

TIME-DEPENDENT DISTANCE RESTRAINTS IN MOLECULAR DYNAMICS SIMULATIONS

Andrew E. TORDA, Ruud M. SCHEEK and Wilfred F. VAN GUNSTEREN

Laboratory of Physical Chemistry, University of Groningen, Nijenborgh 16, 9747 AG Groningen, The Netherlands

Received 27 February 1989

A method for enforcing nuclear Overhauser effect (NOE) distance restraints in molecular dynamics simulations is presented. Rather than model the NOE distance as static, a term is included in the force field such that the distance restraint need only be satisfied as a $\langle r^{-3} \rangle^{-1/3}$ weighted time average over the simulation trajectory. This provides a better approximation of the physical nature of the NOE and reduces the disturbance to the force field due to the artificial term. Tests on a simple model system demonstrate the inadequacy of current methods and show the advantages of this novel approach, resulting in a more extensive search of conformational space.

1. Introduction

Given a set of nuclear Overhauser effect (NOE) distance restraints obtained by NMR spectroscopy, several methods are available for generating molecular structures consistent with the data. These include distance geometry algorithms based on either the metric matrix [1,2] or variable target function methods [3]. Alternatively, energetic considerations may be included by using dynamical simulated annealing [4] or distance-restrained molecular dynamics (MD) [5]. Although there are computational and conceptual differences between the techniques, all of the current methods attempt to minimise the difference between distances in the structure and simple distance bounds that represent the NOEs. In this paper we consider the inadequacies of this model and propose an alternative method which accounts for the dynamic averaging implicit in distances based on NOE measurements [6].

The problem with current methods can be seen by considering a simple example. A molecule may exist in solution in equilibrium between two different conformations, each characterised by a distinct set of NOEs. If however, interconversion between the conformers is rapid on the NMR scale, observed NMR resonances will reflect an averaging of the conformations. Similarly, NOEs from both conformations will be observed simultaneously. Using any of

the techniques mentioned above, one would attempt to generate structures which were consistent with both sets of experimental data. Unfortunately, there may be no single conformation which can satisfy the data, and even if one is generated, it may be highly strained and physically unrealistic. Generalising from this example, it may happen that there are many, rather than two solution conformations. In this case, current procedures may generate conformations even further from the true ensemble responsible for the NOEs.

This situation has recently been observed for several small peptides where apparently contradictory NOEs were measured and the data could only be explained by the presence of several conformations in solution [7,8].

Given these considerations, it is not correct to treat the NOE as a fixed distance bound. Instead, NOE distance information should be used to enforce an average distance bound through time. This can be achieved by imparting particles with a memory of their history with respect to internuclear distances. At the same time, to truly model the physical nature of the NOE, it is necessary to account for the non-linear dependence of the measured NOE intensity on the internuclear distance. In section 2, we describe such a method and its implementation in a MD simulation. Its effect is then tested on a simplified two-dimensional system.

2. Theory

Current methods for generating structures consistent with NMR data include some sort of penalty function, usually quadratic, with respect to violations of measured NOEs. In molecular dynamics simulations, this is done by creating an artificial term in the force field so that the energy of the system increases as violations increase. In the GROMOS force field [9], this is defined as

$$V_{dc}(r) = \frac{1}{2}K_{dc}(r-r_0)^2, \quad \text{if } r > r_0, \\ = 0, \quad \text{if } r \leq r_0, \quad (1)$$

where $V_{dc}(r)$ is the potential due to the distance-restraint term for a given pair of atoms, r is the instantaneous distance between the cross-relaxing nuclei and r_0 is the distance calculated from the measured NOE. The force constant, $\frac{1}{2}K_{dc}$, controls the relative strength of this artificial term in the force field. It should be noted that the full form of the term in the GROMOS force field can be made linear with respect to violations above a threshold and can also be used for repulsive interactions.

We propose a modification to the form of eq. (1) basing it on the time-averaged distance between atoms, denoted by \bar{r}

$$V_{dc}(\bar{r}) = \frac{1}{2}K_{dc}(\bar{r}-r_0)^2, \quad \text{if } \bar{r} > r_0, \\ = 0, \quad \text{if } \bar{r} \leq r_0, \quad (2)$$

It now remains to define the quantity \bar{r} so it can be calculated during a MD simulation.

Since the NOE arises from dipolar interactions between nuclei, the intensity of an NOE is a function which goes as r^{-6} . However, instead of writing the average distance as

$$\bar{r} = \langle r^{-6} \rangle^{-1/6}, \quad (3)$$

where $\langle r^{-6} \rangle$ is an average over time, it has been shown [7,10] that third-power averaging is necessary when the averaging is over a time less than the correlation time for overall molecular tumbling. Neglecting the influence of angular fluctuations on the NOE intensity, eq. (3) must then be written as

$$\bar{r} = \langle r^{-3} \rangle^{-1/3}, \quad (4)$$

where the angular brackets denote an average over time. This can be written in a form more suitable for

incorporation into a MD force field as

$$\bar{r}(t) = \left(\frac{1}{t} \int_0^t r(t')^{-3} dt' \right)^{-1/3}, \quad (5)$$

where t is time.

Eq. (5) is used for the analysis of trajectories in section 4, but it is not suitable for deriving an average over the finite time of a simulation. As time increases, the rate of change of $\bar{r}(t)$ and $V_{dc}(\bar{r}(t))$ will decrease, so the system will become less responsive over the course of a simulation. This problem is avoided by building a decay into the memory of the particle, with a characteristic decay time τ :

$$\bar{r}(t) = \left(\frac{1}{\tau} \int_0^t \exp(-t'/\tau) [r(t-t')]^{-3} dt' \right)^{-1/3}. \quad (6)$$

This has previously been done for electrostatic terms in MD simulations [11,12] so, by direct analogy with ref. [12], one can write

$$\bar{r}(t)^{-3} = \bar{r}(t-\Delta t)^{-3} \exp(-\Delta t/\tau) \\ + r(t)^{-3} [1 - \exp(-\Delta t/\tau)] \quad (7)$$

where Δt is the time step of the integrator in the simulation. In all the simulations described below, eq. (7) was used to calculate $\bar{r}(t)$ for substitution into eq. (2).

3. Model system and methods

For testing the field term described by eqs. (2) and (7), a model system was constructed so as to be as simple as possible while still capable of exhibiting two mutually exclusive conformations. To this end, three Lennard-Jones particles were used, each with a mass of 10 Da and interaction parameters taken from those of a united-atom CH_2 group. The system was made two-dimensional by taking initial coordinates and velocities equal to zero in the third dimension. Two of the particles were fixed, by a strong harmonic restraining force ($K=10000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$), to reference positions 2 nm from each other. The third particle could move freely through the two-dimensional space. Distance bounds of 0.8 nm were then enforced from each of the fixed particles to the

free particle with $K_{dc}=1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$. This system is shown graphically in fig. 1.

All simulations were carried out using software from the GROMOS suite of programs [9]. Velocities were controlled by tight coupling to a temperature bath [13] at 150 K with a coupling constant of 0.005 ps. The time step of the integrator was 0.002 ps and all simulations were carried out for 80 ps. The initial velocity of each fixed particle was set to zero, while that of the free particle is given in section 4.

It should be noted that the chosen parameters were quite arbitrary, but were dependent on each other. The choice of a force constant, K_{dc} , determined the shape of the potential well created by eq. (2) and thus the natural frequency of the system. For a given particle mass, the fluctuations in space were controlled by the temperature.

4 Results

Although the model system is extremely simple, one can see the direct analogy with a real system which can exist in either of two conformations. In the first conformation, the free particle must be within 0.8 nm of the first fixed particle. The second conformation has the free particle within 0.8 nm of the second fixed particle. Because the distance between the two fixed particles is 2.0 nm, the system cannot satisfy both restraints simultaneously.

A series of simulations was performed to assess the influence of the memory decay parameter, τ . The results were then analysed by calculating the average distance between particles according to eq. (5). This method was chosen since it is the behaviour of the system over the whole trajectory which must satisfy the experimental data. These simulations were started by assigning an arbitrary initial velocity of 0.48 ps^{-1} to the free particle only.

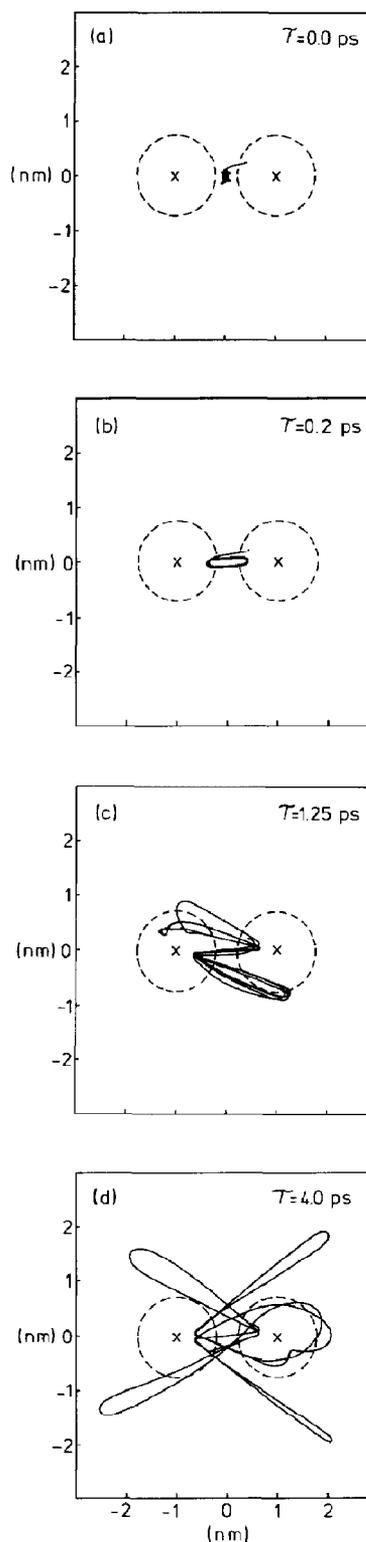


Fig. 1. Free particle trajectories for different values of τ as shown. Crosses mark the positions of the fixed particles and dashed circles show the 0.8 nm distance bounds. In each case, the free particle's starting coordinates were (0.42, 0.24) nm and the simulation ran for 80 ps. The radius of each particle (lowest Lennard-Jones energy) was 0.22 nm. The average violation of the two distance restraints according to eq. (5) was (a) 0.18, (b) 0.07, (c) 0.0, (d) 0.08 nm.

The first simulation was run with $\tau=0$ ps, where eq. (6) reduces to $\bar{r}(t)=r(t)$. This is equivalent to enforcing the NOEs as static distances as in current methods. The average violation of each distance bound at the end of the run was 0.18 nm. This is explained by the trajectory of the free particle shown in fig. 1a. With the static distance bounds, the lowest energy conformation is one where the free particle is trapped exactly in between the two fixed particles. This represents a strained high energy conformation which satisfies neither NOE. It is worth noting that if the system reached the lowest energy conformation, the violation would have been worse at 0.2 nm. The small oscillatory motions of the free particle mean that it is periodically closer to each fixed particle and the $\langle r^{-3} \rangle^{-1/3}$ average is slightly reduced.

The effect of a non-zero time constant is shown in fig. 1b where $\tau=0.2$ ps. The system shows the desired behaviour, attempting to move between the two conformations, but the short time constant does not allow sufficient freedom. At the end of the 80 ps run, each NOE was still, on average, violated by 0.07 nm. Fig. 1c however, shows a run with $\tau=1.25$ ps. At the end of the run, both distance restraints are now satisfied. One is within 0.10 and the other within 0.03 nm. The figure also shows the much greater mobility achieved by the system when τ is long enough.

Finally, fig. 1d shows the effect of a time constant ($\tau=4$ ps) which is too long given the lack of restraints on the system. After 80 ps, one NOE is violated by 0.14 nm and the other by 0.03 nm. It is interesting that this trajectory appears chaotic, but it resulted in smaller distance restraint violations than the trajectory where the system had no memory (fig. 1a).

Further simulations were performed using a range of values for τ from 0.0 to 6.0 ps. The final violations are plotted as a function of τ in fig. 2. The shape of the graph can be explained in terms of the resonance frequency of the system and the response time lag introduced by the memory function. When τ is zero, the free particle simply oscillates within the potential well centered between the fixed particles. At no time can it reach conformation 1 (near the first fixed particle) or conformation 2 (near the second fixed particle). As soon as the memory function is introduced, the walls of the potential themselves begin to oscillate. So, for example, when the free particle moves towards fixed particle 1, the corresponding potential well temporarily shifts away and the system can move to conformation 2. The length of time during which the potential is relaxed depends on τ . If τ is too short, the potential well narrows too fast, preventing transitions between the two conforma-

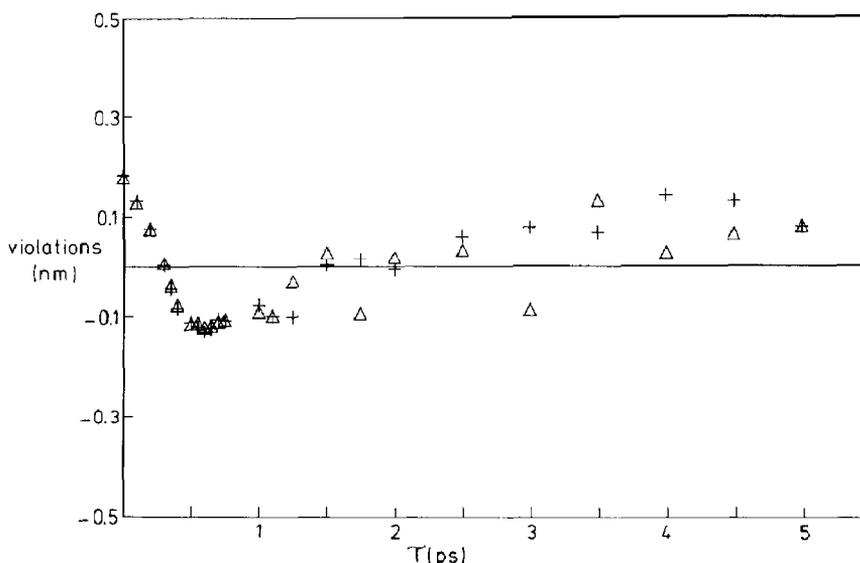


Fig. 2. The effect of τ average distance-restraint violations after 80 ps simulations. Crosses are used for one NOE and triangles for the second NOE.

tions. This accounts for the graph when τ is less than about 0.3 ps. When τ is between 0.3 ps and 1.4 ps, the potential well due to particle 1 allows the system to move completely into conformation 2 before drawing it back and vice versa. In this regime, both distance restraints are satisfied.

For still longer values of τ , the distance-restraint force acts too slowly. Not only can the system move from one allowed conformation to the other, but the potential is relaxed for such long periods of time that the system moves into regions of space not consistent with any of the distance information. In this regime, both distance restraints are violated.

Fig. 2 shows a second effect of the time-dependent force. At small values of τ , the final violation of both distance restraints is the same. For $\tau > 1.3$ ps, the final violations of the two distance restraints begin to diverge from each other. This is not a fault in the form of the potential, but rather represents difficulty

in convergence within the length of the simulation.

This is explained by fig. 3, taken from a simulation with $\tau = 1.25$ ps. The value of $\bar{r}(t)$ calculated from eq. (5) over the whole trajectory up to t , is plotted as a function of t . As τ is lengthened, the time for $\bar{r}(t)$ to converge increases. Furthermore, fig. 3 shows that fluctuations in $\bar{r}(t)$ for the two distance restraints are out of phase with each other. Thus, the apparent divergence in fig. 2 for large values of τ simply results from too short a simulation time.

A final check on the reliability of the results shown in fig. 2 was performed. In another series of simulations, the free particle was assigned random starting points within 1.0 nm of one or the other of the fixed particles. The particle's initial velocity over this series of simulations was also a random choice from a Maxwellian distribution at 150 K. For values of τ less than 1.3 ps, the final violation ($\bar{r}(t) - r_0$) lay on the curve of fig. 2. At longer values of τ , the values began to show the scatter also present in fig. 2. This test showed that the results are independent of the exact starting configuration within the region where the simulation is long enough to achieve convergence for $\bar{r}(t)$.

5. Discussion

Some considerations must be borne in mind when applying these results to a real molecule. The model system's force field contained only Lennard-Jones and distance-restraint terms so the effect of the NOE could be examined in isolation. A real system however, possesses other forces which would tend to damp its motions. So, in the model example, the free particle might move towards one of the fixed particles, representing a change into one of the desired conformations. The particle could then move away again, until the distance-restraint term brought it back. This unrestricted motion caused the final distance-restraint violations to rise as the time constant τ became too long. In a real system, multiple conformations must also represent energetic minima with respect to the rest of the force field. This means that once a molecule reached a desired conformation, there would be steric, covalent and other energetic terms hindering motion away from that conformation. Thus, in more complicated molecules,

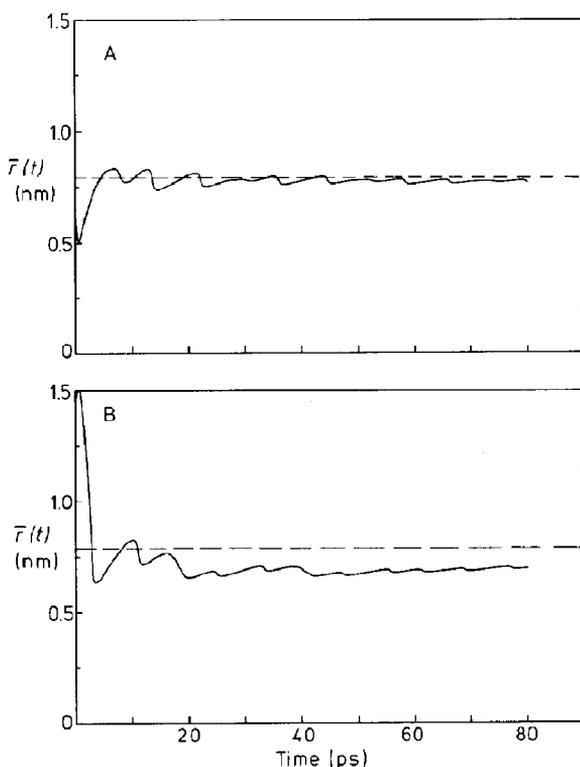


Fig. 3. Value of $\bar{r}(t)$ from eq. (5) calculated at each point in the trajectory. $\tau = 1.25$ ps. The dashed line shows the 0.8 nm distance restraint which was enforced. (a) Calculated for the first NOE. (b) Calculated for the second NOE.

there should be less danger with too large a value for τ . The most important consideration is that τ be long enough to allow for transitions to occur within the simulation, while the simulation itself is long enough to allow convergence of $\bar{r}(t)$.

Another difference between the model and a real molecule is shown by the trajectories in fig. 1, which show distinct periodicity in their motions. This regularity is a unique property of the model for two reasons. Firstly, the potential energy surface is simple and symmetric. The corresponding surface for a real molecule would be more complicated and less regular. Secondly, the motions of the free particle in this study are quite unrestricted. In a real molecule, such motions would be perturbed by constantly changing interactions with other particles in the system.

There is an interesting and subtle effect introduced by this potential since it depends on time as well as conformation. Unlike other terms in the force field, this distance-restraint term does not conserve energy. Even if a conformation is totally static, $\bar{r}(t)$ and consequently, $V_{dc}(\bar{r}(t))$, can change over time. It is then probably essential to use some mechanism for maintaining kinetic energy in the system. This can be done by coupling to a temperature bath with an overall coupling [13] or an individual atom coupling through the Langevin equation [14] or by simple regular velocity rescaling.

This preliminary study has shown that the time-dependent non-linear NOE is not well modelled by a static distance bound. Such fixed distance bounds may be justified during the generation of initial structures by distance geometry methods, but they do not seem appropriate during structural refinement by MD. Application of the method to small molecules, where multiple conformations are thought to exist, will be published elsewhere [15]. Finally, we anticipate increased mobility during simulations when the method is applied to macromolecules. This will result in a more extended search of configurational space for low energy conformations, but also has important consequences for the estimated accuracy of structures derived from NMR data.

Acknowledgement

These investigations were supported in part by the Netherlands' Foundation for Chemical Research (SON) with financial aid from the Netherlands' Technology Foundation. We also wish to thank ICI for their financial support and interest in the project. We are especially grateful to Professor I.D. Kuntz and Professor H.J.C. Berendsen for suggestions and stimulating discussions.

References

- [1] T.F. Havel, I.D. Kuntz and G.M. Crippen, *Bull. Math. Biol.* 45 (1983) 665.
- [2] T.F. Havel and K. Wüthrich, *Bull. Math. Biol.* 46 (1983) 673.
- [3] W. Braun and N. Gö, *J. Mol. Biol.* 186 (1985) 611.
- [4] M. Nilges, G.M. Clore and A.M. Gronenborn, *FEBS Letters* 239 (1988) 129.
- [5] W.F. van Gunsteren, R. Kaptein and E.R.P. Zuiderweg, in: *Proceedings of the NATO/CECAM workshop on nucleic acid conformation and dynamics*, ed. W.K. Olson (Orsay, 1984) p. 79.
- [6] O. Jardetzky, *Biochim. Biophys. Acta* 621 (1980) 227.
- [7] H. Kessler, C. Griesinger, J. Lutz, A. Müller, W.F. van Gunsteren and H.J.C. Berendsen, *J. Am. Chem. Soc.* 110 (1988) 3393.
- [8] H. Pepermans, D. Tourwé, G. van Binst, R. Boelens, R.M. Scheek, W.F. van Gunsteren and R. Kaptein, *Biopolymers* 27 (1988) 323.
- [9] W.F. van Gunsteren and H.J.C. Berendsen, *Groningen Molecular Simulation (GROMOS) library manual* (Biosmos, Groningen, 1988).
- [10] J. Tropp, *J. Chem. Phys.* 72 (1980) 6035.
- [11] W.F. van Gunsteren, H.J.C. Berendsen and J.A.C. Rullman, *Faraday Discussions Chem. Soc.* 66 (1978) 58.
- [12] E. Egbergs and H.J.C. Berendsen, *J. Chem. Phys.* 89 (1988) 3718.
- [13] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola and J.R. Haak, *J. Chem. Phys.* 81 (1984) 3684.
- [14] W.F. van Gunsteren and H.J.C. Berendsen, *Mol. Sim.* 1 (1988) 173.
- [15] A.E. Torda, R.M. Scheek and W.F. van Gunsteren, *J. Mol. Biol.* (1989), to be published.