

Methods for Classical-Mechanical Molecular Simulation in Chemistry: Achievements, Limitations, Perspectives

Wilfred F. van Gunsteren* and Chris Oostenbrink



Cite This: *J. Chem. Inf. Model.* 2024, 64, 6281–6304



Read Online

ACCESS |

Metrics & More

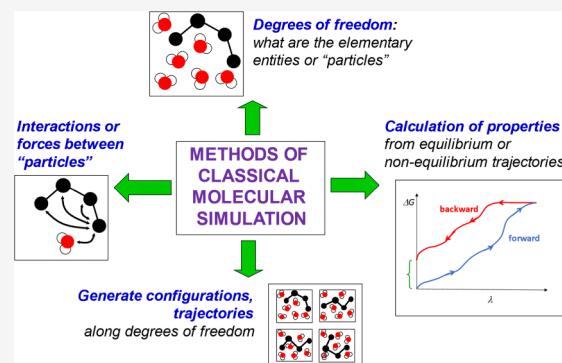
Article Recommendations

ABSTRACT: More than a half century ago it became feasible to simulate, using classical-mechanical equations of motion, the dynamics of molecular systems on a computer. Since then classical-physical molecular simulation has become an integral part of chemical research. It is widely applied in a variety of branches of chemistry and has significantly contributed to the development of chemical knowledge. It offers understanding and interpretation of experimental results, semiquantitative predictions for measurable and nonmeasurable properties of substances, and allows the calculation of properties of molecular systems under conditions that are experimentally inaccessible. Yet, molecular simulation is built on a number of assumptions, approximations, and simplifications which limit its range of applicability and its accuracy. These concern the potential-energy function used, adequate sampling of the vast statistical-mechanical configurational space of a molecular system and the methods used to compute particular properties of chemical systems from statistical-mechanical ensembles. During the past half century various methodological ideas to improve the efficiency and accuracy of classical-physical molecular simulation have been proposed, investigated, evaluated, implemented in general simulation software or were abandoned. The latter because of fundamental flaws or, while being physically sound, computational inefficiency. Some of these methodological ideas are briefly reviewed and the most effective methods are highlighted. Limitations of classical-physical simulation are discussed and perspectives are sketched.

KEYWORDS: molecular dynamics, stochastic dynamics, force field, sampling, properties

1. INTRODUCTION

It is about 60 years ago that Anees Rahman published his ground-breaking computer simulation study of the dynamics and motional correlation in liquid argon.¹ This heralded the advent of computers in physical-chemical research. Since the 1960-ies computer simulation of chemical processes based on classical, i.e. nonquantum-mechanical, physics has been playing an ever larger role in chemistry. It is so widely applied in a variety of branches of chemistry that a comprehensive review of the contribution of classical-physical computer simulation methodology to the development of chemical knowledge is a sheer impossibility. Nevertheless, it seems worthwhile, after 60 years of molecular simulation on computers, to review the various methodological ideas that have been proposed, investigated, evaluated, implemented in general simulation software or abandoned. Some ideas, methods and techniques have become standard tools of molecular simulation or have enhanced our understanding of the processes underlying chemical observations. Others turned out, although physically sound, to be inefficient or to require much computational effort leading to a marginal increase of accuracy, so were not further pursued. Yet it seems useful, a.o. in order to avoid repetition of investigations of



the latter methods, to review methodological ideas, whether or not they turned out to have become standard ingredients of molecular simulation. A third category of methods, critically assessed here, contains techniques that are popular, but lack a sound physical basis, so would not provide physical insight or understanding of the process simulated. Neither the great variety of applications of classical-mechanical simulation in chemistry nor the vast area of quantum-mechanical methodology are reviewed.

Computer simulation of molecular processes based on classical-physical laws essentially rests on three assumptions, approximations or simplifications.

1. The interaction between atoms or particles can be adequately described by a potential-energy function of the positions of the atoms or particles in a system of

Received: May 13, 2024

Revised: July 23, 2024

Accepted: July 23, 2024

Published: August 13, 2024



- interest. This implies that quantum effects must play a minor role in the process of interest.
2. The relevant parts of the vast statistical-mechanical configurational space of a molecular system of interest can be sufficiently sampled in a computer simulation. A great variety of techniques for enhanced sampling have been proposed and investigated.
 3. There are ways to compute particular properties of chemical systems from statistical-mechanical ensembles generated in simulations, either in equilibrium or in nonequilibrium simulations.

These issues of molecular simulation will be discussed distinguishing the following aspects:

1. The degrees of freedom of the molecular system that are being explored, i.e. the level of resolution at which the system is modeled.
2. The functional form of the different terms of the interaction function or force field describing the interactions in the molecular system and the way its parameters are calibrated.
3. Methods to enhance the searching and sampling of the relevant parts of the statistical-mechanical phase or configuration space of the molecular system.
4. Calculation of system properties from statistical-mechanical ensembles using equilibrium, with or without an externally applied field or force, or nonequilibrium simulations.

We have composed reviews of molecular simulation methodology along similar, yet different, lines more than thirty-three and 17 years ago^{2,3} and thought it time for an update with the focus on methodology, not on applications. Since our most recent review, molecular simulation has even more become a standard tool in the molecular sciences, becoming more accessible to nonexperts. This review discusses some of the methodological achievements leading up to this broad accessibility.

2. IDEAS, ACHIEVEMENTS

When describing ideas and achievements reported in the literature during the past half century,⁴ the four above-mentioned basic aspects of molecular simulation will be distinguished: (1) choice of degrees of freedom to be explored, the level of resolution; (2) interaction function or force field; (3) searching and sampling configurational or conformational space; (4) ways to calculate properties from configurational ensembles.

2.1. Level of Spatial Resolution: Atomic, Supra-atomic or Supra-molecular. When simulating a molecular system, one has to choose which degrees of freedom are explicitly treated as variables, i.e. are sampled, and which ones are omitted from the simulation or treated in a mean-field manner. In chemistry, the following levels of spatial resolution or of degrees of freedom explicitly simulated can be distinguished:⁵ (1) nucleons and electrons, (2) nuclei and electrons, (3) atoms, (4) supra-atomic particles, e.g. united CH_n (n = 1, 2, 3) atoms, (5) supra-molecular particles or beads. Levels 1 and 2 require a quantum-mechanical treatment, so are largely left aside here, although current developments to include electronic effects in machine-learned potential-energy surfaces will be shortly mentioned. The interactions at levels 3–5 are commonly modeled as Coulomb, van der Waals, short-range repulsive and bonded interactions.

The elimination of degrees of freedom from a model, in the current case an atomic model, is also called coarse-graining. Figure 1a illustrates that a supra-atomic water model

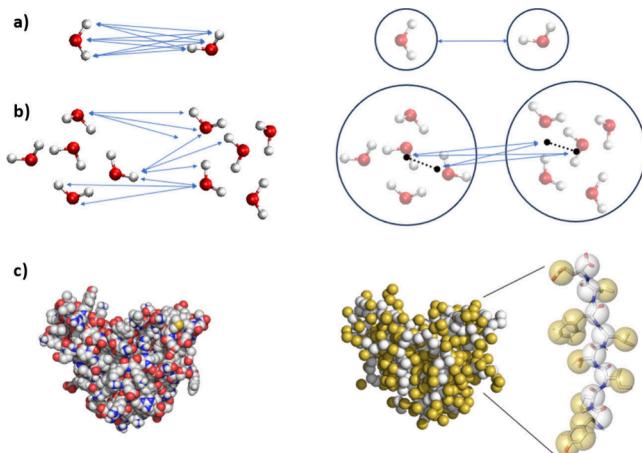


Figure 1. Coarse graining leads to a reduction of the number of interactions. a) The reduction from an atomistic to a supra-atomic water model reduces the interactions between them from 9 to 1. b) The reduction from an atomistic to a supra-molecular polarizable model reduces the number interactions from $15 \times 15 = 225$ to $2 \times 2 = 4$. c) Models of atomistic and coarse-grained protein, the inset shows the mapping of supra-atomistic beads onto the non-hydrogen atoms of the protein.

representing a water molecule as a single Lennard-Jones particle⁶ will reduce the number of pairwise interactions from $3 \times 3 = 9$ for the atomic model to 1 for the supra-atomic model. For a supra-molecular water model that treats five water molecules as a single Lennard-Jones particle with a polarizable charge-on-spring dipole (one additional particle),⁷ the reduction of the number of pairwise interactions is even larger, from $15 \times 15 = 225$ for the atomic model to $2 \times 2 = 4$ for the supra-molecular model (Figure 1b). A second advantage of coarse-graining is that it leads to a smoother potential-energy function compared to the fine-grained one. However, there are conditions that must be fulfilled by degrees of freedom in order that they may be eliminated in a physically correct manner in the coarse-graining process, such that a computationally efficient and yet accurate coarse-grained model is obtained. These are the following.

1. They must be nonessential for the process or property of interest.
2. They must be large in number or computationally intensive, so that the computational gain is substantial enough to offset the loss in accuracy.
3. The interactions governing these degrees of freedom to be eliminated should be largely decoupled from the interactions governing the other degrees of freedom of the system that are to be maintained. This means that the frequency components of the motion along the degrees of freedom to be eliminated must be well separated from the other frequencies occurring in the system, and that the coupling between the two types of motion is weak.⁸
4. Their elimination should allow a simple, efficient representation of the interaction governing the other, remaining degrees of freedom that are to be sampled.

One example of eliminating degrees of freedom or coarse-graining from the atomic level of resolution that turned out to

enhance the efficiency of molecular simulation is the use of united CH_n atoms.⁹ By treating the aliphatic CH , CH_2 and CH_3 groups as united atoms, the number of atomic interaction sites is substantially reduced, up to almost a factor of 10 fewer pairwise nonbonded interactions for lipids, at the cost of losing the dipolar interactions of the CH moieties and the van der Waals interactions of the H atoms. The intramoiety motions of these CH_n groups are largely decoupled from the motions of the other atoms and the torsional interactions involving these H atoms can be incorporated into the corresponding interactions for the torsional angle that does not involve an aliphatic H atom. If the positions of these H atoms are needed, i.e. when calculating quantities such as nuclear Overhauser effect intensities (NOEs), residual dipolar couplings (RDCs) or S^2 order parameters measurable by nuclear magnetic resonance (NMR), the H atom positions can be easily recovered based on the positions of the carbon atom and its non-hydrogen covalently bound neighbors.¹⁰ Thus, all four conditions for appropriate coarse-graining are largely met in this case.

Another example of eliminating degrees of freedom at the atomic level of resolution that turned out to enhance the efficiency of molecular simulation is the use of geometric constraints for small molecules without intramolecular torsional-angle degrees of freedom, such as the solvents water, methanol or chloroform, or the use of bond-length,⁸ and for hydrogens also bond-angle and dihedral-angle,^{11–13} constraints in macromolecules. Bond-length constraints are standardly used in macromolecular simulations, because they satisfy conditions 1 to 4 and allow, through the use of SHAKE,¹⁴ LINCS¹⁵ or other techniques to maintain such constraints, a gain of about a factor of 3 in computational efficiency.

An example of coarse-graining that does not satisfy the mentioned conditions 3 and 4 is the use of an implicit solvent model, i.e. the attempt to mimic the effect of the solvent by a function that is only dependent on the solute coordinates.¹⁶ If the solvent is water, this leads to severe distortions of the energy as a function of the positions of the atoms of the solute. Although the motions of a large solute may cover time scales ranging from femtoseconds to milliseconds and the relaxation times of water molecules are of the order of picoseconds, their motions on picosecond to nanosecond time scales are not decoupled, and thus condition 3 is not satisfied for particular processes. In explicitly simulated aqueous solvents, the nonpolar particles aggregate, and the electrostatic interaction between ions is reduced, leading to dissolution. So-called hydrophobic or nonpolar particles do “like” water molecules, but their interaction with water molecules is less strong than the interaction between water molecules themselves, leading to water excluding the hydrophobes and their subsequent aggregation. Ions with opposite charges do interact with water molecules stronger than with each other, which leads to water molecules surrounding the ions and dissolution of ion pairs. The “hydrophobic effect”, the apparent attraction between nonpolar molecules or repulsion between ions in aqueous solution due to the stronger interaction between the water molecules or between water molecules and ions, cannot be properly modeled in terms of solute and ion coordinates only, because the effective interaction between solute atoms and their entropy is a complex function of the distribution of solvent configurations. Thus, condition 4 is difficult to meet.¹⁷

Another example of coarse-graining that does not satisfy the mentioned conditions 1, 3 and 4 is the use of a supra-atomic model in which several atoms in heterogeneous biomolecules

such as proteins, nucleic acids or carbohydrates are represented by a single particle (Figure 1c), i.e. the attempt to mimic the influence of the atomic heterogeneity by a function that is only dependent on a few internal coarse-grained solute coordinates.¹⁸ In order to counteract the deformation of a protein in such a model, a so-called elastic-network potential-energy term (Elastic Network Model, ENM) is added to the force field. This term introduces linear Hookean springs between all pairs of (coarse-grained) particles in the protein or carbohydrate.¹⁹ In this way the physical basis of such a coarse-grained model is largely eliminated and the intrinsic flexibility of a protein is ignored.¹⁷

2.2. Interaction Function or Force Field: Improvement of Functional Form and Calibration of Parameters. In view of the different levels of approximation of molecular models it is not surprising that the literature contains a great variety of force fields. They can be classified along different lines:² (1) type of compound that is to be mimicked, e.g. polypeptides, polynucleotides, carbohydrates, sugars, organic molecules, inorganic compounds, polymers, etc.; (2) type of environment of the compound of interest; e.g. gas phase, aqueous or nonpolar solution, solids; (3) range of temperatures and pressures covered by the effective interaction; (4) type of interaction terms in the force field; e.g. bond-stretching, bond-angle-bending, torsional-angle terms, two-, three- or many-body nonbonded interaction terms; (5) functional form of the interaction terms, e.g. exponential, ninth or 12th power repulsive nonbonded interaction; (6) type of parameter fitting, that is, to which quantities were parameters fitted, and were experimentally or theoretically obtained values used as target values.

We note that the choice of a particular force field should depend on the system properties one is interested in. Some applications require more refined force fields than others. Moreover, there should be a balance between the level of accuracy or refinement of different parts of a molecular model or force field. Otherwise, the computing effort put into a very detailed and accurate part of the calculation may easily be wasted due to the distorting effect of the cruder parts of the model.²

2.2.1. Functional Form. A molecular force field or effective potential-energy function typically contains different terms, representing different types of interactions:^{2,20} Bond-stretching interactions between atoms linked by one covalent bond (E^b), bond-angle-bending interactions between three atoms linked by two covalent bonds (E^{ba}), torsional-angle- or dihedral-angle-bending interactions between four atoms linked by three covalent bonds (E^{da}), so-called improper (nontorsional) dihedral-angle-bending interactions between four atoms linked by three covalent bonds (E^{ida}), short-range repulsive interactions between atoms linked by three covalent bonds (E^{1-4}), van der Waals interactions between atom pairs not linked by three or less covalent bonds (E^{vdW}), and electrostatic, Coulomb, interactions between (partial) charges of atom pairs not linked by three or less covalent bonds (E^{ele}).

Using machine-learning procedures, a more generalized potential-energy term connecting energy with atomic coordinates may be created based on a selected set of data and a chosen functional form.^{21,22}

2.2.1.1. Short-Range Interactions. Bond-stretching interactions are generally modeled as a harmonic function of the bond length b , in which the minimum energy bond length b_0 , and the force constant K^b vary with the particular type of bond. A computationally more efficient form is a quartic function.¹⁰ Bond-angle-bending (three-body) interactions are also generally modeled as a harmonic function, either of the bond angle θ or of

$\cos(\theta)$, the cosine of θ , in which the minimum energy bond angle θ_0 , and the force constant K^θ vary with the particular type of bond angle. The cosine form is the computationally more efficient term.¹⁰ Potential-energy terms coupling bond lengths and bond angles and using different functional forms, e.g. for representing molecules in the gas phase²³ or the Urey–Bradley form, have also been investigated and are used to better represent molecular vibrations.²⁴ Two forms are used for the (four-body) dihedral-angle interactions: (i) a harmonic term for (improper) dihedral angles ξ that are defined to handle out-of-plane distortions of atoms in planar groups and to maintain the geometry and chirality of a tetrahedral configuration at a particular atom, e.g. at a carbon united atom (E^{ida}); (ii) one or more sinusoidal functions for torsional dihedral angles φ , which may allow for 360 deg rotation (E^{da}). The torsional-angle terms and their parameter values are related to the way the short-range nonbonded (Coulomb and repulsive) interactions between the first and fourth atom of the torsional dihedral angle are modeled (E^{1-4}). The combination of E^{da} and E^{1-4} as a function of a particular torsional dihedral angle φ will basically determine the energy profile of the torsional-angle degree of freedom. Different combinations are employed in current force fields.²⁰ Short-range repulsive nonbonded interactions are generally modeled by a r^{-12} or r^{-9} distance dependence. The simpler the functional form, the more transferable its parameters will be between different compounds.

2.2.1.2. Long-Range, Electrostatic Interactions. The van der Waals interaction has an r^{-6} distance dependence and the Coulomb interaction a r^{-1} distance dependence. This means that the van der Waals energy of an atom with all its neighbors within a radius R_c is proportional to R_c^{-3} , see Table 1. A similar

Table 1. Distance Dependence of Electrostatic Energy $E(r)$ and $E(r)r^2$, Force $f(r)$ and $f(r)r^2$ between Charges and Dipoles^a

Interaction type	$E(r)$	$E(r)r^2$	$E(R_c)$	$f(r)$	$f(r)r^2$	$f(R_c)$
van der Waals	$1/r^6$	$1/r^4$	$1/R_c^3$	$1/r^7$	$1/r^5$	$1/R_c^4$
charge–charge	$1/r$	r	R_c^2	$1/r^2$	1	R_c
charge–dipole	$1/r^2$	1	R_c	$1/r^3$	$1/r$	$\ln(R_c)$
dipole–dipole	$1/r^3$	$1/r$	$\ln(R_c)$	$1/r^4$	$1/r^2$	$1/R_c$

^aThe total energy $E(R_c)$, integrated over a sphere of radius R_c is $E(R_c) \equiv \int_0^{R_c} E(r) 4\pi r^2 dr$. The total force $f(R_c)$, integrated over a sphere of radius R_c is $f(R_c) \equiv \int_0^{R_c} f(r) 4\pi r^2 dr$.

calculation shows that the Coulomb energy of a (partial) charge with all its neighbor charges within a radius R_c is proportional to R_c^2 . The Coulomb energy of an atom with nonzero partial charge with its neighbors (with nonzero partial charges) within a distance R_c grows with growing R_c , as does the corresponding force on the atom. If the molecular model does not involve bare (partial) charges on atoms, but only dipoles and higher multipoles, the electrostatic energy of an atom becomes proportional to $\ln(R_c)$, and the corresponding force proportional to R_c^{-1} . This challenge of handling long-range electrostatic energy and forces in a simulation has been addressed essentially in two different ways.² Each of these are based on distortive assumptions or approximations.

1. The molecular system is assumed to be crystalline, i.e. surrounded by an infinite number of copies of itself, as in the Particle–Particle–Particle-Mesh (P³ M) method,²⁵ later reinvented as the Particle-Mesh-Ewald (PME)

method.²⁶ These methods are called lattice-sum methods, which allow the calculation of the total electrostatic energy, the result depending on the boundary condition, either conducting or vacuum, assumed to exist at infinity.²⁷ Either assumption induces artifacts, as does the imposition of periodicity. For any pair of atoms with nonzero charges in a system, separated by exactly half the length of the periodic box, the effect of the periodicity is to reduce the Coulomb force to zero (Figure 2c, Table 2).

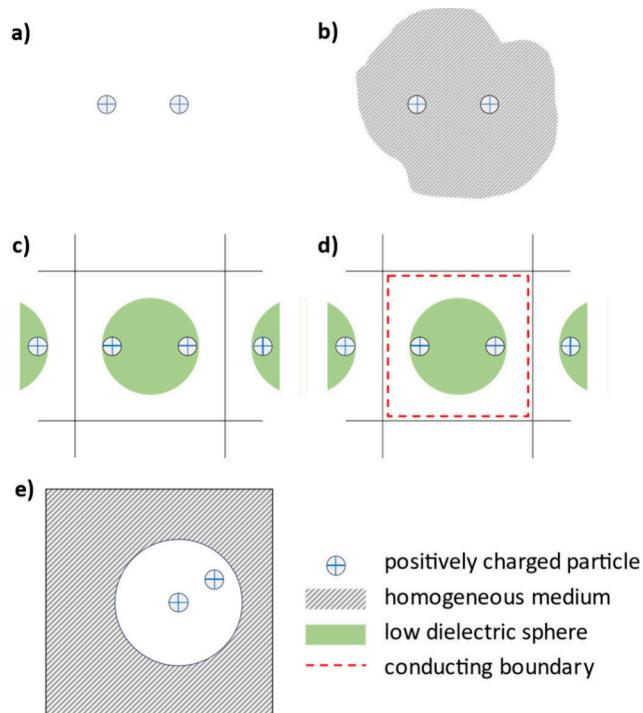


Figure 2. Representations of approaches to compute the electrostatic interactions between two positively charged particles: a) Infinite-box model in vacuum ($\epsilon_r = 1$); b) Infinite-box model in a homogeneous medium with the relative dielectric constant ϵ_r equal to that of a solvent ($2 \leq \epsilon_r \leq 100$); c) Periodic system with lattice-sum electrostatics; d) Periodic system with the electrostatic interactions computed with the conducting box boundary method (CBBM), the conducting box in red;³¹ e) Periodic system with the electrostatic interactions computed based on a cutoff with reaction-field contribution representing a homogeneous dielectric medium outside the cutoff sphere.^{32,33}

Table 2. Coulombic Forces (in kJ/mol/nm) on Two Positively Charged (1e) Particles at an Interatomic Distance of 1.25 nm

Box size (nm)	Two particles on a C ₆₀ Buckminsterfullerene of diameter 1.25 nm in 512 explicit (SPC) water molecules ³¹		
	Two particles only	Fully periodic ^a	Half periodic conducting boundary ^b
Infinite	89 ^c		
Infinite	1 ^d		
2.4863	0	265	314
3.72946	70	304	312

^aSee Figure 2c. ^bSee Figure 2d. ^cSee Figure 2a. ^dSee Figure 2b. Using a relative dielectric constant $\epsilon = 80$.

The same atoms, at the same interatomic distance, but in a larger periodic simulation box, will experience a different

force between them (Table 2). The artificial periodicity thus induces an attractive force between like charges and a repulsive force between opposite charges. One consequence of this is that a lattice-sum approach reduces the magnitude of the calculated free energies of ionic solvation.²⁸ Another effect is an induced polarization of the solvent in systems with a slab geometry.^{29,30} These periodicity artifacts can be avoided by using the so-called conducting box boundary method (CBBM),³¹ in which there is no periodicity regarding electrostatic interactions (Figure 2d, Table 2).

2. The electrostatic interaction due to the medium beyond R_c is represented by a (static) reaction-field approximation,^{32,33} a mean-field approximation (Figure 2e). In reality, the electrostatic reaction field is a response of the medium outside a sphere with radius R_c to the creation of a dipole or a dipole fluctuation within this sphere. This response is physically not immediate but delayed. Thus, introducing a delay in the reaction-field response³⁴ together with stochastic electric-field fluctuations³⁵ would offer a physically more sound method. However, in practice it did not improve significantly the properties of a water test system, so the simpler static version is commonly used in molecular simulation. The reaction-field approximation assumes a dielectrically homogeneous medium beyond R_c . This may be a good approximation when simulating homogeneous liquids, but not for lipid bilayers in contact with water or biomolecules in solution. This consideration led to a refinement of the method.³⁶

2.2.1.3. Accounting for Molecular Polarizability. Polarization plays an important role in the energetics of molecular systems, not the least in biomolecular systems. Yet, most computer simulation studies of such systems do not treat electronic polarizability explicitly, but only implicitly using effective charges, dielectric permittivities or continuum electrostatics methods. The introduction of explicit polarizability into molecular models and force fields may offer a better physical description of the electrostatic interactions.

There are various ways to account for polarizability in (bio)molecular simulation which vary in striking a balance between accuracy on the one hand and simplicity and computational efficiency on the other.³⁷ Computer simulations that include polarizability have been performed for a number of molecular systems, such as ionic liquids, solvents such as water and biomolecules.³⁸ There are basically three ways to model polarizability: (1) by inducing (ideal) point dipoles; (2) by changing the magnitudes of (atomic) charges; (3) by changing the positions of (atomic) charges. The latter type of model is called a Drude Oscillator model,³⁹ a Shell model⁴⁰ or a Charge-On-Spring model.⁴¹ This model is the simplest and computationally most efficient way to model polarizability classically, see e.g.⁴²

2.2.2. Calibration of Force-Field Parameters. Any model contains parameters, the values of which have to be chosen in one way or another.

2.2.2.1. Choice of Target Properties for Parameter Calibration. For parameter calibration one may choose as target properties values of experimentally observable quantities or of quantities that can be calculated from computer simulations of a particular system at another, generally more fine-grained level of modeling. Such a calibration may involve direct matching between the value of a model parameter and its

corresponding value taken from experiment or from a finer-grained model. For example, one may take the O–H bond length and H–O–H bond angle in a model for liquid water from its gas-phase geometry, as obtained from a quantum-mechanical calculation or from its solid-state geometry as obtained from diffraction experiments. A model parameter can also be calibrated in a less direct manner by fitting another property of the model to its finer-grained or experimental counterpart. For example, the partial atomic charges and the repulsive Lennard-Jones parameter of the simple-point-charge (SPC) water model⁴³ were chosen such that the experimental heat of vaporisation and density of liquid water at ambient temperature and pressure were reproduced. Calibration of model parameters against values of quantities obtained from computation or simulation of another, more fine-grained model, as is done when applying force-matching, relative entropy, reverse Monte Carlo, or iterative Boltzmann inversion techniques to calibrate force-field parameters of supra-atomic models, seems a risky strategy.⁵

1. The mentioned techniques project a configurational distribution (partition function) on geometric configurations of chosen, particular degrees of freedom. This implies a loss of free-energy information.
2. The quality of the coarse-grained force field thus relies on the accuracy of the more fine-grained model, which is for molecular systems generally not yet at the level of accuracy of experimental thermodynamic or dielectric screening data. A major consideration to attempt coarse-graining is that a simulation on the basis of a finer-grained model is too expensive to reach sufficient sampling of the configurational space, i.e. to obtain converged results. This makes the quality of the underlying finer-grained model and simulation even more important.
3. Supra-atomic models for larger molecules (except united atom ones) do not fulfill the conditions 3 and 4 mentioned in section 2.1 for physically correct coarse-graining.

The same observations hold for the use of machine learning methods to reproduce energies and forces from a fine-grained level: the potential energy and corresponding forces generated will never be more accurate than the data it was trained on, probably less accurate. Projection of particular quantities from a more fine-grained level to a more coarse-grained level also involves assumptions about the coupling between the motion along the eliminated and preserved degrees of freedom.

On the other hand, not all data considered to be experimental data are really measured, and thus suitable for parameter calibration. One should distinguish between observed, primary experimental data such as intensities of a diffracted beam in a scattering experiment or absorption intensities in a spectroscopic experiment, and derived, secondary “experimental” data that are obtained from observed data by a particular procedure which generally involves model assumptions and approximations that may introduce inaccuracy.⁴⁴ Examples of derived “experimental” data are X-ray structures determined from X-ray diffraction intensities or NMR model structures derived from NMR observables such as NOE intensities, RDCs and 3J -couplings. Use of derived “experimental” data in model parameter calibration is a highly risky strategy due to the uncertainty inherent to such data. Another useful distinction is between atomic, molecular and supra-molecular versus bulk or system properties. For example, the atomic partial charge could be calibrated to reproduce a particular molecular dipole moment

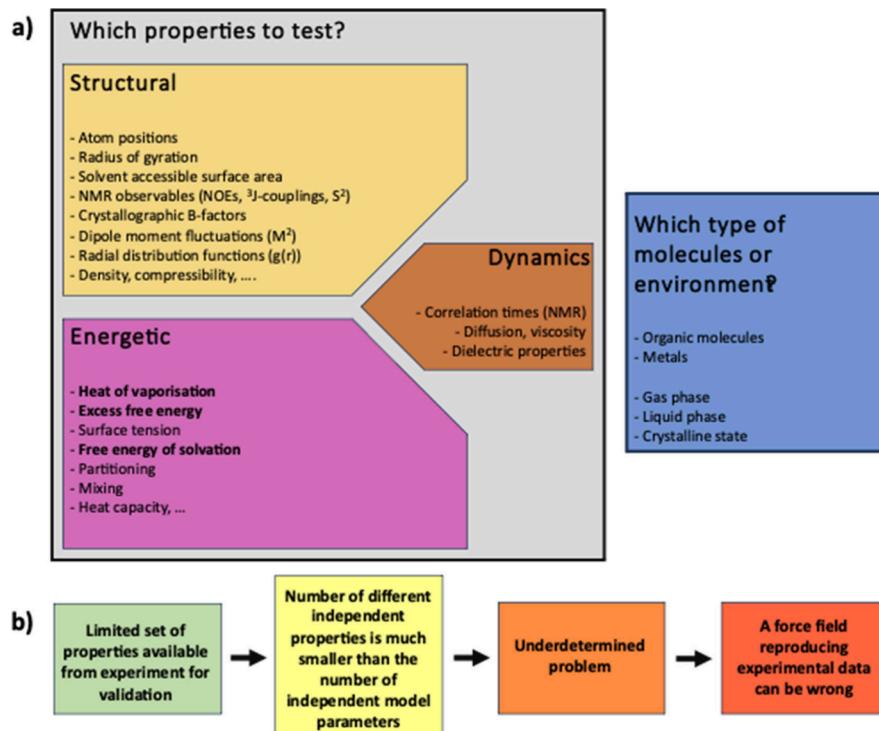


Figure 3. Validation of force-field parameters. a) Selection of properties and environments that can be used to validate force-field parameters. b) Intrinsic problem of force-field parameter calibration against experimental data.

rather than to reproduce the static dielectric permittivity of the bulk liquid.

Finally, we note that any model parameter is an “effective” parameter, i.e. its value has only limited significance per se, but obtains significance in connection with the values of other parameters and the overall properties of the model that were the target of the calibration. For example, the bond angle of a water molecule in the gas phase is not tetrahedral and it is probably also not tetrahedral in the liquid phase. Yet, many models for liquid water use a tetrahedral geometry in order to facilitate the reproduction of the dielectric screening properties of water, which is considered to be more important than the detailed geometry of a water molecule.⁴⁵

2.2.2.2. Relative Weight of Various Properties in Force-Field Parameter Calibration. When calibrating force-field parameters, it is to be determined which properties are of interest, and thus are to be reproduced in the calibration and which properties are of less interest and thus to be relinquished (Figure 3a). Another issue is which properties should get the largest weight in the calibration, for example, the (supramolecular) structural properties or the thermodynamic (system) properties? These questions do not have a universally correct answer, rather, the choice will depend on the intended application(s) of the model. Thus, the weights of the different properties selected for the optimization of the force-field parameters are to be chosen. For example, in biomolecular systems, the following condensed phase properties are of interest:⁵

1. Molecular structure of the solute or structure of the solvent.
2. Thermodynamic properties such as density, heat of vaporisation, excess free energy, surface tension or solvation free energy that carry volume and energetic information. Other thermodynamic quantities that

characterize a response to a change in thermodynamic state point, such as the heat capacity, isothermal compressibility or thermal expansion coefficient, are of secondary importance.

3. Dielectric properties, in particular the static dielectric permittivity that governs the screening of Coulomb interactions by the solvent.
4. Dynamic properties such as diffusion, viscosity, and various molecular or dielectric relaxation times. These are of less importance, because the equilibrium statistical-mechanical ensemble averages of nondynamic quantities generally do not depend on the values of the dynamic ones, and because most molecular processes are thermodynamically, not dynamically driven.

During the past half century, the classical interaction functions or force fields have seen a continuous development in terms of improvement of accuracy and widening of their range of applicability, from organic molecules to proteins, nucleotides, lipids and carbohydrates.^{4,5} The 1960s saw interaction functions for molecules in the gas phase,²³ the 1970s for molecules, such as water, in the condensed phase.⁴⁶ These force fields and models were primarily based on structural data derived from experiment. During the 1980s still widely used interaction functions for liquid water, calibrated using thermodynamic instead of structural data were developed,^{43,47} and general atomic-level interaction functions for biomolecules based on crystal and spectroscopic data became available.^{48–50} During the 1990s these general atomic interaction functions were considerably improved by calibration of their parameters based on free-energy data, i.e. enthalpy and entropy of small molecules.^{51,52} This thermodynamic basis enabled the solution of an experimental paradox by showing that values of NOE intensities from NMR experiments of a short peptide were originating from a conformational equilibrium of left-handed and right-handed

helices.⁵³ The past decades saw the development of polarizable interaction functions, e.g.,^{54–56} and the use of quantum-chemical calculations to obtain partial atomic charges for use in force fields, e.g.,⁵⁷ Current developments involve the use of machine learning methods to directly predict energies from a fine-grained method (e.g., density functional theory) based on the molecular coordinates.^{22,58,59}

The requirements for an effective force field are

1. It should be representative of the quantum-mechanical energy surface and represents the underlying physics appropriately.
2. It should be simple, i.e. contains as few, possibly simple, terms as possible.
3. It should be computationally efficient, avoiding complex derivatives and exponential terms.
4. It should be transferable, containing parameters appropriate for a range of molecules.
5. Its parameters should be calibrated against thermodynamic, free energy (entropy) data.

Figure 3b illustrates why a force field that reproduces many experimental data can still be wrong.

2.3. Searching and Sampling Configurational Space: A Variety of Methods. The first molecular simulations demonstrated that simulation of molecular motion in the condensed phase enhanced our insight of the molecular basis of observed properties and processes. On the other hand, it was clear that the time scale of the simulations, picoseconds, was too short to cover the much longer times that were responsible for physical processes of interest, e.g. protein folding.² This led to a great many ideas and proposals for ways to speed up molecular simulation by enhancing the searching or sampling properties of MD simulation.

Because molecular simulation involves microscopic systems at nonzero temperatures T , the basic theory to describe such a system is quantum or classical statistical mechanics. Consequently, the state of a molecular system is characterized by a statistical-mechanical ensemble of configurations x . At fixed number of molecules, volume, and temperature this is a canonical ensemble, in which the probability $P(x)$ of a molecular or system configuration x is given by the Boltzmann factor $P(x) \sim \exp(-E(x)/k_B T)$, where $E(x)$ denotes the energy of the system and k_B Boltzmann's constant. This implies that the equilibrium properties of the system are determined by those parts of configuration space, for which $E(x)$ is minimal. Therefore, one of the basic challenges to molecular modeling is to develop methodology to efficiently search the molecular energy surface $E(x)$ for regions of low energy. The statistical-mechanical nature of this search problem implies that it cannot be reduced to the problem of finding the global (energy) minimum of the multidimensional function $E(x)$. Statistical-mechanically the free energy $F = E - TS$, composed of an energetic contribution E and an entropic (S) contribution $-TS$, is minimal, not the energy E . The entropy is a measure of the extent of configurational space (x) accessible to the molecular system at a given temperature T . **Figure 4** illustrates that lowest energy does not necessarily mean lowest free energy. Two parts of configurational space x_1 and x_2 may have $E(x_1) \ll E(x_2)$, whereas $F(x_1) > F(x_2)$ due to $S(x_1) \ll S(x_2)$ at the given temperature T . Integrating the probability over all configurations similar to or around x_1 and around x_2 , accordingly leads to a higher overall occurrence of configurations around x_2 than around x_1 . This means that searching for and finding the global

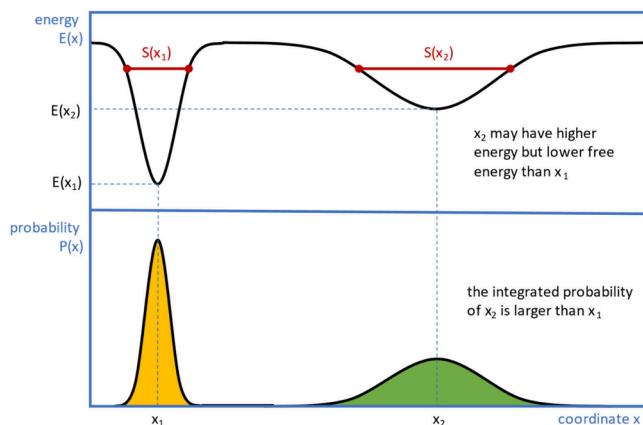


Figure 4. In statistical thermodynamics a state x_2 with a higher potential energy than a state x_1 may still have a higher probability of occurrence and a more favorable free energy.

energy minimum for a molecular system is meaningless when its entropy accounts for a sizable fraction of its free energy.

When considering methods to generate molecular configurations, we distinguish three types based on the characteristics of the set of generated configurations.

1. Methods that generate a series of nonrelated low-energy configurations.
2. Methods that generate a properly (Boltzmann) weighted set of configurations.
3. Methods that generate a (classical) dynamical trajectory of configurations, which are moreover properly (Boltzmann) weighted.

Methods of type 1 should only be used for zero entropy systems, whereas methods of types 2 and 3 yield proper ensembles, so can be used to compute thermodynamic and other equilibrium properties. Only methods of type 3 yield information on dynamical properties of the system.

2.3.1. Equilibrium Simulations. Equilibrium simulations can be performed using a force field representing the interatomic interactions of the molecular systems or using in addition a potential-energy term representing experimentally derived data on the molecular system of interest or using an external field to provoke a response.

2.3.1.1. Integration of the Equations of Motion. Over the past 50 years, a variety of algorithms has been used in molecular dynamics (MD) simulations to integrate the classical equations of motion forward in time. The predictor-corrector Gear-type of algorithms⁶⁰ are suitable when the highest frequency motions are largely harmonic.⁴⁶ When the forces have a more random character, as is the case when bond-length constraints⁸ or in addition bond-angle constraints and dihedral-angle constraints¹³ are applied, simpler algorithms, such as the Störmer-Verlet algorithm^{61,62} or the leapfrog algorithm,²⁵ show a similar performance in terms of energy conservation and stability.²⁸ In the leapfrog algorithm the particle velocities are updated at time points shifted by half a time step Δt from the positions. If velocity-dependent properties are of interest, these can be obtained by interpolation from half time steps to full time steps.⁶³ The error induced by the discretization of the equations of motion based on a truncated Taylor expansion is of third order (Δt)³ in the time step Δt . 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, e.g. for bond lengths or angles, are applied and the temperature of the simulation.^{13,64}

The method of stochastic dynamics (SD) is an extension of MD. A trajectory of the molecular system is generated by integration of the stochastic Langevin equation of motion. Two terms are added to the force in Newton's equations of motion, a stochastic force and a frictional force proportional to a friction coefficient. The stochastic term introduces energy, the frictional term removes (kinetic) energy from the system. These terms are related by a condition for zero energy loss, which keeps the temperature of the system close to a reference value. An SD leapfrog integration scheme to integrate the Langevin equations is available.⁶⁵ SD can be applied to establish a coupling of the individual atom motion to a heat bath.⁶⁶ This is particularly useful in a gas-phase simulation to ensure that the molecule does not get trapped in an energy minimum. SD can also be applied to mimic a solvent effect.⁶⁷ In this case the stochastic term represents collisions of solute atoms with solvent molecules and the frictional term represents the drag exerted by the solvent on the solute atom motion.

An alternative to the use of equations of motion in Cartesian coordinates and imposing constraints through Lagrange multipliers is to employ internal coordinates, bond lengths, bond angles, and torsional angles, as dynamical variables.^{68–75} The classical equations of motion have been formulated by Lagrange in a most general form in terms of generalized coordinates. When using Cartesian coordinates, these equations reduce to Newton's equations of motion. 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 expressed in internal, generalized coordinates are considerably more complex than when expressed in Cartesian coordinates.⁶⁴ 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 value of the coefficients of the different terms in the equations depend on the atomic masses and the molecular topology of the macromolecule considered.⁶⁴ In addition, the calculation of long-range interactions becomes rather cumbersome. Therefore, these equations of motion are generally not used for simulations of condensed-phase molecular systems.

2.3.1.2. Multiple-Time-Step Dynamic Simulation Methods. 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. 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, 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.

Multiple-time-step algorithms come in different flavours.^{2,76–82} In a molecular system, three main frequency ranges can be distinguished: high-frequency (*hf*) bond-stretching forces f^{hf} , low-frequency (*lf*) long-ranged (e.g., Coulomb) nonbonded forces f^{lf} , and the remaining intermediate frequency (*if*) forces 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,^{78,79} in which the longer-ranged nonbonded force f^{lf} is kept constant during n_{lf} time steps Δt , where n_{lf} lies in the range 5–10 for $\Delta t = 1$ or 2 fs (typical update frequency 10–20 fs). This MTS algorithm is

denoted the constant-force multiple-time-step algorithm (*cf-MTS*).⁸² When applying an MTS algorithm to the bond-stretching forces,^{2,80} the high-frequency covalent bond-stretching forces f^{hf} are evaluated at each time point $t + n\Delta t$, with $n = 1, 2, \dots$ When n is a multiple of n_{hf} (n_{hf} is odd, typically 3 or 5), the other (intermediate and low frequency) forces $f^{if} + f^{lf}$ 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*).⁸²

Applying the *cf-MTS* algorithm to nonbonded interactions, it was found to have very little effect on the thermodynamic properties that were used to calibrate the force-field parameters or on the overall dynamics of a large set of protein systems.^{83,84} Applying the *if-MTS* algorithm to nonbonded interactions was found to be less appropriate than the *cf-MTS* one.⁸⁵ However, when applied to the bond-stretching versus other forces, the *if-MTS* algorithm performs better than the *cf-MTS* one,⁸² as expected. However, bond-stretching degrees of freedom are better treated as constrained degrees of freedom.⁸² 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. In addition, it constitutes a better representation of the quantum-mechanical nature of bond-stretching vibrations in molecules than a classically, for example, harmonically, vibrating bond.⁸²

2.3.1.3. Use of Constraints. 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, $\Delta t \ll (\nu_{max})^{-1}$. The highest frequency 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 the four conditions mentioned in Section 2.1 are met.

Bond-length degrees of freedom 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,⁴⁹ metric-tensor effects^{86,87} are insignificant⁸⁸ and algorithms to impose distance constraints do not require excessive computational effort.¹⁴ A factor of 3 in computer time can typically be saved by the application of bond-length constraints in macromolecules.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.⁸⁸ While small solvent molecules without torsional-angle degrees of freedom are commonly simulated as rigid molecules,⁸⁹ 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.⁴⁹ Yet, bond angles involving hydrogen atoms may be constrained without too many artificial side effects.¹³ 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.^{13,90}

Several methods are available to apply distance constraints during a MD simulation based on equations of motion in Cartesian coordinates and solving the constraint equations for the Lagrange multipliers.^{14,15,91–93} For small molecules with three or less constraints, the constraint equations can be solved analytically by inversion of the rather small (maximally 3×3) matrix.⁹⁴ The efficient algorithm SETTLE,⁹⁵ often applied to water molecules, uses this technique. For larger chain molecules, matrix inversion is less efficient than solving the equations for the Lagrange multipliers iteratively. The iterative method SHAKE¹⁴ 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^{11,96,97} and the dihedral angles accordingly with the procedure SHAKEDAC.^{12,97,98}

Constraining bonds works rather well, sometimes even for thermodynamic states quite different from the one used in fitting the bond-length parameters. However, there exist processes and circumstances where the flexibility of bonds can be crucial. But, flexible models add very fast vibrations to the system thereby limiting the time step size. In addition, these vibrations are only loosely coupled to the other degrees of freedom making long equilibration times necessary. Although hard constraints are likely to be a more faithful representation of the quantum-mechanical (quantized) nature of the bond-stretching vibrations than a classical-mechanical harmonic (or quartic) oscillator, as used in nonconstrained simulations, they do not allow for a change in bond lengths during a simulation. In some processes a change in actual average bond lengths, for example, under the influence of pressure, may play an essential role and should therefore be possible in a simulation. To this end flexible-constraint algorithms have been proposed.^{99–102} Using these methods, the bond-length distance constraints are adiabatically adjusted to their current minimum-energy lengths according to the total energy (or total potential energy) of the system at the current time (step). Generally, these methods^{99,100} require multiple energy evaluations at every time step, which makes them an order of magnitude more expensive than hard constraint algorithms. However, using a reasonable approximation this disadvantage could be overcome and a fast flexible-constraint algorithm could be obtained.^{101,102}

2.3.1.4. Control of System Properties. When the classical equations of motion are integrated, the total energy is conserved and the temperature (kinetic energy) is a dependent variable. Its value cannot be imposed on the system. If the volume is also held constant, the pressure is also a dependent variable. For comparison to most experimental data this is not very convenient and a variety of approaches have been described in the literature to perform simulations in which temperature and pressure are independent variables rather than derived properties.¹⁰³ Slow temperature drifts that are an unavoidable result of force truncation and other errors, are compensated, while also rapid transitions to new desired conditions of temperature and pressure are more easily accomplished. The automatic control of system properties, such as temperature and pressure, can be

achieved using four different types of methods, each of which has its advantages and disadvantages:^{2,103}

1. Constraint methods, in which the temperature $T(t)$ at each time point t is set exactly equal to the desired reference temperature T_{ref} by rescaling the velocities by a factor $(T_{ref} / T(t))^{1/2}$.⁴⁶ In a microscopic system, consisting of much less atoms than Avogadro's number, fluctuations in the kinetic energy do occur, and thus should be allowed in a simulation.
2. Stochastic methods,^{104,105} in which the individual atom velocities are randomly modified at each time step such that a Maxwell velocity distribution corresponding to a temperature T_{ref} is obtained. The continuity of the atom velocity trajectories is lost. In addition, artifacts caused by repeating sequences of pseudorandom numbers are to be avoided.
3. Weak coupling methods,¹⁰⁷ in which the average temperature of all atoms is controlled by a first-order coupling of the system to a heat bath. The statistical-mechanical character of the obtained ensemble is unclear. In the limit of very weak coupling, a microcanonical ensemble is generated, while in the limit of very strong coupling, a temperature-constrained (iso-kinetic) ensemble is produced. The size of the kinetic energy fluctuation is determined by the strength of the coupling.¹⁰⁸ Thus, the intensity of the fluctuations cannot be used to calculate thermodynamic properties.¹⁰⁷ An algorithm that scales the velocities of all the particles by a properly chosen random factor, has been shown to sample a canonical distribution and, in spite of its stochastic nature, a quantity can still be defined that remains constant during the simulation.¹⁰⁹
4. Extended system methods,¹¹⁰ in which an extra degree of freedom is added to the atomic degrees of freedom of the system. When using the extended system Lagrangian technique in the form of a Nosé–Hoover thermostat¹¹⁰ without a chain of thermostats, the second-order coupling of the system to the heat bath induces oscillations in the potential energy of the system, the period of which depends on the strength of the coupling to the heat bath.¹⁰⁸ This artifact can be avoided by using a chain of Nosé–Hoover thermostats.¹⁰³

When the actual temperature of a system is very different from the reference temperature, the application of a velocity-scaling or Nosé–Hoover chain thermostat may not be very robust. This is no problem for weak-coupling. Using a constraint, weak coupling, or extended system Lagrangian technique, the average temperature of all particles (or large subsets of particles, e.g. solutes or solvent) is controlled. Controlling the average may lead to an inhomogeneous temperature distribution within the set of particles when some particles are more prone to algorithmic (e.g., nonbonded cutoff, constraints) noise than others. Such inhomogeneities can be suppressed by coupling different subsystems to different heat baths or by using additional stochastic coupling. The pressure can be maintained by the same four types of methods.²

2.3.1.5. Application of Potential-Energy Terms Representing Experimentally Derived Data. The possibility to add an artificial potential-energy term to a physics-based force field is one of the assets of computer simulation. For example, in free-energy difference calculations, the sampling of configurational space can be enhanced or focused by using an umbrella sampling

potential-energy term, see [Section 2.3.1.7](#). Another example is the use of molecular simulation in structure determination or refinement of (bio)molecules, in which the potential-energy function of a molecular system is augmented by a potential-energy term representing experimentally derived data, which favors changes in the structure or conformation of molecules such that the experimentally derived data are reproduced in the simulation. About six decades ago, the idea of combining a model interaction function or force field with experimental data in a single energy minimization to obtain molecular structure matching the experimental data was used when determining molecular structure from X-ray diffraction intensities measured for molecular crystals. The use of a model force field was necessary, because the number of observed X-ray intensities was lower than the number of degrees of freedom in the crystalline unit cell (or asymmetric unit). The model force field was used to restrict the number of possible configurations of the molecular system. In this application the experimentally derived data are complemented by the information content of a relatively simple interaction function, generally containing bonded interaction and repulsive nonbonded interaction terms.¹¹¹ For larger molecules, such as proteins, searching the rather large conformational space was hindered by the many potential-energy barriers present, which could not be passed when applying energy minimization. In the 1980s the use of MD simulation instead of energy minimization solved this problem.^{112,113}

MD simulation of a protein in the gas phase¹¹⁴ and in a crystal surrounded by water¹¹⁵ had shown that the atoms possess considerable mobility, which implied that the experimentally observed diffraction or spectroscopic properties were in fact averages over time and space. In molecular simulation this feature was accounted for by introducing time-¹¹⁶ or molecule-¹¹⁷ averaging into the potential-energy term representing X-ray diffraction data,^{118,119} NMR nuclear Overhauser enhancement (NOE) data^{120,121} or ³J-coupling data.¹²² A ³J-coupling is a multiple-valued function of the corresponding torsional angle, which complicates the determination of the force due to the ³J-coupling restraining potential-energy term.¹²³ Using the local-elevation search technique,¹²⁴ the multiple-barrier problem can be solved.¹²³

For observable quantities Q such as X-ray reflection intensities, NOEs, or ³J-couplings, it is possible to formulate a function relating a Q -value to a particular molecular structure. For other observable quantities, such as S^2 order parameters or residual dipolar couplings (RDCs) derived from NMR experiments, the function relating Q to structure involves some average $\langle f \rangle$ of a function f over a Boltzmann ensemble of structures. This means that a molecular simulation using a potential-energy term involving S^2 order parameters or RDCs must involve the averaging $\langle f \rangle$ in addition to the averaging $\langle Q(\langle f \rangle) \rangle$ of the quantity Q .^{125,126}

The use of molecular simulation when determining molecular structure from experimentally derived data is currently common practice. For a review of the subject we refer to.¹²⁷

2.3.1.6. Search and Sampling Enhancement Techniques. A great many techniques to enhance the sampling of the relevant statistical-mechanical configuration space of molecular systems have been proposed and investigated ([Figure 5](#)).¹²⁸ Here we briefly review a few search and sampling enhancement techniques of theoretical or practical interest.

2.3.1.6.1. Modifying the Force or Molecular Dynamics Step Direction. In MD simulation the direction of a step in

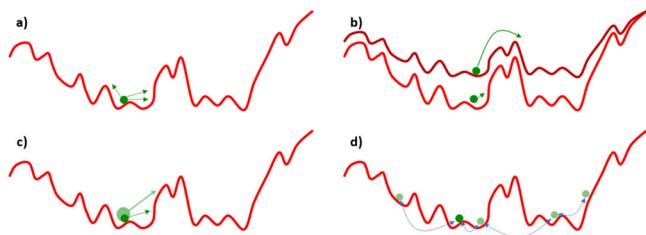


Figure 5. Approaches to enhance sampling: a) Adding stochastic forces to change the search direction regularly. b) Deformation or smoothening of the potential energy surface. c) Scaling of system parameters, e.g. adiabatic decoupling. d) Multicopy simulation with information exchange between simulations, e.g. replica exchange molecular dynamics simulation.

configuration space is determined by the force, the negative derivative of the potential-energy function, and by the inertia of the degrees of freedom, which serves as a short-time memory of the path followed so far. In SD simulation a random component is added to the force, the size of which is determined by the temperature of the system and the atomic masses and friction coefficients. In the potential-energy contour tracing (PECT) algorithm^{129,130} and in the potential-energy annealing conformational search (PEACS) algorithm¹³¹ the energy values are monitored and kept constant (PECT) or annealed (PEACS) to locate saddle points and pass over these. In random acceleration MD (RAMD)¹³² approaches, a randomly oriented force is added to a molecule to facilitate its dissociation from an active site of a protein ([Figure 5a](#)).

2.3.1.6.2. Deformation or Smoothening of the Potential-Energy Hypersurface. The searching or sampling of configurational space can also be enhanced by a deformation or smoothening of the potential-energy surface to reduce barriers.¹²⁸

1. The hard core of atoms, i.e. the strong repulsive interaction between atoms close to each other, is responsible for many barriers on the energy hypersurface of a molecular system. These barriers can be removed by making the repulsive short-range interactions between atoms soft.^{133–136} Soft-core atoms smoothen the energy surface and lead to strongly enhanced sampling.¹³⁷ In diffusion-equation type of approaches¹³⁷ or Gaussian accelerated MD^{138,139} a potential-energy term is added to the force field in order to flatten the energy landscape, lifting up valleys and reducing barriers ([Figure 5b](#)).
2. Incorporation of information on the energy hypersurface obtained during the search into the potential-energy function is another possibility to enhance sampling. Once a local energy minimum is found, it is removed from the energy surface by a suitable local deformation of the potential-energy function. This idea is the basis of “tabu” search,^{140,141} the deflation method,¹⁴² and the local-elevation search method,¹²⁴ later reinvented as metadynamics,¹⁴³ conformational flooding¹⁴⁴ and variations thereof.^{145,146}
3. Another way to introduce a memory into the search is the use of a potential-energy term which is a running average over the atomic trajectories or ensemble generated so far, rather than its instantaneous value.¹¹⁶ Application of this type of time-dependent or ensemble-dependent restraints in protein structure determination based on NMR or X-

ray data leads to much enhanced sampling of the molecular configuration space.^{119,120,147,148}

4. Barriers in the energy hypersurface can be circumvented by an extension of the dimensionality of the configuration space beyond the three Cartesian ones. By performing MD in four-dimensional Cartesian space, energy barriers in three-dimensional space can be circumvented¹⁴⁹ and free energy changes calculated.¹⁵⁰ Yet, this technique appeared not to be the most efficient one to enhance searching or sampling.

2.3.1.6.3. Scaling of System Parameters. Scaling of system parameters, e.g. temperature, atomic masses, potential energy, or forces, has been investigated with an eye to enhance searching and sampling of configurational space.¹²⁸ An example is enhanced sampling by a mean-field approximation in which a molecular system is separated into two parts, A and B, each of which moves in the average field of the other. The initial configuration of the system consists of N_A identical copies of part A and N_B identical copies of part B. The force on atoms in each copy of part A exerted by the atoms in all copies of part B is scaled by a factor $(N_B)^{-1}$, to obtain the mean force exerted by part B on the individual atoms of part A. Correspondingly, the force on atoms in each copy of part B exerted by the atoms in all copies of part A is scaled by a factor $(N_A)^{-1}$, to obtain the mean force exerted by part A on the individual atoms of part B. The forces between the different copies of part A are zero, and so are the forces between the different copies of part B. The MD simulation involves the integration of Newton's equation of motion for all copies of parts A and B simultaneously. Thus, one obtains N_A individual trajectories of part A in the mean field of part B and vice versa. This comes at the loss of correct dynamics: Newton's third law, action is minus reaction, is violated. The technique only enhances efficiency when the system is partitioned into parts of very different sizes, e.g. $\text{size}(A) \ll \text{size}(B)$ and the bigger part is represented by one copy: $N_B = 1$. Mean-field approximations have been proposed in different forms.^{151–156}

Scaling of system parameters may not only modify the velocities of atoms, but also the configurational distribution. This makes these techniques less useful. However, scaling of system parameters of one subsystem, one of two parts of a system, may avoid distortion of the configurational ensemble of the other subsystem in case the motions of the atoms of the two subsystems or two different sets of degrees of freedom, are adiabatically decoupled from each other. The basic idea of adiabatic decoupling is to prevent energy exchange between two sets of degrees of freedom of a system, while maintaining their mutual interaction. This can be achieved by scaling the masses of one set of degrees of freedom such that they differ so much from the masses governing the motions along the other set of degrees of freedom that the transfer of kinetic energy between them, through momentum conserving collisions, becomes very small.¹⁵⁷ Since the masses of the particles only appear in the momentum factor in the statistical-mechanical partition function, at least if no constraints are present in molecules with internal degrees of freedom, the contribution of the momentum can be analytically calculated and the configurational part of the partition function is independent of the masses. Thus, when applying adiabatic decoupling, the nondynamic equilibrium properties of a system remain unaltered. If one is not interested in the dynamical properties of a system, but only in the structural and thermodynamic ones, adiabatic decoupling

with scaling of system parameters for the subsystem of less interest may enhance the sampling of the other subsystem that is of interest (Figure 5c).

We consider a system in which a particular set of N^h (h : heavy) degrees of freedom is adiabatically decoupled from the other N^l (l : light) degrees of freedom of the system by increasing their mass $m^h = s_m m^l$ with $s_m \gg 1$ in such a way that the transfer of kinetic energy between the decoupled degrees of freedom through momentum conserving collisions becomes very small. Here we have denoted the masses of the atoms constituting the N^h degrees of freedom collectively by m^h and the ones of the N^l degrees of freedom collectively by m^l . This decoupling is more easily achieved if $N^h \ll N^l$, for example taking ions as h type and water molecules as l type particles in an ionic system. To enhance the sampling of the N^h h type degrees of freedom, their temperature, potential energy, or force can be scaled.¹⁵⁸ Adiabatic decoupling of the ionic (h type) and water (l type) degrees of freedom using a mass scaling factor $s_m = 100$ and up-scaling the temperature of the translational degrees of freedom of the ionic (h type) degrees of freedom by a factor of 2, $T^h = s_T T^l$ with $s_T = 2$, leads to an increase of a factor of 15 in the diffusion coefficient of the ions.¹⁵⁸ For a test system of 1000 water molecules at 300 K and a density of 997 kg m^{-3} , various fractions $N^h/(N^h + N^l)$ of water molecules were adiabatically decoupled to different degrees.¹⁵⁷ It was found that for $N^h/(N^h + N^l) \leq 0.1$ the diffusion of the N^h molecules could be enhanced by factors up to 35 depending on the method, the ratio $N^h/(N^h + N^l)$, the extent of adiabatic decoupling ($s_m = 10^2$ or 10^3), and the temperature (or force) scaling factor ($s_T = 2$ or 5), at the cost of a slight perturbation of the configurational distribution. The method of adiabatic decoupling combined with temperature or force scaling was also applied in a simulation of the folding equilibrium of a 7- β -peptide in methanol.¹⁵⁹ The peptide degrees of freedom were h type particles and the solvent ones were l type particles. The peptide temperature was up-scaled or the forces felt by the peptide atoms were down-scaled. In this case adiabatic decoupling of the solute degrees of freedom from the solvent ones did not bear fruit when aiming at enhanced sampling of the folding equilibrium of the solute. This result is in contrast to those of a similar study of ways to enhance the sampling of ionic degrees of freedom in aqueous solution,¹⁵⁸ mentioned above. The explanation of these seemingly contradictory results lies in the ratio of the interactions within the h type particle degrees of freedom, V^{hh} , and that between h type and l type particles, V^{hl} . In the ionic system V^{hl} is the dominant interaction whereas in the peptidic system the V^{hh} , the intrapeptidic interaction, plays a larger role with respect to the quantities that were analyzed. In another case¹⁵⁹ in which the potential energy is dominated by the light (solvent) particles, $V^l \ll V^{hl}$, adiabatic decoupling of the solute degrees of freedom from the solvent ones does not lead to a more efficient sampling of the solute degrees of freedom.

2.3.1.6.4. Multicopy Simulation with a Relation between Copies. In the original so-called replica-exchange algorithm multiple copies of the system are simulated by Monte Carlo (MC), MD, or SD, each at a distinct temperature.¹⁶⁰ From time to time copies at adjacent temperatures are exchanged using an exchange probability based on the Boltzmann factor $\exp(-E(x)/k_BT)$ (Figure 5d). This leads in the limit of infinite sampling to Boltzmann-distributed (canonical) ensembles for each temperature.¹⁶¹ So-called multicanonical algorithms are a generalization of this procedure.¹⁶² In Hamiltonian replica exchange, the replicas may connect different end states to

efficiently sample configurations in alchemical free energy calculations.^{163,164} Dynamical information is lost in the exchanges. A variety of schemes of this type has been proposed: generalized-ensemble algorithms,^{162,165,166} local and partial replica-exchange,¹⁶⁷ parallel replica method,¹⁶⁸ combinations of parallel tempering, multicanonical and multiple histogram methods,¹⁶⁹ and broad-histogram MC.^{170,171}

The so-called SWARM type of MD¹⁷² is based on the idea of combining a collection or swarm of copies of the system each with its own trajectory into a cooperative multicopy system that searches configurational space. To build such a cooperative multicopy system, each copy is, in addition to physical forces due to the molecular interaction function, subject to (artificial) forces that drive the trajectory of each copy toward an average of the trajectories of the swarm of copies, in analogy to the fact that intelligent and efficient behavior of a whole swarm of insects can be achieved even in the absence of any particular intelligence or forethought of the individuals. SWARM-MD is less attracted by local minima and is more likely to follow an overall energy gradient toward the global energy minimum. Other multicopy methods can be found in.^{173–175}

2.3.1.7. Guiding the Sampling Using a Biassing Potential-Energy Term, a Reaction or Transition Pathway, or a Hamiltonian Comprising End States. Dynamical processes in biomolecular systems may occur on time scales far beyond the ones that are accessible through standard MD simulations. If these processes are intrinsically slow, i.e. require an extensive sampling of configuration space, not much can be done to speed up their simulation without destroying the dynamics of the system. If, however, these processes are rare, i.e. they do not occur often, but when occurring they are fast, there are possibilities to enhance the sampling of these rare processes. Generally, they are characterized by the need to pass over a high-energy barrier separating two metastable states.

There exist a plethora of methods proposed to enhance the sampling along pathways or between so-called end states of a molecular system.¹²⁸ The oldest approach to sample transitions is to define a reaction coordinate or transition pathway and to sample along this path using a biassing potential-energy term and umbrella sampling.¹⁷⁶ Reaction coordinates can be simple distances or can be defined using more elaborate algorithms, such as a so-called distance field, which in turn can be combined with Hamiltonian replica exchange.¹⁷⁷ A method that requires no knowledge of the transition mechanism or transition state, is transition-path sampling, which finds transition pathways for infrequent events from a definition of the end states only.^{178–180} Although this method is more powerful than traditional reaction-coordinate sampling, the requirement of a proper definition of the two end states restricts its applicability. Other methods to determine and sample transition pathways are referred to in.¹²⁸

A more recently developed method to obtain the free energy difference between two or more states of a molecular system is so-called Enveloping Distribution Sampling.^{181–184} It combines the Hamiltonians or potential-energy functions representing different systems or states A_i ($i = 1, 2, \dots, N$) of a molecular system into one single Hamiltonian or potential-energy function of a reference state R . The reference state R is related to the sum of the Boltzmann factors of the N end-state Hamiltonians. Energy offsets, E_i^R , are introduced to provide more even sampling of all end states,¹⁸⁴ while different approaches exist to increase the frequency by which they are sampled, either using a smoothness parameters s_i or by smoothening (accelerating)

the resulting hybrid potential energy.¹⁸⁵ When using the smoothness parameters, multiple parallel simulations using different values of s_i can be combined in a replica exchange setting.¹⁸⁶ EDS has been applied to calculate octanol–water partition coefficients for a number of solute molecules using a twin-system technique that involves the simultaneous simulation of two computational boxes corresponding to two environments 1 (octanol) and 2 (water) and the simultaneous conversion of solute X to solute Y in box 2 with the conversion of solute Y to solute X in box 1.¹⁸⁷ EDS has also been applied to calculate the free enthalpy differences between different folds of a solute,^{188–192} the calculation of relative binding free energies^{185,193–195} and in combination with other techniques to enhance the sampling.^{185,189,190,194}

2.3.1.8. Application of a Time-Independent, External Field or Force, or Weak Coupling to a Reservoir. A number of properties of molecular systems can be calculated from fluctuations of a quantity in an equilibrium simulation (see Section 2.4 below). An alternative way to calculate particular properties is the application of an external field, force or the coupling of a system quantity to a reservoir. In these cases the system is in a steady or stationary state. Here a few techniques are discussed.

Linear-response theory offers a relation between an external field or driving force, for example an electric field (E) or a temperature (T) or velocity (v) gradient on the one hand and a quantity Q , for example the polarization or energy flux or momentum flux, on the other hand. From the linear relation between two of such quantities the proportionality constant of the relation, for example the dielectric permittivity or the thermal conductance or the shear viscosity, can be calculated. A profile of Q can be imposed in the simulation and the average response E calculated from the simulation, as in the homogeneous deformation method which uses so-called Lees–Edwards periodic boundary conditions.¹⁹⁶ An alternative is to impose a field E and to calculate the average response Q .^{197,198} For a polarizable molecule without permanent dipole moment, such as carbon tetrachloride, the determination of the dielectric permittivity from the dipole moment fluctuations of the system is not feasible. In such a case applying an electric field is the sole alternative.¹⁹⁹

Instead of applying a homogeneous field, a sinusoidal form can be used, which matches the periodicity of the computational box in combination with weak coupling¹⁰⁷ of the temperature (T) or the velocities (v) to an external reservoir of these quantities and measuring the flux of energy or acceleration from the reservoir.²⁰⁰ The steady-state equation relating energy flux or acceleration on the one hand and spatial temperature or velocity profile on the other hand yields an expression for the corresponding transport coefficient, thermal conductance or shear viscosity.²⁰⁰

Of the many search and sampling methods and enhancement techniques reviewed a few are very effective: use of soft-core atoms and adiabatic decoupling, local-elevation simulation and its derivatives, replica-exchange simulation and generalized-ensemble methods. When end states are known, transition path sampling and enveloping distribution sampling are powerful methods.

The methods discussed so far all sample the configurations of a molecular system in equilibrium or nearly in equilibrium, at least if the simulation is sufficiently long, because the potential-energy function is time-independent, except when time-averaging or local-elevation type of force-field terms are present.

We note, however, that the methods that involve a steady or stationary state, discussed here under equilibrium simulations, are often classified as nonequilibrium molecular dynamics (NEMD) methods.

2.3.2. Nonequilibrium Simulations. Nonequilibrium statistical mechanics describes nonequilibrium ensembles and relaxation of a molecular system from a perturbation that is switched off or on at a given point in time toward equilibrium.²⁰¹ It is commonly assumed that in a molecular simulation the system is simulated long enough to have reached equilibrium. For liquids of relatively small molecules this is well possible. For macromolecular systems there may be degrees of freedom that have not reached equilibrium. This can be due to rather high energy or entropy barriers present in the system. A fast-changing parameter in a simulation may also prohibit the system to reach equilibrium. When applying time-averaging restraining of particular quantities in a simulation or in local-elevation sampling and similar methods the system is locally perturbed and continuously relaxing to equilibrium. Yet, the changes in these forces are generally rather small, which allows the system to be considered in equilibrium. When simulation parameters are relatively fast changed during a simulation, this may not be the case. The system is permanently out of equilibrium. **Figure 6**

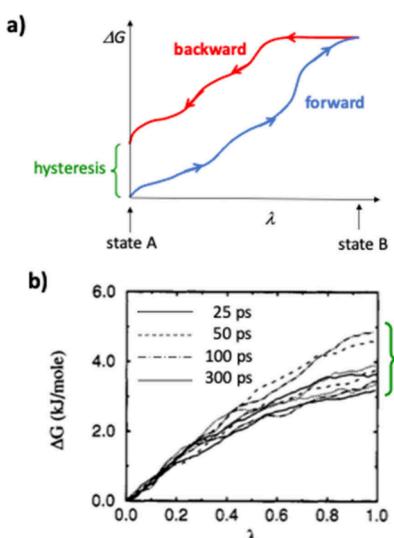


Figure 6. Occurrence of hysteresis in simulations that keep the system out of equilibrium. Schematic representation of a slow-growth simulation, in which the coupling parameter λ is switching the system from state A to state B continuously during a simulation. The hysteresis is the free energy difference between the forward and backward simulations. Adapted with permission from ref.²⁰² Copyright [1994] [ACS].

and **Table 3** illustrate such a case that often occurred in early calculations of free-energy differences using the so-called slow-growth technique, in which a parameter that brings the system from one state to another is slowly changed,²⁰² but not slow enough to let the system reach equilibrium, with as a result hysteresis. The occurrence of hysteresis (nonreversibility) is dependent on two factors: (1) the relaxation time of the system in response to a particular mutation in a molecule and (2) the time needed to sample the states which make up the equilibrium ensemble of the system. If the simulation length is much longer than these two (relaxation) times, no hysteresis will occur and the obtained free energy difference will be correct (adequate

Table 3. Calculated Free Energies and Hysteresis (kJ mol^{-1}) for the Mutation of *p*-Chloro-phenol to *p*-Methyl-phenol in (SPC) Water and within α -Cyclodextrin in Water for Various Simulation Times t^{MD} Using Slow Growth²⁰²

t^{MD} (ps)	in water		within α -cyclodextrin in water	
	ΔG	hysteresis	ΔG	hysteresis
25	-4.33	0.25	3.68	0.51
50	-4.90	1.22	4.16	0.77
100	-4.50	0.21	4.18	1.48
300			3.42	0.51

sampling). If the simulation length similar to these two (relaxation) times, hysteresis will occur and the obtained free energy difference will not be correct (partial sampling). If the simulation length is much shorter than these two (relaxation) times, no hysteresis will occur, but the obtained free energy difference will not be correct (almost no sampling). These three regimes are illustrated in **Table 3**.

An example of true nonequilibrium simulation is the use of the Jarzinsky relation^{203,204} to calculate free energy differences.²⁰⁵ Jarzinsky derived that an equilibrium free energy difference can be obtained from an ensemble of nonequilibrium free energy estimates, $\Delta F^{\text{fg}} = -k_B T \ln(\langle \exp(-\Delta F_i^{\text{fg}}/k_B T) \rangle_i)$, where the average $\langle \dots \rangle_i$ is over many nonequilibrium estimates ΔF_i^{fg} obtained from slow-growth or slow-change simulations started from an appropriately equilibrated sample of initial structures.²⁰⁶ The Jarzinsky equation has seen several applications on various test systems,^{207,208} to calculate the free energy of charging ions²⁰⁹ or solvating small molecules²¹⁰ and in the calculation of potentials of mean force between particles^{211–213} or on conformational changes.²¹⁴ It seems that using a given amount of simulation time, similar accuracies can be obtained using the fast-growth Jarzinsky relation as using standard methodology to calculate free energy differences.²⁰⁵ Two sided approaches include the Crooks method, in which non-equilibrium simulations from both end-states are combined to obtain an equilibrium estimate,²¹⁵ or the Crooks Gaussian Intersection method, which assumes a Gaussian distribution of the nonequilibrium estimates.²¹⁶ An advantage of using nonequilibrium methods is the ease of parallelization, as individual nonequilibrium simulations may be performed on separate computer nodes.

2.4. Calculation of Properties. A variety of properties of molecules or of a molecular system can be calculated from molecular simulations. For some types of properties one technique may be more efficient than another. For example, correlation functions, consisting of a product of a quantity at time t_n and a quantity at time t_m require about N^2 multiplications for a trajectory of N configurations. By applying a Fourier transform to the trajectory (N operations) correlation functions are more efficiently calculated.²¹⁷

2.4.1. Atomic or Molecular Properties. Properties that depend on atomic coordinates, such as radial distribution functions, local mole fractions, self-diffusion, rotational correlation times, can relatively straightforwardly be calculated from atomic trajectories, see e.g.^{218,219} X-ray diffraction intensities require a definition of electron density.²²⁰ Expressions in terms of atomic coordinates for quantities measurable by NMR, such as NOE intensities,²²¹ 3J -couplings,¹²³ S^2 order parameters¹²⁵ and residual dipolar couplings,¹²⁶ are available.

2.4.2. System Properties. Expressions for the excess free energy, mixing enthalpy, excess volume, surface tension, hydrophobic solvation, relative static dielectric permittivity, Debye relaxation time, heat capacity, thermal expansion coefficient, isothermal compressibility are available, see for example,^{218,219}

According to Onsager's regression hypothesis, the relaxation of macroscopic disturbances is governed by the same physical laws as the regression of spontaneous microscopic fluctuations in a system at equilibrium. This means that some system properties can either be calculated from fluctuations of quantities in a simulation or from simulations in which a restraining potential-energy term is applied, a particular quantity is constrained, or an external field is applied (see Section 2.3.1.8), see e.g.²²²

1. In the unrestrained method the quantity Q is calculated from fluctuations of a related quantity M during a single simulation.
2. The quantity Q can also be computed from a series of nonequilibrium simulations, during which the relaxation of the related quantity M toward equilibrium is observed after a time-dependent disturbance was applied or removed. Note that in such a setting, one simulation leads to one observation and appropriate ensemble averages are only obtained when a sufficient number of simulations are performed.
3. Using the umbrella sampling method (probability) distributions $P(M)$ are obtained from multiple simulations involving a biasing potential-energy term to restrain the fluctuations of the quantity M around specific values.
4. Using the constraint method the mean constraint force exerted on M , obtained from multiple simulations at specific constrained values of M is calculated and integrated.

2.4.3. Free-Energy Differences. A great variety of methods to compute free energy differences is available.²²³ Once an appropriate set of configurations has been sampled, various approaches can be used to estimate the difference in free energy: (1) Counting the number of visits of a configuration or state; (2) Using energy differences between two states A and B ; (3) Integrating the derivative of a free energy with respect to a parameter λ defining a pathway from state A to state B ; (4) Estimation of transition probabilities from A to B . We note that these methods have now become standard tools in the search for pharmaceutically active compounds.

3. LIMITATIONS OF CLASSICAL-PHYSICAL SIMULATIONS

The limitations of classical-physical simulations stem from the various assumptions and approximations involved.⁸⁵

3.1. Quantum Effects. Many molecular processes can be reasonably well described using purely classical-mechanical (CM) methods. However, when relatively light particles such as electrons or light nuclei such as the proton are involved, quantum effects will come into play, which can be modeled using a variety of quantum-mechanical techniques. In chemistry, mainly quantum effects involving the motion of atoms, electrons and nuclei are of relevance. Which quantum effects are to be accounted for depends on the molecular degrees of freedom that are considered in a simulation. In principle, the energy of a molecular system is quantized, that is, the system can only adopt

states (energy levels) with discrete values of energy. The electronic energy levels of a molecule are generally very widely spaced, which means that at physiological temperatures only the ground state will be populated. Thus, in atomistic simulation, the atoms and molecules are generally considered to be in their electronic ground state and electronic degrees of freedom are not treated explicitly. When electronic degrees of freedom play an essential role in a molecular process, a quantum-mechanical treatment of such degrees of freedom is required. This is the case when modeling changes of chemical bonds or electronic relaxation processes. Chemical reactions are essentially governed by quantum mechanics. Yet, one may attempt to mimic the result of a chemical reaction in a classical-mechanical simulation using a so-called "reactive" force field. One approach is e.g. to monitor if reactive groups get close in a classical simulation and then perform a stochastic attempt (using a Metropolis MC criterion) to change the reactants into their products classically. If the classical potential-energy function is of good quality and the MC move is carried out correctly, a classical trajectory in which a reaction has occurred is obtained. Such a procedure does not mimic a reaction, but samples the effect of a reaction.

3.2. Treatment of Boundaries. In molecular simulation the system is generally contained in a periodic box. The periodicity may induce distortions in the molecular system. This is in particular the case when electrostatic interactions are calculated using lattice-sum methods (see Section 2.2.1.2). Significant differences in the average end-to-end distance of an octa-alanine peptide in lattice sum simulations with a minimal distance between the peptide and the box wall up to 1.7 nm.²²⁴ Reaction-field methods to calculate long-range electrostatic interactions also suffer from misrepresentation, in particular if the medium beyond the electrostatic interaction cutoff is not homogeneous.

3.3. Coarse Graining. The process of coarse-graining may reduce the usefulness of the model in different ways.⁵

1. The range of thermodynamic state points at which it may be applied is generally reduced.
2. The transferability of model parameters between similar but not identical moieties or compounds may be reduced.
3. The accuracy of various properties may be reduced. The occurrence of specific hydrogen bonds (donating or accepting), for example, will be difficult to monitor if a single bead is used to describe e.g. a glutamine side chain.
4. The physical basis for a particular property or process may be changed, leading to an unphysical mechanism of the process in the coarse-grained model.
5. The reduction of entropy and energy in the system may lead to an unphysical balance between these quantities in the coarse-grained model.

As mentioned in Section 2.1, the combined loss of usefulness on these five counts must be made up for by a much increased computational efficiency of the coarse-grained model. This is often not the case.

3.4. Consistency between Different Levels of Resolution or Modeling. When involving different levels of modeling or grain levels in one simulation, particular problems may arise. For example, when combining an atomic water model (FG) with a supra-molecular one (CG), a mixture of molecules of the two types of water models may not show the same values of properties as the two models themselves. This is illustrated in Figure 7,²²⁵ which show that properties of water change as a

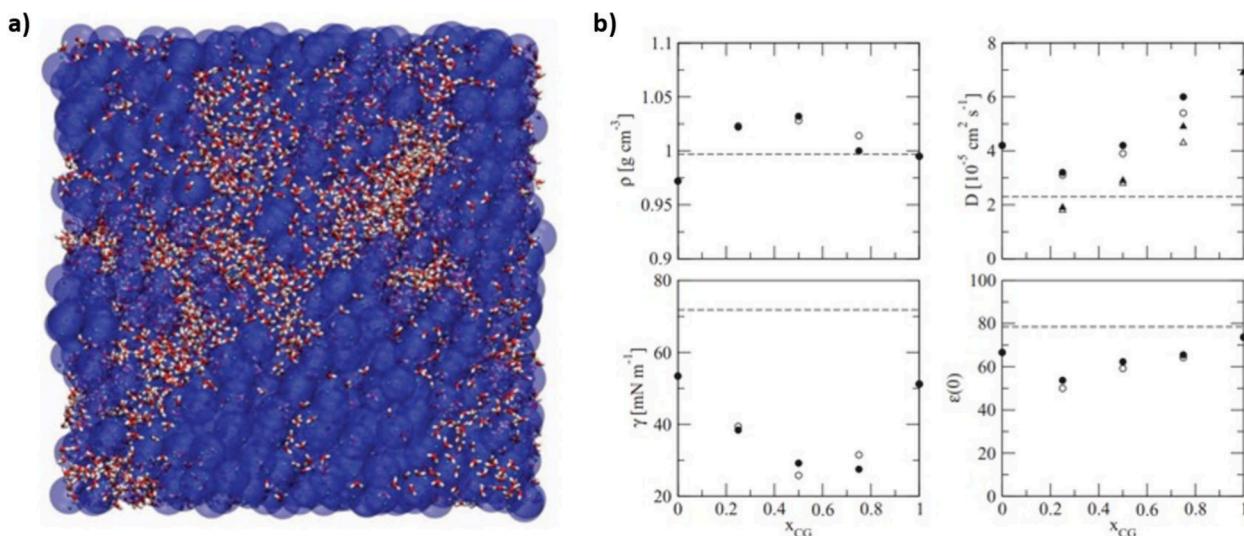


Figure 7. Mixed fine-grained (atomic) and coarse-grained (supra-molecular) simulations of liquid water. a) Visualization of the mixing of both an atomistic and a supra-molecular polarizable model. b) Density ρ ; self-diffusion coefficient of the oxygen atom of fine-grained (SPC) water D_{OW} (circles) and the central site of coarse-grained beads D_{CW} (triangles); surface tension γ ; and relative static dielectric permittivity $\epsilon(0)$ at 298 K and 1 atm as a function of the coarse-grained water mole fraction x_{CG} . The data is shown for different estimates of the effective relative dielectric constant for the interaction between fine- and coarse-grained interactions (filled symbols: 1.9; open symbols: 2.3). The experimental values are shown with the dashed line. Adapted with permission from ref.²²⁵ Copyright [2012] [AIP].

function of the mole fraction x_{CG} of supra-molecular water molecules. Ideally, this should not be the case.

When combining a quantum-chemical model for one part of a molecular system with a classical model for another part, inconsistencies between the forces due to these differently treated subsystems may occur and should be minimized.

Similarly, when simulating a (bio)molecule in (aqueous) solution, the models or force fields used for solute and solvent should be compatible. Otherwise energetic and entropic imbalances may occur in the system.

4. PERSPECTIVES

The challenges and opportunities of improving classical-physical simulations follow from the various limitations, assumptions and approximations involved.⁸⁵

4.1. Force-Field Parameter Calibration. A classical potential-energy function or force field consists of a number of force-field terms for bonded and nonbonded interactions between atoms and constitutes an approximation of the quantum-mechanical energy surface of a molecular system. The functional form of the various force-field terms should be kept as simple as possible and on a physical basis in order to allow the transfer of force-field parameters between similar fragments of large molecules.^{226,227} The balance between the various terms of a force field, that is, their relative weights, is to be chosen such that the total potential energy best represents the underlying quantum-mechanical energy surface, given the classical-mechanical approximation. In this respect, machine-learned potential-energy functions represent a black-box kind of force field that reproduce the quantum-mechanical energy surface as well as possible.^{22,58,59} An increased precision may come at the expense of a lack of physical interpretability of the interaction energy and forces and a lack of transferability between different kinds of molecules.

If solvent molecules are simulated explicitly, the solvent model used should be compatible with the solute model. Since entropy plays an essential role in many molecular equilibria in

the condensed phase, force-field parameters need to be calibrated not only to energetic data such as the heat of vaporisation, but also to measurable free energy data, such as excess free energy, or surface tension of the liquid and free enthalpies of solvation. And, this calibration should be carried out for as many small molecules, for which experimentally measured data are available, as possible.

4.2. Hybrid Quantum/Classical Simulation Techniques. Although molecular systems containing thousands to millions or even billions of atoms can be classical-mechanically (CM) simulated due to an ever-increasing computing power, this is not the case for quantum-chemical calculations, which require orders of magnitude more computing power. A quantum-chemical calculation for not too large solute molecules would be possible, depending on the level of accuracy and thus computing effort required, but for the very many solvent degrees of freedom of a molecular system in the condensed phase, this is a sheer impossibility, in particular considering the time periods of simulation required for many processes. These considerations led to the development of multiresolution hybrid QM/CM simulation techniques, in particular to be applied to molecular systems that contain parts that are fundamentally quantum-mechanical (QM) in nature. Examples are proton transfer reactions which require a hybrid quantum-/classical-dynamical (QD/CM) treatment, because of the small mass of the proton.^{228–230} The further development of multiresolution hybrid QM/CM simulation techniques in which the electrons are treated quantum-chemically,^{231,232} is required, in particular improvements of the manner in which the two alternative, fundamentally different QM and CM representations are combined or merged. Inclusion of a buffer region, facilitated by machine-learned potential-energy functions, was recently suggested.²³³

4.3. Accounting for Molecular Polarizability. When electronic degrees of freedom play an essential role in a molecular process, a quantum-mechanical treatment of such degrees of freedom is mandatory. This is the case when

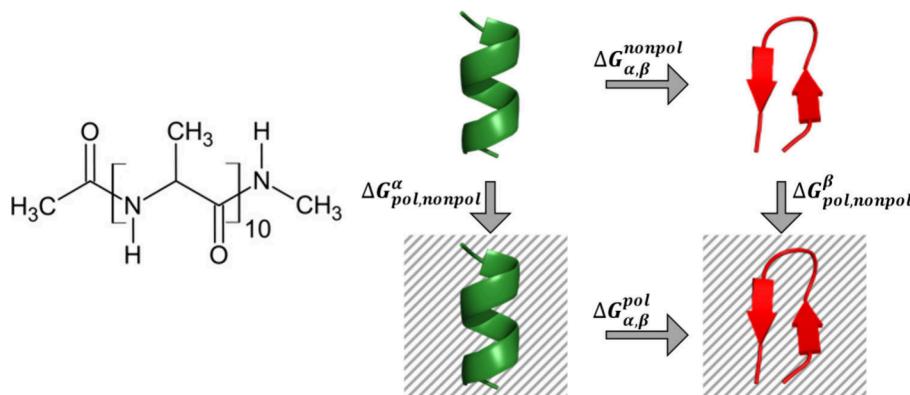


Figure 8. Effect of polarizable versus nonpolarizable solvent models on the relative stability of an α -helical and a β -hairpin structure of an alanine decapeptide. The less polar the solvent is, the more the α -helical structure is stabilized with respect to the β -hairpin structure, up to 20 kJ/mol (Table 4), due to the inclusion of polarizability in the solvent.²³⁶ Left-hand side adapted with permission from ref.²³⁶ Copyright [2015] [ACS].

modeling changes of chemical bonds or electronic relaxation processes. Different levels of accuracy can be achieved, ranging from high accuracy for computationally highly expensive ab initio methodology to computationally much cheaper semi-empirical and valence-bond methods. If the rearrangement of electronic density in response to a change in the molecular environment is small, a classical treatment using atomic or molecular polarizability or charge transfer between atoms may be appropriate. Although it was long realized that polarization degrees of freedom should be taken into account when developing molecular models,^{2,234} only the past decades saw the development of generally applicable polarizable force fields for (bio)macromolecules such as proteins,^{54–56} because of the considerable effort required. One may raise the question though whether it is at all possible to mimic the quantum-chemically determined electron distribution of a protein, a heterogeneous flexible macromolecule, by a set of classical inducible dipoles that will strongly interact because of their proximity to each other. The polarizable force fields for proteins thus need extensive testing which has not yet been done to the same extent as for the nonpolarizable force fields for proteins. A corresponding consideration about a proper classical-physical representation of quantum phenomena led to the representation of bond-stretching, bond-angle and torsional-angle interactions in macromolecular force fields by simple phenomenological terms separate from the general nonbonded interactions.

In view of these considerations, a future alternative may be to treat the entire macromolecular solute quantum-mechanically thereby automatically including polarization. This would avoid the unphysical, distortive effects of nearby induced classical dipoles upon each other. Regarding the many thousands of solvent degrees of freedom needed when explicitly solvating a macromolecule, there is no quantum-chemical alternative to treating electronic polarizability of solvent molecules using classical inducible dipoles. To this end, polarizable models for common solvents such as water, chloroform, dimethyl sulfoxide, methanol, and mixtures thereof were developed that are compatible with nonpolarizable solute models see e.g.²³⁵ and references therein. The use of polarizable solvent models is particularly important when studying folding equilibria of polypeptides in less polar or nonpolar solvents.²³⁶ Figure 8 and Table 4 illustrate that introducing polarizability into a water solvent does not change the relative stability between the α -helix and the β -hairpin, while for methanol and chloroform the α -helix is stabilized by about 1 kJ/mol⁻¹ per residue and for carbon

Table 4. Relative Stability (Free Enthalpy Difference in kJ mol^{-1}) between an α -Helical Structure and a β -Hairpin Structure of an Alanine Decapeptide When Changing a nonpolarizable solvent into a polarisable solvent for the peptide solvated in water, methanol, chloroform, or carbon tetrachloride.²³⁶

solvent	molecular polarizability α ($4\pi\epsilon_0 10^{-3} \text{ nm}^3$)	$(G_{\text{pol}}^{\alpha} - G_{\text{nonpol}}^{\alpha}) - (G_{\text{pol}}^{\beta} - G_{\text{nonpol}}^{\beta})$ (kJ mol^{-1})
H_2O	0.93	0.5 ± 3.7
CH_3OH	1.32	-11.5 ± 3.5
CHCl_3	9.5	-10.2 ± 1.4
CCl_4	10.5	-21.0 ± 0.6

tetrachloride by about 2 kJ mol⁻¹ per residue. These results suggest that the less polar the solvent is, the more the α -helical structure is stabilized with respect to the β -hairpin structure by the use of a polarizable solvent model instead of a nonpolarizable one. This highlights that inclusion of polarizability in models for less polar and nonpolar solvents or for protein-like environments is as important as, if not more important than, including polarizability in models for liquid water. In biomolecular systems, the next most abundant type of molecules, apart from the solvent, are lipids, which contain quite polarizable carbon atoms. This makes these molecules the next type for which a polarizable model compatible with a quantum-chemical treatment of the solute or a nonpolarizable solute force field is to be developed. As aliphatic tails constitute a major part of lipids and membranes, their force-field parameters can be developed using the measured thermodynamic and dielectric properties of liquid alkanes as calibration data.²³⁷

4.4. Representing Entropy in Molecular Simulation.

About 50 years ago, reports on the first molecular dynamics (MD) simulations of liquid water at ambient conditions of temperature and pressure showed a highly mobile picture of water molecules forming and breaking hydrogen bonds and emphasized the need to include motion or configurational variability when aiming at an accurate description of the properties of compounds in the liquid phase. In other words, when proposing a molecular model not only energetic contributions but also entropic contributions to the free energy of a liquid are to be taken into account.²³⁸ For water, the difference in entropy between gas and liquid phase, is about 20 kJ mol⁻¹ about half of the size of its heat of vaporisation of 44 kJ mol⁻¹. This is why molecular models for liquid water and other

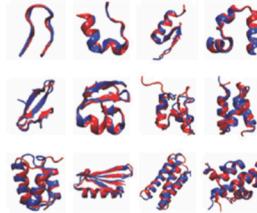
Year	Molecular system (type, size)	Length of the simulation [s]	
1957	First molecular dynamics simulation (hard discs)		
1964	Atomic liquid (argon)	10^{-11}	
1971	Molecular liquid (water)	5×10^{-12}	
1977	Protein in a vacuum	2×10^{-11}	
1983	Protein in water	2×10^{-11}	
1989	Protein-DNA complex in water	10^{-10}	
1997	Polypeptide folding in solvent	10^{-7}	
2001	Micelle formation	10^{-7}	
2010	Folding of a small protein	10^{-3}	
	And the future ...		
2029	Biomolecules in water (folding sooner?)	10^{-3}	
2034	<i>E. Coli</i> bacteria (ca. 10^{11} atoms)	10^9	
2056	Mammalian cell (ca. 10^{15} atoms)	10^9	
2080	Biomolecules in water (as fast as nature)	10^6	
2172	Human body (ca. 10^{27} atoms)	1	

Figure 9. Time-line of hallmarks of molecular simulation using computers, with an extrapolation to prediction as presented in 2006.³ The realization of the predictions for 2010 and 2029 by D.E. Shaw research can be found in.^{242,243} Figure inset: Agreement between simulated (blue) and experimentally determined (red) protein structures. Reprinted with permission from ref.,²⁴³ Copyright [2011] [AAAS].

biomolecular solvents are to be developed and their parameters calibrated based on measured thermodynamic data for such compounds and configurational ensembles generated by MD or Monte Carlo (MC) simulation. Entropic contributions are also non-negligible when considering biomolecular processes such as polypeptide folding or protein–ligand association. The enthalpy and entropy of solvation of biomolecules may vary depending on the cosolvents present in an aqueous solution, thereby leading to different, enthalpy- versus entropy-dominated, mechanisms of partitioning or association. In order to properly model such processes, the molecular models involved should not only represent energetic but also represent entropic effects, the distribution of molecular configurations (Figure 4). In particular, when coarse-graining the solute or solvent model for polypeptide folding simulations, the relative contributions to or balance between enthalpic and entropic contributions to the folding process need to be maintained, assuming that it is correct at the finer-grained level of resolution. Otherwise, the mechanism of folding will not correspond to the real physical one. If the free-energy surface of folding is incorrectly modeled,²³⁸ also the kinetics of folding will be incorrectly represented.²³⁸

4.5. Coarse Graining beyond the Atomic or Molecular Level of Resolution. The scale invariance that lies at the heart of the renormalization group approach to coarse-graining of largely homogeneous polymers is not observed for biopolymers, which are composed of many different, complex structural units that are connected through different types of interactions. In the coarse-graining process, the basic geometry and the balance between the various interactions must be preserved in order to avoid losing characteristic features of these molecules.¹⁷ In addition, entropy plays a non-negligible role in biomolecular processes, which means that the loss of entropy in the process of coarse-graining must be compensated for by a loss in energy in order to maintain the relevant free energy differences. Finally, the reduction of the computational effort when moving from atoms to supra-atomic beads is rather modest compared to that between other levels of resolution.⁵ These considerations lead to the conclusion that coarse-graining at the level of supra-atomic beads (Figure 1c) seems not to pay off for biomolecules such as proteins, DNA, RNA and sugars. Only a limited decrease in the

number of interaction sites is reached at the cost of losing the essential characteristics of such molecules in terms of intramolecular interactions, interactions with the solvent and entropy. Only lipids, which have relatively long homogeneous aliphatic tails, may be able to retain the dominant characteristics of an amphiphilic molecule with a particular geometry and flexibility upon coarse-graining from atoms to supra-atomic beads. Due to the abundance of lipids in membranes, the reduction in computational cost may offset the loss of accuracy.

Since the inclusion of solvent degrees of freedom is essential to properly represent the hydrophobic effect and because the calculation of the solvent–solvent interactions in a simulation of a biomolecule such as a protein or a fragment of DNA in aqueous solution dominates the computational effort, coarse-graining of the solvent degrees of freedom holds some promise to reduce the computational costs, in particular when more than one solvent molecule is subsumed into a supra-molecular bead. In the case of water, such coarse-graining from atoms to supra-molecular beads (Figure 1b) should retain the thermodynamic and dielectric screening properties and hydrogen-bonding capacity of water as much as possible, and a proper ratio between entropy and energy.⁷ This is not the case if a water bead is modeled as a Lennard-Jones particle without charge.^{239–241} Coarse-graining of solvent degrees of freedom in a biomolecular simulation has a good chance of meeting conditions 1 to 4 of Section 2.1, depending on how the coarse-grained interaction between solute and solvent is modeled.

5. CONCLUSION

Over the past half century classical-physical simulation of molecular systems has become a solid part of chemistry. Many methodological ideas were formulated and tested. A number of those did not become standard in molecular simulation, whereas others are commonly used. Examples of the former are (1) Integration of the classical equations of motion for macromolecules expressed in internal, i.e. non-Cartesian, coordinates, because it is, unfortunately, much less efficient than using Cartesian coordinates and constraints applied through Lagrange multipliers. (2) Considering the heterogeneous character of many (bio)molecular systems, the use of mean-field approx-

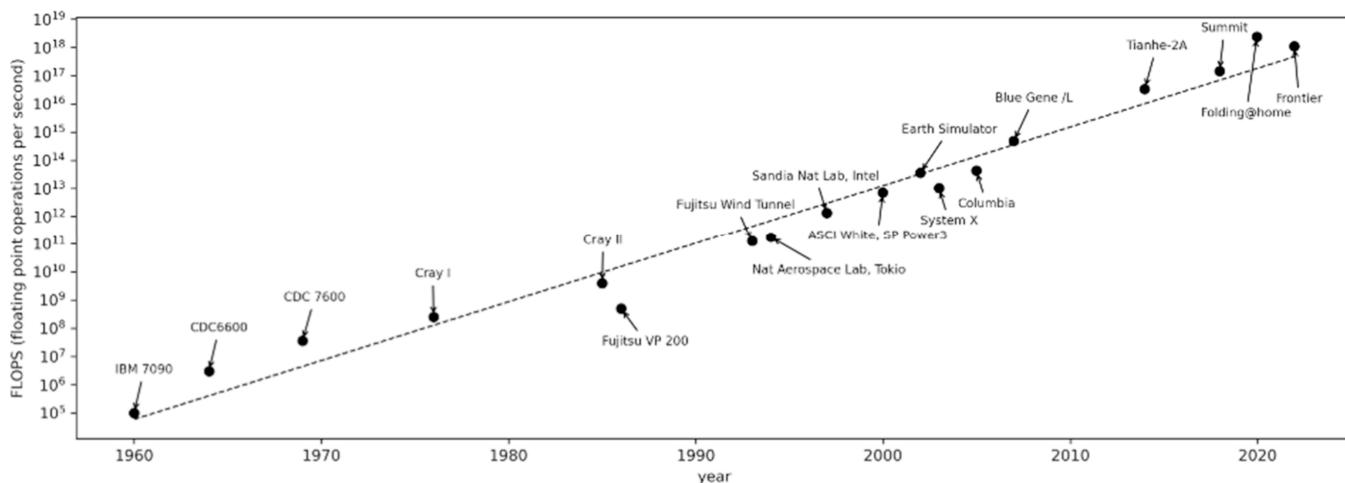


Figure 10. Development of computing power of the most powerful computers.

imations, which generally require a homogeneous system, seems to have severe limitations. (3) Coarse graining beyond the atomic level of modeling for heterogeneous macromolecular systems seems not viable because of the loss of detail and therefore accuracy.

In order to faithfully represent the physical processes underlying chemical phenomena not only the relevant degrees of freedom should be sampled in a simulation, but also the parameters of the classical molecular force field applied should have been calibrated using free energy data on a great variety of compounds representing the different moieties present in macromolecules. In addition, it is advised to keep the force fields as simple as possible in order to enhance the transferability of its parameters between different compounds. The conducting box boundary method (CBBM) to evaluate electrostatic interactions can be used to avoid the periodicity artifacts inherent to the common lattice-sum methods.

The past decades have shown a steady increase of the size of the molecular systems simulated and of the length of molecular simulations, see **Figure 9**. This is primarily due to the development of computing power, about a factor of 10 every five years (**Figure 10**). Yet, techniques proposed to enhance the sampling of the vast configurational space of macromolecular systems have also contributed to an improvement of the efficiency and accuracy of molecular simulation. The most important ones are

1. The application of (bond-length) constraints.
2. Application of constraints allowed much simpler time-integration algorithms, the Störmer/Verlet and the leapfrog algorithms, than the initially used high-order Gear-type algorithms to be used to integrate the equations of motion.
3. The constant-force multiple-time-step integration of the equations of motion applied to the long-ranged part of the nonbonded interactions.
4. Representation of experimentally derived data by an additional term in the molecular interaction function, including time-averaging to properly represent the measurement.
5. Use of restraining terms in the potential-energy function to focus or bias the sampling of configurational space followed by an unbiasing of the generated ensemble.

6. Smoothening of the potential-energy surface, e.g. by use of soft-core nonbonded interactions.
7. Coarse graining to a sensible degree.
8. Local-elevation type of techniques.
9. Adiabatic decoupling type of techniques.
10. Replica-exchange simulation.

Regarding the computation of free-energy differences between different states pathway methods such as umbrella sampling, transition path sampling or enveloping distribution sampling seem most useful.

Over the past half century much has been achieved which made molecular simulation more accurate. Yet, there still is ample room for further development of classical simulation methodology with an eye to enhancing its efficiency and its accuracy. For the latter the development of more accurate potential-energy functions (force fields) for (bio)molecular simulation, at any level of resolution, remains the most central and greatest challenge.

■ ASSOCIATED CONTENT

Data Availability Statement

The manuscript contains no newly generated data, all data can be found through references to the literature.

■ AUTHOR INFORMATION

Corresponding Author

Wilfred F. van Gunsteren – *Institute for Molecular Physical Science, Swiss Federal Institute of Technology, ETH, CH-8093 Zurich, Switzerland; orcid.org/0000-0002-9583-7019; Email: wfgvn@ethz.ch*

Author

Chris Oostenbrink – *Institute of Molecular Modelling and Simulation, BOKU University, 1190 Vienna, Austria; Christian Doppler Laboratory for Molecular Informatics in the Biosciences, BOKU University, 1190 Vienna, Austria; orcid.org/0000-0002-4232-2556*

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jcim.4c00823>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The help of Michael Gillhofer in making the figures is gratefully acknowledged. Financial support by the Austrian Federal Ministry for Digital and Economic Affairs, the National Foundation for Research, Technology and Development, the Christian Doppler Research Association, BASF and Boehringer Ingelheim is gratefully acknowledged.

REFERENCES

- (1) Rahman, A. Correlations in the Motions of Atoms in Liquid Argon. *Phys. Rev. A* **1964**, *136*, 405–411.
- (2) van Gunsteren, W. F.; Berendsen, H. J. C. 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.
- (3) van Gunsteren, W. F.; Bakowies, D.; Baron, R.; Chandrasekhar, I.; Christen, M.; Daura, X.; Gee, P.; Geerke, D. P.; Glättli, A.; Hünenberger, P. H.; Kastenholz, M. A.; Oostenbrink, C.; Schenk, M.; Trzesniak, D.; van der Vegt, N. F. A.; Yu, H. B. Biomolecular Modelling: Goals, Problems, Perspectives. *Angew. Chem., Int. Ed.* **2006**, *45*, 4064–4092; *Angew. Chem.* **2006**, *118*, 4168–4198.
- (4) van Gunsteren, W. F.; Dolenc, J. Thirty-Five Years of Biomolecular Simulation: Development of Methodology, Force Fields, and Software. *Mol. Simul.* **2012**, *38*, 1271–1281.
- (5) Riniker, S.; Allison, J. R.; van Gunsteren, W. F. On Developing Coarse-Grained Models for Biomolecular Simulation: a Review. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12423–12430.
- (6) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111*, 7812–7824.
- (7) Riniker, S.; van Gunsteren, W. F. A Simple, Efficient Polarisable Coarse-Grained Water Model for Molecular Dynamics Simulations. *J. Chem. Phys.* **2011**, *134*, No. 084110; *J. Chem. Phys.* **2017**, *146*, No. 129901. Erratum:
- (8) van Gunsteren, W. F.; Berendsen, H. J. C. Algorithms for Macromolecular Dynamics and Constraint Dynamics. *Mol. Phys.* **1977**, *34*, 1311–1327.
- (9) Levitt, M.; Lifson, S. Refinement of Protein Conformations Using a Macromolecular Energy Minimization Procedure. *J. Mol. Biol.* **1969**, *46*, 269–279.
- (10) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; vdf Hochschulverlag AG an der ETH Zürich and Biomas b.v.: Zürich, Switzerland, Groningen, The Netherlands, 1996.
- (11) Pechlaner, M.; Dorta, A. P.; Lin, Z.; Rusu, V. H.; van Gunsteren, W. F. A Method to Apply Bond-Angle Constraints in Molecular Dynamics Simulation. *J. Comput. Chem.* **2021**, *42*, 418–434.
- (12) Pechlaner, M.; van Gunsteren, W. F. Algorithms to Apply Dihedral-Angle Constraints in Molecular or Stochastic Dynamics Simulations. *J. Chem. Phys.* **2020**, *152*, No. 024109.
- (13) Pechlaner, M.; van Gunsteren, W. F. On the Use of Intra-Molecular Distance and Angle Constraints to Lengthen the Time Step in Molecular and Stochastic Dynamics Simulations of Proteins. *Proteins: Struct. Funct. Bioinf.* **2022**, *90*, 543–559.
- (14) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. 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.
- (15) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
- (16) Pechlaner, M.; van Gunsteren, W. F.; Hansen, N.; Smith, L. J. Molecular Dynamics Simulation or Structure Refinement of Proteins: Are Solvent Molecules Required? A Case Study Using Hen Lysozyme. *Eur. Biophys. J.* **2022**, *51*, 265–282.
- (17) Müller, M.; Katsov, K.; Schick, M. Biological and Synthetic Membranes: What Can Be Learned from a Coarse-Grained Description? *Phys. Rep.* **2006**, *434*, 113–176.
- (18) Souza, P. C. T.; Alessandri, R.; Barnoud, J.; Thallmair, S.; Faustino, I.; Grünewald, F.; Patmanidis, I.; Abdizadeh, H.; Bruininks, B. M. H.; Wassenaar, T. A.; Kroon, P. C.; Melcr, J.; Nieto, V.; Corradi, V.; Khan, H. M.; Domański, J.; Javanainen, M.; Martinez-Seara, H.; Reuter, N.; Best, R. B.; Vattulainen, I.; Monticelli, L.; Periole, X.; Tieleman, D. P.; de Vries, A. H.; Marrink, S. J. Martini 3: A General Purpose Force Field for Coarse-Grained Molecular Dynamics. *Nat. Methods* **2021**, *18*, 382–388.
- (19) Togashi, Y.; Flechsig, H. Coarse-Grained Protein Dynamics Studies Using Elastic Network Models. *Int. J. Mol. Sci.* **2018**, *19*, 3899–3916.
- (20) Riniker, S. Fixed-Charge Atomistic Force Fields for Molecular Dynamics Simulations in the Condensed Phase: An Overview. *J. Chem. Inf. Model.* **2018**, *58*, 565–578.
- (21) Blank, T. B.; Brown, S. D.; Calhoun, A. W.; Doren, D. J. Neural network models of potential energy surfaces. *J. Chem. Phys.* **1995**, *103*, 4129–4137.
- (22) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98*, No. 146401.
- (23) Allinger, N. L. Conformational Analysis. 130. MM2. A Hydrocarbon Force Field Utilizing V1 and V2 Torsional Terms. *J. Am. Chem. Soc.* **1977**, *99*, 8127–8134.
- (24) MacKerell, A. D.; Wiorkiewicz-Kuczera, J.; Karplus, M. An All-Atom Empirical Energy Function for the Simulation of Nucleic Acids. *J. Am. Chem. Soc.* **1995**, *117*, 11946–11975.
- (25) Hockney, R. W.; Eastwood, J. W. *Computer Simulation Using Particles*; McGraw-Hill: New York, USA, 1981.
- (26) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N-log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (27) Reif, M. M.; Oostenbrink, C. Toward the Correction of Effective Electrostatic Forces in Explicit-Solvent Molecular Dynamics Simulations: Restraints on Solvent-Generated Electrostatic Potential and Solvent Polarization. *Theor. Chem. Acc.* **2015**, *134*, 2–21.
- (28) Hünenberger, P. H.; McCammon, J. A. Ewald Artifacts in Computer Simulations of Ionic Solvation and Ion-Ion Interaction: A Continuum Electrostatics Study. *J. Chem. Phys.* **1999**, *110*, 1856–1872.
- (29) Yeh, I. C.; Berkowitz, M. L. Ewald Summation for Systems with Slab Geometry. *J. Chem. Phys.* **1999**, *111*, 3155–3162.
- (30) Galicia-Andrés, E.; Petrov, D.; Gerzabek, M. H.; Oostenbrink, C.; Tunega, D. On the Adsorption Mechanism of Humic Substances on Kaolinite and Their Microscopic Structure. *Langmuir* **2019**, *35*, 15086–15099.
- (31) Luty, B. A.; van Gunsteren, W. F. Calculating Electrostatic Interactions Using the Particle-Particle Particle-Mesh Method with Nonperiodic Long-Range Interactions. *J. Phys. Chem.* **1996**, *100*, 2581–2587.
- (32) Barker, J. A.; Watts, R. O. Monte Carlo Studies of the Dielectric Properties of Water-Like Models. *Mol. Phys.* **1973**, *26*, 789–792.
- (33) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. A Generalized Reaction-Field Method for Molecular Dynamics Simulations. *J. Chem. Phys.* **1995**, *102*, 5451–5459.
- (34) van Gunsteren, W. F.; Berendsen, H. J. C.; Rullmann, J. A. C. Inclusion of Reaction Fields in Molecular Dynamics: Application to Liquid Water. *Faraday Disc. Chem. Soc.* **1978**, *66*, 58–70.
- (35) Tironi, I. G.; Luty, B. A.; van Gunsteren, W. F. Space-Time Correlated Reaction Field: A Stochastic Dynamical Approach to the Dielectric Continuum. *J. Chem. Phys.* **1997**, *106*, 6068–6075.
- (36) Sidler, D.; Frasch, S.; Cristòfol-Clough, M.; Riniker, S. Anisotropic Reaction-Field Correction for Long-Range Electrostatic Interactions in Molecular Dynamics Simulations. *J. Chem. Phys.* **2018**, *148*, No. 234105.
- (37) Yu, H.; van Gunsteren, W. F. Accounting for Polarization in Molecular Simulation. *Comput. Phys. Commun.* **2005**, *172*, 69–85.

- (38) Lemkul, J. A.; Huang, J.; Roux, B.; MacKerell, A. D., Jr An Empirical Polarizable Force Field Based on the Classical Drude Oscillator Model: Development History and Recent Applications. *Chem. Rev.* **2016**, *116*, 4983–5013.
- (39) Drude, P. *The Theory of Optics*; Longmans, Green, and Co.: New York, USA, 1902.
- (40) Born, M.; Huang, K. *Dynamic Theory of Crystal Lattices*; Oxford University Press: Oxford, UK, 1954.
- (41) Straatsma, T. P.; McCammon, J. A. Molecular Dynamics Simulations with Interaction Potentials Including Polarization Development of a Noniterative Method and Application to Water. *Mol. Simulation* **1990**, *5*, 181–192.
- (42) Bachmann, S. J.; van Gunsteren, W. F. On the Compatibility of Polarisable and Non-Polarisable Models for Liquid Water. *Mol. Phys.* **2014**, *112*, 2761–2780.
- (43) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. Interaction Models for Water in Relation to Protein Hydration. In *Intermolecular Forces*; Pullman, B., ed.; Reidel: Dordrecht, The Netherlands, 1981; pp 331–342.
- (44) van Gunsteren, W. F.; Dolenc, J.; Mark, A. E. Molecular Simulation as an Aid to Experimentalists. *Curr. Opin. Struct. Biology* **2008**, *18*, 149–153.
- (45) Ouyang, J. F.; Bettens, R. P. A. Modelling Water: A Lifetime Enigma. *Chimia* **2015**, *69*, 104–111.
- (46) Rahman, A.; Stillinger, F. H. Molecular Dynamics Study of Liquid Water. *J. Chem. Phys.* **1971**, *55*, 3336–3359.
- (47) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (48) Weiner, P. K.; Kollman, P. A. AMBER – Assisted Model-Building with Energy Refinement – A General Program for Modeling Molecules and their Interactions. *J. Comput. Chem.* **1981**, *2*, 287–303.
- (49) van Gunsteren, W. F.; Karplus, M. Effect of Constraints on the Dynamics of Macromolecules. *Macromolecules* **1982**, *15*, 1528–1544.
- (50) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM – A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem.* **1983**, *4*, 187–217.
- (51) Daura, X.; Mark, A. E.; van Gunsteren, W. F. Parametrization of Aliphatic CH_n United Atoms of GROMOS96 Force Field. *J. Comput. Chem.* **1998**, *19*, 535–547.
- (52) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: the GROMOS Force-Field Parameter Sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656–1676.
- (53) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. Peptide Folding: When Simulation Meets Experiment. *Angew. Chemie Int. Ed.* **1999**, *38*, 236–240; *Angew. Chem.* **1999**, *111*, 249–253.
- (54) Patel, S.; Brooks, C. L., III CHARMM Fluctuating Charge Force Field for Proteins: I Parameterization and Application to Bulk Organic Liquid Simulations. *J. Comput. Chem.* **2004**, *25*, 1–15.
- (55) Anisimov, V. M.; Lamoureux, G.; Vorobyov, I. V.; Huang, N.; Roux, B.; MacKerell, A. D., Jr. Determination of Electrostatic Parameters for a Polarizable Force Field Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* **2005**, *1*, 153–168.
- (56) Ponder, J. W.; Wu, C. J.; Ren, P. Y.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **2010**, *114*, 2549–2564.
- (57) Stroet, M.; Kozlara, K. B.; Malde, A. K.; Mark, A. E. Optimization of Empirical Force Fields by Parameter Space Mapping: A Single-Step Perturbation Approach. *J. Chem. Theory Comput.* **2017**, *13*, 6201–6212.
- (58) Smith, J. S.; Isayev, O.; Roitberg, A. E. ANI-1: An Extensible Neural Network Potential with DFT Accuracy at Force Field Computational Cost. *Chem. Sci.* **2017**, *8*, 3192–3203.
- (59) Schütt, K. T.; Sauceda, H. E.; Kindermans, P. J.; Tkatchenko, A.; Müller, K. R. SchNet – A Deep Learning Architecture for Molecules and Materials. *J. Chem. Phys.* **2018**, *148*, No. 241722.
- (60) Gear, C. W. *Numerical Initial Value Problems in Ordinary Differential Equations*; Prentice-Hall: Upper Saddle River (NJ), USA, 1971.
- (61) Verlet, L. Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.* **1967**, *159*, 98–10.
- (62) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes: The Art of Scientific Computing* (3rd ed.); Cambridge University Press: New York, USA, 2007.
- (63) Cuendet, M. A.; van Gunsteren, W. F. On the Calculation of Velocity-Dependent Properties in Molecular Dynamics Simulations Using the Leap-Frog Integration Algorithm. *J. Chem. Phys.* **2007**, *127*, No. 184102.
- (64) Stocker, U.; Juchli, D.; van Gunsteren, W. F. 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.
- (65) van Gunsteren, W. F.; Berendsen, H. J. C. A Leap-Frog Algorithm for Stochastic Dynamics. *Mol. Simul.* **1988**, *1*, 173–185.
- (66) Berkowitz, M.; McCammon, J. A. Molecular Dynamics with Stochastic Boundary Conditions. *Chem. Phys. Lett.* **1982**, *90*, 215–217.
- (67) Shi, Y. Y.; Lu, W.; van Gunsteren, W. F. On the Approximation of Solvent Effects on the Conformation and Dynamics of Cyclosporin A by Stochastic Dynamics Simulation Techniques. *Mol. Simul.* **1988**, *1*, 369–388.
- (68) Wittenburg, J. *Dynamics of Systems of Rigid Bodies*; Teubner: Stuttgart, Germany, 1977.
- (69) Katz, H.; Roderich Walter, T.; Somorjai, R. L. Rotational Dynamics of Large Molecules. *Computers & Chemistry* **1979**, *3*, 25–32.
- (70) Bae, D. S.; Haug, E. J. A Recursive Formulation for Constrained Mechanical System Dynamics: Part I. Open Loop Systems. *J. Struct. Mech.* **1987**, *15*, 359–382.
- (71) Bae, D. S.; Haug, E. J. A Recursive Formulation for Constrained Mechanical System Dynamics: Part II. Closed Loop Systems. *J. Struct. Mech.* **1987**, *15*, 481–506.
- (72) Mazur, A. K.; Dorofeev, V. E.; Abagyan, R. A. Derivation and Testing of Explicit Equations of Motion for Polymers Described by Internal Coordinates. *J. Comput. Phys.* **1991**, *92*, 261–272.
- (73) Jain, A.; Vaidehi, N.; Rodriguez, G. A Fast Recursive Algorithm for Molecular Dynamics Simulation. *J. Comput. Phys.* **1993**, *106*, 258–268.
- (74) Turner, J. D.; Weiner, P. K.; Chun, H. M.; Lupi, V.; Gallion, S.; Singh In *Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications*, Vol 2; Springer: The Netherlands, 1993; pp 535–555.
- (75) Mathiowetz, A. M.; Jain, A.; Karasawa, N.; Goddard, W. A., 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.
- (76) Finney, J. L. Long-Range Forces in Molecular Dynamics Calculations on Water. *J. Comput. Phys.* **1978**, *28*, 92–102.
- (77) Streett, W. B.; Tildesley, D. J.; Saville, G. Multiple Time-Step Methods in Molecular Dynamics. *Mol. Phys.* **1978**, *35*, 639–648.
- (78) Berendsen, H. J. C.; van Gunsteren, W. F.; Zwinderman, H. R. J.; Geurtsen, R. G. Simulations of Proteins in Water. *Ann. N.Y. Acad. Sci.* **1986**, *482*, 269–285.
- (79) van Gunsteren, W. F.; Berendsen, H. J. C.; Geurtsen, R. G.; Zwinderman, H. R. J. 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.
- (80) Teleman, O.; Jönsson, B. Vectorizing a General Purpose Molecular Dynamics Simulation Program. *J. Comput. Chem.* **1986**, *7*, 58–66.

- (81) Kräutler, V.; Hünenberger, P. H. A Multiple-Timestep Algorithm Compatible with a Large Number of Distance Classes and an Arbitrary Distance Dependence of the Timestep Size for the Fast Evaluation of Non-Bonded Interactions in Molecular Simulations. *J. Comput. Chem.* **2006**, *27*, 1163–1176.
- (82) Pechlaner, M.; Oostenbrink, C.; van Gunsteren, W. F. 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.
- (83) Diem, M.; Oostenbrink, C. The Effect of Using a Twin-Range Cutoff Scheme for Nonbonded Interactions: Implications for Force-Field Parametrization? *J. Chem. Theory Comput.* **2020**, *16*, 5985–5990.
- (84) Diem, M.; Oostenbrink, C. The Effect of Different Cutoff Schemes in Molecular Simulations of Proteins. *J. Comput. Chem.* **2020**, *41*, 2740–2749.
- (85) van Gunsteren, W. F.; Daura, X.; Fuchs, P. F. J.; Hansen, N.; Horta, B. A. C.; Hünenberger, P. H.; Mark, A. E.; Pechlaner, M.; Riniker, S.; Oostenbrink, C. On the Effect of the Various Assumptions and Approximations used in Molecular Simulation on the Properties of Bio-Molecular Systems: Overview and Perspective on Issues. *ChemPhysChem* **2021**, *22*, 264–282.
- (86) Fixman, M. Classical Statistical Mechanics of Constraints: A Theorem and Application to Polymers. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 3050–3053.
- (87) Fixman, M. Simulation of Polymer Dynamics. I. General Theory. *J. Chem. Phys.* **1978**, *69*, 1527–1537.
- (88) van Gunsteren, W. F. Constrained Dynamics of Flexible Molecules. *Mol. Phys.* **1980**, *40*, 1015–1019.
- (89) Tironi, I. G.; Brunne, R. M.; van Gunsteren, W. F. On the Relative Merits of Flexible versus Rigid Models for Use in Computer Simulations of Molecular Liquids. *Chem. Phys. Lett.* **1996**, *250*, 19–24.
- (90) Feenstra, K. A.; Hess, B.; Berendsen, H. J. C. Improving Efficiency of Large Time-Scale Molecular Dynamics Simulations of Hydrogen-Rich Systems. *J. Comput. Chem.* **1999**, *20*, 786–798.
- (91) Andersen, H. C. Rattle: A “Velocity” Version of the Shake Algorithm for Molecular Dynamics Calculations. *J. Comput. Phys.* **1983**, *52*, 24–34.
- (92) Barth, E.; Kuczera, K.; Leimkuhler, B.; Skeel, R. D. Algorithms for Constrained Molecular Dynamics. *J. Comput. Chem.* **1995**, *16*, 1192–1209.
- (93) Tao, P.; Wu, X.; Brooks, B. R. Maintain Rigid Structures in Verlet Based Cartesian Molecular Dynamics Simulations. *J. Chem. Phys.* **2012**, *137*, No. 134110.
- (94) Kräutler, V.; van Gunsteren, W. F.; Hünenberger, P. H. A Fast SHAKE Algorithm to Solve Distance Constraint Equations for Small Molecules in Molecular Dynamics Simulations. *J. Comput. Chem.* **2001**, *22*, 501–508.
- (95) Miyamoto, S.; Kollman, P. A. SETTLE: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.* **1992**, *13*, 952–962.
- (96) Gonnet, P.; Walther, J. H.; Koumoutsakos, P. θ -SHAKE: an Extension to SHAKE for the Explicit Treatment of Angular Constraints. *Comput. Phys. Commun.* **2009**, *180*, 360–364.
- (97) Dubbeldam, D.; Oxford, G. A. E.; Krishna, R.; Broadbelt, L. J.; Snurr, R. Q. Distance and Angular Holonomic Constraints in Molecular Simulations. *J. Chem. Phys.* **2010**, *133*, No. 034114.
- (98) Christen, M.; Kunz, A. P. E.; van Gunsteren, W. F. Sampling of Rare Events Using Hidden Restraints. *J. Phys. Chem. B* **2006**, *110*, 8488–8498; *J. Phys. Chem. B* **2008**, *112*, 11446. Erratum:
- (99) Zhou, J.; Reich, S.; Brooks, B. R. Elastic Molecular Dynamics with Self-Consistent Flexible Constraints. *J. Chem. Phys.* **2000**, *112*, 7919–7929.
- (100) Hess, B.; Saint-Martin, H.; Berendsen, H. J. C. Flexible Constraints: An Adiabatic Treatment of Quantum Degrees of Freedom, with Application to the Flexible and Polarizable Mobile Charge Densities in Harmonic Oscillators Model for Water. *J. Chem. Phys.* **2002**, *116*, 9602–9610.
- (101) Christen, M.; van Gunsteren, W. F. An Approximate but Fast Method to Impose Flexible Distance Constraints in Molecular Dynamics Simulations. *J. Chem. Phys.* **2005**, *122*, No. 144106.
- (102) Christen, M.; Christ, C.; van Gunsteren, W. F. Free Energy Calculations Using Flexible-Constrained, Hard-Constrained and Non-Constrained MD Simulations. *ChemPhysChem* **2007**, *8*, 1557–1564.
- (103) Hünenberger, P. H. Thermostat Algorithms for Molecular Dynamics Simulations. *Adv. Polym. Sci.* **2005**, *173*, 105–149.
- (104) Andersen, H. C. Molecular Dynamics Simulations at Constant Pressure and/or Temperature. *J. Chem. Phys.* **1980**, *72*, 2384–2393.
- (105) Schneider, T.; Stoll, E. Molecular-Dynamics Study of a Three-Dimensional One-Component Model for Distortive Phase Transitions. *Phys. Rev. B* **1978**, *17*, 1302–1322.
- (106) Cerutti, D. S.; Duke, R.; Freddolino, P. L.; Fan, H.; Lybrand, T. P. Vulnerability in Popular Molecular Dynamics Packages Concerning Langevin and Andersen Dynamics. *J. Chem. Theory Comput.* **2008**, *4*, 1669–1680.
- (107) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (108) Lin, Z.; van Gunsteren, W. F. On the Use of a Weak-Coupling Thermostat in Replica-Exchange Molecular Dynamics Simulations. *J. Chem. Phys.* **2015**, *143*, No. 034110.
- (109) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling Through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, No. 014101.
- (110) Nosé, S. A Molecular Dynamics Method for Simulations in the Canonical Ensemble. *Mol. Phys.* **1984**, *52*, 255–268.
- (111) Hendrickson, W. A.; Konnert, J. H. Stereochemically Restricted Crystallographic Least-Squares Refinement of Macromolecular Structures. In *Biomolecular Structure, Conformation, Function and Evolution*, Vol. 1; Srinivasan, R., ed.; Pergamon: Oxford, 1981; pp 43–57.
- (112) van Gunsteren, W. F.; Kaptein, R.; Zuiderweg, E. R. P. Use of Molecular Dynamics Computer Simulations When Determining Protein Structure by 2D NMR. In *Proceedings NATO/CECAM Workshop on Nucleic Acid Conformation and Dynamics*; Olson, W. K., ed.; CECAM: Orsay, 1984; pp 79–92.
- (113) Kaptein, R.; Zuiderweg, E. R. P.; Scheek, R. M.; Boelens, R.; van Gunsteren, W. F. A Protein Structure from Nuclear Magnetic Resonance Data lac Repressor Headpiece. *J. Mol. Biol.* **1985**, *182*, 179–182.
- (114) McCammon, J. A.; Gelin, B. R.; Karplus, M. Dynamics of Folded Proteins. *Nature (London)* **1977**, *267*, 585–590.
- (115) van Gunsteren, W. F.; Berendsen, H. J. C.; Hermans, J.; Hol, W. G. J.; Postma, J. P. M. Computer Simulation of the Dynamics of Hydrated Protein Crystals and its Comparison with X-Ray Data. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 4315–4319.
- (116) Torda, A. E.; Scheek, R. M.; van Gunsteren, W. F. Time-Dependent Distance Restraints in Molecular Dynamics Simulations. *Chem. Phys. Lett.* **1989**, *157*, 289–294.
- (117) Fennen, J.; Torda, A. E.; van Gunsteren, W. F. Structure Refinement with Molecular Dynamics and a Boltzmann-Weighted Ensemble. *J. Biomol. NMR* **1995**, *6*, 163–170.
- (118) Gros, P.; van Gunsteren, W. F.; Hol, W. G. J. Inclusion of Thermal Motion in Crystallographic Structures by Restrained Molecular Dynamics. *Science* **1990**, *249*, 1149–1152.
- (119) Schiffer, C. A.; van Gunsteren, W. F. Accessibility and Order of Water Sites in and Around Proteins: A Crystallographic Time-Averaging Study. *Proteins: Struct. Funct. Genet.* **1999**, *36*, 501–511.
- (120) Torda, A. E.; Scheek, R. M.; van Gunsteren, W. F. Time-averaged Nuclear Overhauser Effect Distance Restraints Applied to Tendamistat. *J. Mol. Biol.* **1990**, *214*, 223–235.
- (121) Schiffer, C. A.; Huber, R.; Wüthrich, K.; van Gunsteren, W. F. Simultaneous Refinement of the Structure of BPTI Against NMR Data Measured in Solution and X-ray Diffraction Data Measured in Single Crystals. *J. Mol. Biol.* **1994**, *241*, 588–599.
- (122) Torda, A. E.; Brunne, R. M.; Huber, T.; Kessler, H.; van Gunsteren, W. F. Structure Refinement Using Time-Averaged J-Coupling Constant Restraints. *J. Biomol. NMR* **1993**, *3*, 55–66.

- (123) Smith, L. J.; van Gunsteren, W. F.; Hansen, N. On the Use of Time-Averaging Restraints when Deriving Biomolecular Structure from J_3 -coupling Values Obtained from NMR Experiments. *J. Biomol. NMR* **2016**, *66*, 69–83.
- (124) Huber, T.; Torda, A. E.; van Gunsteren, W. F. Local Elevation: A Method for Improving the Searching Properties of Molecular Dynamics Simulation. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 695–708.
- (125) Hansen, N.; Heller, F.; Schmid, N.; van Gunsteren, W. F. Time-Averaged Order Parameter Restraints in Molecular Dynamics Simulations. *J. Biomol. NMR* **2014**, *60*, 169–187.
- (126) van Gunsteren, W. F.; Pechlaner, M.; Smith, L. J.; Stankiewicz, B.; Hansen, N. A Method to Derive Structural Information on Molecules from Residual Dipolar Coupling NMR Data. *J. Phys. Chem. B* **2022**, *126*, 3867–3888.
- (127) van Gunsteren, W. F.; Allison, J. R.; Daura, X.; Dolenc, J.; Hansen, N.; Mark, A. E.; Oostenbrink, C.; Rusu, V. H.; Smith, L. J. Deriving Structural Information from Experimentally Measured Data on Biomolecules. *Angew. Chem., Int. Ed.* **2016**, *55*, 15990–16010; *Angew. Chem.* **2016**, *128*, 16222–16244.
- (128) Christen, M.; van Gunsteren, W. F. On Searching in, Sampling of, and Dynamically Moving Through Conformational Space of Biomolecular Systems: A Review. *J. Comput. Chem.* **2008**, *29*, 157–166.
- (129) Cotterill, R. M. J.; Madsen, J. K. Potential Energy Contour Tracing: An Efficient Way of Exploring Configuration Hyperspace. In *Characterising Complex Systems*; Bohr, H., ed.; World Scientific, Singapore, 1990; pp 177–191.
- (130) Byrne, D.; Li, J.; Platt, E.; Robson, B.; Weiner, P. Novel Algorithms for Searching Conformational Space. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 67–82.
- (131) van Schaik, R. C.; van Gunsteren, W. F.; Berendsen, H. J. C. Conformational Search by Potential Energy Annealing: Algorithm and Application to Cyclosporin A. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 97–112.
- (132) Kokh, D. B.; Doser, B.; Richter, S.; Ormersbach, F.; Cheng, X.; Wade, R. C. A workflow for exploring ligand dissociation from a macromolecule: Efficient random acceleration molecular dynamics simulation and interaction fingerprint analysis of ligand trajectories. *J. Chem. Phys.* **2020**, *153*, No. 125102.
- (133) Levitt, M. Protein Folding by Restrained Energy Minimization and Molecular Dynamics. *J. Mol. Biol.* **1983**, *170*, 723–764.
- (134) Beutler, T. C.; Mark, A. E.; van Schaik, R.; Gerber, P. R.; van Gunsteren, W. F. Avoiding Singularities and Numerical Instabilities in Free Energy Calculations Based on Molecular Simulations. *Chem. Phys. Lett.* **1994**, *222*, 529–539.
- (135) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. Separation-Shifted Scaling, a New Scaling Method for Lennard-Jones Interactions in Thermodynamic Integration. *J. Chem. Phys.* **1994**, *100*, 9025–9031.
- (136) Hornak, V.; Simmerling, C. Development of Softcore Potential Functions for Overcoming Steric Barriers in Molecular Dynamics Simulations. *J. Mol. Graph. Model.* **2004**, *22*, 405–413.
- (137) Huber, T.; Torda, A. E.; van Gunsteren, W. F. Structure Optimization Combining Soft-Core Interaction Functions, the Diffusion Equation Method, and Molecular Dynamics. *J. Phys. Chem. A* **1997**, *101*, 5926–5930.
- (138) Miao, Y.; Feher, V. A.; McCammon, J. A. Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation. *J. Chem. Theory Comput.* **2015**, *11*, 3584–3595.
- (139) Gracia Carmona, O.; Oostenbrink, C. Flexible Gaussian Accelerated Molecular Dynamics to Enhance Biological Sampling. *J. Chem. Theory Comput.* **2023**, *19*, 6521–6531.
- (140) Glover, F. Tabu Search Part I. *ORSA J. Computing* **1989**, *1*, 190–206.
- (141) Glover, F. Tabu Search Part II. *ORSA J. Computing* **1990**, *2*, 4–32.
- (142) Crippen, G. M.; Scheraga, H. A. Minimization of Polypeptide Energy, VIII. Application of the Deflation Technique to a Dipeptide. *Proc. Natl. Acad. Sci. U.S.A.* **1969**, *64*, 42–49.
- (143) Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562–12566.
- (144) Grubmüller, H. Predicting Slow Structural Transitions in Macromolecular Systems: Conformational Flooding. *Phys. Rev. E* **1995**, *52*, 2893–2906.
- (145) Rahman, J. A.; Tully, J. C. Puddle-Jumping: A Flexible Sampling Algorithm for Rare Event Systems. *Chem. Phys.* **2002**, *285*, 277–287.
- (146) Hamelberg, D.; Mongan, J.; McCammon, J. A. Accelerated Molecular Dynamics: A Promising and Efficient Simulation Method for Biomolecules. *J. Chem. Phys.* **2004**, *120*, 11919–11929.
- (147) van Gunsteren, W. F.; Brunne, R. M.; Gros, P.; van Schaik, R. C.; Schiffer, C. A.; Torda, A. E. Accounting for Molecular Mobility in Structure Determination Based on Nuclear Magnetic Resonance Spectroscopic and X-Ray Diffraction Data. In *Methods in Enzymology: Nuclear Magnetic Resonance*, Vol. 239; James, T.L., Oppenheimer, N.J., eds.; Academic Press: New York, USA, 1994; pp 619–654.
- (148) Daura, X.; Antes, I.; van Gunsteren, W. F.; Thiel, W.; Mark, A. E. The Effect of Motional Averaging on the Calculation of NMR-Derived Structural Properties. *Proteins: Struct. Funct. Genet.* **1999**, *36*, 542–555.
- (149) van Schaik, R. C.; Berendsen, H. J. C.; Torda, A. E.; van Gunsteren, W. F. A Structure Refinement Method Based on Molecular Dynamics in Four Spatial Dimensions. *J. Mol. Biol.* **1993**, *234*, 751–762.
- (150) Beutler, T. C.; van Gunsteren, W. F. Molecular Dynamics Free Energy Calculation in Four Dimensions. *J. Chem. Phys.* **1994**, *101*, 1417–1422.
- (151) Elber, R.; Karplus, M. Enhanced Sampling in Molecular Dynamics: Use of the Time-Dependent Hartree Approximation for a Simulation of Carbon Monoxide Diffusion Through Myoglobin. *J. Am. Chem. Soc.* **1990**, *112*, 9161–9175.
- (152) Straub, J. E.; Karplus, M. Energy Equipartitioning in the Classical Time-Dependent Hartree Approximation. *J. Chem. Phys.* **1991**, *94*, 6737–6739.
- (153) Zheng, Q. A.; Rosenfeld, R.; Vajda, S.; DeLisi, C. Determining Protein Loop Conformation Using Scaling-Relaxation Techniques. *Protein Sci.* **1993**, *2*, 1242–1248.
- (154) Olszewski, K. A.; Piela, L.; Scheraga, H. A. Mean Field Theory as a Tool for Intramolecular Conformational Optimization. 1. Tests on Terminally-Blocked Alanine and Met-Enkephalin. *J. Phys. Chem.* **1992**, *96*, 4672–4676.
- (155) Simmerling, C.; Miller, J. L.; Kollman, P. A. Combined Locally Enhanced Sampling and Particle Mesh Ewald as a Strategy to Locate the Experimental Structure of a Nonhelical Nucleic Acid. *J. Am. Chem. Soc.* **1998**, *120*, 7149–7155.
- (156) Liu, H.; Duan, Z.; Luo, Q.; Shi, Y. Structure-Based Ligand Design by Dynamically Assembling Molecular Building Blocks at Binding Site. *Prot. Struct. Funct. Bioinf.* **1999**, *36*, 462–470.
- (157) Kunz, A. P. E.; Liu, H.; van Gunsteren, W. F. Enhanced Sampling of Particular Degrees of Freedom in Molecular Systems Based on Adiabatic Decoupling and Temperature or Force Scaling. *J. Chem. Phys.* **2011**, *135*, No. 104106.
- (158) Kunz, A. P. E.; van Gunsteren, W. F. Enhancing the Configurational Sampling of Ions in Aqueous Solution Using Adiabatic Decoupling with Translational Temperature Scaling. *J. Phys. Chem. B* **2011**, *115*, 2931–2936.
- (159) Kunz, A. P. E.; Lin, Z.; van Gunsteren, W. F. Test of a Method for Sampling the Internal Degrees of Freedom of a Flexible Solute Molecule Based on Adiabatic Decoupling and Temperature or Force Scaling. *Mol. Phys.* **2012**, *110*, 407–417.
- (160) Liu, P.; Kim, B.; Friesner, R. A.; Berne, B. J. Replica Exchange with Solute Tempering: A Method for Sampling Biological Systems in Explicit Water. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13749–13754.
- (161) Sugita, Y.; Okamoto, Y. Replica-Exchange Molecular Dynamics Method for Protein Folding. *Chem. Phys. Lett.* **1999**, *314*, 141–151.
- (162) Okamoto, Y. Generalized-Ensemble Algorithms: Enhanced Sampling Techniques for Monte Carlo and Molecular Dynamics Simulations. *J. Mol. Graph. Model.* **2004**, *22*, 425–439.
- (163) Sugita, Y.; Kitao, A.; Okamoto, Y. Multidimensional Replica-Exchange Method for Free-Energy Calculations. *J. Chem. Phys.* **2000**, *113*, 6042–6051.

- (164) Woods, C. J.; Essex, J. W.; King, M. A. The Development of Replica-Exchange-Based Free-Energy Methods. *J. Phys. Chem. B* **2003**, *107*, 13703–13710.
- (165) Zhou, R. H.; Berne, B. J.; Germain, R. The Free Energy Landscape for β Hairpin Folding in Explicit Water. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14931–14936.
- (166) Mitsutake, A.; Sugita, Y.; Okamoto, Y. Generalized-Ensemble Algorithms for Molecular Simulations of Biopolymers. *Biopolymers (Peptide Science)* **2001**, *60*, 96–123.
- (167) Cheng, X. L.; Cui, G. L.; Hornak, V.; Simmerling, C. Modified Replica Exchange Simulation Methods for Local Structure Refinement. *J. Phys. Chem. B* **2005**, *109*, 8220–8230.
- (168) Voter, A. F. Parallel Replica Method for Dynamics of Infrequent Events. *Phys. Rev. B* **1998**, *57*, No. R13985.
- (169) Calvo, F.; Doye, J. P. K. Entropic Tempering: A method for Overcoming Quasi Ergodicity in Simulation. *Phys. Rev. E* **2000**, *63*, No. 010902.
- (170) Trebst, S.; Huse, D. A.; Troyer, M. Optimizing the Ensemble for Equilibration in Broad-Histogram Monte Carlo Simulations. *Phys. Rev. E* **2004**, *70*, No. 046701.
- (171) Trebst, S.; Gull, E.; Troyer, M. Optimized Ensemble Monte Carlo Simulations of Dense Lennard-Jones Fluids. *J. Chem. Phys.* **2005**, *123*, No. 204501.
- (172) Huber, T.; van Gunsteren, W. F. SWARM-MD: Searching Conformational Space by Cooperative Molecular Dynamics. *J. Phys. Chem. A* **1998**, *102*, 5937–5943.
- (173) Hixson, C. A.; Wheeler, R. A. Rigorous Classical-Mechanical Derivation of a Multiple-Copy Algorithm for Sampling Statistical Mechanical Ensembles. *Phys. Rev. E* **2001**, *64*, No. 026701.
- (174) Huber, G. A.; McCammon, J. A. Weighted-Ensemble Simulated Annealing: Faster Optimization on Hierarchical Energy Surfaces. *Phys. Rev. E* **1997**, *55*, 4822.
- (175) Shirts, M. R.; Pande, V. S. Mathematical Analysis of Coupled Parallel Simulations. *Phys. Rev. Lett.* **2001**, *86*, 4983–4987.
- (176) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23*, 187–199.
- (177) de Ruiter, A.; Oostenbrink, C. Protein–Ligand Binding from Distancefield Distances and Hamiltonian Replica Exchange Simulations. *J. Chem. Theory Comput.* **2013**, *9*, 883–892.
- (178) Dellago, C.; Bolhuis, P. G.; Csaik, F. S.; Chandler, D. Transition Path Sampling and the Calculation of Rate Constants. *J. Chem. Phys.* **1998**, *108*, 1964–1977.
- (179) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. Transition Path Sampling: Throwing Ropes Over Rough Mountain Passes. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291–318.
- (180) Dellago, C.; Bolhuis, P. G.; Geissler, P. L. Transition Path Sampling. *Adv. Chem. Phys.* **2002**, *123*, 1–78.
- (181) Christ, C. D.; van Gunsteren, W. F. Enveloping Distribution Sampling: A Method to Calculate Free Energy Differences from a Single Simulation. *J. Chem. Phys.* **2007**, *126*, No. 184110.
- (182) Christ, C. D.; van Gunsteren, W. F. Multiple Free Energies from a Single Simulation: Extending Enveloping Distribution Sampling to Non-Overlapping Phase-Space Distributions. *J. Chem. Phys.* **2008**, *128*, No. 174112; *J. Chem. Phys.* **2011**, *134*, No. 229901. Erratum:
- (183) Christ, C. D.; van Gunsteren, W. F. Simple, Efficient, and Reliable Computation of Multiple Free Energy Differences from a Single Simulation: A Reference Hamiltonian Parameter Update Scheme for Enveloping Distribution Sampling (EDS). *J. Chem. Theory Comput.* **2009**, *5*, 276–286.
- (184) Christ, C. D.; van Gunsteren, W. F. Comparison of Three Enveloping Distribution Sampling Hamiltonians for the Estimation of Multiple Free Energy Differences from a Single Simulation. *J. Comput. Chem.* **2009**, *30*, 1664–1679.
- (185) Perthold, J. W.; Oostenbrink, C. Accelerated Enveloping Distribution Sampling: Enabling Sampling of Multiple End States while Preserving Local Energy Minima. *J. Phys. Chem. B* **2018**, *122*, 5030–5037.
- (186) Sidler, D.; Schwaninger, A.; Riniker, S. Replica Exchange Enveloping Distribution Sampling (RE-EDS): A Robust Method to Estimate Multiple Free-Energy Differences from a Single Simulation. *J. Chem. Phys.* **2016**, *145*, No. 154114.
- (187) Hansen, N.; Hünenberger, P. H.; van Gunsteren, W. F. Efficient Combination of Environment Change and Alchemical Perturbation Within the Enveloping Distribution Sampling (EDS) Scheme: Twin System EDS and Application to the Determination of Octanol-Water Partition Coefficients. *J. Chem. Theory Comput.* **2013**, *9*, 1334–1346.
- (188) Lin, Z.; Liu, H.; Riniker, S.; van Gunsteren, W. F. On the Use of Enveloping Distribution Sampling (EDS) to Compute Free Enthalpy Differences between Different Conformational States of Molecules: Application to α , π Helices. *J. Chem. Theory Comput.* **2011**, *7*, 3884–3897.
- (189) Lin, Z.; Timmerscheidt, T. A.; van Gunsteren, W. F. Using Enveloping Distribution Sampling (EDS) to Compute the Free Enthalpy Difference between Right and Left-Handed Helices of a β -Peptide in Solution. *J. Chem. Phys.* **2012**, *137*, No. 064108.
- (190) Lin, Z.; van Gunsteren, W. F. Combination of Enveloping Distribution Sampling (EDS) of a Soft-Core Reference-State Hamiltonian with One-Step Perturbation to Predict the Effect of Side Chain Substitution on the Relative Stability of Right- and Left-Helical Folds of β -Peptides. *J. Chem. Theory Comput.* **2013**, *9*, 126–134.
- (191) Lin, Z.; Necula, C.; van Gunsteren, W. F. Using Enveloping Distribution Sampling to Compute the Folding Free Enthalpy of a β -Peptide with a very Unstable Folded Conformation in Solution: The Advantage of Focused Sampling Using EDS. *Chem. Phys.* **2014**, *428*, 156–163.
- (192) Lin, Z.; van Gunsteren, W. F. A Comparison of Pathway Independent and Pathway Dependent Methods in the Calculation of Conformational Free Enthalpy Differences. *Protein Sci.* **2016**, *25*, 184–191.
- (193) Sidler, D.; Cristófol-Clough, M.; Riniker, S. Efficient Round-Trip Time Optimization for Replica-Exchange Enveloping Distribution Sampling (RE-EDS). *J. Chem. Theory Comput.* **2017**, *13*, 3020–3030.
- (194) Gracia Carmona, O.; Oostenbrink, C. Accelerated Enveloping Distribution Sampling (AEDS) Allows for Efficient Sampling of Orthogonal Degrees of Freedom. *J. Chem. Inf. Model.* **2023**, *63*, 197–207.
- (195) Perthold, J. W.; Petrov, D.; Oostenbrink, C. Toward Automated Free Energy Calculation with Accelerated Enveloping Distribution Sampling (A-EDS). *J. Chem. Inf. Model.* **2020**, *60*, 5395–5406.
- (196) Lees, A. W.; Edwards, S. F. The Computer Study of Transport Processes Under Extreme Conditions. *J. Phys. C* **1972**, *5*, 1921–1929.
- (197) Müller-Plathe, F. Reversing the Perturbation in Nonequilibrium Molecular Dynamics: An Easy Way to Calculate the Shear Viscosity of Fluids. *Phys. Rev. E* **1999**, *59*, 4894–4898.
- (198) Riniker, S.; Kunz, A. P. E.; van Gunsteren, W. F. On the Calculation of the Dielectric Permittivity and Relaxation Time of Molecular Models in the Liquid Phase. *J. Chem. Theory Comput.* **2011**, *7*, 1469–1475.
- (199) Kunz, A. P. E.; Eichenberger, A.; van Gunsteren, W. F. A Simple, Efficient Polarisable Molecular Model for Liquid Carbon Tetrachloride. *Mol. Phys.* **2011**, *109*, 365–372.
- (200) Berendsen, H. J. C. Transport Properties Computed by Linear Response Through Weak Coupling to a Bath. In *Computer Simulation in Materials Science*; Meyer, M., Pontikis, V., eds.; Kluwer Academic Publishers, 1991; pp 139–155.
- (201) Evans, D. J.; Morriss, G. P. *Statistical Mechanics of Non-Equilibrium Liquids*; Academic: London, UK, 1990.
- (202) Mark, A. E.; van Helden, S. P.; Smith, P. E.; Janssen, L. H. M.; van Gunsteren, W. F. Convergence Properties of Free Energy Calculations: α -Cyclodextrin Complexes as a Case Study. *J. Am. Chem. Soc.* **1994**, *116*, 6293–6302.
- (203) Jarzynski, C. Nonequilibrium Equality for Free Energy Differences. *Phys. Rev. Lett.* **1997**, *78*, 2690–2693.
- (204) Jarzynski, C. Equilibrium Free-Energy Differences from Nonequilibrium Measurements: A Master-Equation Approach. *Phys. Rev. E* **1997**, *56*, 5018–5035.

- (205) Oostenbrink, C.; van Gunsteren, W. F. Calculating Zeros: Non-Equilibrium Free Energy Calculations. *Chem. Phys.* **2006**, *323*, 102–108.
- (206) Hendrix, D. A.; Jarzynski, C. A “Fast Growth” Method of Computing Free Energy Differences. *J. Chem. Phys.* **2001**, *114*, 5974–5981.
- (207) Gore, J.; Ritort, F.; Bustamante, C. Bias and Error in Estimates of Equilibrium Free-Energy Differences from Nonequilibrium Measurements. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12564–12569.
- (208) Ytreberg, F. M.; Zuckerman, D. M. Efficient Use of Nonequilibrium Measurement to Estimate Free Energy Differences for Molecular Systems. *J. Comput. Chem.* **2004**, *25*, 1749–1759.
- (209) Hummer, G. Fast-growth Thermodynamic Integration: Results for Sodium Ion Hydration. *Mol. Simul.* **2002**, *28*, 81–90.
- (210) Hu, H.; Yun, R. H.; Hermans, J. Reversibility of Free Energy Simulations: Slow Growth May Have a Unique Advantage. (With a Note on Use of Ewald Summation). *Mol. Simul.* **2002**, *28*, 67–80.
- (211) Hummer, G. Fast-Growth Thermodynamic Integration: Error and Efficiency Analysis. *J. Chem. Phys.* **2001**, *114*, 7330–7337.
- (212) Rodriguez-Gomez, D.; Darve, E.; Pohorille, A. Assessing the Efficiency of Free Energy Calculation Methods. *J. Chem. Phys.* **2004**, *120*, 3563–3578.
- (213) Braun, O.; Hanke, A.; Seifert, U. Probing Molecular Free Energy Landscapes by Periodic Loading. *Phys. Rev. Lett.* **2004**, *93*, No. 158105.
- (214) Park, S.; Schulten, K. Calculating Potentials of Mean Force from Steered Molecular Dynamics Simulations. *J. Chem. Phys.* **2004**, *120*, 5946–5961.
- (215) Crooks, G. E. Path-Ensemble Averages in Systems Driven far from Equilibrium. *Phys. Rev. E* **2000**, *61*, 2361–2366.
- (216) Goette, M.; Grubmüller, H. Accuracy and Convergence of Free Energy Differences Calculated from Nonequilibrium Switching Processes. *J. Comput. Chem.* **2009**, *30*, 447–456.
- (217) Kestemont, E.; van Craen, J. On the Computation of Correlation Functions in Molecular Dynamics Experiments. *J. Comput. Phys.* **1976**, *22*, 451–458.
- (218) Glättli, A.; Daura, X.; van Gunsteren, W. F. Derivation of an Improved Simple Point Charge Model for Liquid Water: SPC/A and SPC/L. *J. Chem. Phys.* **2002**, *116*, 9811–9828.
- (219) Bachmann, S. J.; van Gunsteren, W. F. An Improved Polarisable Water Model for Use in Biomolecular Simulation. *J. Chem. Phys.* **2014**, *141*, No. 22D515.
- (220) Fujinaga, M.; Gros, P.; van Gunsteren, W. F. Testing the Method of Crystallographic Refinement Using Molecular Dynamics. *J. Appl. Crystallogr.* **1989**, *22*, 1–8.
- (221) Peter, C.; Daura, X.; van Gunsteren, W. F. Calculation of NMR-Relaxation Parameters for Flexible Molecules from Molecular Dynamics Simulations. *J. Biomol. NMR* **2001**, *20*, 297–310.
- (222) Heinz, T. N.; van Gunsteren, W. F.; Hünenberger, P. H. Comparison of Four Methods to Compute the Dielectric Permittivity of Liquids from Molecular Dynamics Simulations. *J. Chem. Phys.* **2001**, *115*, 1125–1136.
- (223) Christ, C. D.; Mark, A. E.; van Gunsteren, W. F. Basic Ingredients of Free Energy Calculations: A Review. *J. Comput. Chem.* **2010**, *31*, 1569–1582.
- (224) Petrov, D.; Perthold, J. W.; Oostenbrink, C.; de Groot, B. L.; Gapsys, V. Guidelines for free-energy calculations using charge changes. *J. Chem. Theory Comput.* **2024**, *20*, 914–925.
- (225) Riniker, S.; van Gunsteren, W. F. Mixing Coarse-Grained and Fine-Grained Water in Molecular Dynamics Simulations of a Single System. *J. Chem. Phys.* **2012**, *137*, No. 044120.
- (226) Reif, M. M.; Hunenberger, P. H.; Oostenbrink, C. New Interaction Parameters for Charged Amino Acid Side Chains in the GROMOS Force Field. *J. Chem. Theory Comput.* **2012**, *8*, 3705–3723.
- (227) Horta, B. A. C.; Merz, P. T.; Fuchs, P. F. J.; Dolenc, J.; Riniker, S.; Hünenberger, P. H. A GROMOS-Compatible Force Field for Small Organic Molecules in the Condensed Phase: The 2016H66 Parameter Set. *J. Chem. Theory Comput.* **2016**, *12*, 3825–3850.
- (228) Berendsen, H. J. C.; Mavri, J. Quantum Simulation of Reaction Dynamics by Density Matrix Evolution. *J. Phys. Chem.* **1993**, *97*, 13464–13468.
- (229) Mavri, J.; Berendsen, H. J. C.; van Gunsteren, W. F. Influence of Solvent on Intramolecular Proton Transfer in Hydrogen Malonate. Molecular Dynamics Simulation Study of Tunneling by Density Matrix Evolution and Nonequilibrium Solvation. *J. Phys. Chem.* **1993**, *97*, 13469–13476.
- (230) Billeter, S. R.; van Gunsteren, W. F. Computer Simulation of Proton Transfers of Small Acids in Water. *J. Phys. Chem. A* **2000**, *104*, 3276–3286.
- (231) Meier, K.; Schmid, N.; van Gunsteren, W. F. Interfacing the GROMOS (Bio) Molecular Simulation Software to Quantum-Chemical Program Packages. *J. Comput. Chem.* **2012**, *33*, 2108–2117.
- (232) Meier, K.; Choutko, A.; Dolenc, J.; Eichenberger, A. P.; Riniker, S.; van Gunsteren, W. F. Multi-Resolution Simulation of Biomolecular Systems: A Review of Methodological Issues. *Angew. Chem., Int. Ed.* **2013**, *52*, 2820; *Angew. Chem.* **2013**, *125*, 2888.
- (233) Lier, B.; Poliak, P.; Marquetand, P.; Westermayr, J.; Oostenbrink, C. BuRNN: Buffer Region Neural Network Approach for Polarizable-Embedding Neural Network/Molecular Mechanics Simulations. *J. Phys. Chem. Lett.* **2022**, *13*, 3812–3818.
- (234) Berendsen, H. J. C. Models for Protein Dynamics. In *Report of the CECAM Workshop on Models for Protein Dynamics*; CECAM: Orsay, France, 1976.
- (235) Rusu, V. H.; Bachmann, S. J.; van Gunsteren, W. F. GROMOS Polarisable Model for Acetone. *Mol. Phys.* **2016**, *114*, 845–854.
- (236) Lin, Z.; van Gunsteren, W. F. On the Effects of Polarisable Solvent Models upon the Relative Stability of an α -Helical and a β -Hairpin Structure of an Alanine Deca-Peptide. *J. Chem. Theory Comput.* **2015**, *11*, 1983–1986.
- (237) Szklarczyk, O. M.; Bachmann, S. J.; van Gunsteren, W. F. A Polarisable Empirical Force Field for Molecular Dynamics Simulation of Liquid Hydrocarbons. *J. Comput. Chem.* **2014**, *35*, 789–801.
- (238) Huang, W.; van Gunsteren, W. F. Challenge of Representing Entropy at Different Levels of Resolution in Molecular Simulation. *J. Phys. Chem. B* **2015**, *119*, 753–763.
- (239) Marrink, S. J.; de Vries, A. H.; Mark, A. E. Coarse Grained Model for Semiquantitative Lipid Simulations. *J. Phys. Chem. B* **2004**, *108*, 750–760.
- (240) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111*, 7812–7824.
- (241) Shinoda, W.; DeVane, R.; Klein, M. L. Multi-Property Fitting and Parameterization of a Coarse Grained Model for Aqueous Surfactants. *Mol. Simul.* **2007**, *33*, 27–36.
- (242) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; Wriggers, W. Atomic-Level Characterization of the Structural Dynamics of Proteins. *Science* **2010**, *330*, 341–346.
- (243) Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E. How Fast-Folding Proteins Fold. *Science* **2011**, *334*, 517–520.