Received: 12 April 2021 | Revised: 27 August 2021 | Accepted: 21 September 2021

DOI: 10.1002/prot.26251

RESEARCH ARTICLE

PROTEINS WILEY

On the use of intra-molecular distance and angle constraints to lengthen the time step in molecular and stochastic dynamics simulations of proteins

Maria Pechlaner | Wilfred F. van Gunsteren

Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, Zurich, Switzerland

Correspondence
Maria Pechlaner, Laboratory of Physical Chemistry, ETH Zurich, Vladimir-Prelog-Weg 10, 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. An example of such a time-saving technique is the application of constraints to particular degrees of freedom when integrating Newton's or Langevin's equations of motion in molecular dynamics (MD) or stochastic dynamics (SD) simulations, respectively. The application of bond-length constraints is standard practice in protein simulations and allows for a lengthening of the time step by a factor of three. Applying recently proposed algorithms to constrain bond angles or dihedral angles, it is investigated, using the protein trypsin inhibitor as test molecule, whether bond angles and dihedral angles involving hydrogen atoms or even stiff proper (torsional) dihedral angles as well as improper ones (maintaining particular tetrahedral or planar geometries) may be constrained without generating too many artificial side effects. Constraining the relative positions of the hydrogen atoms in the protein allows for a lengthening of the time step by a factor of two. Additionally constraining the improper dihedral angles and the stiff proper (torsional) dihedral angles in the protein does not allow for an increase of the MD or SD time step.

KEYWORDS
classical equations of motion, constraints, molecular dynamics, stochastic dynamics, trypsin inhibitor

1 | INTRODUCTION

The length of the time step Δt in a molecular dynamics 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 or possibly bond-angle or particular torsional-angle vibrations. This then allows for a longer time step Δt. However, although such internal vibrations are often not of primary interest, the application of constrained dynamics only makes sense physically and computationally when[1]

1. the frequencies of the frozen (constrained) degrees of freedom are (considerably) higher than those of the remaining ones, thereby allowing for a (substantial) increase of Δt,
2. the frozen degrees of freedom are only weakly coupled to the remaining ones, that is, when the molecular motion is not significantly affected by the application of the constraints,

3. so-called metric-tensor effects[2–4] play a minor role,
4. the algorithm by which the constraints are imposed on the molecular system does not require excessive computational effort.

Bond-length degrees of freedom in proteins 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,[5] metric-tensor effects are insignificant,[6] and algorithms to impose distance constraints do not require excessive computational effort.[7] A factor of three in computer time can typically be saved by the application of bond-length constraints.[8]

Bond-angle constraints in molecules containing torsional-angle degrees of freedom do not satisfy the above-mentioned four conditions as well as bond-length constraints. Their vibrational frequencies are lower, their motions are less decoupled from other motions in a molecule, and metric-tensor effects can be significant.[6] While small solvent molecules without torsional-angle degrees of freedom are commonly simulated as rigid molecules,[9] in proteins, the effect of bond-angle constraints applied to all bond angles present in the molecule is significant: Flexibility and entropy are halved, and the number of torsional-angle transitions is reduced.[5] Yet, bond angles involving hydrogen atoms might be constrained without too many artificial side effects. In addition, torsional-angle degrees of freedom, proper (torsional) ones as well as improper ones (maintaining particular tetrahedral or planar geometries), may also be constrained in order to remove their motions from the molecular system.

In bio-molecular systems a hierarchy of motional frequencies originating in different types of interatomic interactions can be distinguished.[10–13] In order of decreasing frequency or increasing smoothness of the corresponding atom–atom interaction term in the bio-molecular force field we have (table 2 of Reference 13): (i) Bond-stretching vibrations with an approximate oscillation or relaxation time τ_I < 10 fs for bonds involving a hydrogen atom and τ_I < 20 fs for bonds involving only carbon or heavier atoms. (ii) Bond-angle bending vibrations with τ_{II} < 20 fs for bond angles involving hydrogen atoms and τ_{II} < 40 fs for bond angles involving only carbon or heavier atoms. (iii) Improper dihedral angle vibrations due to force field terms used to impose the proper chirality at chiral CH1 united atoms or to impose planarity on ring structures with conjugated double bonds with τ_{III} < 30 fs. (iv) Torsional-angle vibrations around double bonds (e.g., peptide bonds) with τ_{IV} < 30 fs. (v) Water or solvent librational vibrations with τ_V < 30 fs. (vi) Torsional-angle vibrations around single bonds with τ_{VI} < 40 fs for torsional angles involving a hydrogen atom and τ_{VI} < 80 fs for torsional angles involving only carbon or heavier atoms. (vii) Motions dominated by (not softened) van der Waals contacts and short-range (e.g., hydrogen bonding) Coulomb interactions with τ_{VII} < 150 fs. (viii) Motions dominated by long-range Coulomb (ionic, dielectric) interactions with relaxation time τ_{VIII} < 2000 fs. It is this hierarchy that is exploited to enhance the efficiency of a simulation through the use of longer time steps Δt.[14]

Several methods are available to apply distance constraints during a MD simulation based on equations of motion in Cartesian coordinates,[7,15–18] of which the SHAKE method[7] 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. Bond angles can be constrained in a similar manner using the procedure SHAKEBAC,[19–21] and the dihedral angles accordingly with the procedure SHAKEDAC.[20,22,23]

An alternative to the use of equations of motion in Cartesian coordinates and imposing constraints through Lagrange multipliers[19–23] is to employ internal coordinates, bond lengths, bond angles, and torsional angles.[24–32] The classical equations of motion have been formulated by Lagrange in a most general form

$$\frac{d}{dt}\left(\frac{\partial L(q,\dot{q})}{\partial \dot{q}_i}\right)=\frac{\partial L(q,\dot{q})}{\partial q_i}, \qquad i=1,2,...,N_{df} \tag{2}$$

where q_i denote the generalized coordinates, $\dot{q}_i$ their time derivatives and the Lagrangean $L(q,\dot{q})$ is the kinetic energy $K(\dot{q})$ minus the potential energy $V(q)$ of the system which contains N_{df} degrees of freedom. When using Cartesian coordinates $q \equiv x$, one has $K(\dot{x})=(1/2)m\dot{x}^2$ and Equation (2) reduce to Newton's equations of motion (for N_{df} degrees of freedom).

$$m_i\frac{d^2x_i}{dt^2}=\frac{-\partial V(x_1,x_2,...,x_{N_{df}})}{\partial x_i}, \qquad i=1,2,...,N_{df} \tag{3}$$

For branched polymers, such as proteins, the choice of internal coordinates, bond lengths, bond angles, and torsional angles seems to be natural. However, the equations of classical dynamics (2) expressed in the N_{df} internal, generalized coordinates $q_i=\theta_i$.[13]

$$\sum_{j=1}^{N_{df}} a_{ij}\frac{d^2\theta_i}{dt^2}=\frac{-\partial V(\theta_1,\theta_2,...,\theta_{N_{df}})}{\partial \theta_i}-\sum_{j=1}^{N_{df}} b_{ij}\left(\frac{d\theta_j}{dt}\right)^2 - \sum_{j=1}^{N_{df}}\sum_{k=1}^{N_{df}} c_{ijk}\left(\frac{d\theta_j}{dt}\right)\left(\frac{d\theta_k}{dt}\right),\ i=1,2,...,N_{df} \tag{4}$$

are considerably more complex than when expressed in Cartesian coordinates (Equation 3). They contain two additional summations over the number of degrees of freedom and two additional quadratic (i.e., nonlinear) terms in the generalized velocities, and the coefficients a_{ij}, b_{ij}, and c_{ij} depend on the atomic masses and the molecular topology of the protein considered. Therefore, these equations of motion will not be used here.

In the present article, it is investigated whether the application of bond-angle and dihedral-angle constraints, in combination with the commonly applied bond-length constraints, in a molecular dynamics (MD) or stochastic dynamics (SD) simulation of a protein may allow for a lengthening of the discrete time step Δt used to numerically integrate the equations of motion, thereby keeping the distortive effects of constraining particular degrees of freedom upon the motion of the

nonconstrained degrees of freedom small. In view of the above-mentioned hierarchy of motions in proteins, three sets of constraints were used (see Tables 1 and 2):

1. All bond lengths in the protein (constraint set *BLC*).
2. One bond angle and one improper dihedral angle per hydrogen atom that, in combination with the constrained bond to the hydrogen atom, rigidify the position of the hydrogen atom with respect to its covalently bound neighbor nonhydrogen atoms (constraint set *BADAC[H]*), see Table 1. Stiff torsional dihedral-angle degrees of freedom involving a hydrogen atom are also constrained. For hydrogen atoms in C–O–H groups, in Ser, Thr, and Tyr amino-acid residues, only the O–H bond length and the C–O–H bond angle are constrained, its C–C–O–H dihedral angle is not constrained.
3. All other (not involving hydrogen atoms) improper dihedral angles and the stiff peptide-plane torsional angle and other stiff torsional angles in side chains are constrained (constraint set *ISDAC*), see Table 2.

The small protein bovine pancreatic trypsin inhibitor (BPTI) was chosen as test case, as in an earlier study[5] of the effect of constraining bond-length and bond-angle degrees of freedom in MD simulation. Since the protein model of Reference 5 did not contain explicit hydrogen atoms, the effect of constraining their relative positions was not investigated. The effect of constraining the geometry of the peptide plane and other stiff torsional-angle degrees of freedom was also not investigated.

The protein is simulated in three different ways:

1. In vacuo using MD simulation (*MD_vac*), in order to test the degree of conservation of total energy, linear and rotational momentum[8] of the protein as function of the type of constraint (set *BLC*, set *BLC* + *BADAC(H)*, or set *BLC* + *BADAC(H)* + *ISDAC*) applied and the size of the MD time step Δt. 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 geometric error induced by the application of constraints is governed by tol_{DC} for the bond lengths, by tol_{BAC} for the bond angles, and by tol_{DAC} for the dihedral or torsional angles. Values of $tol_{DC} = 10^{-5}$ and 10^{-4}, and $tol_{BAC} = tol_{DAC} = 0.001^\circ$ and 0.01° were used 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. Solvated in water using MD simulation (*MD_aq*), that is, in a rectangular periodic box with minimum image periodic boundary conditions, with many (6295) explicit rigid water molecules, the standard way of protein simulation, in order to test the behavior of dynamical protein properties as function of the type of constraint (set *BLC*, set *BLC* + *BADAC(H)*, or set *BLC* + *BADAC(H)* + *ISDAC*) applied and the size of the MD time step Δt. Again, values of $tol_{DC} = 10^{-4}$ and 10^{-5} were used for the bond lengths and $tol_{BAC} = tol_{DAC} = 0.01^\circ$ and 0.001° for the bond angles and dihedral angles when solving the constraint equations. 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.
3. In vacuo using SD simulation (*SD_vac*), that is, replacing explicit water molecules by frictional and stochastic forces on the protein atoms, which roughly mimic the solvation effect[33,34] in order to test the behavior of various protein properties as function of the type of constraint (set *BLC*, set *BLC* + *BADAC(H)*, or set *BLC* + *BADAC(H)* + *ISDAC*) applied and the size of the SD time step Δt. This type of simulation may be used for very large protein complexes, such as the nucleosome, which would require the presence of very many water molecules in the system to fill a periodic box containing the complex.

2 | METHODS

2.1 | Potential energy function or force field

When solvated in water (*MD_aq*), the protein was modeled using the GROMOS bio-molecular force field 54A7.[35,36] The rigid simple point charge (SPC) model[37] was used to describe the 6295 water molecules in the rectangular periodic box. When simulating the protein in vacuo (*MD_vac*, *SD_vac*), the GROMOS bio-molecular force field 54B7[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.

When applying bond-length constraints, that is, constraint set *BLC*, the corresponding bond-stretching terms of the force field are not evaluated.

When applying the bond-angle and dihedral-angle constraint set *BADAC(H)*, the bond-angle bending terms of the force field are not evaluated for *all* bond angles involving a hydrogen atom, and the improper dihedral-angle terms and the *stiff* torsional dihedral-angle terms of the force field are not evaluated for the dihedral angles involving a hydrogen atom. A stiff proper (torsional) dihedral angle has a force constant $k^{(\varphi)} > 15$ kJ/mole in the cosine force-field term.[38,39] We note that the force field contains more bond-angle terms involving hydrogen atoms than bond angles involving a hydrogen atom that are constrained. The force field also contains more covalent interaction terms (bond-stretching, bond-angle bending, dihedral-angle bending) force-field terms than the corresponding number of degrees of freedom. This implies that the number of constraints may be less than the number of force-field terms corresponding to the constraints that must not be evaluated when the constraints are applied. Each hydrogen atom should have not more than three constraints, for example,

TABLE 1 Bond angles and dihedral angles involving hydrogen atoms in the various amino-acid residues that define the set of constraints *BADAC(H)*

Amino-acid residue	Hydrogen atom	Constrained bond angle	Value constraint θ^0 (°)	Constrained dihedral angle	Value constraint ξ^0 or φ^0 (°)
N-terminus	H1	H1-N-CA	109.5		
	H1, H2	H1-N-H2	109.5		
	H2	H2-N-CA	109.5		
	H2, H3	H2-N-H3	109.5		
	H3	H3-N-CA	109.5		
Backbone (not Pro)	H	H-N-CA	115.0	N-C(-1)-CA-H	0.0
Arg	HE	HE-NE-CD	116.0	NE-CD-CZ-HE	0.0
	HH11	HH11-NH1-CZ	120.0	NE-CZ-NH1-HH11	0.0 or 180.0
	HH12	HH12-NH1-CZ	120.0	NH1-HH11-HH12-CZ	0.0
	HH21	HH21-NH2-CZ	120.0	NE-CZ-NH2-HH21	0.0 or 180.0
	HH22	HH22-NH2-CZ	120.0	NH2-HH21-HH22-CZ	0.0
Asn	HD21	HD21-ND2-CG	120.0	ND2-HD21-HD22-CG	0.0
	HD22	HD22-ND2-CG	120.0	CB-CG-ND2-HD21	0.0 or 180.0
Gln	HE21	HE21-NE2-CD	120.0	NE2-HE21-HE22-CD	0.0
	HE22	HE22-NE2-CD	120.0	CG-CD-NE2-HE21	0.0 or 180.0
Hisa	HD1	HD1-ND1-CG	126.0	ND1-CG-CE1-HD1	0.0
	HD2	HD2-CD2-CG	126.0	CD2-CG-NE2-HD2	0.0
	HE1	HE1-CE1-ND1	126.0	CE1-ND1-NE2-HE1	0.0
Hisb	HD2	HD2-CD2-CG	126.0	CD2-CG-NE2-HD2	0.0
	HE1	HE1-CE1-ND1	126.0	CE1-ND1-NE2-HE1	0.0
	HE2	HE2-NE2-CD2	126.0	NE2-CD2-CE1-HE2	0.0
Hish	HD1	HD1-ND1-CG	126.0	ND1-CG-CE1-HD1	0.0
	HD2	HD2-CD2-CG	126.0	CD2-CG-NE2-HD2	0.0
	HE1	HE1-CE1-ND1	126.0	CE1-ND1-NE2-HE1	0.0
	HE2	HE2-NE2-CD2	126.0	NE2-CD2-CE1-HE2	0.0
Lysh	HZ1	HZ1-NZ-CE	109.5		
	HZ1, HZ2	HZ1-NZ-HZ2	109.5		
	HZ2	HZ2-NZ-CE	109.5		
	HZ2, HZ3	HZ2-NZ-HZ3	109.5		
	HZ3	HZ3-NZ-CE	109.5		
Phe	HD1	HD1-CD1-CG	120.0	CD1-CG-CE1-HD1	0.0
	HD2	HD2-CD2-CG	120.0	CD2-CG-CE2-HD2	0.0
	HE1	HE1-CE1-CD1	120.0	CE1-CD1-CZ-HE1	0.0
	HE2	HE2-CE2-CD2	120.0	CE2-CD2-CZ-HE2	0.0
	HZ	HZ-CZ-CE1	120.0	CZ-CE1-CE2-HZ	0.0
Ser	HG	HG-OG-CB	109.5		
Thr	HG1	HG1-OG1-CB	109.5		
Trp	HD1	HD1-CD1-CG	126.0	CD1-CG-NE1-HD1	0.0
	HE1	HE1-NE1CD1	126.0	NE1-CD1-CE2-HE1	0.0
	HE3	HE3-CE3-CD2	120.0	CE3-CD2-CZ3-HE3	0.0
	HZ2	HZ2-CZ2-CE2	120.0	CZ2-CE2-CH2-HZ2	0.0
	HZ3	HZ3-CZ3-CE3	120.0	CZ3-CE3-CH2-HZ3	0.0
	HH2	HH2-CH2-CZ2	120.0	CH2-CZ2-CZ3-HH2	0.0

TABLE 1 (Continued)

Amino-acid residue	Hydrogen atom	Constrained bond angle	Value constraint θ^0 (°)	Constrained dihedral angle	Value constraint ξ^0 or φ^0 (°)
Tyr	HD1	HD1-CD1-CG	120.0	CD1-CG-CE1-HD1	0.0
	HD2	HD2-CD2-CG	120.0	CD2-CG-CE2-HD2	0.0
	HE1	HE1-CE1-CD1	120.0	CE1-CD1-CZ-HE1	0.0
	HE2	HE2-CE2-CD2	120.0	CE2-CD2-CZ-HE2	0.0
	HH	HH-OH-CZ	109.5		

Note: Values for the constraints were taken from the GROMOS force field 54A7.[36,38] An atom of the preceding residue in the polypeptide chain is indicated by (−1). The values of the improper dihedral angle ξ^0 are 0° or 35.26439°. The values for the stiff (force constant >15 kJ/mol in the cosine force field term) proper (torsional) dihedral angle φ^0 are 0° or 180° for planar groups.

one bond-length, one bond-angle and one (improper) dihedral-angle constraint. For hydrogens in a –C–O–H group, there are only two constraints, the bond length and the bond angle. For hydrogens in a –C–N–H_3 group, there are only three bond-length and five bond-angle constraints.

When applying the improper and proper (torsional) angle constraint set *ISDAC*, the corresponding improper and stiff torsional dihedral-angle bending terms of the force field must not be evaluated. Again, the force field contains more dihedral bending force-field terms than the corresponding number of constrained degrees of freedom.

2.2 | Constraints

The implementation into simulation software of the constraints and the omission of particular force-field terms is not trivial. In the GROMOS simulation software[40] use of the bond-length constraint set *BLC* is simple. The constraints correspond exactly with the bond-stretching force-field terms, and thus the constraints are taken from these force-field terms. In the GROMOS simulation software, each type of force-field term can be switched on or off. So when applying the *BLC* constraints, the bond-stretching force-field terms are switched off.

The use of the bond-angle and dihedral-angle constraints set *BADAC(H)* is less straightforward. In that case, the constraints of Table 1 are used in the form of a constraint list for the functions SHAKEBAC[21] and SHAKEDAC[23] that constrain bond angles and dihedral angles, respectively. In addition, all bond-angle terms and improper dihedral-angle terms involving hydrogen atoms are switched off. For the proper (torsional) dihedral angles involving hydrogens, this cannot be done, because only the stiff ones should not be evaluated. So the stiff dihedral angles involving hydrogen atoms should be removed either from the protein molecular topology or from the amino-acid residue building blocks[38] used to construct the protein molecular topology.

The use of the improper and stiff proper (torsional) dihedral-angle constraints set *ISDAC* is also less straightforward. In that case, the constraints of Table 2 are used in the form of a constraint list for the function SHAKEDAC. For some dihedral angles the choice of 0° or 180° as reference value for the constraint will depend on the actual value of the dihedral angle in the protein structure used in the simulation. In addition, all improper dihedral-angle terms not involving hydrogen atoms are switched off. For the proper (torsional) dihedral angles not involving hydrogens, this cannot be done, because only the stiff ones should not be evaluated. So the stiff dihedral angles not involving hydrogen atoms should be removed either from the protein molecular topology or from the amino-acid residue building blocks[38] used to construct the protein molecular topology.

The bond lengths in the protein and water molecules and the bond-angle distance of the water molecules were kept rigid with a relative geometric precision of $tol_{DC} = 10^{-4}$ or 10^{-5} using the SHAKE algorithm.[7] The bond angles and the dihedral angles in the protein were kept rigid with a precision of $tol_{BAC} = 0.01°$ or 0.001° using the SHAKEBAC algorithm[21] and $tol_{DAC} = 0.01°$ or 0.001° using the SHAKEDAC algorithm,[23] respectively. The procedure to determine the Lagrange multipliers that determine the constraints may fail (i) because reference vectors (distances) and vectors being modified to satisfy the constraint distances or angles may become orthogonal (i.e., the atoms are moved much due to large forces) or (ii) because no convergence of the iterations over all constraints can be obtained within a set limit of 1000 iterations.

2.3 | Treatment of long-ranged interactions

When simulating the protein in vacuo using MD (*MD_vac*), the nonbonded interactions are calculated for all atom pairs in the protein. When simulating the protein solvated in a periodic box filled with water molecules (*MD_aq*), the nonbonded interactions were calculated using a triple-range method[41,42] with cut-off radii of 0.8/1.4 nm. Short-range (within 0.8 nm) van der Waals and electrostatic interactions were evaluated every time step based on a charge-group pair list.[39] Medium-range van der Waals and electrostatic interactions, between pairs at a distance larger than 0.8 nm and shorter than 1.4 nm, were evaluated every 10 fs, at which time point the pair list was updated, and kept constant between updates. Outside the larger cut-off radius (1.4 nm) a reaction-field approximation[43,44] 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 was $\varepsilon_{cs} = 1$.[45] Minimum-image periodic boundary conditions

TABLE 2 Improper and proper (torsional) dihedral angles not involving hydrogen atoms in the various amino-acid residues that define the set of constraints *ISDAC*

Amino-acid residue	Constrained dihedral angle	Value constraint ξ^0 or φ^0 (°)
Backbone	C-CA-N(+1)-O	0.0
	CA(−1)-C(−1)-N-CA	0.0 or 180.0
Backbone (not Gly)	CA-N-C-CB	35.26
Arg	CZ-NH1-NH2-NE	0.0
	CD-NE-CZ-NH1	0.0 or 180.0
Asn	CG-OD1-ND2-CB	0.0
Asp	CG-OD1-OD2-CB	0.0
Cys1	CB(1)-SG(1)-SG(2)-CB(2)	90.0 or 270.0
Gln	CD-OE1-NE2-CG	0.0
Glu	CD-OE1-OE2-CG	0.0
Hisa/b/h	CG-ND1-CD2-CB	0.0
	ND1-CG-CD2-NE2	0.0
	CD2-CG-ND1-CE1	0.0
Ile	CB-CG1-CG2-CA	35.26
Leu	CB-CD1-CD2-CG	35.26
Phe	CG-CD1-CD2-CB	0.0
	CD1-CG-CD2-CE2	0.0
	CD2-CG-CD1-CE1	0.0
	CG-CD1-CE1-CZ	0.0
Thr	CB-OG1-CG2-CA	35.26
Trp	CG-CD1-CD2-CB	0.0
	CD1-CG-CD2-CE2	0.0
	CD2-CG-CD1-NE1	0.0
	CD2-CE2-CE3-CG	0.0
	CE2-CD2-CZ2-NE1	0.0
	CD2-CE2-CZ2-CH2	0.0
	CE2-CD2-CE3-CZ3	0.0
Tyr	CG-CD1-CD2-CB	0.0
	CD1-CG-CD2-CE2	0.0
	CD2-CG-CD1-CE1	0.0
	CG-CD1-CE1-CZ	0.0
	CZ-CE1-CE2-OH	0.0
Val	CA-CG1-CG2-CB	35.26
C-terminus	C-O1-O2-CA	0.0

Note: Values for the constraints were taken from the GROMOS force field 54A7.[36,38] An atom of the next residue in the polypeptide chain is indicated by (+1), of the preceding residue by (−1). Cys1: first Cys residue that forms a S-S bridge with residue Cys2. The atoms of Cys1 and Cys2 are indicated with (1) and (2), respectively. The values of the improper dihedral angle ξ^0 are 0° or 35.26439°. The values for the stiff (force constant >15 kJ/mol in the cosine force field term) proper (torsional) dihedral angle φ^0 are either 0° or 180° for planar groups, and 90° or 270° for the S–S bridge.

were applied. The triple-range method was also used in the SD simulations in vacuo (*SD_vac*) in order to keep the conditions as close as possible to the ones for the explicitly solvated protein.

2.4 | Simulation set-up and equilibration

The protein was simulated using the GROMOS bio-molecular simulation software.[40,46] The leap-frog algorithm[47,48] was used to integrate Newton's or Langevin's equations of motion.

The initial structure of the protein BPTI, including four internally hydrogen bonded water molecules, was taken from the Protein Data Bank (PDB),[49] entry *1bpi*.

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 and SD simulations in vacuo (*MD_vac*, *SD_vac*). In order to exclude the influence of the fast librational motions of water molecules, the four internal water molecules were not included in these simulations.

The MD simulation of the protein solvated in a periodic box with explicit water molecules required the addition of water molecules. For the simulation of the protein in aqueous solution (*MD_aq*), the protein 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 6295 water molecules for the initial protein structure. 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[40] with a force constant of 25 000 kJmol^{-1} nm^{-2}. The resulting protein-water configuration was used as initial configuration for the MD simulation of the protein solvated in explicit water (*MD_aq*).

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 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 short 20 ps simulations at temperatures 60, 120, 180, 240, and 300 K, at constant volume. At 300 K, the equilibration was extended to 2 ns. 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[50] with a relaxation or coupling time $\tau_T = 0.1$ ps. Solute and solvent, if present, were separately coupled to the heat bath. In the *MD_aq* simulations, the pressure was kept at 1 atm using the weak-coupling algorithm[50] with a coupling time $\tau_p = 0.5$ ps and an isothermal compressibility $\kappa_T = 4.575\ 10^{-4}$ (kJ mol^{-1} nm^{-3})$^{-1}$.

Following this equilibration procedure, the simulations were performed at a reference temperature of 300 K and, if in aqueous solution (*MD_aq*), a reference pressure of 1 atm. The center of mass motion of the system was removed after equilibration of the system

and then, for the simulations of the protein in explicit water (*MD_aq*) and using SD in vacuo (*SD_vac*), every 2 ps.

After the equilibration of the protein in vacuo, in the MD simulations in vacuo (*MD_vac*) the coupling to the heat bath was removed and translation of and rotation around the center of mass of the protein (if present) were not removed anymore in order to avoid perturbation of total energy conservation.

After the equilibration of the protein in vacuo, in the SD simulations in vacuo (*SD_vac*) the temperature is maintained by the Langevin equations or thermostat, and by weak coupling to a heat bath ($\tau_T = 0.1$ ps), the latter in order to control the temperature of atoms that have a friction coefficient equal to zero, whose temperature is thus not controlled by the Langevin thermostat.

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.

2.5 | MD simulation 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 algorithms to impose the constraints (tol_{DC}, tol_{BAC}, and tol_{DAC}), and the size of the MD time step Δt. This is investigated using MD 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, bond-angle and dihedral-angle constraints was made very small by requiring precisions $tol_{DC} = 10^{-5}$, $tol_{BAC} = 0.001^\circ$, and $tol_{DAC} = 0.001^\circ$. 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. 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 of 1, 2, 4, 6, and 8 fs were tested.

2.6 | MD simulation 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. These simulations were performed for 10 ns with time steps of 2, 4, and 6 fs and the constraint precisions $tol_{DC} = 10^{-4}$ or 10^{-5}, tol_{BAC}, $= 0.01^\circ$ or 0.001°, and $tol_{DAC} = 0.01^\circ$ or 0.001°. Translational motion of the center of mass of the system was removed every 2 ps. Configurations were saved every 0.1 ps and were used to analyze various dynamical properties as function of the type of constraints applied.

2.7 | SD simulation in vacuo without explicit water, but with frictional and stochastic forces

In stochastic dynamics (SD) simulation the Langevin equations of motion in Cartesian coordinates

$$m_i \frac{d^2 x_i}{dt^2} = \frac{-\partial V\left(x_1, x_2, \ldots, x_{N_{df}}\right)}{\partial x_i} + f_i^{stoch} - m_i \gamma_i v_i, \qquad i = 1, 2, \ldots, N_{df} \tag{5}$$

are integrated forward in time. The atomic velocities are indicated by v_i. The stochastic force $f_i^{stoch}(t)$ and the atomic friction coefficient γ_i will only be sizable for protein atoms at the surface. Therefore, they are taken dependent on the number of neighbor atoms.[33]

$$\gamma_i(t) = \gamma_{solv}\, \omega_i(t) \tag{6}$$

with

$$\omega_i(t) = \max\left(0, 1 - N_i^{nb}(t)/N^{nbref}\right) \tag{7}$$

where $N_i^{nb}(t)$ denotes the number of nonhydrogen neighbor atoms of the protein atom i within 0.3 nm radius, and N^{nbref} was defined as an upper limit of six neighbor protein atoms at which solvent forces on solute atom i are assumed to vanish. For water as solvent $\gamma_{solv} = 91$ ps^{-1}, and $\omega_i(t)$ was updated every 1 ps during the simulation.[33] The SD simulations were performed for 10 ns with time steps of 2, 4, and 6 fs and the constraint precisions $tol_{DC} = 10^{-4}$ or 10^{-5}, tol_{BAC}, $= 0.01^\circ$ or 0.001°, and $tol_{DAC} = 0.01^\circ$ or 0.001°. Translational motion of the center of mass of the system was removed every 2 ps. Configurations were saved every 0.1 ps and were used to analyze various dynamical properties as function of the type of constraints applied.

2.8 | 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 (*RMS*) fluctuation of an energy E is defined as

$$\Delta E \equiv \left(< (E(t) - < E >_t)^2 >_t \right)^{1/2}, \tag{8}$$

TABLE 3 Energy conservation in 100 ps MD simulations of the protein BPTI in vacuo (*MD_vac*) using three different sets of constraints, set *BLC*, set *BADAC(H)*, and set *ISDAC*

Constraints	tol_{DC}	$tol_{BAC} = tol_{DAC}$ (°)	Δt (fs)	E_{tot}	ΔE_{tot}	E_{tot}^{drift}	ΔE_{tot}^{drift}	E_{kin}	ΔE_{kin}	E_{kin}^{drift}	ΔE_{kin}^{drift}	T (K)	$\Delta E_{tot}/\Delta E_{kin}$	$\Delta E_{tot}^{drift}/\Delta E_{kin}^{drift}$
BLC	10^{-5}	–	1	−1493	15	−0.52	0.60	1471	45	−0.22	44	298	0.34	0.014
			2	−1511	23	−0.79	2.43	1466	44	−0.32	43	297	0.52	0.056
			4	−1485	36	−1.20	10.26	1466	43	−0.50	41	297	0.84	0.253
			6,8	*	*	*	*	*	*	*	*	*	*	*
BLC + *BADAC (H)*		0.001	1	−2116	5	−0.18	0.17	1152	40	0.11	40	308	0.13	0.004
			2	−2155	7	−0.26	0.51	1114	39	−0.24	38	298	0.19	0.013
			4	−2202	11	−0.36	2.22	1116	38	−0.08	38	298	0.28	0.058
			6	−2185	14	−0.47	4.76	1106	36	−0.27	36	295	0.40	0.134
			8	−2106	16	−0.45	9.05	1136	35	−0.05	35	303	0.46	0.260
		0.01	1	−2200	8	−0.27	0.24	1107	39	−0.17	39	295	0.20	0.006
			2	−2163	16	−0.55	0.59	1128	39	−0.23	38	301	0.41	0.016
			4	−2192	18	−0.61	2.17	1110	37	−0.26	37	296	0.47	0.059
			6	−2214	27	−0.91	5.03	1107	38	−0.50	35	296	0.70	0.142
			8	−2156	15	−0.43	9.32	1136	34	−0.02	34	303	0.46	0.276
BLC + *BADAC (H)* + *ISDAC*		0.001	1–8	#	#	#	#	#	#	#	#	#	#	#
		0.01	1	#	#	#	#	#	#	#	#	#	#	#
			2	−2820	138	−4.79	3.15	724	68	−2.12	29	261	2.04	0.108
			4	−2913	164	−5.66	7.63	680	77	−2.49	27	245	2.13	0.284
			6	−2947	198	−6.86	7.72	666	90	−3.00	26	240	2.20	0.298
			8	−2993	219	−7.58	13.47	639	103	−3.46	25	230	2.13	0.534
BLC	10^{-4}	–	1	−1756	169	−5.85	4.64	1339	84	−2.58	40	271	2.00	0.117
			2	−1893	249	−8.63	14.64	1281	121	−3.98	37	259	2.06	0.394
			4	−2031	302	−10.45	18.27	1216	140	−4.70	35	246	2.16	0.529
			6,8	*	*	*	*	*	*	*	*	*	*	*
BLC + *BADAC (H)*		0.001	1	−2166	41	−1.42	0.57	1122	41	−0.52	38	300	0.99	0.015
			2	−2331	93	−3.23	1.90	1053	55	−1.44	36	281	1.70	0.053
			4	−2379	137	−4.73	5.02	1017	75	−2.31	34	271	1.83	0.149
			6	−2423	164	−5.68	7.95	995	82	−2.61	31	266	2.01	0.253
			8	−2454	186	−6.43	10.28	997	89	−2.89	31	266	2.09	0.333
		0.01	1	−2226	39	−1.34	1.10	1094	42	−0.64	38	292	0.92	0.029
			2	−2363	95	−3.28	0.89	1033	59	−1.61	36	276	1.61	0.025
			4	−2441	138	−4.79	4.40	985	75	−2.32	33	263	1.85	0.134
			6	−2444	179	−6.19	8.44	988	90	−2.92	33	264	1.98	0.259
			8	−2489	179	−6.21	10.40	977	88	−2.89	29	261	2.03	0.358
BLC + *BADAC (H)* + *ISDAC*		0.001	1–8	#	#	#	#	#	#	#	#	#	#	#
		0.01	1	#	#	#	#	#	#	#	#	#	#	#
			2	−2919	172	−5.95	5.68	674	81	−2.62	27	243	2.13	0.207
			4	−3017	239	−8.25	17.33	625	116	−3.92	27	225	2.06	0.654
			6	−3103	273	−9.41	24.94	595	124	−4.22	24	214	2.20	1.020
			8	−3122	303	−10.45	28.48	589	140	−4.78	24	212	2.17	1.195

Note: The bond-length constraints are imposed with a relative geometric precision of tol_{DC}, the bond-angle constraints with a precision of tol_{BAC}, and the dihedral-angle constraints with a precision of tol_{DAC}. *, no constraint solution, vectors orthogonal; #, no constraint solution, no convergence; Nonbonded interaction cut-off $R_{cp} = R_{cl}$, 100 nm (infinity); Δt, leap-frog integration time step; 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 0.02 ps. Energies in kJ mol^{-1}. Energy drifts in kJ mol^{-1} ps^{-1}.

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 $E^{\text{line}}(t)$ that is least-squares fitted to $E(t)$ for a chosen period of time. The quantity

$$\Delta E^{\text{drift}} \equiv \left(< \left(E(t) - E^{\text{line}}(t) \right)^2 >_t \right)^{1/2} \quad (9)$$

represents the deviation of the actual energy from the line, $E^{\text{line}}(t)$, 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 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 bond angles and of torsional angles.[5] From a time series of a quantity $Q(t)$, a normalized time correlation function,

$$C_Q(t) = \frac{< Q(\tau) \cdot Q(\tau + t) >_\tau}{< Q(\tau) \cdot Q(\tau) >_\tau} \quad (10)$$

was calculated using the Fast Fourier Transform technique.[51,52] For these calculations, 25 ps towards 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.

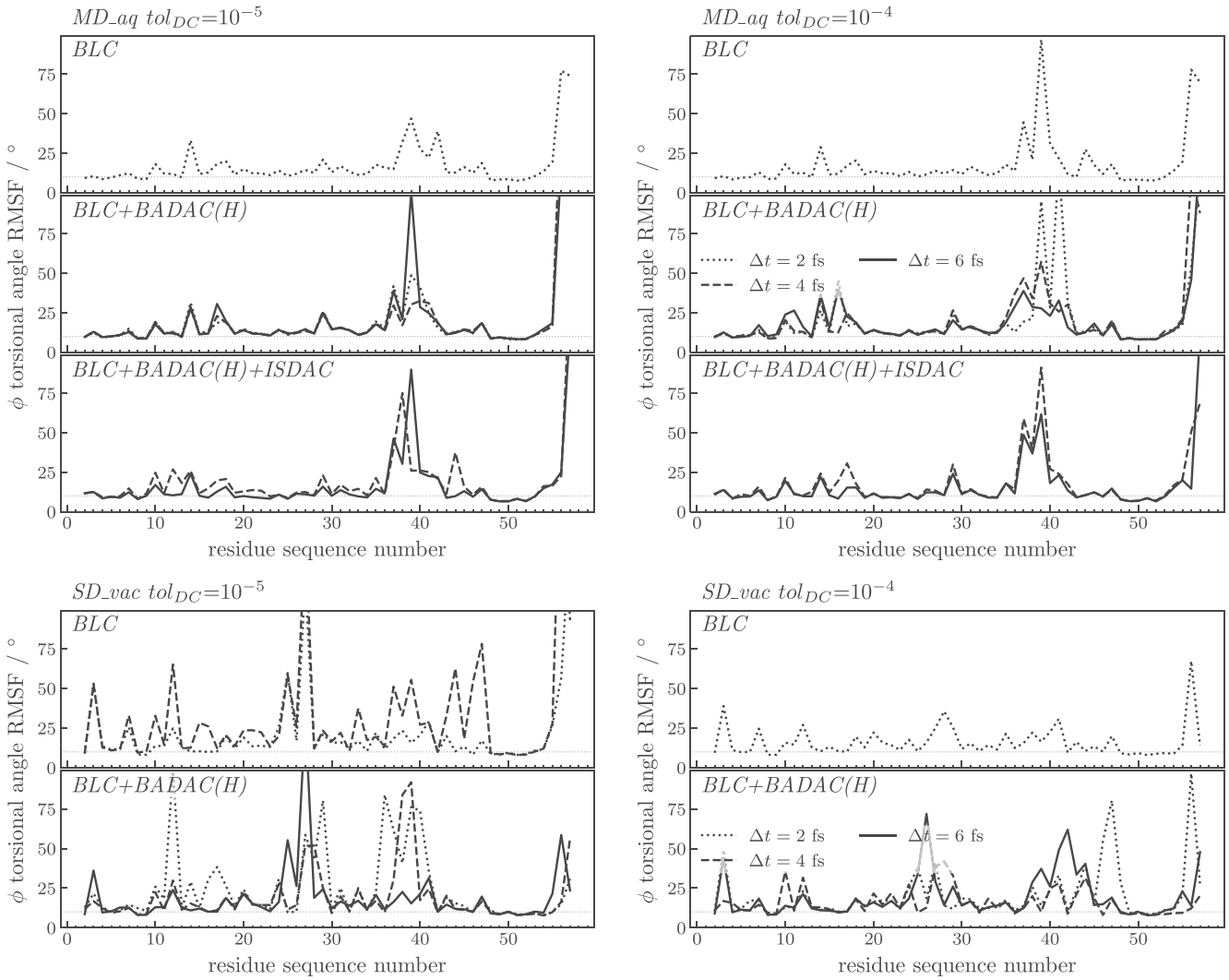

FIGURE 1 Root-mean-square fluctuations (in degree) of the φ-angles in the backbone of the protein BPTI as function of residue number in 10 ns MD simulations solvated in (SPC) water (*MD_aq*) and from 10 ns SD simulations in vacuo (*SD_vac*) using three different sets of constraints, set *BLC*, set *BLC* + *BADAC(H)*, and set *BLC* + *BADAC(H)* + *ISDAC*, and for different time steps Δt. Upper three panels: simulations *MD_aq*. Lower two panels: simulations *SD_vac*. Left panels: $tol_{DC} = 10^{-5}$. Right panels: $tol_{DC} = 10^{-4}$. $tol_{BAC} = tol_{DAC} = 0.01°$. Dotted lines: $\Delta t = 2$ fs. Dashed lines: $\Delta t = 4$ fs. Solid lines: $\Delta t = 6$ fs

3 | RESULTS

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

In Table 3, the average total and kinetic energy and their fluctuations are shown for the protein in vacuo (*MD_vac*) using various time steps and different values of the geometric precision by which the bond-length (tol_{DC}), bond-angle (tol_{BAC}), and dihedral-angle (tol_{DAC}) constraints are maintained. Constraining all bond lengths (*BLC*) in the protein with a relative precision of $tol_{DC} = 10^{-5}$, the ratio $\Delta E_{\text{tot}}^{\text{drift}}/\Delta E_{\text{kin}}^{\text{drift}}$ of the fluctuation of the total energy around the total energy drift ($\Delta E_{\text{tot}}^{\text{drift}}$) and the fluctuation of the kinetic energy around the kinetic energy drift ($\Delta E_{\text{kin}}^{\text{drift}}$) changes from 0.014 for a time step $\Delta t = 1$ fs to 0.253 for a time step $\Delta t = 4$ fs. Using a lower precision of maintaining the bond-length constraints, $tol_{DC} = 10^{-4}$, the ratio $\Delta E_{\text{tot}}^{\text{drift}}/\Delta E_{\text{kin}}^{\text{drift}}$ changes from 0.117 for a time step $\Delta t = 1$ fs to 0.529 for a time step $\Delta t = 4$ fs. So for these time steps, the precision of maintaining the bond-length constraints governs total energy conservation. Yet, using $tol_{DC} = 10^{-5}$ and $\Delta t = 4$ fs, the ratio $\Delta E_{\text{tot}}^{\text{drift}}/\Delta E_{\text{kin}}^{\text{drift}}$ is larger, 0.253, than using $tol_{DC} = 10^{-4}$ and $\Delta t = 1$ fs. For larger time steps, the integration error dominates the constraint error. Requiring a $\Delta E_{\text{tot}}^{\text{drift}}/\Delta E_{\text{kin}}^{\text{drift}}$ ratio of about 1/20 for total energy conservation, use of $tol_{DC} = 10^{-5}$ and $\Delta t = 2$ fs leading to a ratio of 0.056 seems possible. We note that in[8] it was found that $tol_{DC} = 10^{-4}$ and $\Delta t = 2$ fs would lead to sufficient energy conservation. Since the protein BPTI in[8] was modeled using only united atoms, that is, without explicit hydrogens, the higher-frequency bond-length motions of the light hydrogen atoms were missing in the simulation, which allowed a lower precision for the bond-length constraints to be used for a given total energy conservation value.

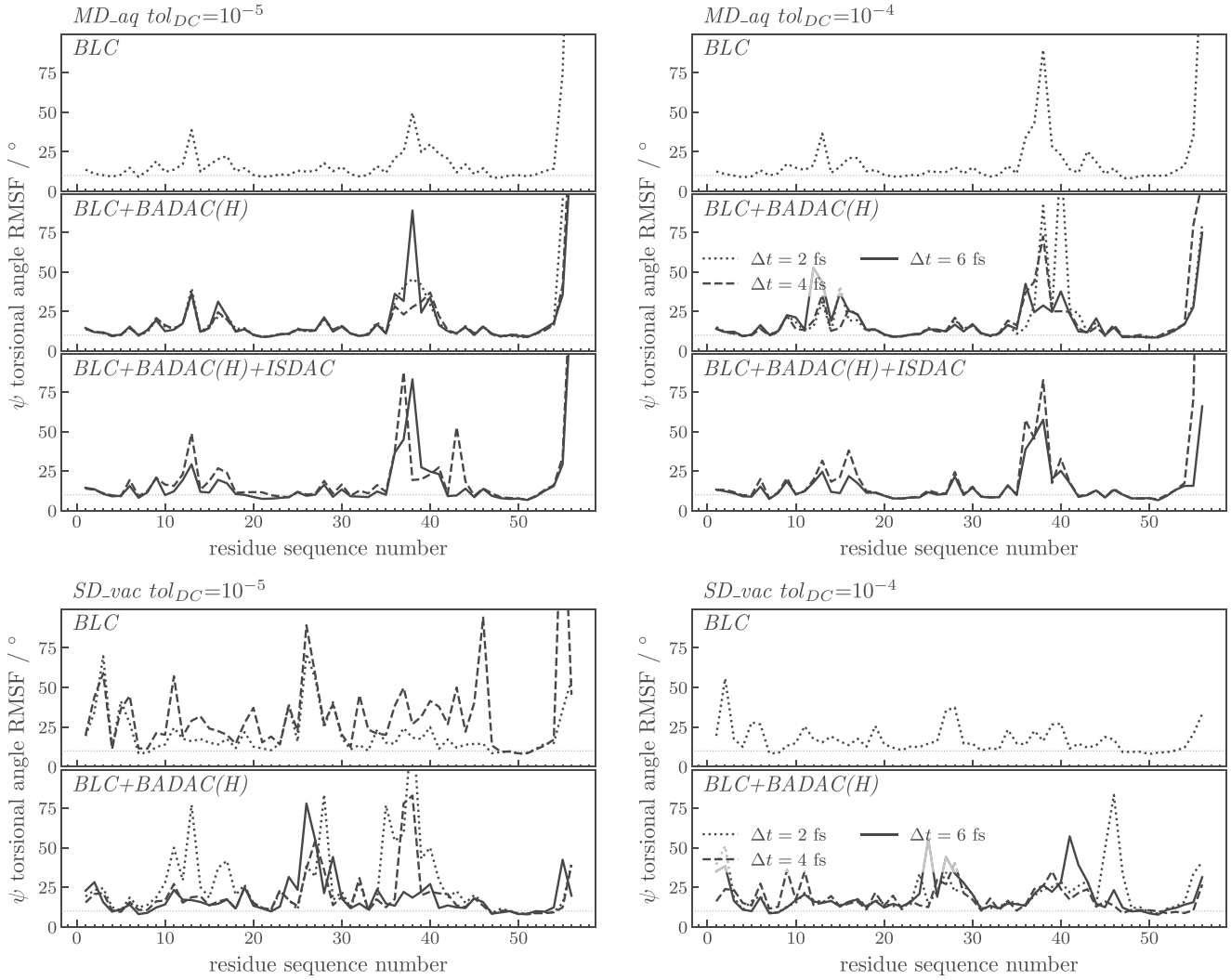

FIGURE 2 Root-mean-square fluctuations (in degree) of the ψ-angles in the backbone of the protein BPTI as function of residue number in 10 ns MD simulations solvated in (SPC) water (*MD_aq*) and from 10 ns SD simulations in vacuo (*SD_vac*) using three different sets of constraints, set *BLC*, set *BLC* + *BADAC(H)*, and set *BLC* + *BADAC(H)* + *ISDAC*, and for different time steps Δt. Upper three panels: simulations *MD_aq*. Lower two panels: simulations *SD_vac*. Left panels: $tol_{DC} = 10^{-5}$. Right panels: $tol_{DC} = 10^{-4}$. $tol_{BAC} = tol_{DAC} = 0.01°$. Dotted lines: $\Delta t = 2$ fs. Dashed lines: $\Delta t = 4$ fs. Solid lines: $\Delta t = 6$ fs

When also constraining the relative positions of the hydrogen atoms, using constraint set *BADAC(H)* in addition to set *BLC*, the ratio $\Delta E_{\text{tot}}^{\text{drift}}/\Delta E_{\text{kin}}^{\text{drift}}$ becomes much smaller. Using a high precision maintaining the constraints, $tol_{DC} = 10^{-5}$ and $tol_{BAC} = tol_{DAC} = 0.001^\circ$, the ratio $\Delta E_{\text{tot}}^{\text{drift}}/\Delta E_{\text{kin}}^{\text{drift}}$ changes from 0.004 for a time step $\Delta t = 1$ fs to 0.260 for a time step $\Delta t = 8$ fs. Lowering the precision of the bond-angle and dihedral-angle constraints to $tol_{BAC} = tol_{DAC} = 0.01^\circ$, the ratio $\Delta E_{\text{tot}}^{\text{drift}}/\Delta E_{\text{kin}}^{\text{drift}}$ becomes only slightly larger comparing the same time step sizes, except for the small time step $\Delta t = 1$ fs. For a time step $\Delta t = 4$ fs, the ratio $\Delta E_{\text{tot}}^{\text{drift}}/\Delta E_{\text{kin}}^{\text{drift}}$ is 0.059, which value is comparable to the value 0.056 obtained using only bond-length constraints (*BLC*) and $\Delta t = 2$ fs. Lowering the precision of the bond-length constraints from $tol_{DC} = 10^{-5}$ to $tol_{DC} = 10^{-4}$, leads, as expected, to larger $\Delta E_{\text{tot}}^{\text{drift}}/\Delta E_{\text{kin}}^{\text{drift}}$ values, in particular for the smaller time steps. Interestingly, the additional use of a lower precision $tol_{BAC} = tol_{DAC} = 0.01^\circ$ instead of 0.001° leads to a lower ratio $\Delta E_{\text{tot}}^{\text{drift}}/\Delta E_{\text{kin}}^{\text{drift}}$ for $\Delta t = 2$ and 4 fs, and to a comparable ratio for $\Delta t = 6$ fs. Yet using $tol_{DC} = 10^{-4}$, the time step should not be larger than $\Delta t = 2$ fs, with a $\Delta E_{\text{tot}}^{\text{drift}}/\Delta E_{\text{kin}}^{\text{drift}}$ ratio of 0.053 or 0.025 for $tol_{BAC} = tol_{DAC} = 0.001^\circ$ or 0.01°, respectively, because using $\Delta t = 4$ fs, the ratio becomes 0.149 or 0.134.

When also constraining the improper dihedral angles and the stiff proper (torsional) dihedral angles in the protein, using constraint set *ISDAC* in addition to the sets *BADAC(H)* and *BLC*, the ratio $\Delta E_{\text{tot}}^{\text{drift}}/\Delta E_{\text{kin}}^{\text{drift}}$ becomes rather large, ranging from 0.108 to 1.195, for all time steps and constraint precision values tol_{DC}, tol_{BAC} and tol_{DAC} investigated. For $tol_{BAC} = tol_{DAC} = 0.001^\circ$ and $\Delta t = 1$–8 fs and for $tol_{BAC} = tol_{DAC} = 0.01^\circ$ and $\Delta t = 1$ fs the algorithms to maintain the various constraints do not even converge. This is due to the large number of constraints that involve many of the same atoms and thus the

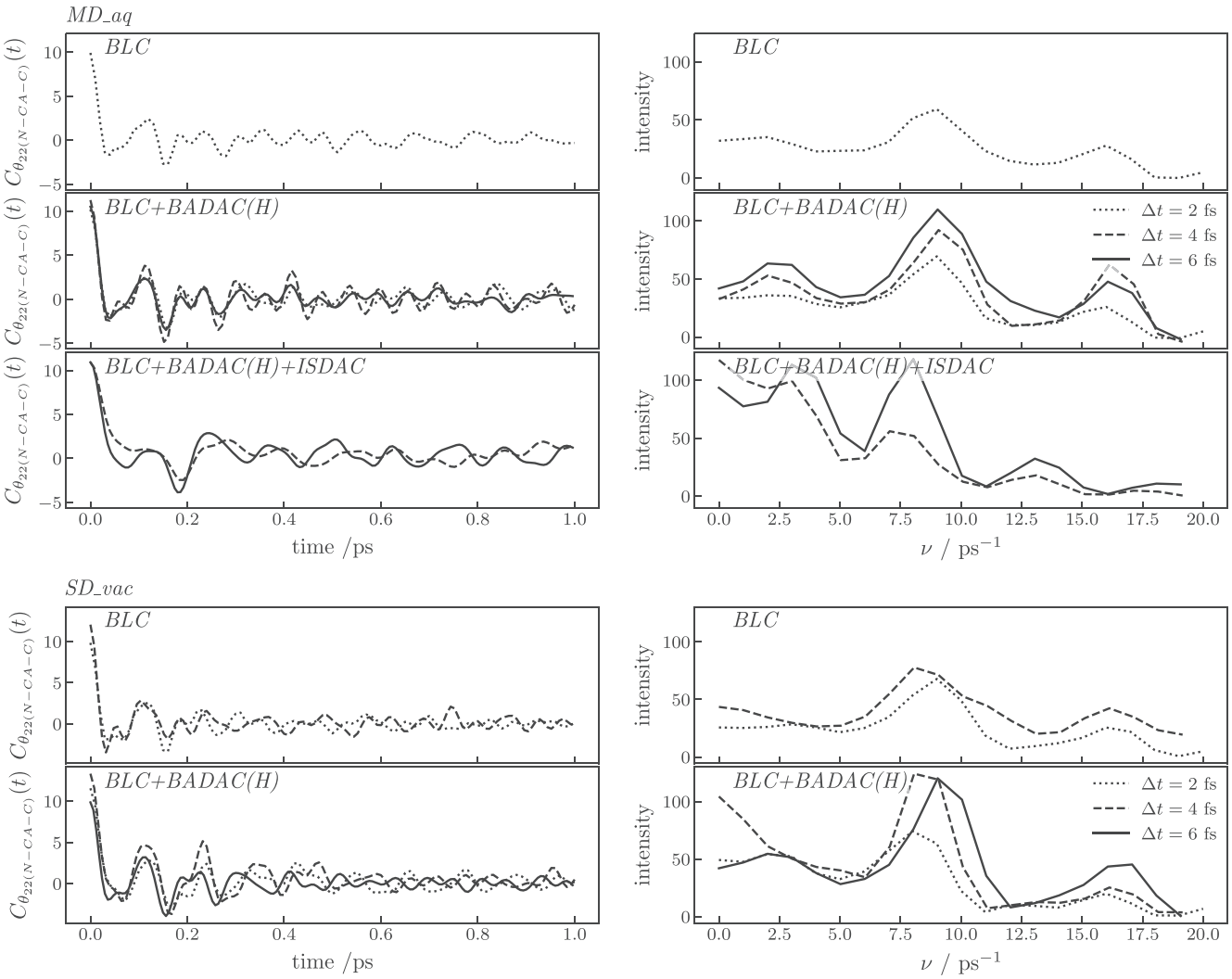

FIGURE 3 Auto-correlation function (left panels) and spectral density (right panels) of the bond angle θ(N-CA-C) of residue Phe 22 in the backbone of the protein BPTI in 10 ns MD simulations solvated in (SPC) water (*MD_aq*) and from 10 ns SD simulations in vacuo (*SD_vac*) using three different sets of constraints, set *BLC*, set *BLC* + *BADAC(H)*, and set *BLC* + *BADAC(H)* + *ISDAC*, and for different time steps Δt. Upper three panels: simulations *MD_aq*. Lower two panels: simulations *SD_vac*. Dotted lines: $\Delta t = 2$ fs. Dashed lines: $\Delta t = 4$ fs. Solid lines: $\Delta t = 6$ fs. $tol_{DC} = 10^{-5}$. $tol_{BAC} = tol_{DAC} = 0.01^\circ$. Configurations from 25 ps towards 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 have trouble to find Lagrange multiplier values that make the molecular structure satisfy all constraints to within the specified precision. So the use of improper dihedral-angle and stiff proper (torsional) dihedral-angle constraints in the protein does not allow an increase of the time step in a MD simulation using Cartesian coordinates.

3.2 | MD simulation of a protein in aqueous solution to test dynamical protein properties

Using only bond-length constraints (set *BLC*), simulations in aqueous solution fail at $\Delta t = 4$ and 6 fs, because vectors become orthogonal during the execution of the SHAKE algorithm, an indication that disruptive forces occur. In simulations with constraint set *ISDAC* with $\Delta t = 2$ fs SHAKE does not converge.

Figures 1 and 2 (upper part) show the root-mean-square fluctuations of the backbone φ and ψ torsional angles as function of residue number applying the different sets of constraints and time steps Δt. The larger peaks are 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. For example, in the simulation with constraint set *BLC*, $tol_{DC} = 10^{-4}$ and $tol_{BAC} = tol_{DAC} = 0.01°$, with $\Delta t = 2$ fs, the peptide plane between residues 39 and 40 changes orientation which induces a correlated change in $\psi(39)$ and $\varphi(40)$.

The application of constraints when integrating the equations of motion would primarily affect the motions along the nonconstrained degrees of freedom that are close or adjacent to the constrained ones. The influence of the different combinations of the three sets of

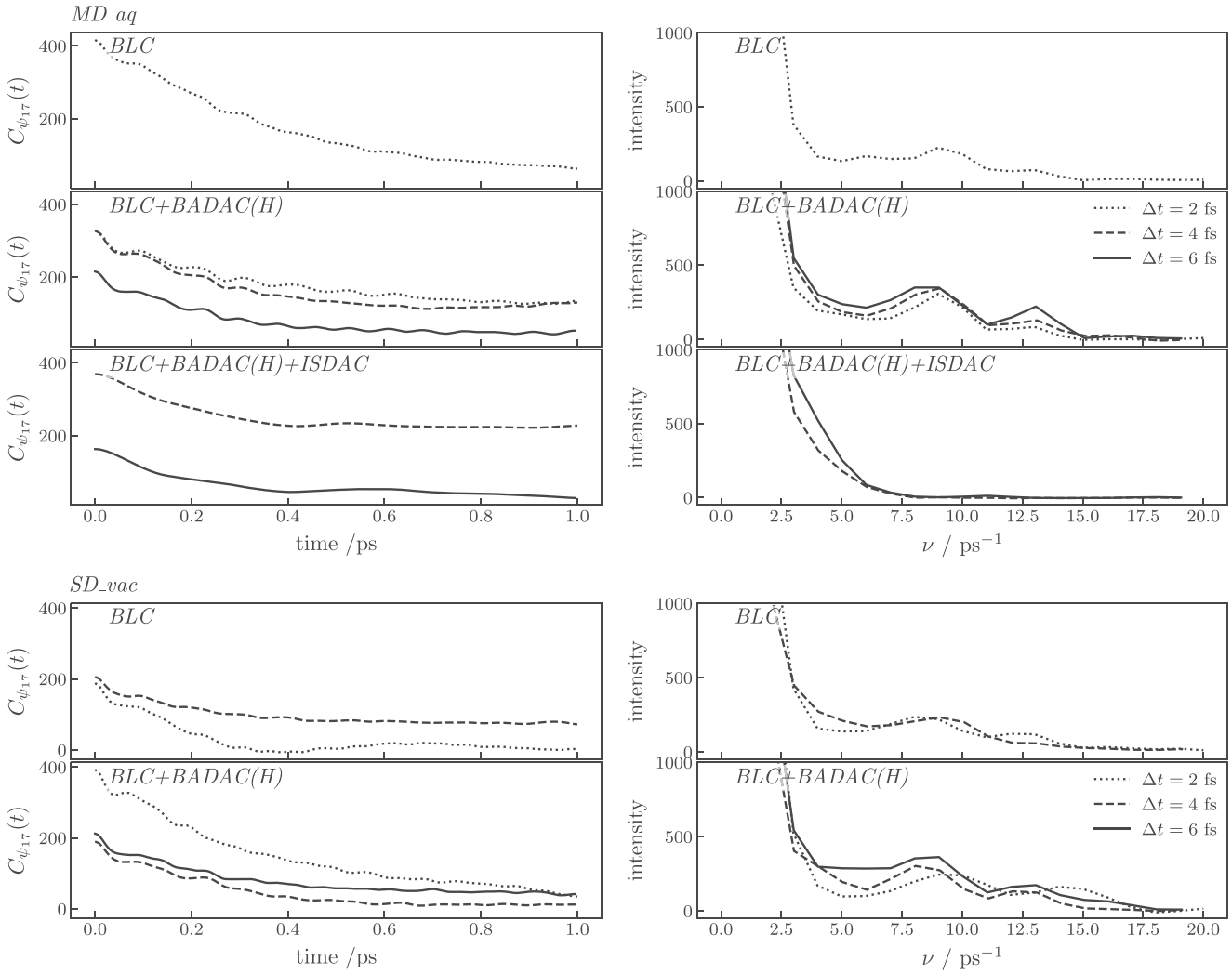

FIGURE 4 Auto-correlation function (left panels) and spectral density (right panels) of the torsional angle ψ(N-CA-C-N) of residue Arg 17 in the backbone of the protein BPTI in 10 ns MD simulations solvated in (SPC) water (*MD_aq*) and from 10 ns SD simulations in vacuo (*SD_vac*) using three different sets of constraints, set *BLC*, set *BLC* + *BADAC(H)*, and set *BLC* + *BADAC(H)* + *ISDAC*, and for different time steps Δt. Upper three panels: simulations *MD_aq*. Lower two panels: simulations *SD_vac*. Dotted lines: $\Delta t = 2$ fs. Dashed lines: $\Delta t = 4$ fs. Solid lines: $\Delta t = 6$ fs. $tol_{DC} = 10^{-5}$. $tol_{BAC} = tol_{DAC} = 0.01°$. Configurations from 25 ps towards 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

constraints can be inferred from Figure 3 (upper part), 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 for the different sets of constraints and time steps Δt. The curves resulting from the application of the constraint sets *BLC* and *BLC + BADAC(H)* are almost identical. When also constraining the improper dihedral angles and the stiff proper (torsional) dihedral angles in the protein, using constraint set *ISDAC* in addition to the sets *BADAC(H)* and *BLC*, slightly larger differences are observed.

Figures 4 and 5 (upper part) 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 the protein applying the different sets of constraints and time steps Δt. Using the constraint sets *BLC* and set *BLC + BADAC(H)*, 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 6 for the side-chain torsional angle χ_2(CA-CB-CG-CD) of residue Arg 39. When also constraining the improper dihedral angles and the stiff proper (torsional) dihedral angles in the protein, using constraint set *ISDAC* in addition to the sets *BADAC(H)* and *BLC*,

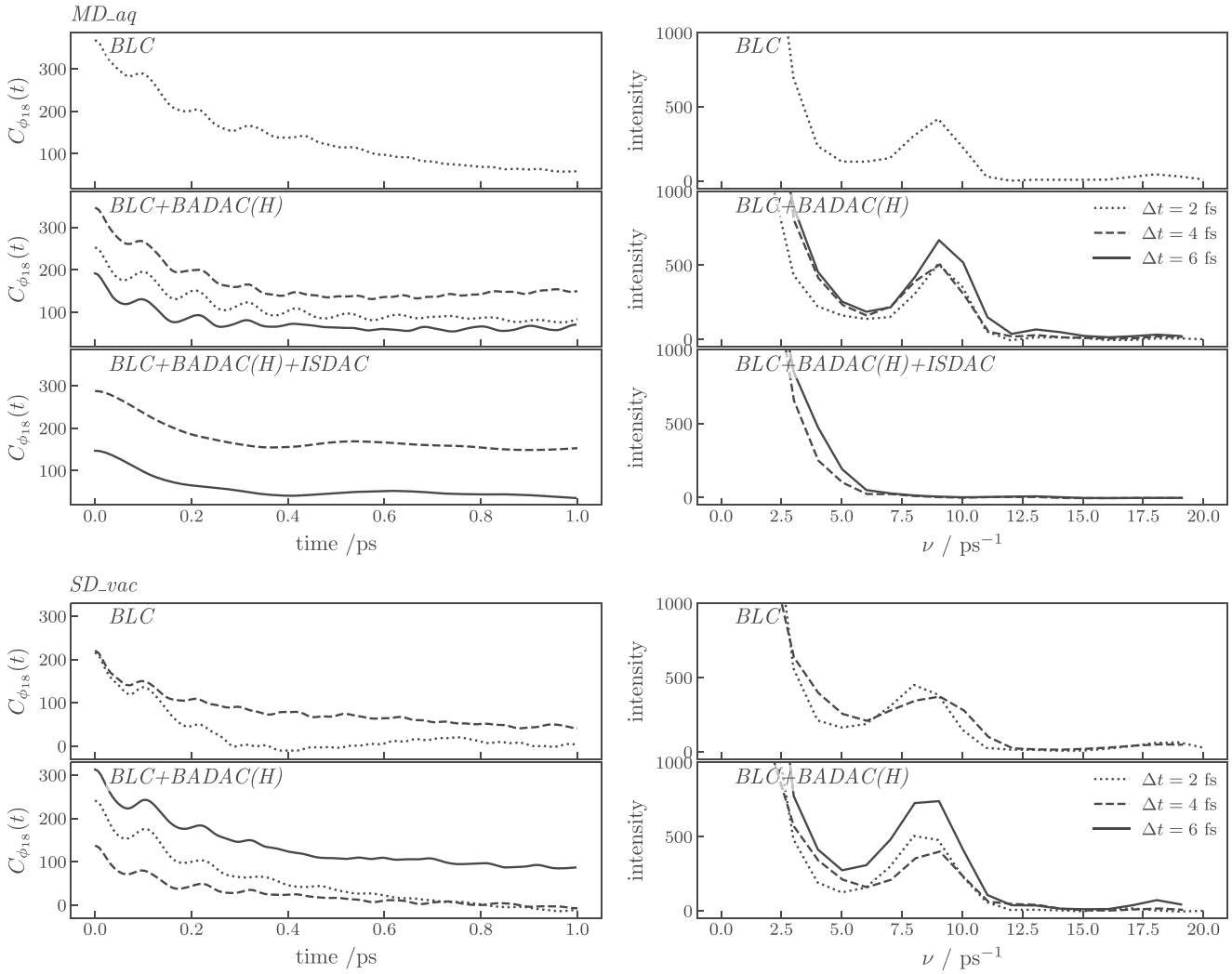

FIGURE 5 Auto-correlation function (left panels) and spectral density (right panels) of the torsional angle φ(C-N-CA-C) of residue Ile 18 in the backbone of the protein BPTI in 10 ns MD simulations solvated in (SPC) water (*MD_aq*) and from 10 ns SD simulations in vacuo (*SD_vac*) using three different sets of constraints, set *BLC*, set *BLC + BADAC(H)*, and set *BLC + BADAC(H) + ISDAC*, and for different time steps Δt. Upper three panels: simulations *MD_aq*. Lower two panels: simulations *SD_vac*. Dotted lines: $\Delta t = 2$ fs. Dashed lines: $\Delta t = 4$ fs. Solid lines: $\Delta t = 6$ fs. $tol_{DC} = 10^{-5}$. $tol_{BAC} = tol_{DAC} = 0.01^{\circ}$. Configurations from 25 ps towards 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 peak in the spectral densities of the torsional angles ψ(N-CA-C-N) of residue Arg 17 and φ(C-N-CA-C) of residue Ile 18 at about 9 ps^{-1} is gone.

In summary, when comparing the motions along nonconstrained degrees of freedom in the MD simulations of the protein in aqueous solution using the different combinations of the two sets of constraints, set *BLC* and set *BADAC(H)*, no large differences are observed. This indicates that the motions along these constrained degrees of freedom are not strongly coupled to the motions along the other degrees of freedom of the protein. This allows constraining the relative positions of the hydrogen atoms in the protein, using constraint set *BADAC(H)* in addition to set *BLC*, and a time step $\Delta t = 4$ fs. When also constraining the improper dihedral angles and the stiff proper (torsional) dihedral angles in the protein, using constraint set *ISDAC* in addition to the sets *BADAC(H)* and *BLC*, larger variations in mobility are observed between the different time step sizes.

In a study[53] similar to the current one it was observed that the presence of many water molecules, as in liquid water, limits the size of the MD time step to about 2 fs.

3.3 | SD simulation of a protein in vacuo to test dynamical protein properties

In a stochastic dynamics simulation of a protein, the influence of the solvent molecules on the protein structure and dynamics is modeled using a mean-field approximation by stochastic and frictional forces acting on the protein atoms, see Equations (5–7). Due to the absence of specific interactions, such as hydrogen bonding, between protein

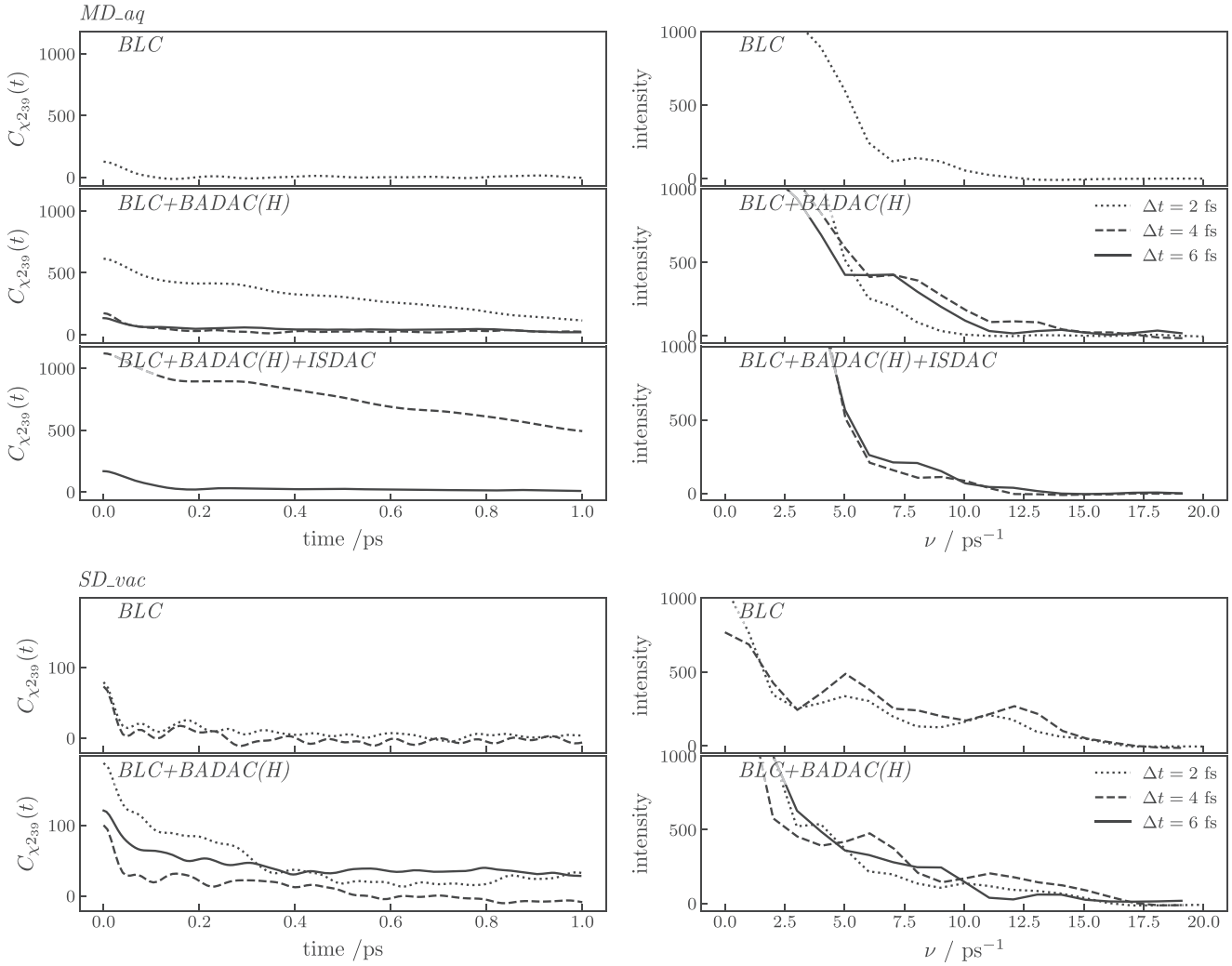

FIGURE 6 Auto-correlation function (left panels) and spectral density (right panels) of the torsional angle χ_2(CA-CB-CG-CD) of residue Arg 39 of the protein BPTI in 10 ns MD simulations solvated in (SPC) water (*MD_aq*) and from 10 ns SD simulations in vacuo (*SD_vac*) using three different sets of constraints, set *BLC*, set *BLC* + *BADAC(H)*, and set *BLC* + *BADAC(H)* + *ISDAC*, and for different time steps Δt. Upper three panels: simulations *MD_aq*. Lower two panels: simulations *SD_vac*. Dotted lines: $\Delta t = 2$ fs. Dashed lines: $\Delta t = 4$ fs. Solid lines: $\Delta t = 6$ fs. $tol_{DC} = 10^{-5}$. $tol_{BAC} = tol_{DAC} = 0.01°$. Configurations from 25 ps towards 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

atoms and water molecules, the protein structure and mobility may differ from that in a simulation with explicit water molecules, see for example,.[33,34] This can be addressed by adding an additional force-field term, that in a mean-field manner approximates the influence of the solvent molecules. Since in the present study the focus is on the influence of constraining particular degrees of freedom on the dynamics and mobility of the protein atoms, such mean-field force-field terms were not used in order to minimize the changes in the force field between the three simulation types, MD in vacuo (*MD_vac*), MD in aqueous solution (*MD_aq*) and SD in vacuo (*SD_vac*).

Figures 1 and 2 (lower part) show the root-mean-square fluctuations of the backbone φ and ψ torsional angles as function of residue number applying the different sets of constraints and time steps Δt. Constraining all bond lengths (*BLC*) in the protein, the simulations fail to converge for $tol_{DC} = 10^{-5}$ and $\Delta t = 6$ fs and for $tol_{DC} = 10^{-4}$ and $\Delta t = 4$ and 6 fs. The torsional-angle fluctuations are more or less comparable, the differences being due to differences in torsional-angle transitions occurring during the simulations. All SD simulations using constraint set ISDAC fail due to vectors becoming orthogonal in the SHAKE procedure.

When also constraining the relative positions of the hydrogen atoms, using constraint set *BADAC(H)* in addition to set *BLC*, the larger time steps can be used. Constraining the improper dihedral angles and the stiff proper (torsional) dihedral angles in the protein, using constraint set *ISDAC* in addition to the sets *BADAC(H)* and *BLC* in the SD simulations, no convergence could be obtained.

Figure 3 (lower part) 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 for the different sets of constraints and time steps Δt. The curves resulting from the application of the constraint sets *BLC* and *BLC* + *BADAC(H)*, are similar. For the large time steps $\Delta t = 4$ and 6 fs, the peak at about 9 ps^{-1} is more pronounced when constraining the relative hydrogen positions. The auto-correlation functions and spectral densities do resemble those for the MD simulations in aqueous solution (*MD_aq*).

Figures 4 and 5 (lower part) 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 the protein applying the different sets of constraints and time steps Δt. Using the constraint sets *BLC* and set *BLC* + *BADAC(H)*, 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. Again, the auto-correlation functions and spectral densities do resemble those obtained from the MD simulations in aqueous solution (*MD_aq*).

In summary, when comparing the motions along nonconstrained degrees of freedom in the SD simulations of the protein in vacuo using the two sets of constraints *BLC* and *BLC* + *BADAC(H)*, no large differences are observed. This indicates that the motions along these constrained degrees of freedom are not strongly coupled to the motions along the other degrees of freedom of the protein. This allows constraining the relative positions of the hydrogen atoms in the protein, using constraint set *BADAC(H)* in addition to set *BLC*, and a time step $\Delta t = 4$ fs.

4 | CONCLUSIONS

The application of bond-angle and dihedral-angle constraints, in addition to bond-length constraints, in molecular (MD) and stochastic (SD) dynamics simulations of a protein, bovine pancreatic trypsin inhibitor (BPTI), is investigated with an eye to lengthening the time step by which the classical equations of motion are integrated forward in time. Using Newton's equations of motion the extent of conservation of the total energy will be determined by the numerical integration algorithm (the leap-frog algorithm), the precision of the algorithms to impose the bond-length (tol_{DC}), bond-angle (tol_{BAC}) and dihedral-angle (tol_{DAC}) constraints, and the size of the MD time step Δt. For the MD simulations of the protein solvated in water and for the SD simulations of the protein in vacuo it is investigated whether dynamical properties of selected nonconstrained degrees of freedom are affected by the application of the various types of constraints.

Using total energy conservation as criterion, constraining all bond lengths in the protein with a relative geometric precision of $tol_{DC} = 10^{-5}$ allows for a time step $\Delta t = 2$ fs. In one of the earliest protein simulation studies,[8] using a simple protein model containing no hydrogen atoms, only united atoms, it was found that a time step of $\Delta t = 2$ fs could be used when the bond-length constraints between the united atoms were maintained with a precision of $tol_{DC} = 10^{-4}$. Since in current models of proteins the light hydrogen atoms are explicitly treated, their relatively fast motion requires a higher precision, $tol_{DC} = 10^{-5}$, when constraining the bond lengths in a protein.

Constraining the relative positions of the hydrogen atoms, in addition to constraining all bond lengths in the protein, allows for a lengthening of the time step by a factor two to $\Delta t = 4$ fs, without much distortion of the dynamics of the protein.

The computational effort of the evaluation of the bond-stretching and bond-angle bending forces at a number of small time steps Δt is comparable to that of maintaining the bond-length (and bond-angle) constraints at the larger time step Δt.

Constraining the improper dihedral angles and the stiff proper (torsional) dihedral angles in the protein, in addition to constraining all bond lengths and the relative positions of the hydrogen atoms, does not allow for a lengthening of the time step in a MD or SD simulation using Cartesian coordinates. This is due to the large number of constraints that involve many of the same atoms and thus the algorithms that maintain the constraints have trouble to find Lagrange multiplier values that make the molecular structure satisfy all constraints to within the specified precision.

When simulating a protein solvated in water, it may not be possible to use a 4 fs time step while constraining the relative positions of the hydrogen atoms, because the relatively fast librational motions of the rigid water molecules present may not allow for a time step longer than 2 fs. Furthermore, in simulations including water in a periodic box the nonbonded interaction cut-off is the dominant factor for

nonconservation of energy and leads to changes in some dynamic water properties at time steps higher than 2 fs.[53]

Yet, when performing SD simulation in vacuo, including a proper potential of mean force mimicking the influence of the solvent upon the protein motion, a time step of 4 fs can be used when constraining the relative positions of the hydrogen atoms.

ACKNOWLEDGMENT

Open access funding provided by Eidgenossische Technische Hochschule Zurich.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1002/prot.26251.

DATA AVAILABILITY STATEMENT

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

REFERENCES

1. van Gunsteren WF, Berendsen HJC. Computer simulation of molecular dynamics: methodology, applications and perspectives in chemistry. *Angew. Chem. Int. Ed. Engl*. 1990;29:992-1023. Angew Chem 1990, 102, 1020–1055.
2. Fixman M. Classical statistical mechanics of constraints: a theorem and application to polymers. *Proc Natl Acad Sci U S A*. 1974;71:3050-3053.
3. Fixman M. Simulation of polymer dynamics. I. General theory. *J Chem Phys*. 1978;69:1527-1537.
4. H.J.C. Berendsen, W.F. van Gunsteren, Molecular dynamics with constraints, in: "*The Physics of Superionic Conductors and Electrode Materials*", J.W. Perram, ed., NATO ASI Series 1983, *B92*, 221–240 (Plenum Press).
5. van Gunsteren WF, Karplus M. Effect of constraints on the dynamics of macromolecules. *Macromolecules*. 1982;15:1528-1544.
6. van Gunsteren WF. Constrained dynamics of flexible molecules. *Mol Phys*. 1980;40:1015-1019.
7. Ryckaert J-P, Ciccotti G, Berendsen HJC. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys*. 1977;23:327-341.
8. van Gunsteren WF, Berendsen HJC. Algorithms for macromolecular dynamics and constraint dynamics. *Mol. Phys*. 1977;34:1311-1327.
9. Tironi IG, Brunne RM, van Gunsteren WF. On the relative merits of flexible versus rigid models for use in computer simulations of molecular liquids. *Chem Phys Lett*. 1996;250:19-24.
10. Mazur AK. Hierarchy of motions in protein dynamics. *J. Phys. Chem. B*. 1998;102:473.
11. van Gunsteren WF, Bonvin AMJJ, Daura X, Smith LJ. Aspects of modeling biomolecular structure on the basis of spectroscopic or diffraction data. *Structure Computation and Dynamics in Protein NMR, Biological Magnetic Resonance*. Vol 17. Plenum Press; 1999:3-35.
12. Feenstra K, Hess B, Berendsen HJC. Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *J Comput Chem*. 1999;20:786-798.
13. Stocker U, Juchli D, van Gunsteren WF. Increasing the time step and efficiency of molecular dynamics simulations: optimal solutions for equilibrium simulations or structure refinement of large biomolecules. *Mol Simul*. 2003;29:123-138.
14. van Gunsteren WF. Computer simulation of biomolecular systems: overview of time-saving techniques. In: Lavery R, Rivail J-L, Smith J, eds. *Advances in Biomolecular Simulations*. American Inst. of Physics (A.I.P.) Conference Proceedings. Vol 239. American Institute of Physics; 1991:131-146.
15. Andersen HC. Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *J Comput Phys*. 1983;52:24-34.
16. Barth E, Kuczera K, Leimkuhler B, Skeel RD. Algorithms for constrained molecular dynamics. *J Comput Chem*. 1995;16:1192-1209.
17. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: a linear constraint solver for molecular simulations. *J Comput Chem*. 1997;18:1463-1472.
18. Tao P, Wu X, Brooks BR. Maintain rigid structures in Verlet based Cartesian molecular dynamics simulations. *J Chem Phys*. 2012;137:134110.
19. Gonnet P, Walther JH, Koumoutsakos P. θ-SHAKE: an extension to SHAKE for the explicit treatment of angular constraints. *Comput Phys Commun*. 2009;180:360-364.
20. Dubbeldam D, Oxford GAE, Krishna R, Broadbelt LJ, Snurr RQ. Distance and angular holonomic constraints in molecular simulations. *J Chem Phys*. 2010;133:034114.
21. Pechlaner M, Dorta AP, Lin Z, Rusu VH, van Gunsteren WF. A method to apply bond-angle constraints in molecular dynamics simulation. *J Comput Chem*. 2021;42:418-434.
22. Christen M, Kunz A-PE, van Gunsteren WF. Sampling of rare events using hidden restraints. *J. Phys. Chem. B*. 2006;110:8488-8498. Erratum: *J Phys Chem B* 2008, 112, 11446.
23. Pechlaner M, van Gunsteren WF. Algorithms to apply dihedral-angle constraints in molecular or stochastic dynamics simulations. *J Chem Phys*. 2020;152:024109.
24. Rice LM, Brünger AT. Torsion angle dynamics: reduced variable conformational sampling enhances crystallographic structure refinement. *Proteins: Struct. Funct. Genet*. 1994;19:277-290.
25. Katz H, Roderich Walter T, Somorjai RL. Rotational dynamics of large molecules. *Comput Chem*. 1979;3:25-32.
26. Mazur AK, Dorofeev VE, Abagyan RA. Derivation and testing of explicit equations of motion for polymers described by internal coordinates. *J Comput Phys*. 1991;92:261-272.
27. Wittenburg J. *Dynamics of Systems of Rigid Bodies*. Teubner, Stuttgart; 1977.
28. Bae D-S, Haug EJ. A recursive formulation for constrained mechanical system dynamics: part I. open loop systems. *J. Struct Mech*. 1987;15:359-382.
29. Bae D-S, Haug EJ. A recursive formulation for constrained mechanical system dynamics: part II. Closed loop systems. *J. Struct. Mech*. 1987;15:481-506.
30. Jain A, Vaidehi N, Rodriguez G. A fast recursive algorithm for molecular dynamics simulation. *J Comput Phys*. 1993;106:258-268.
31. Turner JD, Weiner PK, Chun HM, Lupi V, Gallion S, Singh UC. *Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications*. Vol 2. Netherlands: Springer; 1993:535-555.
32. Mathiowetz AM, Jain A, Karasawa N, Goddard WA III. Protein simulations using techniques suitable for very large systems: the cell multipole method for nonbond interactions and the Newton–Euler inverse mass operator method for internal coordinate dynamics. *Proteins: Struct Funct Genet*. 1994;20:227-247.
33. Shi YY, Wang L, van Gunsteren WF. On the approximation of solvent effects on the conformation and dynamics of Cyclosporin A by stochastic dynamics simulation techniques. *Mol Simul*. 1988;1:369-383.
34. van Gunsteren WF, Luque FJ, Timms D, Torda AE. Molecular mechanics in biology: from structure to function, taking account of solvation. *Ann Rev Biophys Biomol Structure*. 1994;23:847-863.

35. Poger D, van Gunsteren WF, Mark AE. A new force field for simulating phosphatidylcholine bilayers. *J Comput Chem*. 2010;31:1117-1125.
36. Schmid N, Eichenberger AP, Choutko A, et al. Definition and testing of the GROMOS force-field versions: 54A7 and 54B7. *Eur Biophys J*. 2011;40:843-856.
37. Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J. Interaction models for water in relation to protein hydration. In: Pullman B, ed. *Intermolecular Forces*. Reidel; 1981:331-342.
38. W. F. van Gunsteren, et al. The GROMOS software for (Bio)molecular simulation. Vol. 3. *Force field and topology data set*; 2011. http://www.gromos.net/.
39. W. F. van Gunsteren, et al. The GROMOS software for (Bio)molecular simulation. Vol. 2. *Algorithms and formulae for modelling of molecular systems*; 2011. http://www.gromos.net/.
40. W. F. van Gunsteren et al. The GROMOS Software for (Bio)Molecular Simulation. Vol. 1-9. 2011. http://www.gromos.net/.
41. Berendsen HJC, van Gunsteren WF, Zwinderman HRJ, Geurtsen RG. Simulations of proteins in water. *Ann N Y Acad Sci*. 1986;482:269-285.
42. van Gunsteren WF, Berendsen HJC, Geurtsen RG, Zwinderman HRJ. A molecular dynamics computer simulation of an Eight-Base-pair DNA fragment in aqueous solution: comparison with experimental two-dimensional NMR data. *Ann N Y Acad Sci*. 1986;482:287-303.
43. Barker JA, Watts RO. Monte Carlo studies of the dielectric properties of water-like models. *Mol. Phys*. 1973;26:789-792.
44. Tironi IG, Sperb R, Smith PE, van Gunsteren WF. A generalized reaction field method for molecular dynamics simulations. *J Chem Phys*. 1995;102:5451-5459.
45. Kunz APE, Allison JR, Geerke DP, et al. New functionalities in the GROMOS biomolecular simulation software. *J Comput Chem*. 2012;33:340-353.
46. Schmid N, Christ CD, Christen M, Eichenberger AP, van Gunsteren WF. Architecture, implementation and parallelization of the GROMOS software for biomolecular simulation. *Comp Phys Commun*. 2012;183:890-903.
47. Hockney RW, Eastwood JW. *Computer Simulation Using Particles*. McGraw-Hill; 1981.
48. van Gunsteren WF, Berendsen HJC. A leap-frog algorithm for stochastic dynamics. *Mol Simul*. 1988;1:173-185.
49. Berman HM, Westbrook J, Feng Z, et al. The Protein Data Bank. *Nucleic Acids Res*. 2000;28:235-242. www.pdb.org
50. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. *J Chem Phys*. 1984;81:3684-3690.
51. Futrelle RP, McGinty DJ. Calculation of spectra and correlation functions from molecular dynamics data using the fast Fourier transform. *Chem Phys Lett*. 1971;12:285-287.
52. W. F. van Gunsteren et al. The GROMOS software for (bio)molecular simulation. Vol. 6. *Technical details*; 2011. http://www.gromos.net/.
53. 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.

How to cite this article: Pechlaner M, van Gunsteren WF. On the use of intra-molecular distance and angle constraints to lengthen the time step in molecular and stochastic dynamics simulations of proteins. *Proteins*. 2022;90(2):543-559. doi: 10.1002/prot.26251