Received: 26 February 2021 | Revised: 30 March 2021 | Accepted: 6 April 2021

DOI: 10.1002/jcc.26541

FULL PAPER

Journal of COMPUTATIONAL CHEMISTRY WILEY

On the use of multiple-time-step algorithms to save computing effort in molecular dynamics simulations of proteins

Maria Pechlaner[1] | Chris Oostenbrink[2] | Wilfred F. van Gunsteren[1]

[1]Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, Zurich, Switzerland

[2]Institute of Molecular Modelling and Simulation, University of Natural Resources and Life Sciences, Vienna, Austria

Correspondence
Maria Pechlaner, Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, ETH, 8093 Zurich, Switzerland.
Email: maria.pechlaner@chem.ethz.ch

Abstract

Computer simulation of proteins in aqueous solution at the atomic level of resolution is still limited in time span and system size due to limited computing power available and thus employs a variety of time-saving techniques that trade some accuracy against computational effort. Examples of such time-saving techniques are the application of constraints to particular degrees of freedom or the use of a multiple-time-step (MTS) algorithm distinguishing between particular forces when integrating Newton's equations of motion. The application of two types of MTS algorithms to bond-stretching forces versus the remaining forces in molecular dynamics (MD) simulations of a protein in aqueous solution or of liquid water is investigated and the results in terms of total energy conservation and the influence on various other properties are compared to those of MD simulations of the same systems using bond-length, and for water bond-angle, constraints. At comparable computational effort, the use of bond-length constraints in proteins leads to better energy conservation and less distorted properties than the two MTS algorithms investigated.

KEYWORDS
constraints, molecular dynamics, multiple-time-step integration, trypsin inhibitor, water

1 | INTRODUCTION

Since the first computer simulations of a protein more than four decades ago,[1,2] the method of molecular dynamics (MD) has found wide application in chemistry, biochemistry and molecular biology.[3–5] Due to the ever-increasing computing power, larger and larger molecular systems, containing many thousands of atoms, could be simulated over longer and longer time periods, up till milliseconds. This was also made possible by the use of a variety of time-saving techniques.[6] Such techniques trade a reduced accuracy against a reduced computational effort. In view of the still limited size of the systems simulated and the limited time-length of generally accessible molecular simulations, time-saving techniques are still indispensable in MD simulations of proteins and bio-molecular complexes. Examples of such techniques are the omission of degrees of freedom, for example, by coarse-graining the atomic molecular model into a supra-atomic or supra-molecular one, the application of a cut-off distance for nonbonded interactions, and the application of constraints or the use of a multiple-time-step (MTS) algorithm when integrating Newton's equations of motion forward in time. The idea behind the latter two techniques is to remove or to separately integrate the fast oscillatory motions of lower interest in a protein, for example, the bond-stretching vibrations, thereby reducing the computational effort considerably. The application of bond-length constraints has long become a standard practice in bio-molecular simulation, whereas the application of an MTS algorithm to integrate bond-stretching vibrational forces separately from other forces in a molecule has not yet become standard practice. An advantage of the use of bond-length constraints over the application of an MTS algorithm to the bond-stretching forces is the following. Bond-stretching vibrations are governed by quantum mechanics. At ambient temperature and atmospheric pressure, the bond-stretching degrees of freedom are usually in their ground states. A bond in its ground state is best approximated by a constrained (fixed) bond length, rather than by a classically, for example, harmonically, vibrating bond, which yields the wrong

configurational distribution and heat capacity.[7] In addition, when applying an MTS algorithm to bond-stretching forces, the molecular configurations sampled between updates of the other (nonbond-stretching) forces can be highly unphysical, which may induce spurious resonances between the bond-stretching and other degrees of freedom, leading to inappropriate configurations being sampled. On the other hand, the algorithms to impose bond-length constraints on a protein inhibit an optimal use of modern distributed-processor computer systems, which require parallel processing of computing tasks in a simulation, whereas an MTS algorithm applied to bond-stretching forces allows straightforward parallelization. The additional computational effort required by each of these two techniques is comparable, in any case much less than the effort required to evaluate the nonbonded interactions.

Another aspect of the use of a time-saving technique is whether and to which extent it distorts the dynamical and statistical-mechanical properties of the simulated system. The length of the time step Δt in a molecular dynamics (MD) simulation is limited by the highest frequency (ν_{max}) motions occurring in the system,

$$\Delta t < < 1/\nu_{max}. \quad (1)$$

ν_{max} can be decreased by freezing high-frequency internal vibrations, such as bond-length vibrations, viz. by the application of distance constraints. This then allows for a longer time step Δt. Alternatively, the high-frequency internal forces can be discretely integrated forward in time using a smaller time step Δt than the time step $\Delta t'$ that is used to integrate the other forces,

$$\Delta t < < \Delta t'. \quad (2)$$

Although such internal vibrations are often not of primary interest, the application of constrained dynamics or an MTS algorithm only makes sense physically and computationally when.[3,8]

1. the vibrational frequencies of the frozen (constrained) or fast-moving degrees of freedom are (substantially) higher than those of the remaining ones, thereby allowing for a (substantial) increase of Δt, or for a $\Delta t'$ (substantially) larger than Δt,
2. the motions along the frozen or fast-moving higher-frequency degrees of freedom are only weakly coupled to the motions along the nonconstrained or lower-frequency degrees of freedom, that is, when the molecular motion is not significantly affected by the application of the constraints or the separate integration of the equations of motion of the higher-frequency degrees of freedom,
3. the computing effort to impose the constraints or to compute the higher-frequency (bond-stretching) forces is much smaller than that of calculating the nonconstrained lower-frequency (nonbonded) forces.

Bond-length degrees of freedom in molecules largely satisfy these conditions. Their vibrational frequencies are higher than those of the other degrees of freedom, their oscillations are largely decoupled from the other motions in a molecule,[9] and algorithms to impose distance constraints do not require excessive computational effort.[2,10] A factor of three in computing time can typically be saved by the application of bond-length constraints in protein simulations.[2]

Several methods are available to apply distance constraints during a MD simulation based on equations of motion in Cartesian coordinates,[10–14] of which the SHAKE method[10] is the oldest, simplest and a very robust technique. In SHAKE, there is a limit to the maximal displacement, induced by the unconstrained forces, allowed for the atoms involved in each individual constraint. This local convergence criterion means that SHAKE will fail to converge when the forces acting on the specific atoms in a given constraint become very large. Thus, a SHAKE failure can be used to detect an error in the simulation, specifically the presence and location of unphysically large forces.

Multiple-time-step algorithms come in different flavors.[3,15–20] In a molecular system, three main frequency ranges can be distinguished: high-frequency (*hf*) bond-stretching forces $\vec{f}^{hf}$, low-frequency (*lf*) long-ranged (e.g., Coulomb) nonbonded forces $\vec{f}^{lf}$, and the remaining intermediate frequency (*if*) forces $\vec{f}^{if}$. The contributions of the different forces to the atomic trajectories may be integrated using different time steps. An example is the twin-range method,[17,18] in which the longer-ranged nonbonded force is kept constant during n'' time steps Δt, where n'' lies in the range 5–10 for Δt = 1 or 2 fs (typical update frequency 10–20 fs). This MTS algorithm is denoted the constant-force multiple-time-step algorithm (cf-*MTS*). When applying an MTS algorithm to the bond-stretching forces,[3,19] the high-frequency covalent bond-stretching forces $\vec{f}^{hf}$ are evaluated at each time point $t + n\Delta t$, with n = 0, 1, 2, ... When n is a multiple of n_{hf} (n_{hf} is odd, typically 3 or 5), the other (intermediate and low frequency) forces $\vec{f}^{if} + \vec{f}^{lf}$ are applied, but multiplied by a factor n_{hf} in order to compensate for their omission at the intermediate steps. This MTS algorithm is denoted the impulse-force multiple-time-step algorithm (*if-MTS*).

The difference between the constant force and the impulse force multiple-time-step algorithms can be illustrated by comparing their equations using the leap-frog algorithm to integrate the equations of motion.[21] In the leap-frog algorithm the particle velocities $\vec{v}_i$ are updated at time points shifted by half a time step Δt from the positions $\vec{r}_i$,

$$\vec{v}_i\left(t+\frac{1}{2}\Delta t\right) = \vec{v}_i\left(t-\frac{1}{2}\Delta t\right) + m_i^{-1}\vec{f}_i\left(\vec{r}^{\,N}(t)\right)\Delta t + O\left((\Delta t)^3\right) \quad i = 1,2,\ldots N, \quad (3)$$

$$\vec{r}_i(t+\Delta t) = \vec{r}_i(t) + \vec{v}_i\left(t+\frac{1}{2}\Delta t\right)\Delta t + O\left((\Delta t)^3\right) \quad i = 1,2,\ldots N. \quad (4)$$

If velocity-dependent properties are of interest, these can be obtained by interpolation from half time steps to full time steps.[22] The error induced by the discretization of the equations of motion based on a truncated Taylor expansion is of order $(\Delta t)^3$. Thus, the time step Δt should be chosen much smaller than the shortest oscillation period in the system. This depends on the masses of the particles, the curvature of the potential-energy surface (force field), whether geometric constraints, for example, for bond lengths or angles, are applied and the temperature of the simulation.[2,23]

Using the leap-frog time-integration algorithm, an MTS scheme can be derived by explicitly performing n_{hf} steps Δt, from $t - (½n_{hf})\Delta t$ till $t + (½n_{hf})\Delta t$ for the velocities, see Figure 1,

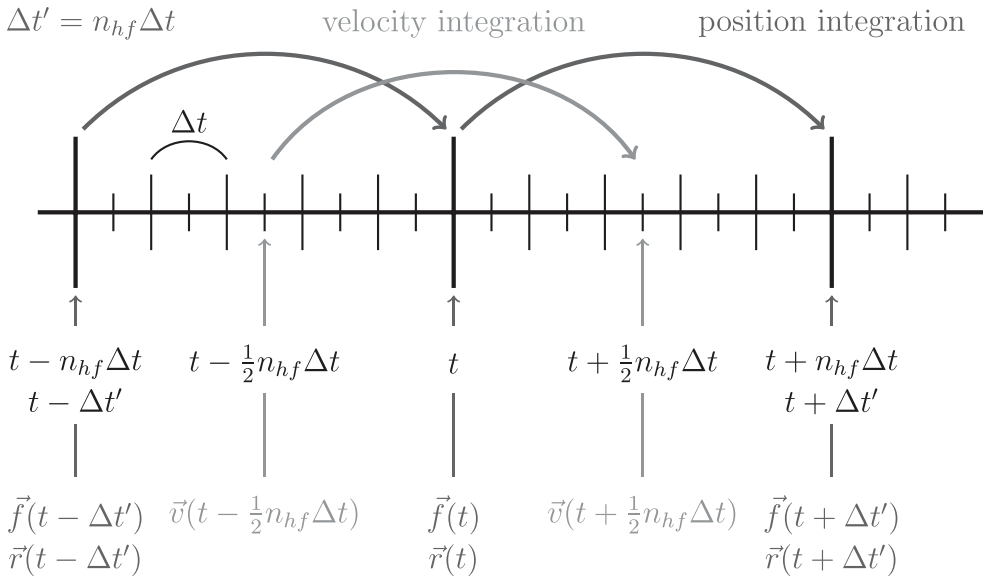

FIGURE 1 Leap-frog multiple-time-step integration scheme for classical equations of motion

$$\vec{v}_i\left(t+\left(\frac{1}{2}n_{hf}\right)\Delta t\right) = \vec{v}_i\left(t-\left(\frac{1}{2}n_{hf}\right)\Delta t\right) + m_i^{-1}\left\{\sum_{n'=1}^{(n_{hf}-1)/2}\vec{f}_i\left(t-n'\Delta t\right) + \vec{f}_i(t) + \sum_{n'=1}^{(n_{hf}-1)/2}\vec{f}_i\left(t+n'\Delta t\right)\right\}\Delta t, \tag{5}$$

and n_{hf} steps Δt, from t till $t + n_{hf}\Delta t$ for the positions,

$$\vec{r}_i(t+n_{hf}\Delta t) = \vec{r}_i(t) + \sum_{n''=1}^{n_{hf}}\vec{v}_i(t+(n''-1/2)\Delta t)\Delta t. \tag{6}$$

Using for $1 \le n'' \le n_{hf}$ the velocity leap-frog formula

$$\vec{v}_i\left(t+\left(n''-1/2\right)\Delta t\right) = \vec{v}_i\left(t-\left(\frac{1}{2}n_{hf}\right)\Delta t\right) + m_i^{-1}\left\{\sum_{n'=1}^{(n_{hf}-1)/2}\vec{f}_i\left(t-n'\Delta t\right) + \vec{f}_i(t) + \sum_{n'=1}^{(n''-1)}\vec{f}_i\left(t+n'\Delta t\right)\right\}\Delta t, \tag{7}$$

one finds

$$\vec{r}_i(t+n_{hf}\Delta t) = \vec{r}_i(t) + \sum_{n''=1}^{n_{hf}}\vec{v}_i\left(t-\left(\frac{1}{2}n_{hf}\right)\Delta t\right)\Delta t + m_i^{-1}\left\{\sum_{n''=1}^{n_{hf}}\sum_{n'=1}^{(n_{hf}-1)/2}\vec{f}_i(t-n'\Delta t)(\Delta t)^2 + \sum_{n''=1}^{n_{hf}}\vec{f}_i(t)(\Delta t)^2 + \sum_{n''=1}^{n_{hf}}\sum_{n'=1}^{(n''-1)}\vec{f}_i(t+n'\Delta t)(\Delta t)^2\right\}. \tag{8}$$

or

$$\vec{r}_i(t+n_{hf}\Delta t) = \vec{r}_i(t) + \vec{v}_i\left(t-\left(\frac{1}{2}n_{hf}\right)\Delta t\right)(n_{hf}\Delta t) + m_i^{-1}\left\{\sum_{n'=1}^{(n_{hf}-1)/2}\vec{f}_i(t-n'\Delta t)(n_{hf}\Delta t) + \vec{f}_i(t)(n_{hf}\Delta t) + \frac{1}{n_{hf}}\sum_{n'=1}^{(n_{hf}-1)}(n_{hf}-n')\vec{f}_i(t+n'\Delta t)(n_{hf}\Delta t)\right\}\Delta t. \tag{9}$$

Based on Equations (5) and (9) the force can be split into a high-frequency force component, which is updated every n'th time step Δt, and the intermediate to low-frequency force component, which is updated every n_{hf}th time step $\Delta t' = n_{hf}\Delta t$, and is treated differently between the constant-force and the impulse-force MTS algorithms. These algorithms can be obtained from Equations (5) and (9) as follows.

By inserting a low-frequency force that is constant from $t - ([n_{hf} - 1]/2)\Delta t$ till $t - \Delta t$ (i.e., $[n_{hf} - 1]/2$ terms) and from t till $t + (n_{hf} - 1)\Delta t$ (i.e., n_{hf} terms),

$$\vec{f}_i(t-((n_{hf}-1)/2)\Delta t) = \vec{f}_i(t-((n_{hf}-3)/2)\Delta t) = \ldots = \vec{f}_i(t-\Delta t) = \vec{f}_i(t-n_{hf}\Delta t), \tag{10}$$

and

$$\vec{f}_i(t) = \vec{f}_i(t+\Delta t) = \vec{f}_i(t+2\Delta t) = \ldots = \vec{f}_i(t+(n_{hf}-1)\Delta t), \tag{11}$$

we find for the constant-force algorithm

$$\vec{v}_i\left(t+\left(\frac{1}{2}n_{hf}\right)\Delta t\right) = \vec{v}_i\left(t-\left(\frac{1}{2}n_{hf}\right)\Delta t\right) + m_i^{-1}\left\{\frac{1}{2}\left(\left(1-\frac{1}{n_{hf}}\right)\vec{f}_i(t-n_{hf}\Delta t) + \left(1+\frac{1}{n_{hf}}\right)\vec{f}_i(t)\right)\right\}(n_{hf}\Delta t), \tag{12}$$

and

$$\vec{r}_i(t+n_{hf}\Delta t)=\vec{r}_i(t)+\vec{v}_i\left(t-\left(\frac{1}{2}n_{hf}\right)\Delta t\right)(n_{hf}\Delta t) + m_i^{-1}\left\{\frac{1}{2}\left(\left(1-\frac{1}{n_{hf}}\right)\vec{f}_i(t-n_{hf}\Delta t)+\left(1+\frac{1}{n_{hf}}\right)\vec{f}_i(t)\right)\right\}(n_{hf}\Delta t)^2. \quad (13)$$

If the force changes slowly between $t - n_{hf}\Delta t = t - \Delta t'$ and t, approximating $\vec{f}_i(t-\Delta t')$ by $\vec{f}_i(t)$ in Equations (12) and (13), one recovers the leap-frog equations (3) and (4) because $n_{hf}\Delta t = \Delta t'$. So, in the limit of slowly (per period $n_{hf}\Delta t = \Delta t'$) varying forces, this MTS algorithm is time-invariant and symplectic.

By using an impulse force, that is, inserting at time points $t - n_{hf}\Delta t$ and t, the force at those time points multiplied by n_{hf} and setting the force to zero at the other time points, Equations (5) and (9) become,

$$\vec{v}_i\left(t+\left(\frac{1}{2}n_{hf}\right)\Delta t\right)=\vec{v}_i\left(t-\left(\frac{1}{2}n_{hf}\right)\Delta t\right)+m_i^{-1}\vec{f}_i(t)(n_{hf}\Delta t) \quad (14)$$

and

$$\vec{r}_i(t+n_{hf}\Delta t)=\vec{r}_i(t)+\vec{v}_i\left(t-\left(\frac{1}{2}n_{hf}\right)\Delta t\right)(n_{hf}\Delta t)+m_i^{-1}\vec{f}_i(t)(n_{hf}\Delta t)^2, \quad (15)$$

which are the leap-frog equations (3) and (4), because $n_{hf}\Delta t = \Delta t'$. So, this MTS algorithm is time-invariant and symplectic.

The constant-force MTS algorithm is expected to perform well when the long-ranged nonbonded forces are the slowly changing, lower-frequency ones.[24–27] In this case, it is assumed that the interactions between atoms in close spatial proximity vary more rapidly than those at a distance. The application to different nonbonded force distance ranges is rather stable and relatively insensitive to the time between updates of the long-ranged forces.[20] The accuracy of the constant-force MTS algorithm is likely to suffer less from noncompliance of a molecular system with the condition of weak coupling between the motions along the higher-frequency degrees of freedom and the motions along the lower-frequency degrees of freedom than the impulse-force MTS algorithm, because of the nonphysical multiplication of the lower-frequency forces by a factor n_{hf} in the latter algorithm. The configurations sampled between impulses can be highly nonphysical, and the impulse force can induce resonances between the high frequency and the other degrees of freedom within the system leading to inappropriate configurations being sampled.[8,28–32] It was found,[32] that an impulse-force MTS algorithm should not be used to independently integrate nonbonded interactions that belong to different distance ranges, so-called distance classes, when persisting opposing forces exist between the separated components of the system. The net force on a conservative system is zero and therefore some degree of opposing force between separated components is unavoidable. In contrast, the impulse-force MTS algorithm may perform well when bond-stretching degrees of freedom are the fast oscillating, higher-frequency ones.

In the present article, it is investigated whether the application of an MTS algorithm, a constant-force or an impulse-force one, to the bond-stretching degrees of freedom of a protein in aqueous solution in a molecular dynamics simulation is as efficient as the application of bond-length constraints, thereby keeping the distortive effects upon the motion of the nonconstrained degrees of freedom small. The small protein bovine pancreatic trypsin inhibitor (BPTI) and water were chosen as test molecules.

For the protein BPTI, in vacuo or solvated in simple-point-charge (SPC) water, four different types of simulation were performed:

1. Bond-stretching forces are integrated using the same time step as for all other forces (*no-MTS*).
2. Bond-stretching forces of all bonds are integrated with a shorter time step than all other forces using an impulse-force MTS algorithm (*if-MTS*).
3. Bond-stretching forces of all bonds are integrated with a shorter time step than all other forces using a constant-force MTS algorithm (cf-*MTS*).
4. Bond lengths of all bonds are constrained when integrating the equations of motion (*c-no-MTS*).

For a single water molecule in vacuo and for liquid water, 1000 molecules in a periodic box, two different types of simulation were performed:

1. The bond-stretching and bond-angle bending forces of a flexible water model, SPC/F,[7] are integrated forward in time (*no-MTS*).
2. The two bond lengths and the bond angle of each water molecule are constrained, using the SPC water model,[33] when integrating the equations of motion (*c-no-MTS*).

For liquid water, 1000 molecules in a periodic box, two additional types of simulations were performed:

3. Bond-stretching and bond-angle bending forces of a flexible water model, SPC/F, are integrated with a shorter time step than all other forces using an impulse-force MTS algorithm (*if-MTS*).
4. Bond-stretching and bond-angle bending forces of a flexible water model, SPC/F, are integrated with a shorter time step than all other forces using a constant-force MTS algorithm (cf-*MTS*).

Five systems were simulated:

1. BPTI in vacuo (*protein_vac*), in order to test the degree of conservation of total energy, linear and rotational momentum[2] of the protein as function of the type of simulation (unconstrained vs. MTS vs. constrained), the size of the time step Δt and the number of time steps n_{hf} between the calculation of the intermediate and low frequency forces $\bar{\bar{f}}^{if} + \bar{\bar{f}}^{lf}$. The vacuum boundary condition was chosen, instead of the commonly used periodic boundary condition, in order to be able to use a very large (100 nm), de facto infinite, nonbonded interaction cut-off distance. This eliminates the nonbonded interaction cut-off error, because all atom pairs are used in the force calculation. The error induced by the application of bond-length constraints was made very small by requiring a relative geometric precision of $tol_{DC} = 10^{-6}$ when solving the

constraint equations. In order to monitor energy conservation, the system was not coupled to heat or pressure baths and the motions of the center of mass and around the center of mass were not removed during the simulation.

2. A single water molecule in vacuo (*water_vac*), using a flexible (SPC/F) and a rigid (SPC) water model, in order to test the degree of conservation of total energy, linear and rotational momentum as function of the type of simulation (unconstrained vs. constrained) and the size of the time step Δt. The error induced by the application of constraints to maintain the rigid geometry of the SPC water model was made very small by requiring a relative geometric precision of $tol_{DC} = 10^{-6}$ when solving the constraint equations. In order to monitor energy conservation, the system was not coupled to heat or pressure baths and the motions of the center of mass and around the center of mass were not removed during the simulation.
3. Liquid water (*water_liq_ener*), 1000 molecules in a cubic periodic box with minimum image periodic boundary conditions using a flexible (SPC/F) and a rigid (SPC) water model, in order to test the conservation of the total energy in simulations of liquid water as function of the type of simulation (unconstrained vs. MTS vs. constrained), the size of the time step Δt and the number of time steps n_{hf} between the calculation of the (nonbonded interaction) intermediate and low frequency forces $\vec{f}^{\,if} + \vec{f}^{\,lf}$. The computational effort of evaluating the nonbonded interactions is often reduced by the application of a constant-force multiple-time-step scheme. In the present simulations of liquid water this was not done in order to reduce the error induced by the lower frequency of updating part of the nonbonded interactions. The error induced by the application of constraints to maintain the rigid geometry of the SPC water model was made very small by requiring a relative geometric precision of $tol_{DC} = 10^{-6}$ when solving the constraint equations. In order to monitor energy conservation, the system was not coupled to heat or pressure baths and the motion of the center of mass of the system was not removed during the simulation.
4. Liquid water (*water_liq_prop*), 1000 molecules in a cubic periodic box with minimum image periodic boundary conditions using a flexible (SPC/F) and a rigid (SPC) water model, in order to test the properties of liquid water as function of the type of simulation (unconstrained vs. MTS vs. constrained), the size of the time step Δt and the number of time steps n_{hf} between the calculation of the (nonbonded interaction) intermediate and low frequency forces $\vec{f}^{\,if} + \vec{f}^{\,lf}$. The nonbonded interactions were handled as in the *water_liq_ener* simulations. The error induced by the application of constraints to maintain the rigid geometry of the SPC water model was made very small by requiring a relative geometric precision of $tol_{DC} = 10^{-6}$ when solving the constraint equations. To allow for longer simulations to evaluate the properties of the liquid, the system was coupled to heat and pressure baths and the motion of the center of mass of the system was removed during the simulation.
5. BPTI in aqueous solution (*protein_aq*), that is, in a rectangular periodic box with minimum image periodic boundary conditions, with many (6291) explicit rigid water molecules, the standard way of protein simulation, in order to test the behavior of various protein properties as function of the type of simulation (unconstrained vs. MTS vs. constrained protein), the size of the time step Δt and the number of time steps n_{hf} between the calculation of the intermediate and low frequency forces $\vec{f}^{\,if} + \vec{f}^{\,lf}$. No constant-force multiple-time-step scheme was used when evaluating the nonbonded interactions. A geometric precision of $tol_{DC} = 10^{-4}$ was used when solving the constraint equations, as in standard MD simulations. The protein and water were separately coupled to a heat bath and the system was coupled to a pressure bath. The motion of the center of mass of the system was removed during the simulation.

2 | METHOD

2.1 | Potential energy function or force field, molecular model

When solvated in water (*protein_aq*), the protein was modeled using the GROMOS bio-molecular force field 54A7.[34–36] The protein contains 114 (nonaliphatic) hydrogen atoms and 490 nonhydrogen atoms. When simulating the protein in vacuo (*protein_vac*), the GROMOS bio-molecular force field 54B7[37,38] was used. The A-version of a GROMOS force field is the basic force field designed for molecules in solution or in crystalline form. The B-version is derived from the A-version in order to be used for simulating molecules in vacuo, where the dielectric screening effect of the environment is neglected. The atomic charges and van der Waals parameters are changed such that atom charge groups with a nonzero total charge are neutralized while maintaining the hydrogen-bonding capacity of the individual atoms. The 6291 solvent molecules solvating the protein in the rectangular periodic box (*protein_aq*) were modeled using the rigid simple point charge (SPC) model.[33]

When applying an MTS algorithm to the bond-stretching degrees of freedom, the bond-stretching forces are to be computed 3–5 (n_{hf}) times as frequently as the other forces. The evaluation of the (harmonic) bond energy term, commonly used in bio-molecular force fields,

$$(1/2)K^{b,h}\left(b(t)-b^0\right)^2, \tag{16}$$

where $K^{b,h}$ is the force constant, $b(t)$ the actual bond length at time t, that is, the distance $r_{ij}(t)$ between the atoms i and j forming the bond, and b^0 is the ideal bond length, and the evaluation of its derivative.

$$K^{b,h}\left(r_{ij}(t)-b^0\right)\vec{r}_{ij}(t)/r_{ij}(t), \tag{17}$$

require the evaluation of the square root $r_{ij} \equiv \left|\vec{r}_{ij}\cdot\vec{r}_{ij}\right|^{1/2} \equiv \left(x_{ij}^2+y_{ij}^2+z_{ij}^2\right)^{1/2}$, with $\vec{r}_{ij} \equiv \vec{r}_i - \vec{r}_j$. A physically rather insignificant change of the harmonic (i.e., quadratic) bond-stretching interaction term (16) to

$$(1/2)K^{b,q}\left((b(t))^2-\left(b^0\right)^2\right)^2/\left(4\left(b^0\right)^2\right) = (1/2)K^{b,q}\left(r_{ij}(t)-b^0\right)^2\left(r_{ij}(t)+b^0\right)^2/\left(2b^0\right)^2, \quad (18)$$

a quartic bond-stretching interaction term,[6,37,38] avoids the evaluation of the square root, both in the energy term (18) as well as in its derivative

$$K^{b,q}\left((r_{ij}(t))^2-\left(b^0\right)^2\right)\vec{r}_{ij}(t)/\left(2b^0\right)^2, \quad (19)$$

which saves computing effort in the multiple-time-step algorithm.

In the simulations of liquid water (*water_liq*) the flexible SPC model, SPC/F,[7] was used as well as the rigid SPC one.[33]

2.2 | Constraints

In the GROMOS simulation software,[39] the use of bond-length constraints is simple. The constraints for the protein correspond exactly with the bond-stretching force-field terms, and thus the constraints (ideal lengths of the bonds) are taken from these force-field terms. When applying bond-length constraints, the corresponding bond-stretching term of the force field is not evaluated. The bond lengths in the protein and water molecules and the bond angle of the SPC water molecules were kept rigid using the SHAKE algorithm[10] with a relative geometric precision of $tol_{DC} = 10^{-4}$ in case of the simulations of the solvated protein (*protein_aq*) or 10^{-6} in case of the simulations of the protein in vacuo (*protein_vac*) and of the simulations of water (*water_vac*, *water_liq_ener*, *water_liq_prop*).

2.3 | Treatment of long-ranged interactions

When simulating the protein in vacuo (*protein_vac*) using MD, the nonbonded interactions are calculated for all atom pairs in the protein. When simulating liquid water (*water_liq_ener*, *water_liq_prop*) or the protein solvated in a periodic box filled with water molecules, the nonbonded interactions were calculated using a single cut-off radius of 1.4 nm. Outside the cut-off radius a reaction-field approximation[40,41] with a relative dielectric permittivity $\varepsilon_{RF} = 78.5$ and a ionic strength of zero ($\kappa_{RF} = 0$) was used, The relative dielectric permittivity in the cut-off sphere $\varepsilon_{cs} = 1$.[42] Minimum-image periodic boundary conditions were applied.

2.4 | Simulation set-up, heat-up, and equilibration

The protein was simulated using the GROMOS bio-molecular simulation software.[39,43] The leap-frog algorithm[21] was used to integrate Newton's equations of motion.

The initial structure of the protein BPTI, including four internally hydrogen bonded water molecules in case of the solvated protein (*protein_aq*), was taken from the Protein Data Bank (PDB),[44] entry 1*bpi*.

The protein initial structure was first energy minimized in vacuo to release possible strain induced by small differences in bond lengths, bond angles, improper dihedral angles, and short nonbonded contacts between the force-field parameters and the X-ray structure. The resulting protein configuration was used as initial configuration for the MD simulations in vacuo (*protein_vac*).

The initial atomic velocities of the protein in vacuo (*protein_vac*) were sampled from a Maxwell distribution at $T = 60$ K and the translation of and the rotation around the center of mass of the system were removed. The equilibration scheme consisted of five consecutive, short 20 ps simulations at temperatures 60, 120, 180, 240, and 300 K. The temperature was kept constant using the weak-coupling algorithm[45] with a relaxation or coupling time $\tau_T = 0.1$ ps and the translation of and the rotation around the center of mass of the system were removed every 2 ps. At 300 K, the equilibration was extended for 2 ns. After the equilibration, the coupling to the heat bath was removed and translation of and rotation around the center of mass of the system (if present) were not removed anymore in order to avoid perturbation of energy conservation.

The initial atomic velocities of the single water molecule in vacuo (*water_vac*) were sampled from a Maxwell distribution at $T = 298.15$ K and the translation of and the rotation around the center of mass of the system were *not* removed. The equilibration scheme consisted of a 20 ps simulation at a temperature of 298.15 K. The temperature was kept constant using the weak-coupling algorithm[45] with a relaxation or coupling time $\tau_T = 0.1$ ps. The translational and rotational plus internal (the latter only present for the SPC/F model) kinetic energy were separately coupled to the heat bath. After the equilibration, the couplings to the heat bath were removed and translation of and rotation around the center of mass of the molecule were again *not* removed in order to avoid perturbation of energy conservation.

The simulations of liquid water (*water_liq_ener*, *water_liq_prop*) were started from an equilibrated cubic periodic box filled with 1000 water molecules. In the *water_liq_prop* simulations using the flexible SPC/F water model, the translational and rotational plus internal degrees of freedom of the molecules were separately coupled to the heat bath.

The MD simulation of the protein solvated in a periodic box with explicit water molecules (*protein_aq*) required the addition of water molecules. The protein, including the four internal water molecules, was put into a rectangular box filled with water molecules, such that the minimum solute-wall distance was 1.0 nm, and water molecules closer than 0.23 nm from the solute were removed. This resulted in a box with 6291 water molecules for the initial protein structures. In order to relax unfavorable contacts between atoms of the solute and the solvent, a second energy minimization was performed for the protein in the periodic box with water while keeping the atoms of the solute harmonically position-restrained[38] with a force constant of 25,000 kJ mol^{-1} nm^{-2}. The resulting protein-water configuration was

TABLE 1 Energy conservation in 100 ps MD simulations of the protein BPTI in vacuo using no MTS algorithm, the *if-MTS* algorithm or bond-length constraints for the bond-stretching degrees of freedom of all bonds, without coupling to temperature and pressure baths, and without removal of translation of and rotation around the center of mass, as function of the MD time step Δt and the number of time steps n_{hf} at which only the bond-stretching forces are evaluated and integrated

MTS or constraints	n_{hf}	Δt (fs)	E_{bond}	ΔE_{bond}	E_{tot}	ΔE_{tot}	E_{tot}^{drift}	ΔE_{tot}^{drift}	E_{kin}	ΔE_{kin}	E_{kin}^{drift}	ΔE_{kin}^{drift}	T (K)	$\Delta E_{tot}/\Delta E_{kin}$	$\Delta E_{tot}^{drift}/\Delta E_{kin}^{drift}$
-	-	0.1	818	43	14	0.033	0.00008	0.033	2252	54	0.13	54	299.9	0.00061	0.00061
		0.2	826	42	4	0.128	−0.00007	0.128	2257	54	0.20	54	300.6	0.0024	0.0024
		0.5	819	42	58	0.806	−0.001	0.805	2261	53	0.07	53	301.1	0.015	0.015
		1.0	828	43	72	3.210	−0.002	3.209	2272	54	0.23	53	302.7	0.059	0.061
		2.0	846	44	15	15.104	−0.108	14.778	2254	48	0.02	48	300.3	0.31	0.31
if	3	0.1	820	42	18	2.321	0.004	2.318	2244	55	0.41	54	298.9	0.042	0.043
		0.2	805	42	23	4.596	0.006	4.593	2247	54	−0.03	54	299.3	0.085	0.085
		0.5	818	42	48	11.803	−0.012	11.798	2264	51	0.05	50	301.5	0.23	0.24
		1.0	813	43	35	22.627	0.033	22.606	2264	47	0.12	47	301.6	0.48	0.48
		2.0	-	-	-	-	-	-	-	-	-	-	-	-	-
if	5	0.1	808	42	0	4.510	0.0003	4.510	2246	54	−0.12	54	299.2	0.084	0.084
		0.2	811	42	51	8.818	0.003	8.818	2269	52	0.09	52	302.2	0.17	0.17
		0.5	805	41	26	19.862	0.00008	19.862	2246	49	−0.29	48	299.2	0.41	0.42
		1.0	-	-	-	-	-	-	-	-	-	-	-	-	-
		2.0	-	-	-	-	-	-	-	-	-	-	-	-	-
constraints	-	0.1	0	0	−1430	0.103	0.004	0.012	1507	44	0.11	44	305.1	0.0023	0.00027
		0.2	0	0	−1413	0.323	−0.011	0.029	1501	44	0.13	44	304.0	0.0073	0.00068
		0.5	0	0	−1489	0.744	−0.025	0.151	1459	43	0.02	43	295.3	0.017	0.0035
		1.0	0	0	−1424	1.410	−0.043	0.660	1509	44	0.05	44	305.5	0.032	0.015
		2.0	0	0	−1451	3.143	−0.071	2.386	1479	43	−0.04	42	299.5	0.073	0.057
		5.0	0	0	-	-	-	-	-	-	-	-	-	-	-
		10.0	0	0	-	-	-	-	-	-	-	-	-	-	-

Note: The bond-length constraints are imposed with a relative geometric precision of $tol_{DC} = 10^{-6}$. Nonbonded interaction cut-off $R_{cp} = R_{cl}$: 100 nm (infinity). E_{bond}: Bond-stretching energy. ΔE_{bond}: Fluctuation of E_{bond}. E_{tot}: Total energy. ΔE_{tot}: Fluctuation of E_{tot}. E_{tot}^{drift}: Total energy drift. ΔE_{tot}^{drift}: Fluctuation around E_{tot}^{drift}. E_{kin}: Kinetic energy. ΔE_{kin}: Fluctuation of E_{kin}. E_{kin}^{drift}: Kinetic energy drift. ΔE_{kin}^{drift}: Fluctuation around E_{kin}^{drift}. T: Temperature. All values are averages calculated from the same number of trajectory structures separated by approximately 0.02 ps. Energies in kJ mol^{-1}. Energy drifts in kJ mol^{-1} ps^{-1}.

used as initial configuration for the MD simulation of the protein solvated in explicit water.

In order to avoid artificial deformations in the protein structure due to relatively high-energy atomic interactions still present in the system, the MD simulations of the protein were started at T = 60 K and then the temperature was slowly raised to T = 300 K. Initial atomic velocities were sampled from a Maxwell distribution at T = 60 K and the translation of and the rotation around the center of mass of the system were removed. The equilibration scheme consisted of five consecutive, short 20 ps simulations at temperatures 60, 120, 180, 240, and 300 K, at constant volume in the case of the periodic system. During the first four of the equilibration periods, the solute atoms were harmonically restrained to their positions in the initial structures with force constants of 25,000, 2500, 250, and 25 kJ mol^{-1} nm^{-2}. The temperature was kept constant using the weak-coupling algorithm[45] with a relaxation or coupling time τ_T = 0.1 ps. Solute and solvent were separately coupled to the heat bath. Following this equilibration procedure, the simulations were performed at a reference temperature of 300 K and a reference pressure of 1 atm (*protein_aq*). The pressure was kept constant using the weak-coupling algorithm[45] with a coupling time τ_p = 0.5 ps and an isothermal compressibility κ_T = 4.575 10^{-4} (kJ mol^{-1} nm^{-3})$^{-1}$. The center of mass motion of the system was removed after equilibration of the system and then, for the simulations of liquid water (*water_liq_prop*) and of the protein in explicit water (*protein_aq*), every 0.1 ps.

The required equilibration time before analysis of the trajectories depends mainly on the coupling between and the range of the interactions, the systems size, and the initial configuration and velocities. We note, however, that any change in system composition, set of constraints, boundary condition or size of the time step is to be followed by some further equilibration, here 20 ps, in order to let the system adapt to the changed circumstances before analyzing ensemble averages or time series and correlation functions.

2.5 | MD simulation of a protein in vacuo to test conservation properties

When integrating Newton's equations of motion forward in time t, the total energy $E_{tot}(t)$, the total translational kinetic energy $E_{kin,trans}(t)$, the total translational momentum $\boldsymbol{P}_{trans}(t)$, and, in case of a single molecule in vacuo, the total rotational kinetic energy $E_{kin,rot}(t)$ and the total rotational momentum $\boldsymbol{L}_{rot}(t)$ must be conserved. The extent of conservation will be determined by the numerical integration algorithm (the leap-frog algorithm), the precision of the algorithm to impose the constraints (tol_{DC}), the size of the MD time steps Δt and $\Delta t' = n_{hf}\Delta t$. This is investigated using MD simulation in vacuo. This boundary condition was chosen in order to be able to use a very large ($R_{cl} = R_{cp}$ = 100 nm), de facto infinite, nonbonded interaction cut-off distance. This eliminates the nonbonded cut-off error because all atom pairs are used in the force calculation. The error induced by the application of bond-length constraints was made very small by requiring a precision $tol_{DC} = 10^{-6}$. To test the conservation of total energy and momenta, which should apply at every MD time step, long simulations are not required. MD simulations of 100 ps were performed. Configurations were saved every 0.02 ps, approximately due to

TABLE 2 Energy conservation in 20 MD simulations of 100 ps of a single water molecule in vacuo using the flexible SPC/F model or the rigid SPC model, without coupling to temperature and pressure baths, and without removal of translation of and rotation around the center of mass, as function of the MD time step Δt

Flexible or constraints	Δt (fs)	E_{int}	ΔE_{int}	E_{tot}	ΔE_{tot}	E_{tot}^{drift} ($\times 10^{-3}$)	ΔE_{tot}^{drift} ($\times 10^{-3}$)	E_{kin}	ΔE_{kin}	E_{kin}^{drift} ($\times 10^{-3}$)	ΔE_{kin}^{drift} ($*10^{-3}$)	T (K)	SR	$\Delta E_{tot}/\Delta E_{kin}$
Flexible	0.1	0.04	0.014	11.19	0.00002	0	0.02	11.16	0.014	0.0002	14	298.2	1.0	0.0014
	0.2	0.04	0.014	11.20	0.00006	0	0.06	11.15	0.014	0.0002	14	298.1	1.0	0.0043
	0.5	0.04	0.014	11.20	0.0004	0	0.4	11.15	0.014	0.0004	14	298.1	1.0	0.029
	1.0	0.04	0.016	11.20	0.002	0	2	11.15	0.015	0.0002	15	298.1	1.0	0.13
	2.0	0.14	0.080	11.29	0.067	0.007	67	11.16	0.023	0.0002	23	298.2	0.6	2.9
Constraints	0.1	0	0	7.44	0.00002	0.0006	0.004	7.44	0.00002	0.0006	0.004	298.2	1.0	1
	0.2	0	0	7.44	0.00002	0.0005	0.004	7.44	0.00002	0.0005	0.004	298.1	1.0	1
	0.5	0	0	7.44	0.00003	0.001	0.004	7.44	0.00003	0.001	0.004	298.1	1.0	1
	1.0	0	0	7.44	0.00006	0.002	0.004	7.44	0.00006	0.002	0.004	298.1	1.0	1
	2.0	0	0	7.44	0.00009	0.003	0.005	7.44	0.00009	0.003	0.005	298.1	1.0	1
	5.0	0	0	7.44	0.0002	0.008	0.005	7.44	0.0002	0.008	0.005	298.1	1.0	1
	10.0	0	0	7.44	0.0005	0.02	0.006	7.44	0.0005	0.02	0.006	298.1	1.0	1

Note: Relative geometric precision of the distance constraints (SPC): $tol_{DC} = 10^{-6}$. E_{int}: Bond-stretching and bond-angle bending energy. ΔE_{int}: Fluctuation of E_{int}. E_{tot}: Total energy. ΔE_{tot}: Fluctuation of E_{tot}. E_{tot}^{drift}: Total energy drift. ΔE_{tot}^{drift}: Fluctuation around E_{tot}^{drift}. E_{kin}: Kinetic energy. ΔE_{kin}: Fluctuation of E_{kin}. E_{kin}^{drift}: Kinetic energy drift. ΔE_{kin}^{drift}: Fluctuation around E_{kin}^{drift}. T: Temperature. SR: Success ratio, fraction of successful simulations. All values are averages over the 20 independent simulations, calculated from the same number of trajectory structures of each simulation separated by 0.02 *ps*. Energies in kJ mol^{-1}. Energy drifts in kJ mol^{-1} ps^{-1}.

the different sizes of the time steps used. No coupling of the temperature to a heat bath was present and center of mass translation and molecular rotation were not removed during the simulation. Time steps Δt of 0.1, 0.2, 0.5, 1.0 and 2.0 fs were tested using n_{hf} values of 1, 3 and 5, so for $\Delta t'$ values of 0.3, 0.6, 1.5, 3.0 and 6.0 fs in case n_{hf} = 3, and for $\Delta t'$ values of 0.5, 1.0, 2.5, 5.0 and 10.0 fs in case n_{hf} = 5. In the simulations using constraints, longer time steps Δt = 5.0 and 10.0 fs were also investigated. In total 32 simulations (*protein_vac*) were performed. The simulations applying the cf-*MTS* algorithm were not stable, so their results were omitted from Table 1. As expected, also in some other cases the simulations were not stable when using longer time steps (Table 1).

2.6 | MD simulation of a single water molecule in vacuo to test conservation properties

For a single water molecule in vacuo, there are only intramolecular bond-stretching and bond-angle bending forces when using the flexible SPC/F model and no forces at all using the rigid SPC model, apart from inertial forces. MD simulations for these two water models were thus only performed without and with constraints, respectively, using different time steps (no MTS algorithms), in total 12 simulations (*water_vac*). In this case, configurations were saved about every 0.02 ps and the results were averaged over 20 MD simulations of 100 ps each (Table 2).

2.7 | MD simulations of liquid water under periodic boundary conditions

MD simulations of liquid water are commonly performed using periodic spatial boundary conditions and keeping the temperature and pressure constant. The algorithm and parameters of the temperature and pressure coupling were mentioned above. However, when testing total energy conservation, the system is not coupled to a heat bath and a pressure bath, and the center of mass motion of the system is

TABLE 3 Energy conservation in 20 ps MD simulations of liquid water, 1000 molecules in a cubic periodic box using the minimum-image convention to calculate the forces, without coupling to temperature and pressure baths, and without removal of center of mass translation, using no MTS algorithm or the *if-MTS* algorithm for the flexible SPC/F model or three distance constraints for the rigid SPC model, as function of the MD time step Δt and the number of time steps n_{hf} at which only the bond-stretching and bond-angle forces (SPC/F) are evaluated and integrated

MTS or constraints	n_{hf}	Δt (fs)	E_{int}	ΔE_{int}	E_{tot}	ΔE_{tot}	E_{tot}^{drift}	ΔE_{tot}^{drift}	E_{kin}	ΔE_{kin}	E_{kin}^{drift}	ΔE_{kin}^{drift}	T_{tr} (K)	T_{ir} (K)
-	-	0.1	5.50	0.134	−22.81	0.780	0.135	0.036	11.58	0.323	0.051	0.138	315.8	306.3
		0.2	5.52	0.121	−22.55	0.812	0.140	0.045	11.64	0.321	0.051	0.129	317.0	308.4
		0.5	5.59	0.136	−22.41	0.818	0.142	0.027	11.71	0.325	0.051	0.132	317.7	310.7
		1.0	5.45	0.130	−22.35	0.871	0.151	0.030	11.69	0.336	0.054	0.129	323.5	307.0
		2.0	-	-	-	-	-	-	-	-	-	-	-	-
if	3	0.1	5.56	0.133	−22.57	0.828	0.143	0.029	11.66	0.323	0.051	0.133	316.4	309.2
		0.2	5.47	0.131	−22.47	0.833	0.144	0.043	11.69	0.338	0.053	0.137	320.7	308.1
		0.5	5.17	0.128	−22.61	0.814	0.141	0.063	11.68	0.324	0.052	0.127	318.1	309.2
		1.0	4.04	0.122	−23.05	0.820	0.142	0.064	11.62	0.327	0.052	0.127	312.7	309.6
		2.0	-	-	-	-	-	-	-	-	-	-	-	-
if	5	0.1	5.48	0.154	−22.59	0.786	0.136	0.046	11.65	0.317	0.050	0.132	317.7	308.2
		0.2	5.40	0.133	−22.44	0.927	0.160	0.046	11.73	0.372	0.060	0.139	320.1	310.3
		0.5	4.37	0.111	−23.11	0.799	0.138	0.069	11.60	0.318	0.051	0.118	315.2	307.6
		1.0	11.52	1.369	−11.38	3.494	0.585	0.885	18.25	1.504	0.255	0.304	208.2	627.5
		2.0	-	-	-	-	-	-	-	-	-	-	-	-
constraints	-	0.1	0	0	−33.21	0.603	0.104	0.068	7.75	0.227	0.034	0.117	310.6	310.4
		0.2	0	0	−33.02	0.733	0.127	0.040	7.80	0.285	0.045	0.118	312.5	312.6
		0.5	0	0	−33.18	0.649	0.112	0.021	7.74	0.244	0.037	0.116	310.6	310.1
		1.0	0	0	−32.98	0.680	0.118	0.019	7.83	0.249	0.038	0.112	313.4	314.5
		2.0	0	0	−33.20	0.616	0.107	0.021	7.78	0.210	0.030	0.115	311.4	312.4
		5.0	0	0	−33.30	0.614	0.106	0.039	7.75	0.226	0.034	0.114	305.0	316.6
		10.0	-	-	-	-	-	-	-	-	-	-	-	-

Note: Relative geometric precision of the distance constraints (SPC): $tol_{DC} = 10^{-6}$. Nonbonded interaction cut-off radius $R_{cp} = R_{cl}$ = 1.4 nm. Outside a sphere of radius $R_{RF} = R_{cl}$, a homogeneous continuum dielectric with ε_{RF} = 78.5, κ_{RF} = 0 and ε_{cs} = 1 is assumed to be present. E_{int}: Bond-stretching and bond-angle bending energy. ΔE_{int}: Fluctuation of E_{int}. E_{tot}: Total energy. ΔE_{tot}: Fluctuation of E_{tot}. E_{tot}^{drift}: Total energy drift. ΔE_{tot}^{drift}: Fluctuation around E_{tot}^{drift}. E_{kin}: Kinetic energy. ΔE_{kin}: Fluctuation of E_{kin}. E_{kin}^{drift}: Kinetic energy drift. ΔE_{kin}^{drift}: Fluctuation around E_{kin}^{drift}. All energies given per molecule. T_{tr}: Temperature of the translational degrees of freedom. T_{ir}: Temperature of the internal and rotational degrees of freedom. All values are averages calculated from trajectory structures separated by approximately 0.02 *ps*. Energies in kJ mol^{-1}. Energy drifts in kJ mol^{-1} ps^{-1}.

not removed during the simulation (*water_liq_ener*). The nonbonded interactions were calculated using a single cut-off radius of 1.4 nm. Minimum-image periodic boundary conditions were applied. Configurations were saved about every 0.02 ps. These 32 simulations were performed for 20 ps (Table 3). The simulations applying the cf-*MTS* algorithm were not stable, so their results were omitted from Table 3.

In the simulations to evaluate the properties of liquid water (*water_liq_prop*) the nonbonded interactions were also calculated using a single cut-off radius of 1.4 nm. Outside the cut-off radius a reaction-field approximation[40,41] with a relative dielectric permittivity of 78.5 was used. Minimum-image periodic boundary conditions were applied. The temperature and pressure were kept constant using the weak-

TABLE 4 Selected properties of liquid water from 1 ns MD simulations of 1000 molecules in a cubic periodic box using the minimum-image convention to calculate the forces, with coupling to temperature (τ_T = 0.1 ps, T_{ref} = 298.15 K) and pressure (τ_p = 0.5 ps, p_{ref} = 1 atm) baths, using no MTS algorithm or two different (*if* and cf) MTS algorithms for the flexible SPC/F model or three distance constraints for the rigid SPC model, as function of the MD time step Δt and the number of time steps n_{hf} at which only the bond-stretching and bond-angle bending forces (SPC/F) are evaluated and integrated

Algorithm			Thermodynamic properties						Dynamic properties		
MTS or constraints	n_{hf}	Δt (fs)	T_{tr} (K)	T_{ir} (K)	ρ (kg m^{-3})	ΔH_{vap} (kJ mol^{-1})	E_{pot} (kJ mol^{-1})	E_{int} (kJ mol^{-1})	$<D>$ (10^{-3} nm^2 ps^{-1})	$<\tau_2^{OH}>$ (ps)	$<\tau^{H\text{-}bond}>$ (ps)
Experiment	-	-	298.2	298.2	997	44.0			2.3	1.95	
-	-	0.1	299.0	298.1	979.4	46.6	−35.12	5.56	5.4	1.3	0.53
		0.2	299.0	298.0	979.9	46.7	−35.12	5.56	5.6	1.4	0.53
		0.5	298.9	298.1	980.1	46.6	−35.13	5.56	5.5	1.3	0.53
		1.0	298.8	298.1	980.2	46.7	−35.12	5.56	5.4	1.3	0.53
		2.0	-	-	-	-	-	-	-	-	-
if	3	0.1	299.0	298.1	980.0	46.6	−35.13	5.54	5.6	1.4	0.53
		0.2	298.9	298.1	979.9	46.7	−35.14	5.50	5.4	1.4	0.53
		0.5	298.7	298.3	978.7	46.3	−35.21	5.19	5.6	1.4	0.53
		1.0	297.8	298.8	974.2	45.5	−35.48	3.97	5.8	1.2	0.52
		2.0	-	-	-	-	-	-	-	-	-
if	5	0.1	299.0	298.1	979.8	46.6	−35.14	5.51	5.6	1.4	0.53
		0.2	298.8	298.1	979.3	46.6	−35.19	5.36	5.7	1.2	0.53
		0.5	298.2	298.6	976.1	45.8	−35.50	4.33	5.7	1.2	0.53
		1.0	264.8	479.0	1004.9	47.4	−32.05	9.31	3.2	2.4	0.67
		2.0	-	-	-	-	-	-	-	-	-
cf	3	0.1	332.0	295.7	927.6	42.6	−32.60	4.02	8.1	0.7	0.44
		0.2	339.7	308.1	914.8	41.8	−31.75	4.05	9.4	0.7	0.43
		0.5	362.9	351.4	867.5	39.2	−28.88	4.34	10.5	1.9	0.39
		1.0	402.9	418.9	740.4	33.7	−23.03	4.62	15.0	3.6	0.32
		2.0	-	-	-	-	-	-	-	-	-
cf	5	0.1	339.8	308.0	914.4	41.7	−31.75	4.05	9.0	1.0	0.43
		0.2	355.0	336.2	885.8	40.2	−29.91	4.24	9.7	8.3	0.40
		0.5	403.6	418.2	737.2	33.5	−22.98	4.59	14.2	6.0	0.33
		1.0	321.0	318.6	54.6	8.6	−1.70	0.82	443.6	1.5	0.25
		2.0	-	-	-	-	-	-	-	-	-
constraints	-	0.1	298.2	298.7	972.0	43.8	−41.59	0	4.2	1.7	0.57
		0.2	298.3	298.5	972.0	43.8	−41.59	0	4.2	1.7	0.57
		0.5	298.3	298.4	971.4	43.8	−41.60	0	4.3	1.7	0.57
		1.0	298.0	298.8	971.2	43.8	−41.59	0	4.5	1.7	0.57
		2.0	297.5	299.2	972.0	43.9	−41.61	0	4.3	1.7	0.57
		5.0	292.5	304.5	973.7	44.0	−41.70	0	4.2	1.7	0.58
		10.0	-	-	-	-	-	-	-	-	-

Note: Removal of center of mass translation every 2 ps. Relative geometric precision of the constraints (SPC): $tol_{DC} = 10^{-6}$. Nonbonded interaction cut-off radius $R_{cp} = R_{cl}$ = 1.4 nm. Outside a sphere of radius $R_{RF} = R_{cl}$, a homogeneous continuum dielectric with ε_{RF} = 78.5, κ_{RF} = 0 and ε_{cs} = 1 is assumed to be present. T_{tr}: Temperature of the translational degrees of freedom. T_{ir}: Temperature of the internal and rotational degrees of freedom. ρ: Mass density. ΔH_{vap}: Heat of vaporization. E_{pot}: Potential energy per molecule. E_{int}: Bond-stretching and bond-angle bending energy per molecule. $<D>$: Average molecular diffusion coefficient. $<\tau_2^{OH}>$: Average molecular rotational correlation time of an O-H bond. $<\tau^{H\text{-}bond}>$: Average lifetime of water–water hydrogen bonds. All values are averages calculated from trajectory structures separated by approximately 0.1 *ps*.

coupling algorithm. In the simulations using the flexible SPC/F water model, the motions of the translational and rotational plus internal degrees of freedom of the molecules were separately coupled to the heat bath. These simulations were performed for 1 ns. Time steps Δt of 0.1, 0.2, 0.5, 1.0, and 2.0 fs were tested using n_{hf} values of 1, 3 and 5, so for $\Delta t'$ values of 0.3, 0.6, 1.5, 3.0, and 6.0 fs in case n_{hf} = 3, and for $\Delta t'$ values of 0.5, 1.0, 2.5, 5.0 and 10.0 fs in case n_{hf} = 5. In the simulations using constraints, longer time steps Δt = 5.0 and 10.0 fs were also used. When constraints were applied, $tol_{DC} = 10^{-6}$ was used. Translational motion of the center of mass of the system was removed every 2 ps. Configurations were saved about every 0.1 ps and were used to analyze various properties of liquid water as function of the type of MTS algorithm, time-step size and whether constraints are applied. In total 32 simulations were performed (Table 4).

TABLE 5 Internal energy of a single flexible (SPC/F) water molecule in the gas phase from 1 ns SD simulations with friction coefficient γ = 1 ps^{-1} as function of the time step Δt

Δt (fs)	E_{int}	ΔE_{int}	E_{tot}	ΔE_{tot}	E_{kin}	ΔE_{kin}	T (K)
0.1	3.74	3.01	14.99	5.84	11.25	5.07	300.7
0.2	3.79	3.04	15.08	6.15	11.28	5.32	301.6
0.5	3.69	3.00	14.83	5.96	11.14	5.16	297.7
1.0	3.81	3.03	14.97	5.91	11.16	4.99	298.3

Note: E_{int}: Bond-stretching and bond-angle bending energy. ΔE_{int}: Fluctuation of E_{int}. E_{tot}: Total energy. ΔE_{tot}: Fluctuation of E_{tot}. E_{kin}: Kinetic energy. ΔE_{kin}: Fluctuation of E_{kin}. T: Temperature. All values are averages calculated from trajectory structures separated by approximately 0.1 *ps*. Energies in kJ mol^{-1}.

2.8 | MD simulation of a protein solvated in explicit water under periodic boundary conditions

MD simulations of proteins in explicit water are commonly performed using periodic spatial boundary conditions and keeping the temperature and pressure constant. The algorithm and parameters of the temperature and pressure coupling were mentioned above. The nonbonded interactions were calculated using a single cut-off radius of 1.4 nm. Outside the cut-off radius a reaction-field approximation[40,41] with a relative dielectric permittivity of 78.5 was used. Minimum-image periodic boundary conditions were applied. The rigid SPC model was used for all water molecules in all simulations in order to only change the way the protein is modeled, flexible versus constrained, leaving the water molecules unchanged. These simulations were performed for 1 ns. Time steps Δt of 0.2, 0.5, 1.0, and 2.0 fs were tested. Since the constant force MTS algorithm did not perform well for the protein in vacuo, only the impulse-force MTS algorithm was applied using n_{hf} values of 3 and 5, that is, for $\Delta t'$ values of 0.6, 1.5, and 3.0 fs in case n_{hf} = 3, and for $\Delta t'$ values of 1.0, and 2.5 fs in case n_{hf} = 5. When constraints were applied, $tol_{DC} = 10^{-4}$ was used and the time steps tested were 0.2, 0.5, 1.0, and 2.0 fs. Translational motion of the center of mass of the system was removed every 2 ps. Configurations were saved about every 0.1 ps and were used to

TABLE 6 Atom-positional root-mean-square fluctuations (in nm) presented as averages over atoms of the same name[53] present in the protein, calculated from 1 ns MD simulations of the protein BPTI solvated in (SPC) water (*protein_aq*) using no MTS algorithm, the impulse-force MTS (*if-MTS*) algorithm or bond-length constraints for the protein, as function of the MD time step Δt and the number of time steps n_{hf} at which only the bond-stretching forces in the protein are evaluated and integrated

Algorithm			Protein atom names									
MTS or constraints	n_{hf}	Δt (fs)	*N*	CA	C	O	CB	CG, OG, SG	CD, ND, OD, SD	CE, NE, OE	CZ, NZ	CH, NH, OH
-	-	0.2	0.069	0.074	0.074	0.084	0.084	0.087	0.120	0.172	0.143	0.104
		0.5	0.074	0.081	0.083	0.091	0.088	0.088	0.130	0.183	0.162	0.105
		1.0	0.071	0.076	0.075	0.085	0.087	0.094	0.134	0184	0.159	0.101
		2.0	-	-	-	-	-	-	-	-	-	-
if	3	0.2	0.080	0.087	0.086	0.094	0.095	0.101	0.146	0.200	0.172	0.108
		0.5	0.069	0.074	0.074	0.082	0.083	0.088	0.123	0.173	0.151	0.102
		1.0	0.073	0.078	0.078	0.088	0.087	0.089	0.132	0.182	0.157	0.107
if	5	0.2	0.078	0.084	0.084	0.089	0.094	0.099	0.145	0.203	0.171	0.103
		0.5	0.071	0.076	0.076	0.085	0.085	0.094	0.129	0.186	0.157	0.103
constraints	-	0.2	0.068	0.073	0.073	0.083	0.085	0.092	0.125	0.180	0.154	0.097
		0.5	0.072	0.077	0.077	0.084	0.085	0.090	0.127	0.185	0.160	0.104
		1.0	0.069	0.073	0.072	0.082	0.084	0.090	0.134	0.184	0.158	0.102
		2.0	0.067	0.071	0.071	0.081	0.081	0.087	0.123	0.177	0.153	0.097

Note: Removal of center of mass translation every 2 ps. Relative geometric precision of the constraints: $tol_{DC} = 10^{-4}$. Nonbonded interaction cut-off radius $R_{cp} = R_{cl}$ = 1.4 nm. Outside a sphere of radius $R_{RF} = R_{cl}$, a homogeneous continuum dielectric with ε_{RF} = 78.5, κ_{RF} = 0 and ε_{cs} = 1 is assumed to be present. All values are averages calculated from trajectory structures separated by approximately 0.1 *ps*.

analyze various properties as function of the type of simulation (unconstrained vs. MTS vs. constrained) and time-step size. In total 13 simulations (*protein_aq*) were performed (Tables 5 and 6).

2.9 | Analysis of trajectory structures

Total energy conservation can be evaluated by comparing the fluctuation of the total energy with that of the kinetic energy. The former should be much smaller than the latter. The root-mean-square fluctuation (*RMSF*) of an energy E is defined as

$$\Delta E \equiv \left(< (E(t) - < E >_t)^2 >_t \right)^{1/2}, \tag{20}$$

where $<\ldots>_t$ indicates an average over time t. The drift E^{drift} of an energy E can be defined as the slope of the line $E^{line}(t)$ that is least-squares fitted to $E(t)$ for a chosen period of time. The quantity

$$\Delta E^{drift} \equiv \left(< \left(E(t) - E^{line}(t) \right)^2 >_t \right)^{1/2} \tag{21}$$

represents the deviation of the actual energy from the line representing the drift. ΔE^{drift} represents the short-time-scale fluctuation of E. This quantity may thus be better suited than ΔE to evaluate the extent of total energy conservation while integrating the equations of motion.

The properties of liquid water were calculated as specified in Reference 46. For the flexible SPC/F model of water, the molar heat of vaporization was calculated according to the following formula[47]:

$$\begin{aligned} \Delta H_{vap}(T) &= +U_{gas}(T) - U_{liquid}(T) + p\Delta V + Q \\ &= +U_{gas}(T) - U_{liquid}(T) + RT + Q, \end{aligned} \tag{22}$$

where U_{gas} is the computed potential energy per molecule in the gas phase, U_{liquid} is the computed potential energy per molecule in the liquid phase, p is the pressure, and ΔV is the molar volume change between liquid and gas. R is the gas constant and T is the absolute temperature. Q is a quantum correction that accounts (*i*) for the difference in vibrational energy of a water molecule between the liquid and the gas phases and (*ii*) for the difference between the vibrational energies calculated quantum-mechanically and classically. At room temperature Q = –0.23 kJ/mol.[47] For the rigid SPC model of water, U_{gas} = 0. For the flexible SPC/F model of water, the value for U_{gas} was obtained from a simulation using a Langevin thermostat, that is, solving Langevin's equations of motion[48] with a friction coefficient γ = 1 ps^{-1}.[49] The U_{gas} values obtained without an MTS algorithm were also used when calculating the heat of vaporization for the MTS simulations.

Atom-positional root-mean-square fluctuations, that is, around their average positions, in the MD trajectories were calculated after superposition of the positions of the backbone atoms (N, C_α, C) of residues 2–56 in the fit in order to eliminate the effect of overall translation and rotation of the protein upon the fluctuations.

Hydrogen bonds were identified according to a geometric criterion: a hydrogen bond was assumed to exist if the hydrogen-acceptor distance was smaller than 0.25 nm and the donor-hydrogen-acceptor angle was larger than 135°. When applied to each time frame of the MD trajectories, this definition yields the percentage of occurrence of a hydrogen bond. Using this criterion, transitions between different hydrogen bonds can be monitored. Transitions of a hydrogen bond from one to another configuration sometimes show a diffuse pattern: atomic positional fluctuations at the transition state may generate many counted transitions due to a strict application of the hydrogen bond criterion. Therefore, we have filtered out the effect of the short-time fluctuations by counting a transition only when a specific hydrogen bond has ceased to exist during 10-time frames (0.197 ps).[50] In this way, hydrogen-bond lifetimes were calculated, which have a lower bound of 0.197 ps and an upper bound of 1 ns, the time period covered by the MD simulations.

The time evolution of structural features that would be sensitive to the way the bond-stretching forces are integrated or to whether bond-length constraints are applied, was examined in terms of auto-correlation functions and spectral densities of fluctuations of atomic positions, of bond angles and of torsional angles.[9] From a time series of a quantity $Q(t)$, a normalized time correlation function,

$$C_Q(t) = \frac{< Q(\tau) \cdot Q(\tau + t) >_\tau}{< Q(\tau) \cdot Q(\tau) >_\tau} \tag{23}$$

was calculated using the Fast Fourier Transform technique.[51,52] For these calculations, 25 ps toward the end of the simulations were repeated while saving configurations every 0.01 ps instead of 0.1 ps in order to obtain a finer resolution of the auto-correlation functions. When calculating the spectral density, only the first 2% of the auto-correlation function was used.

3 | RESULTS

3.1 | MD simulation of a protein in vacuo to test conservation properties

In Table 1, the average total and kinetic energy and their fluctuations are shown for the protein in vacuo using various time steps and different ways to integrate the equations of motion. Using neither an MTS algorithm nor constraints (*no-MTS*), the ratio of the fluctuation of the total energy, ΔE_{tot}, and the fluctuation of the kinetic energy, ΔE_{kin}, changes from about 0.00061 for a time step Δt = 0.1 fs to 0.31 for Δt = 2.0 fs. Using bond-length constraints (*c-no-MTS*), this ratio grows from 0.0023 for Δt = 0.1 fs to 0.073 for Δt = 2.0 fs. For the larger time steps, using bond-length constraints leads to better total energy conservation than without constraints, as expected, due to the elimination of the high-frequency motions from the protein. For time steps smaller than 0.5 fs, the precision of $tol_{DC} = 10^{-6}$ by which the constraints were maintained, becomes the dominating factor for total

energy nonconservation. In order to better conserve total energy for these small time steps, smaller values of tol_{DC} would have to be used.

Using the impulse-force MTS algorithm (*if-MTS*) with n_{hf} = 3, the ratio $\Delta E_{tot}/\Delta E_{kin}$ changes from 0.042 for $\Delta t'$ = 0.3 fs to 0.48 for $\Delta t'$ = 3.0 fs, and with n_{hf} = 5, it changes from 0.084 for $\Delta t'$ = 0.5 fs to 0.41 for $\Delta t'$ = 2.5 fs. Using the constant-force MTS algorithm (*cf-MTS*), the total energy rises such that no stable simulation was obtained for the time steps tested.

When comparing the *if-MTS* algorithm with the use of constraints, the latter conserves the total energy better than the former. The ratio $\Delta E_{tot}/\Delta E_{kin}$ is about 0.085 for $\Delta t'$ = 0.6 fs (n_{hf} = 3) and $\Delta t'$ = 0.5 fs (n_{hf} = 5) to be compared with a value 0.017 for $\Delta t'$ = 0.5 fs (*c-no-MTS*) when bond-length constraints are applied. Comparing these algorithms for larger time steps, the use of constraints also leads to better energy conservation. With n_{hf} = 3, the ratio $\Delta E_{tot}/\Delta E_{kin}$ is 0.23 for $\Delta t'$ = 1.5 fs and 0.48 for $\Delta t'$ = 3.0 fs. With n_{hf} = 5, the ratio $\Delta E_{tot}/\Delta E_{kin}$ is 0.17 for $\Delta t'$ = 1.0 fs and 0.41 for $\Delta t'$ = 2.5 fs. These values are much larger than the ones obtained using constraints: a value 0.032 for $\Delta t'$ = 1.0 fs and 0.073 for $\Delta t'$ = 2.0 fs (*c-no-MTS*).

When applying bond-length constraints, the total energy of the protein is better conserved than when applying a multiple-time-step algorithm, no matter whether an impulse-force one or a constant-force one is used, when considering comparable time step sizes $\Delta t'$. The computational effort of the evaluation of the bond-stretching forces at the small time steps Δt is comparable to that of maintaining the bond-length constraints at the larger time steps $\Delta t'$.

In the presence of bond-length constraints there are no bond-stretching degrees of freedom, so the total energy is rather different from the one in the other simulations. Yet, the potential energy of the nonbond-stretching force-field terms in the flexible model, $E_{tot} - E_{bond} - E_{kin}$ is with about −3050 kJ/mol only slightly lower than that in the constrained model, for which $E_{tot} - E_{kin}$ is about −2930 kJ/mol. This is expected because of the additional (bond-stretching) degrees of freedom in the flexible model.

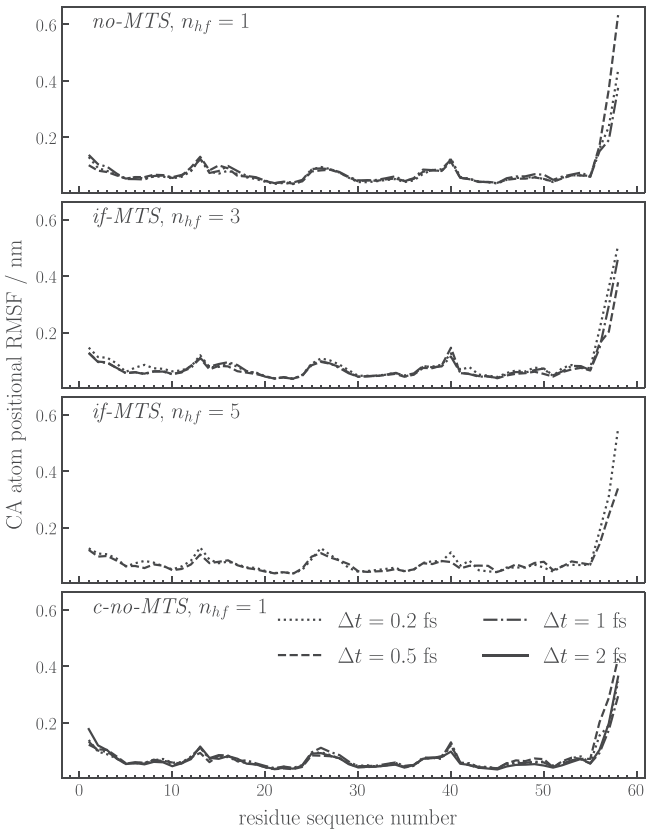

FIGURE 2 Root-mean-square fluctuations (in nm) of the CA atoms of the protein BPTI as function of residue number in 1 ns MD simulations (*protein_aq*) based on different algorithms and time steps Δt. Upper panel: *no-MTS* algorithm. Middle panels: *if-MTS* algorithm with n_{hf} = 3 and n_{hf} = 5, respectively. Lower panel: *c-no-MTS* algorithm. Dotted lines: Δt = 0.2 fs. Dashed lines: Δt = 0.5 fs. Dot-dashed lines: Δt = 1.0 fs. Solid lines: Δt = 2.0 fs

3.2 | MD simulation of a single water molecule in vacuo to test conservation properties

In Table 2, the average total and kinetic energy and their fluctuations are shown for a single water molecule in vacuo using the flexible SPC/F model and the rigid SPC model for various time steps. Using neither an MTS algorithm nor constraints (SPC/F, *no-MTS*), the ratio $\Delta E_{tot}/\Delta E_{kin}$ of the fluctuation of the total energy, ΔE_{tot}, and the fluctuation of the kinetic energy, ΔE_{kin}, changes from 0.0014 for a time step Δt = 0.1 fs to 2.9 for Δt = 2.0 fs. Using bond-length and bond-angle constraints (SPC, *c-no-MTS*), the potential energy is zero, so the kinetic energy is equal to the total energy, which means $\Delta E_{tot}/\Delta E_{kin}$ = 1. In this case one may compare the ratio $\Delta E_{tot}/E_{tot} = \Delta E_{kin}/E_{kin}$. Using neither an MTS algorithm nor constraints (SPC/F, *no-MTS*), this ratio changes from about 0.0000018 for a time step Δt = 0.1 fs to 0.0059 for Δt = 2.0 fs. Using bond-length and bond-angle constraints (SPC, *c-no-MTS*), this ratio changes from 0.0000027 for a time step Δt = 0.1 fs to 0.000012 for Δt = 2.0 fs and 0.000067 for Δt = 10.0 fs.

As expected, using a rigid model for water, the total energy is much better conserved than using a flexible one.

3.3 | MD simulations of liquid water in a periodic box to test conservation properties

In Table 3, the average total and kinetic energy and their fluctuations are shown for liquid water, 1000 molecules in a periodic box, using no MTS algorithm or the *if-MTS* algorithm for the flexible SPC/F model or three distance constraints for the rigid SPC model, as function of the MD time step Δt and the number of time steps n_{hf} at which only the bond-stretching and bond-angle forces (SPC/F) are evaluated and integrated. In this case, the nonbonded interactions, van der Waals and electrostatic, dominate the potential energy and the noise induced by the use of a nonbonded interaction cut-off radius (1.4 nm) is the dominant factor causing total energy nonconservation. Thus the ratio $\Delta E_{tot}/\Delta E_{kin}$ does not change much when varying the time step

TABLE 7 Torsional-angle root-mean-square fluctuations $\Delta\varphi$, etc. (RMSF; in degree) and their standard deviations *stdev*($\Delta\varphi$), etc. presented as averages over the angles of the same type or name[53] present in the protein, calculated from 1 ns MD simulations of the protein BPTI solvated in (SPC) water (*protein_aq*) using no MTS algorithm, the impulse-force MTS (*if-MTS*) algorithm or bond-length constraints for the protein, as function of the MD time step Δt and the number of time steps n_{hf} at which only the bond-stretching forces in the protein are evaluated and integrated

Algorithm			Protein torsional angle							
MTS or constraints	n_{hf}	Δt (fs)	$\Delta\varphi$	*stdev*($\Delta\varphi$)	$\Delta\psi$	*stdev*($\Delta\psi$)	$\Delta\chi_1$	*stdev*($\Delta\chi_1$)	$\Delta\chi_2$	*stdev*($\Delta'\chi_2$)
-	-	0.2	16.5	12.0	18.2	21.2	21.4	15.2	36.0	19.5
		0.5	17.9	14.1	20.4	21.6	24.7	23.0	39.2	21.8
		1.0	17.0	13.2	17.2	12.3	24.2	16.1	37.8	16.3
		2.0	-	-	-	-	-	-	-	-
if	3	0.2	16.7	11.0	16.7	9.4	26.3	22.9	49.4	35.3
		0.5	15.0	8.6	15.5	11.0	25.8	29.0	39.2	25.5
		1.0	17.1	13.2	16.8	10.3	21.0	12.7	42.3	26.9
if	5	0.2	15.6	8.6	17.4	17.5	27.0	23.9	39.2	15.9
		0.5	16.4	11.6	17.4	16.2	26.1	22.0	36.0	14.2
constraints	-	0.2	16.0	11.5	15.5	8.0	23.6	15.7	36.1	14.4
		0.5	17.4	24.2	17.9	24.4	25.3	27.3	43.5	30.9
		1.0	15.3	8.8	15.4	8.1	26.1	20.4	48.7	35.2
		2.0	15.7	11.2	16.4	14.2	21.6	14.7	35.2	19.7

Note: Removal of center of mass translation every 2 ps. Relative geometric precision of the constraints: $tol_{DC} = 10^{-4}$. Nonbonded interaction cut-off radius $R_{cp} = R_{cl} = 1.4$ nm. Outside a sphere of radius $R_{RF} = R_{cl}$, a homogeneous continuum dielectric with $\varepsilon_{RF} = 78.5$, $\kappa_{RF} = 0$ and $\varepsilon_{cs} = 1$ is assumed to be present. All values are calculated from trajectory structures separated by approximately 0.1 *ps*.

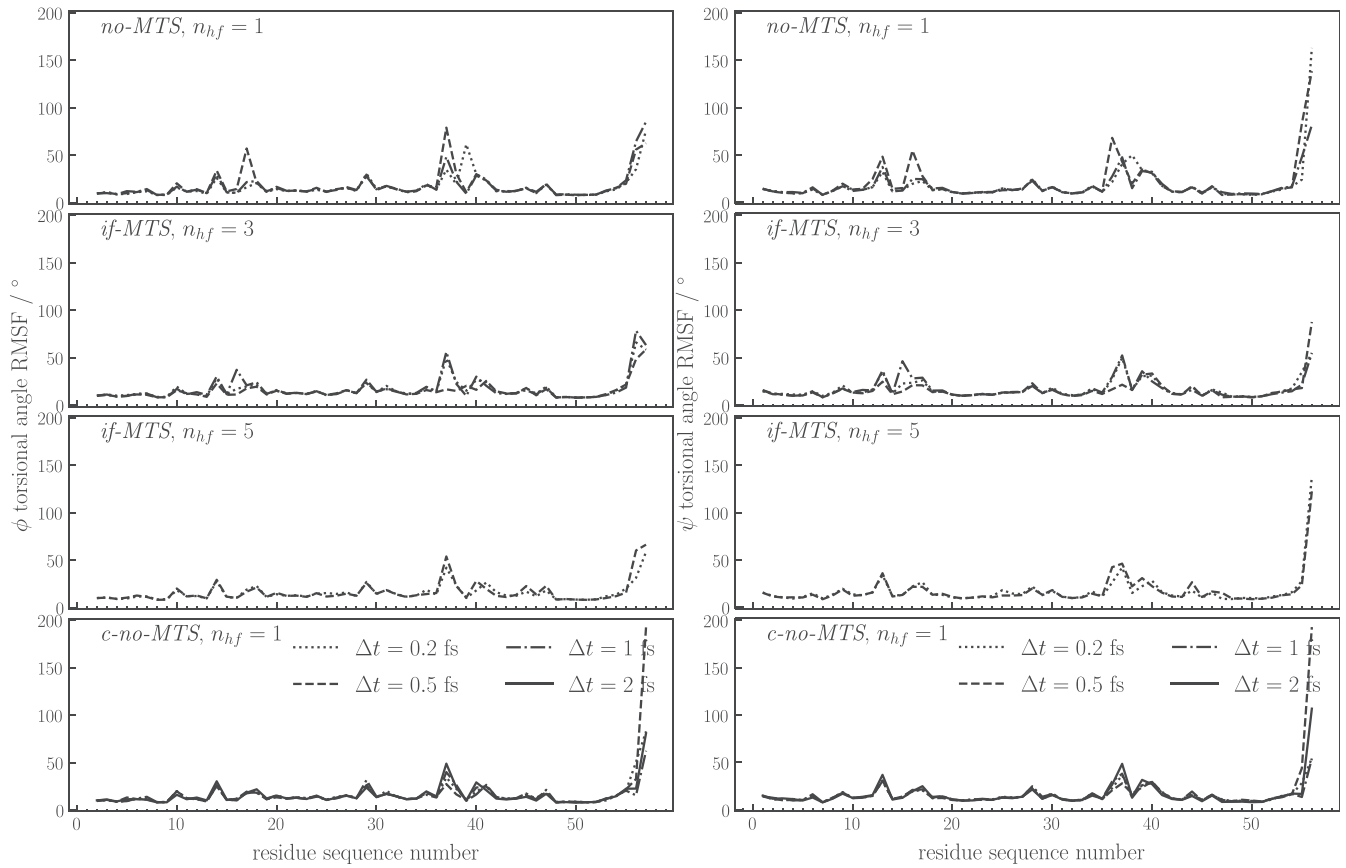

FIGURE 3 Root-mean-square fluctuations (in degree) of the φ-(left panels) and ψ-(right panels) angles in the backbone of the protein BPTI as function of residue number in 1 ns MD simulations (*protein_aq*) based on different algorithms and time steps Δt. Upper panel: *no-MTS* algorithm. Middle panels: *if-MTS* algorithm with $n_{hf} = 3$ and $n_{hf} = 5$, respectively. Lower panel: *c-no-MTS* algorithm. Dotted lines: $\Delta t = 0.2$ fs. Dashed lines: $\Delta t = 0.5$ fs. Dot-dashed lines: $\Delta t = 1.0$ fs. Solid lines: $\Delta t = 2.0$ fs

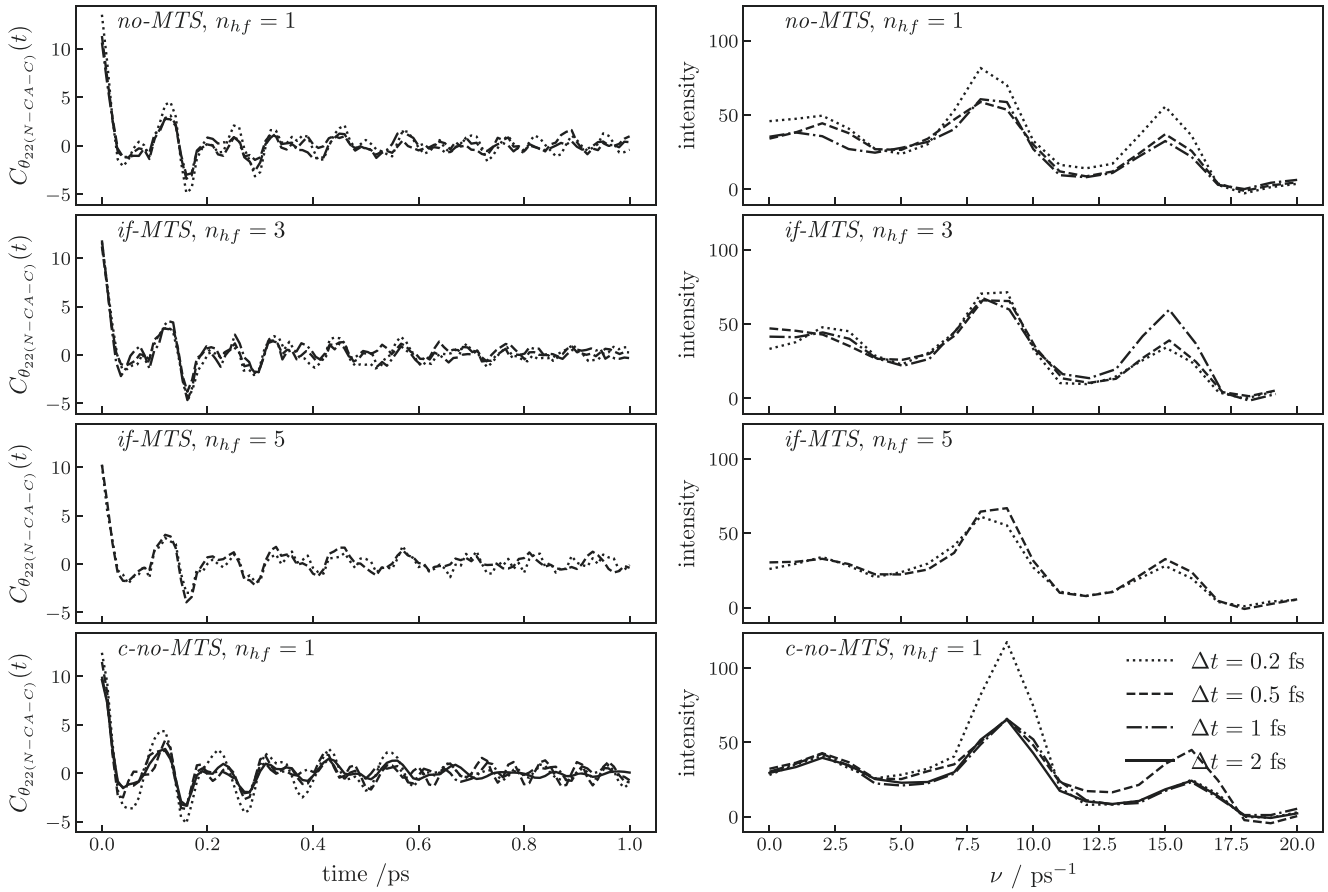

FIGURE 4 Auto-correlation function (left panels) and spectral density (right panels) of the bond angle θ(N-CA-C) of residue Phe 22 in the backbone of the protein BPTI in 1 ns MD simulations (*protein_aq*) based on different algorithms and time steps Δt. Upper panel: *no-MTS* algorithm. Middle panels: *if-MTS* algorithm with n_{hf} = 3 and n_{hf} = 5, respectively. Lower panel: *c-no-MTS* algorithm. Dotted lines: Δt = 0.2 fs. Dashed lines: Δt = 0.5 fs. Dot-dashed lines: Δt = 1.0 fs. Solid lines: Δt = 2.0 fs. Configurations from 25 ps toward the end of the simulations, separated by 0.01 ps were used to calculate the auto-correlation functions and only the first 2% of the auto-correlation function was used to calculate the spectral density

Δt and the n_{hf} values, except for the *if-MTS* algorithm with Δt = 1.0 fs and n_{hf} = 5, or $\Delta t'$ = 5.0 fs. For this time step, using constraints leads to better total energy conservation. As before, the cf-*MTS* algorithm did not yield stable simulations.

As for the protein, when applying bond-length (and bond-angle in this case) constraints, the total energy of liquid water is better conserved than when applying a multiple-time-step algorithm, no matter whether an impulse-force one or a constant-force one is used, when considering comparable time step sizes $\Delta t'$. The computational effort of the evaluation of the bond-stretching and bond-angle bending forces at the small time steps Δt is comparable to that of maintaining the bond-length (and bond-angle) constraints at the larger time steps $\Delta t'$.

3.4 | MD simulations of liquid water in a periodic box to test properties of the liquid

In Table 4, some thermodynamic and dynamic properties of liquid water, 1000 molecules in a periodic box simulated at constant temperature and pressure, are shown when applying no MTS algorithm or two different (*if* and cf) MTS algorithms for the flexible SPC/F model or three distance constraints for the rigid SPC model. Using neither an MTS algorithm nor constraints (*no-MTS*), the properties are stable up till a time step Δt = 1.0 fs. Using bond-length and bond-angle constraints (*c-no-MTS*), this is the case up till Δt = 2.0 fs. For a larger time step, Δt = 5.0 fs, the translational and rotational temperatures start to diverge. The rotational degrees of freedom of a water molecule are more influenced by the nonbonded cut-off noise than the translational ones. When coupling these degrees of freedom jointly to a single heat bath, only the total temperature is scaled to stay close to the reference temperature of the heat bath. This leads to a lower translational temperature and a higher rotational temperature. For Δt = 5.0 fs, some properties of liquid water begin to change too.

Using the impulse-force MTS algorithm (*if-MTS*) with n_{hf} = 3, properties begin to deviate at $\Delta t'$ = 3.0 fs, and with n_{hf} = 5, at $\Delta t'$ = 5.0 fs. Using the constant-force MTS algorithm (cf-*MTS*), the simulations are stable due to the temperature coupling, which removes the excess heat generated. The various properties of the

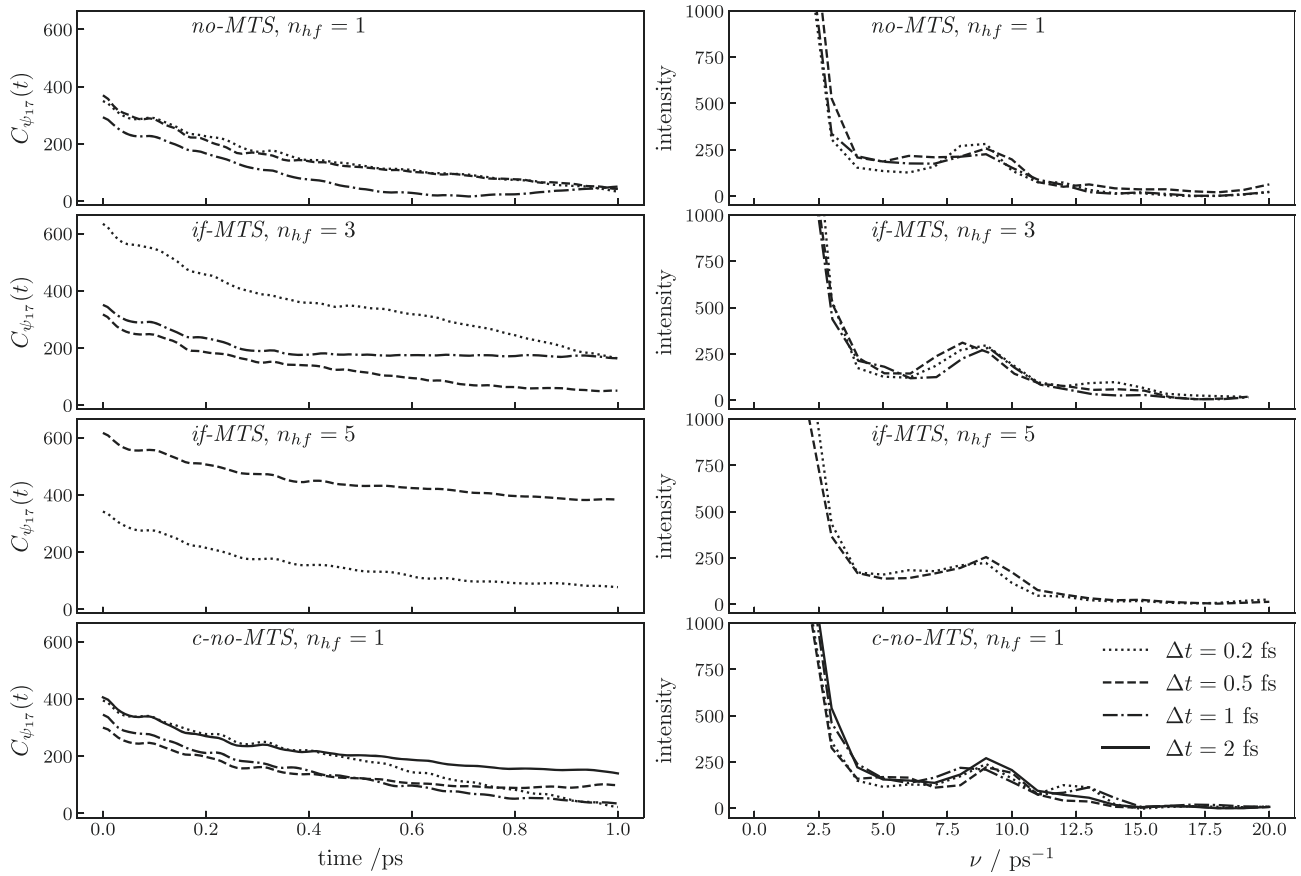

FIGURE 5 Auto-correlation function (left panels) and spectral density (right panels) of the torsional angle ψ(N-CA-C-N) of residue Arg 17 in the backbone of the protein BPTI in 1 ns MD simulations (*protein_aq*) based on different algorithms and time steps Δt. Upper panel: *no-MTS* algorithm. Middle panels: *if-MTS* algorithm with n_{hf} = 3 and n_{hf} = 5, respectively. Lower panel: *c-no-MTS* algorithm. Dotted lines: Δt = 0.2 fs. Dashed lines: Δt = 0.5 fs. Dot-dashed lines: Δt = 1.0 fs. Solid lines: Δt = 2.0 fs. Configurations from 25 ps toward the end of the simulations, separated by 0.01 ps were used to calculate the auto-correlation functions and only the first 2% of the auto-correlation function was used to calculate the spectral density

liquid are strongly dependent on the size of the time step. The density and heat of vaporization become too low, the diffusion too fast, and hydrogen bond lifetimes become shorter. The rigid water model shows slightly slower dynamics than the flexible model, which is not surprising.

In Table 5, the internal energy of a flexible water molecule in the gas phase is shown as obtained from simulations using a Langevin thermostat. In the liquid phase (Table 3), the internal energy E_{int} in the flexible SPC/F model is per molecule larger than for a single molecule in vacuo (Table 5). This is due to the nonbonded interactions with other water molecules in the periodic box. The dipole–dipole interactions lead to slightly larger bond lengths and to a smaller bond angle, which increases the molecular dipole moment, and thus to larger bond-stretching and bond-angle bending energies.[7]

When comparing the *if-MTS* algorithm with the use of constraints, the latter yields slower translational and rotational dynamics of the water molecules. The rigid model also reproduces the density and heat of vaporization of liquid water better. This is no surprise, because the parameters of the rigid SPC model were calibrated to reproduce these two quantities,[33] whereas those of the flexible SPC/F model were not.[7]

3.5 | MD simulations of a protein in aqueous solution to test protein properties

In Table 6, the mobility (atom-positional RMSFs) of the different non-hydrogen atoms in BPTI bearing the same name,[53] derived from 1 ns MD simulations of the protein solvated in (SPC) water (*protein_aq*), is shown for the three different algorithms to integrate Newton's equations of motion and different sizes of the MD time step Δt. The mobility of the atoms is rather comparable for the three algorithms, with the *if-MTS* algorithm tending to slightly enhanced mobility. In Figure 2, the positional root-mean-square fluctuations (in nm) of the CA atoms of BPTI are shown as function of residue number for the different algorithms and time steps Δt. The patterns are very similar.

Table 7 shows the mobility of the backbone torsional angles φ and ψ, and the first two side-chain torsional angles χ_1 and χ_2, that is, root-mean-square fluctuations $\Delta\varphi$, $\Delta\psi$, $\Delta\chi_1$, and $\Delta\chi_2$, and their standard deviations, presented as averages over the angles of the same type or name[53] present in the protein, for the three different integration algorithms and different sizes of the MD time step Δt. The mobility of these torsional angles is rather comparable for the three

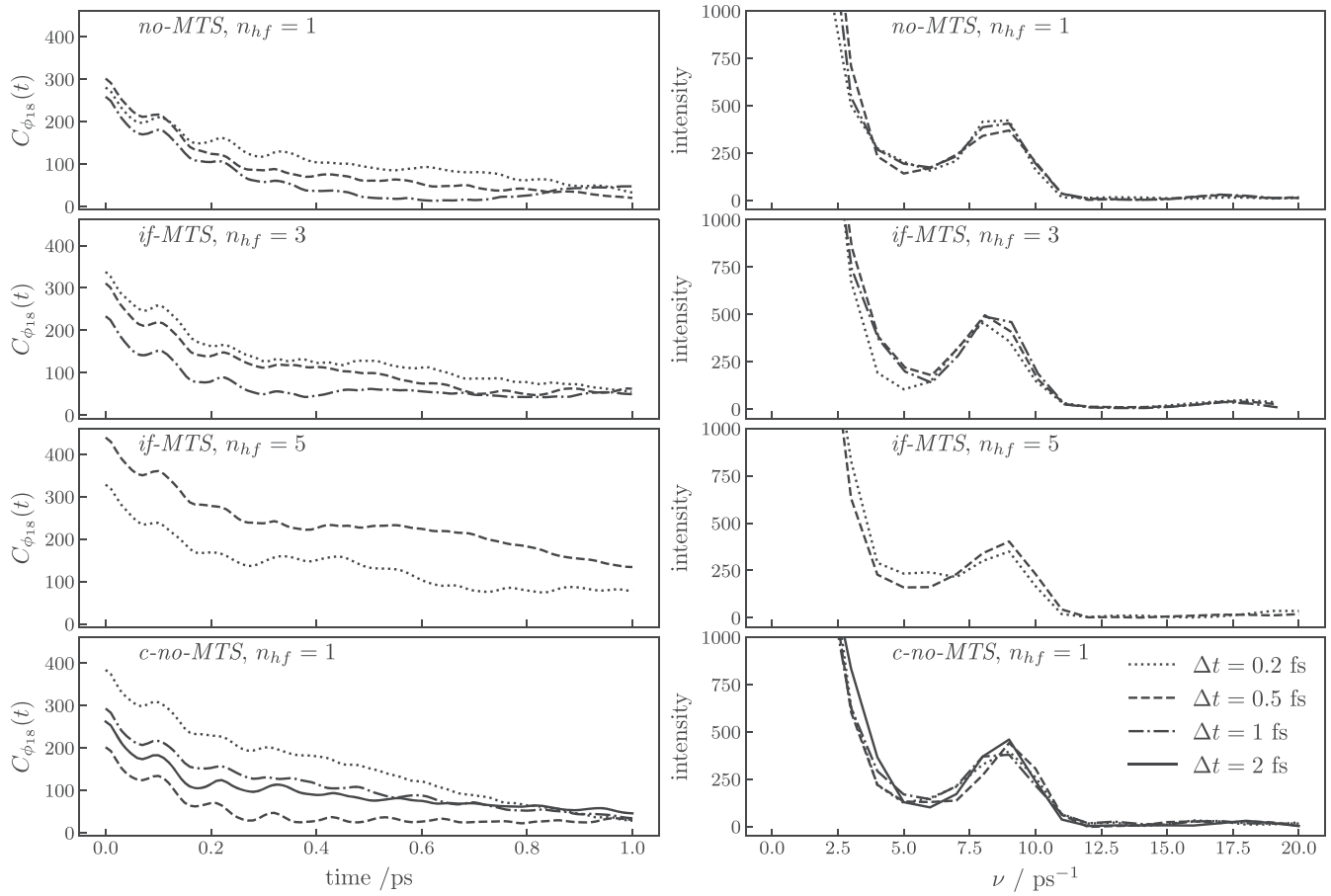

FIGURE 6 Auto-correlation function (left panels) and spectral density (right panels) of the torsional angle φ(C-N-CA-C) of residue Ile 18 in the backbone of the protein BPTI in 1 ns MD simulations (*protein_aq*) based on different algorithms and time steps Δt. Upper panel: *no-MTS* algorithm. Middle panels: *if-MTS* algorithm with n_{hf} = 3 and n_{hf} = 5, respectively. Lower panel: *c-no-MTS* algorithm. Dotted lines: Δt = 0.2 fs. Dashed lines: Δt = 0.5 fs. Dot-dashed lines: Δt = 1.0 fs. Solid lines: Δt = 2.0 fs. Configurations from 25 ps toward the end of the simulations, separated by 0.01 ps were used to calculate the auto-correlation functions and only the first 2% of the auto-correlation function was used to calculate the spectral density

algorithms. The variation between the values for different time steps is due to the occurrence of relatively rare (on the nanosecond time scale) torsional-angle transitions over relatively low barriers separating the different minima of the torsional-angle potential-energy terms in the force field used. This is illustrated in Figure 3, which shows the root-mean-square fluctuations of the backbone φ and ψ torsional angles as function of residue number for the different algorithms and time steps Δt. The larger peaks are due to such transitions. For example, in simulation *if-MTS* with n_{hf} = 5 with Δt = 0.2 fs, the peptide plane between residues 46 and 47 changes orientation which induces a correlated change in ψ(46) and φ(47). A similar change of orientation of a peptide plane is observed for ψ(39) and φ(40) in simulation *c-no-MTS* with Δt = 0.5 fs.

The different treatments of the bond-length degrees of freedom when integrating the equations of motion would primarily affect the motions along the (adjacent) bond-angle degrees of freedom. The influence of the three different treatments can be inferred from Figure 4, which shows the auto-correlation function and spectral density of the bond angle θ(N-CA-C) of residue Phe 22 in the backbone of the protein BPTI for the different algorithms and time steps Δt. The different curves are almost identical.

Figures 5 and 6 show the auto-correlation function and spectral density of the torsional angles ψ(N-CA-C-N) of residue Arg 17 and φ(C-N-CA-C) of residue Ile 18 in the backbone of BPTI for the different algorithms and time steps Δt. The spectral densities are rather similar, while the auto-correlation functions show differences in the longer time (beyond 0.2 ps) correlation. This is due to torsional-angle transitions occurring rarely on the simulated time scale. The difference of the angle at time t with its average is much larger when a transition occurs than when this is not the case. When a transition occurs, the difference of the angle with its average is thus only slowly reduced, leading to a slow decay of the auto-correlation function. In case there is no transition, the difference of the angle with its average is much smaller and changes much more rapidly, leading to a much faster decay of the auto-correlation function. The effect of relatively rare torsional-angle transitions on the auto-correlation function is even more prominent for side-chain torsional angles, as is illustrated in Figure 7 for the side-chain torsional angle χ_2(CA-CB-CG-CD) of residue Arg 39.

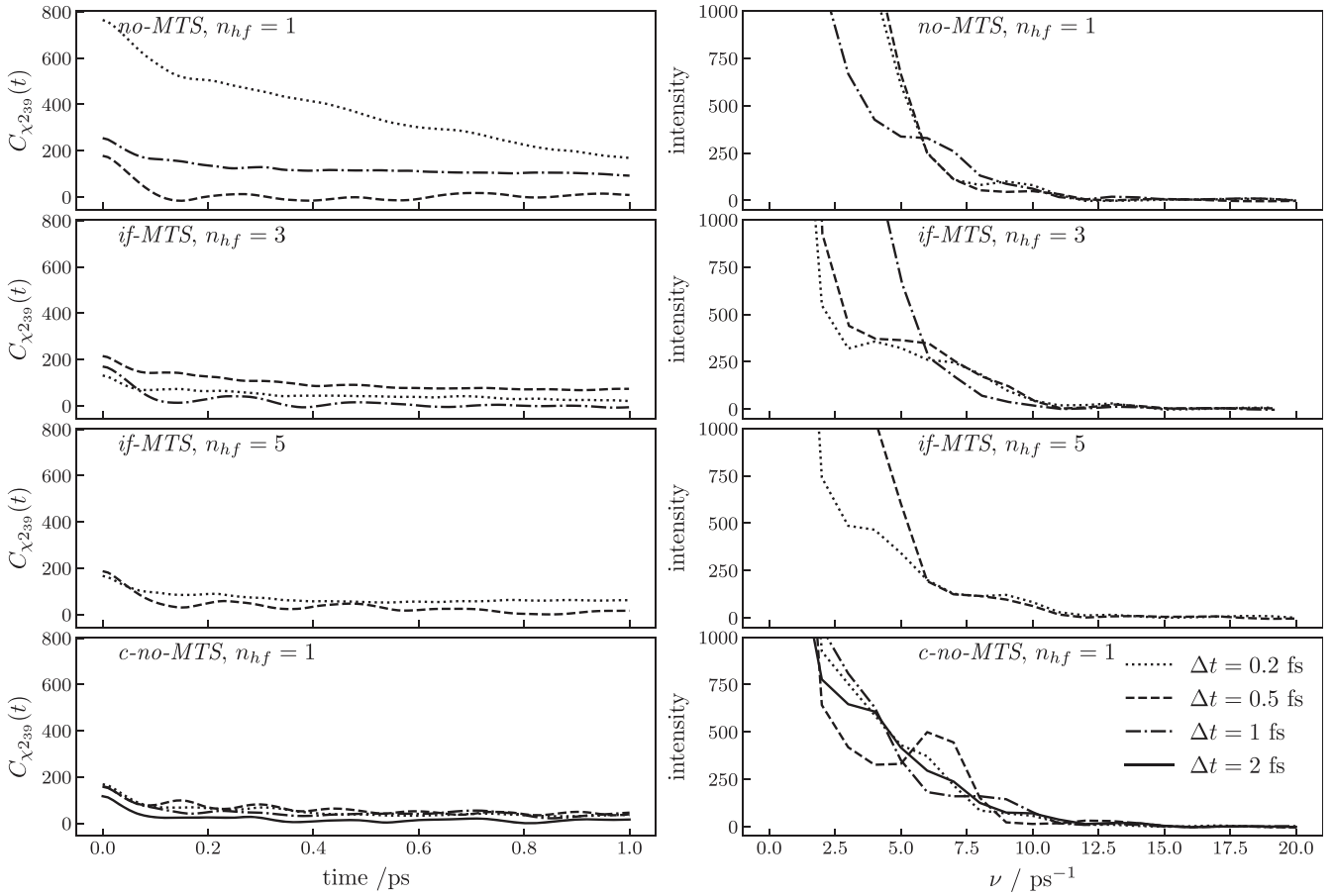

FIGURE 7 Auto-correlation function (left panels) and spectral density (right panels) of the torsional angle χ_2(CA-CB-CG-CD) of residue Arg 39 of the protein BPTI in 1 ns MD simulations (*protein_aq*) based on different algorithms and time steps Δt. Upper panel: *no-MTS* algorithm. Middle panels: *if-MTS* algorithm with n_{hf} = 3 and n_{hf} = 5, respectively. Lower panel: *c-no-MTS* algorithm. Dotted lines: Δt = 0.2 fs. Dashed lines: Δt = 0.5 fs. Dot-dashed lines: Δt = 1.0 fs. Solid lines: Δt = 2.0 fs. Configurations from 25 ps toward the end of the simulations, separated by 0.01 ps were used to calculate the auto-correlation functions and only the first 2% of the auto-correlation function was used to calculate the spectral density

The average hydrogen-bond lifetimes, protein–protein and protein-water, did not show dependence on the time step.

In summary, when comparing the motions along nonbond-stretching degrees of freedom in the simulations using the three different integration algorithms, no significant differences are observed. This indicates that the motions along the bond-stretching degrees of freedom are virtually decoupled from the motions along the other degrees of freedom of the protein. Regarding the use of bond-length constraints, this conclusion was also obtained from 25 ps MD simulations of BPTI in vacuo using a simpler force field without explicit hydrogen atoms.[9]

4 | CONCLUSIONS

The application of two different multiple-time-step (MTS) algorithms, an impulse-force one (*if-MTS*) and a constant-force one (cf-*MTS*), to integrate the bond-stretching forces in a protein with a smaller molecular dynamics (MD) time step as used for the other forces in the molecule, is compared to standard (*no-MTS*) simulation and to the use of bond-length constraints in a simulation (*c-no-MTS*). The protein bovine pancreatic trypsin inhibitor (BPTI), in vacuo and in aqueous solution, and liquid water were used as test systems.

When applied to different ranges of nonbonded interactions, the *if-MTS* algorithm was found to be less appropriate than the cf-*MTS* one,[25–27] but when applied to the bond-stretching versus other forces the opposite is the case. However, bond-stretching degrees of freedom are better treated as constrained degrees of freedom.

When applying bond-length constraints, the total energy of the protein is better conserved than when applying a multiple-time-step algorithm, no matter whether an impulse-force one or a constant-force one is used, when considering comparable time step sizes $\Delta t'$. The computational effort of the evaluation of the bond-stretching forces at the small time steps Δt is comparable to that of maintaining the bond-length constraints at the larger time steps $\Delta t'$. When applying bond-length and bond-angle constraints to water molecules, the same observations hold. The total energy of liquid water is better

conserved than when applying a multiple-time-step algorithm, no matter whether an impulse-force one or a constant-force one is used and the computational effort of the evaluation of the bond-stretching and bond-angle bending forces at the small time steps Δt is comparable to that of maintaining the bond-length and bond-angle constraints at the larger time steps $\Delta t'$. When comparing the motions along nonbond-stretching degrees of freedom in the simulations of the protein BPTI in aqueous solution, use of the three different integration algorithms shows no significant differences in dynamical properties between the three different integration algorithms. This indicates that the motions along the bond-stretching degrees of freedom are virtually decoupled from the motions along the other degrees of freedom of the protein, which is a condition for the application of constraints or an MTS algorithm. Yet, the application of bond-length constraints leads to less distortion of the dynamics of the atoms in the molecules than the application of a multiple-time-step algorithm, and constitutes a better representation of the quantum-mechanical nature of bond-stretching vibrations in proteins than a classically, for example, harmonically, vibrating bond.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Maria Pechlaner https://orcid.org/0000-0002-1615-5600
Chris Oostenbrink https://orcid.org/0000-0002-4232-2556

REFERENCES

[1] J. A. McCammon, B. R. Gelin, M. Karplus, *Nature* **1977**, *267*, 585.
[2] W. F. van Gunsteren, H. J. C. Berendsen, *Mol. Phys.* **1977**, *34*, 1311.
[3] W. F. van Gunsteren, H. J. C. Berendsen, *Am. Ethnol.* **1990**, *102*, 1020.
[4] W. F. van Gunsteren, D. Bakowies, R. Baron, I. Chandrasekhar, M. Christen, X. Daura, P. Gee, D. P. Geerke, A. Glättli, P. H. Hünenberger, M. A. Kastenholz, C. Oostenbrink, M. Schenk, D. Trzesniak, N. F. A. van der Vegt, H. B. Yu, *Angew. Chem* **2006**, *118*, 4168.
[5] W. F. van Gunsteren, X. Daura, N. Hansen, A. E. Mark, C. Oostenbrink, S. Riniker, L. J. Smith, *Am. Ethnol.* **2018**, *130*, 894.
[6] W. F. van Gunsteren, Computer Simulation of Biomolecular Systems: Overview of time-saving techniques, *AIP Conference Proceedings*, **1991**, *239*, 131.
[7] I. G. Tironi, R. M. Brunne, W. F. van Gunsteren, *Chem. Phys. Lett.* **1996**, *250*, 19.
[8] S. Reißer, D. Poger, M. Stroet, A. E. Mark, *J. Chem. Theory Comput.* **2017**, *13*, 2367.
[9] W. F. van Gunsteren, M. Karplus, *Macromolecules* **1982**, *15*, 1528.
[10] J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327.
[11] H. C. Andersen, *J. Comput. Phys.* **1983**, *52*, 24.
[12] E. Barth, K. Kuczera, B. Leimkuhler, R. D. Skeel, *J. Comput. Chem.* **1995**, *16*, 1192.
[13] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, *J. Comput. Chem.* **1997**, *18*, 1463.
[14] P. Tao, X. Wu, B. R. Brooks, *J. Chem. Phys.* **2012**, *137*, 134110.
[15] J. L. Finney, *J. Comput. Phys.* **1978**, *28*, 92.
[16] W. B. Street, D. J. Tildesley, G. Saville, *Mol. Phys.* **1978**, *35*, 639.
[17] H. J. C. Berendsen, W. F. van Gunsteren, H. R. J. Zwinderman, R. G. Geurtsen, *Ann. N. Y. Acad. Sci.* **1986**, *482*, 269.
[18] W. F. van Gunsteren, H. J. C. Berendsen, R. G. Geurtsen, H. R. J. Zwinderman, *Ann. N. Y. Acad. Sci.* **1986**, *482*, 287.
[19] O. Teleman, B. Jönsson, *J. Comput. Chem.* **1986**, *7*, 58.
[20] V. Kräutler, P. H. Hünenberger, *J. Comput. Chem.* **2006**, *27*, 1163.
[21] R. W. Hockney, J. W. Eastwood, *Computer Simulation Using Particles*, McGraw-Hill, New York **1981**.
[22] M. A. Cuendet, W. F. van Gunsteren, *J. Chem. Phys.* **2007**, *127*, 184102.
[23] U. Stocker, D. Juchli, W. F. van Gunsteren, *Mol. Sim.* **2003**, *29*, 123.
[24] A. M. J. J. Bonvin, A. E. Mark, W. F. van Gunsteren, *Comput. Phys. Commun.* **2000**, *128*, 550.
[25] M. Diem, C. Oostenbrink, *J. Chem. Theory Comput.* **2020**, *16*, 5985.
[26] M. Diem, C. Oostenbrink, *J. Comput. Chem.* **2020**, *41*, 2740.
[27] W. F. van Gunsteren, X. Daura, P. F. J. Fuchs, N. Hansen, B. A. C. Horta, P. H. Hünenberger, A. E. Mark, M. Pechlaner, S. Riniker, C. Oostenbrink, *ChemPhysChem* **2021**, *22*, 264.
[28] H. Grubmüller, H. Heller, A. Windemuth, K. Schulten, *Mol. Simul.* **1991**, *6*, 121.
[29] Q. Ma, J. A. Izaguirre, R. D. Skeel, *SIAM J. Sci. Comput.* **2003**, *24*, 1951.
[30] J. A. Morrone, R. Zhou, B. J. Berne, *J. Chem. Theory Comput.* **2010**, *6*, 1798.
[31] T. C. Bishop, R. D. Skeel, K. Schulten, *J. Comput. Chem.* **1997**, *18*, 1785.
[32] S. J. Stuart, R. Zhou, B. J. Berne, *J. Chem. Phys.* **1996**, *105*, 1426.
[33] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, J. Hermans, Interaction models for water in relation to protein hydration. in *Intermolecular Forces* (Ed: B. Pullman), Reidel, Dordrecht, The Netherlands **1981**, p. 331.
[34] D. Poger, W. F. van Gunsteren, A. E. Mark, *J. Comput. Chem.* **2010**, *31*, 1117.
[35] N. Schmid, A. P. Eichenberger, A. Choutko, S. Riniker, M. Winger, A. E. Mark, W. F. van Gunsteren, *Eur. Biophys. J.* **2011**, *40*, 843.
[36] W. F. van Gunsteren, et al., *The GROMOS Software for (Bio)Molecular Simulation. Volume 3: Force Field and Topology Data Set* **2011**. http://www.gromos.net/. Accessed 4 July 2019.
[37] W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott, I. G. Tironi, *Biomolecular Simulation: The GROMOS96 Manual and User Guide*, Vdf Hochschulverlag AG an der ETH Zürich, Zürich, Switzerland **1996**.
[38] W. F. van Gunsteren, et al., *The GROMOS Software for (Bio)Molecular Simulation. Volume 2: Algorithms and Formulae for Modelling of Molecular Systems* **2011**. http://www.gromos.net/. Accessed 4 July 2019.
[39] W. F. van Gunsteren, et al., *The GROMOS Software for (Bio)Molecular Simulation. Volumes 1–9* **2011** http://www.gromos.net/. Accessed 4 July 2019.
[40] J. A. Barker, R. O. Watts, *Mol. Phys.* **1973**, *26*, 789.
[41] I. G. Tironi, R. Sperb, P. E. Smith, W. F. van Gunsteren, *J. Chem. Phys.* **1995**, *102*, 5451.
[42] A. P. E. Kunz, J. R. Allison, D. P. Geerke, B. A. C. Horta, P. H. Hünenberger, S. Riniker, N. Schmid, W. F. van Gunsteren, *J. Comput. Chem.* **2012**, *33*, 340.
[43] N. Schmid, C. D. Christ, M. Christen, A. P. Eichenberger, W. F. van Gunsteren, *Comp. Phys. Commun.* **2012**, *183*, 890.
[44] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res.* **2000**, *28*, 235. www.pdb.org.
[45] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684.
[46] A. Glättli, X. Daura, W. F. van Gunsteren, *J. Chem. Phys.* **2002**, *116*, 9811.
[47] J. P. M. Postma, Ph.D. thesis, Rijksuniversiteit, Groningen, The Netherlands **1985**.

[48] W. F. van Gunsteren, H. J. C. Berendsen, *Mol. Simul.* **1988**, *1*, 173.
[49] Z. Lin, W. F. van Gunsteren, *J. Chem. Phys.* **2015**, *143*, 034110.
[50] J. E. H. Koehler, W. Saenger, W. F. van Gunsteren, *Eur. Biophys. J.* **1988**, *16*, 153.
[51] R. P. Futrelle, D. J. McGinty, *Chem. Phys. Lett.* **1971**, *12*, 285.
[52] W. F. van Gunsteren, et al., *The GROMOS Software for (Bio)Molecular Simulation. Volume 6: Technical Details* **2011**. http://www.gromos.net/. Accessed 4 July 2019.
[53] IUPAC-IUB, *Biochemistry* **1970**, *9*, 3471.

How to cite this article: Pechlaner M, Oostenbrink C, van Gunsteren WF. On the use of multiple-time-step algorithms to save computing effort in molecular dynamics simulations of proteins. *J Comput Chem*. 2021;42:1263–1282. https://doi.org/10.1002/jcc.26541