Fixman compensating potential for general branched molecules

Abhinandan Jain,[1,a)] Saugat Kandel,[2] Jeffrey Wagner,[2] Adrien Larsen,[2] and Nagarajan Vaidehi[2,b)]
[1]*Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, California 91109, USA*
[2]*Division of Immunology, Beckman Research Institute of the City of Hope, Duarte, California 91010, USA*

(Received 26 September 2013; accepted 29 November 2013; published online 26 December 2013)

The technique of constraining high frequency modes of molecular motion is an effective way to increase simulation time scale and improve conformational sampling in molecular dynamics simulations. However, it has been shown that constraints on higher frequency modes such as bond lengths and bond angles stiffen the molecular model, thereby introducing systematic biases in the statistical behavior of the simulations. Fixman proposed a compensating potential to remove such biases in the thermodynamic and kinetic properties calculated from dynamics simulations. Previous implementations of the Fixman potential have been limited to only short serial chain systems. In this paper, we present a spatial operator algebra based algorithm to calculate the Fixman potential and its gradient within constrained dynamics simulations for branched topology molecules of any size. Our numerical studies on molecules of increasing complexity validate our algorithm by demonstrating recovery of the dihedral angle probability distribution function for systems that range in complexity from serial chains to protein molecules. We observe that the Fixman compensating potential recovers the free energy surface of a serial chain polymer, thus annulling the biases caused by constraining the bond lengths and bond angles. The inclusion of Fixman potential entails only a modest increase in the computational cost in these simulations. We believe that this work represents the first instance where the Fixman potential has been used for general branched systems, and establishes the viability for its use in constrained dynamics simulations of proteins and other macromolecules. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4851315]

I. INTRODUCTION

Rigid constraints on the higher frequency degrees of freedom, such as bond lengths and bond angles, are often used with molecular dynamics (MD) simulations. The SHAKE, RATTLE, and torsional MD techniques are notable examples of the use of such constraints.[1,2] Torsional MD simulations can help speed up simulations and enhance conformational sampling in the low-frequency torsional degrees of freedom. Constraints also provide a powerful avenue to vary model coarseness. Significant progress has been made in developing algorithms and techniques to effectively tackle the increased complexity of constrained MD,[3] and in applying these MD techniques to protein folding, domain motion, and structure refinement problems.[4–6]

Scheraga, Fixman, and other researchers[7–11] recognized that the use of rigid constraints leads to systematic biases in the statistical behavior of MD simulations of the system, leading to errors in calculated thermodynamic and kinetic properties. Fixman's work in particular has provided considerable insight into the effect of rigid constraints on MD simulations. Fixman proposed the addition of a compensating potential, referred to as the *Fixman potential*, to remove such biases from constrained MD simulations. The Fixman potential depends on the determinant of the mass matrix for the constrained molecular model.

The mass matrix determinant is computationally difficult to evaluate, since it is usually large in size, complex in structure, and configuration dependent. Several researchers have conducted numerical studies to evaluate the effectiveness of the Fixman potential in correcting for the biases in the properties calculated from constrained dynamics. However, due to the computational difficulty in calculating the Fixman potential, these investigations have mostly focused on small and idealized serial chain systems. In these studies, the application of the Fixman potential has been shown to recover the uniform probability distribution function (pdf) for the torsion angles in simulations with rigid constraints.[12–19]

More recently, Patriciu *et al.* have studied the conformational dependence of the Fixman potential for serial chains, and demonstrated computationally that the maximum variation in the Fixman potential increases with increase in the chain length.[20] Echenique *et al.* have studied the relative magnitude of the Fixman potential and the quantum mechanical potential energy for alanine dipeptide,[21,22] and concluded that the Fixman potential becomes significant in simulations of peptides with more than 2 residues. In addition, Brooks and Abagyan have independently proposed alternate methods to circumvent the bias arising from the stiffness in the constrained model[23,24] by deriving corrective torsional potentials to match the flexible model potential energy surface. However, these corrections are *ad hoc* and system dependent and not generalizeable. Although the inclusion of the Fixman potential holds great promise for improving

a)Electronic mail: Abhi.Jain@jpl.nasa.gov
b)Electronic mail: nvaidehi@coh.org

the accuracy of constrained MD simulation methods, the use of such methods has been held back by the serious lack of computationally tractable algorithms to calculate the Fixman potential, especially for large and realistic macromolecules.

Previously, we have used *spatial operator algebra (SOA)* techniques from multibody dynamics to develop the *Generalized Newton-Euler Inverse Mass Operator (GNEIMO)* constrained molecular dynamics method[3,25] that reduces the cost of solving the equations of motion in internal coordinates for molecular systems with rigid constraints. The computational cost of the GNEIMO method scales linearly with the number of degrees of freedom, instead of the cubic cost growth of prior methods.[26] Later, SOA techniques were used to extend the GNEIMO method to develop the tractable GNEIMO-Fixman method for the computation of the Fixman potential.[27] More significantly, the GNEIMO-Fixman method is also able to compute the partial derivatives of the Fixman potential. These partial derivatives define the additional forces, i.e., the Fixman torque, to be applied within constrained MD simulations. The GNEIMO-Fixman method can be used for branched molecules of arbitrary size with only a modest increase in computational cost.

In this work, we have evaluated the GNEIMO-Fixman method by carrying out thorough validation studies and demonstrations on the use of the GNEIMO-Fixman method for branched and realistic molecular systems. We have performed Langevin dynamics simulations of serial and branched molecular systems of different levels of complexity to validate the GNEIMO-Fixman methods. We believe that this work represents the first time that such studies for the Fixman potential have been carried out for general branched molecules without restrictions on the size or the topology of the molecule. The modest computational cost for using the GNEIMO-Fixman method makes possible the routine inclusion of the Fixman potential in constrained MD simulations. The improved accuracy obtained by including the Fixman potential opens up the possibility of a broader use of constrained MD simulations.

II. COMPUTATIONAL METHODS

A. Unconstrained (flexible) models

Cartesian, or absolute coordinates, are the common choice for coordinates in all-atom, unconstrained MD simulations. Curvilinear *bond/angle/torsion (BAT)* coordinates are another important option for describing a molecule's configuration. BAT coordinates are instances of *relative* coordinates and are also often referred to as *internal* coordinates. The importance of BAT coordinates stems from the fact that the conformational motion of molecules is typically determined by the motion of the torsional angles to a much larger degree than by the motion of the bond angle and bond length degrees of freedom. Thus, BAT coordinates offer a natural way for analyses and models to focus on the dominant motion. For this reason, entropy analyses in absolute coordinate MD often involves the transformation of covariance data into BAT coordinates.[28]

With motion dominated by torsion angles, a natural path to simplifying molecular models is to use rigid constraints to freeze and eliminate some of the bond length and bond angle BAT coordinates from the model. *Torsional MD* is a well known example of such constrained MD where all the bond and angle coordinates are frozen using rigid constraints, and only the torsional coordinates are allowed to vary. Rigid constraints are also often used in entropy and free energy evaluations.[29–32]

Unconstrained systems are often referred to as *flexible* models. We partition the 3n BAT coordinates into $\mathcal{N}$ unconstrained coordinates, denoted as α, and $(3n - \mathcal{N})$ coordinates, denoted as q, that are to be constrained. In the flexible model, both q and α are allowed to vary, though the variation in the q coordinates is often limited by the bond potentials. In the constrained model, the q coordinates have a fixed value, q_0, and only the α coordinates are allowed to vary.

With p denoting the canonical momentum coordinates for the unconstrained system, the Hamiltonian, $\mathcal{H}(\alpha, q, p)$, has the form

$$\mathcal{H}(\alpha, q, p) = \frac{1}{2} p^* \mathcal{M}_B^{-1}(\alpha, q) p + \mathcal{U}(\alpha, q), \tag{1}$$

where $\mathcal{M}_B \in \mathcal{R}^{3n\times 3n}$ denotes the BAT coordinates mass matrix for the unconstrained system, $\mathcal{U}$ the force field potential function, and n the number of atoms. For a temperature, T, the ensemble partition function, $\mathcal{Z}(T)$, is given by

$$\mathcal{Z}(T) = c_1 \int dp\, d\alpha\, dq\, e^{-\mathcal{H}(\alpha,q,p)/kT}, \tag{2}$$

where c_1 is a scaling constant, k the Boltzmann constant, and T the temperature.

Using Eq. (1) in Eq. (2), and integrating over the momentum coordinates yields the following expression for the configuration space partition function:[33]

$$\mathcal{Z}(T) = c_2 \int d\alpha\, dq\, \det\left\{\mathcal{M}_B^{\frac{1}{2}}(\alpha, q)\right\} e^{-\mathcal{U}(\alpha,q)/kT}. \tag{3}$$

The probability density function (pdf), $\rho(\alpha, q)$, for the unconstrained configuration coordinates thus is

$$\rho(\alpha, q) \quad \propto \quad \det\left\{\mathcal{M}_B^{\frac{1}{2}}(\alpha, q)\right\} e^{-\mathcal{U}(\alpha,q)/kT}. \tag{4}$$

Note that

$$\mathcal{M}_B = \mathcal{J}_B^* \mathcal{M}_c \mathcal{J}_B, \tag{5}$$

where $\mathcal{J}_B \in \mathcal{R}^{3n\times 3n}$ is the Jacobian for the transformation between the BAT and Cartesian coordinates, and $\mathcal{M}_c \in \mathcal{R}^{3n\times 3n}$ is the constant Cartesian coordinates diagonal mass matrix consisting of the atom masses along the diagonal. Go and Scheraga[10,34] developed the following simple expression for $\det\{\mathcal{J}_B\}$ for general branched systems,

$$\det\{\mathcal{J}_B\} = \sin\theta_{ex}\, d_2^2 \prod_{i=3}^{n} d_i^2 \sin\theta_i, \tag{6}$$

where d_i denote the $(n-1)$ bond lengths, θ_i denote the $(n-2)$ bond angles, and θ_{ex} denotes an overall molecule orientation coordinate. A remarkable property of this expression is that it is independent of the torsional angles, and only depends on

bond lengths and angles. Using this expression in Eq. (5), we obtain

$$\det\{\mathcal{M}_B\} = \det\{\mathcal{J}_B\}^2 \prod_{i=1}^{n} \mathfrak{m}_i^3 = \sin^2\theta_{ex}\, d_2^4 \prod_{i=3}^{n} d_i^4 \sin^2\theta_i \prod_{i=1}^{n} \mathfrak{m}_i^3, \quad (7)$$

where $\mathfrak{m}_i$ denotes the mass of the ith atom. This expression is simple to evaluate. When the potential $\mathcal{U}(\alpha, q)$ is independent of the torsion angles (e.g., for bond angle and bond length potentials), the right-hand side of Eq. (4) is entirely independent of torsion angles. For this case, the pdf for any torsion angle β_i is thus uniform and given by

$$\rho(\beta_i) = \frac{1}{2\pi}. \quad (8)$$

In other words, in the presence of just bond angle and bond length potentials, all torsional angles have uniform pdfs for unconstrained molecular models.

B. Constrained models

In constrained models, hard constraints are used to freeze the $(3n - \mathcal{N})$ coordinates q at a constant value, q_0. The constrained model configuration partition function has the form

$$\mathcal{Z}(T) = c_3 \int d\alpha\ \det\left\{\mathcal{M}^{\frac{1}{2}}(\alpha)\right\} e^{-\mathcal{U}(\alpha, q_0)/kT}, \quad (9)$$

where $\mathcal{M}(\alpha) \in \mathcal{R}^{\mathcal{N}\times\mathcal{N}}$ denotes the mass matrix for the constrained model. This matrix is the unconstrained α coordinates sub-block of the $\mathcal{M}_B$ flexible mass matrix. The pdf for the constrained configuration coordinates for the constrained model thus has the form

$$\rho(\alpha) \quad \propto \quad \det\left\{\mathcal{M}^{\frac{1}{2}}(\alpha)\right\} e^{-\mathcal{U}(\alpha, q_0)/kT}. \quad (10)$$

Unlike $\det\{\mathcal{M}_B\}$, $\det\{\mathcal{M}\}$ does depend on the torsion angles. Comparing Eqs. (4) and (10), it is evident that the pdfs for the flexible and constrained cases are different. This results in systematic differences in the statistical behavior of the flexible and constrained models. Thus, while the torsion angles have uniform pdf in the presence of just bond angle and bond length potentials for the flexible model, the pdf for the torsion angles is no longer uniform for constrained models.

In order to remove the statistical biases introduced by the constraints, Fixman[7,8] proposed the use of the following $\mathcal{U}'(\alpha)$ modified potential function

$$\mathcal{U}'(\alpha) \triangleq \mathcal{U}(\alpha, q_0) + \mathcal{U}_f(\alpha) \quad \text{where}$$
$$\mathcal{U}_f(\alpha) \triangleq \frac{1}{2} kT \ln \frac{\det\{\mathcal{M}(\alpha)\}}{\det\{\mathcal{M}_B(\alpha, q_0)\}} \quad (11)$$

in place of $\mathcal{U}(\alpha, q_0)$ in constrained MD simulations. The "$\triangleq$" symbol above denotes "defined as." The compensating potential, $\mathcal{U}_f$, is referred to as the *Fixman potential*. Substituting $\mathcal{U}'$ into Eq. (10) transforms it into agreement with Eq. (4) for $q = q_0$. The compensated pdfs have the form

$$\rho(\alpha) \quad \propto \quad \det\left\{\mathcal{M}_B^{\frac{1}{2}}(\alpha, q_0)\right\} e^{-\mathcal{U}(\alpha, q_0)/kT}. \quad (12)$$

Thus, the use of the Fixman compensating potential with the constrained model removes the statistical biases for the α coordinates introduced by the constraints.

The contribution of the $\mathcal{M}_B$ term to the $\mathcal{U}_f$ Fixman potential in Eq. (11) is easy to evaluate using Eq. (7). Indeed, when the α unconstrained coordinates are just torsion angles, this contribution does not depend on α, and is constant. For this case, the Fixman potential expression in Eq. (11) and the compensated pdf $\rho(\alpha)$ in Eq. (12) simplify to

$$\mathcal{U}_f(\alpha) = c_f + \frac{1}{2} \ln\det\{\mathcal{M}(\alpha)\} \quad \text{and} \quad \rho(\alpha) \propto e^{-\mathcal{U}(\alpha, q_0)/kT}, \quad (13)$$

where c_f is a constant. *Without any loss in generality, we assume from here on that the α unconstrained coordinates are indeed just the torsion angles.* This assumption will help simplify the rest of the development in this paper.

The key computational challenge for evaluating the Fixman potential lies with the $\det\{\mathcal{M}\}$ term. The $\mathcal{M}$ matrix is dense and configuration dependent, and the associated Jacobian matrices are non-square. For very small systems, it is possible to derive explicit symbolic expressions for the determinant.[12,13] However, this approach does not scale well and is intractable for even moderate size systems.

Fixman[7] developed a method to evaluate the mass matrix determinant for serial-chain topology systems (i.e., systems with no branches). His method exploits the sparse structure of the constrained sub-block of $\mathcal{M}_B^{-1}$. This expression has been used in Refs. 13 and 14 and others for numerical studies of the Fixman potential for small serial topology constrained MD. Fixman's formula becomes more complex to use for larger serial systems, and cannot be used for branched topology systems. The lack of practical, general-purpose computational methods for evaluating the Fixman potential has limited its use to academic studies over the years.

Use of the Fixman potential in a constrained dynamics simulation requires the evaluation of the Fixman torque for each of the unconstrained coordinates and applying these torques for each angle coordinate when solving the equations of motion. The Fixman torque for an unconstrained coordinate angle is the partial derivative of the Fixman potential with respect to the coordinate angle. Thus, the Fixman torque, $\mathcal{T}(i)$, for the ith coordinate α_i is given by

$$\mathcal{T}(i) = -\frac{\partial \mathcal{U}_f(\alpha)}{\partial \alpha_i}. \quad (14)$$

Not surprisingly, given the challenges in evaluating just the Fixman potential, no general method has been available for computing the Fixman torque. The sole exception to this is the GNEIMO-Fixman method described in Sec. II C.

C. GNEIMO-Fixman method for calculating the Fixman potential

Jain[27] developed the GNEIMO-Fixman general method for evaluating the Fixman potential and torque, for arbitrary size, serial, and branched constrained models. This method

makes use of the *SOA* analytical techniques for the factorization and inversion of the constrained mass matrix $\mathcal{M}$. For the sake of completeness we paraphrase the main ideas here, and refer the reader to Refs. 3 and 35 for detailed definitions and derivations. The key spatial operator expressions for the factorization and inversion of the mass matrix[3,35] are as follows:

$$\begin{aligned} \mathcal{M} &= \mathrm{H}\phi\mathrm{M}\phi^{*}\mathrm{H}^{*}, \\ \mathcal{M} &= [\mathrm{I}+\mathrm{H}\phi\mathcal{K}]\mathcal{D}[\mathrm{I}+\mathrm{H}\phi\mathcal{K}]^{*}, \\ [\mathrm{I}+\mathrm{H}\phi\mathcal{K}]^{-1} &= [\mathrm{I}-\mathrm{H}\psi\mathcal{K}], \\ \mathcal{M}^{-1} &= [\mathrm{I}-\mathrm{H}\psi\mathcal{K}]^{*}\mathcal{D}^{-1}[\mathrm{I}-\mathrm{H}\psi\mathcal{K}]. \end{aligned} \tag{15}$$

The "*" superscript above denotes matrix transpose. The first expression is known as the Newton-Euler operator factorization of the mass matrix $\mathcal{M}$ in terms of the block-diagonal H hinge articulation, the lower-triangular ϕ rigid body propagation, and the block-diagonal M link spatial inertia operators. The block diagonal terms of H define the torsional axes of rotation for each torsional degree of freedom. The block diagonal elements of M are the spatial inertias for each of the clusters. The ϕ(i, j) block element propagates spatial forces on the *j*th cluster rigidly to the ith cluster. While this factorization has non-square factors, the second expression describes an alternative factorization involving only square factors with block diagonal $\mathcal{D}$ and block lower-triangular $[\mathrm{I}+\mathrm{H}\phi\mathcal{K}]$ matrices. This factorization involves additional spatial operators that are associated with the *articulated body (AB)* recursive algorithm[3] for the system. The next expression describes an analytical expression for the inverse of the $[\mathrm{I}+\mathrm{H}\phi\mathcal{K}]$ operator. Using this leads to the final analytical expression for the inverse of the mass matrix. These operator expressions hold generally for branched systems, irrespective of the number of bodies, the types of hinges, and the specific topological structure of the system.[35] The last expression with the operator factorization of the mass matrix inverse forms the basis for the recursive GNEIMO constrained dynamics algorithm[3] whose computational cost scales just linearly with the number of degrees of freedom in the constrained model.

From the second expression in Eq. (15), it follows that

$$\begin{aligned} \det\{\mathcal{M}\} &= \det\{\mathrm{I}+\mathrm{H}\phi\mathcal{K}\}^{2}\det\{\mathcal{D}\} \\ &= \prod_{\mathrm{i}=0}^{\mathcal{N}-6}\det\{\mathcal{D}(\mathrm{i})\} = \det\{\mathcal{D}(0)\}\prod_{\mathrm{i}=1}^{\mathcal{N}-6}\mathcal{D}(\mathrm{i}). \end{aligned} \tag{16}$$

The above expression uses the fact that $\det\{\mathrm{I}+\mathrm{H}\phi\mathcal{K}\}=1$ and that $\mathcal{D}$ is block-diagonal. For molecular models, all of the $\mathcal{D}$(i) elements are scalars except for the base cluster $\mathcal{D}(0)$ which is a 6×6 matrix. The $\mathcal{D}$(i) terms are available as a by product of the GNEIMO constrained dynamics algorithm. Substituting Eq. (16) into Eq. (13), the Fixman potential expression becomes

$$\mathcal{U}_{\mathrm{f}}(\alpha) = \mathrm{c}_{\mathrm{f}} + \frac{1}{2}\ln\det\{\mathcal{D}(0)\} + \frac{1}{2}\sum_{\mathrm{i}=1}^{\mathcal{N}-6}\ln\mathcal{D}(\mathrm{i}). \tag{17}$$

Since all of the $\mathcal{D}$(i) are available from the GNEIMO algorithm, Eq. (17) provides a simple way, and with negligible additional cost, to extend the GNEIMO constrained dynamics algorithm to also evaluate the Fixman potential for arbitrary sized branched molecular systems! This is the GNEIMO-Fixman method for evaluating the Fixman potential as originally described in Ref. 27

D. GNEIMO-Fixman method for calculating the Fixman torque

Combining Eqs. (13) and (14), the expression for the Fixman torque for the coordinate α_{i} is

$$\mathcal{T}(\mathrm{i}) = -\frac{1}{2}\frac{\partial\ln\det\{\mathcal{M}(\alpha)\}}{\partial\alpha_{\mathrm{i}}}. \tag{18}$$

Jain[27] developed a simple expression for this Fixman torque. We state here standard results from matrix calculus that play a key role in these simplifications. For any smooth scalar function g(X(y)) of a matrix $\mathrm{X}\in\mathcal{R}^{\mathrm{m}\times\mathrm{n}}$, we have the identity

$$\frac{\partial\mathrm{g}(\mathrm{X}(\mathrm{y}))}{\partial\mathrm{y}} = \mathrm{Trace}\left\{\left[\frac{\partial\mathrm{g}}{\partial\mathrm{X}}\right]^{*}\frac{\partial\mathrm{X}(\mathrm{y})}{\partial\mathrm{y}}\right\}, \tag{19}$$

where $\frac{\partial\mathrm{g}}{\partial\mathrm{X}}$ and $\frac{\partial\mathrm{X}(\mathrm{y})}{\partial\mathrm{y}}$ are m × n matrices whose elements are defined as

$$\frac{\partial\mathrm{g}}{\partial\mathrm{X}}(\mathrm{i},\mathrm{j}) \triangleq \frac{\partial\mathrm{g}}{\partial\mathrm{X}(\mathrm{i},\mathrm{j})} \quad\text{and}\quad \frac{\partial\mathrm{X}}{\partial\mathrm{y}}(\mathrm{i},\mathrm{j}) \triangleq \frac{\partial\mathrm{X}(\mathrm{i},\mathrm{j})}{\partial\mathrm{y}}. \tag{20}$$

Another useful matrix calculus identity is[36]

$$\frac{\partial\ln\det\{\mathrm{X}\}}{\partial\mathrm{X}} = \{\mathrm{X}^{*}\}^{-1}. \tag{21}$$

References 27 and 35 use these matrix calculus expressions, and other SOA analysis, to derive explicit expressions for the Fixman torque. Reference 35 uses Eq. (17) as a point of departure while Ref. 27 starts with Eq. (18). For the sake of completeness we include the final expressions for the Fixman torque, but refer the reader to the original references for the derivation and definition details. The Fixman torque for the ith torsion angle is[27]

$$\mathcal{T}(\mathrm{i}) = -\,\mathrm{Trace}\{\mathcal{P}(\mathrm{i})\Upsilon(\mathrm{i})\,\widetilde{\mathrm{H}}^{*}_{\omega}(\mathrm{i})\}. \tag{22}$$

The $\mathcal{P}$(i), Υ(i), and $\widetilde{\mathrm{H}}^{*}_{\omega}$(i) are all 6×6 matrices. A slightly simpler version of Eq. (22) is

$$\begin{aligned} \mathcal{T}(\mathrm{i}) &= \mathrm{h}^{*}(\mathrm{i})\,\mathcal{F}[\mathrm{Q}_{11}+\mathrm{Q}_{22}] \quad \text{with} \\ \mathcal{P}(\mathrm{i})\Upsilon(\mathrm{i}) &= \begin{bmatrix}\mathrm{Q}_{11} & \mathrm{Q}_{12}\\ \mathrm{Q}_{21} & \mathrm{Q}_{22}\end{bmatrix},\quad \mathrm{Q}_{\mathrm{ij}}\in\mathcal{R}^{3\times 3}. \end{aligned} \tag{23}$$

In the above, $\mathcal{F}$(A) maps an arbitrary 3×3 matrix A into a 3-vector v such that the A − A* skew-symmetric matrix is the 3×3 cross-product matrix for v. Reference 35 includes a generalization of the above expression for the case where the α coordinates happen to also include bond length coordinates.

It is worth emphasizing that the expression for the torque in Eq. (23) holds for branched systems of any size. It forms the basis of the GNEIMO-Fixman method for evaluating Fixman torques. The $\mathcal{P}$(i) matrix for each cluster is once again a by-product of the GNEIMO algorithm for solving the constrained dynamics. The computation of the 6×6 Υ(i) matrices for the clusters on the other hand requires an additional

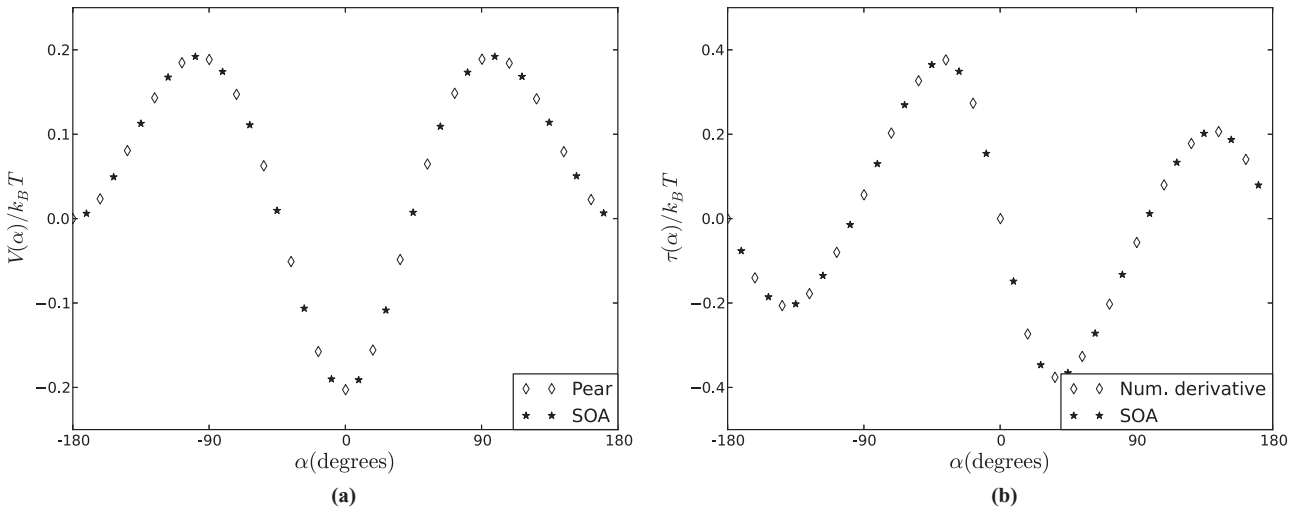


FIG. 1. Fixman potential and torque plots for a three-bond (C4) chain with 90° bond angle. (a) Comparison of Fixman potential values from the analytical expression in Eq. (24) and the SOA based expression in Eq. (17). (b) Comparison of Fixman torque values from the SOA based expression in Eq. (23) and those from numerically differentiating the Fixman potential from Eq. (17).

recursive scatter algorithm that starts at the base cluster.[27,35] This recursion also makes use of quantities available from the GNEIMO dynamics algorithm. Its computational cost scales linearly with the number of torsion degrees of freedom. Thus, the overall cost of computing and including the Fixman torque into the simulation remains of linear computational complexity, and adds only marginally to the GNEIMO constrained dynamics computational cost as seen later in Sec. III. The generality and low cost of the GNEIMO-Fixman method for evaluating $\mathcal{T}$(i) is in stark contrast with the symbolic techniques previously employed[8] that are practical for only small serial systems. The GNEIMO-Fixman method makes possible the routine inclusion of the Fixman potential in constrained MD simulations. We now present our results on validating and testing the effect of the Fixman potential and torque on constrained MD simulations of serial and branched molecules.

III. RESULTS AND DISCUSSION

A. Numerical validation of the GNEIMO-Fixman potential and torque calculations

The GNEIMO-Fixman expressions in Eqs. (17) and (23) can be used to recursively compute the Fixman potential and torque, respectively, for any unconstrained degree of freedom in a serial or branched molecule. We describe here numerical results that validate the correctness of these expressions for our computations. We begin with the Fixman potential, since there exists prior work that we can compare against.

Pear and Weiner[12,13] derived symbolic closed form expressions for the Fixman potential for the C4 idealized 3 bond serial chain system with fixed bond lengths, and with the bond angles set to 90°. For this system, they calculated the mass matrix determinant to be

$$\det\{\mathcal{M}(\alpha)\} = c_5(35 + 4\cos(\alpha) - 16\cos^2(\alpha) + \cos^4(\alpha), \tag{24}$$

where α is the single torsion angle, and c_5 is a constant dependent on the bond lengths and the masses in the system. A comparison of the Fixman potential values obtained from Eq. (24) with those calculated using Eq. (17) for the GNEIMO-Fixman method, presented in Figure 1(a), shows excellent agreement between them.

Patriciu *et al.*[20] used an algorithmic technique, similar to the one in Jain,[27] to calculate the Fixman potential for idealized serial chains of arbitrary length. Figure 2 shows the contour of our Fixman potential values for the two torsion angles in the C5 four-bond serial chain system. These are in excellent agreement with the contour plot presented by Patriciu *et al.*[20] The close numerical match with prior published data provides positive validation evidence for the GNEIMO-Fixman techniques for the Fixman potential – for at least serial systems.

Since there is no prior data for the Fixman torque in literature, we validated our expression for the Fixman torque, Eq. (23), by comparing its calculated values to those obtained by numerically differentiating the Fixman potential values from Eq. (17). Figure 1(b) shows that the values obtained via the two different methods are identical for the C4

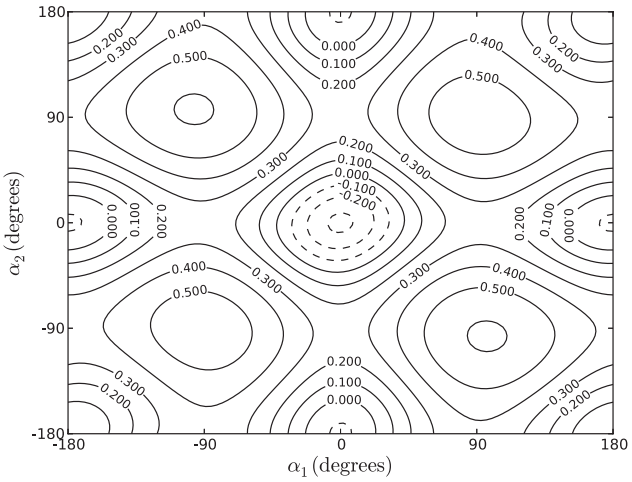


FIG. 2. Contour plot of the Fixman potential for the pair of torsion angles for a C5 chain with 109° bond angle.

system. We also carried out similar validation studies for the Fixman torque for larger systems. Figure S1 in the supplementary material[44] shows the contour plot of the Fixman torque for two torsional angles in C5 where we observe similar agreement.

B. Descriptions of the MD simulations

We carried out MD simulations to study the effectiveness of the Fixman potential in recovering the thermodynamic pdf of torsion angles for both serial and branched systems. The effect of Fixman potential can be most clearly seen by comparing the pdfs calculated from torsional MD simulations with Fixman potential to the pdfs from unconstrained MD simulations. To enhance conformational sampling, we used Langevin dynamics for all the simulations. All the simulations were performed for isolated single molecules. The simulations were carried out at 300 K temperature and with a damping constant of 0.01/fs. We performed three sets of simulations, namely, FLEXIBLE, TORSIONAL, and FIXMAN, for each system. These are:

1. FLEXIBLE: Flexible Cartesian Langevin dynamics simulation with only bond length and bond angle potentials. The non-bonded Coulombic and van der Waals potentials were turned off in these simulations. The $\mathcal{U}(\alpha, \mathrm{q})$ potential here does not depend on torsional angles, and therefore as discussed earlier, every α_{i} torsion coordinate in FLEXIBLE simulations has an expected uniform pdf given by Eq. (8):
$$\rho(\alpha_{\mathrm{i}}) = \frac{1}{2\pi}.$$
2. TORSIONAL: Constrained Langevin dynamics simulation *without* Fixman potential. These are torsional dynamics simulations where the bond length and angle coordinates are fixed as hard constraints and hence these potentials were not used. The non-bonded Coulombic and van der Waals potentials were turned off in these simulations. For this case the expected torsion pdfs are given by Eqs. (10) and (16):
$$\rho(\alpha) \propto \det\left\{\mathcal{M}^{\frac{1}{2}}(\alpha)\right\} = \left[\det\{\mathcal{D}(0)\}\prod_{\mathrm{i}=1}^{\mathcal{N}-6}\mathcal{D}(\mathrm{i})\right]^{\frac{1}{2}}. \tag{25}$$
3. FIXMAN: Constrained Langevin dynamics simulation *with* Fixman potential. These are the same simulations as TORSIONAL, but with the additional Fixman potential included. In this case the expected torsion pdfs given by Eq. (13) once again simplify to the uniform pdf
$$\rho(\alpha_{\mathrm{i}}) = \frac{1}{2\pi}.$$

The FLEXIBLE simulations do not use any hard constraints, and establish the expected uniform pdf for the torsion angles. The TORSIONAL simulations measure the pdf deviations from the uniform distribution stemming from the constraints. The probability distributions for the torsions in these simulations are governed by Eq. (25). The FIXMAN simulations assess the extent to which the use of the Fixman potential in the constrained simulations recover the uniform distributions for the FLEXIBLE simulations.

For the unconstrained FLEXIBLE simulations, the Langevin dynamics has the form
$$\mathrm{M}\ddot{\mathrm{x}} = -\nabla_{\mathrm{x}}\mathcal{U} - \gamma\mathrm{M}\dot{\mathrm{x}} + \sqrt{2\gamma\mathrm{kT}}\,\mathrm{M}^{\frac{1}{2}}\mathrm{dW}, \tag{26}$$
where x denotes the Cartesian coordinates, γ is the damping coefficient, dW is a vector of independent Weiner processes, and $\mathcal{U}$ the bond angle and bond length potentials. The BBK algorithm[37] was used for the numerical integration of the equations of motion.

C. MD simulation results for serial chains

We performed FLEXIBLE, TORSIONAL, and FIXMAN simulations for C4, C5, C11, and C15 serial chain systems. Each simulation was run with a timestep of 1 fs for a total of 50 ns, and the coordinates and energy were recorded every 100 steps. In these simulations, the beads were assigned equal masses of 14 amu, and equal bond lengths of 1.54 Å and 90° bond angle. For the FLEXIBLE simulation, the bond and angle spring constants were set to 83.66 kcal/Å ^{2} and 43.46 kcal, respectively.

For the C4 system, Pear and Weiner[12] have previously shown that the constrained model with the Fixman potential recovers the uniform pdf for the torsion angles. Figure 3(b) shows the pdf for various values of the single torsion angle in C4 system. As expected the FLEXIBLE simulations show a uniform pdf while the TORSIONAL simulations show a bimodal biased distribution. The dashed lines in Figures 3(a) and 3(c) represent the predicted distributions from Eq. (8), and in Figure 3(b) from Eq. (25). The constrained model's torsion angle has a bimodal probability distribution with

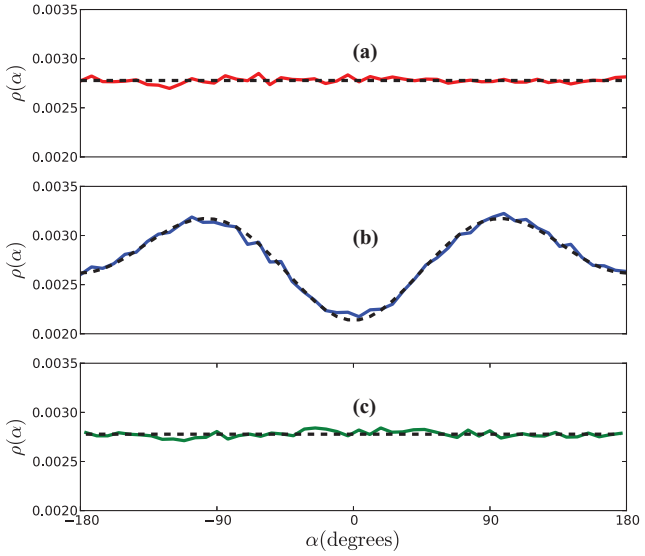


FIG. 3. Torsion angle distributions for the three-bond C4 chain, for (a) FLEXIBLE, (b) TORSIONAL, and (c) FIXMAN simulations. The dashed lines represent the predicted distributions, with the ones in (a) and (c) being uniform pdfs from Eq. (8), and the one in (b) from Eq. (25). The bin size for the histograms is $\mathrm{d}\alpha = 7.2°$.

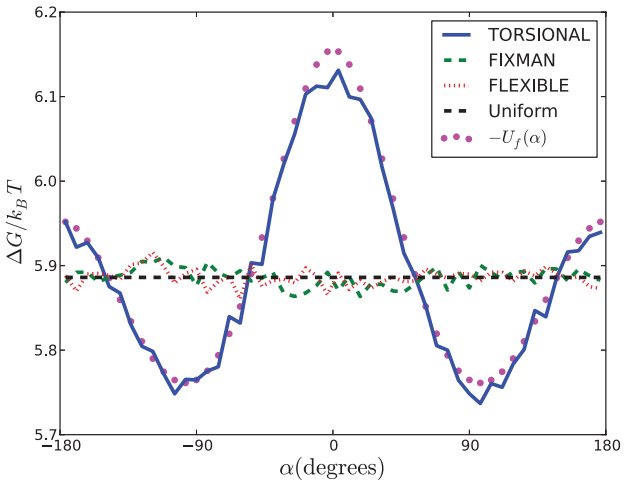


FIG. 4. Free energy distribution in a three-bond chain C4 system, plotted as a function of the torsion angle. The negative of the Fixman potential, $-\mathcal{U}_f$, is very close to the free energy in the TORSIONAL simulation, demonstrating that the application of the Fixman potential corrects the bias in the free energy landscape of the TORSIONAL model. The free energies were calculated using bins of size 7.2° for the torsion angles.

maxima at ≈±83° and minima at 0° and ±180° as seen in Figure 3(b). The presence of such a bimodal structure in the pdf illustrates the inherent bias in TORSIONAL simulations. This bias can affect the torsion angle probability distribution function, the transition rates between conformational states and the free energy surface. The FIXMAN simulation recovers the uniform pdf as seen in the FLEXIBLE simulation (Figure 3). This effectively demonstrates that the application of the Fixman potential compensates for the bias and removes the distortions of the potential energy surface introduced by the hard constraints in this system.

The corrective feature of the Fixman potential can also be observed in the free energy landscape derived from the pdf of the simulation (Figure 4). The free energy was calculated using the equation $\Delta G(\alpha_i)/kT = -\ln(\rho(\alpha_i))$. As such, the uniform free energy distribution shown in this figure is equal to $-\ln(1/360)$. A comparison of the free energy in the TORSIONAL simulation of the C4 system (shown in blue line in Figure 4) to the negative of the Fixman potential energy from the FIXMAN simulation (dotted pink line) shows that the measured value of the Fixman potential is almost equal to and opposite in sign to the inherent free energy in the TORSIONAL simulation, and that the net free energy in the FIXMAN simulation agrees with that of the FLEXIBLE simulation (broken green and red lines).

In addition to the three-bond C4 chain, we performed Langevin dynamics simulations and calculated the pdf for all the torsion angles for a *four-bond* serial chain C5 (Figure S2 in the supplementary material[44]), a *ten-bond* serial chain C11 (Figure S3 in the supplementary material[44]), and *fourteen-bond serial chain C15*. For every torsion angle in all of these cases, the application of the Fixman potential recovered the ideal uniform distribution of the flexible model. Figure 5 shows the computed distributions for the C15 TORSIONAL simulations, as well as the recovered uniform distributions for the same system in the FIXMAN simulations. This again demonstrates that the Fixman potential corrects for the bias generated in constrained dynamics.

The three-dimensional structures corresponding to the global minimum and maximum of the Fixman potential, as extracted from the trajectories for C5 and C11 systems are shown, in Figures 6(a) and 6(b). Our results are consistent with the corresponding results presented by Patriciu *et al.*,[20] who observed that, for serial chains, the minimum of the Fixman potential occurs when the structure is in a

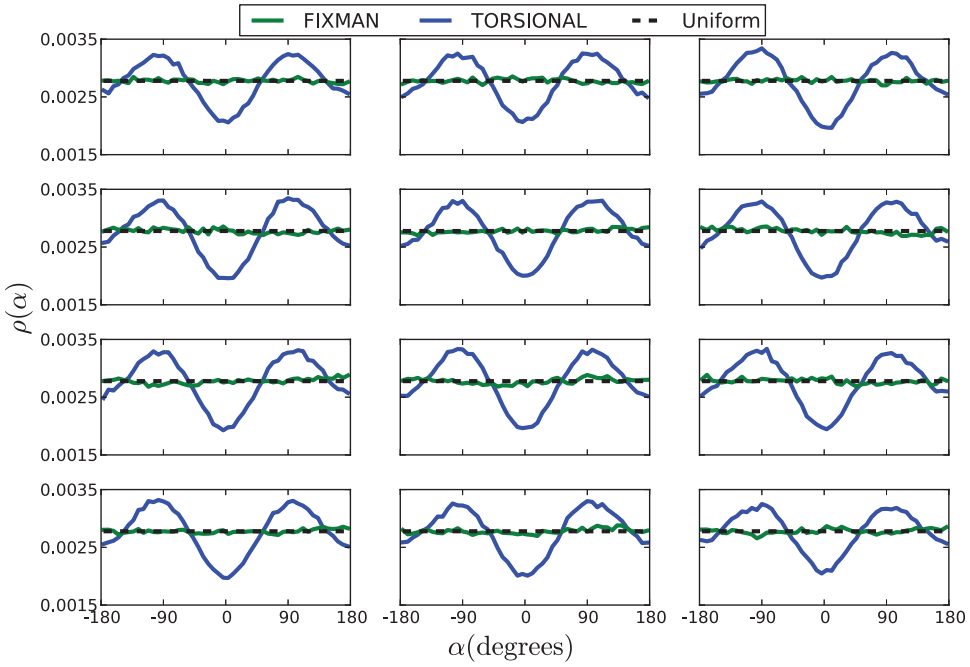
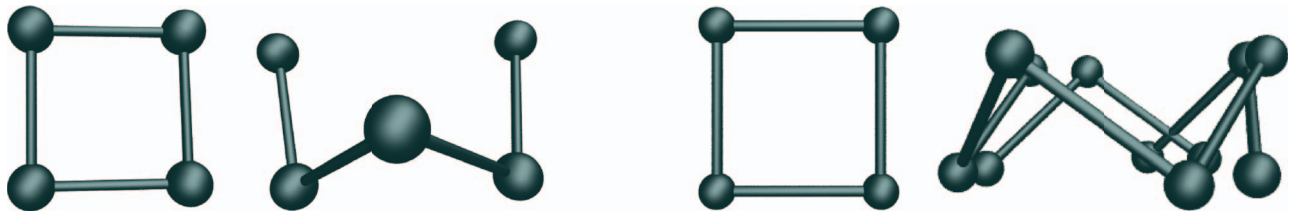


FIG. 5. Probability distribution functions for torsion angles in a fourteen-bond C15 chain. The application of the Fixman potential in the TORSIONAL model recovers the uniform pdf for all the torsion angles. The bin size for the histograms is $d\alpha = 7.2°$.



(a) (b)

FIG. 6. The global minimum and maximum configurations of the Fixman potential for (a) the C5 and (b) the C11 molecule correspond to the self-intersecting planar and self-intersecting spatial conformations, respectively. In the self-intersecting planar conformation, each torsion angle is 0°. In the self-intersecting spatial conformation, each torsion angle is ≈±83°.

self-intersecting planar configuration, while the maximum occurs when the structure is in a self-intersecting spatial configuration. In the self-intersecting planar conformations, the value of each torsion angle corresponds to the global minimum of the Fixman potential for a three-bond chain, with the torsion angle 0°, as presented in Figure 1(a). Similarly, in the self-intersecting spatial conformation, the value of each torsion corresponds to one of the maxima of the Fixman potential for a three-bond chain, ≈±83°.

D. MD simulations for branched systems

In Sec. III C, we simulated idealized serial chains consisting of identical masses, bond lengths, and bond angles. Natural polymers, in contrast, are much more complex – they are generally asymmetric, branched systems consisting of large numbers of atoms with varying masses, bond lengths, and bond angles. The application of the Fixman potential on constrained MD simulations of such branched systems has, to our knowledge, not been carried out so far.

To investigate the effect of the application of the Fixman potential in generalized branched systems, we performed Langevin dynamics FLEXIBLE, TORSIONAL, and FIXMAN simulations on three realistic branched peptides, namely, alanine dipeptide, valine dipeptide, and a ten amino acid peptide chignolin.[38] The pdf of the torsion angles were calculated from the trajectories obtained from these simulations.

1. Small branched peptides

In this section, we look at the equilibrium statistical properties of two small branched peptide systems, alanine dipeptide, and valine dipeptide, whose chemical structures are shown in Figures 7(a) and 7(b). The peptides are modeled as a collection of rigid bodies, referred to as *clusters*, connected by flexible hinges. Each rigid cluster is a collection of atoms within which all bond lengths and bond angles are frozen via rigid holonomic constraints, and the hinges are the torsions connecting these rigid clusters. All the terminal bonds are treated as rigid, as are the aromatic ring moieties present in the side chains of phenylalanine, tyrosine, tryptophan, and histidine. In this clustering scheme the alanine and valine dipeptides have eight and ten clusters, respectively. The bond length and bond angle parameters are from the AMBER99SB forcefield.[39] For each of the FLEXIBLE, TORSIONAL, and FIXMAN sets, we ran 3 separate simulations, each with a timestep of 1 fs, and a simulation length of 100 ns, for a total of 300 ns of simulation data. The coordinates and energy were recorded every 200 steps.

Figures 7(c) and 7(d) show the pdfs for the $C^1 - N^2 - C^2_\alpha - C^2$ (marked by green and red circles in Figs. 7(a) and 7(b), respectively) torsion for both the alanine dipeptide and valine dipeptide molecules. We observe that, for both these systems, the TORSIONAL simulations produce bimodal distributions, while the FIXMAN simulations recover the uniform pdf from the FLEXIBLE simulations. To quantify the effectiveness of the Fixman potential in recovering the pdf for each torsion in the system, we calculated the average root mean square deviation of the pdfs obtained from the simulations, from the uniform pdf. These RMS deviations are denoted R_{flex}, R_{tor}, and R_{Fix} for the FLEXIBLE, TORSIONAL, and FIXMAN simulation pdfs, respectively.

For alanine dipeptide, we find that $R_{tor} = 1.0 \times 10^{-3}$, $R_{Fix} = 8.7 \times 10^{-5}$, and $R_{flex} = 6.5 \times 10^{-5}$. For valine dipeptide, we find that $R_{tor} = 8.4 \times 10^{-4}$, $R_{Fix} = 7.9 \times 10^{-5}$, and $R_{flex} = 6.6 \times 10^{-5}$. This again shows that, with the application of the Fixman potential, we recover the uniform pdfs as observed for the flexible model.

In addition to individual torsions, we also evaluated the joint probability distribution functions for pairs of torsion angles for the three sets of FLEXIBLE, TORSIONAL, and FIXMAN simulations. A plot of the joint pdf of $C^1_\alpha - C^1 - N^2 - C^2_\alpha$ torsion (α_1) versus the $C^1 - N^2 - C^2_\alpha - C^2$ torsion (α_2) is presented in Figure 8. The joint pdf of backbone torsion angles from FLEXIBLE simulations (Figure 8(a)) is close to flat, and so is the joint pdf for FIXMAN (Figure 8(c)). The joint pdf from TORSIONAL simulations on the other hand are biased as shown in Figure 8(b). This demonstrates that the Fixman potential recovers not just the single torsion pdfs, but also joint torsion pdfs for pairs of torsion angles.

2. Moderate sized branched peptide

We tested the effects of the Fixman potential for a larger branched peptide, namely, the ten amino acid β-hairpin peptide called chignolin (Figure 9). Again, TORSIONAL, FLEXIBLE, and the FIXMAN Langevin simulations were performed. The terminal bonds and rings were clustered, as in the case of the alanine dipeptide, for the TORSIONAL and FIXMAN simulations. We used the nuclear magnetic resonance

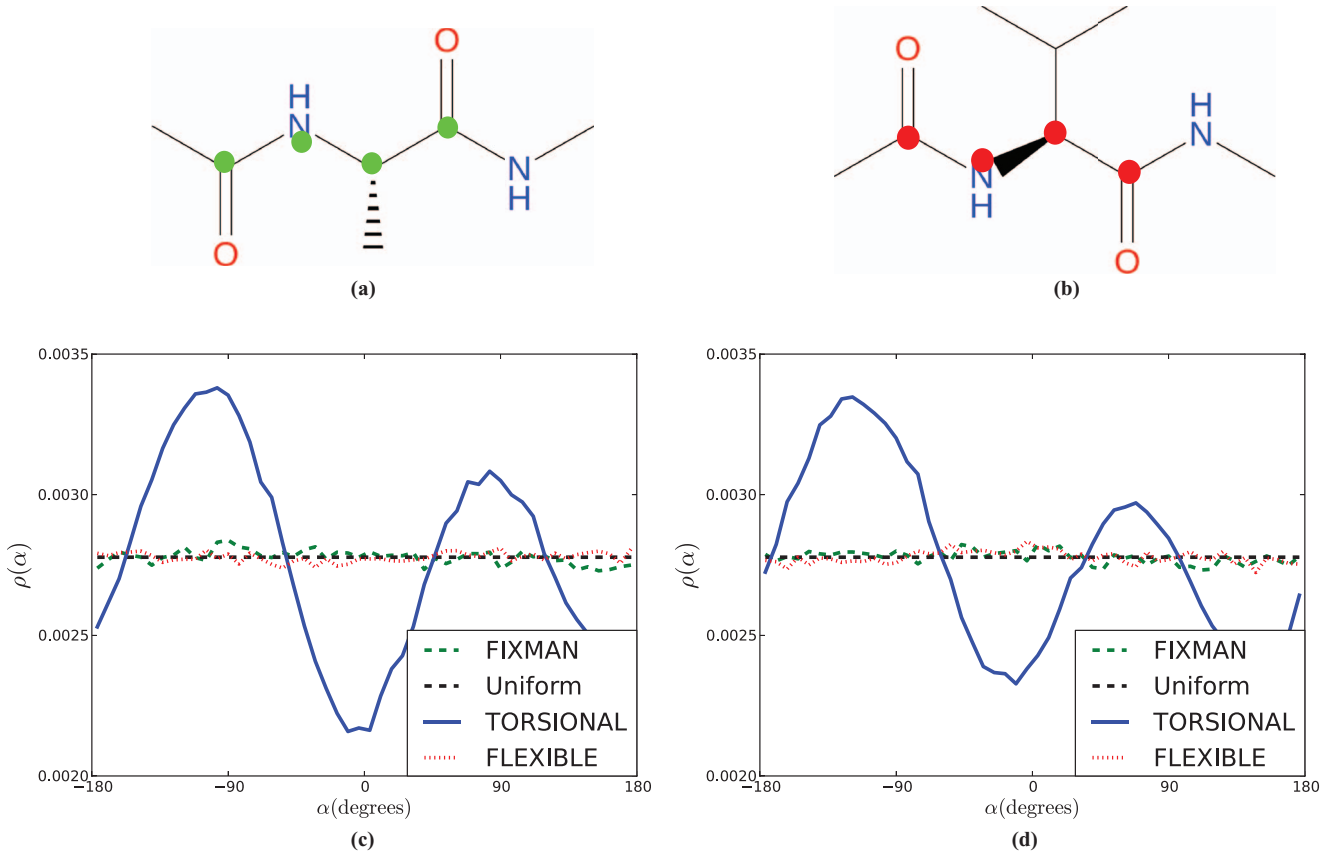


FIG. 7. Structures of (a) alanine dipeptide and (b) valine dipeptide. The atoms highlighted with the green circles in (a) and the red circles in (b) constitute the $C^1 - N^2 - C^2_\alpha - C^2$ torsion in the respective systems. Probability distributions for these torsions for alanine dipeptide and valine dipeptide, respectively, in (c) and (d). The application of the Fixman potential in the TORSIONAL model recovers the uniform pdf observed for the FLEXIBLE model. The bin size for the histograms is $d\alpha = 7.2^\circ$.

(NMR) structure from the protein data bank (pdb ID: 1UAO) as the starting structure for the simulations. For each of the TORSIONAL and FIXMAN sets, we ran four different simulations, each with a timestep of 1 fs and a simulation length of 100 ns, for a total of 300 ns of data. The coordinates and energy were recorded every 200 steps.

The torsion angle pdfs for the torsions labeled $C^1 - N^2 - C^2_\alpha - C^2$, $C^4 - N^5 - C^5_\alpha - C^5$, and $C^7 - C^8 - C^8_\alpha - C^8$ are shown in Figure 9. The atoms comprising these three torsions are shown as colored spheres in Figure 9(a). These three torsions are located at different parts of the peptide. We find that, in all three cases, the TORSIONAL simulations produce a bimodal probability distribution. With the introduction of the Fixman potential, however, the probability distributions become nearly flat as we expect for a fully flexible model.

For chignolin, the pdf deviation measures were $R_{tor} = 1.2 \times 10^{-3}$ and $R_{Fix} = 8.6 \times 10^{-5}$. Again, we find that, with the application of the Fixman potential, the pdfs for the FIXMAN model are very close to the uniform pdf.

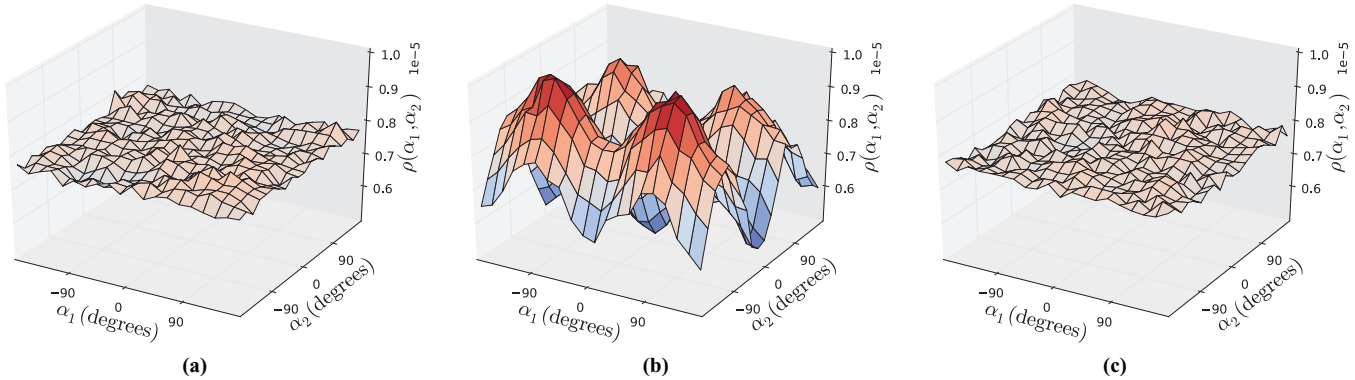


FIG. 8. Joint probability distribution function of the $C^1_\alpha - C^1 - N^2 - C^2_\alpha$ (α_1) and $C^1 - N^2 - C^2_\alpha - C^2$ (α_2) torsions for alanine dipeptide. (a) FLEXIBLE simulations, (b) TORSIONAL simulations, and (c) FIXMAN simulations. The bin size for each axis is $d\alpha = 18^\circ$.

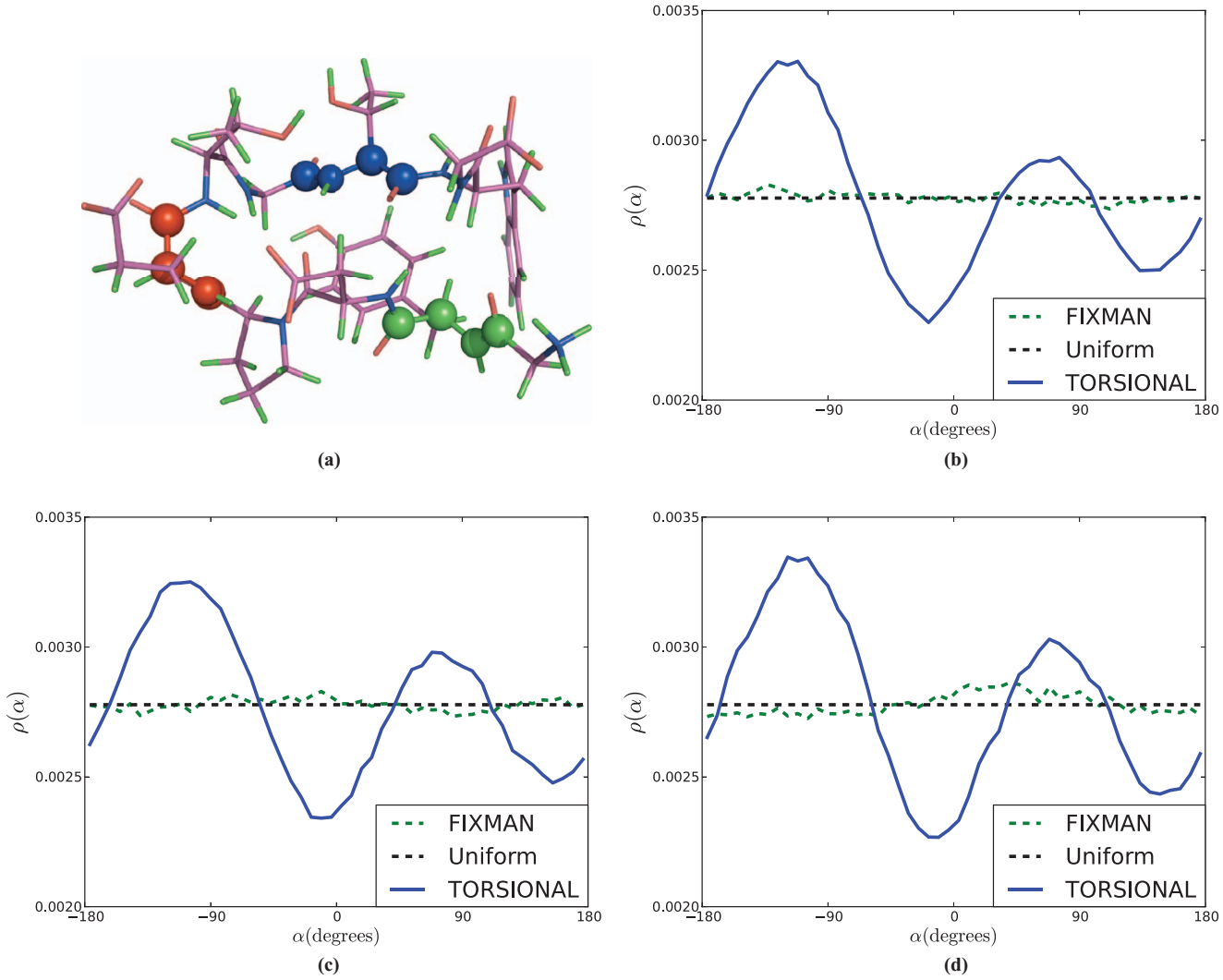


FIG. 9. The structure of the ten-residue β-hairpin protein, chignolin in (a). The green, red, and blue sphere atoms constitute the $C^1 - N^2 - C^2_\alpha - C^2$ torsion, the $C^4 - N^5 - C^5_\alpha - C^5$ torsion, and the $C^7 - N^8 - C^8_\alpha - C^8$ torsion, respectively. The application of the Fixman potential in the TORSIONAL model recovers the expected uniform pdf as see in (b) for the $C^1 - N^2 - C^2_\alpha - C^2$ torsion, (c) for the $C^4 - N^5 - C^5_\alpha - C^5$ torsion, and (d) for the $C^7 - N^8 - C^8_\alpha - C^8$ torsion. The bin size for the histograms is $d\alpha = 7.2^\circ$.

E. Computational cost

So far, we have focused on validating the correctness of the GNEIMO-Fixman method in evaluating the Fixman potential and torque, and on verifying that their use does indeed correct the statistical biases observed in constrained TORSIONAL simulations. Any practical use of the Fixman potential introduces another important consideration – the additional computational cost involved in calculating the Fixman potential and torque. Figure 10 shows that the computational cost for the GNEIMO dynamics solution scales linearly with system size, and that the introduction of the GNEIMO-Fixman method leads to a 24% increase in the computational cost. This increase is modest in large part because the GNEIMO-Fixman methods use several by-products from the GNEIMO method for solving the constrained dynamics. Most of the additional cost arises from the recursive evaluation of the $\Upsilon(i)$ matrix for each cluster.

We now examine the computational cost implications for regular all-atom molecular dynamics simulations. In our experience, each integration step for TORSIONAL simulations with all-atom force fields is typically twice as expensive (for a moderate sized system such as calmodulin with about 150 residues) as corresponding all-atom FLEXIBLE simulations due to the added complexity of solving the constrained equations of motion. However, this added cost is more than offset by the ability to use time steps of the order of 5–20 fs with TORSIONAL simulations compared with the 0.5–2 fs typical for FLEXIBLE simulations. The story changes only slightly with the inclusion of the Fixman potential. As observed from Figure 10, each time step in a GNEIMO-Fixman simulation with full force fields costs about 2.24 times (instead of 2 times for TORSIONAL) the cost of FLEXIBLE simulation time steps. The modest increase in computational cost is once again easily overcome by the larger time step size in GNEIMO-Fixman simulations making possible the

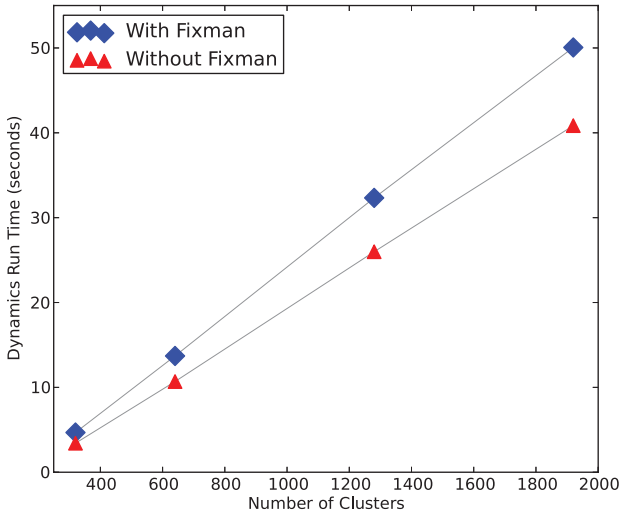


FIG. 10. The computational time for the dynamics evaluation versus the number of clusters for the GNEIMO method with and without the GNEIMO-Fixman method. Including GNEIMO-Fixman increases GNEIMO costs by 24%.

routine inclusion of the Fixman correction potential within constrained MD simulations.

IV. CONCLUSIONS

The Fixman compensating potential has been known to be critical in improving the accuracy of constrained MD simulations. However, the lack of practical algorithms to evaluate the Fixman potential has been a significant impediment for the development and application of constrained MD simulations. We have previously developed computationally efficient algorithms to solve the equations of motion for constrained MD simulations, and have demonstrated their application to various important problems in protein dynamics.[3–6,25,40–43] In this paper we have studied the SOA based GNEIMO-Fixman algorithm for calculating the Fixman compensating potential and torque.[27] We have run numerical experiments to cross-validate the GNEIMO-Fixman algorithm against prior published data for serial topology systems. We have further demonstrated the effectiveness of the application of the Fixman potential and torque in recovering the probability distribution function of single and joint torsion angles for simplistic models such as C4, C5, C11, C15 as well as realistic branched molecules – alanine and valine dipeptide and a ten amino acid peptide, chignolin.

To the best of our knowledge, this is the first demonstration of the use of the Fixman potential and torque for realistic branched molecular systems of arbitrary size. Our results confirm the validity of our computational algorithms, and also verified the modest computational cost for using them. Our results establish the viability of using the Fixman potential routinely in constrained MD simulations to remove undesired statistical biases from these simulations. In future work we expect to study the interplay between the Fixman potential and all-atom forcefields for large and realistic molecules. We also will investigate possibilities for improving absolute entropy and free energy computations using constrained dynamics simulations.

ACKNOWLEDGMENTS

This work was supported by funding from NIH Grant No. RO1GM082896. Part of the research described in this paper was performed at the Jet Propulsion Laboratory (JPL), California Institute of Technology, under contract with the National Aeronautics and Space Administration.

[1] J. P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen, J. Comput. Phys. **23**, 327 (1977).
[2] H. C. Andersen, J. Comput. Phys. **52**, 24 (1983).
[3] A. Jain, N. Vaidehi, and G. Rodriguez, J. Comput. Phys. **106**, 258 (1993).
[4] N. Vaidehi, A. Jain, and W. A. Goddard III, J. Phys. Chem. **100**, 10508 (1996).
[5] G. S. Balaraman, I.-H. Park, A. Jain, and N. Vaidehi, J. Phys. Chem. B **115**, 7588 (2011).
[6] I.-H. Park, V. K. Gangupomu, J. R. Wagner, A. Jain, and N. Vaidehi, J. Phys. Chem. B **116**, 2365 (2012).
[7] M. Fixman, Proc. Natl. Acad. Sci. U.S.A. **71**, 3050 (1974).
[8] M. Fixman, J. Chem. Phys. **69**, 1527 (1978).
[9] N. Go and H. A. Scheraga, J. Chem. Phys. **51**, 4751 (1969).
[10] N. Go and H. A. Scheraga, Macromolecules **9**, 535 (1976).
[11] W. F. V. Gunsteren and H. J. C. Berendsen, Mol. Phys. **34**, 1311 (1977).
[12] M. R. Pear and J. H. Weiner, J. Chem. Phys. **71**, 212 (1979).
[13] M. R. Pear and J. H. Weiner, J. Chem. Phys. **72**, 3939 (1980).
[14] W. F. V. Gunsteren, Mol. Phys. **40**, 1015 (1980).
[15] D. Perchak, J. Skolnick, and R. Yaris, Macromolecules **18**, 519 (1985).
[16] E. Helfand, J. Chem. Phys. **71**, 5000 (1979).
[17] D. Chandler and B. J. Berne, J. Chem. Phys. **71**, 5386 (1979).
[18] T. Mülders and W. Swegat, Mol. Phys. **94**, 395 (1998).
[19] M. Pasquali and D. C. Morse, J. Chem. Phys. **116**, 1834 (2002).
[20] A. Patriciu, G. S. Chirikjian, and R. V. Pappu, J. Chem. Phys. **121**, 12708 (2004).
[21] P. Echenique and I. Calvo, J. Comput. Chem. **27**, 1748–1755 (2006).
[22] P. Echenique, I. Calvo, and J. L. Alonso, J. Comput. Chem. **27**, 1733 (2006).
[23] J. Chen, W. Im, and C. L. Brooks III, J. Comput. Chem. **26**, 1565 (2005).
[24] V. Katritch, M. Totrov, and R. Abagyan, J. Comput. Chem. **24**, 254 (2003).
[25] J. R. Wagner, G. S. Balaraman, M. J. M. Niesen, A. B. Larsen, A. Jain, and N. Vaidehi, J. Comput. Chem. **34**, 904 (2013).
[26] A. K. Mazur and R. Abagyan, J. Biomol. Struct. Dyn. **6**, 815 (1989).
[27] A. Jain, J. Comput. Phys. **136**, 289 (1997).
[28] M. Karplus and J. N. Kushick, Macromolecules **14**, 325 (1981).
[29] R. Baron, W. F. V. Gunsteren, and P. H. Hünenberger, Trends Phys. Chem. **11**, 87 (2006).
[30] W. K. den Otter and W. J. Briels, J. Chem. Phys. **109**, 4139 (1998).
[31] M. Sprik and G. Ciccotti, J. Chem. Phys. **109**, 7737 (1998).
[32] W. K. den Otter, J. Chem. Theory Comput. **9**, 3861 (2013).
[33] S. Boresch and M. Karplus, J. Chem. Phys. **105**, 5145 (1996).
[34] V. Hnizdo and M. K. Gilson, Entropy **12**, 578 (2010).
[35] A. Jain, *Robot and Multibody Dynamics: Analysis and Algorithms* (Springer, 2011), p. 512.
[36] A. Graham, *Kronecker Products and Matrix Calculus: With Applications* (Halsted Press, 1981).
[37] R. D. Skeel and J. A. Izaguirre, Mol. Phys. **100**, 3885 (2002).
[38] S. Honda, K. Yamasaki, Y. Sawada, and H. Morii, Structure (London) **12**, 1507 (2004).
[39] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling, Proteins: Struct., Funct., Bioinf. **65**, 712 (2006).
[40] R. A. Bertsch, N. Vaidehi, S. I. Chan, and W. A. Goddard III, Proteins: Struct., Funct., Bioinf. **33**, 343 (1998).
[41] N. Vaidehi and W. A. Goddard III, J. Phys. Chem. **104**, 2375 (2000).
[42] V. K. Gangupomu, J. R. Wagner, I.-H. Park, A. Jain, and N. Vaidehi, Biophys. J. **104**, 1999 (2013).
[43] A. Jain, I.-H. Park, and N. Vaidehi, J. Chem. Theory Comput. **8**, 2581–2587 (2012).
[44] See supplementary material at http://dx.doi.org/10.1063/1.4851315 for Fixman torque contour plots for C5, and pdfs for torsion angles for C5 and a ten-bond chain.