The GROMOS Biomolecular Simulation Program Package

Walter R. P. Scott, Philippe H. Hünenberger, Ilario G. Tironi, Alan E. Mark, Salomon R. Billeter, Jens Fennen, Andrew E. Torda, Thomas Huber, Peter Krüger, and Wilfred F. van Gunsteren*

Laboratory of Physical Chemistry, ETH Zentrum, CH-8092 Zürich, Switzerland

Received: October 27, 1998; In Final Form: March 1, 1999

We present the newest version of the GROningen MOlecular Simulation program package, GROMOS96. GROMOS96 has been developed for the dynamic modelling of (bio)molecules using the methods of molecular dynamics, stochastic dynamics, and energy minimization as well as the path-integral formalism. An overview of its functionality is given, highlighting methodology not present in the last major release, GROMOS87. The organization of the code is outlined, and reliability, testing, and efficiency issues involved in the design of this large (73 000 lines of FORTRAN77 code) and complex package are discussed. Finally, we present two applications illustrating new functionality: local elevation simulation and molecular dynamics in four spatial dimensions.

1. Introduction

The GROMOS96[1] package has been developed to facilitate research efforts in the field of biomolecular simulation in a university environment. The package acts as a test bed for developing new simulation methods, as well as providing the tools required for routine modelling applications. The criteria for a feature to be included in GROMOS are the following: (1) scientific interest, (2) extent of the use by the scientific community, (3) demonstrated usefulness or efficiency, (4) well-defined and correct formulae and algorithms, (5) ease of implementation, (6) computational efficiency, and (7) interest of our research group.

The GROMOS package has been developed by W. F. van Gunsteren and co-workers since 1978 (at Harvard University (USA), University of Groningen (The Netherlands), and ETH Zürich (Switzerland)). Because it has been designed for ease of extendability and the complete source code is made available to research establishments for a nominal fee, it has found widespread use (hundreds of licences in over 40 countries on all continents). [For details, see the GROMOS homepage at http://igc.ethz.ch/gromos/ or contact Biomos b.v., Laboratory of Physical Chemistry, ETH Zürich, ETH Zentrum, CH-8092 Zürich, Switzerland; Fax +41-1-6321039, e-mail: biomos@igc.phys.chem.ethz.ch.]

GROMOS96 contains a complete rewrite of the molecular dynamics (MD), stochastic dynamics (SD), and energy minimization (EM) part of the last version of GROMOS, GROMOS87,[2] in order to incorporate new functionality. GROMOS96 was designed and implemented over a span of approximately 20 months (October 1994−June 1996).

In the next section, we discuss the organization of the source code and the block structure used in the data files. Section 3 lists the programs and types of file in the package, section 4 presents the time-stepping algorithms used in detail, in section 5 the force field is briefly discussed, and in section 6 reliability and speed issues are discussed. Section 7 contains examples of applications.

2. Code Organization

GROMOS96 is designed along modular lines, with separate files containing programs and subroutines, and header files defining common blocks and array size parameters. The introduction of header files facilitates modifying the array sizes in the code considerably: a parameter may be changed in a header file, and all programs affected by the change can be recompiled using the standard Unix make utility.

The GROMOS96 data file format departs from that used in previous GROMOS versions. The new file format has been designed to be extendable to facilitate the adding of new functionality. This has been achieved by the introduction of a block concept: each GROMOS96 file consists of a sequence of blocks. Each block is preceded by a header which uniquely identifies the contents of the block. Based on the block header, a program can decide whether to read or skip the data. It is possible, for example, for a programmer to modify the simulation program (PROMD) to write a new block containing additional data to the coordinate trajectory file. This will not affect any standard analysis programs that read that trajectory file. They will simply ignore the unknown block type. The block concept is implemented slightly differently for formatted and unformatted files. In the former case, a block header consists simply of the block name of maximum length of 16 characters, e.g. "POSITION". The data follows on subsequent lines in free format. Any line starting with a "#" character is ignored. The end of the block is marked by a line with the word "END". In the unformatted case, a header consists of the same blockname *and an integer* which denotes the number of FORTRAN records in the block. By convention, GROMOS96 expects the first block in a file to contain a title (TITLE block). GROMOS96 automatically detects whether a file is formatted or unformatted on opening it.

GROMOS96 has been written in standard FORTRAN77 with the exception of the use of include files. The code has been compiled and run successfully on a wide range of machines

* Corresponding author. Phone: +41 1 632 5501. Fax: +41 1 632 1039. Email: wfvgn@igc.phys.chem.ethz.ch.

10.1021/jp984217f CCC: $18.00
Published on Web 04/24/1999

ranging from PC's to supercomputers. There are two versions of the code targeted toward specific machines: a vectorized version for Cray computers and a shared memory parallel version for SGI Power Challenge computers.

3. Overview of Functionality

GROMOS96 consists of 40 programs grouped by functionality as follows:

1. Topology builders

- PROGMT builds a molecular topology by chaining molecular topology building blocks together
- PROMMT merges two topology files together
- PROCMT cuts or reduces a molecular topology file according to a distance criterion or an atom list
- PROPMT converts a classical molecular topology into a path-integral molecular topology

2. Coordinate and data resequencers and reformatters (to GROMOS96)

- PROCS1 converts format and order of atom coordinates from many formats, including Brookhaven Protein Data Bank (PDB) format,[3] to that used in PROCS2
- PROCS2 converts format and order of atom coordinates from many formats, including PDB, to that used in GROMOS96
- PROCDR converts atom−atom distance constraints from DISGEO,[4] DISMAN,[5] or DIANA[6] format into GROMOS96 form

3. Coordinate generators

- PROGCA generates atomic Cartesian coordinates from a set of internal coordinates
- PROSSC substitutes missing coordinates of protein side chains with standard coordinates from a library
- PROGCH generates or deletes hydrogen atom coordinates of a molecule
- PROGWH generates or deletes hydrogen atom coordinates of water molecules
- PROCRY generates crystal atom coordinates for molecules plus solvent by performing rotations and translations
- PROBOX puts a simulation box around a molecule and fills it with solvent molecules
- PROION replaces solvent molecules by an ion (used to neutralize a system's total charge)
- PROGCB expands a configuration of classical atoms to a configuration with the path-integral atoms represented by a number of beads.

4. Minimizers and simulators

- PROMD performs an energy minimization, a molecular dynamics simulation, or a stochastic dynamics simulation

5. Analyzers

- PROAVX calculates the mean and root mean square (rms) fluctuation of specified atomic coordinates from a sequence of configurations
- PROAVQ calculates for a quantity (bond, bond angle, dihedral angle) the mean, the rms fluctuation, third and fourth moment from a sequence of configurations
- PROAJC calculates nuclear magnetic resonance spin−spin *J*-coupling constants and averages from a sequence of configurations
- PROADR calculates atom−atom restraint distances and averages from a sequence of configurations
- PROAVN calculates and averages the number of neighbor atoms surrounding each solute atom for a sequence of configurations
- PROAHB calculates and averages hydrogen bonds for a sequence of configurations
- PROMHB monitors the occurrence of hydrogen bonds and hydrogen bond lifetimes for a sequence of configurations
- PROCOC calculates the occupancies for a set of sites for a sequence of configurations
- PROAVS calculates the means and rms fluctuations of the atomic coordinates from a sequence of solvent configurations
- PROCOX compares two configurations: absolute and rms differences of selected atom coordinates, radius of gyration, atom density
- PROCOQ compares two configurations: differences in bond lengths, bond angles, improper dihedrals, dihedral angles, and corresponding energies.
- PROCOB compares two sets of isotropic atomic rms positional fluctuations or crystallographic *B* factors
- PROCAB compares two sets of anisotropic atomic rms positional fluctuations or crystallographic *B* factors
- PROCOD compares two sets of configurations in terms of atom−atom distances
- PRONBL generates a neighbor list for a single configuration for a number of solute molecules in a variety of ways
- PROCHB generates a list of hydrogen bonds for a number of solute molecules plus solvent molecules
- PROCPS for a single configuration, determines for each solvent molecule the solute atom nearest to it
- PROCOS compares two solvent configurations plus rms positional fluctuations or crystallographic *B* factors
- PROTCF calculates distributions, mean, second, third and fourth moment, the auto or cross time correlation function and the corresponding spectral density of a number of quantities for a set of configurations
- PROFEE calculates free energy differences between two states of a system from a reference trajectory
- PROPIC contracts a trajectory of path-integral configurations containing beads representing path-integral quantum atoms into a configuration of classical atoms

6. Coordinate reformatters (from GROMOS)

- PROMCF merges a number of trajectory files into one trajectory file. The atom sequence may be changed and atoms removed
- PROPSF converts atomic Cartesian coordinates from a number of trajectory files into oblique contravariant fractional coordinates suitable for structure factor calculations
- PROPDF reads atomic Cartesian coordinates and writes these in non-GROMOS format.

A distinction is made between two types of data in GROMOS96: topological information and configurational information. In the former, lists of covalent bonds, angles, masses, charges, etc. for the molecules in the system are defined. This data is stored once for multiple identical molecules and remains constant during a simulation, with the exception of free energy calculations. In the latter, all coordinate-dependent or derived quantities, such as coordinates, velocities, atom−atom distances, dihedral angles, energies, etc. are defined.

Often a system to be simulated will consist of one or a small number of solute molecules surrounded by a large number of simple solvent molecules. The GROMOS96 topology therefore consists of two parts: a general part called (for historic reasons) the solute, although it can contain any collection of molecules including solvent molecules, and a restricted part called the solvent. Restrictions are placed on solvent molecules to improve overall performance. In a typical application, for instance the simulation of a solvated protein, the majority of the computational time is spent on computing interactions between the many solvent molecules. The restrictions allow for fast code to be

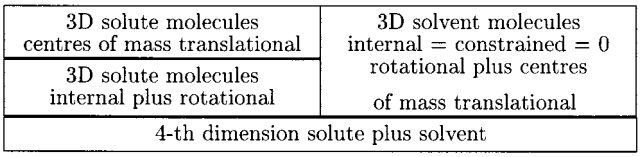

Figure 1. Degrees of freedom which may be coupled separately to different temperature baths in GROMOS96. All degrees of freedom in three dimensions may also be coupled together in all combinations.

written for the solvent-solvent nonbonded interactions. The restrictions for a molecule to be defined as solvent include the following:

- the molecule must be rigid: no internal interactions such as bondstretch, bond-angle bending, (improper) dihedral torsion, and intramolecular nonbonded interactions are allowed.
- the internal structure of a solvent molecule is maintained by application of distance constraint forces (procedure SHAKE[7]) between atoms.
- the molecule must consist of one charge group (see section 5.2). The coordinates of the first atom of a solvent molecule are taken to represent the position of the charge group for the purpose of generating a charge group pair list.
- position constraining cannot be applied to a solvent molecule.
- distance restraining cannot be applied to atoms of solvent molecules.
- solvent parameters cannot be changed using a molecular topology perturbation file for obtaining free energy differences.

If a solvent molecule does not satisfy these conditions, its topological data must be included in the solute part of the topology.

GROMOS96 allows only one type of solvent molecule in the solvent part of the topology at any given time. In a system with mixed solvent, one solvent type would be defined in the solute part of the topology.

4. Simulation Algorithms

In this section the finite time-stepping algorithms for molecular dynamics (MD) and stochastic dynamics (SD) used in GROMOS96 are presented. Both MD and SD simulations use a leapfrog integration scheme modified to include optional bond constraints with the SHAKE method,[7,8] coupling to external baths[9] and with stochastic forces.[10] The following additions make the algorithms more complex.

1. The center of mass translational degrees of freedom in three dimensions of the solute molecules may be coupled to a separate heat bath (Figure 1). This is necessary in path-integral simulations, as the motions of the beads are very weakly coupled to the rest of the system due to the difference in mass between the path integral beads and the classical atoms. The same distinction between degrees of freedom is not made for solvent, as only atoms of the solute part of the topology can be treated as path-integral beads. When performing a simulation in four dimensions, the temperature in the fourth dimension is also coupled separately. This is so that the reference temperature in the fourth dimension may be changed in conjunction with the force constant of the restraining potential energy function in the fourth dimension when forcing the system back into three dimensions (see subsection 7.2 for an example of this). For the stepping algorithm, this simply means that one integrates the force in four dimensions instead of three per atom, using a different scaling factor for the velocity in the fourth dimension.

2. Position *constraining* of atoms is possible in GROMOS96 in addition to position *restraining* available in previous GROMOS versions. In the former, a positionally constrained atom is kept rigidly at a place of reference. As a consequence, the system loses three degrees of freedom. Atoms in four dimensions cannot be positionally constrained. In the latter, a harmonic potential is applied to keep the restrained atom in the vicinity of the reference position. In this case, no degrees of freedom are removed from the system.

In the following description, the positions $\vec{r}(t) = \vec{r}_i(t)$, the velocities $\vec{v}_i(t - (\Delta t/2))$, and the forces $\vec{f}(t) = \vec{f}_i(t)$ are given for all atoms $i = 1...N$, where N is the total number of atoms in the system.

4.1. Molecular Dynamics Algorithm. Preconditions:

1. $E_{\mathrm{kin}}^{\mathrm{solute,3D}}(t - (\Delta t/2))$, $E_{\mathrm{kin}}^{\mathrm{solvent,3D}}(t - (\Delta t/2))$, $E_{\mathrm{kin}}^{\mathrm{4thD}}(t - (\Delta t/2))$ and the corresponding temperatures have been calculated. The initial step is $n = 0$.

2. The side lengths of the (periodic) computational box are given.

3. The positions $\vec{r}(t)$ satisfy the constraints and the velocities $\vec{v}(t - (\Delta t/2))$ do not contain components along the constraints.

4. Solvent molecules are not split by periodic boundaries. The same condition holds for the atoms of a solute charge group.

Algorithm:

1. Apply the periodic boundary conditions to put the solute charge groups and the solvent molecules into the central computational box.

2. Calculate the solute molecular center of mass translational kinetic energy

$$E_{\mathrm{kin,mol}}^{\mathrm{tr,solute}}\left(t - \frac{\Delta t}{2}\right) = \frac{1}{2}\sum_{\alpha=1}^{N_M^{\mathrm{solute}}} M_\alpha \vec{V}_\alpha^{\,2}\left(t - \frac{\Delta t}{2}\right) \tag{1}$$

and the solute internal and rotational kinetic energy

$$E_{\mathrm{kin,mol}}^{\mathrm{int,rot,solute}}\left(t - \frac{\Delta t}{2}\right) = \frac{1}{2}\sum_{i=1}^{N^{\mathrm{solute}}} m_i \vec{v}_i^{\,\mathrm{int,rot}^2}\left(t - \frac{\Delta t}{2}\right) = E_{\mathrm{kin}}^{\mathrm{solute,3D}}\left(t - \frac{\Delta t}{2}\right) - E_{\mathrm{kin,mol}}^{\mathrm{tr,solute}}\left(t - \frac{\Delta t}{2}\right) \tag{2}$$

where N^{solute} is the number of solute atoms, N_M^{solute} is the number of solute molecules, and

$$\vec{V}_\alpha\left(t - \frac{\Delta t}{2}\right) = \frac{1}{M_\alpha}\sum_{i=1}^{N_\alpha} m_{i\alpha}\vec{v}_{i\alpha}\left(t - \frac{\Delta t}{2}\right) \tag{3}$$

is the center of mass velocity,

$$M_\alpha = \sum_{i=1}^{N_\alpha} m_{i\alpha} \tag{4}$$

is the mass of the solute molecule $\alpha = 1, ..., N_M$ and

$$\vec{v}_{i\alpha}^{\,\mathrm{int,rot}}\left(t - \frac{\Delta t}{2}\right) = \vec{v}_{i\alpha}\left(t - \frac{\Delta t}{2}\right) - \vec{V}_\alpha\left(t - \frac{\Delta t}{2}\right) \tag{5}$$

is the atomic velocity relative to the molecular center of mass.

The double index (e.g. $m_{i\alpha}$) denotes the atom i of solute molecule α which consists of N_α atoms.

3. If the virial is to be calculated, calculate the molecular center of mass positions

$$\vec{R}_\alpha = \frac{1}{M_\alpha}\sum_{i=1}^{N_\alpha}\vec{r}_{i\alpha}m_{i\alpha} \quad (6)$$

the *x*-, *y*-, and *z*-components of the total molecular center of mass translational kinetic energy $E^{tr}_{kin,mol}(t - (\Delta t/2))$ and the relative positions

$$\vec{r}^{\,int}_{i\alpha}(t) = \vec{r}_{i\alpha}(t) - \vec{R}_\alpha(t) \quad (7)$$

for solute and solvent molecules.

4. Calculate the unconstrained forces from the potential energy function $V(\vec{r}(t))$

$$\vec{f}_i(t) = -\frac{\partial V(\vec{r}(t))}{\partial \vec{r}_i} \quad (8)$$

using the nearest image or minimum image convention in case of periodic boundary conditions and at the same time calculate the molecular virial

$$\Xi(t) = -\frac{1}{2}\sum_{i\alpha<j\beta}^{N}[\vec{r}^{\,NI}_{i\alpha j\beta} - \vec{r}^{\,int}_{i\alpha}(t) + \vec{r}^{\,int}_{j\beta}(t)]\cdot\vec{f}_{i\alpha j\beta}(t) \quad (9)$$

and its components in *x*-, *y*-, and *z*-directions if required. The summation in the above formula runs over all pairs of atoms that are selected by the cutoff criterion used, and for which atoms *i* and *j* are in different molecules, and the superscript NI indicates the nearest-neighbor image vector. Note that, as the *molecular* virial is calculated, there is no contribution to the virial from the covalent terms.

5. If the calculation of the pressure is required, calculate the volume of the periodic box $V_b(t)$ and the pressure

$$P(t) = \frac{2\left[E^{3D}_{kin}\left(t - \frac{\Delta t}{2}\right) - \Xi(t)\right]}{3V_b(t)} \quad (10)$$

where $E^{3D}_{kin}(t - (\Delta t/2)) = E^{solute,3D}_{kin}(t - (\Delta t/2)) + E^{solvent,3D}_{kin}(t - (\Delta t/2))$.

6. If required, write the positions $\vec{r}(t)$, the velocities $\vec{v}(t - (\Delta t/2))$, the box dimensions and the time and step number to trajectory files.

7. Determine the (unconstrained) velocities

$$\vec{v}_i\left(t + \frac{\Delta t}{2}\right) = \vec{v}_i\left(t - \frac{\Delta t}{2}\right) + \frac{1}{m_i}\vec{f}_i(t)\Delta t \quad (11)$$

8. Determine the translational centre of mass velocities of solute molecules $\vec{V}_\alpha(t + (\Delta t/2))$ using eq 3 and the atomic velocities with respect to the molecular centres of mass $\vec{v}^{\,int,rot}_{i\alpha}(t + (\Delta t/2))$ using equation 5.

9. Scale the velocities $\vec{v}_i(t + (\Delta t/2))$ of the atoms that are coupled to a temperature bath with the appropriate scaling factor. The scaling factors are based on the kinetic energies at time $(t - (\Delta t/2))$.

10. Determine the unconstrained positions

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \vec{v}_i\left(t + \frac{\Delta t}{2}\right)\Delta t \quad (12)$$

11. Make the positions $\vec{r}_i(t + \Delta t)$ satisfy the constraints by applying the SHAKE algorithm.

12. Calculate the constrained velocities

$$\vec{v}_i\left(t + \frac{\Delta t}{2}\right) = \frac{\vec{r}_i(t + \Delta t) - \vec{r}_i(t)}{\Delta t} \quad (13)$$

13. Calculate the kinetic energies $E^{solute,3D}_{kin}(t + (\Delta t/2))$, $E^{solvent,3D}_{kin}(t + (\Delta t/2))$, $E^{4thD}_{kin}(t + (\Delta t/2))$.

14. If required, write the energies, volume, pressure, scaling data, and free energy data to trajectory files.

15. If required, print the energies, volume, pressure, scaling data, and free energy data to standard output.

16. When using pressure coupling, scale the atomic positions $\vec{r}_i(t + \Delta t)$ and box lengths with the appropriate pressure scaling factor. If positionally restrained or constrained atoms are in the system, also scale their reference positions.

17. Increase the time $t_{n+1} = t_n + \Delta t$ and the step number by one. If required, change the free energy perturbation coupling parameters λ to $\lambda + \Delta\lambda$ and μ to $\mu = \mu + \Delta\mu$.

4.2. Stochastic Dynamics Algorithm. In stochastic dynamics (SD), an external force $\vec{f}^{\,ext}_i(t)$ due to degrees of freedom not explicitly treated in the simulation is added to the internal force $\vec{f}^{\,int}_i(t)$ exerted on atom *i* derived from the potential energy function.

In the method discussed here, $\vec{f}^{\,ext}_i(t)$ is composed of a mean external force $\vec{f}^{\,mean}_i(t)$, its fluctuations in time $\vec{f}^{\,stoch}_i(t)$, and its frictional effect $\vec{f}^{\,fric}_i(t) = m_i\gamma_i\vec{v}_i(t)$

$$\vec{f}^{\,ext}_i(t) = \vec{f}^{\,mean}_i(t) + \vec{f}^{\,stoch}_i(t) - \vec{f}^{\,fric}_i(t) \quad (14)$$

where γ_i is the atomic friction coefficient. This leads to the Langevin equation of motion

$$d\vec{v}_i(t)/dt = \frac{1}{m_i}[\vec{f}^{\,int}_i(t) + \vec{f}^{\,mean}_i(t) + \vec{f}^{\,stoch}_i(t)] - \gamma_i\vec{v}_i(t) \quad (15)$$

A major application of SD is to approximate the effect of explicit solvent molecules on the dynamics of macromolecules. The macromolecule is simulated in vacuo with SD, without periodic boundary conditions, resulting in a considerable gain in speed, but at the loss of detailed solvent effects. Because of the vacuum, calculating the pressure and coupling to a pressure bath is meaningless, and is not implemented for SD simulations. The atomic friction coefficients γ_i can be set to a specific value for each atom or in accordance with an approximation to the solvent-accessible surface area.[11] The SD algorithm is the following. Preconditions

1. The positions $\vec{r}(t) = \vec{r}_i(t)$, the velocities $\vec{v}(t - (\Delta t/2)) = \vec{v}_i(t - (\Delta t/2))$ and the stochastic integrals $\vec{R}_i(t - (\Delta t/2); \Delta t/2)$ are given. The definition of the stochastic integral is

$$\vec{R}_i\left(t; \frac{\Delta t}{2}\right) = \frac{1}{m_i\gamma_i}\int_t^{t+\Delta t/2}\left[1 - \exp\left(-\gamma_i\left(t + \frac{\Delta t}{2} - t'\right)\right)\right]\vec{f}^{\,stoch}_i(t')\,dt' \quad (16)$$

$E^{solute,3D}_{kin}(t - (\Delta t/2))$, $E^{solvent,3D}_{kin}(t - (\Delta t/2))$, $E^{4thD}_{kin}(t - (\Delta t/2))$ and the corresponding temperatures have been calculated. The initial step is $n = 0$.

2. The positions $\vec{r}(t)$ satisfy the constraints and the velocities $\vec{v}(t - (\Delta t/2))$ do not contain components along the constraints.

Algorithm

1. Calculate the unconstrained forces from the potential energy function $V(\vec{r}(t))$

$$\vec{f}_i(t) = -\frac{\partial V(\vec{r}(t))}{\partial \vec{r}_i} \quad (17)$$

2. If required, write the positions $\vec{r}(t)$, the velocities $\vec{v}(t - (\Delta t/2))$, the time and step number to trajectory files.

3. Sample the components of a vector $\vec{V}'_i$ from a Gaussian distribution with zero mean and width σ_2^2 and determine

$$\vec{V}_i\left(t;-\frac{\Delta t}{2}\right) = \sigma_3 \vec{R}_i\left(t-\frac{\Delta t}{2};\frac{\Delta t}{2}\right) + \vec{V}_i' \tag{18}$$

where

$$\sigma_2^2 = \frac{k_B T_{ref} B\left(\frac{\gamma_i \Delta t}{2}\right) C\left(\frac{\gamma_i \Delta t}{2}\right)}{m_i} \tag{19}$$

$$\sigma_3 = \frac{\gamma_i D(\gamma_i \Delta t/2)}{C(\gamma_i \Delta t/2)} \tag{20}$$

$$B\left(\frac{\gamma_i \Delta t}{2}\right) = \gamma_i \Delta t[\exp(\gamma_i \Delta t) - 1] - 4\left[\exp\left(\frac{\gamma_i \Delta t}{2}\right) - 1\right]^2 \tag{21}$$

$$C\left(\frac{\gamma_i \Delta t}{2}\right) = \gamma_i \Delta t - 3 + 4\exp\left(-\frac{\gamma_i \Delta t}{2}\right) - \exp(-\gamma_i \Delta t) \tag{22}$$

and

$$D\left(\frac{\gamma_i \Delta t}{2}\right) = 2 - \exp\left(\frac{\gamma_i \Delta t}{2}\right) - \exp\left(-\frac{\gamma_i \Delta t}{2}\right) \tag{23}$$

4. Sample the components of a vector $\vec{V}_i(t;\ \Delta t/2)$ from a Gaussian distribution with zero mean and width ρ_1^2, where

$$\rho_1^2 = \frac{k_B T_{ref}[1 - \exp(-\gamma_i \Delta t)]}{m_i} \tag{24}$$

The definition of this integral is

$$\vec{V}_i\left(t;\frac{\Delta t}{2}\right) = \frac{1}{m_i}\exp\left(-\frac{\gamma_i \Delta t}{2}\right) \int_t^{t+(\Delta t/2)} \exp(1 - \gamma_i(t - t'))\vec{f}_i^{\,stoch}(t')\,dt' \tag{25}$$

5. Determine the unconstrained velocities

$$\vec{v}_i\left(t+\frac{\Delta t}{2}\right) = \vec{v}_i\left(t-\frac{\Delta t}{2}\right)\exp(-\gamma_i \Delta t) + \frac{\vec{f}_i(t)\Delta t[1 - \exp(-\gamma_i \Delta t)]}{m_i \gamma_i \Delta t} - \exp(-\gamma_i \Delta t)\vec{V}_i\left(t;-\frac{\Delta t}{2}\right) + \vec{V}_i\left(t;\frac{\Delta t}{2}\right) \tag{26}$$

6. Scale the velocities $\vec{v}_i(t + (\Delta t/2))$ of the atoms that are coupled to a temperature bath with the appropriate scaling factor. The scaling factors are based on the kinetic energies at time $t - (\Delta t/2)$.

7. Calculate the new positions excluding the contributions of the stochastic integrals where

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \vec{v}_i\left(t+\frac{\Delta t}{2}\right)\Delta t E\left(\frac{\gamma_i \Delta t}{2}\right) \tag{27}$$

where

$$E\left(\frac{\gamma_i \Delta t}{2}\right) = \frac{\exp\left(\frac{\gamma_i \Delta t}{2}\right) - \exp\left(-\frac{\gamma_i \Delta t}{2}\right)}{\gamma_i \Delta t} \tag{28}$$

8. Make the velocities satisfy the constraints by first applying the SHAKE algorithm to the coordinates $\vec{r}_i(t + \Delta t)$ and then calculating the constrained velocities

$$\vec{v}_i\left(t+\frac{\Delta t}{2}\right) = \frac{\vec{r}_i(t + \Delta t) - \vec{r}_i(t)}{\Delta t E\left(\frac{\gamma_i \Delta t}{2}\right)} \tag{29}$$

9. Calculate the kinetic energies $E_{kin}^{solute,3D}(t + (\Delta t/2))$, $E_{kin}^{solvent,3D}(t + (\Delta t/2))$, $E_{kin}^{4thD}(t + (\Delta t/2))$.

10. If required, write the energies, scaling data, and free energy data to trajectory files.

11. If required, print the energies, scaling data, and free energy data to standard output.

12. Sample the components of a vector $\vec{R}'_i$ from a Gaussian distribution with zero mean and width ρ_2^2 and determine

$$\vec{R}_i\left(t+\frac{\Delta t}{2};-\frac{\Delta t}{2}\right) = \rho_3 \vec{V}_i\left(t;\frac{\Delta t}{2}\right) + \vec{R}'_i \tag{30}$$

where

$$\rho_2^2 = \frac{k_B T_{ref} B\left(\frac{-\gamma_i \Delta t}{2}\right)}{m_i \gamma_i^2 (1 - \exp(-\gamma_i \Delta t))} \tag{31}$$

and

$$\rho_3 = -\frac{D\left(\frac{-\gamma_i \Delta t}{2}\right)}{\gamma_i (1 - \exp(-\gamma_i \Delta t))} \tag{32}$$

13. Sample the components of a vector $\vec{R}_i(t + (\Delta t/2);\ \Delta t/2)$ from a Gaussian distribution with zero mean and width σ_1^2, where

$$\sigma_1^2 = \frac{k_B T_{ref} C\left(\frac{\gamma_i \Delta t}{2}\right)}{m_i \gamma_i^{\,2}} \tag{33}$$

14. Add the stochastic integrals to obtain the unconstrained positions

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t + \Delta t) - \vec{R}_i\left(t+\frac{\Delta t}{2};-\frac{\Delta t}{2}\right) + \vec{R}_i\left(t+\frac{\Delta t}{2};\frac{\Delta t}{2}\right) \tag{34}$$

15. Make the positions $\vec{r}_i(t + \Delta t)$ satisfy the constraints by applying the SHAKE algorithm.

16. Increase the time $t_{n+1} = t_n + \Delta t$ and the step number by one. If required, change the free energy perturbation coupling parameters λ to $\lambda + \Delta\lambda$ and μ to $\mu = \mu + \Delta\mu$.

4.3. Free Energy Differences by Extrapolation. One important addition to the GROMOS96 suite of programs is the program PROFEE. It can be used to obtain free energy differences from a precalculated trajectory using the perturbation formula[12]

$$\Delta F_{BA} = F_B - F_A = -k_B T \ln\left\langle \exp\left(-\frac{H(1) - H(0)}{k_B T}\right)\right\rangle_{\lambda=0} \quad (35)$$

between two states A and B of a system, where the Hamiltonian is

$$H(\lambda) = (1 - \lambda)H_A + \lambda H_B \quad (36)$$

and λ is the coupling parameter. The reliability of the value obtained for ΔF_{BA} depends largely on whether the configurations sampled at $\lambda = 0$ are representative of the system at $\lambda = 1$. This is not the case when atoms are created or deleted in the perturbation process. It has been shown, however, that the introduction of a soft-core interaction function term as described in subsection 5.2 into the Hamiltonian for $\lambda = 0$ at sites where atoms are to be created or deleted can act as an umbrella or biasing potential. The ensemble thus produced at this nonphysical (A) state can be representative of both physical end states (B_1) and (B_2) and reliable results obtained. This concept can easily be extended to calculate the free energy difference between a number of similar systems B_1, ..., B_n. Thus, after having performed a single, comparatively time consuming, simulation at a judicially chosen nonphysical reference (A) state, the difference in free energy between the reference state and a number of physical (B) states of interest can be readily obtained by analysis of the trajectory.[13] The free energy difference between any physical states (B_i) and (B_j) can then be determined by the construction of an appropriate thermodynamic cycle.

4.4. Dynamics in Four Dimensions. Efficient sampling of phase space of a system is often hampered by the existence of barriers in the potential energy surface. One way to enhance sampling is to introduce a fourth spatial dimension in which some or all of the atoms can move, thus allowing the system to circumvent the barriers existing in three dimensions. The method has been demonstrated to be useful when performing structure refinement simulations based on NMR data[14] and in free energy calculations.[15] For the latter case, a brief outline of the concepts involved is presented. For a system of N particles in three spatial dimensions, the Hamiltonian can be written as

$$H_{3D}(p,q) = V_{3D}(q) + K_{3D}(p) \quad (37)$$

where q are the coordinates and p the momenta, and V_{3D} and K_{3D} are the potential and kinetic energies in three dimensions, respectively. If the three (x, y, and z) dimensions are not coupled to the fourth (w) dimension, then the same Hamiltonian can be written in four dimensions:

$$H_{4D}(p,q;0) = V_{3D}(q) + K_{3D}(p) + V_{4thD}(q) + K_{4thD}(p) \quad (38)$$

in which the suffix *4thD* denotes that that term is independent of the first three dimensions. A similar, but different, Hamiltonian can be defined in which the potential energy is truly evaluated in four dimensions:

$$H_{4D}(p,q;1) = V_{4D}(q) + K_{3D}(p) + V_{4thD}(q) + K_{4thD}(p) \quad (39)$$

A free energy difference between $H_{4D}(p, q; 0)$ and $H_{4D}(p, q; 1)$ can be calculated by thermodynamic integration. The parametrized Hamiltonian is then

$$H_{4D}(p,q;\mu) = \mu V_{4D}(q) + (1 - \mu)V_{3D}(q) + K_{3D}(p) + V_{4thD}(q) + K_{4thD}(p) \quad (40)$$

and the free energy difference in the canonical ensemble is

$$\Delta F_{3D\rightarrow 4D} = F_{4D}(\mu{=}1) - F_{4D}(\mu{=}0) = \int_0^1 \left\langle \frac{\partial H_{4D}(p,q;\mu)}{\partial \mu}\right\rangle_\mu d\mu \quad (41)$$

where the brackets $\langle\ \rangle$ denote the corresponding ensemble at a fixed point μ. The free energies of the system described by $H_{4D}(p,q;0)$ and $H_{3D}(p,q)$ differ by the contributions in the fourth dimension $V_{4thD}(q) + K_{4thD}(p)$. As $V_{4thD}(q)$ is chosen to be harmonic (see section 5), the difference can be calculated analytically.

4.5. Path Integral Simulations. A classical Hamiltonian can be used to study the statistical−mechanical equilibrium properties of a quantum-mechanical system. The method is based on an isomorphism[16] between a classical system of rings consisting of particles (beads) connected by harmonic oscillators and a path-integral quantum-mechanical description of the quantum system[17] The resulting classical system can be simulated with small changes to the conventional functional forms of the potential energy function terms to obtain average properties of a quantum system.[18]

5. Force Field

In this section, we discuss aspects of the force field concerned with functional form and new functionality. Parametrization aspects have been discussed in a separate publication.[19]

The potential energy function used in GROMOS96 is

$$V(\vec{r}(t);\lambda,\mu) = V^{phys}(\vec{r}(t);\lambda,\mu) + V^{special}(\vec{r}(t)) \quad (42)$$

where we distinguish between the standard physical atomic interaction

$$\begin{aligned} V^{phys}(\vec{r}(t);\lambda,\mu) &= V^{bon}(\vec{r}(t);\lambda,\mu) + V^{nonb}(\vec{r}(t);\lambda,\mu) = \\ & V^{bond}(\vec{r}(t);\lambda,\mu) + V^{angle}(\vec{r}(t);\lambda,\mu) + V^{har}(\vec{r}(t);\lambda) + \\ & V^{trig}(\vec{r}(t);\lambda,\mu) + V^{nonb}(\vec{r}(t);\lambda,\mu) = \mu[V^{bond}(\vec{r}^{\,4D}(t);\lambda) + \\ & V^{angle}(\vec{r}^{\,4D}(t);\lambda) + V^{trig}(\vec{r}^{\,4D}(t);\lambda)] + \\ & (1-\mu)[V^{bond}(\vec{r}^{\,3D}(t);\lambda) + V^{angle}(\vec{r}^{\,3D}(t);\lambda) + \\ & V^{trig}(\vec{r}^{\,3D}(t);\lambda)] + V^{har}(\vec{r}^{\,3D}(t);\lambda) + V^{nonb}(\vec{r}(t);\lambda,\mu) \end{aligned} \quad (43)$$

and nonphysical terms in

$$\begin{aligned} V^{special}(\vec{r}(t)) = V^{pr}(\vec{r}^{\,3D}(t)) + V^{dr}(\vec{r}(t)) + V^{dlr}(\vec{r}^{\,3D}(t)) + \\ V^{Jr}(\vec{r}^{\,3D}(t)) + V^{le}(\vec{r}^{\,3D}(t)) + V^{fdr}(\vec{r}^{\,4thD}(t)) \end{aligned} \quad (44)$$

which are included for a particular purpose, e.g. restraining functions. The superscript 4D over $\vec{r}$ indicates that the position vector $\vec{r}$ is to be taken as a 4-dimensional vector. Likewise, the superscript 3D over $\vec{r}$ indicates that only the 3-dimensional part (x-, y-, z-components) of the position vector $\vec{r}$ is to be taken. If only the 4th dimensional (w) component of the position vector $\vec{r}$ is meant, this is indicated by the superscript *4thD* over $\vec{r}$. The various terms of V are explained below. λ is the coupling parameter used in free energy perturbation calculations[20] and μ is an additional parameter involving the coupling between the 3D (x,y,z) dimensions and the 4th (w) dimension.[15] We have

$$\lambda = \lambda_A = 0 \qquad \text{state A}$$

$$\lambda = \lambda_B = 1 \qquad \text{state B}$$

and

$$\mu = \mu_A = 0 \qquad \text{no coupling between } \vec{r}^{\,3D}(t) \text{ and } \vec{r}^{\,4thD}(t)$$

$$\mu = \mu_B = 1 \qquad \text{interaction function uses } \vec{r}^{\,4D}(t)$$

5.1. Bonded Terms. The bonded terms, in three or four dimensions, read

$$V^{bond}(\vec{r}(t);\lambda) = \frac{1}{4}\sum_{n=1}^{N_b}\{[(1-\lambda)K_{b_n}^A + \lambda K_{b_n}^B][b_n^2(t) - ((1-\lambda)b_{0_n}^A + \lambda b_{0_n}^B)^2]^2\} \quad (45)$$

$$V^{angle}(\vec{r}(t);\lambda) = \frac{1}{2}\sum_{n=1}^{N_\theta}\{[(1-\lambda)K_{\theta_n}^A + \lambda K_{\theta_n}^B][\cos(\theta_n(t)) - ((1-\lambda)\cos(\theta_{0_n}^A) + \lambda\cos(\theta_{0_n}^B))]^2\} \quad (46)$$

$$V^{har}(\vec{r}(t);\lambda) = \frac{1}{2}\sum_{n=1}^{N_\xi}[(1-\lambda)K_{\xi_n}^A + \lambda K_{\xi_n}^B][\xi_n(t) - ((1-\lambda)\xi_{0_n}^A + \lambda\xi_{0_n}^B)]^2 \quad (47)$$

and

$$V^{trig}(\vec{r}(t);\lambda) = \sum_{n=1}^{N_\phi}\{(1-\lambda)K_{\phi_n}^A[1+\cos(\delta_n^A)\cos(m_n^A\phi_n(t))] + \lambda K_{\phi_n}^B[1+\cos(\delta_n^B)\cos(m_n^B\phi_n(t))]\} \quad (48)$$

where N_b, N_θ, N_ξ and N_ϕ denote the number of bonds, bond angles, harmonic or improper dihedral angles and trigonometric dihedral angles, $K_{b_n}^A$, $K_{b_n}^B$, $K_{\theta_n}^A$, $K_{\theta_n}^B$, $K_{\xi_n}^A$, $K_{\xi_n}^B$, and $K_{\phi_n}^A$, K_{θ}^B the respective force constants in state A or B, $b_{0_n}^A$, $b_{0_n}^B$, $\theta_{0_n}^A$, $\theta_{0_n}^{B^n}$, $\xi_{0_n}^A$, $\xi_{0_n}^B$ denote equilibrium quantities, δ_n^A,$\delta_n^B = 0$ or π denote phase shifts and m_n^A, $m_n^B = 1, 2, ..., 6$ are multiplicities. Equations 45−48 reduce to

$$V^{bond}(\vec{r}(t)) = \frac{1}{4}\sum_{n=1}^{N_b}K_{b_n}[b_n^2(t) - b_{0_n}^2]^2 \quad (49)$$

$$V^{angle}(\vec{r}(t)) = \frac{1}{2}\sum_{n=1}^{N_\theta}K_{\theta_n}[\cos(\theta_n(t)) - \cos(\theta_{0_n})]^2 \quad (50)$$

$$V^{har}(\vec{r}(t)) = \frac{1}{2}\sum_{n=1}^{N_\xi}K_{\xi_n}[\xi_n(t) - \xi_{0_n}]^2 \quad (51)$$

$$V^{trig}(\vec{r}(t)) = \sum_{n=1}^{N_\phi}K_{\phi_n}[1+\cos(\delta_n)\cos(m_n\phi_n(t))] \quad (52)$$

for $\lambda = \mu = 0$ and $K_{b_n}^A = K_{b_n}$, $K_{\theta_n}^A = K_{\theta_n}$, $K_{\xi_n}^A = K_{\xi_n}$, $K_{\phi_n}^A = K_{\phi_n}$, $b_{0_n}^A = b_{0_n}$, $\theta_{0_n}^A = \theta_{0_n}$, $\xi_{0_n}^A = \xi_{0_n}$, $m_n^A = m_n$, and $\delta_n^A = \delta_n$.

5.2. Nonbonded Terms. The nonbonded terms are

$$V^{nonb}(\vec{r}(t);\lambda,\mu) = \sum_{\substack{\text{nonbonded}\\ \text{pairs } (i,j)}}\{(1-\mu)^m[\lambda^n V^{LJCRF}(\vec{r}_{ij}^{\,3D}(t),\vec{r}_{ij}^{\,3D}(t);B;(1-\lambda),\mu) + (1-\lambda)^n V^{LJCRF}(\vec{r}_{ij}^{\,3D}(t),\vec{r}_{ij}^{\,3D}(t);A;\lambda,\mu)] + \mu^m[\lambda^n V^{LJCRF}(\vec{r}_{ij}^{\,4D}(t),\vec{r}_{ij}^{\,3D}(t);B;(1-\lambda),0) + (1-\lambda)^n V^{LJCRF}(\vec{r}_{ij}^{\,4D}(t),\vec{r}_{ij}^{\,3D}(t);A;\lambda,0)]\} \quad (53)$$

with

$$V^{LJCRF}(\vec{r}_{ij}^{\,4D}(t),\vec{r}_{ij}^{\,3D}(t);X;\lambda,\mu) = V^{LJ}(\vec{r}_{ij}^{\,4D}(t);X;\lambda,\mu) + V^{LJ}(\vec{r}_{ij}^{\,4D}(t),\vec{r}_{ij}^{\,3D}(t);X;\lambda,\mu) = \frac{1}{\alpha_{LJ}(i,j)(\lambda^2+\mu^2)C_{126}^X(i,j) + r_{ij}^{4D}(t)^6}\times\left[\frac{C_{12}^X(i,j)}{\alpha_{LJ}(i,j)(\lambda^2+\mu^2)C_{126}^X(i,j) + r_{ij}^{4D}(t)^6} - C_6^X(i,j)\right] + \frac{q_i^X q_j^X}{4\pi\epsilon_0\epsilon_1}\left[\frac{1}{[\alpha_C(i,j)(\lambda^2+\mu^2) + r_{ij}^{4D}(t)^2]^{1/2}} - \frac{^1/_2 C_{rf}\vec{r}_{ij}^{\,3D}(t)^2}{[\alpha_C(i,j)(\lambda^2+\mu^2) + R_{rf}^{\,2}]^{3/2}} - \frac{1-{}^1/_2 C_{rf}}{R_{rf}}\right] \quad (54)$$

where $X = A$, B, m, and n are positive integers and q_i^X and q_j^X the atomic partial charges in state X. $\alpha_{LJ}(i, j)$ and $\alpha_C(i, j)$ are the soft-core parameters for the Lennard-Jones and charge interactions,[21] respectively, which avoid the singularity at $r_{ij}^{3D} = 0$.

$$\alpha_{LJ}(i,j) = \begin{cases}\alpha_{LJ} & \text{if } i \text{ or } j \text{ is flagged as being a soft-core atom}\\ 0 & \text{otherwise}\end{cases} \quad (55)$$

$$\alpha_C(i,j) = \begin{cases}\alpha_C & \text{if } i \text{ or } j \text{ is flagged as being a soft-core atom}\\ 0 & \text{otherwise}\end{cases} \quad (56)$$

where α_{LJ} and α_C are the global soft-core parameters.

$$C_{126}^X(i,j) = \begin{cases}\dfrac{C_{12}^X(i,j)}{C_6^X(i,j)} & \text{if } C_6^X(i,j) \neq 0\\ 0 & \text{if } C_6^X(i,j) = 0\end{cases} \quad (57)$$

with $C_6^X(i, j)$ and $C_{12}^X(i, j)$ the r^{-6} and r^{-12} Lennard-Jones parameters for atom pair (i, j) in state X and

$$C_{rf} = \frac{(2\epsilon_1 - 2\epsilon_2)(1+\kappa R_{rf}) - \epsilon_2(\kappa R_{rf})^2}{(\epsilon_1 + 2\epsilon_2)(1+\kappa R_{rf}) + \epsilon_2(\kappa R_{rf})^2} \quad (58)$$

with ϵ_0 the dielectric permittivity in vacuum. The relative permittivity inside the sphere determined by the Poisson−Boltzmann reaction field[22] cutoff radius R_{rf} is ϵ_1 and the medium outside the sphere has relative permittivity ϵ_2 and an ionic strength characterized by the inverse Debye screening length κ. The nonbonded interactions reduce to

$$V^{nonb}(\vec{r}(t)) = \sum_{\substack{\text{nonbonded}\\ \text{pairs } (i,j)}}V^{LJ}(\vec{r}(t)) + V^{CRF}(\vec{r}(t)) = \sum_{\substack{\text{nonbonded}\\ \text{pairs } (i,j)}}\left[\frac{C_{12}(i,j)}{r_{ij}^{4D}(t)^6} - C_6(i,j)\right]\frac{1}{r_{ij}^{4D}(t)^6} + \sum_{\substack{\text{nonbonded}\\ \text{pairs } (i,j)}}\frac{q_i q_j}{4\pi\epsilon_0\epsilon_1}\left[\frac{1}{r_{ij}^{4D}(t)} - \frac{C_{rf}r_{ij}^{3D}(t)^2}{2R_{rf}^{\,3}} - \frac{1-{}^1/_2 C_{rf}}{R_{rf}}\right] \quad (59)$$

when both i and j are not perturbed, or when $\lambda = \mu = 0$. The summation over the nonbonded pairs (i,j) requires some clarification. No nonbonded interaction is calculated for atom pairs which are connected by one or two bonds or in special cases three bonds (excluded neighbors) and slightly weaker Lennard-

Jones interactions are used for certain atom types three bonds apart. GROMOS96 makes use of the charge-group concept in which the charges of atoms belonging to one charge group add up to zero or an integral value in case of charged atom groups. In the former case, the electrostatic interaction potential is dipolar ($\propto 1/r^3$) which leads to a significant reduction in its range. GROMOS96 makes use of the twin-range method[23] in which a long-range force is calculated with a cutoff distance of R_{cl} when the charge group pairlist is constructed and held constant between pairlist updates. The nonbonded interactions are evaluated at every intermediate step using the shorter cutoff R_{cp} to which the long-range contribution is added. The Poisson−Boltzmann reaction field terms represent the force on an atom i induced by atom j when assuming an electrostatic continuum.[22] This last term is always calculated in three dimensions.

The nonphysical interaction terms in GROMOS96 are

$$V^{\text{special}}(\vec{r}(t)) = V^{\text{pr}}(\vec{r}^{\,3\text{D}}(t)) + V^{\text{dr}}(\vec{r}(t)) + V^{\text{dlr}}(\vec{r}^{\,3\text{D}}(t)) + V^{\text{Jr}}(\vec{r}^{\,3\text{D}}(t)) + V^{\text{le}}(\vec{r}^{\,3\text{D}}(t)) + V^{\text{fdr}}(\vec{r}^{\,4\text{thD}}(t)) \quad (60)$$

which we discuss in the following paragraphs.

5.3. Position Restraining. The position restraining potential energy term is of the usual harmonic form:

$$V^{\text{pr}}(\vec{r}^{\,3\text{D}}(t)) = {}^1/_2 \sum_{n=1}^{N_n^{\text{pr}}} K_n^{\text{pr}} [\vec{r}_n^{\,3\text{D}}(t) - \vec{r}_{0_n}]^2 \quad (61)$$

where $\vec{r}_{0_n}$ and $\vec{r}_{0_n}^{\,3\text{D}}(t)$ denote the reference and actual position, respectively, of the nth position restraint. The harmonic oscillator force constant K_n^{pr} can be chosen to be equal for all atoms, or to be inversely proportional to the individual atomic B factors. Position restraining is always performed in three dimensions.

5.4. Distance Restraining. The distance restraining potential energy term[24−28] used in GROMOS96

$$V^{\text{dr}}(\vec{r}(t)) = \sum_{m=1}^{N_m^{\text{dr}}} V_m^{\text{dr}}(r_{nn'}^{4\text{D}}(t), K_m^{\text{dr}}, r_m^0, \Delta r^{\text{h}}) \quad (62)$$

involving atoms n and n' is harmonic for small deviations from the atom−atom reference distance r_m^0 and linear (positive or negative) for larger deviations. It can be of two forms: attractive

$$V_m^{\text{dr}}(r_{nn'}^{4\text{D}}(t), K_m^{\text{dr}}, r_m^0, \Delta r^{\text{h}}) = \begin{cases} 0 & 0 < r_{nn'}^{4\text{D}}(t) < r_m^0 \\ {}^1/_2 K_m^{\text{dr}} [r_{nn'}^{4\text{D}}(t) - r_m^0]^2 & r_m^0 < r_{nn'}^{4\text{D}}(t) < r_m^0 + \Delta r^{\text{h}} \\ + K_m^{\text{dr}} [r_{nn'}^{4\text{D}}(t) - r_m^0 - {}^1/_2 \Delta r^{\text{h}}] \Delta r^{\text{h}} & r_m^0 + \Delta r^{\text{h}} < r_{nn'}^{4\text{D}}(t) \end{cases} \quad (63)$$

or repulsive

$$V_m^{\text{dr}}(r_{nn'}^{4\text{D}}(t), K_m^{\text{dr}}, r_m^0, \Delta r^{\text{h}}) = \begin{cases} - K_m^{\text{dr}} [r_{nn'}^{4\text{D}}(t) - r_m^0 + {}^1/_2 \Delta r^{\text{h}}] \Delta r^{\text{h}} & 0 < r_{nn'}^{4\text{D}}(t) < r_m^0 - \Delta r^{\text{h}} \\ {}^1/_2 K_m^{\text{dr}} [r_{nn'}^{4\text{D}}(t) - r_m^0]^2 & r_m^0 - \Delta r^{\text{h}} < r_{nn'}^{4\text{D}}(t) < r_m^0 \\ 0 & r_m^0 < r_{nn'}^{4\text{D}}(t) \end{cases} \quad (64)$$

where $r_{nn'}^{4\text{D}}(t)$ denotes the actual atom−atom distance. $r_m^0 + \Delta r^{\text{h}}$ (attractive case) and $r_m^0 - \Delta r^{\text{h}}$ (repulsive case) are the distances at which the interaction term changes from a harmonic to a linear function of $r_{nn'}^{4\text{D}}(t)$.

The force constant K_m^{dr} can be chosen to be equal for all atoms or proportional to atomic weight factors. Distance restraining can be performed in four dimensions.

When the given atom-atom distance restraints r_m^0 have been derived from nuclear overhauser effect (NOE) cross peak intensities originating from nuclei n and n', they represent an average over space and time. In order to account for this, it is possible to replace the instantaneous distance $r_{nn'}^{4\text{D}}(t)$ with an average, so that only this average is restrained. In GROMOS96 time averaging is implemented as described in ref 26. The ensemble is taken as the time trajectory average with a memory relaxation time τ_{dr}:

$$\langle r_{nn'}{}^{-3} \rangle = \overline{r_{nn'}{}^{-3}(t;\tau_{\text{dr}})} \equiv \frac{1}{[\tau_{\text{dr}}[1 - \exp(-t/\tau_{\text{dr}})]]} \int_0^t \exp\left(-\frac{t-t'}{\tau_{\text{dr}}}\right) r_{nn'}{}^{-3}(t')\, \mathrm{d}t' \quad (65)$$

5.5. Dihedral Angle Restraining. The dihedral angle restraining potential is of the form

$$V^{\text{dlr}}(\vec{r}^{\,3\text{D}}(t)) = {}^1/_2 \sum_{n=1}^{N_{\text{dlr}}} K_n^{\text{dlr}} [\phi_n(t) - \phi_n^0]^2 \quad (66)$$

which is the same as the improper dihedral term. It is always calculated in three dimensions. The dihedral angle ϕ_n, is defined by specifying the atoms i, j, k and l.

5.6. *J*-Coupling Constant Restraining. J-coupling constant restraining can be used to restrain the spin-spin ${}^3J$ coupling constant ${}^3J_{mm'}$ between two nuclei m and m' to a given experimental value ${}^3J_{mm'}^0$ or ${}^3J_n^0$. The ${}^3J$ coupling constant depends on the value of the dihedral angle $\theta(m, j, k, m') = \theta_n$ involving the three covalent bonds connecting the atoms m and m' via atoms j and k through the Karplus relation

$${}^3J_n = {}^3J_{mm'} = a \cos^2(\theta(m,j,k,m')) + b \cos(\theta(m,j,k,m')) + c \quad (67)$$

where a, b, and c are the empirically derived Karplus constants. The potential energy term for J-coupling constant restraining[29] in GROMOS96 is of the form

$$V^{\text{Jr}}(\vec{r}^{\,3\text{D}}(t)) = \sum_{n=1}^{N_{\text{Jr}}} V_n^{\text{Jr}}(\overline{{}^3J(\theta_n;\tau_{\text{Jr}})}\,; K_n^{\text{Jr}}, {}^3J_n^0) \quad (68)$$

where

$$\overline{{}^3J(\theta_n;\tau_{\text{Jr}})} = \frac{1}{[\tau_{\text{Jr}}[1 - \exp(-t/\tau_{\text{Jr}})]]} \int_0^t \exp\left(-\frac{t-t'}{\tau_{\text{Jr}}}\right) {}^3J_n(t')\, \mathrm{d}t' \quad (69)$$

is a weighted time average defined in a similar way to that used in distance restraining and τ_{Jr} is the memory relaxation time. The functional form of $V_n^{\text{Jr}}(\overline{{}^3J(\theta_n;\tau_{\text{Jr}})}\,; K_n^{\text{Jr}}, {}^3J_n^0)$ can be chosen to be harmonic

$$V_n^{\text{Jr}}(\overline{{}^3J(\theta_n;\tau_{\text{Jr}})};\ K_n^{\text{Jr}}, {}^3J_n^0) = {}^1/_2 K_n^{\text{Jr}} [\overline{{}^3J(\theta_n;\tau_{\text{Jr}})} - {}^3J_n^0]^2 \quad (70)$$

half-harmonic (attractive)

$$V_n^{\mathrm{Jr}}(\overline{{}^3\mathcal{J}(\theta_n;\tau_{\mathrm{Jr}})}\,;K_n^{\mathrm{Jr}},\,{}^3J_n^0) = \begin{cases} 0 & \overline{{}^3\mathcal{J}(\theta_n;\tau_{\mathrm{Jr}})} < {}^3J_n^0 \\ {}^1/_2K_n^{\mathrm{Jr}}\,[\overline{{}^3\mathcal{J}(\theta_n;\tau_{\mathrm{Jr}})} - {}^3J_n^0]^2 & \overline{{}^3\mathcal{J}(\theta_n;\tau_{\mathrm{Jr}})} > {}^3J_n^0 \end{cases} \quad (71)$$

or half-harmonic (repulsive)

$$V_n^{\mathrm{Jr}}(\overline{{}^3\mathcal{J}(\theta_n;\tau_{\mathrm{Jr}})};\;K_n^{\mathrm{Jr}},\,{}^3J_n^0) = \begin{cases} {}^1/_2K_n^{\mathrm{Jr}}\,[\overline{{}^3\mathcal{J}(\theta_n;\tau_{\mathrm{Jr}})} - {}^3J_n^0]^2 & \overline{{}^3\mathcal{J}(\theta_n;\tau_{\mathrm{Jr}})} < {}^3J_n^0 \\ 0 & \overline{{}^3\mathcal{J}(\theta_n;\tau_{\mathrm{Jr}})} > {}^3J_n^0 \end{cases} \quad (72)$$

$V^{\mathrm{Jr}}(\vec{r}^{\,3\mathrm{D}}(t))$ is always calculated in three dimensions.

5.7. Local Elevation. Local elevation[30] is a searching method in which a record of visited conformations is kept during the simulation in order to energetically penalize them and thus force the system to search other areas of phase space, overcoming barriers in the original energy surface. In GROMOS96, a conformation is defined by a set of discretized dihedral angles. The potential energy term reads

$$V^{\mathrm{le}}(\vec{r}^{\,3\mathrm{D}}(t)) = E_0^{\mathrm{le}}N(\phi^{\mathrm{m}})\prod_{n=1}^{N_{\mathrm{le}}}V_n^{\mathrm{le}}(\phi_n;\Delta\phi^0) \quad (73)$$

where the product runs over a set of N_{le} selected dihedral angles ϕ_n which are defined by four atoms i_n, j_n, k_n, and l_n forming the dihedral angle $\phi(i_n,j_n,k_n,l_n)$. The range of a dihedral angle ϕ_n, $\in[0 \ldots 360]$ is divided into N_0 equal segments, where segment number $m_n = 1 \ldots N_0$ is defined by

$$(m_n - {}^3/_2)\Delta\phi^0 < \phi_n < (m_n - {}^1/_2)\Delta\phi^0 \quad (74)$$

where $\Delta\phi^0 = 360/N_0$. The midpoint angle value of segment m_n is $\phi_n^{\mathrm{m}} = (m - 1)\Delta\phi^0$. A local elevation energy (73) is determined by the set of local elevation dihedral angles $\vec{\phi} = (\phi(i_1, j_1, k_1, l_1), \ldots, \phi(i_{N_{\mathrm{le}}}, j_{N_{\mathrm{le}}}, k_{N_{\mathrm{le}}}, l_{N_{\mathrm{le}}}))$ and its corresponding discretized configuration, the midpoint angles of the segments $\vec{\phi}^m = (\phi_1^m, \ldots, \phi_{N_{\mathrm{le}}}^m)$ or, alternatively, by the set of corresponding segment numbers $\vec{m} = (m_1, \ldots, m_{N_0})$ that the dihedral angles $\phi(i_n, j_n, k_n, l_n)$ fall into. $N(\phi^m)$ is the number of times the conformation $\vec{\phi}^m$ has been visited during the simulation and the factor E_0^{le} determines the increase of penalization per resampling of a conformation. The factor in (73) is

$$V_n^{\mathrm{le}}(\phi_n;\Delta\phi^0) = \prod_{m'=1}^{N_0}\exp\left(\frac{-(\Delta(\phi_n;\,\phi_n^{m'}))^2}{2(\Delta\phi^0)^2}\right) \quad (75)$$

where

$$\Delta(\phi_n;\,\phi_n^m) = \begin{cases} \phi_n - \phi_n^m & |\phi_n - \phi_n^m| < \Delta\phi^0/2 \\ 0 & |\phi_n - \phi_n^m| > \Delta\phi^0/2 \end{cases} \quad (76)$$

Local elevation potential energy is always calculated in three dimensions.

5.8. Fourth Dimension Restraining. When performing MD or SD in four dimensional space, one may wish to restrict the sampling of the fourth dimension, since, in the end, the molecular system must be forced back into three dimensions. The fourth dimension restraining potential energy term reads

$$V^{\mathrm{fdr}}(\vec{r}^{\,4\mathrm{thD}}(t)) = {}^1/_2\sum_{n=1}^{N_{\mathrm{fdr}}}K_n^{\mathrm{fdr}}w_n^{\,2}(t) \quad (77)$$

where $w_n(t)$ is the actual fourth dimension coordinate of the nth atom in four dimensions.

6. Reliability and Efficiency

The design and implementation of a complex program package by a design team has its own set of problems which must be addressed. The development team met weekly to review progress. All implementation decisions were recorded in a specification document. This document proved useful in several respects.

- New members of the development team were provided with a means to assess the status of the project.
- It gave a concise overview of the project and a better understanding how parts of the code should interact.
- It provided a framework for the logical development of the project. By maintaining a record of past decisions wasteful backtracking and revision could be minimized.

Testing of the code took a variety of forms:

- Independent code reading. Code that any one person wrote was reviewed by at least one other person. Integration of code contributed by individual programmers was performed by one person in order to ensure consistency. At the same time, the quality of the code and its documentation was evaluated. Programming guidelines were drawn up at the beginning of the project in order to encourage a uniform coding style and to improve legibility and ensure portability.
- Independent code testing. Code provided by any one person was tested (i.e., compiled and executed) for correctness by another person. In some cases, individual subroutines were tested by writing special "driver programs" in order to test the subroutines in isolation with many input combinations.
- Portability checking. GROMOS96 is designed to run on a wide variety of platforms, a feature which was repeatedly tested throughout its development. Compiling and running on different platforms during development was used to find machine or compiler dependencies.
- Consistency checking. The same functionality is sometimes performed by one portion of code in a more general form than in another. A consistency check can be carried out by performing test runs using the more general code with program input chosen such as to produce the same results as in the simplified case. For example, the SD algorithm, with its stochastic behavior, is not easy to test conclusively. However, it must produce the same results as the MD algorithm for $\gamma_i = 0$.

A wide range of tests were performed, including the following:

- Each individual force term (bonds, bond angles, etc.) was numerically integrated and checked for agreement with the potential energy.
- The sum of all atomic forces must be zero when no position restraining or constraining is used.
- When performing an MD simulation without temperature or pressure scaling, and with the pairlist updated every step, energy must be conserved using sufficiently small time steps.
- When performing a free energy calculation with the state A equivalent to the state B, the simulation results must be the same for any value of λ. Note that this does not hold for simulations involving soft-core interactions.
- Comparison of results produced by GROMOS96 with those produced by GROMOS87. A direct comparison was not always

TABLE 1: GROMOS96 Benchmark Suite[a]

benchmark	molecules	Nsm	Nsa	Nsolvm	Nsolva	Na	boundary	NMD
I	cyclosporin A	1	90	0	0	90	vacuo	1000
II	cyclosporin A (water)	1	90	764	2298	2382	octa	100
III	thrombin	1	3078	0	0	3078	vacuo	100
IV	thrombin (water)	1	3078	5427	16281	19359	octa	10
V	H_2O (medium)	0	0	1728	5184	5184	cubic	100
VI	H_2O (large)	0	0	13824	41472	41472	cubic	10

[a] Nsm, number of solute molecules; Nsa, number of solute atoms; Nsolvm, number of solvent molecules; Nsolva, number of solvent atoms; Na, number of atoms; NMD, number of MD time steps; vacuo, vacuum boundary condition; octa, periodic truncated octahedron; cubic, periodic cubic box; $R_{cp} = R_{cl} = 0.8$ nm for II−VI ($R_{cl} = 1.4$ nm for IV); $R_{cp} = R_{cl} = \infty$ for I. Pair list update every 5 time steps. Time step = 0.002 ps. (*N*, *V*, *T*) simulation.

TABLE 2: GROMOS96 Benchmark Results (in s)[a]

	I cyclosporin A	II cyclosporin A (water)	III thrombin	IV thrombin (water)	V H_2O (medium)	V H_2O (large)
Cray J90-32/16	25.1	70.1	87.4	97.4[b]	136.1	138.3
DEC 21264 AS8400 (575 MHz)	1.3	7.7	5.4	11.1[b]	15.8	22.8
DEC 21164						
1 × AS8400 (300 MHz) CPU	6.2	29.5	20.5	53.5	63.4	82.8
4 × AS8400 (300 MHz) CPU	3.9	10.0	9.5	17.3	20.4	24.3
HP 9000-735 (99 MHz)	18.0	96.9	75.7	218.4	210.5	374.2
IBM 6000/397 (160 MHz)	8.7	51.8	38.6	114.8	104.3	173.6
Pentium II (400 MHz)	3.9	24.6	15.7	43.6	50.5	59.8
SGI Octane (195 MHz)	4.2	19.7	16.2	33.4	41.5	54.4
SGI Power Challenge						
1 × R8000 (75 MHz) CPU	13.7	40.9	48.0	66.9	75.7	107.5
4 × R8000 (75 MHz) CPU	6.4	15.9	20.1	21.5	30.3	37.7
SUN Ultra-10 (300 MHz)	6.4	26.8	22.6	68.4	53.0	73.4
SUN E4000						
1 × Ultra (336 MHz) CPU	4.5	17.9	15.8	33.7	36.1	49.4
4 × Ultra (336 MHz) CPU	2.5	6.3	7.1	9.9	12.5	15.3

[a] Three versions of GROMOS96 are currently distributed, a sequential version, a shared-memory parallel version, and a vectorized version. The latter two incorporate special (fast) nonbonded interaction subroutines. Timings are given for the most efficient version on each platform. This was the parallel version on all machines listed with the exception of the Cray on which the vector version is faster. [b] Pair list updated every 10 steps.

possible due to the new functionality of GROMOS96. In addition, direct comparisons were hampered by the fact that functional forms for some covalent interaction terms are different in the two GROMOS versions.

In order to compare the performance of the code on a number of machines, and in order to estimate the amount of time a particular simulation might take, a set of benchmarks has been defined (see Table 1) covering a large range of system size. The large H_2O system is 8 times the size of the smaller system which allows the testing of a machine's caching performance. Table 2 shows benchmark results for a number of computers. Note, while these timings are indicative, performance is dependent on the precise configuration of the machines.

7. Examples of Application

7.1. Local Elevation Conformational Search for Loop 33−43 of Ribonuclease A. The loop consisting of residues 33−43 in the structure of the 124-residue protein ribonuclease A (entry 6RSA in the Brookhaven Protein Data Bank[3]) was chosen to demonstrate the utility of the local elevation conformational search technique. All simulations were performed with the main chain atoms of residues 1−32 and 44−124 positionally restrained with a force constant of 100 kJ mol^{-1} nm^{-2}. The disulfide bridge between Cys 40 and Cys 95 was reduced. Simulations in vacuo were performed using the 43B1 force field whereas those in water were performed using the 43A1 force field.[1] For the local elevation simulations, a force constant of $E_0^{le} = 1500$ kJ mol^{-1} was chosen, N_0 was set to 16 (see section 5.7) and 10 (artificial) torsional dihedral angles (residues 32-33-34-35, 33-34-35-36, ..., 41-42-43-44) were defined between

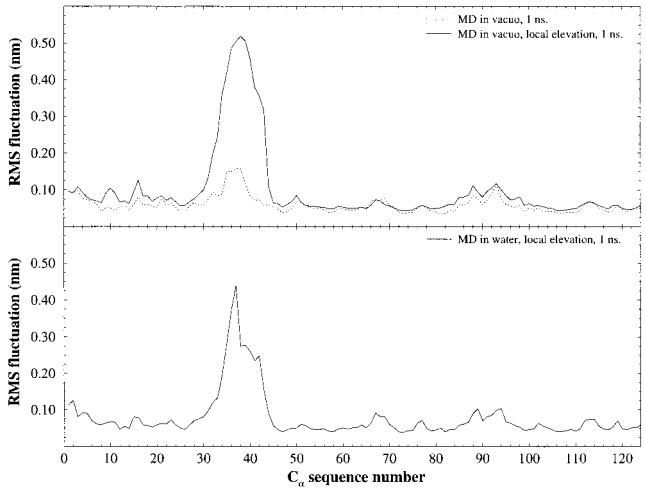

Figure 2. Root mean square fluctuations of atomic positions in ribonuclease A. Upper graph: simulations in vacuo with and without local elevation search. The simulation using local elevation search produces larger positional fluctuations than without, which is indicative of the larger conformational space searched. Lower graph: simulation in solvent with local elevation search. The fluctuations are smaller than those encountered in the vacuum simulations.

C_α atoms of the loop under study. Two simulations in vacuo, with and without the local elevation potential energy term, were performed. In addition, one simulation with the local elevation potential energy term under truncated octahedron periodic boundary conditions and with 5774 water molecules was performed. All simulations were 1 ns in length. The loop is comprised primarily of polar residues and is exposed to solvent.

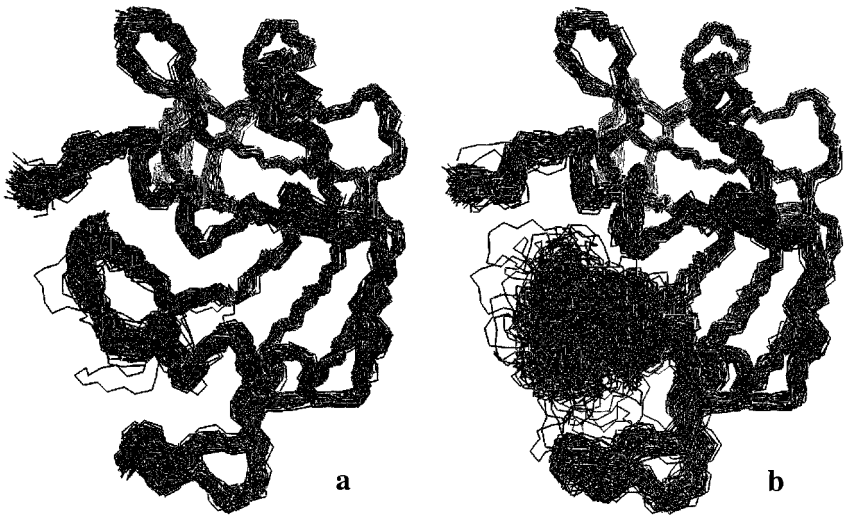

Figure 3. Superposition of 200 conformations of ribonuclease A taken at 5 ps intervals (total of 1 ns simulation time) in vacuo: (a) simulation without local elevation; (b) with local elevation search in the loop consisting of residues 33−43. The larger conformational space searched is apparent.

Solvent effects are therefore expected to play an important role in the conformations that are energetically favored. Figure 2 (upper graph) shows the root mean square fluctuations of the main chain atoms of the protein. It is clear from Figure 2 that a much broader range of configurations are sampled when the local elevation potential energy term is applied. The explicit treatment of solvent molecules reduces the efficiency of the sampling somewhat, as expected (Figure 2, lower graph). This is also evident in Figure 3 which shows an overlay of configurations generated during the simulations of 1ns in length.

7.2. Conformational Search in Four Dimensions for Cyclosporin A. Cyclosporin A is a cyclic peptide of 11 residues which adopts essentially two different conformations depending on whether it is bound to cyclophilin or free in solution.[31] Interconversion between these conformers is difficult to observe during the course of a simulation in three dimensions due to steric hindrance of the side chain atoms. By allowing the molecule to move temporarily into a fourth dimension, the interconversion is much eased, which demonstrates the utility of SD or MD simulation in four dimensions.

The structure (PDB entry 1CYA) of cyclosporin A bound to cyclophilin determined by NMR[31] and an X-ray derived structure[32] which is very similar to the conformation in apolar solvents[32] were taken as representatives of the two conformations. All simulations were performed using the stochastic dynamics method at a temperature of 300 K using a time step of 0.002 ps in vacuo using the 43B1 force field and an atomic friction coefficient of 91 ps^{-1} for all atoms. No bond constraints or temperature coupling in three dimensions were employed and no cutoffs were used for long-range interactions.

A set of artificial distance restraints was generated by comparing the interatomic distance d_{ij}^{A} and d_{ij}^{B} of the two conformers A (1CYA) and B (X-ray) for all atoms i and j. In the case of $d_{ij}^{A} < d_{ij}^{B}$ a repulsive distance restraint was postulated (eq 64), otherwise an attractive distance restraint was postulated (eq 63). As the molecule was modeled using 90 atoms, this resulted in a total of 4005 (2318 attractive, 1687 repulsive) distance restraints which were applied with a force constant of $K_m^{dr} = 10$ kJ mol^{-1} nm^{-2} during the simulations in order to force the system to adopt the X-ray conformation (B).

For the simulations in four dimensions, the bond and nonbonded interaction terms were calculated in four dimensions, whereas bond angle and dihedral angle interactions and distance restraints were calculated in three dimensions. Only the side chain atoms were specified as being allowed to move into the fourth dimension. The temperature in the fourth dimension was coupled to a temperature bath of 30 K with a relaxation time of 0.1 ps.

The following simulations were performed in three dimensions:

1. Starting from the 1CYA structure, an energy minimization was performed.

2. Starting from the minimized structure, initial atomic velocities were taken from a Maxwell distribution at 300 K . An SD simulation of 1 ns in length was performed with the distance restraints in place.

The following simulations were performed in four dimensions:

1. Starting from the 1CYA structure, an energy minimization was performed. A value of $K_n^{fdr} = 1000$ kJ mol^{-1} nm^{-2} was used for the harmonic restraint in the fourth dimension (see eq 77 in subsection 5.8).

2. Starting from the minimized structure, initial atomic velocities were taken from a Maxwell distribution at 300 K for the three-dimensional atomic components. A corresponding distribution was sampled at 30 K for the fourth dimension. An SD simulation of 500 ps was performed without distance restraints and a value of $K_n^{fdr} = 10$ kJ mol^{-1} nm^{-2}.

3. Starting from the final configuration of the previous step, an SD simulation of 1 ns was performed with distance restraints and a value of $K_n^{fdr} = 10$ kJ mol^{-1} nm^{-2}.

4. Starting from the final configuration of the previous step, an SD simulation of 1 ns was performed with distance restraints. The K_n^{fdr} was linearly increased from l0 to 100 kJ mol^{-1} nm^{-2}. At the same time, the reference temperature of the temperature bath in the fourth dimension was scaled linearly from 30 to 10 K.

The most difficult part of a study in four dimensions is moving the system from three dimensions to four dimensions and back. A system which is in a local minumum in three dimensions is very unlikely to be in one in four dimensions. As a consequence, the system will often severely heat up until the integrator fails if a simulation is started directly from a structure in 3D. It is therefore advantageous to first minimize the system in four dimensions before attempting an SD

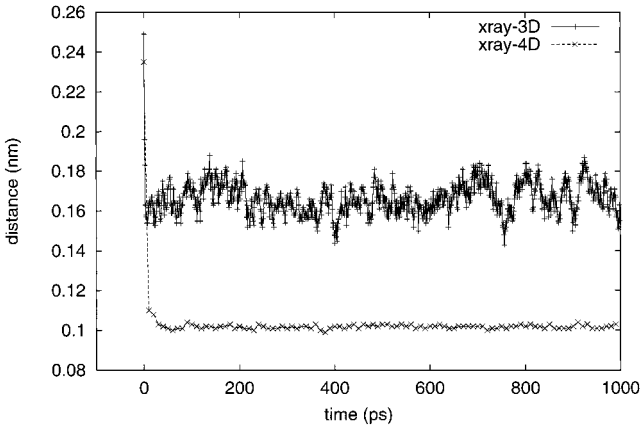

Figure 4. Root mean square distance calculated in three dimensions between the simulated conformation of cyclosporin A to the target X-ray structure in simulations in three and four spatial dimensions. In three dimensions, the system is limited in its movement by steric hindrance of the side chains. In four dimensions, these barriers are overcome to a large extent.

simulation (step 1). Note the relatively high value for K_n^{fdr} in this step. Step 2 allows the system to relax in four dimensions before the distance restraints are applied in step 3. The extent to which a system will move into four dimensions is governed by K_n^{fdr} and the temperature in the fourth dimension. Choosing a suitable combination of values is done by trial and error. Figure 4 compares the all atom 3D root mean square distance to the (target) X-ray structure as a function of time for the simulation in step 2 in three dimensions and for the simulation in step 3 in four dimensions. The distance to the target structure is considerably smaller in the 4D simulation. It was found that the value of K_m^{dr} has little bearing on how closely the X-ray structure can be approximated (data not shown).

Acknowledgment. The authors thank Xavier Daura for providing the material discussed in subsection 7.1.

References and Notes

(1) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide.* VdF: Hochschulverlag AG an der ETH Zürich and BIOMOS b.v, Zürich, Groningen, 1996; ISBN 3 7281 2422 2.

(2) van Gunsteren, W. F.; Berendsen, H. J. C., *Groningen Molecular Simulation (GROMOS) Library Manual;* BIOMOS b.v., University of Groningen, Groningen, 1987.

(3) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535−542.

(4) Havel, T.; Wüthrich, K. *Bull. Math. Biol.* **1984**, *46*, 673−698.

(5) Braun, W.; Go, N. *J. Mol. Biol.* **1985**, *186*, 611−626.

(6) Güntert, P.; Braun, W.; Wüthrich, K. *J. Mol. Biol.* **1991**, *217*, 517−530.

(7) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.

(8) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311−1327.

(9) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(10) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Simul.* **1988**, *1*, 173−185.

(11) Yun-yu, S.; Lu, W.; van Gunsteren, W. F. *Mol. Simul.* **1988**, *1*, 369−383.

(12) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420−1426.

(13) Liu, H.; Mark, A. E.; van Gunsteren, W. F. *J. Phys. Chem.* **1996**, *100*, 9485−9494.

(14) van Schaik, R. C.; Berendsen, H. J. C.; Torda, A. E.; van Gunsteren, W. F. *J. Mol. Biol.* **1993**, *234*, 751−762.

(15) Beutler, T. C.; van Gunsteren, W. F. *J. Chem. Phys.* **1994**, *101*, 1417−1422.

(16) Chandler, D.; Wolynes, P. G. *J. Chem. Phys.* **1981**, *74*, 4078−4095.

(17) Feynman, R. P.; Hibbs, A. R. *Path Integrals and Quantum Mechanics;* McGraw Hill: New York, 1965.

(18) Billeter, S. R.; King, P. M.; van Gunsteren, W. F. **1994**, *J. Chem. Phys. 100*, 6692−6699.

(19) Daura, X.; Mark, A. E.; van Gunsteren, W. F. **1998**, *J. Comput. Chem. 19*, 535−547.

(20) van Gunsteren, W. F.; Beutler, T. C.; Fraternali, F.; King, P. M.; Mark, A. E.; Smith, P. E. *Computer Simulation of Biomolecular Systems, Theoretical and Experimental Applications;* Escom Science Publishers: Leiden, The Netherlands, 1993; Vol. 2, pp 315−348.

(21) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1994**, *222*, 529−539.

(22) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 5451−5459.

(23) Berendsen, H. J. C. in *Molecular Dynamics and Protein Structure;* Polycrystal Book Service: Western Springs, IL, 1985; pp 18−22.

(24) van Gunsteren, W. F.; Boelens, R.; Kaptein, R.; Scheek, R. M.; Zuiderweg, E. R. P. In *Molecular Dynamics and Protein Structure;* Polycrystal Book Service: Western Springs, IL, 1985; pp 92−99.

(25) Torda, A. E.; van Gunsteren, W. F. *Reviews in Computational Chemistry, Volume III;* VCH Publishers Inc.: New York, 1992; pp 143−172.

(26) Torda, A. E.; Scheek, R. M.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1989**, *157*, 289−294.

(27) van Gunsteren, W. F.; Brunne, R. M.; Gros, P.; van Schaik, R. C.; Schiffer, C. A.; Torda, A. E. *Methods in Enzymology: Nuclear Magnetic Resonance;* Academic Press: New York, Vol. 239, pp 619−654.

(28) Nanzer, A. P.; van Gunsteren, W. F.; Torda, A. E. *J. Biomol. NMR* **1995**, *6*, 313−320.

(29) Torda, A. E.; Brunne, R. M.; Huber, T.; Kessler, H.; van Gunsteren, W. F. *J. Biomol. NMR* **1993**, *3*, 55−66.

(30) Huber, T.; Torda, A. E.; van Gunsteren, W. F. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 695−708.

(31) Spitzfaden, C.; Weber, H.-P.; Braun, W.; Kallen, J.; Wider, G.; Widmer, H.; Walkinshaw, M. D.; Wüthrich, K. *FEBS Lett.* **1992**, *300*, 291−300.

(32) Loosli, H. R.; Kessler, H.; Oschkinat, H.; Weber, H.-P.; Petcher, T. J.; Widmer, A. *Helv. Chim. Acta* **1985**, *68*, 682−704.