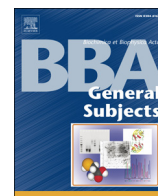




Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: [www.elsevier.com/locate/bbagen](http://www.elsevier.com/locate/bbagen)

# Correcting for the free energy costs of bond or angle constraints in molecular dynamics simulations<sup>☆</sup>

Gerhard König<sup>\*</sup>, Bernard R. Brooks

Laboratory of Computational Biology, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD 20892, United States

## ARTICLE INFO

### Article history:

Received 30 June 2014

Received in revised form 28 August 2014

Accepted 1 September 2014

Available online xxx

### Keywords:

Free energy calculation

Constraint correction

Molecular dynamics simulation

Normal mode analysis

Bennett's acceptance ratio

## ABSTRACT

**Background:** Free energy simulations are an important tool in the arsenal of computational biophysics, allowing the calculation of thermodynamic properties of binding or enzymatic reactions. This paper introduces methods to increase the accuracy and precision of free energy calculations by calculating the free energy costs of constraints during post-processing. The primary purpose of employing constraints for these free energy methods is to increase the phase space overlap between ensembles, which is required for accuracy and convergence.

**Methods:** The free energy costs of applying or removing constraints are calculated as additional explicit steps in the free energy cycle. The new techniques focus on hard degrees of freedom and use both gradients and Hessian estimation. Enthalpy, vibrational entropy, and Jacobian free energy terms are considered.

**Results:** We demonstrate the utility of this method with simple classical systems involving harmonic and anharmonic oscillators, four-atomic benchmark systems, an alchemical mutation of ethane to methanol, and free energy simulations between alanine and serine. The errors for the analytical test cases are all below 0.0007 kcal/mol, and the accuracy of the free energy results of ethane to methanol is improved from 0.15 to 0.04 kcal/mol. For the alanine to serine case, the phase space overlaps of the unconstrained simulations range between 0.15 and 0.9%. The introduction of constraints increases the overlap up to 2.05%. On average, the overlap increases by 94% relative to the unconstrained value and precision is doubled.

**Conclusions:** The approach reduces errors arising from constraints by about an order of magnitude. Free energy simulations benefit from the use of constraints through enhanced convergence and higher precision.

**General significance:** The primary utility of this approach is to calculate free energies for systems with disparate energy surfaces and bonded terms, especially in multi-scale molecular mechanics/quantum mechanics simulations. This article is part of a Special Issue entitled Recent developments of molecular dynamics.

© 2014 Published by Elsevier B.V.

## 1. Introduction

Constraints are used in most molecular dynamics simulations, since the maximum length of the time step for integrating the Newtonian equations of motion is restricted by the frequency of the fastest motions in the system. Imposing constraints that remove the associated rapid vibrational modes makes those degrees of freedom rigid. Thus, it is possible to use longer time steps without losing the conservation of energy or significantly distorting the desired ensemble. Thus, bond constraints can reduce the required computer time by a factor of three [1]. Furthermore, constraints can be a very valuable tool to improve the convergence and efficiency of free energy simulations, e.g., when using the Simplified Confinement [2] or Confine-and-Release methods [3].

Several algorithms are available to impose constraints in molecular dynamics simulations [4]. Among the most widely used methods is

the simple two body constraint SHAKE [5]. SHAKE constrains bonds and angles by finding forces that maintain the right geometry with the iterative Gauss–Seidel method. However, constraining too many bond distances that are coupled to each other through SHAKE is not practical due to the recursive nature of this method (e.g., ring structures). Prominent alternative approaches include the pairwise Rattle algorithm [6], which also includes the velocities of the atoms, or the LINCS algorithm [7]. Recently, the Shape rigid body integrator [8] has been developed that combines the accuracy of SHAKE with the possibility to use higher numbers and more types of constraints (i.e., rigid bodies of three or more centers, thus including bonds, angles and dihedrals).

Notably, imposing constraints can affect the outcome of the simulation by restricting the phase space of specified degrees of freedom to a choice selected by the user. Van Gunsteren reported quenching of dihedral angle transitions if both bonds and angles are constrained [9,10]. Furthermore, the efficiency of constrained simulations is lowered by angle constraints [1]. Toxvaerd pointed out that angle constraints change the trans-gauche transition rate of decane drastically [11]. Tobias and Brooks [12] determined that constraints shift the frequencies of the normal modes in biomolecules, but only in the region between

<sup>☆</sup> This article is part of a Special Issue entitled Recent developments of molecular dynamics.

<sup>\*</sup> Corresponding author at: NIH, NHLBI, Laboratory of Computational Biology, 5635 Fishers Lane T-900 Suite, Rockville, MD 20852.

E-mail address: [gerhard.koenig@nih.gov](mailto:gerhard.koenig@nih.gov) (G. König).

100 and 1400  $\text{cm}^{-1}$ . Similarly, Hinsen and Kneller found that the dynamics of polypeptides are slightly perturbed by constraints [13]. Recently, Echenique et al. [14] carefully analyzed and quantified the difference between unconstrained models and several types of constraints. Moreover, deviations between 0.2 and 0.5 kcal/mol were found in solvation free energy calculations when bond length constraints were employed [15]. Deviations of such magnitude highlight the need to properly account for free energy changes due to constraints in free energy simulations.

A number of publications have focused on methods for calculating constraint corrections [16–19]; good overviews are provided by Boresch and Karplus [20] or Wang and Hermans [21]. The most prominent approach for such corrections is the use of the average constraint force [22]. However, this approach only seems to be valid for bond length constraints [20]. The potential of mean force corrections of Pearlman and Kollman [23,24] compute the contribution to the free energy for systems with constrained bond terms. However, this requires additional simulations. Furthermore, most of the currently existing constraint corrections have been developed for free energy calculations with Thermodynamic Integration [25], and, therefore, cannot be used straightforwardly in the increasingly popular Bennett's acceptance ratio (BAR) [26], multistate-BAR (MBAR) [27], or Non-Boltzmann Bennett (NBB) [28,29] methods.

For free energy simulations with BAR, the primary utility of including constraints is to increase the phase space overlap between two distinct energy surfaces. Since the variance of the free energy estimate is directly linked to the phase space overlap, constraints allow free energy simulations to be more accurate and to converge more quickly (thus saving computational time). This is of particular utility for coupling disparate energy surfaces, such as QM/MM with MM, where small changes in bond lengths and angles can significantly decrease the necessary overlap needed for an accurate result. This increased accuracy and faster convergence rate offsets the costs associated with post-processing constraints. Here, we attempt to provide a simple framework to correct for the effect of constraints in free energy calculations with BAR by postprocessing the involved trajectories. In particular, we achieve this by explicitly calculating the free energy costs of adding and removing constraints as additional steps in the free energy cycle. We achieve this by employing a gradient calculation combined with normal mode analysis [30] to approximate the contributions to the partition function associated with the constrained degrees of freedom. Notably, this approach is also compatible with most quantum-packages and QM/MM [31–37].

Our approach is inspired by Gō and Scheraga [38], who showed that bond lengths and angles can be treated as functions of the dihedral angles. This implies that the optimal position of a bond or angle at each frame of a trajectory can be determined after the simulation has been conducted with an suboptimal configuration of that particular bond or angle. In particular, they regarded proteins to be a collection of independent normal modes. Some of the normal modes can be considered hard if their force constant is high enough, while all other normal modes are considered soft. The hard normal modes tend to be insensitive to the changes in conformation, so they can be regarded as functions of the soft variables (changing the bond length and bond angles after the conformation changes). This approximation is of course only valid close to the equilibrium conformation. Using this procedure, the hard modes can be considered to be frozen (constrained) during the dynamics, while the soft modes are treated classically to approximate the partition function. Thus, to account for the changes of the hard degrees of freedom, it is possible to use a post-processing step that changes the bond lengths and angles according to their (soft) environment, and the free energy decrease associated with that change.

The remainder of this paper is organized as follows. First, we outline how to calculate the free energy costs of bond and angle constraints (referred to as “constraint correction”), starting with a simple harmonic oscillator, and progressing to the analysis of trajectories in free energy simulations with multiple constraints (Section 2). Methodological

details of the benchmark systems and simulations are presented in Section 3. We then present the results for harmonic and anharmonic oscillators, four-atomic benchmark systems, water boxes of varying size, an alchemical mutation of ethane to methanol, and free energy simulations between alanine and serine (Section 4, see Fig. 1). We conclude with a short discussion concerning the practical advantages of using constraints in free energy calculations in Section 5. A description of the implementation of the constraint correction in CHARMM, as well as a discussion of alternative ways to approximate the Hessian can be found in the Supplementary material.

## 2. Theory

To illustrate the problem of calculating the free energy of releasing a constraint, we start with a simple classical harmonic oscillator with an internal coordinate  $q$ . For this system, the potential energy  $U$  is given by

$$U(\Delta q) = U_0 + \frac{K}{2} \Delta q^2, \quad (1)$$

where  $\Delta q$  is the deviation from the energy minimum,  $U_0$  is the zero point energy that depends on the reference point and  $K$  is the force constant of the bond or angle. Given this potential energy function, the partition function  $Z$  of the harmonic oscillator ( $Z^{h.o.}$ ) can be obtained by integrating over the single degree of freedom,  $q$ ,

$$Z^{h.o.} = \int e^{-\beta U(\Delta q)} d\Delta q = e^{-\beta U_0} \int e^{-\beta \frac{K}{2} \Delta q^2} d\Delta q. \quad (2)$$

The absolute free energy of the harmonic oscillator is then

$$G^{h.o.} = -\beta^{-1} \ln Z^{h.o.} = U_0 - \beta^{-1} \ln \int e^{-\beta \frac{K}{2} \Delta q^2} d\Delta q \quad (3)$$

$$= U_0 - \beta^{-1} \ln \sqrt{\frac{2\pi}{\beta K}}, \quad (4)$$

where we have made use of the Gaussian integral.

If we introduce a constraint in the harmonic oscillator, we completely remove the degree of freedom  $q$ , so the integral in the partition function becomes a single point, corresponding to a Dirac delta  $\delta(x)$ . Making use of the integral of the Dirac delta,  $\int \delta(x) dx = 1$ , the partition function of the constrained system becomes

$$Z_{cons}^{h.o.} = e^{-\beta U(\Delta q_{cons})}, \quad (5)$$

and, correspondingly, the free energy is reduced to

$$G_{cons}^{h.o.} = -\beta^{-1} \ln Z_{cons}^{h.o.} = U(\Delta q_{cons}). \quad (6)$$

Accordingly the free energy of imposing a constraint corresponds to

$$\Delta G_{cons}^{h.o.} = G_{cons}^{h.o.} - G^{h.o.} = U(\Delta q_{cons}) - U_0 + \beta^{-1} \ln \sqrt{\frac{2\pi}{\beta K}}. \quad (7)$$

For convenience, it is possible to divide the total energy into two contributions. First, the enthalpic contribution  $\Delta H$ ,

$$\Delta H^{h.o.} = U(\Delta q_{cons}) - U_0, \quad (8)$$

which is independent of the temperature. Notably,  $U(\Delta q_{cons})$  inherently contains  $U_0$ .

Download English Version:

<https://daneshyari.com/en/article/10800047>

Download Persian Version:

<https://daneshyari.com/article/10800047>

[Daneshyari.com](https://daneshyari.com)