
The Flying Ice Cube: Velocity Rescaling in Molecular Dynamics Leads to Violation of Energy Equipartition

STEPHEN C. HARVEY,¹ ROBERT K.-Z. TAN,¹
THOMAS E. CHEATHAM III^{2,*}

¹*Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, Birmingham, Alabama 35294*

²*Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143*

Received 31 July 1997; accepted 4 December 1997

ABSTRACT: This article describes an unexpected phenomenon encountered during MD simulations: velocity rescaling using standard protocols can systematically change the proportion of total kinetic energy (KE) found in motions associated with the various degrees of freedom. Under these conditions, the simulation violates the principle of equipartition of energy, which requires a mean kinetic energy of $RT/2$ in each degree of freedom. A particularly pathological form of this problem occurs if one does not periodically remove the net translation of (and rotation about) the center of mass. In this case, almost all of the kinetic energy is converted into these two kinds of motion, producing a system with almost no kinetic energy associated with the internal degrees of freedom. We call this phenomenon "the flying ice cube." We present a mathematical analysis of a simple diatomic system with two degrees of freedom, to document the origin of the problem. We then present examples from three kinds of MD simulations, one being an *in vacuo* simulation on a diatomic system, one involving a low resolution model of DNA *in vacuo*, and the third using a traditional all-atom DNA model with full solvation, periodic boundary

*Present address: National Institutes of Health, Bethesda, MD
20892-5626

Correspondence to: S. C. Harvey; e-mail: harvey@neptune.cmc.uab.edu

Contract/grant sponsor: National Institutes of Health; contract/grant number: HL-34343, CA-25644

Contract/grant sponsor: Pittsburgh Supercomputing Center; contract/grant number: MCA935017P

conditions, and the particle mesh Ewald method for treating long-range electrostatics. Finally, we discuss methods for avoiding the problem.
© 1998 John Wiley & Sons, Inc. J. Comput Chem 19: 726–740, 1998

Keywords: molecular dynamics; energy equipartition; macromolecular simulations; constant temperature simulations

Introduction

Molecular dynamics (MD) simulations^{1,2} are widely used to investigate questions of macromolecular structure and dynamics. Properly parameterized, they can provide important kinetic and thermodynamic information on a wide range of problems.

The MD algorithm numerically integrates Newton's equations of motion. Once the trajectory is begun, a free MD simulation should correspond to an ensemble that has constant total energy. There are two problems with such simulations. First, the gradual accumulation of unavoidable numerical errors may lead to drifts in the total energy. Second, experiments are usually conducted at constant temperature and pressure, which corresponds to the isothermic–isobaric ensemble, not a constant energy ensemble.

To correct for these problems, standard MD packages offer different tools for maintaining constant temperature. Among these are periodic velocity rescaling and other methods that uniformly modify the atomic velocities. For periodic velocity rescaling, the total kinetic energy is determined at regular intervals, and the instantaneous effect temperature, T , is calculated by noting that the total kinetic energy is $3NRT/2$, where N is the number of atoms, R is the universal gas constant, and the factor of three arises from the fact that each atom has three degrees of freedom. (Strictly speaking, this relationship only holds for the average kinetic energy, calculated over a sufficiently long time that the system represents the appropriate thermodynamic ensemble.) The ratio between the target temperature and the instantaneous temperature is used to calculate a factor by which all velocities are multiplied. One can use a factor that instantaneously adjusts the temperature to the target value, or one can use some fraction of this factor, treating the system as if it were coupled to a temperature bath with a characteristic relaxation time for the transmission of energy between the bath and the system.³

We have found that periodic velocity rescaling leads to an unexpected problem: a gradual bleeding of KE from high frequency motions such as bond stretching and angle bending into low-frequency motions. This represents a violation of the principal of equipartition of energy, which requires that each degree of freedom has the same mean kinetic energy. We made this discovery independently of one another on two very different kinds of systems. In one, AMBER⁴ was used to simulate an all-atom model with full solvation, whereas, in the other, *in vacuo* simulations were carried out on a low-resolution model of DNA using YAMMP.⁵ A similar phenomenon was independently observed at Wesleyan University (M. A. Young and D. L. Beveridge, personal communication) during all-atom simulations with full solvation. We therefore concluded that this problem is inherent in velocity rescaling.

This article has three sections. In the first, we present an analytical treatment of the effects of velocity rescaling on a very simple model system. This provides an introduction to the phenomenon and a rigorous explanation for its cause. The second section presents some manifestations of this problem that we have observed during different kinds of MD simulations. The final section discusses precautionary measures that can be taken to avoid or minimize problems associated with velocity rescaling.

Analysis of Problem

Let us consider what happens during an MD simulation with periodic velocity rescaling in a simple one-dimensional system with two particles, which we can treat analytically. The system has two degrees of freedom. One of these corresponds to the translational motion of the center of mass, which has kinetic energy E_0 . The other corresponds to the vibrational motion.

Let us suppose that the potential energy of the single internal degree of freedom, x , depends quadratically on the deviation of x from its ideal

value, x_0 . The most common example of this is a bond stretching term with force constant k . (Following convention, we have incorporated the Hooke's law factor of 0.5 into k .) We denote the potential energy of this degree of freedom by U :

$$U(x) = k(x - x_0)^2 \quad (1)$$

The kinetic energy for the internal degree of freedom is denoted by E_i , where the subscript indicates "internal." x and its derivatives are oscillatory, and so is E_i . We denote the angular frequency of the kinetic energy oscillations by ω . It is easily shown that:

$$E_i(t) = \varepsilon(1 + \sin \omega t)$$

where ε indicates the mean kinetic energy of the internal degree of freedom, $\langle E_i \rangle$. We use $E(t)$ (without a subscript) to indicate the total kinetic energy:

$$E(t) = E_0 + E_i(t) = E_0 + \varepsilon(1 + \sin \omega t)$$

and we designate the average kinetic energy by $\bar{E} = E_0 + \varepsilon$. The purpose of periodically reassigning or rescaling velocities is, of course, to remove drifts in \bar{E} , so that, over the duration of the MD simulation, the mean kinetic energy is close to the value required by the temperature specified for the simulation. This requires that:

$$\bar{\bar{E}} = \left(\frac{n}{2}\right)RT$$

where n is the number of degrees of freedom (two in the present example), and the double bar indicates an average over the duration of the MD simulation.

In the discussion that follows, we will assume that rescaling follows the original procedure developed about 20 years ago and used in many programs. In this protocol, the instantaneous kinetic energy is calculated, and then velocities are rescaled by a factor $\sqrt{\alpha}$, where α is equal to the desired mean kinetic energy divided by the instantaneous kinetic energy. In the present situation, with two degrees of freedom:

$$\alpha = RT/E$$

Note that, because E is time-dependent, α also depends on time.

The results described in what follows are no different if, rather than instantaneously jumping the kinetic energy all the way to its target value,

rescaling is more gradual and is based on coupling to a temperature bath.³ The rescaling factor in the coupling protocol is only a fraction of the value of α specified here, but the direction of rescaling (increasing or decreasing velocities) is the same, and the effect over very long simulations is also the same.

The effect of rescaling is to immediately give the system an instantaneous kinetic energy corresponding to the target temperature. However, even in the absence of numerical errors, the mean kinetic energy during the interval of free MD from this rescaling to the next rescaling will not be equal to RT . This is because rescaling can happen at any point in the cycle of the intramolecular motion, and α is calculated from the instantaneous value of the kinetic energy. The system has constant total energy, but kinetic energy oscillates. If the rescaling occurs when the internal coordinate is far from its ideal value, the kinetic energy will be small (and α will be larger), whereas, if rescaling happens when the internal coordinate is near its ideal value, the kinetic energy will be large (and α will be smaller).

One would expect that, if the kinetic energy oscillates about a value near RT , there should be about a 50% probability that the instantaneous kinetic energy is less than RT , so that $\alpha > 1$, and the probability that $\alpha < 1$ would also be nearly 50%. Furthermore, because E_i and E_0 are scaled by the same amount, one would intuitively expect that there should be no change in the fraction of total kinetic energy that is found in the internal degrees of freedom versus the kinetic energy associated with the center of mass motion. In other words, one would expect the principal of equipartition of energy to be preserved. The purpose of this work is to show that this is not the case.

This unexpected result arises from the fact that the rescaling factor α depends on the instantaneous value of the kinetic energy at the time of rescaling, and α has different effects on the two kinds of kinetic energy. Let us use primes to designate the kinetic energies after rescaling, and angular brackets to denote expected values. First, the expected value of the kinetic energy for the center of mass motion is

$$\langle E'_0 \rangle = \langle \alpha E_0 \rangle = \langle \alpha \rangle E_0 \quad (2)$$

where the removal of E_0 from the angular brackets is possible because the center of mass motion is uniform throughout the oscillatory cycle. In contrast, the kinetic energy associated with the inter-

nal degree of freedom has expectation value:

$$\langle E'_i \rangle = \langle \alpha E_i \rangle \quad (3)$$

and E_i cannot be taken outside the angular brackets, because it is time-dependent.

α depends on the point in the cycle at which rescaling takes place. So, to calculate expectation values, we consider α to be a time-dependent variable and derive the expectation values by averaging over one full cycle of the bond vibration. Thus:

$$\begin{aligned} \alpha(t) &= \frac{RT}{E_0 + \varepsilon(1 + \sin \omega t)} = \frac{RT}{\bar{E} + \varepsilon \sin \omega t} \\ &= \frac{RT}{\bar{E}} \frac{1}{1 + c \sin \omega t} \end{aligned} \quad (4)$$

where we have introduced the abbreviation $c = \varepsilon/\bar{E}$. Using ωt as the variable of integration and averaging over one cycle, we get:

$$\begin{aligned} \langle \alpha \rangle &= \frac{1}{2\pi} \frac{RT}{\bar{E}} \int_0^{2\pi} \frac{1}{1 + c \sin(\omega t)} d(\omega t) \\ &= \frac{RT}{\bar{E}} (1 - c^2)^{-1/2} \end{aligned}$$

Combining this result with eq. (2) gives, on rearranging:

$$\frac{\langle E'_0 \rangle}{E_0} = \frac{RT}{\bar{E}} (1 - c^2)^{-1/2} \quad (5)$$

This gives a hint at the problem: if the mean kinetic energy is exactly on target (i.e., if $\bar{E} = RT$), we would anticipate that the kinetic energy of translation should be preserved (i.e., $\langle E'_0 \rangle/E_0 = 1$), but this is not so. The scaling factor, which should be equal to one in this case, is actually $(1 - c^2)^{-1/2}$. Because c is always less than one, the expectation value after rescaling, $\langle E'_0 \rangle$, is larger than E_0 , at least when $\bar{E} = RT$. This results in an increase in the kinetic energy of translation.

Let us now calculate the expectation value for the kinetic energy associated with the internal degree of freedom. Substituting eqs. (1) and (4) into eq. (3) gives:

$$\langle E'_i \rangle = \langle \alpha E_i \rangle = \left\langle \frac{RT}{\bar{E}} \frac{\varepsilon(1 + \sin \omega t)}{1 + c \sin \omega t} \right\rangle$$

Dividing through by the mean kinetic energy ε , and calculating the average over one full cycle

gives:

$$\frac{\langle E'_i \rangle}{\langle E_i \rangle} = \frac{\langle \alpha E_i \rangle}{\varepsilon} = \frac{1}{2\pi} \frac{RT}{\bar{E}} \int_0^{2\pi} \frac{1 + \sin(\omega t)}{1 + c \sin(\omega t)} d(\omega t) \quad (6)$$

We are unable to evaluate this integral analytically, but it can be integrated numerically using *Mathematica* (Wolfram Research, Inc., Champaign, IL). When divided by the normalization factor 2π , a value less than one is always obtained. If the mean energy were right on target (i.e., if $\bar{E} = RT$), then we would expect that the kinetic energy associated with the internal degree of freedom should remain unchanged (i.e., $\langle \alpha E_i \rangle/\varepsilon = 1$), but this is not the case: this component of the kinetic energy would actually be reduced.

Table I compares the effects of rescaling on the translational kinetic energy, eq. (5), with its effect on the kinetic energy of the internal degree of freedom, eq. (6). The independent variable is the fraction of the total kinetic energy associated with the internal degree of freedom over the time interval since the previous rescaling, $c = \varepsilon/\bar{E}$. For all

TABLE I. Effect of Rescaling on Distribution of Kinetic Energy As a Function of Fraction of Total Kinetic Energy Associated with Internal Degree of Freedom ($c = \varepsilon/\bar{E}$).

c	$\frac{\langle E'_i \rangle}{\langle E_i \rangle}$	$\frac{\langle E'_0 \rangle}{E_0}$
0.01	0.99	1.00005
0.1	0.95	1.005
0.2	0.90	1.02
0.3	0.85	1.05
0.4	0.79	1.09
0.5	0.73	1.15
0.6	0.67	1.25
0.7	0.59	1.40
0.8	0.50	1.67
0.9	0.37	2.29
0.99	0.13	7.09

On rescaling, the kinetic energy of the internal degree of freedom, E_i , is decreased (column 2), whereas the kinetic energy of the overall translation, E_0 , is increased (column 3). Unprimed and primed values are prior to and immediately after rescaling respectively. Angular brackets indicate either the mean value (prior to rescaling) or the expectation value (after rescaling). E_0 is not oscillatory, so $\langle E_0 \rangle = E_0$. The boldface values at $c = 0.5$ indicate what would happen if equipartition were instantaneously satisfied prior to rescaling. See text for discussion of this case.

values of c over the physically accessible range, $0 < c < 1$, there is a net conversion of kinetic energy from the internal degree of freedom to the translation of the center of mass.

With repeated rescalings, there will be a gradual loss of vibrational kinetic energy, and a corresponding increase in the translational kinetic energy. Eventually, the amplitude of the vibrations shrinks to nearly zero, as does the effective temperature for the internal motions, and the molecule becomes a "flying ice cube." This phenomenon is demonstrated in the simulation on the model diatomic system in the next section.

In more complex systems, repeated velocity rescaling will produce a gradual loss of kinetic energy associated with the high-frequency motions, and a concomitant increase in kinetic energy associated with the zero frequency motions (global translation and rotation). Other degrees of freedom may also experience a gradual buildup of kinetic energy, if the periods of their characteristic vibrations are long by comparison to the time interval between successive rescalings. For such low frequency motions, the expectation value for the relative energy after rescaling is given by an expression resembling eq. (6), whereas the expectation value for higher frequency motions resembles eq. (5). Thus, a given degree of freedom will see a net gain or loss of kinetic energy after many rescalings, depending on whether it is a low frequency or high frequency motion, relative to the frequency of rescaling.

Examples from MD Simulations

MODEL DIATOMIC SYSTEM: ETHANE (UNITED ATOM MODEL)

We have carried out MD simulations on a united atom model for ethane using the YAMMP package.⁵ Because the simulations were done in three-dimensional space, the system has six degrees of freedom. As a consequence, the mathematical analysis done in the previous section does not apply exactly, but the net transfer of kinetic energy from the high-frequency vibrational motion to the zero frequency motions is clearly demonstrated.

The model consists of two pseudoatoms, each representing a methyl group and having a mass of 15 amu. The force field parameters of eq. (1) are $x_0 = 1.54 \text{ \AA}$ (ideal bond length) and $k = 240 \text{ kcal/mol} \cdot \text{\AA}^2$ (bond force constant). With a target

temperature of 300 K and six degrees of freedom, the total target kinetic energy is $6RT/2 = 1.8 \text{ kcal/mol}$.

The period of free vibration for this system is:

$$\tau = 2\pi\sqrt{\mu/k}$$

where μ is the reduced mass, $\mu = m_1m_2/(m_1 + m_2) = m/2$, and $m = m_1 = m_2$ is the mass of the methyl group. Substituting the above values into these relationships gives $\tau = 38.4 \text{ fs}$.

A timestep of 1 fs was used in the simulations, and velocity rescaling was done at intervals of 100 fs. This interval is longer than the period of vibration and is not simply related to the period, so that, in a very long simulation, rescaling occurs at all possible points in the oscillation cycle. The effect of rescaling on the vibrational kinetic energy is therefore described by eq. (6).

Figures 1–3 show the time dependence of the bond length, kinetic energy, and potential energy, respectively. If we examine only the kinetic energy, it appears that rescaling is doing exactly what it should: depending on the exact point in the cycle where rescaling occurs, the mean kinetic energy during the next free interval of MD may be larger or smaller than in the previous interval, but the mean kinetic energy, averaged over the total trajectory, will converge to 1.8 kcal/mol after many rescalings. On the other hand, the energy associated with the internal degree of freedom (bond stretching) is gradually reduced, as is seen in the declining amplitude of the bond length vibration (Fig. 1) and in the decaying mean potential energy (Fig. 3). Note that there are occasions when the rescaling causes the internal energy to increase slightly (at $t = 0.1, 0.2, 0.5, 0.9 \text{ ps}$), but when it drops, it often does so sharply (at $t = 0.3, 0.8 \text{ ps}$).

Lest there be any doubt that the bond vibrational energy continues to decline, we have run another simulation with identical parameters, but a different initial random number seed, and saved the structures over five 200-fs windows, with successive windows separated by 100 ps. A plot of bond length versus time for this simulation (Fig. 4) shows that the amplitude of the bond vibration is sharply reduced after only 100 ps, and it is essentially zero after about 300 ps. There is no vibrational kinetic energy, and all the kinetic energy is in the translation of the center of mass and rotation about the center of mass. Note that the mean bond length (1.543 Å), is larger than the ideal bond length (1.540 Å), because of the centrifugal force of the molecular rotation.

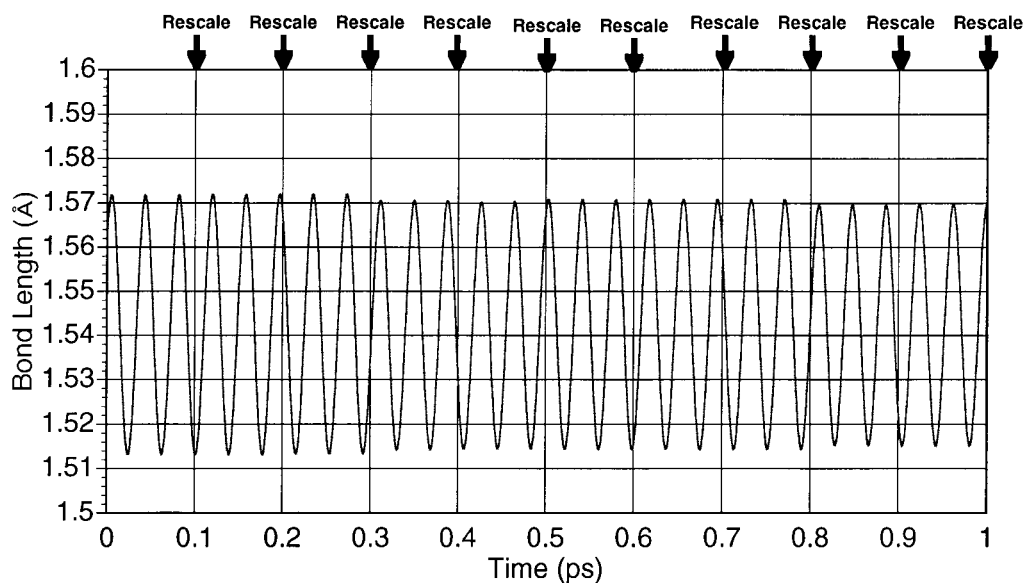


FIGURE 1. Time dependence of the bond length of the ethane model with two united atoms, during an MD simulation with periodic rescaling. The gradual decay of the vibrational amplitude is evident. The period of the oscillation is 38.4 fs. The overall molecular rotation produces a centrifugal force that tends to stretch the bond, so the mean bond length is slightly greater than the ideal bond length of 1.54 Å.

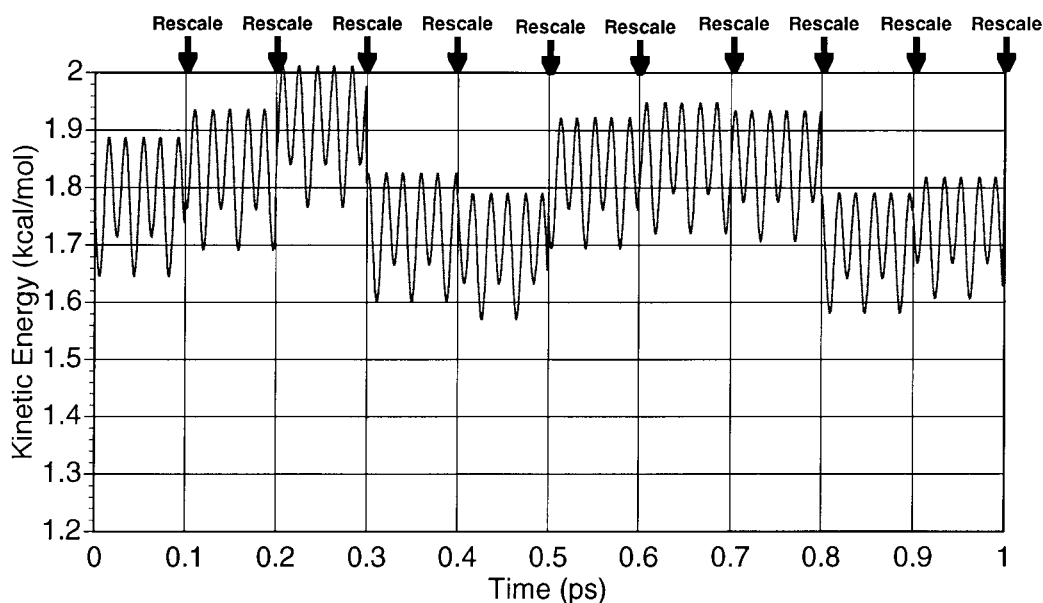


FIGURE 2. Time dependence of the kinetic energy of the ethane model. The KE goes through two peaks and two valleys during each oscillatory cycle. The maxima occur when the potential energy goes to zero (i.e., the bond length is at its ideal value). The minima in KE occur at the extreme values of the bond length. Conservation of angular momentum causes the rotational kinetic energy to be larger when the bond is fully compressed than when it is fully stretched. (Compare the valleys in this figure to the peaks and valleys in Fig. 1.)

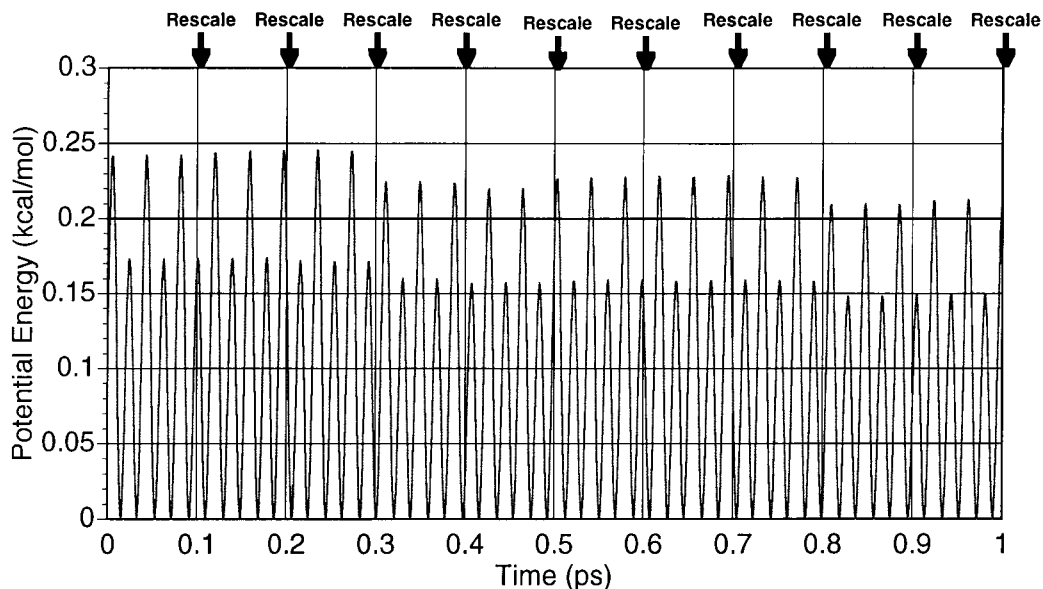


FIGURE 3. Time dependence of the potential energy of the ethane model. The PE goes through two peaks and two valleys during each oscillatory cycle. $PE = 0$ when the bond length has its ideal value (i.e., when $x = x_0$). The centrifugal force of overall rotation causes the mean bond length to be greater than the ideal length, so the potential energy is larger when the bond is stretched than it is when it is compressed. (Compare the peaks in this figure to the peaks and valleys in Fig. 1.)

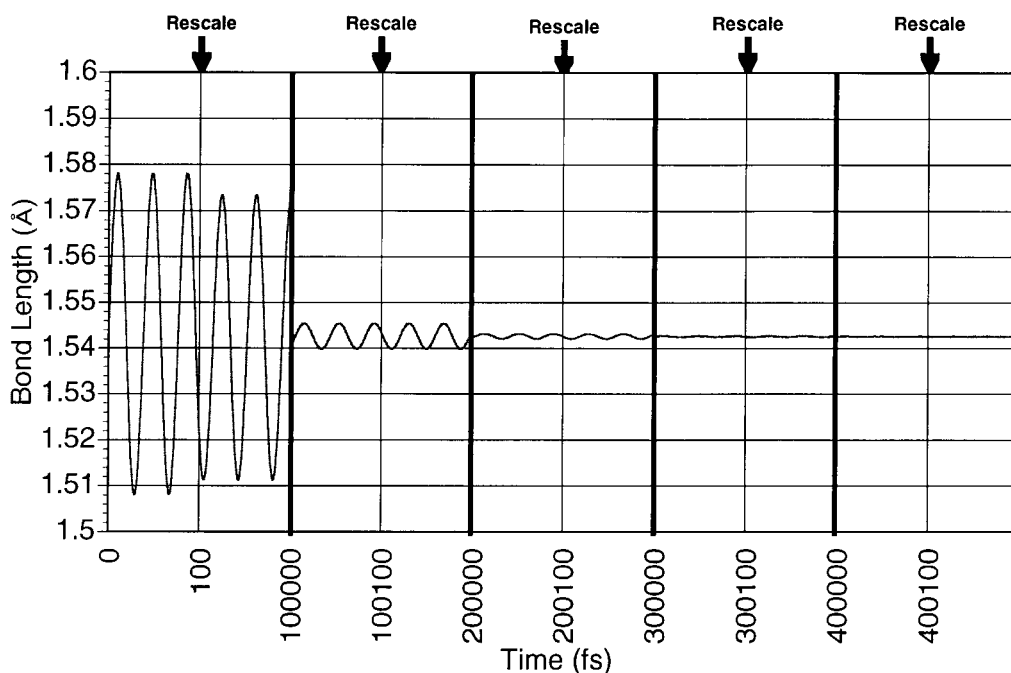


FIGURE 4. Time dependence of the bond length of the ethane model in an MD simulation covering 400.2 ps, with periodic velocity rescaling. Five 200-fs windows at intervals of 100 ps are shown, and the rescaling event at the center of each window is indicated. The vibrational energy is severely damped by 100 ps and is essentially zero by 300 ps.

MORE COMPLEX ELASTIC SYSTEM: SUPERCOILED DNA

In closing the section "Analysis of the Problem," we argued that, in addition to the buildup of kinetic energy in zero frequency motions, energy might also increase in other motions whose characteristic frequencies are smaller than the frequency of velocity rescaling. We have encountered interesting examples of this phenomenon in simulations on supercoiled DNA.

Double-helical DNA molecules in the size range of several hundred to several thousand basepairs (bp) are elastic polymers. Given the elastic moduli of DNA, it is easier to bend the molecule than it is to twist it. As a consequence, when twisting deformations are applied to DNAs in this size range, they are inclined to relieve the torsional stresses by bending, producing nonplanar writhed configurations.

A *relaxed closed circular DNA* is formed when the molecule is covalently closed into a circle in such a way that no torsional stress is introduced; the equilibrium conformation is circular, and the linking number (the number of times the two strands wrap around one another) has its optimal value, Lk_0 . If the molecule is closed into a circle with any other linking number (i.e., if $Lk \neq Lk_0$), then the equilibrium conformation is *supercoiled*. It resembles a figure eight, except that there are multiple crossings. The number of crossings is approximately equal to the writhe, Wr , which is proportional to the linking deficit. For a molecule with the elastic properties of DNA, $Wr \approx 0.7(Lk - Lk_0)$.

We have developed a reduced representation model for double-helical DNA that is useful for studying the structure of supercoiled molecules.⁶ The model, called 3DNA, uses three pseudoatoms to define the plane of each basepair. Successive basepairs are held together by pseudobonds, and all of the terms in the force field are harmonic (the energy depends quadratically on the coordinate). These include bonds, angles, improper torsions, and nonbonded volume exclusion terms. Force constants for the bonds, angles, and torsions were chosen to mimic the known elastic moduli for DNA bending and twisting.

This is thus a complex elastic model, and the many degrees of freedom have a wide range of characteristic frequencies. If velocities are rescaled often enough, the analysis of the previous section would predict a gradual transfer of kinetic energy

from the high-frequency vibrational motions to the zero-frequency and low-frequency degrees of freedom.

We have carried out a series of long molecular dynamics simulations (500 ns or more) on large closed circular DNAs. For both relaxed and supercoiled molecules, a wide range of conformations is observed when the simulations are run using free MD, MD with infrequent velocity rescalings, or MD with infrequent velocity reassignments. Any of these procedures generates a thermodynamic ensemble of structures whose average properties generally differ significantly from those of the equilibrium structure.⁷ In particular, relaxed molecules almost never show an open, circular conformation, and interwound molecules are rarely as extended as their equilibrium structures.

The behavior changes drastically when frequent velocity rescaling is used, as the kinetic energy from hundreds or thousands of high-frequency vibrational modes is bled into the zero-frequency and low-frequency motions. Here we describe results of MD with coupling to a constant temperature bath, following the algorithm of Berendsen et al.³ Because of the large masses and small force constants of the 3DNA model, the simulations described here use a timestep of 31.25 fs. Berendsen rescaling is done at every timestep, with a characteristic time for coupling to the temperature bath of 1 ps (32 timesteps).

For a relaxed closed circular molecule, the buildup of rotational kinetic energy causes the structure to open into a spinning circular conformation, like a wheel. The behavior of underwound molecules is determined partly by the magnitude of the linking deficit (how tightly the molecule is interwound), and partly by chance (the relationship between the initial distribution of velocities and the gradually accumulating centrifugal force). Sometimes these molecules open into a spinning circular conformation, like the relaxed molecules; this is particularly likely if the linking deficit is small. More highly underwound molecules are prone to developing extended interwound conformations, closely resembling the equilibrium structure. An example of this is shown in Figure 5.

With 1500 bp and 4500 pseudoatoms, the molecule in Figure 5 has 13,500 degrees of freedom (df). After several hundred nanoseconds, the kinetic energy associated with all of these is mostly converted into global translational and rotational motions, which comprise only 6 df . But there is one other zero-frequency mode that is also excited

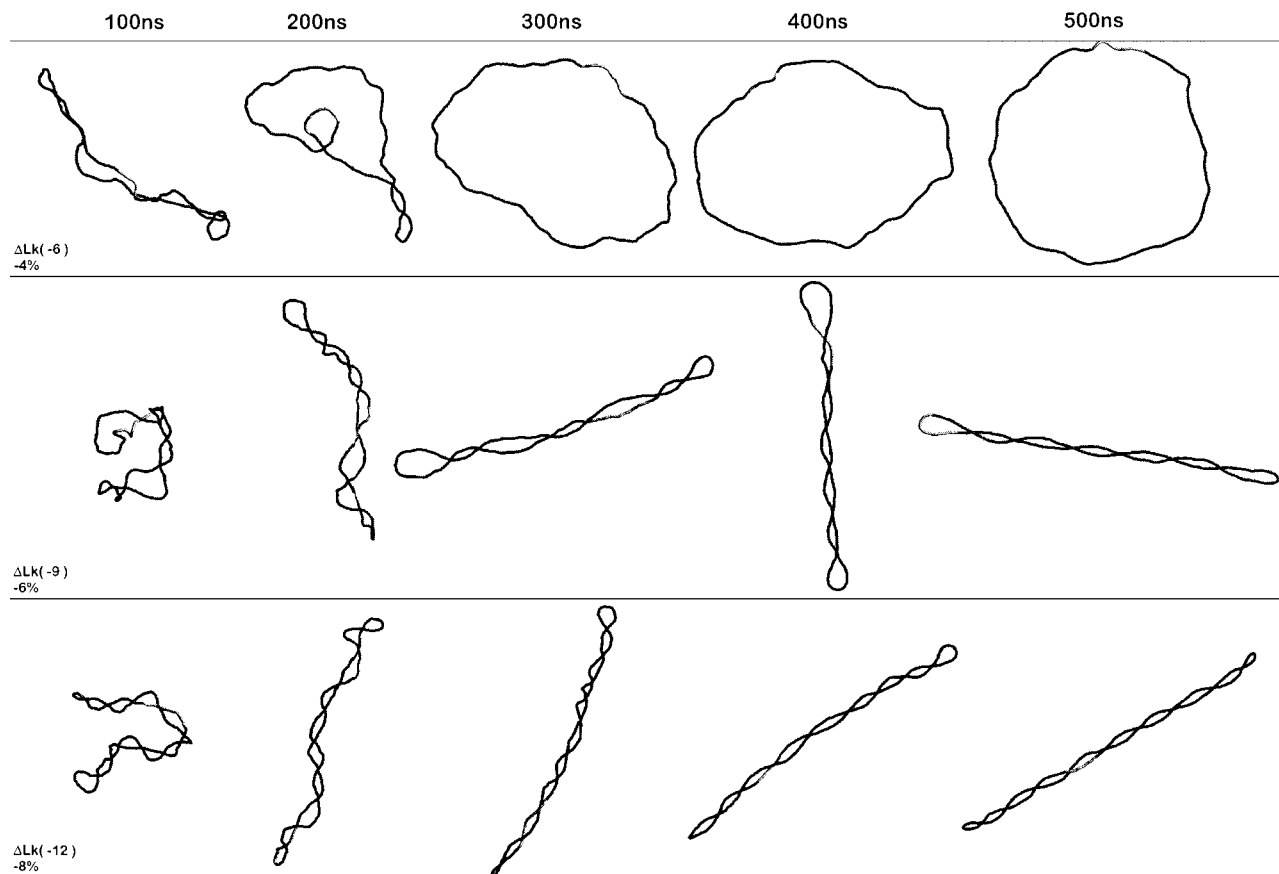


FIGURE 5. Evolution of the structure of model supercoiled DNAs in long MD simulations with coupling to a constant temperature bath. The model has 1500 bp. The expected writhe is about 70% of the linking deficit. This is indeed the case after a long time for two cases ($\Delta Lk = -9$ and $\Delta Lk = -12$), but the rotational motion of the molecule with $\Delta Lk = -4$ causes it to open into a circular form. After several hundred nanoseconds, almost all the kinetic energy in all three models is associated with the translation of the center of mass and rotational motion about the center of mass.

by the rescaling. This motion, often called *slithering*, is best described by considering one of the interwound configurations at $t = 500$ ns (Fig. 5). Although the conformation appears fixed, the position of a given basepair within the structure is time-dependent, and slithering refers to the motion of a given basepair around the track defined by the interwound conformation. A formal way of following the slithering motion is the *loop tip profile*, which plots the numbers identifying the two basepairs at the two ends of the interwound structure, as a function of time.⁸

Loop tip motions in real DNA will have an irregular, diffusive character. Loop tips appear and disappear, due to changes in the global folding, and the motion of a given loop tip will be diffusive as long as the loop persists. Such behavior is seen in MD simulations using our model if velocities are only occasionally rescaled, or if velocities are

reassigned.^{8,9} Repeated velocity rescaling changes this behavior, however (Fig. 6). After many rescalings, there is sufficient kinetic energy associated with the slithering degree of freedom that the motion becomes persistent and monotonic. The basepairs race around the interwound configuration like a long toy train on an interwound track.

We have observed another interesting artificial behavior in interwound DNA models with 600 bp, when the bending elastic modulus of a 120-bp segment is substantially reduced relative to that of normal DNA. This model mimics a closed circular DNA with a 120-bp insert of a triplet repeat sequence, because such sequences have a smaller persistence length (lower bending modulus) than normal DNA.¹⁰ Two examples are shown in Figure 7. In each of them, the discontinuity in elastic properties at the boundary between the insert and the rest of the DNA creates an obstacle

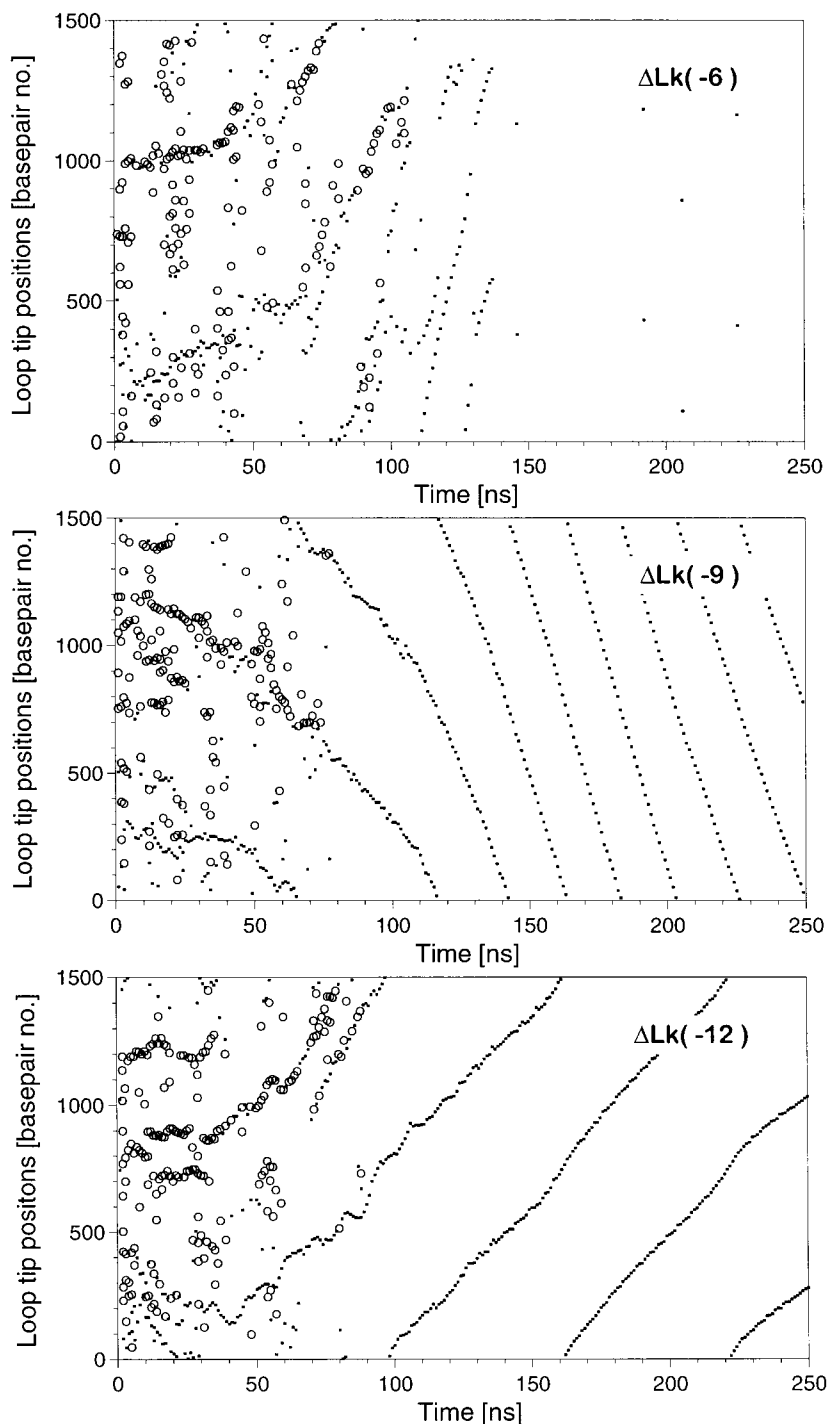


FIGURE 6. Slithering motion in the model supercoiled DNAs of Figure 5. The initially random, diffusive slither eventually gives way to persistent, monotonic slithering motion along the interwound configurations for $\Delta Lk = -9$ and $\Delta Lk = -12$. There are no defined loop tips for $\Delta Lk = -4$ after the molecule transforms into an open circle.

to the slithering motion. Slithering proceeds monotonically, as in Figure 6, until this barrier is hit, at which point the direction of slither reverses. The reversal is classified as a low-frequency motion,

because it has a characteristic time on the order of 10 ns, which is much larger than the characteristic time used to couple the system to the temperature bath (1 ps).

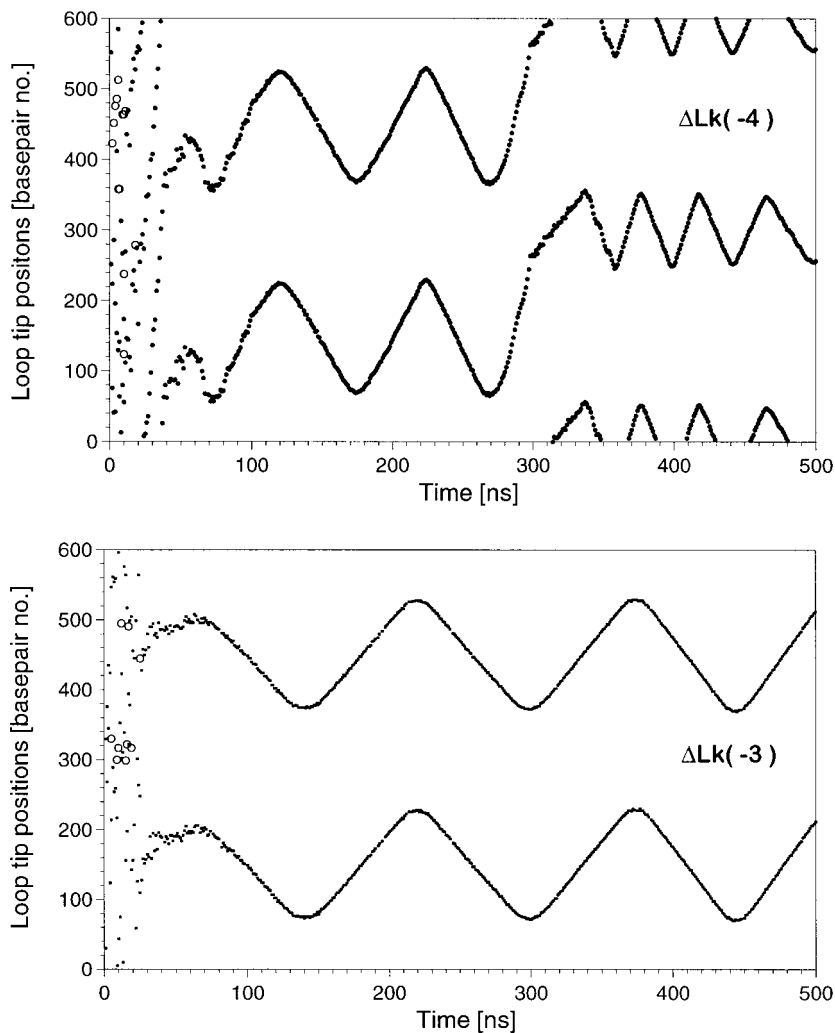


FIGURE 7. Slithering motions in interwound closed circular DNAs with 600 bp, containing a more flexible 120-bp insert. The persistence length of the 480 bp of normal DNA is 500 Å, whereas that of the insert is 300 Å. After an initial period characterized by slithering motions with a random, diffusive character, energy and momentum build up in this degree of freedom. The molecule slithers monotonically and with essentially constant velocity until the boundary between the two domains of differing elasticity reaches the loop, at which point the slithering usually reverses direction.

ALL-ATOM MOLECULAR DYNAMICS SIMULATIONS: CONDENSED PHASE SIMULATIONS WITH PERIODIC BOUNDARY CONDITIONS

Despite an increase in the number of interacting particles and therefore a greater number of means to couple energy between internal degrees of freedom, all-atom molecular dynamics simulations of condensed phase systems still offer examples of the flying ice cube syndrome. As discussed in the previous sections, if velocity scaling occurs at an interval longer than the period of any given vibration in the system, energy from these high-frequency modes may be transferred into lower

frequency modes. This is probably not a major problem in the simulation of condensed phase systems as long as the energy of these modes can couple back into the system; for example, through collisions among the molecules and if there is a large enough number of low-frequency modes that the drain of energy into each mode is very small.

Unfortunately, many of the methods employed in the simulation of condensed phase systems do not conserve energy in common usage unless particular care is taken. Specific examples include infrequent pairlist updates in simulations where the electrostatic interactions are subject to a cutoff, cooling due to integration timesteps that are too

large, SHAKE tolerances that are not low enough, and Berendsen pressure coupling,³ among other possible sources. Cutoffs applied to the van der Waals interactions and infrequent pairlist updates can also lead to heating or cooling.

Small energy drains in the system can lead to disastrous effects on the expected dynamics when any method that uniformly scales velocities is applied to control temperature. Gradually over the course of the simulation, the drain into low-frequency modes continues until essentially all of the motion is stored in these modes. This problem is particularly acute if the center of mass translational motion is not periodically removed. Because the center of mass translational energy cannot couple back into the system, the center of mass translational energy will grow with repeated rescaling. This leads to the flying ice cube and the trapping of the system in a local energy minimum.

We have previously reported all-atom MD simulations on the DNA duplex d[CCAACGTTGG]₂ run with 18 sodium ions in water,^{11,12} and we show here that energy equipartition can be violated with frequent velocity rescaling. The simulations were run using the particle mesh Ewald (PME) method within AMBER 4.1, with repeating boundary conditions. The simulation was run with cubic B-spline interpolation, a grid spacing of ~ 1 Å, an Ewald coefficient of ~ 0.31 , SHAKE on hydrogens (tolerance 0.0005), a pairlist update every 10 steps, and a cutoff of 9 Å. Constant temperature and pressure were maintained by coupling to a bath (coupling time 0.2 ps, 1 atm, 300 K), following the Berendsen method.³ The simulation represents approximately 9300 atoms in a periodic cell of $\sim 55 \times 41 \times 41$ Å³.

Figure 8 shows the atomic fluctuations over 100 ps windows for all the DNA atoms in the simulation. The fluctuations over the first 100 ps (dotted line) and the 100-ps interval after the first nanosecond (not shown) are comparable. However, after 2.3 ns, the fluctuations on most of the atoms drop by nearly half (solid line), and much of the kinetic energy is associated with translation of the center of mass. Note that not all atoms have their motion damped: four peaks are evident for atoms whose motion is relatively unaffected. These represent the hydrogens on the thymine methyl groups. Methyl rotation is opposed by only relatively weak forces, so the frequency of rotational oscillations of methyl groups is low. As the energy gets transferred into low-frequency modes, the methyl groups spin faster and faster.

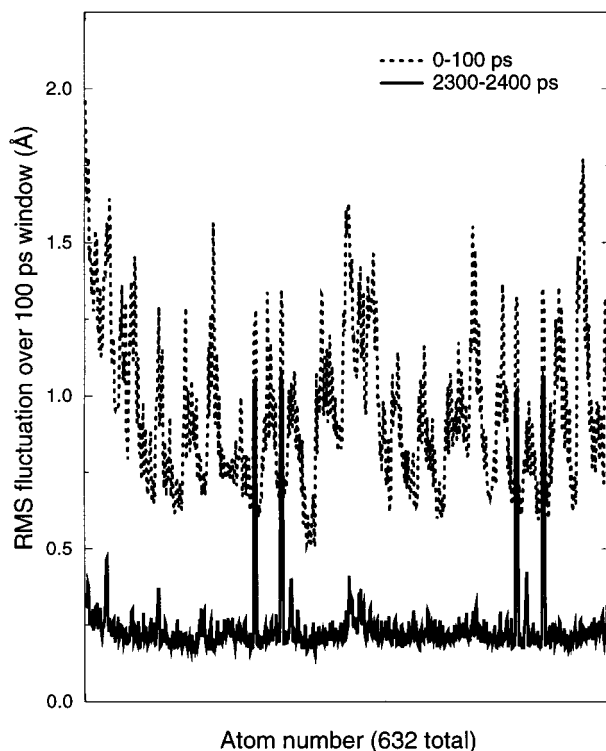


FIGURE 8. Root-mean-square amplitude of fluctuations in DNA atomic positions, as a function of atom number, for an all-atom MD simulation with periodic velocity rescaling. The DNA molecule has 632 atoms. Heating and equilibration are completed prior to time $t = 0$. The fluctuations over the first 100 ps of the simulation after equilibration are represented by the dashed line, whereas the solid line indicates the fluctuations over a 100-ps interval that begins at $t = 2.3$ ns. Note the loss of kinetic energy in these degrees of freedom.

This system shows the flying ice cube phenomenon in a traditional MD simulation with all-atom representation. It provides examples of energy drains into a zero-frequency motion (center of mass translation) and into a low-frequency mode (rotation of the methyl groups). It should be noted that the effect is not due to some peculiarity in the interaction between the macromolecule and the solvent, because simulations on 125 TIP3P water molecules lead to a complete drain into the center of mass translational kinetic energy in ~ 1.5 ns (T. E. Cheatham and P. A. Kollman, unpublished).

Discussion

The MD algorithm numerically integrates Newton's equations of motion. It is widely agreed that, to avoid instabilities, dissipation, and systematic

drifts, the Hamiltonian must be conservative, and the integration method should be reversible and symplectic.^{13–16} The most common integrators, such as the leapfrog¹⁷ and velocity Verlet¹⁸ algorithms, meet these conditions, so that these problems are avoided, at least in principle, in the limit of very small timesteps.

Nevertheless, there are a number of problems that emerge in MD simulations. Despite small timesteps, the gradual accumulation of unavoidable numerical errors can lead to drifts in the total energy over very long simulations. More insidious effects come from drifts in the total energy due to systematic errors in the simulation. These generally relate to methodological inaccuracies, such as improper truncation of the long-range interactions,¹⁹ and to tradeoffs that arise in methods designed to speed up the calculations, such as the use of low-order multipole expansions for long-range electrostatics.²⁰ In addition, the experiments with which simulations are compared are normally conducted at constant temperature and pressure, which corresponds to the isothermic–isobaric ensemble, not the constant energy ensemble of a free MD simulation.

These problems are usually attacked in MD simulations by making occasional adjustments to the atomic velocities. Velocity rescaling is one approach, whether by instantaneously adjusting the temperature to the target temperature,^{1,2} or by coupling to a temperature bath,³ which represents a gradual relaxation back toward the target temperature. In many laboratories, these procedures are done routinely, following historical protocols, without concern about rigor or possible side effects. With regard to rigor, rescaling is often done whenever the instantaneous temperature departs from the target value by some margin; that margin is often set without calculating the largest fluctuation that would be statistically expected to occur for the system being modeled. With regard to side effects, the analysis and examples presented here demonstrate that velocity rescaling can introduce artifacts in the structural, dynamic, and energetic properties of the system.

The phenomenon that we have described is distantly related to the temperature echoes first described by Grest et al.²¹ and investigated by a number of others.^{22–25} In an equilibrated MD simulation, if the atomic velocities are instantaneously set to zero (quenched) at time t_0 , the phases of the vibrational modes are all identical as the system resumes its motion after the quench. If the dynam-

ics are allowed to evolve freely for a time t_1 , and then a second quench (a quench echo) is applied, all vibrational modes lose kinetic energy except those whose phase is zero or π at the time of the second quench. The unaffected modes are those whose frequencies ω_i satisfy the relationship $\omega_i = n\pi/t_1$, where n is an integer. The first quench creates coherent vibrations, whereas the second removes kinetic energy from all modes except those whose periods are submultiples of t_1 . As a consequence, at times $t = t_0 + 2t_1$, $t = t_0 + 3t_1$, $t = t_0 + 4t_1, \dots$ there will be a dip in the kinetic energy (a temperature echo). The approach has been generalized to the inclusion of both cooling and heating pulses,²³ and random but correlated sets of velocities can be used for the two pulses to generate the phase correlations without perturbing the system temperature.²⁵ Temperature echoes can be used to examine normal modes, the density of states, anharmonic effects, and other dynamic properties in Lennard–Jones glasses^{21,22} and in macromolecular systems.^{23,24}

The temperature echo phenomenon manifests itself after only two velocity reassignments. Traditionally, zero velocities are reassigned,^{21,24} but random velocities may be used as long as the two sets of velocities are correlated.²⁵ In contrast, the phenomenon reported here occurs after repeated velocity rescalings, and there is a permanent shift in the distribution of kinetic energy among different modes, rather than any kind of echo. Our analysis and examples show that this redistribution occurs whether the rescaling interval is quite long or as small as one timestep (i.e., when the Berendsen temperature bath protocol is used).

Conclusions and Protective Measures

The gradual drain of kinetic energy from high-frequency motions into zero- and low-frequency motions is an unavoidable artifact of repeated velocity rescaling. It has been observed in all-atom simulations and in simulations using reduced representations (pseudoatoms), and the origin of the problem is now understood. As seen in the simulations described, this can greatly affect the structural, kinetic, and thermodynamic properties of the system.

There are several possible approaches to prevent the artifacts caused by periodic velocity rescaling. Here we discuss three.

VELOCITY REASSIGNMENT

Although velocity rescaling is very popular, constant temperature can be maintained equally well by other methods. Andersen²⁶ discussed protocols in which atomic velocities are periodically reassigned from a Maxwellian distribution with the appropriate temperature. This procedure, offered in many MD packages, is equivalent to subjecting each atom to a stochastic collision. (A discussion of the subtle differences in the thermodynamic ensembles defined by the velocity rescaling and reassignment algorithms is given in Appendix 3 of the monograph by McCammon and Harvey.¹)

Maintaining temperature by reassigning velocities, rather than rescaling, is one way to eliminate the problems described in this study. Simulations on supercoiled DNA with the 3DNA model do not suffer from buildup of energy in low-frequency modes if velocity reassignment is used to maintain temperature (D. Sprous, R. K.-Z. Tan and S. C. Harvey, unpublished). As another example, simulations of 125 TIP3P waters with velocity reassignments showed no appreciable artifact during 10 ns of simulation (T. E. Cheatham and P. A. Kollman, unpublished). We also note that frequent velocity reassignment during the heating and equilibration phases of MD simulations also helps to distribute energy throughout the structure, avoiding potential hot spots due to nonuniform distribution of structural distortions in the initial structure.

REMOVAL OF MOTION OF CENTER OF MASS

Because velocity rescaling converts much of the kinetic energy of the system into motion of the center of mass, artifactual behavior can be reduced by periodically removing the translational motion of the center of mass, and rotation about it. Unfortunately, this still leaves the simulation subject to violation of energy equipartition, because low-frequency modes can still be excited by repeated velocity reassignments, as shown in both the all-atom and reduced representation simulations on DNA, described earlier. Furthermore, if only translational motion of the center of mass is removed, the excitation of the zero-frequency rotational mode around the center of mass can cause substantial distortion of the structure, due to the growing centrifugal force. This is demonstrated by the distortion of the mean bondlength in the diatomic molecule simulation (Fig. 1), by the asymmetric character of the bondlength vibration in the

same simulation (Figs. 2–4), and by the regularization of the structure of the supercoiled DNA (Fig. 5).

We regard removal of translational and rotational motions of the entire system as a desirable, but not sufficient, precaution.

MODIFICATIONS TO ALGORITHMS THAT USE VELOCITY RESCALING

Because velocity rescaling has a long history, and those who wish to continue to use it should consider modifying it. One possibility would be to measure the mean temperature over some defined interval, then use that value to calculate a rescaling factor. Current protocols measure the instantaneous kinetic energy and calculate the corresponding temperature. By averaging over a suitable interval, much smaller scaling factors would result, and one would uncouple the direction of scaling from the instantaneous kinetic energy, which is the source of the problem.

Within current algorithms, the problem can be reduced by much less frequent rescaling, or, equivalently, by the use of very long coupling times in the Berendsen temperature bath protocol.³ Rescaling should probably be supplemented by occasionally reassigning velocities. If motion of the center of mass and global rotational motion are also periodically removed, this combination should substantially ameliorate the problems identified here.

Acknowledgments

Matt Young and David Beveridge have described similar problems with velocity rescaling to us (personal communication); we thank them, Dennis Sprous, Tom Darden, and Bernard Brooks for stimulating discussions. We thank Tamar Schlick for reading a preliminary version of the manuscript and offering several useful suggestions, in particular for urging us to include discussions on symplectic integrators and quench echo dynamics.

References

1. J. A. McCammon and S. C. Harvey, *Dynamics of Proteins and Nucleic Acids*, Cambridge University Press, Cambridge, U.K., 1987.
2. C. L. Brooks, M. Karplus, and B. M. Pettit, *Proteins: A*

- Theoretical Perspective of Dynamics, Structure and Thermodynamics*, Wiley-Interscience, New York, 1988.
3. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, *J. Chem. Phys.*, **81**, 3684 (1984).
 4. P. K. Weiner and P. A. Kollman, *J. Comput. Chem.*, **2**, 287 (1981).
 5. R. K.-Z. Tan and S. C. Harvey, *J. Comput. Chem.*, **14**, 455 (1993).
 6. R. K.-Z. Tan and S. C. Harvey, *J. Mol. Biol.*, **205**, 573 (1989).
 7. D. Sprous, R. K.-Z. Tan, and S. C. Harvey, *Biopolymers*, **39**, 243 (1996).
 8. R. K.-Z. Tan, D. Sprous, and S. C. Harvey, *Biopolymers*, **39**, 259 (1996).
 9. D. Sprous and S. C. Harvey, *Biophys. J.*, **70**, 1893 (1996).
 10. A. Bacolla, R. Gellibolian, M. Shimizu, S. Amirhaeri, S. Kang, K. Ohshima, J. E. Larson, S. C. Harvey, B. D. Stollar, and R. D. Wells, *J. Biol. Chem.*, **272**, 16783 (1997).
 11. T. E. Cheatham III and P. A. Kollman, *J. Mol. Biol.*, **259**, 434 (1996).
 12. T. E. Cheatham III and P. A. Kollman, *J. Am. Chem. Soc.*, **119**, 4805 (1997).
 13. M. Tuckerman, B. J. Berne, and G. J. Martyna, *J. Chem. Phys.*, **97**, 1990 (1992).
 14. J. J. Biesiadecki and R. D. Skeel, *J. Comp. Phys.*, **109**, 318 (1993).
 15. J. M. Sanz-Serna and M. P. Calvo, *Numerical Hamiltonian Problems*, Chapman & Hall, London, 1994.
 16. T. Schlick, E. Barth, and M. Mandziuk, *Annu. Rev. Biophys. Biomol. Struct.*, **26**, 181 (1997).
 17. L. Verlet, *Phys. Rev.*, **159**, 98 (1967).
 18. W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, *J. Comp. Phys.*, **76**, 637 (1982).
 19. P. J. Steinbach and B. R. Brooks, *J. Comput. Chem.*, **15**, 667 (1994).
 20. T. C. Bishop, R. D. Skeel, and K. Schulten, *J. Comput. Chem.*, **18**, 1785 (1997).
 21. G. S. Grest, S. R. Nagel, and A. Rahman, *Solid State Commun.*, **36**, 875 (1980).
 22. S. R. Nagel, A. Rahman, and G. S. Grest, *Phys. Rev. Lett.*, **47**, 1665 (1981).
 23. O. M. Becker and M. Karplus, *Phys. Rev. Lett.*, **70**, 3514 (1983).
 24. D. Xu, K. Schulten, O. M. Becker, and M. Karplus, *J. Chem. Phys.*, **103**, 3112 (1995).
 25. D. Xu and K. Schulten, *J. Chem. Phys.*, **103**, 3124 (1995).
 26. H. C. Andersen, *J. Chem. Phys.*, **72**, 2384 (1980).