Calculation of molecular free energies in classical potentials

Asaf Farhi\(^1\) and Bipin Singh\(^2\)

\(^1\) Raymond and Beverly Sackler School of Physics and Astronomy, Faculty of Exact Sciences, Tel Aviv University, IL-6997801 Tel Aviv, Israel

\(^2\) Center for Computational Natural Sciences and Bioinformatics (CCNSB), International Institute of Information Technology Hyderabad (IIIT-H), Gachibowli, Hyderabad, 500032, India

E-mail: asaffarhi@post.tau.ac.il

Supplementary material for this article is available online

Keywords: free energy, molecules, classical, non-interacting systems, molecular dynamics, Monte Carlo

Abstract

Free energies of molecules can be calculated by quantum chemistry computations or by normal mode classical calculations. However, the first can be computationally impractical for large molecules and the second is based on the assumption of harmonic dynamics. We present a novel, accurate and complete calculation of molecular free energies in standard classical potentials. In this method we transform the molecule by relaxing potential terms which depend on the coordinates of a group of atoms in that molecule and calculate the free energy difference associated with the transformation. Then, since the transformed molecule can be treated as non-interacting systems, the free energy associated with these atoms is analytically or numerically calculated. This two-step calculation can be applied to calculate free energies of molecules or free energy difference between (possibly large) molecules in a general environment. We demonstrate the method in free energy calculations for methanethiol and butane molecules in vacuum and solvent. We suggest the potential application of free energy calculation of chemical reactions in classical molecular simulations.

Introduction

Free energy calculations have a variety of applications which include binding, solvation and chemical reactions. For the applications of solvation and binding equilibrium and non-equilibrium methods are used. In equilibrium methods (alchemical transformations) a molecule is transformed into another through non-realistic intermediate systems. The average of the derivative of the interpolating Hamiltonian is calculated for each intermediate system and by numerically integrating, the free energy associated with the transformation is calculated. The free energies of these transformations are then used to calculate the free energy of the desired molecular process\([1–5]\). In non-equilibrium methods the work in the process of switching between the two Hamiltonians is measured\([6, 7]\). For the application of chemical reactions the free energy difference between the reactants and products needs to be calculated. Free energies of molecules can be calculated by quantum computations\([8]\) or by normal mode classical calculations\([9]\) (see Entropies section). While quantum mechanics (QM) calculations are highly accurate they are computationally very demanding for large molecules. Normal mode calculations can be very efficient but are based on harmonic dynamics\([10]\).

Molecular modeling includes covalent bond, bond angle, dihedral angle, improper dihedral angle, electrostatic and van der Waals (vdW) potentials (see\([11–13]\) and appendix A). Covalent bond, bond angle and dihedral angle terms depend on the coordinates of two, three and four nearest covalently linked atoms respectively. Electrostatic and vdW potentials are between every atom pair in the system.

In previous studies, the free energy associated with the covalent bond and bond angle terms in molecules and dihedral angle term in the context of restraints were directly calculated by assuming that they are quadratic and by using the rigid rotor approximation and the Herschbach, Johnston and Rapp technique\([14–16]\). However, the free energies associated with dihedral potential terms in the realistic non-quadratic potentials and with non-quadratic covalent bond and bond angle terms\([17]\) have not been calculated. In addition, it has not been
suggested to calculate molecular free energies of submolecules with coupled bond angle degrees of freedom such as the methyl group.

Here we present a direct free energy calculation which is exact and does not require the potential terms to be quadratic. We show that the partition function associated with a dihedral angle in molecules can be decoupled from the partition function of the rest of the system independently of the potential function (even though the dihedral potential term depends on coordinates of atoms in the system) and calculate analytically free energies of realistic dihedral terms. Moreover, we show that free energy associated with submolecules with coupled bond angle degrees of freedom such as the methyl group can be calculated accurately.

In section 1 we present the first step of the calculation in which we transform the molecule by relaxing all the interactions of a group of atoms in that molecule, except ones which depend on a relative spherical coordinate and calculate the free energy difference associated with the transformation. The relaxed potentials include vDW, electrostatic and all other types of potentials except bond stretching, bond angle and dihedral angle potentials. In section 2 we present the second step of the calculation in which the free energy associated with the remaining potentials (which depend on a relative spherical coordinate-bond stretching, bond angle and dihedral angle potentials) of this group of atoms is calculated analytically or numerically. This two-step calculation is applicable to any environment using the definition above, wherein for the case of a solvent in the first step we also relax the interactions of these atoms with the solvent molecules. In section 3 we demonstrate the method in free energy calculations for methanethiol and butane molecules. In section 4 we suggest the potential application of free energy calculation of chemical reactions. In section 5 we discuss the method.

Alchemical free energy transformation (as in section 1) is a standard procedure in molecular free energy calculations [12, 13, 18–20] and is usually used in the context of solvation and binding or in the context of chemical reactions in combination with QM calculations [8]. Here we apply it in the context of free energy calculations of molecules or submolecules in a general environment, where the second calculation is used to obtain the free energy difference between two molecules with submolecule in common. We combine the transformation with novel direct free energy calculations explained in section 2, followed by the novel application of free energy calculation of chemical reactions presented in section 3. Here we do not need to assume harmonic dynamics.

1. The transformation

We now explain the first step of the calculation in which we transform the molecule by relaxing potential terms that depend on the coordinates of a group of atoms in that molecule and calculate the free energy difference associated with the transformation using thermodynamic integration (TI) [3, 21, 22]. Alternatively, it can be performed using methods such as exponential averaging/ free energy perturbation (FEP) [23], Bennett acceptance ratio (BAR) [24] and weighted Histogram analysis method [25]. We denote the Hamiltonian with the terms that are removed in the transformation by \( H_0 \) and the Hamiltonian with the other terms by \( H_r \). We define the Hamiltonian \( H_0 \), following the definition above and the potentials in appendix A as the improper dihedral, electrostatic and vDW terms which depend on the coordinates of the group of selected atoms. The \( \lambda \) dependent Hamiltonian that transforms between the systems can be written as follows:

\[
H(\lambda) = \lambda H_0 + H_r,
\]

where \( \lambda \in [0, 1] \). The free energy difference between the original and transformed systems can be written as:

\[
\Delta F_{A \rightarrow A'} = -k_B T \int_0^1 \frac{d \ln Z(\beta, \lambda)}{d \lambda} d \lambda = \int_0^1 \langle H_0 \rangle d \lambda,
\]

where \( A' \) is the transformed system in which \( \lambda = 0 \). Simulating intermediate systems at \( \lambda \) values in the range \([0, 1]\) and calculating \( \langle H_0 \rangle \), followed by numerical integration, will enable us to calculate the free energy difference. It is noted that this calculation is valid for any environment (solvent, etc.) since the interactions of the group of atoms with the environment are also relaxed in the transformation. In addition, soft core potentials are used in the vDW and electrostatic interactions to avoid divergences of \( \langle H_0 \rangle \) at \( \lambda \to 0 \) [26, 27].

As an example we present in figure 1 three systems in the transformation of the molecule phenol with the selected atoms O and H1. The Hamiltonian \( H_r \) includes the vDW and electrostatic interactions of the atoms O and H1 with all the atoms in the system, and the improper dihedral term associated with \( \phi_{C1-C2-O-C3} \).

2. The remaining free energy difference

We now turn to explain how the free energy associated with the selected atoms can be calculated. We first switch to relative coordinates and then to spherical coordinates.
\[ \Omega = \prod_{i=1}^{n} dr_i = \prod_{i=2}^{k} dr_i \prod_{i=k+1}^{n} dr_i \]
\[ = dr'_1 \prod_{i=2}^{k} dr_i \prod_{i=k+1}^{n} r_i^2 \sin \theta_i dr_i d\theta_i d\phi_i, \]

where \( r_i \equiv r'_i - r'_{i-1} \), which will be chosen as the position of atoms relative to a covalently bound atom. \( k + 1 \) represents the first atom in the group of selected atoms (e.g. the atom O in the example of phenol).

The transformed molecule \( A' \) is first divided into elements of covalent bonds, bond junctions, dihedral terms and of more complex structures that include molecular rings. The dihedral potential term depends on the angle between two planes which is equal to the \( \phi \) angle in spherical coordinates. This correspondence can be understood by recalling the definition of the angle between two planes as the angle between the vectors in these planes that are perpendicular to the intersection line of these planes. Hence, the covalent bond, bond angle and dihedral angle terms can be expressed in terms of the spherical variables \( r \), \( \theta \) and \( \phi \) defined with respect to the relevant atoms.

For the example of phenol we can write the partition function of transformed phenol as follows

\[ Z_{A'} = \int e^{-\beta H_1} e^{-\beta H_{\text{sol}}} e^{-\beta H_{\text{OH}}} e^{-\beta H_{\text{other}}} \sin \theta_{\text{H1}} e^{-\beta H_{\text{H1}}} \sin \theta_{\text{H2}} e^{-\beta H_{\text{H2}}} \sin \theta_{\text{H3}} e^{-\beta H_{\text{H3}}} d\Omega G d\Omega_1 d\Omega_2 d\theta_{\text{H1}} d\phi_{\text{H1}}, \]

where the degrees of freedom of atom H and O are denoted by \( \eta_H \), \( \theta_{\text{H1}} \), \( \phi_{\text{H1}} \) and \( t_O, \theta_{\text{O}}, \phi_{\text{O}} \) respectively. All the degrees of freedom of all the other atoms in the system (molecule + solvent molecules in the case of explicit solvent) and \( \theta_{\text{O}}, \phi_{\text{O}} \) are denoted by \( \Omega \). We define \( \theta_{\text{H1}} \) and \( \phi_{\text{H1}} \) as follows: \( \theta_{\text{H1}} \equiv \theta_{\text{H1-0-C2}}, \phi_{\text{H1}} \equiv \phi_{\text{H1-0-C2-C1}} \).

Our goal in this example is to calculate the free energy associated with all the degrees of freedom except for the ones denoted by \( \Omega \). The free energy of the subsystem with the degrees of freedom denoted by \( \Omega \) can for example then cancel out with an identical subsystem when calculating the free energy difference between two molecules with a submolecule in common. Note that for example \( \phi_{\text{H1}} \) does depend on the coordinates of the C1 atom which is included in \( \Omega \) and therefore the decomposition of the partition functions is not trivial.

As will become clear, the integration in each element is independent of the others. Thus the integrals can be performed separately and then multiplied to yield the partition function and hence the free energy difference.

We write the free energy factors explicitly:

### 2.1. Covalent bond

The partition function of a covalent bond represented by a quadratic potential can be written as follows (for non-quadratic terms integrated analytically see appendix E)

\[ Z_c = \frac{1}{\beta} \int e^{\frac{\beta k}{2} \Delta^2} r^2 dr = \frac{\pi^2 (2d^2 \beta k_c + 1) (e (d \sqrt{\beta k_c}) + 1) + 2 \pi e^{-\beta k_c} d (d \sqrt{\beta k_c}) \beta k_c)}{\beta^2 (\beta k_c)^y}, \]

where \( l \) is an arbitrary length (\( l^2 \) cancels out in comparisons). To account for the bond-dissociation energy/bond energy one has to differentiate between the potential at \( r = 0 \) and \( k_c \rightarrow 0 \) (free atoms) either by introducing a factor \([28-30]\) or by using potentials such as the Morse potential (numerical integration) \([31]\).

---

**Figure 1.** Systems in the transformation of phenol at \( \lambda = 0, 0.5, 1 \). The semi transparent/ transparent atoms represent atoms with interactions partly/ totally removed.
2.2. Two bonds junctions

When considering the case of a linear/bent molecular shapes, it can be seen that when varying the bond angle, the rest of the molecule moves as a rigid body. Since the spherical coordinates representation includes three independent variables, varying a bond angle is decoupled from all the other degrees of freedom of the molecule. Hence the calculation of free energy factor associated with the bond angle potential term is straightforward. For a quadratic term we can write:

\[
Z_0 = \int e^{-i\theta_0} \sin \theta d\theta = \int e^{-\frac{i}{2}k_\theta(\theta-\theta_0)^2} \sin \theta d\theta = \frac{1}{2\sqrt{2k_\theta}} e^{-i\theta_0} \times
\]

\[
\int_{\pi}^{\pi} \left\{ 1 + i \theta_0 \sqrt{2k_\theta} + i \theta_0 \sqrt{2k_\theta} \right\}
\]

This expression is real for positive and real values of \(k_\theta\), \(\beta\) and \(\theta_0\).

When varying a dihedral angle, the potential term value depends on the orientation of first bond (which determines the axis from which the dihedral angle is measured). However, since the integration has to be performed over all the range \([0, 2\pi]\), integrating with respect to the \(\phi\) angle will yield a factor which is independent of the location of the first bond. Thus, the integration does not depend on the direction of the first bond and is straightforward. For the commonly used dihedral angle potential we get

\[
Z_4 = \int e^{-i\theta_0}(1 + \cos(n\phi - \phi_0)) d\phi = 2\pi e^{-i\theta_0} I_0(\sqrt{\frac{\theta_0}{k_\theta}}),
\]

where \(I_0\) is the modified Bessel function of the first kind. Note that this result does not depend on \(n\), which is an integer.

It can now be seen that the partition function of transformed phenol can be decomposed (exactly) into the partition function associated with the \(\Omega\) degrees of freedom and the covalent bond, bond angle and dihedral angle partition functions.

2.3. Three or more bonds junctions

Molecule shapes can include a monomer (covalent bond) that splits into more than one monomer. Such examples are the trigonal planar, tetrahedral trigonal pyramidal etc. These cases will necessitate numerical integration which can be performed using the Spherical law of cosines:

\[
\cos(\theta_{12}) = \cos(\theta_1) \cos(\theta_2) + \sin(\theta_1) \sin(\theta_2) \cos(\Delta \phi),
\]

where \(\theta_1, \theta_2\) denote the bond angles of two bonds with respect to the principal bond and \(\theta_{12}, \Delta \phi\) denote the bond angle and the difference in the angle between these two bonds respectively. Usually in these cases there is one dihedral term, which we denote by \(\phi_1\), that depends on the angle defined by one of the bonds, the principal bond and a previous bond. Since the integration over the other degrees of freedom yields a factor that is independent of the value of \(\phi_1\), the integrations are decoupled. Thus, the integration for the case of one monomer that splits into two can be written as follows:

\[
Z = \int e^{-\beta(V_{1i} + V_{2j} + V_{3k} + V_{4l})} \sin \theta_1 \sin \theta_2 \sin \theta_1 d\theta_1 d\phi_1 d\phi_2 = Z_4 Z_{3b},
\]

where \(Z_4\) is according to the definition in equation (3) and

\[
Z_{3b} = \int e^{-\frac{2}{\alpha}b^2(\theta_1 - \theta_0)^2} \sin(\theta_1) \sin(\theta_2) d\theta_1 d\theta_2.
\]

For the general case it can be written as follows:

\[
Z_{n_b} = \int \prod_{i} e^{-\frac{2}{\alpha}b^2(\theta_i - \theta_0)^2} \prod_{i>j} e^{-\frac{2}{\alpha}b^2(\theta_i - \theta_j)^2} \prod_{i} \sin \theta_i d\theta_i \prod_{i>j} d\theta_i,
\]

where \(\theta_0\) can be calculated from equation (4) and \(\phi_1\), which has to be substituted in \(\Delta \phi\) in this equation, is an arbitrary constant. In case there are energy terms that introduce complexity they can be relaxed in the transformation. For an explicit example with a bond junction see appendix B.

The free energy factors can be substituted in:

\[
\frac{\Delta F_{\alpha \to \alpha'}}{k_B T} = \left( \sum_i \ln Z_{\alpha_i} + \sum_i \ln Z_{\alpha'_{i}} + \sum_i \ln Z_{\alpha''_{i}} \right),
\]

to give the free energy difference, where \(\alpha'\) denotes a submolecule of the transformed molecule that is not necessarily realistic whose free energy will not be calculated.

In addition, the partition function of complex structures may be calculated numerically. In many cases the complex structures need to be compared to identical ones, eliminating the need for these calculations. Thus, we
can write in terms of the partition functions:

\[ Z \rightarrow Z_{\text{int}} \prod_{i=1}^{l} Z_{c_i} \prod_{j=1}^{m} Z_{d_j} \prod_{k=1}^{p} Z_{2b_k} \prod_{l=1}^{q} Z_{3b_l} \prod_{r=1}^{r} Z_{\text{complex}_r}, \]

where \( Z_{\text{int}} \) represents the partition function of the submolecule that is fully interacting and \( Z_{c_i} \) and \( Z_{d_j} \) represent the \( i \)th covalent bond and dihedral angle partition function respectively. \( Z_{2b_k} \) and \( Z_{3b_l} \) represent the \( k \)th two bond and three or more bond junctions respectively and \( Z_{\text{complex}_r} \) represents the \( r \)th complex structure partition function. The arrow represents the transformation \( \lambda = 1 \rightarrow 0 \).

We can also choose the group of selected atoms as all the atoms of the molecule, and calculate the free energy in a similar manner. The partition functions in this case can be written as follows:

\[ Z \rightarrow Z_{\text{fp}} \prod_{i=1}^{l} Z_{c_i} \prod_{j=1}^{m} Z_{d_j} \prod_{k=1}^{p} Z_{2b_k} \prod_{l=1}^{q} Z_{3b_l} \prod_{r=1}^{r} Z_{\text{complex}_r}, \]

where \( Z_{\text{fp}} \) denotes the partition function of a free particle.

In figure 2 we present the free energy decomposition of transformed phenol (A'). The degrees of freedom associated with each system are written on top. \( F \) denotes the free energy and \( A'' \) which is the reference (possibly imaginary) submolecule, is presented in the first term in the decomposition. It should be noted that the \( \theta_0 \) and \( \phi_0 \) degrees of freedom are associated with the submolecule \( A'' \) and the \( \eta_3 \) degree of freedom is associated with the second system in the decomposition. The potential terms that depend on the coordinates of the O and H atoms in the five decomposed systems are: 1 : \( V_b(\theta_{C1-C2-O}), V_b(\theta_{C2-C3-O}), V_b(\phi_{C3-C4-C2-O}) \), 2 : \( V_r(\eta_3), V_r(\eta_3) \), 3 : \( V_r(\eta_3), V_r(\eta_3) \), 4 : \( V_b(\theta_{C2-O-H}) \), 5 : \( V_b(\phi_{C1-C2-O-H}) \). The reason that the degrees of freedom \( \theta_0 \) and \( \phi_0 \) are associated with the first system in the decomposition is that the corresponding potential terms usually do not depend on the atom (e.g. O). The free energies of the second and third systems are calculated according to equation (1) and the free energies of the fourth and fifth systems are calculated according to equations (2) and (3) respectively.

3. Demonstrations

The covalent bond free energy calculation has been demonstrated and compared with numerical integration for two molecules of two atoms in a spherical potential (see appendix C). In addition, the free energies for methanethiol (CH₃SH) and butane (C₄H₁₀) molecules were calculated in vacuum and water environments. In the first step we removed the vdW and Coulomb interactions in a transformation and calculated the free energy difference. We then performed the second step of the calculation using the method and compared to the results obtained by MD simulations. It should be noted that C₄H₁₀ has a very low solubility in water and our goal was to calculate free energies when the molecule is in an environment (other than vacuum). In the second step of the calculation Helmholtz free energy (denoted by \( F \)) is calculated. Since we expect \( \Delta F_{\text{method}} = \Delta F_{\text{MD}} \) and usually \( \Delta F \approx \Delta G \) (\( G \) is Gibbs free energy) we compared \( \Delta F_{\text{method}} \) to \( \Delta F_{\text{MD}} \) obtained in constant pressure and temperature (NPT) MD simulations. In case that the difference between the results was larger than 0.15 kcal mol⁻¹ we also compared the results to \( \Delta F_{\text{MD}} \) obtained by performing constant temperature and volume (NVT) MD simulations to show that it originates from a difference between \( F \) and \( G \). This occurred only once (0.485 kcal mol⁻¹) and when we compared \( \Delta F_{\text{method}} \) to \( \Delta F_{\text{MD}} \) the free energy difference was 0.038 kcal mol⁻¹.
3.1. Free energy calculations for methanethiol (CH₃SH)

In the calculations in vacuum due to the fact that the only existing potentials are bond stretching, bond angle and dihedral angle potentials we performed only the second step of the calculation (in solvent both steps were performed). We compared the results obtained according to the equations in section 2 to those obtained by MD simulations in vacuum and water environments and compared the running times.

We performed the calculations at $T = 300$ K and used the GROMOS ATB force field in which each molecule is individually parametrized [17] (see figure 3). The force field parametrization was according to bond stretching and bond angle potentials which are slightly different from the standard ones and we performed our calculations accordingly. We give the equations for the potentials here for completeness:

$$V_\text{c} = \frac{1}{2} c \left( r^2 - r_0^2 \right)^2, \quad V_\text{b} = \frac{1}{2} k_\text{b} \left( \cos \theta - \cos \theta_0 \right)^2, \quad V_\text{d} \left( \phi_{\text{dih}} \right) = k_\text{d} \left( 1 + \cos (n \phi - \phi_0) \right).$$

Our goal is to demonstrate the free energy calculations associated with the potentials which depend on the relative spherical coordinates presented in section 2. This calculation is for atoms which do not interact via vdW and electrostatic interactions, after that these interactions have been removed in a transformation. In vacuum state since only atoms that are distanced four covalent bonds or more interact via vdW and electrostatic interactions, for methanethiol these interactions do not exist. Thus the free energy of the molecule can be decomposed into the free energies associated with the bond stretching terms, the bond angle term, the dihedral angle term and the CH₃ bond junction. When the molecule is in a solvent we first removed the vdW and Coulomb interactions of the atoms involved in a transformation (first step of the calculation) and then calculated the free energies associated with the bond and dihedral angle terms (second step of the calculation). We compared the results and the running times in the second step of the calculation to the ones obtained by MD simulations in vacuum and water environments.

3.1.1. Results

We calculated the free energies associated with removing the the vdW and Coulomb interactions of atom 1 and atoms 4–6 in water environment and obtained

$$\Delta G_{\text{vdW} \rightarrow \text{Coulomb} \ 1 \ \text{MD water}} = -7.266 \pm 0.019 \text{ kcal mol}^{-1},$$
$$\Delta G_{\text{vdW} \rightarrow \text{Coulomb} \ 4–6 \ \text{MD water}} = -14.91 \pm 0.019 \text{ kcal mol}^{-1}.$$ 

After performing these transformations we could calculate the free energy associated with the $\phi_{1234}$ and $\theta_{123}$ and the CH₃ bond junction respectively. There was no need to perform these transformations in vacuum ($\Delta G = 0$).

The free energy contributions in the second step were calculated by substituting the ATB bond stretching and bond angle potentials in the corresponding partition functions presented in section 2 and integrating (see appendix E) and by using the partition functions for the dihedral potential derived in equation (3). The results for the free energies associated with all the components of the molecule defined in section 2 are presented below:

$$E_{12} = 5.33 \text{ kcal mol}^{-1}, \quad E_{23} = 5 \text{ kcal mol}^{-1},$$
$$E_{34} = E_{35} = E_{36} = 5.54 \text{ kcal mol}^{-1}, \quad E_{b \ 123} = 0.97 \text{ kcal mol}^{-1},$$
$$E_{1234} = -0.58 \text{ kcal mol}^{-1}, \quad E_{\text{CH}_3} = 5.39 \text{ kcal mol}^{-1},$$

where we have used $F = -k_B T \ln Z$ and $E_{\text{CH}_3}$ is the free energy associated with the CH₃ bond junction. These results are valid both for vacuum and solvent.
The free energy associated with the bond and dihedral angle terms can be calculated using MD simulations by removing terms in a transformation and subtracting the free energy of the corresponding element without the potentials (this is simply $F = -k_B T \ln \Omega$, where $\Omega$ is phase space volume). The free energy associated with the bond stretching terms on the other hand are less practical to calculate by MD simulations since removing these terms will result in atoms which are unbound to the molecule.

To verify the analysis and the calculations we first performed 21 MD simulations of intermediate systems which interpolate between the molecule with the dihedral term and without the dihedral term associated with $\phi_{1234}$ both in vacuum/dilute gas and in solvent. The calculated free energy differences associated with these transformations were

$$\Delta G_d \text{MD vacuum} = -0.5095 \pm 0.0023 \text{ kcal mol}^{-1},$$

$$\Delta G_d \text{MD water} = -0.5095 \pm 0.0023 \text{ kcal mol}^{-1}.$$ 

These results can be compared to

$$\Delta F_d \text{equation(3)} = -k_B T \ln \Omega_d - F_d = -k_B T \ln 2\pi + k_B T \ln 2 e^{-\beta k_d} I_0(\beta k_d)$$

$$= k_B T \ln [e^{-\beta k_d} I_0(\beta k_d)] = -0.507982 \text{ kcal mol}^{-1},$$

obtained by simply substituting $k_d$ and $\beta$ in this equation which is based on equation (3). This validates both the free energy calculation and the partition function decomposition in equation (3).

We then performed transformations in which the bond angle term associated with $\theta_{123}$ is relaxed. We performed the transformation both using the bond angle potential of the force field and by treating the parameters as if they were given for the standard quadratic bond angle term. The results obtained from the MD simulations in vacuum and water for the quadratic potential are the following:

$$\Delta G_b \text{MD vacuum quadratic} = -1.39 \pm 0.0023 \text{ kcal mol}^{-1},$$

$$\Delta G_b \text{MD water quadratic} = -1.39 \pm 0.01 \text{ kcal mol}^{-1},$$

These results are in agreement with the calculation according to equation (2):

$$\Delta F_b \text{ Eq. (2)} = -k_B T \ln 2 - F_b \text{ quadratic} = -1.38152 \text{ kcal mol}^{-1}.$$ 

The results obtained from MD simulations in vacuum and water for the ATB bond angle potential are the following:

$$\Delta G_b \text{MD vacuum cos} = -1.388 \pm 0.0023 \text{ kcal mol}^{-1},$$

$$\Delta G_b \text{MD water cos} = -1.421 \pm 0.024 \text{ kcal mol}^{-1}.$$ 

These results are in agreement with the calculation according to the same potential:

$$\Delta F_b \text{cos} = -k_B T \ln 2 - F_b \text{cos} = -1.38152 \text{ kcal mol}^{-1}.$$ 

We then performed the transformation of the CH$_3$ bond junction in which the bond angle terms associated with the angles $\theta_{234}$, $\theta_{235}$, $\theta_{236}$, $\theta_{335}$, $\theta_{336}$ and $\theta_{336}$ were relaxed. The results obtained from MD simulations for the quadratic potential are the following:

$$\Delta G_{\text{CH}_3} \text{MD vacuum quadratic} = -9.019 \pm 0.007 \text{ kcal mol}^{-1},$$

$$\Delta G_{\text{CH}_3} \text{MD water quadratic} = -9.038 \pm 0.012 \text{ kcal mol}^{-1}.$$ 

These results are in agreement with result obtained by numerically integrating according to equation (7)

$$\Delta F_{\text{CH}_3} \text{equation (7)} = -k_B T [\ln (32\pi^2) - \ln (Z_{\text{CH}_3 \text{quadratic}})] = -9.02383 \pm 0.0021 \text{ kcal mol}^{-1}.$$ 

We also calculated the free energy difference associated with the CH$_3$ transformation using the ATB force field bond angle potential. The results obtained from MD simulations are the following:

$$\Delta G_{\text{CH}_3} \text{MD vacuum cos} = -8.83 \pm 0.0047 \text{ kcal mol}^{-1},$$

$$\Delta G_{\text{CH}_3} \text{MD water cos} = -8.82 \pm 0.014 \text{ kcal mol}^{-1}.$$ 

These results are in agreement with the result obtained by the numerical integration

$$\Delta F_{\text{CH}_3} \text{cos} = -8.804 \pm 0.00365 \text{ kcal mol}^{-1}.$$ 

The running times of the calculations in the first step were 0 s for vacuum (no transformation) and 1.27 × 21 = 26.7 h and 1.38 × 21 = 29 h for removing vdW and Coulomb terms of atom 1 and atoms 4–6, respectively, in water environment (Six-Core AMD Opteron(tm) processor 2427, 2.19 Ghz).

The calculations of the free energies associated with the bond stretching, bond angle and dihedral angle terms according to section 2 were immediate $\sim 10^{-8}$–$10^{-7}$ with eight digits of precision as they are merely substitutions in equations (1)–(3) (for the ATB force field potentials equations (1) and (2) were slightly different). MD simulations in a total running time of 3.5 × 21 = 73.5 min and 1 × 21 = 21 h (Six-Core AMD Opteron (tm) processor 2427, 2.19 Ghz) for vacuum and water environments respectively, yielded a precision of 2–3 digits, that may be lower in practice due to the spacing between intermediates.
The running time of the numerical integration in the free energy calculation of the CH$_3$ bond junction according to equation (7) was 2.34 sec on a 2 GHz dual core intel processor. The total running time of the MD simulations was $3.5 \times 21 = 73.5$ min and $9 \times 21 = 189$ h (Six-Core AMD Opteron(tm) processor 2427, 2.19 Ghz) for vacuum and water environments respectively.

In conclusion, the calculations according to equations (2)–(7) were in agreement with the results obtained by performing transformations using MD simulations. The computations using these equations (second step of the calculation) were much faster for the same computational resources. This originates from the fact that using the method we perform calculations on subsystems with a small number of degrees of freedom, compared with the MD calculations which include all the degrees of freedom.

When using the method the dominant running times are of the first step of the calculation, which are reported above. Note that calculating free energies of a covalent bond can be performed by a direct calculation (see for example equations (1) and (E1)) and removing a covalent bond term in MD transformation will involve breaking the covalent bond and difficult sampling. The same applies to calculating free energies of molecular rings, which may be performed only by a direct calculation. Thus, the complete free energy calculation is difficult or impractical to perform using only MD transformations. In addition, the running times of the method were compared in some cases (e.g. bond junctions) to the running times of novel MD transformations.

3.2. Free energy calculations for butane (C$_4$H$_{10}$)

We performed free energy calculations for butane in vacuum and water at $T = 300$ K using the method and by performing MD simulations. In figure 4 we present butane molecule with the ATB bond and dihedral angle potential (force field) constants [17]. The dihedral terms in this molecule can have high energy barriers, which lead to difficult sampling. This can be overcome by using techniques such as parallel tempering/ replica exchange [33–35], umbrella sampling [25] or transforming variables without introducing a bias [36, 37]. In our simulations and for the purpose of our calculations we could perform the sampling without using these techniques.

We first removed the potentials which do not depend on the spherical coordinates of atoms 1–4 in vacuum and in water environments. The Gibbs free energies associated with these transformations were the following:

\[
\Delta G_{\text{vdW} - \text{Coulomb MD vacuum}} = -3.368 \pm 0.0024 \text{ kcal mol}^{-1},
\]

\[
\Delta G_{\text{vdW} - \text{Coulomb MD water}} = -3.348 \pm 0.0024 \text{ kcal mol}^{-1}.
\]

We then calculated the free energies associated with the four bonds junction which involves atoms 1–5 and with the dihedral angle $\phi_{1-2-5-8}$. These free energies were calculated using the method and by transforming the molecule and performing MD simulations in vacuum and water environments.
We obtained the following results:

\[ \Delta G_{C_{3}H_{8}, \text{MD vacuum}} = -8.442 \pm 0.0024 \text{ kcal mol}^{-1}, \]
\[ \Delta G_{C_{3}H_{8}, \text{MD water}} = -8.535 \pm 0.0035 \text{ kcal mol}^{-1}, \]
\[ \Delta F_{C_{3}H_{8}, \text{method}} = -8.4089 \pm 0.0051 \text{ kcal mol}^{-1}, \]
\[ \Delta G_{d 1 \rightarrow 2 \rightarrow 5 \rightarrow 8, \text{MD vacuum}} = -0.777 \pm 0.00 \text{ kcal mol}^{-1}, \]
\[ \Delta G_{d 1 \rightarrow 2 \rightarrow 5 \rightarrow 8, \text{MD water}} = -0.837 \pm 0.0024 \text{ kcal mol}^{-1}, \]
\[ \Delta F_{d 1 \rightarrow 2 \rightarrow 5 \rightarrow 8, \text{method}} = -0.762 \text{ kcal mol}^{-1}. \]

We then removed in a transformation the potentials which do not depend on the spherical coordinates of atoms 5–7 in vacuum and water environments and obtained:

\[ \Delta G_{vdW-Coulomb, \text{MD vacuum}} = 7.311 \pm 0.0072 \text{ kcal mol}^{-1}, \]
\[ \Delta G_{vdW-Coulomb, \text{MD water}} = 14.166 \pm 0.012 \text{ kcal mol}^{-1}. \]

We calculated the free energies associated with the four bonds junction which involves atoms 2, 5–8 and with the dihedral angle \( \phi_{2 \rightarrow 5 \rightarrow 8 \rightarrow 11} \). These free energies were calculated using the method and by transforming the molecule and performing MD simulations in vacuum and water. We obtained the following results:

\[ \Delta G_{C_{3}H_{8}, \text{MD vacuum}} = 9.582 \pm 0.0096 \text{ kcal mol}^{-1}, \]
\[ \Delta G_{C_{3}H_{8}, \text{MD water}} = 9.981 \pm 0.0143 \text{ kcal mol}^{-1}, \]
\[ \Delta F_{C_{3}H_{8}, \text{method}} = 9.4958 \pm 0.0064 \text{ kcal mol}^{-1}, \]
\[ \Delta G_{d 2 \rightarrow 5 \rightarrow 8 \rightarrow 11, \text{MD vacuum}} = -0.762 \pm 0.0024 \text{ kcal mol}^{-1}, \]
\[ \Delta G_{d 2 \rightarrow 5 \rightarrow 8 \rightarrow 11, \text{MD water}} = -0.770 \pm 0.00 \text{ kcal mol}^{-1}, \]
\[ \Delta F_{d 2 \rightarrow 5 \rightarrow 8 \rightarrow 11, \text{method}} = -0.762 \text{ kcal mol}^{-1}, \]

where we have calculated \( \Delta F_{vdW-Coulomb, \text{MD water}} \) to show that the difference between the results originates from the difference between \( F \) and \( G \).

### 4. Potential application

Here we suggest the potential application of calculating free energies of chemical reactions using classical molecular simulations followed by analytic or numerical calculations. We will calculate the free energies of only small parts of the molecules, possibly in solvent, to obtain the free energy of the chemical reaction which can involve large molecules.

To obtain the equilibrium constant of a chemical reaction the standard Gibbs free energies (which are similar to Helmholtz free energies in many cases [32]) are usually calculated. The standard state is the hypothetical state with the standard state concentration but exhibiting infinite-dilution behavior (the interactions between e.g. the solute molecules are negligible). Hence, when we are interested in the standard free energy of a substance a single copy of the molecule of interest can be simulated either in vacuum or solvent environments.

The idea is to transform the reactants and the products, between which the free energy difference is calculated, into molecules that have the same partition functions up to factors that can be calculated. First, we match reactant molecules with product molecules that have submolecules in common if possible so that the free energy of these submolecules will cancel out. Then, we transform the molecules by relaxing potential terms of the atoms that are different between the molecules and calculate the free energy differences associated with each transformation. Finally, the free energy factors associated with the different atoms are calculated and we can deduce the free energy of the chemical reaction. For example, given the chemical reaction

\[ A + B \rightarrow C + D, \]

we can match, if possible, molecules \( A \), \( B \) to molecules \( C \), \( D \). Then we transform each of the molecules to \( A', B', C' \) and \( D' \) and calculate the free energy differences \( \Delta F_{A \rightarrow A'}, \Delta F_{B \rightarrow B'}, \Delta F_{C \rightarrow C'} \) and \( \Delta F_{D \rightarrow D'} \) respectively. Finally, we calculate the free energy factors \( \Delta F_{A \rightarrow A''}, \Delta F_{B' \rightarrow B''}, \Delta F_{C' \rightarrow C''} \) and \( \Delta F_{D' \rightarrow D''} \) as previously explained.

Thus, for the standard state free energy we get:

\[ \Delta F_{A + B \rightarrow C + D} = \Delta F_{A \rightarrow A'} + \Delta F_{B \rightarrow B'} - \Delta F_{C \rightarrow C'} - \Delta F_{D \rightarrow D'} + \Delta F_{A' \rightarrow A''} + \Delta F_{B' \rightarrow B''} - \Delta F_{C' \rightarrow C''} - \Delta F_{D' \rightarrow D''}. \]

\[ \text{New J. Phys. 18 (2016) 023039 A Farhi and B Singh} \]
An example for such a calculation is given in Figure 5. The molecules that participate in the chemical reaction and the transformed molecules are presented on top and bottom, respectively. In this case molecule B is matched with C and $F_B^c = F_C^c$, $F_A^r$ and $F_D^r$, which are the free energy of a free particle, are equal.

5. Discussion

To apply the method, the force field should contain potentials that depend on a relative spherical coordinate (bond stretching, bond angle and dihedral angle), which is the case in most force fields. In addition in order for the free energies associated with bond stretching potentials to fully match experimental results, the force field should account for bond energies. An example for a force field which satisfies both requirements is MM3 [30]. Free energies associated with bond stretching, bond angle, dihedral angle and bond junctions can be readily calculated. It is assumed that free energy associated with an aromatic ring can be calculated. The free energy calculation of more complicated molecular rings needs to be assessed and treating them as a non-interacting system is expected to simplify their free energy calculations (also in other direct calculation methods). Fortunately, molecular rings are in many cases stable so the free energy associated with them often does not need to be calculated. Currently, the software which supports MM3 force field enables to perform alchemical transformations (first step of the calculation) only for single atoms. We intend to extend the code and demonstrate the method in a free energy calculation of a chemical reaction in our next project.

We presented an accurate and complete method for calculating molecular free energies in classical potentials. We suggested the potential application of free energy calculation of chemical reactions in classical potentials. This application can find use in organic chemistry, biochemistry and drug discovery.

Acknowledgments

D J Bergman, A Nitzan, O Hod, D Harries and G Falkovich are acknowledged for the useful comments. H Suchowski is acknowledged for funding the publication.

Appendix A. Molecular potentials

Here we briefly describe the molecular potentials.

The covalent bond between two atoms is modeled with the following potential:

$$V_c = k_c (r - d)^2,$$  \hspace{1cm} (A1)

where $r$ is the distance between the atoms. The bond angle term between three atoms is usually modeled with the following potential:
The commonly used dihedral angles potential is of the following type:

\[ V_{\theta}(\theta) = \frac{1}{2} k_\theta (\theta - \theta_0)^2. \]  

The improper dihedral term that is used to enforce planarity is defined as follows:

\[ V(\phi_{ijkl}) = k_{ijkl} (1 + \cos(n\phi_{ijkl} - \phi)), \]

where \( \phi_{ijkl} \) is defined by the angle between the plane formed by the \( ijk \) atoms and the plane formed by the \( jkl \) atoms. The coefficients for these interactions are determined by the atom charge and vdW coefficients of both atoms.

### Appendix B. An example with a bond junction

To illustrate a case with a bond junction we write the partition function of transformed Toluene (see figure 7), where as in the case of phenol, the Hamiltonian \( H_t \) includes the improper dihedral, electrostatic and vdW potential terms which depend on the coordinates of the group of selected atoms.

\[
Z_{Toluene'} = \int e^{-\beta H_t} \prod_{i=1}^{3} e^{-\beta H_{\text{elec}}(\Omega_i)} e^{-\beta H_{\text{vdW}}(\theta_i)} e^{-\beta H_{\text{covalent}}(\theta_{ij} - \theta_{kl}) - \beta H_{\text{covalent}}(\theta_{ik} - \theta_{jm}) - \beta H_{\text{covalent}}(\theta_{im} - \theta_{jn})} \\
\times e^{-\beta H_{\text{vdW}}(\phi_{ijkl})} d\Omega_i d\theta_i d\phi_{ijkl}.
\]

It can be readily seen that the partition function can be decomposed into the one associated with the \( \Omega \) degrees of freedom and the covalent bond, dihedral angle and three bonds junction partition functions.

### Appendix C. Demonstration of the method—bond stretching potential

In order to demonstrate the method it was used with all its ingredients in MC simulation to calculate the free energy difference between the systems \( A \) and \( B \) and this value was compared to the one calculated by numerical integration. Then, in order to assess the efficiency of the method, the free energy difference between the systems was calculated using TI combined with Hamiltonian parallel tempering (for sampling rugged energy landscape).

The compared systems are composed of a molecule of two atoms in which one atom is fixed to the origin and the second one is bound to the first by a covalent bond. The second atom in each system is in a \( \theta \) dependent potential (in spherical coordinates), containing \( \theta^{-12} \) term to represent the vdW repulsive term used in molecular modeling. The potential barrier was chosen to be of typical value of systems with tens of atoms, having rugged energy landscape. The covalent bond length difference was chosen to represent systems with few different atom lengths (the values of the pairs of spring constant and covalent bond length were taken from molecular...
simulation software). The following potential and parameters were used:

\[
V_A = \frac{1}{2} k_A (\eta_2 - r_{eq_A})^2 - 5.5 k_B T \cdot \left[ \sin(4\theta) - \frac{10^{-8}}{\left(\theta - \frac{3\pi}{8}\right)^2} \right],
\]

\[
V_B = \frac{1}{2} k_B (\eta_2 - r_{eq_B})^2 + 5 k_B T \cdot \left[ \sin(4\theta) + \frac{10^{-8}}{\left(\theta - \frac{5\pi}{8}\right)^2} \right],
\]

\[
r_{eq_A} = 2.1 \text{ Å}, \quad r_{eq_B} = 1.3 \text{ Å}, \quad k_{eq_A} = 123 \text{ kcal mol}^{-1} \text{ Å}^{-2}, \quad k_{eq_B} = 428 \text{ kcal mol}^{-1} \text{ Å}^{-2}.
\]

Here we used the partition function of two atoms with a covalent bond term represented by a quadratic potential that can be written as follows:

\[
Z(\beta) = \frac{1}{\beta} \int_{-\infty}^{\infty} e^{-\beta k_c l^2 - \phi} d\beta = \frac{4\pi}{\beta} \int_{0}^{\infty} e^{-\beta k_c l^2} r^2 r^2 dr,
\]

\[
= \pi \left[ (2d^2/\beta k_c + 1) (\text{erf}(d/\sqrt{\beta k_c}) + 1) + 2\sqrt{\pi} e^{-\beta k_c d^2} d/\sqrt{\beta k_c} \right],
\]

where \( l \) is an arbitrary length. This was used to calculate the free energy difference between the systems with the coupling terms relaxed.

The comparison of the method to the numerical integration yielded exactly the same results.

Appendix D. Simulation protocol

We have used Gromacs simulation package [38] and Gromos53a7 force field parameters (from ATB server [17]) along with spc water model for the simulations. For all the simulations a time step of 2 fs was used. The cubic or dodecahedron box with a minimum distance of 1 nm between the solute and the box edge was considered during the simulations. After minimization, the systems were equilibrated (only for simulations in water) in NVT and NPT ensembles for 100 ps. All the production simulations were performed under NPT ensemble, except the bond angle transformation for C4H9 in water, where NVT ensemble was also used due to the difference between \( \Delta G_{MD} \) and \( \Delta G_{method} \) in both vacuum and water, the systems were simulated at each of the 21 equi-spaced intermediate states including the initial (\( \lambda = 0 \)) and final states (\( \lambda = 1 \)). The length of simulation was 20 ns in vacuum, 1 ns for CH3SH in water and 10 ns for C4H10 in water. However, in the case of the CH3 bond angle transformation in water (in CH3SH), the length of the simulation was 20 ns at each intermediate state. We have computed the free energy change using the Bennett’s acceptance ratio method in which a series of individual free energies is combined into a free energy estimate [24].
Appendix E. Additional partition function calculations

We substituted the (ATB force field) bond stretching and bond angle potentials in the corresponding equations in section 2 and integrated. We obtained the following expressions

$$Z_c = \frac{\pi e^{-\frac{a}{4}}}{8\beta k_b \rho_0} \left\{ a \left[ I_4 \left( \frac{a}{4} \right) + I_2 \left( \frac{a}{4} \right) \right] + aL_\frac{1}{4} \left( \frac{a}{4} \right) + (a + 2)I_4 \left( \frac{a}{4} \right) \right\},$$

(E1)

where $a \equiv \beta k_b \rho_0^4$ and $I_n(x)$ is the modified Bessel function of the first kind

$$Z_0 = \frac{\pi}{\sqrt{2\beta k_b}} \left\{ \text{erf} \left[ \sqrt{\frac{2\beta k_b}{\sin^2 \left( \frac{\theta_0}{2} \right)} \right] + \text{erf} \left[ \frac{\beta k_b}{\sqrt{2}} \left( \cos \left( \theta_0 \right) + 1 \right) \right] \right\},$$

(E2)

$$Z_{nb} = \int \prod_{i} \int \prod_{i>j} e^{-\frac{1}{2}k_b^4 \cos \theta_i \cos \theta_j} \int \prod_{i} \sin \theta_i d\theta_i \int \prod_{i>j} d\theta_{ij},$$

(E3)

References

[1] Zuckerman D M 2011 Ann. Rev. Biophys. 40 41
[2] Chipot C and Pohorille A 2007 Free Energy Calculations: Theory and Applications in Chemistry and Biology vol 86 (Berlin: Springer)
[3] Frenkel D and Smit B 1996 Understanding Molecular Simulation: From Algorithms to Applications (Orlando, FL: Academic)
[4] Binder K and Heermann D W 2010 Monte Carlo Simulation in Statistical Physics: An Introduction (Berlin: Springer)
[5] Allen M P and Tildesley D 1987 Computer Simulation of Liquids (Oxford: Oxford Science Publications)
[6] Jarzynski C 1997 Phys. Rev. Lett. 78 2690
[7] Crooks G 2000 Phys. Rev. E 61 2361
[8] Jorgensen W L, Briggs J M and Gao J 1987 J. Am. Chem. Soc. 109 6857
[9] L J H and Allinger N L 1989 J. Am. Chem. Soc. 111 8566
[10] Hayward S and Go N 1995 Ann. Rev. Phys. Chem. 46 223
[11] Mayo S L, Olafson B D and Goodard W A 1990 J. Phys. Chem. 94 8897
[12] Phillips J C, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel R D, Kale L and Shulten K 2005 J. Comput. Chem. 26 1781
[13] Hess B, Kutzner C, Van Der Spoel D and Lindahl E 2008 J. Chem. Theory Comput. 4 435
[14] Herschbach D R, Johnston H S and Rapp D 1959 J. Chem. Phys. 31 1652
[15] Borensch S and Karplus M 1999 J. Phys. Chem. A 103 103
[16] Borensch T, Tettinger F, Leitgeb M and Karplus M 2003 J. Phys. Chem. B 107 9535
[17] Malde A K, Zuo L, Breeze M, Stroet M, Poger D, Nair P C, Oostenbrink C and Mark A E 2011 J. Chem. Theory Comput. 7 4026
[18] Cuendet M A and Tuckerman M E 2012 J. Chem. Theory Comput. 8 3504
[19] Salomon-Ferrer R, Case D A and Walker R C 2013 Wiley Interdisciplinary Reviews: Computational Molecular Science vol 3, p 198
[20] Mughnai M L and Elber R 2012 J. Chem. Theory Comput. 8 3022
[21] Kirkwood J 1935 J. Chem. Phys. 3 300
[22] Strassmann T and McCammon J 1991 J. Chem. Phys. 95 1175
[23] Zwanzig R 1954 J. Chem. Phys. 22 1420
[24] Bennett C 1976 J. Comput. Phys. 22 245
[25] Kumar S, Rosenberg J, Bouzida D, Swendsen R and Kollman P 1992 J. Comput. Chem. 13 1011
[26] Beutler T C, Mark A E, van Schaik N, Gerber P R and van Gunsteren W F 1994 Chem. Phys. Lett. 222 529
[27] Farhi A 2011 arXiv:1212.4081
[28] Farhi A, Hed G, Bon M, Caticha N, Mak C and Domany E 2013 Physica A 392 5836
[29] Zavitsas A A 1987 J. Phys. Chem. 91 5573
[30] Blankensby S I and Ellison G B 2003 Acc. Chem. Res. 36 255
[31] Allinger N L, Yuh Y H and Lii J H 1989 J. Am. Chem. Soc. 111 8551
[32] Morse F M 1929 Phys. Rev. 34 17
[33] Abraham M, Van Der Spoel D, Lindahl E and Hess B 2015 GROMACS User Manual Version 5
[34] Ferrenberg A and Swendsen R 1988 Phys. Rev. Lett. 64 2635
[35] Hansen M 1997 Chem. Phys. Lett. 281 140
[36] Earl D and Deem M 2005 Phys. Chem. Chem. Phys. 7 3910
[37] Zhu Z, Tuckerman M E, Samuelson S O and Martyna G J 2002 Phys. Rev. Lett. 88 100201
[38] Minarly P, Tuckerman M E and Martyna G J 2008 SIAM J. Sci. Comput. 30 2055
[39] Protin S et al 2013 Bioinformatics