simulations require a means of selecting states according to a given probability distribution. That is, a sequence of states must be generated such that, given a long enough sequence, the probability distribution of the generated states is equal to the *a priori* probability distribution. This task is accomplished with the aid of a Markov process.

A Markov process is a mechanism which, given a state μ of a particular system, generates another state ν of the same system.[134] The new state is selected randomly. The same input state does not always yield the same output state, and the two states may even be identical.

A Markov process is defined in terms of a set of transition probabilities $P(\mu \rightarrow \nu)$ which are required to obey several properties.[134] In addition to being independent of time, the $P(\mu \rightarrow \nu)$ depend only on the properties of the current states μ and ν. The Markov transition probabilities must be entirely unaffected by the past and future of the system's pathway.

Another property of the transition probabilities is that

$$\sum_{\nu} P(\mu \rightarrow \nu) = 1. \tag{3.20}$$

Equation (3.20) requires that the total probability of finding the particle after transition from state μ is unity or, equivalently, that the particle does, in fact, go somewhere.

Markov processes are required to fulfil the condition of ergodicity. This condition ensures that the system will reach every state of non-zero probability if enough state transitions are generated. Naturally, this has to be the case if an *a priori* probability distribution is to be reproduced. A state generating mechanism can hardly be expected to return a non-zero probability for a state to which it prohibits transition. However, ergodicity does not require that every $P(\mu \rightarrow \nu)$ be non-zero. It merely states that for two states there exist at least one transition pathway of non-zero probability.

The final requirement of the Markov transition probabilities is that they obey detailed balance,

$$\rho_\mu P(\mu \rightarrow \nu) = \rho_\nu P(\nu \rightarrow \mu) \tag{3.21}$$

where ρ_μ and ρ_ν are the probabilities of the system existing, respectively, in states μ and ν. Detailed balance basically ensures that a system obeying the dynamics of a particular Markov process ultimately reaches equilibrium. The rate of change from a state μ to another state ν is the same as the rate in the other direction.

Equation (3.21) also provides the means by which a Markov process may be obtained for a given problem. By rearranging this equation and recalling the *a priori* path integral probability distribution, we conclude that Markov transition probabilities for a path integral Monte Carlo simulation must obey

$$\frac{P(\mu \rightarrow \nu)}{P(\nu \rightarrow \mu)} = \frac{\rho_\nu}{\rho_\mu} = \frac{\rho\left(\vec{x}_\nu\right)}{\rho\left(\vec{x}_\mu\right)} = e^{-\beta[(V_{int}(\vec{x}_\nu)+V_{ext}(\vec{x}_\nu))-(V_{int}(\vec{x}_\mu)+V_{ext}(\vec{x}_\mu))]}. \tag{3.22}$$

Within these given constraints, any set of transition probabilities may legitimately be employed. All of the elements necessary to perform a Monte Carlo

simulation have now been described. However, it remains an unfortunate reality that the task of finding an efficient Markov process that obeys Equation (3.22) is often a nontrivial process.

3.6 Acceptance Ratios

For many problems physical problems, the *a priori* selection of a Markov process that exactly obeys the condition of Equation (3.22) can be a difficult task. Fortunately, a very convenient solution to this problem accompanies the introduction of an acceptance ratio.[134] Assume $P(\mu \rightarrow \nu)$ takes the form

$$P(\mu \rightarrow \nu) = g(\mu \rightarrow \nu)\, A(\mu \rightarrow \nu) \tag{3.23}$$

where $g(\mu \rightarrow \nu)$ is called the selection probability and $A(\mu \rightarrow \nu)$ is the acceptance ratio. The term $g(\mu \rightarrow \nu)$, much like $P(\mu \rightarrow \nu)$, is the probability that the system will consider proceeding to state ν from the current state μ. However, the selection probability $g(\mu \rightarrow \nu)$ is not constrained to obey Equation (3.22). The acceptance ratio $A(\mu \rightarrow \nu)$ determines the probability that the system will actually accept a transition from state μ to state ν proposed by $g(\mu \rightarrow \nu)$. The acceptance ratio is easily chosen so that the entire transition probability produces the correct probability distribution.

It is important to expand on the fundamental change in strategy introduced with the acceptance ratio. Prior to the consideration of Equation (3.23), the system invariably performed the transition to the state selected by the Markov process. However, the acceptance probability simplifies the responsibilities of the state generation mechanism by allowing the system to actually perform state transitions only as often as is necessary to reproduce the desired probability distribution.

Although the introduction of the acceptance ratio greatly reduces the difficulty of selecting a valid Markov process for the Monte Carlo simulation, an important factor must still be considered. In order for the system to efficiently sample the states of the system, state transitions must be accepted as frequently as possible. That is, the selection probability and acceptance ratio should be chosen so that $A(\mu \rightarrow \nu)$ is as close to unity as possible. In this study, we employ a very common form of the acceptance ratio known as the Metropolis algorithm.[130]

3.7 Metropolis Algorithm

Suppose we choose our selection probabilities so that the systems changes from a state μ to a state ν with a probability equal to that of transition in the other direction. That is,

$$g(\mu \rightarrow \nu) = g(\nu \rightarrow \mu)\,. \tag{3.24}$$

This is a very convenient approach for PIMC. It is very easy to simply allow the beads to move in an arbitrary direction. Then, the acceptance ratio is chosen so that

$$\Delta E_{\nu\mu} = \left(V_{int}\left(\vec{x}_\nu\right) + V_{ext}\left(\vec{x}_\nu\right)\right) - \left(V_{int}\left(\vec{x}_\mu\right) + V_{ext}\left(\vec{x}_\mu\right)\right) \tag{3.25}$$

$$A(\mu \rightarrow \nu) = \begin{cases} e^{-\beta \Delta E_{\nu\mu}} & \text{if } \Delta E_{\nu\mu} > 0 \\ 1 & \text{otherwise} \end{cases} \tag{3.26}$$

Note that $A(\mu \rightarrow \nu)$ will assume unity for half of the transition performed by the system. This will lead to facile exploration of expectation values. It is easily verified that, using the Metropolis algorithm,

$$\frac{P(\mu \rightarrow \nu)}{P(\nu \rightarrow \mu)} = \frac{g(\mu \rightarrow \nu)\, A(\mu \rightarrow \nu)}{g(\nu \rightarrow \mu)\, A(\nu \rightarrow \mu)} = \frac{A(\mu \rightarrow \nu)}{A(\nu \rightarrow \mu)} = e^{-\beta \Delta E_{\nu\mu}} \tag{3.27}$$

as required by Equation (3.22).

Chapter 4

Torsional PIMC Theory

4.1 Introduction

In Chapter 2, a thorough derivation was performed of the path integral formalism for a single particle not subjected to periodic boundary conditions. However, the objective of this research is to use the PIMC technique to model the many torsional degrees of freedom that exist in large molecules. The molecular problem will require a formalism for a many-particle system, and each particle must conform to a periodic boundary condition on the interval $[0, 2\pi)$.

4.2 Many-Particle Torsional Path Integrals

The derivation of many-particle torsional path integrals in this section will be obtained by extending the single-particle translational case provided in Section (2.3).

For a system of N torsions, the kinetic energy operator T takes the form

$$T = \sum_{j=1}^{N} \frac{-\hbar^2}{2I_j} \frac{\partial^2}{\partial \theta_j^2} \tag{4.1}$$

where I_j is the moment of inertia of the j^{th} torsion. The N dimensional position vector $\left|\vec{\theta}\right\rangle$ where $\theta_j \in [0, 2\pi)$ comprise the eigenfunctions of the potential operator V with eigenfunctions $V\left(\vec{\theta}\right)$. These changes lead to the following form for the exact partition function Q:[119]

$$Q(\beta) = \lim_{P\to\infty} \left(\frac{P}{2\pi\beta\hbar^2}\right)^{NP/2} I^{1/2} \prod_{t=1}^{P} \prod_{j=1}^{N} \left[\sum_{n_{j,t}=-\infty}^{\infty} \int_0^{2\pi} \mathrm{d}\theta_{j,t}\right] e^{-\beta\left(V_{\vec{n_j}}^{int} + V^{ext}\right)} \tag{4.2}$$

where P is the number of Trotter beads for the system, $\theta_{j,1} = \theta_{j,P+1}$, $I^{1/2} = \prod_{k=1}^{N} I_k^{1/2}$, and with

$$V_{\vec{n_j}}^{int} = \frac{P}{2\beta^2\hbar^2} \sum_{t=1}^{P} \sum_{j=1}^{N} I_j \left(\theta_{j,t} - \theta_{j,t+1} + 2\pi n_{j,t}\right)^2 \text{ and } V^{ext} = \frac{1}{P} \sum_{t=1}^{P} V\left(\vec{\theta}_t\right). \tag{4.3}$$

Clearly, Equations (4.2) and (4.3) are the many-particle torsional analog of Equations (2.24) and (2.25). However, the complicated nature of these current equations warrants some discussion. The *classical isomorphism* for the many-particle case requires each particle j to be replaced by a ring polymer of size P. The terms $V_{\vec{n_j}}^{int}$ and V^{ext} in Equation (4.3) maintain the same form as in Equation (2.25), except for the appearance of integer $n_{j,t}$ in the equation for $V_{\vec{n_j}}^{int}$. This integer term is a manifestation of the constraint that the $\theta_{j,t}$ lie in the interval $[0, 2\pi)$. It accounts for the fact that the distance between two consecutive Trotter beads for a particle j, namely $\theta_{j,t}$ and $\theta_{j,t+1}$, may assume not only the explicit value $\theta_{j,t} - \theta_{j,t+1}$ but all of the values $\theta_{j,t} - \theta_{j,t+1} + 2\pi n_{j,t}$. That is, in travelling from one Trotter bead to the next, the chain may wrap around the $[0, 2\pi)$ interval $n_{j,t}$ times.

4.2.1 Winding Number Representation

A variable transformation is now performed to a more compact set of variables. Substituting

$$\begin{aligned}
\tilde{\theta}_{j,1} &= \theta_{j,1} \\
\tilde{\theta}_{j,2} &= \theta_{j,2} - 2\pi n_{j,1} \\
\tilde{\theta}_{j,3} &= \theta_{j,3} - 2\pi n_{j,1} - 2\pi n_{j,2} \\
&\vdots \\
\tilde{\theta}_{j,t} &= \theta_{j,2} - 2\pi \sum_{t'=1}^{t-1} n_{j,t'}
\end{aligned} \tag{4.4}$$

leads to a cancellation of all of the $2\pi n_{j,t}$ terms for $t = 1, \ldots, P-1$:

$$(\theta_{j,t} - \theta_{j,t+1} + 2\pi n_{j,t})^2 \rightarrow \left(\tilde{\theta}_{j,t} - \tilde{\theta}_{j,t+1}\right)^2 . \tag{4.5}$$

However, the P^{th} term yields

$$(\theta_{j,P} - \theta_{j,P+1} + 2\pi n_{j,P})^2 \rightarrow \left(\tilde{\theta}_{j,P} - \tilde{\theta}_{j,t+P} + 2\pi n_j\right)^2 \tag{4.6}$$

where $n_j = \sum_{t'=1}^{P} n_{j,t'}$ is known as the winding number. This new set of variables is known as the Winding Number Representation and leads to the following form for Q:[120]

$$Q(\beta) = \lim_{P\to\infty} \left(\frac{P}{2\pi\beta\hbar^2}\right)^{NP/2} I^{1/2} \prod_{j=1}^{N} \sum_{n_j=-\infty}^{\infty} \int_0^{2\pi} \mathrm{d}\tilde{\theta}_{j,1} \prod_{t=2}^{P} \left[\int_{-\infty}^{\infty} \mathrm{d}\tilde{\theta}_{j,t}\right] e^{-\beta\left(V_{\vec{n}}^{int} + V^{ext}\right)} \tag{4.7}$$

with

$$V_{\vec{n}}^{int} = \frac{P}{2\beta^2\hbar^2} \sum_{t=1}^{P} \sum_{j=1}^{N} I_j \left(\tilde{\theta}_{j,t} - \tilde{\theta}_{j,t+1} + 2\pi n_j \delta_{t,P}\right)^2 \text{ and } V^{ext} = \frac{1}{P} \sum_{t=1}^{P} V\left(\vec{\tilde{\theta}}_t\right) . \tag{4.8}$$

The periodicity of the potential renders the expression for V^{ext} unchanged from Equation (4.3). The winding number n_j in this representation is a property of the

entire ring of Trotter beads for a given particle j. It represents the net number of times that the ring loops around the interval $[0, 2\pi)$. Introduction of the winding number in this fashion frees all particles from the $[0, 2\pi)$ boundary constraint, except for particle $j = 1$. The Winding Number Representation for rotational path integrals has been utilised extensively by Marx and coworkers in the study of dimers rotating on a surface.[119,120]

4.2.2 Uncoupled Winding Number Representation

Another change of variables is still necessary before we arrive at the path integral formulation most readily applicable to molecular torsions. The following variable substitution was first suggested by Cao.[30]

$$\bar{\theta}_{j,t} = \tilde{\theta}_{j,t} + 2\pi n_j \left(\frac{t-1}{P}\right) \qquad t = 1, \ldots, P \tag{4.9}$$

Transformation to the Uncoupled Winding Number Representation leaves the following computationally useful form of the partition function:

$$Q(\beta) = \lim_{P\to\infty} \left(\frac{P}{2\pi\beta\hbar^2}\right)^{NP/2} I^{1/2} \prod_{j=1}^{N} \int_0^{2\pi} \mathrm{d}\tilde{\theta}_{j,1} \prod_{t=2}^{P} \left[\int_{-\infty}^{\infty} \mathrm{d}\bar{\theta}_{j,t}\right] \sum_{n_j=-\infty}^{\infty} e^{-\beta\left(V^{int}+V_{\vec{n}}^{ext}+V_{\vec{n}}^{wind}\right)} \tag{4.10}$$

The terms V^{int}, $V_{\vec{n}}^{ext}$, and $V_{\vec{n}}^{wind}$ assume the form

$$\begin{aligned} V^{int} &= \frac{P}{2\beta^2\hbar^2} \sum_{j=1}^{N} I_j \left[\sum_{t=1}^{P} \left(\bar{\theta}_{j,t} - \bar{\theta}_{j,t+1}\right)^2\right], \\ V_{\vec{n}}^{ext} &= \frac{1}{P} \sum_{t=1}^{P} V\left(\vec{\theta}_t - \frac{2\pi\vec{n}\,(t-1)}{P}\right), \text{ and} \\ V_{\vec{n}}^{wind} &= \frac{1}{2\beta^2\hbar^2} \sum_{j=1}^{N} I_j \left(2\pi n_j\right)^2. \end{aligned} \tag{4.11}$$

The Uncoupled Winding Number Representation is a tremendously useful formulation of torsional PIMC theory, largely because of its strong similarities to the

formalism of translational PIMC. Note that the expressions for V^{int} and $V_{\vec{n}}^{ext}$ are virtually unchanged from their form for the case of translational PIMC in Equation (2.25). Also, the winding number terms n_j in Equation (4.10) are, aside from the coordinate shift in $V_{\vec{n}}^{ext}$, decoupled from the angular coordinates. Computationally, of course, the expression for Q in Equation (4.10) is implemented by choosing maximum values of P and n_j. Suitable maximum values of the n_j are apparent from the uncoupled expression because the entire winding number contribution for a given particle falls away exponentially with I_j and n_j^2. A final computational boon of the uncoupled representation regards selection of a suitable Monte Carlo sampling distribution. Equation (4.10) suggests that, just as in translational PIMC, a probability distribution of $\rho = \frac{1}{Q} e^{-\beta\left(V^{int}+V_{\vec{0}}^{ext}\right)}$ may be appropriate.[30]

Chapter 5

Calculation Details

5.1 Introduction

The molecular systems considered in this study were ethane, n-butane, n-octane, and the peptide enkephalin. These molecules include, respectively, 1, 3, 7, and 33 internal torsions.

Two potentials are employed in this study. The first is an analytical potential used only for the ethane molecule. The analytical ethane potential is a 1-D sinusoidal function of the torsional angle θ. Parametrizing the maxima of the potential to the MM3Pro force field yielded the function $V(\theta) = C\,(\cos(3\theta) + 1)$, where $C = 1.33$ kcal mol^{-1}.[7] The second potential was used for all of the molecules studied. This latter potential is the MM3Pro atom-atom force field as implemented in the TINKER molecular mechanics package.[7,139]

The torsional moment of inertia I_i about a bond between atoms k and l was defined in terms of the moments of inertia, $\tilde{I}_k$ and $\tilde{I}_l$, of two unsymmetrical tops

that rotate with respect to one another about the bond.[39,138] That is,

$$I_i = \frac{\tilde{I}_k \tilde{I}_l}{\tilde{I}_k + \tilde{I}_l}. \tag{5.1}$$

Each of the two tops correspond to the molecular moieties attached to the atoms terminating the bond, so

$$\tilde{I}_k = \sum_{j<k} m_j \left(r_{jk}^2 - \left[\frac{\vec{r}_{jk} \cdot \vec{r}_{kl}}{|\vec{r}_{kl}|} \right]^2 \right) \quad \text{and} \quad \tilde{I}_l = \sum_{j>l} m_j \left(r_{jl}^2 - \left[\frac{\vec{r}_{jl} \cdot \vec{r}_{kl}}{|\vec{r}_{kl}|} \right]^2 \right). \tag{5.2}$$

In Equation (5.2),

$$\vec{r}_{kl} = \vec{r}_k - \vec{r}_l, \tag{5.3}$$

$\vec{r}_k$ is the position vector of atom k with mass m_k, and the atoms are indexed in a fashion appropriate to their connectivity. In this study, the torsional moments of inertia are calculated only once and assumed to remain constant throughout the simulation. The moments of inertia are calculated using the starting geometry of the molecule at the global minimum of the potential energy surface. Although this approach is consistent with previous studies, it should be pointed out that moments of inertia will, in general, vary as a function of the torsional angle vector.[39,41]

The exchange of indistinguishable hydrogen atoms arising from the three-fold symmetry of the methyl torsions requires, in principle, the consideration of nuclear exhange symmetry.[126] These rotations only allow even permutations of the fermionic hydrogens, so the total nuclear-rotational wavefunction must be totally symmetric. Although the effects of nuclear exchange symmetry can be relevant in methyl torsions at very low temperatures, they are expected to be negligible at the temperatures of interest in this study and are not further considered.[126,140]

The internal energy was determined for a variety of molecular systems using three different methodologies. The remainder of the chapter is dedicated to discussing these techniques.

5.2 PIMC Method

Although the formal theory of torsional PIMC has been thoroughly discussed, no general summary of the algorithm has thus far been presented. In the following section, we include an outline of the torsional PIMC algorithm implemented in our code. The PIMC code is currently capable of calculating the internal energy of any system of coupled torsions.

5.2.1 Algorithm Summary

The probability distribution that is sampled in our implementation of the Uncoupled Winding Number Representation of torsional PIMC was originally proposed in Section (4.2.2).[30]

$$\rho = \frac{1}{Q} e^{-\beta\left(V^{int}+V_{\vec{0}}^{ext}\right)} \tag{5.4}$$

where

$$\begin{aligned} V^{int} &= \frac{P}{2\beta^2\hbar^2}\sum_{j=1}^{N} I_j \left[\sum_{t=1}^{P}\left(\bar{\theta}_{j,t}-\bar{\theta}_{j,t+1}\right)^2\right] \text{ and} \\ V_{\vec{0}}^{ext} &= \frac{1}{P}\sum_{t=1}^{P} V\left(\vec{\theta}_t\right). \end{aligned} \tag{5.5}$$

Starting with a given configuration of beads $\vec{x}_\mu$, a new configuration of beads $\vec{x}_\nu$ is proposed according to an appropriate selection probability. The selection probability $g(\mu \rightarrow \nu)$ employed in this study for molecular torsions allows each bead to move arbitrarily within the range of adjustable chain and bead step sizes. This mechanism of generating new configurations obeys Equation (3.24).

It is then necessary to determine whether to accept the proposed configuration step so that the sampled probability distribution in Equation (5.4) is achieved. This

is done by calculating the term

$$\Delta E_{\nu\mu} = \left(V^{int}\left(\vec{x}_\nu\right) + V_{\vec{0}}^{ext}\left(\vec{x}_\nu\right)\right) - \left(V^{int}\left(\vec{x}_\mu\right) + V_{\vec{0}}^{ext}\left(\vec{x}_\mu\right)\right) \tag{5.6}$$

from Section (3.7) and employing the Metropolis Algorithm. If the configuration change is accepted, set $\vec{x}_\mu = \vec{x}_\nu$. Otherwise, $\vec{x}_\mu$ remains unchanged.

Now use the configuration $\vec{x}_\mu$ and an appropriate torsional PIMC energy estimator to sample the internal energy of the configuration. The many-particle torsional PIMC analog of the Kinetic Energy Estimator presented in Section (3.3) is

$$E^{PI} = \prod_{j=1}^{N}\left[\sum_{n_j=-n_j^{max}}^{n_j^{max}}\right]\epsilon_{\vec{n}}^{PI} e^{-\Delta S_{\vec{n}}} \tag{5.7}$$

where n_j^{max} is the maximum winding number for a given particle j, and

$$\begin{aligned}
\Delta S_{\vec{n}} &= \beta\left(\left[V^{int} + V_{\vec{n}}^{ext} + V_{\vec{n}}^{wind}\right] - \left[V^{int} + V_{\vec{0}}^{ext} + V_{\vec{0}}^{wind}\right]\right) \\
&= \beta\left(V_{\vec{n}}^{ext} + V_{\vec{n}}^{wind} - V_{\vec{0}}^{ext}\right) \\
&= \beta\left(\frac{1}{P}\sum_{t=1}^{P}\left[V\left(\vec{\theta}_t - \frac{2\pi\vec{n}\,(t-1)}{P}\right) - V\left(\vec{\theta}_t\right)\right] + \frac{1}{2\beta^2\hbar^2}\sum_{j=1}^{N} I_j\left(2\pi n_j\right)^2\right)
\end{aligned} \tag{5.8}$$

and

$$\epsilon_{\vec{n}}^{PI} = \frac{NP}{2\beta} - V^{int} + V_{\vec{n}}^{ext} - V_{\vec{n}}^{wind}. \tag{5.9}$$

The equations for V^{int}, $V_{\vec{n}}^{ext}$, and $V_{\vec{n}}^{wind}$ were provided in Equation (4.11). Also, calculate the weight of the configuration

$$\omega = \prod_{j=1}^{N}\left[\sum_{n_j=-n_j^{max}}^{n_j^{max}}\right] e^{-\Delta S_{\vec{n}}}. \tag{5.10}$$

The average internal energy expectation value is

$$\langle E\rangle = \frac{\sum_{configs} E^{PI}}{\sum_{configs}\omega}. \tag{5.11}$$

The scaling of the method is dominated by the term:

$$\prod_{j=1}^{N} \left[2n_j^{max} + 1\right] PN^{MC} \tag{5.12}$$

where N^{MC} is the number of steps in the Monte Carlo simulation.

5.2.2 Simplification of Winding Number Term

As was noted in Section (4.2.2), an advantage of the Uncoupled Winding Number Representation is the easily determined influence of winding numbers in the evaluation of expectation values for a given system. Equation (5.7) displays that contributions to the internal energy from a given torsion in a given configuration fall away exponentially with the term

$$\beta V_j^{wind} = \frac{1}{2\beta\hbar^2} I_j \left(2\pi n_j\right)^2 . \tag{5.13}$$

However, for even the lightest molecular torsions the term I_j will be of the order 10^3 in atomic units. Thus, at reasonable temperatures,

$$\beta V_j^{wind} = \frac{1}{2\beta\hbar^2} I_j \left(2\pi n_j\right)^2 \sim \frac{1}{2(1000)(1)^2}(1000)\left(2\pi n_j\right)^2 \sim 20n_j^2. \tag{5.14}$$

Clearly, therefore, the only exponential terms large enough to contribute to the expectation value will thus correspond to $n_j = 0$, so we can reliably set $n_{max} = 0$ for molecular calculations. This dramatically improves the scaling of the method, as is clear from Equation (5.12).

5.2.3 Parameter Discussion

The following is a list of all computational parameters that emerge in our implementation of the torsional PIMC code. The considerations employed in selecting the numerical value of each parameter are described.

Beta: This parameter determines the temperature at which the calculation is performed since $\beta = 1/kT$. For all calculations in this study, a temperature of 273.15 K was employed.

Path Integral Parameters:

Number of Trotter Beads: The number of Trotter beads is equivalent to the number of path segments approximating the trace in the path integral discretized canonical partition function. As shown in Equation (2.19), the path integral approximation improves with the number of Trotter beads P. However, even the facile torsional degrees of freedom are not expected to exhibit the amount of quantum behaviour that requires very large numbers of Trotter beads. This arises from the simple fact that the torsional moments of inertia, even for methyl groups and small substituents, are too large to exhibit massive quantum behaviour. The "appropriate" value of this quantity is explored extensively in this study as it is directly proportionate to the computational cost of the method.

Maximum Winding Number: As described in Section (5.2.2), this parameter can be confidently set to zero. All numerical tests corroborated this supposition.

Monte Carlo Parameters:

Chain and Bead Step Sizes: In compliance with the selection probability employed for PIMC calculations in this study, a new configuration is obtained for a set of torsions after both the *relative* position of the beads for each ring and the *total* position of the ring are altered in a random fashion. The bead step size controls the change in the relative bead positions. During a configuration step, each bead is allowed to move, with equal probability, by any angle less than that defined by the bead step size. The chain step size controls the subsequent change in the total position of the ring of beads. In this step, every bead in the ring is moved by the

same angle chosen randomly but constrained to be smaller than that defined by the chain step size. At the beginning of a PIMC calculation, a chain and bead step size must be chosen for each torsion in the molecular system, a nontrivial process that will be discussed in Section (5.2.4). The selected bead and chain step sizes are held fixed throughout the PIMC calculation to ensure that the MC simulation obeys detailed balance.

Number of Calculation Cycles: Before a PIMC calculation is performed, a geometry optimisation is performed to obtain the lowest energy geometry of the molecular system. At the beginning of the PIMC calculation, the positions of the beads are determined with respect to this geometry. The first bead of every ring is allowed to take the optimal position. The rest of the beads in each ring are then randomly scattered within the neighbouring angles defined by the bead step size.

It is potentially problematic to perform an entire PIMC calculation as one, long, single trajectory. Conceivably, the molecule could work itself into a local minimum and be unable to get out, resulting in the calculation of an artificially high internal energy and likewise damaging other calculated expectation values. To prevent this possibility, the PIMC calculation is performed in a series of cycles. After each cycle, the geometry is reset to the lowest-energy structure, and the PIMC calculation is restarted. At the start of the new cycle, all step size information remains unchanged. Naturally, the number of cycles required for a PIMC calculation is necessarily related to the probability of the system getting hung in an bad region of configuration space and also to the number of configuration steps performed in each cycle. The calculations in this study were each performed for 20 cycles, a value large enough to give reasonably small variance in the calculated values.

Number of Steps per Cycle: This parameter controls the total number of con-

figuration steps that will be performed within a given PIMC cycle. Its value must be chosen to be large enough to yield equilibrated Monte Carlo results. The criterion we employed to evaluate this convergence was the variance in the expectation values calculated within each PIMC cycle. A value of 10^5 was found to be large enough for most PIMC calculations performed in this study.

Sampling Frequency: Monte Carlo simulations hinge, via the employment of a Markov process, on the random generation of configurations or states. However, consecutively generated configurations in the PIMC simulation are more likely to exhibit similar properties than two configurations taken at substantially different points in the calculation. To prevent correlations of this sort, a sampling frequency is defined so that the calculation will only sample the property of interest after a certain number of cycles has been performed. In this study, the internal energy is sampled only once in every 20 cycles.

5.2.4 Chain and Beads Step Size Determination Process

It was noted in Section (5.2.3) that the determination of the chain and bead step sizes for a given system could prove a very difficult task. Although one suspects that the magnitude of these terms should correlate with the magnitude of the moments of inertia, the actual relationship is related to the complex potential energy hypersurface of the system. Selecting reasonable step sizes are critical to the efficiency of the MC simulation. If the step sizes are too large, a large proportion of the steps will lead to unreasonable structures rejected by the Metropolis algorithm. If the step sizes are too small, the simulation will be slow in exploring the state space. It is common in metropolis MC algorithms to choose step sizes that facilitate an acceptance rate of approximately 0.5.

The step size determination process is divided into two parts. The first optimises the bead and chain step sizes of each torsion with respect to the step sizes of the other torsions. The second part holds step sizes of the torsions in relative proportion and simply adjusts their magnitude to obtain an acceptance ratio of 0.5. In both parts, we run a set of normal PIMC steps and meanwhile adjust the bead and chain step size parameters. However, the steps are then discarded, so no MC samples are recorded during this step size optimisation process.

For the first step size determination part, each torsion is assigned to one of several groups according to its moment of inertia. Torsions with similar moments of inertia are placed in the same inertia group. For each of these groups, starting guesses of the chain and bead step sizes are assigned. Then an iterative process is employed to, in some sense, equilibrate each parameter with respect to both its individual group and also the system as a whole.

For each inertia group, a particular number of Monte Carlo steps is performed in which one torsional degree of freedom is moved at a time. The corresponding bead step size and chain step size are adjusted to obtain an acceptance ration of 0.5. One at a time, each torsion in the inertia group is treated in this fashion. If the differences between the updated and the original step size guesses fall below a given threshold, the procedure moves on. Otherwise, the procedure continues cycling over the inertia group until a satisfactory equilibration is achieved. This local equilibration process is performed for each inertia group.

After each inertia group has been locally equilibrated, the calculated step size parameters for the entire system are compared to those originally guessed. If the difference falls below a specified threshold, then the first part of the step size parameter selection process is completed. Otherwise, the local equilibration process

described above is repeated until a satisfactory global equilibration is obtained.

For the calculations in the study, the local equilibration cycle is performed with a threshold of 0.1 for the local cycle. The global cycle, unfortunately, is often reluctant to converge so is simply performed 20 times.

After having optimised the bead and chain step size for each torsion with respect to those for the other torsions, it is important it ensure that the magnitudes of these parameter vectors yield the desired acceptance probability. To do so, a second set of discarded PIMC steps is performed in which the torsional parameters are held in proportion. That is, only the magnitude, not the direction, of the step size parameter vectors is optimised. In the study, this was done for 500 discarded steps, and the parameters were adjusted every 10 steps to maintain an acceptance rate of 0.5.

5.3 Variational Method

For simple model systems, the exact quantum result of statistical mechanical properties can be found variationally. A variational calculation was performed in this study for the single-torsion ethane system modelled by a potential energy surface depending only on the torsional angle. Using an excessively large basis set of the wavefunctions for a free particle subjected to periodic boundary constraints, the eigenfunctions and eigenvalues of the model system were obtained to high numerical accuracy.

The comparison of the torsional PIMC results to the exact variational results for small systems provides an important check of the newly implemented PIMC technique. It also lends clear insight to the nature of the Monte Carlo numerical

convergence and the appropriate values for the PIMC parameters.

5.4 Harmonic Oscillator Approximation Method

If we assume that each molecular torsion is adequately described by a single harmonic oscillator well, then quantum mechanical expectation values of the system can easily be obtained to high accuracy. This approach is employed to explore the degree to which the anharmonicity of the torsional degrees of freedom (which is correctly accounted for using the PIMC algorithm) is important in large molecular systems. The PIMC approach is considerably more costly than employing the harmonic oscillator approximation, so it is worth while to determine whether the additional computational expense is necessary.

By performing a Taylor expansion about the local minimum of each torsion, we readily obtain the spring constant k_j for the harmonic oscillator corresponding to the molecular torsion j.

$$k_j = \frac{E_j^+ + E_j^- - 2E_{min}}{h^2} \tag{5.15}$$

where E_{min} is the energy of the system corresponding to the local minimum of each torsion, and E_j^+ and E_j^- are the energies of the system for which only the j^{th} torsions have been, respectively, displaced from their local minima by the small angular distance h.

Now, defining $\vec{n}$ to be the vector of quantum numbers for the harmonic oscillators representing the torsional system, we obtain the energy for a given state $E_{\vec{n}}$ according to

$$E_{\vec{n}} = \sum_{j=1}^{N} \left(n_j + \frac{1}{2} \right) \hbar\omega_j \tag{5.16}$$

where N is the number of torsions, $\omega_j = \sqrt{\frac{k_j}{I_j}}$, and I_j is the moment of inertia for the j^{th} torsion.

Conveniently, the partition function and internal energy expectation value may be obtained analytically for the system of uncoupled harmonic oscillators. Repeated application of the summation formula

$$\sum_{n=0}^{\infty} \alpha^n = \frac{1}{1-\alpha}, \qquad 0 < \alpha < 1, \tag{5.17}$$

yields for the partition function

$$Q = \prod_{j=1}^{N} \left[\sum_{n_j=0}^{\infty} \right] e^{-\beta E_{\vec{n}}} = \prod_{j=1}^{N} \frac{e^{\frac{-\beta\hbar\omega}{2}}}{1 - e^{-\beta\hbar\omega}}. \tag{5.18}$$

The internal energy expectation value is then derived from the partition function:

$$\langle E \rangle = -\frac{\partial \log Q}{\partial \beta} = \sum_{j=1}^{N} \frac{\hbar\omega_j}{2} + \sum_{j=1}^{N} \frac{\hbar\omega_j e^{-\beta\hbar\omega}}{1 - e^{-\beta\hbar\omega}}. \tag{5.19}$$

Note that, in addition to neglecting the anharmonicity of the molecular torsions, this formulation of the harmonic oscillator approximation neglects the coupling between the individual oscillators.

Chapter 6

Results and Discussion

6.1 Introduction

In this study, the internal energies of the molecules ethane, n-butane, n-octane, and enkephalin were determined using the methods described in Chapter 5. Whereas the PIMC technique and the harmonic oscillator approximation were utilised for each molecule, the exact variational technique was employed only for the ethane molecule.

6.2 Simple Ethane Model

The ethane molecule served as a simple, yet realistic, system with which the various methods could be tested and compared. The PIMC technique, the variational technique, and the harmonic oscillator approximation were all employed for the simple ethane model. The analytical potential energy surface used for ethane was describe

in Section (5.1). A comparison of the PIMC technique and a quasi-harmonic technique for crystalline polyethylene has been previously reported.[117,148] However, the torsional path integral formulation employed in this study is computationally less expensive and will ultimately enable the study of larger molecular systems.

Figure (6.1) displays the numerical results for the internal energy at $273.15K$ obtained with the various methods. For the analytical potential, the variational result was easily obtained to high accuracy, and the reported variational value of 0.870 kcal mol^{-1} is presumed to be exact for the model. The PIMC algorithm was performed using one, two, three, four, five, and ten Trotter beads. As explained in Section (3.3.1), the PIMC method performed with one Trotter bead yields the classical statistical mechanics results. The one-bead PIMC result of 0.585 kcal mol^{-1} is accordingly labelled in the figure. Also included in Figure (6.1) is the value 0.883 kcal mol^{-1} of the internal energy predicted with the harmonic oscillator approximation.

Two important features of Figure (6.1) require discussion. The first is that, as required, the internal energy calculated using the PIMC algorithm converged with the increase of Trotter beads to the exact variational result. Whereas the majority of the quantum contribution to the internal energy is recovered upon inclusion of a second Trotter bead, greater than five beads is required, for this simple ethane model, to recover the exact result. Using ten Trotter beads, the PIMC method has converged to within numerical accuracy for the simple problem.

Of the entire alkane series, ethane is, in a sense, the most quantum mechanical. It is a standard result from the theory of quantum mechanical tunnelling that the wavefunction of a particle in a classically forbidden region diminishes exponentially with the square root of the mass of the particle.[150] Translating this observation to

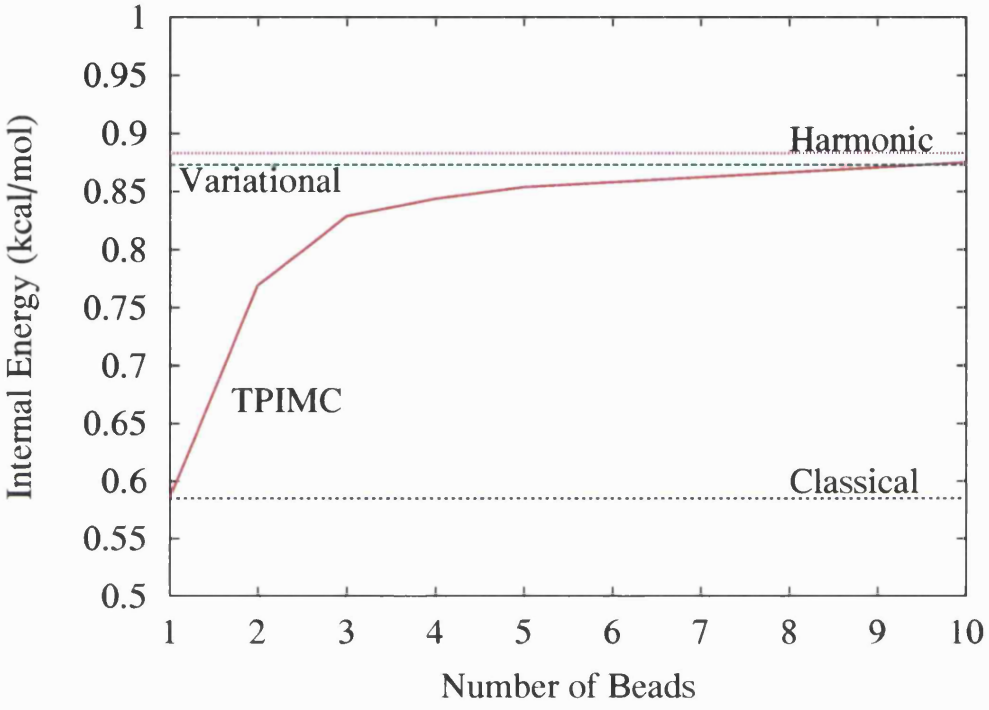


Figure 6.1: Calculated Internal Energy for Ethane Model at 273.15 K in kcal mol^{-1}

the current application, we note that the degree of quantum behaviour of torsions with large moments of inertia is expected to be drastically smaller than that of torsions with small moments of inertia. Suppose that I_1 and I_2 are the moments of inertia of two substituent groups about the bond that connects them. The moment of inertia about this torsion is then $\frac{I_1 I_2}{I_1+I_2}$. Presume, momentarily, that only the terminal methyl groups are small enough in an unbranched alkane to exhibit quantum behaviour. Define I_m to be the moment of inertia for a methyl substituent, and define I_x to be the considerably larger moment of inertia for the rest of the molecule. Therefore, for all unbranched alkanes longer than ethane, the molecule will exhibit two torsions with moment of inertia $I_{tor} = \frac{I_m I_x}{I_m+I_x} \sim \frac{I_m I_x}{I_x} = I_m$. On the other hand, the lone torsion in ethane will exhibit a moment of inertia equal to that of $\frac{I_m}{2}$. This single torsion in the ethane molecule is expected to contribute more to the quantum mechanical portion of the internal energy than the combined contributions of the two methyl torsions in a longer alkane chain.

Of course, the above analysis of moments of inertia is approximate, but it does lend credence to the likelihood that larger molecules will not require as many Trotter beads as the simple ethane model required for the PIMC value to converge to the exact quantum mechanical result. For larger systems, the PIMC value of internal energy is reported for the same array of Trotter beads, and we note the number of beads necessary to observe the plateau of convergence.

A second important feature of Figure (6.1) is that the harmonic oscillator performs very well in comparison to the exact variational result. As one might have predicted from a comparison of the harmonic and sinusoidal potentials, the steeper walls of the harmonic potential drive the calculated internal energy above the exact variational result. However, the relationship between the harmonic approximation

and variational internal energies is more complicated than this result suggests. In Figure (6.2), the calculated internal energies for the simple ethane model are presented as a function of temperature. In this figure, the PIMC results are obtained with 15 Trotter beads, and the classical results are obtained from the PIMC technique by utilising a single Trotter bead.

For both the low and high temperature extremes, the harmonic approximation behaves as expected. In the low temperature extreme, only the ground state of the system is occupied. The steeper curvature of the harmonic potential increases the approximated ground state energy with respect to the exact result. Figure (6.2) shows that the harmonic approximation internal energy of 0.7928 kcal mol^{-1} overestimates the exact variational result of 0.7212 kcal mol^{-1} as $T \rightarrow 0$. Note that, as predicted by the Trotter product formula in Equation (2.19), the PIMC result does not converge to the variational result in the low temperature extreme. Both the PIMC and classical results converge to the classical result of 0 kcal mol^{-1} as $T \rightarrow 0$. For any value of temperature greater than absolute zero, however, the PIMC result can be made to recover the variational result simply by increasing the number of Trotter beads in the calculation.

In the high temperature limit, the harmonic and variational results naturally diverge as the harmonic approximation of the sinusoidal potential deteriorates. With increasing quantum number, the exact wavefunctions of the sinusoidal potential converge to those of the free particle, not the harmonic oscillator. Increased thermal population of the high energy oscillator states in the harmonic approximation drive the approximated internal energy steadily higher than the exact result. Both the PIMC and classical results correctly converge to the exact result in the high temperature extreme as the error term in Trotter product formula goes to zero. As seen

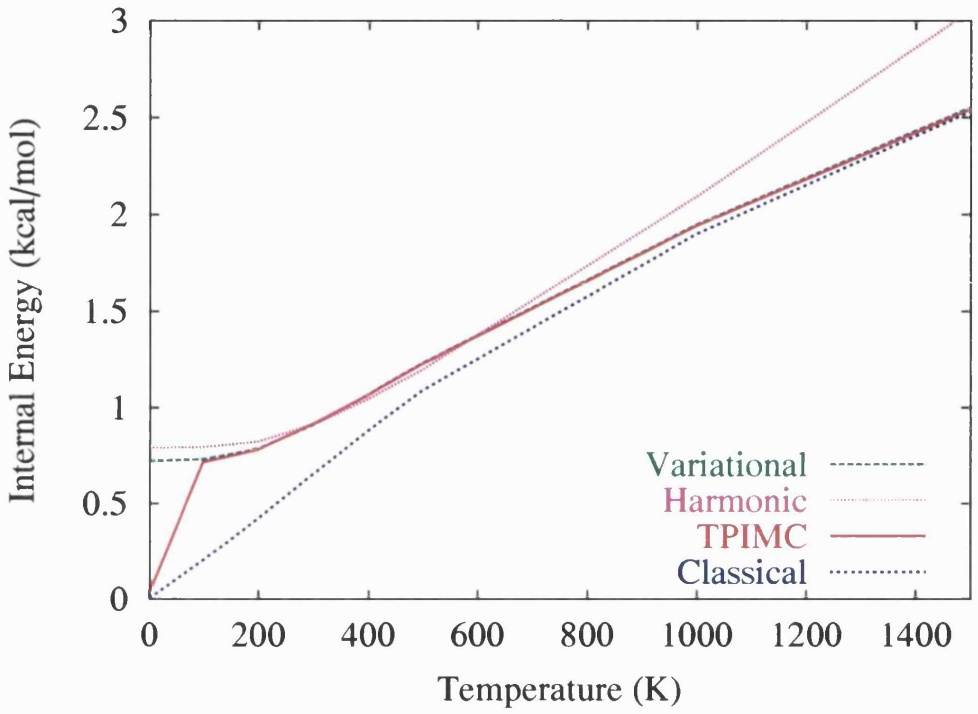


Figure 6.2: Internal Energy for Ethane Model Calculated at Various Temperatures

in Figure (6.2), the variational, PIMC, and classical internal energies all approach the value of 2.553 kcal mol^{-1} at 1500 K.

There exists in Figure (6.2) an intermediate temperature range for which the internal energy calculated with the harmonic approximation falls below the exact result. This temperature range is expanded in Figure (6.3). Also presented in Figure (6.3) are the harmonic approximation and exact variational results for the ethane model in which all of the hydrogen atoms have been replaced with deuterium atoms.

For the simple ethane model, the harmonic approximation falls below the exact internal energy between $300K$ and $600K$. Although the energies of the harmonic oscillator states are regularly spaced, such is not the case for the eigenfunctions of the sinusoidal potential. These exact eigenfunctions become more dense with increasing energy. Thus, there may exist a temperature range for which the increased number of energetically accessible quantum states in the sinusoidal potential raises the exact variational internal energy in excess of the harmonic approximation internal energy. Interestingly, this effect becomes more pronounced with increasing moments of inertia. It may be seen in Figure (6.3) that the temperature range for which the deuterated ethane exact and harmonic internal energies are inverted is extended to between $200K$ and $700K$. Figure (6.3) suggests that the harmonic approximation will become increasingly weak for larger moments of inertia.

It is worthwhile to note that in Figure (6.2), the PIMC results accurately reproduce the exact internal energies for all temperatures except those in the low-energy extreme. Performance of the PIMC technique at low temperatures can be improved by simply increasing the number of Trotter beads.

Using the PIMC technique, it is straightforward to obtain geometric information regarding the torsional angle distribution. In particular, quantum deviation

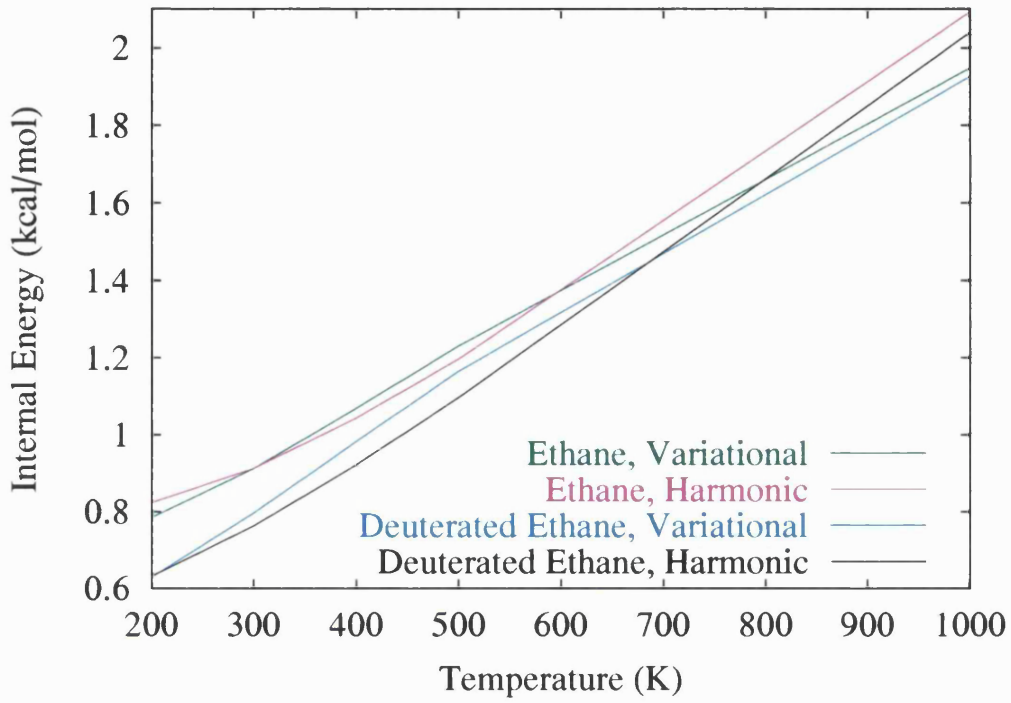


Figure 6.3: Internal Energies for Ethane and Deuterated Ethane Models Calculated at Various Temperatures

from the classical distribution is of interest. Although a more thorough analysis of the impact of quantum mechanics on intramolecular structure will be performed in our future research, a useful illustration can be provided here. Consider a harmonic oscillator potential fit to one of the three equivalent minima on the model ethane potential energy surface. Elementary expressions enable analytical calculation of both the classical and quantum root mean square (rms) values of the torsional angle position for this simplified system. This rms angle provides a measure of the spreading of the torsional angle distribution away from its mean value at the bottom of the potential well. The classical and quantum rms angles are 12.2 degrees and 15.5 degrees, respectively. The quantum tunnelling is expected to be even larger for the realistic periodic potential that exhibits coupled minima. This considerable torsional spreading provides evidence for the importance of quantum effects in intramolecular torsions.

6.3 Larger Molecules

In this section, we report the torsional PIMC and harmonic approximation results for the four molecules: ethane, *n*-butane, *n*-octane, and enkephalin. The MM3Pro force field was used for the potential of these molecules, and energy evaluations were performed by interfacing the TINKER molecular mechanics code to our torsional PIMC program.[7,139] Geometry optimisations were performed using TINKER to obtain suitable structures for the start of the PIMC calculations. In no case was a lower energy structure found during the course of the Monte Carlo simulation. The moments of inertia for each molecule were calculated at the beginning of the PIMC code and assumed to remain unchanged throughout the Monte Carlo simulation.

Table 6.1: Calculated Moments of Inertia of Ethane, n-Butane, and n-Octane in a.u. $\times 10^3$

	Ethane	n-Butane	n-Octane
1	2.99	5.70	5.94
2		34.8	65.3
3		5.70	104
4			177
5			104
6			65.3
7			5.94

The unbranched alkanes ethane, n-butane, and n-octane exhibit, respectively, one, three, and seven torsions. The calculated values of these moments of inertia are reported in Table (6.1). Note, in agreement with the suggestion in the previous section, the diminutive magnitude of the moment of inertia of ethane in comparison to those of the other molecules.

Enkephalin is a considerably larger molecule than the three considered alkanes. Displayed in Figure (6.4), it contains 75 atoms comprising 33 torsional degrees of freedom. However, only three of these torsions have moments of inertia less than 10^4 a.u. This feature portends less dramatic quantum mechanical behaviour for enkephalin than would be expected for more heavily branched molecules exhibiting a greater number of terminal methyl groups.

The internal energies of ethane, n-butane, n-octane, and enkephalin calculated

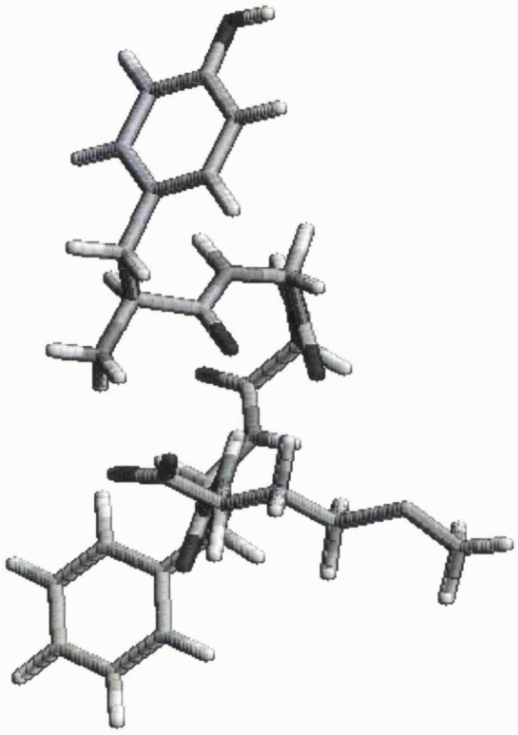


Figure 6.4: Structure of the Enkephalin Molecule

Table 6.2: Calculated Internal Energies[a] at 273.15 K in kcal mol^{-1}

P	Ethane	n-Butane	n-Octane	Enkephalin
1	0.594(1)	1.97(2)	4.99(5)	17.9(2)
2	0.758(1)	2.27(2)	5.26(8)	19.1(1)
3	0.810(1)	2.33(2)	5.32(8)	19.0(3)
4	0.830(2)	2.32(2)	5.34(7)	19.4(1)
5	0.841(2)	2.37(2)	5.35(7)	19.5(1)
10	0.855(4)	2.43(2)	5.34(8)	

[a]Numbers in parentheses indicate standard deviation in last significant digit.

with the torsion PIMC method are reported in Table (6.2). All calculations were performed at a temperature of $273.15K$, and the values are reported in kcal mol^{-1}. For the three alkanes, PIMC calculations were performed with one, two, three, four, five, and ten Trotter beads. The enkephalin internal energies were determined with only one, two, three, four, and five Trotter beads.

First of all, it is important to compare the results obtained for ethane with the MM3Pro potential implemented in TINKER, as opposed to the sinusoidal potential parametrized to MM3Pro discussed in the previous section. For the true MM3Pro potential, the PIMC method predicts classical and quantum energies of 0.594 kcal mol^{-1} and 0.86 kcal mol^{-1}, respectively, which are in reasonable agreement with the values of 0.585 kcal mol^{-1} and 0.875 kcal mol^{-1} reported for the model potential. This suggests correct implementation of the PIMC code for general molecular systems.

For each model, we note the number of beads required before the calculated

internal energy remains constant (within numerical accuracy) with respect to an increase in the number of beads. As observed for the model ethane system, at least five Trotter beads are required for the ethane internal energy to reach convergence. The other molecules, however, are more classical in behaviour. Neither *n*-butane, *n*-octane, nor enkephalin exhibit numerically significant changes in internal energy beyond the third Trotter bead. That is, a calculation with one Trotter bead is performed to obtain the classical internal energy, and another calculation with only two or three Trotter beads is needed to obtain the quantum internal energy.

Despite the small number of required Trotter beads, there is a substantial quantum contribution to the internal energy for every studied molecule. The quantum mechanical contributions to the internal energy of ethane and *n*-butane are, respectively, seen to comprise 31% and 19% of the total calculated internal energies. The significance of this proportion diminishes rapidly with the size of the molecules, however, and both *n*-octane and enkephalin are seen to have quantum contributions of only $\sim$ 7%. Nonetheless, the 1 kcal mol^{-1} of quantum internal energy in enkephalin is far from negligible and strongly suggests the need for consideration of quantum mechanical behaviour in large biomolecules.

In the study of large molecular systems, it is useful to know when the inclusion of a quantum mechanical description of nuclear motion is necessary. The de Broglie thermal wavelength λ_j of a given torsion provides such an indication of the relevance of quantum effects,[170]

$$\lambda_j = \left(\frac{2\pi\beta\hbar^2}{I_j}\right)^{1/2}. \tag{6.1}$$

Classical mechanics is sufficient when the de Broglie thermal wavelength is small in comparison to the characteristic length scale of the potential.[36] For the ethane torsion and a representative methyl torsion, the respective standard temperature

Table 6.3: PIMC and Harmonic Oscillator Approximation Internal Energies for Ethane, n-Butane, n-Octane and Enkephalin at 273.15K in kcal mol^{-1}

	Ethane	n-Butane	n-Octane	Enkephalin
PIMC	0.86	2.43	5.3	19.5
Harm. Osc.	0.86	2.08	4.27	19.14

wavelengths of 1.56 radians and 1.10 radians confirm the need for quantum mechanics. Equation (6.1) suggests that even torsions corresponding to substantially larger moments of inertia will require consideration of quantum mechanics at standard temperature.

Another comparison of the PIMC technique and the harmonic oscillator approximation is possible with the data presented in Table (6.3), which includes the PIMC results for ethane, n-butane, and n-octane with ten Trotter beads and the PIMC result for enkephalin with five Trotter beads. Also displayed are the results obtained with the harmonic oscillator approximation.

Similarly to the results reported in Figure (6.1) for the simple ethane model, the harmonic oscillator results in Table (6.3) perform well in comparison to the PIMC results. Nonetheless, as the size of the molecules increase, the larger moments of inertia lead the harmonic approximation to underestimate the actual internal energy (see Figure (6.3)). Although the impact of the deviation of the harmonic approximation is still relatively small for the molecules presented in Table (6.3), the systematic error could become more significant for larger molecules.

Chapter 7

Conclusions and Future Work

In this research, we successfully apply the uncoupled winding number formulation of path integral Monte Carlo theory to the torsional degrees of freedom in the molecules ethane, n-butane, n-octane, and enkephalin. This torsional PIMC technique offers a significant reduction in computational cost for systems in which vibrational degrees of freedom may be safely neglected. Employment of the PIMC method is simplified by the observation that contributions to calculated properties will be negligible for winding numbers greater than zero. For a simple ethane model potential, the PIMC result recovers the exact internal energy value obtained with a variational technique. For n-butane, n-octane, and enkephalin, the PIMC converged to the quantum mechanical limit with only two or three Trotter beads. All studied molecules exhibited significant quantum mechanical contributions to their internal energy expectation values according to the PIMC technique.

Although the harmonic oscillator approximation method for calculating the internal energy performs well for the molecules presented in this study, there exists a tendency in larger molecules for this approximation to underestimate the results

obtained with the more rigorous PIMC technique. It was illustrated with the simple ethane model that this underestimation is related to the size of the molecular moments of inertia, suggesting a systematic shortcoming in the application of the harmonic oscillator approximation to larger systems. Moreover, the harmonic approximation does not readily consider the coupling between torsions intrinsically accounted for with the PIMC technique.

Having confirmed the utility and applicability of the torsional PIMC technique to large molecular systems, extension of the approach to new applications is now conceivable. Firstly, the author intends to develop a Monte Carlo sampling algorithm that will enable the calculation of quantum free energies. Free energy is more difficult to calculate than the internal energy because it relies on the explicit evaluation of the canonical partition function, but the importance of free energies makes effort in this direction worthwhile.

Improved Monte Carlo convergence is another aspect of our calculations that will require improvement. The data regarding PIMC calculations of the enkephalin molecule clearly indicated that numerical accuracy was deteriorating at a disturbing rate. Without improvements in convergence, larger calculations will not be feasible. Encouragingly, some possible improvements are already being considered. Namely, the essentially "free-particle-like" or "harmonic-oscillator-like" behaviour of particular torsions suggests that an importance sampling scheme based on these *a priori* probability distributions will bear fruit. Also, it is recognised by the author that Monte Carlo convergence will improve with the reduction of Monte Carlo parameters being adjusted in the simulation. Replacement of the bead and chain step size parameters with a single step size parameter could go a long way to reducing statistical uncertainty.

Improved Monte Carlo convergence will bring numerous chemical problems within grasp. Obvious examples include single-molecule quantum free energy calculations on small proteins. By trivially constraining the motion of unimportant torsions, even larger proteins will be tractable. Straightforward extension of the torsional path integral Monte Carlo code to enable multi-molecular systems will allow calculations on ligand-docking and solvation effects. Consideration of quantum effects on torsional degrees of freedom will be particularly important for chemical problems such as these in which bonding connectivity is not altered but the conformational constraints on molecules are altered.

Chapter 8

Acknowledgements

I wish to thank Prof. Clary for his continued assistance and patience. I would also like to thank the rest of the Clary group, particularly Adam Farebrother for his help in introducing me to PIMC theory. I am also grateful to Dr. Jianshu Cao and Dr. Seogjoo Jang for their helpful discussions regarding the uncoupled winding number representation of torsional PIMC. Finally, I would like to thank the Marshall Aid Commemoration Commission for their generous funding.

Bibliography

[1] C Alhambra, K Byun, and J Gao. *ACS Symp. Ser.*, 712:35, 1998.

[2] C Alhambra, J C Corchado, M L Sanchez, J Gao, and D G Truhlar. *J. Amer. Chem. Soc.*, 122:8197, 2000.

[3] C Alhambra, J C Corchado, M L Sanchez, M Garcia-Viloca, J Gao, and D G Truhlar. *J. Phys. Chem. B*, 105:11326, 2001.

[4] C Alhambra, J Gao, J C Corchado, J Villa, and D G Truhlar. *J. Amer. Chem. Soc.*, 121:2253, 1999.

[5] C Alhambra, M L Sanchez, J Corchado, J Gao, and D G Truhlar. *Chem. Phys. Lett.*, 347:512, 2001.

[6] M P Allen and D J Tildesley. *Computer Simulation of Liquids.* Clarendon Press, Oxford, 1987.

[7] N L Allinger, F B Li, and L Q Yan. *J. Comp. Chem.*, 11:848, 1990.

[8] N L Allinger, F B Li, and L Q Yan. *J. Comp. Chem.*, 11:868, 1990.

[9] N L Allinger, Y H Yuh, and J H Lii. *J. Amer. Chem. Soc.*, 111:8551, 1989.

[10] J B Anderson. *J. Chem. Phys.*, 63:1499, 1975.

[11] J B Anderson. *Int. Rev. Phys. Chem.*, 11:85, 1995.

[12] P W Atkins. *Molecular Quantum Mechanics.* Oxford University Press, Oxford, 1987.

[13] E Balog, A L Hughes, and G J Martyna. *J. Chem. Phys.*, 112:870, 2000.

[14] J A Barker. *J. Chem. Phys.*, 70:2914, 1979.

[15] P A Bash, U C Singh, F K Brown, R Langridge, and P A Kollman. *Science*, 235:574, 1987.

[16] P A Bash, U C Singh, R Langridge, and P A Kollman. *Science*, 236:564, 1987.

[17] D M Benoit and D C Clary. *J. Chem. Phys.*, 113:5193, 2000.

[18] D M Benoit and D C Clary. *J. Phys. Chem. A*, 104:5590, 2000.

[19] M Benoit, D Marx, and M Parrinello. *Nature*, 392:258, 1996.

[20] U Berkert and N L Allinger. *Molecular Mechanics.* American Chemical Society, Washington, DC, 1982.

[21] B J Berne and D Thirumalai. *Annu. Rev. Phys. Chem.*, 37:401, 1986.

[22] B J Berne and D Thirumalai. *Quantum Simulations of Condensed Matter Phenomena.* World Scientific, Singapore, 1990.

[23] S F Billeter, S P Webb, T Iordanov, P K Agarwal, and S Hammes-Schiffer. *J. Chem. Phys.*, 114:6925, 2001.

[24] D Blume, M Mladenovic, M Lewerenz, and K B Whaley. *J. Chem. Phys.*, 110:5789, 1999.

[25] B R Brooks, R E Bruccoleri, B D Olafson, D J States, S Swaminathan, and M Karplus. *J. Comp. Chem.*, 4:187, 1983.

[26] C L Brooks and D A Case. *Chem. Rev.*, 93:2487, 1993.

[27] C L Brooks III, M Karplus, and B R Pettitt. *Adv. Chem. Phys.*, LXXI:1, 1988.

[28] V Buch. *J. Chem. Phys.*, 97:726, 1992.

[29] M P Calvo and J M Sanz-Sarna. *Numerical Hamiltonian Problems.* Chapman and Hall, London, 1994.

[30] J Cao. *Phys. Rev. E*, 49:882, 1994.

[31] J Cao and G A Voth. *J. Chem. Phys.*, 100:5106, 1994.

[32] R Car and M Parrinello. *Phys. Rev. Lett.*, 55:2471, 1985.

[33] D M Ceperley. *Adv. Chem. Phys.*, 93:1, 1996.

[34] D M Ceperley and M H Kalos. *Monte Carlo Methods in Statistical Physics.* Springer, Berlin, 1981.

[35] D Chandler. *J. Chem. Phys.*, 68:2959, 1978.

[36] D Chandler. *Introduction to Modern Statistical Mechanics.* Oxford University Press, Oxford, 1987.

[37] D Chandler. *Liquids, Freezing and Glass Transition.* Elsevier, Amsterdam, 1991.

[38] D Chandler and P G Wolynes. *J. Chem. Phys.*, 74:4078, 1981.

[39] Y Y Chuang and D G Truhlar. *J. Chem. Phys.*, 112:1221, 2000.

[40] T Clark. *A Handbook of Computational Chemistry.* Wiley, New York, 1985.

[41] D C Clary. *J. Chem. Phys.*, 114:9725, 2001.

[42] W D Cornell, P Cieplak, C I Bayly, I R Gould, K M Merz, D M Ferguson, D C Spellmeyer, T Fox, J W Caldwell, and P A Kollman. *J. Amer. Chem. Soc.*, 117:5179, 1995.

[43] C J Cramer and D G Truhlar. *J. Comput.-Aided Mol. Des.*, 6:629, 1992.

[44] C J Cramer and D G Truhlar. *Science*, 256:213, 1992.

[45] C Dellago, P G Bolhuis, and D Chandler. *J. Chem. Phys.*, 108:9236, 1998.

[46] C Dellago, P G Bolhuis, and D Chandler. *J. Chem. Phys.*, 110:6617, 1999.

[47] C Dellago, P G Bolhuis, F S Csajka, and D Chandler. *J. Chem. Phys.*, 108:1964, 1998.

[48] M J S Dewar and W Thiel. *J. Amer. Chem. Soc.*, 99:4899, 1977.

[49] M J S Dewar, E G Zoebisch, E F Healy, and J J P Stewart. *J. Amer. Chem. Soc.*, 107:3902, 1985.

[50] T G Dewey. *Phys. Rev. E*, 60:4652, 1999.

[51] P A M Dirac. *Physikalische Z. der Sowjetunion*, 3:64, 1933.

[52] M Diraison, G J Martyna, and M E Tuckerman. *J. Chem. Phys.*, 111:1096, 1999.

[53] S L Dixon and K M Merz, Jr. *J. Chem. Phys.*, 104:6643, 1996.

[54] C E Dykstra and T A Van Voorhis. *J. Comp. Chem.*, 18:702, 1997.

[55] R Elber. *Curr. Opin. Struct. Biol.*, 6:232, 1996.

[56] J Espinosa-Garcia, J C Corchado, and D G Truhlar. *J. Amer. Chem. Soc.*, 119:9891, 1997.

[57] I Feierberg, V Luzhkov, and J Aqvist. *J. Biol. Chem.*, 275:22657, 2000.

[58] R P Feynman. *Rev. Mod. Phys.*, 20:367, 1948.

[59] R P Feynman. *Statistical Mechanics.* Benjamin, New York, 1972.

[60] R P Feynman and A R Hibbs. *Quantum Mechanics and Path Integrals.* McGraw-Hill, New York, 1965.

[61] M J Field, P A Bash, and M Karplus. *J. Comp. Chem.*, 11:700, 1990.

[62] M Freindorf and J Gao. *J. Comp. Chem.*, 17:386, 1996.

[63] F Gago and W G Richards. *Mol. Pharmacol.*, 37:341, 1990.

[64] G Galli and M Parrinello. *Comput. Mater. Sci.*, 3:283, 1991.

[65] J Gao. *J. Phys. Chem.*, 96:537, 1992.

[66] J Gao. *J. Amer. Chem. Soc.*, 115:2930, 1993.

[67] J Gao. *ACS Symp. Ser.*, 569:8, 1994.

[68] J Gao. *J. Amer. Chem. Soc.*, 116:9324, 1994.

[69] J Gao. *in Reviews in Computational Chemistry, edited by K. B. Lipkowitz, D. B. Boyd.* VCH, New York, 1995.

[70] J Gao. *Acc. Chem. Res.*, 29:298, 1996.

[71] J Gao. *J. Comp. Chem.*, 18:1062, 1997.

[72] J Gao, L Cho, and A Auerbach. *Biophys. J.*, 65:43, 1993.

[73] J Gao and J J Pavelites. *J. Amer. Chem. Soc.*, 114:1912, 1993.

[74] J Gao and M A Thompson. *Combined Quantum Mechanical and Molecular Mechanical Methods.* Am. Chem. Soc., Washington, D C, 1998.

[75] J Gao and D G Truhlar. *Annu. Rev. Phys. Chem.*, 53:467, 2002.

[76] J Gao and X F Xia. *Science*, 258:631, 1992.

[77] M Garcia-Viloca, C Alhambra, D G Truhlar, and J Gao. *J. Chem. Phys.*, 114:9953, 2001.

[78] M J Gillan. *Phil. Mag. A*, 58:257, 1988.

[79] J K Gregory and D C Clary. *J. Phys. Chem.*, 100:18014, 1996.

[80] J K Gregory, D C Clary, K Liu, M G Brown, and R J Saykally. *Science*, 275:814, 1997.

[81] H Grubmuller, H Heller, A Windemeth, and K Schulten. *Mol. Simul.*, 6:121, 1991.

[82] A L Gughes, G J Martyna, and M Tuckerman. *J. Chem. Phys.*, 110:3275, 1999.

[83] P Guntert. *Q. Rev. Biophys.*, 31:145, 1998.

[84] K D M Harris and M Tremayne. *Chem. Mater.*, 8:2554, 1996.

[85] M F Herman, E J Bruskin, and B J Berne. *J. Chem. Phys.*, 76:5150, 1982.

[86] G Herzberg. *Infrared and Raman Spectra of Polyatomic Molecules.* Van Nostrand, London, 1945.

[87] A Hinchliffe. *Modelling Molecular Structures.* Wiley, New York, 1996.

[88] J K Hwang and A Warshel. *J. Phys. Chem.*, 97:10053, 1993.

[89] J K Hwang and A Warshel. *J. Amer. Chem. Soc.*, 118:11745, 1996.

[90] S Jang and G A Voth. *J. Chem. Phys.*, 112:8747, 2000.

[91] S Jang and G A Voth. *J. Chem. Phys.*, 114:1944, 2001.

[92] W Jorgensen and C Ravimohan. *J. Chem. Phys.*, 83:3050, 1985.

[93] W L Jorgensen. *Acc. Chem. Res.*, 22:184, 1989.

[94] W L Jorgensen and D L Severance. *J. Amer. Chem. Soc.*, 112:4768, 1990.

[95] W L Jorgensen and J Tirado-Rives. *J. Amer. Chem. Soc.*, 110:1657, 1988.

[96] M H Kalos. *Phys. Rev. A*, 2:250, 1970.

[97] M H Kalos, D Levesque, and L Verlet. *Phys. Rev. A*, 9:2178, 1974.

[98] M Karplus and G A Petsko. *Nature*, 347:631, 1990.

[99] M Karplus and E I Shaknovich. *Protein Folding.* Freeman, San Francisco, 1992.

[100] H Kleinert. *Path Integrals in Quantum Mechanics, Statistics and Polymer Physics.* World Scientific, Singapore, 1990.

[101] P A Kollman. *Chem. Rev.*, 93:2395, 1993.

[102] P A Kollman and K M Merz. *Acc. Chem. Res.*, 23:246, 1990.

[103] A R Leach. *Molecular Modelling Principles and Applications.* Longman, Singapore, 1986.

[104] C T Lee, W T Yang, and R G Parr. *Phys. Rev. B*, 37:785, 1988.

[105] T-S Lee, D M York, and W T Yang. *J. Chem. Phys.*, 105:2744, 1996.

[106] Y H Lee and B J Berne. *J. Phys. Chem. A*, 104:86, 2000.

[107] Y H Lee and B J Berne. *J. Phys. Chem. A*, 105:459, 2001.

[108] J P Lewis, C W Carter, Jr., J Hermans, W Pan, T-S Lee, and W T Yang. *J. Amer. Chem. Soc.*, 120:5407, 1998.

[109] J P Lewis, S B Liu, T-S Lee, and W T Yang. *J. Comp. Phys.*, 151:242, 1999.

[110] J H Lii and N L Allinger. *J. Amer. Chem. Soc.*, 111:8566, 1989.

[111] J H Lii and N L Allinger. *J. Amer. Chem. Soc.*, 111:8576, 1989.

[112] K Liu, M G Brown, C Carter, R J Saykally, J K Gregory, and D C Clary. *Nature*, 381:501, 1996.

[113] J Lobaugh and G A Voth. *Chem. Phys. Lett.*, 198:311, 1992.

[114] A D MacKerell, D Bashford, M Bellott, R L Dunbrack, J D Evanseck, M J Field, S Fischer, J Gao, H Guo, S Ha, D Joseph-McCarthy, L Kuchnir, K Kuczera, F T K Lau, C Mattos, S Michnick, T Ngo, D T Nguyen, B Prodhom, W E Rieher, B Roux, M Schlenkrich, J C Smith, R Stote, J Straub, M Watanabe, J Wiorkiewicz-Kuczera, D Yin, and M Karplus. *J. Phys. Chem. B*, 102:3586, 1998.

[115] A D MacKerell, Jr., J Wiorkiewicz-Kuczera, and M Karplus. *J. Amer. Chem. Soc.*, 117:11946, 1995.

[116] D E Makarov and M Topaler. *Phys. Rev. E*, 52:178, 1995.

[117] R Martoňák, W Paul, and K Binder. *J. Chem. Phys.*, 108:10278, 1998.

[118] G J Martyna and J Cao. *J. Chem. Phys.*, 104:2028, 1996.

[119] D Marx and M H Muser. *J. Phys.: Condens. Matter*, 11:R117, 1999.

[120] D Marx and P Nielaba. *Phys. Rev. A*, 45:8968, 1992.

[121] D Marx and M Parrinello. *Z. Phys. B*, 95:143, 1994.

[122] D Marx and M Parrinello. *Nature*, 375:216, 1995.

[123] D Marx and M Parrinello. *J. Chem. Phys.*, 104:4077, 1996.

[124] D Marx and M Parrinello. *Science*, 271:179, 1996.

[125] D Marx, M E Tuckerman, J Hutter, and M Parrinello. *Nature*, 397:601, 1999.

[126] D A McQuarrie. *Statistical Mechanics.* Harper and Row, New York, 1976.

[127] K M Merz and P A Kollman. *J. Amer. Chem. Soc.*, 111:5649, 1989.

[128] M Messina, G K Schenter, and B C Garrett. *J. Chem. Phys.*, 98:8525, 1993.

[129] M Messina, G K Schenter, and B C Garrett. *J. Chem. Phys.*, 103:3430, 1995.

[130] N Metropolis, A W Rosenbluth, M N Rosenbluth, A H Teller, and E Teller. *J. Chem. Phys.*, 21:1087, 1953.

[131] G Mills, G K Schenter, and D E Makarov. *Chem. Phys. Lett.*, 278:91, 1997.

[132] Y Mo and J Gao. *J. Phys. Chem. A*, 104:3012, 2000.

[133] G Nadig, L C Van Zant, S L Dixon, and K M Merz, Jr. *ACS Symp. Ser.*, 721:439, 1999.

[134] M Newman and G T Barkema. *Monte Carlo Methods in Statistical Physics.* Clarendon, New York, 1990.

[135] R G Parr and W Yang. *Density-Functional Theory of Atoms and Molecules.* Oxford University Press, Oxford, 1989.

[136] D Pearlman, J Caldwell, U C Singh, P K Weiner, and P Kollman. *Amber 4.0.* 1992.

[137] L Piela, J Kastrowicki, and H Scheraga. *J. Phys. Chem.*, 93:3339, 1989.

[138] K S Pitzer. *J. Chem. Phys.*, 14:239, 1946.

[139] J W Ponder. *TINKER Software Tools for Molecular Design, Version 3.8.* 2000.

[140] W Press. *Single-Particle Rotations in Molecular Crystals.* Springer, Berlin, 1981.

[141] B G Rao and U C Singh. *J. Amer. Chem. Soc.*, 112:3803, 1990.

[142] B G Rao and U C Singh. *J. Amer. Chem. Soc.*, 113:4381, 1991.

[143] S N Rao, U C Singh, P A Bash, and P A Kollman. *Nature*, 328:551, 1987.

[144] A K Rappe and C J Casewit. *Molecular Mechanics Across Chemistry.* University Science Books, Sausalito, 1997.

[145] D K Remler and P A Madden. *Mol. Phys.*, 70:921, 1990.

[146] B Roux and J M Karplus. *J. Amer. Chem. Soc.*, 115:3250, 1993.

[147] K J Runge, M P Surh, C Mailhiot, and E L Pollock. *Phys. Rev. Lett.*, 69:3527, 1992.

[148] G C Rutledge, D J Lacks, R Martoňák, and K Binder. *J. Chem. Phys.*, 108:10278, 1998.

[149] Jang S, Schwieters C D, and Voth G A. *J. Phys. Chem. A*, 103:9527, 1999.

[150] J J Sakurai. *Modern Quantum Mechanics.* Addison-Wesley, New York, 1994.

[151] T Schlick, E Barth, and M Mandziuk. *Annu. Rev. Biophys. Biomol. Struct.*, 26:181, 1997.

[152] L S Schulman. *Techniques and Applications of Path Integrations.* Wiley, New York, 1981.

[153] S Schweber. *Rev. Mod. Phys.*, 58:449, 1986.

[154] T Simonson, C F Wong, and A T Brunger. *J. Phys. Chem. A*, 101:1935, 1997.

[155] U C Singh and P A Kollman. *J. Comp. Chem.*, 7:718, 1986.

[156] R V Stanton, D S Hartsough, and K M Merz, Jr. *J. Phys. Chem.*, 97:11868, 1993.

[157] J J P Stewart. *J. Comput.-Aided Mol. Des.*, 4:1, 1990.

[158] J J P Stewart. *Theochem*, 401:195, 1997.

[159] A J Stone. *The Theory of Intermolecular Forces.* Oxford University Press, Oxford, 1996.

[160] J E Straub. *New Developments in Theoretical Study of Proteins.* World Scientific, Singapore, 1995.

[161] M A Suhm and R O Watts. *Phys. Rep.*, 204:293, 1991.

[162] M Suzuki. *Commun. Math. Phys.*, 56:183, 1976.

[163] M Suzuki. *Quantum Monte Carlo Methods in Equilibrium and Nonequilibrium Systems.* Springer, Berlin, 1987.

[164] A Szabo and N S Ostlund. *Modern Quantum Chemistry.* Dover, London, 1996.

[165] O Tapia, J M Lluch, R Cardenas, and J Andres. *J. Amer. Chem. Soc.*, 111:829, 1989.

[166] A Thomas, D Jourand, C Bret, P Amara, and M J Field. *J. Amer. Chem. Soc.*, 121:9693, 1999.

[167] S J Titmuss, P L Cummins, A A Bliznyuk, A P Rendell, and J E Gready. *Chem. Phys. Lett.*, 320:169, 2000.

[168] D J Tobias, G J Martyna, and M L Klein. *J. Chem. Phys.*, 101:4177, 1994.

[169] H F Trotter. *Proc. Am. Math. Soc.*, 10:545, 1959.

[170] D G Truhlar. *J. Comp. Chem.*, 12:266, 1991.

[171] D G Truhlar, J L Gao, C Alhambra, M Garcia-Viloca, J Corchado, M L Sanchez, and J Vill. *Acc. Chem. Res.*, 35:341, 2002.

[172] D G Truhlar, B C Garrett, and S J Klippenstein. *J. Phys. Chem.*, 100:12771, 1996.

[173] D G Truhlar, W L Hase, and J T Hynes. *J. Phys. Chem.*, 87:2664, 1983.

[174] D G Truhlar, W L Hase, and J T Hynes. *J. Phys. Chem.*, 87:5523, 1983.

[175] M Tuckerman, B J Berne, G J Martyna, and M L Klein. *J. Chem. Phys.*, 99:2796, 1993.

[176] M Tuckerman, G J Martyna, and B J Berne. *J. Chem. Phys.*, 97:1990, 1992.

[177] M E Tuckerman and G J Martyna. *J. Phys. Chem.*, 104:159, 2000.

[178] M E Tuckerman, D Marx, M L Klein, and M Parrinello. *J. Chem. Phys.*, 104:5579, 1996.

[179] M E Tuckerman, D Marx, and M Parrinello. *Nature*, 417:925, 2002.

[180] ME Tuckerman and D Marx. *Phys. Rev. Lett.*, 86:4946, 2001.

[181] J P Valleua and G M Torrie. *in Modern Theoretical Chemistry, edited by B J Berne.* Plenum, New York, 1977.

[182] A Van der Vaart and K M Merz, Jr. *J. Phys. Chem. A*, 103:3321, 1999.

[183] A Van der Vaart and K M Merz, Jr. *J. Amer. Chem. Soc.*, 121:9182, 1999.

[184] W F Van Gunsteren. *Protein Eng.*, 2:5, 1988.

[185] W F Van Gunsteren and A E Mark. *Eur. J. Biochem.*, 204:947, 1992.

[186] W F Van Gunsteren and P K Weiner. *Computer Simulation of Biomolecular Systems.* Escom, Leiden, 1989.

[187] A Vedani and D W Huhta. *J. Amer. Chem. Soc.*, 112:4759, 1990.

[188] J Villa and A Warshel. *J. Phys. Chem. B*, 105:7887, 2001.

[189] G A Voth, D Chandler, and W H Miller. *J. Chem. Phys.*, 91:7749, 1989.

[190] R C Wade and J A McCammon. *J. Mol. Biol.*, 225:679, 1992.

[191] R C Wade and J A McCammon. *J. Mol. Biol.*, 225:697, 1992.

[192] W Wang, O Donini, C M Reyes, and P A Kollman. *Annu. Rev. Biophys. Biomol. Struct.*, 30:211, 2001.

[193] A Warshel and M Levitt. *J. Mol. Biol.*, 103:227, 1976.

[194] P K Weiner and P A Kollman. *J. Comp. Chem.*, 2:287, 1981.

[195] S J Weiner, P A Kollman, D A Case, U C Singh, C Ghio, G Alagona, S Profeta, and P Weiner. *J. Amer. Chem. Soc.*, 106:765, 1984.

[196] P A Whitlock and M H Kalos. *J. Comp. Phys.*, 30:361, 1979.

[197] Hehre WJ, L Radom, PvR Schleyer, and J Pople. *Density-Functional Theory of Atoms and Molecules.* Wiley, New York, 1986.

[198] P G Wolynes, J N Onuchic, and D Thirumalai. *Science*, 267:1619, 1995.

[199] C F Wong and J A McCammon. *J. Amer. Chem. Soc.*, 108:3830, 1986.

[200] D M York, T-S Lee, and W T Yang. *Phys. Rev. Lett.*, 80:5011, 1998.

[201] R Zwanzig. *J. Chem. Phys.*, 22:1420, 1954.