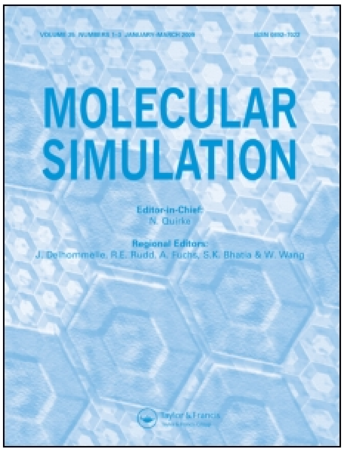

Molecular Simulation

Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Path Dependence of Free Energy Components in Thermodynamic Integration

M. Zacharias[a][b]; T. P. Straatsma[a]
[a] Department of Chemistry, University of Houston, Houston, TX, USA [b] Health Sciences Center, University of Colorado, Denver, CO

To cite this Article Zacharias, M. and Straatsma, T. P.(1995) 'Path Dependence of Free Energy Components in Thermodynamic Integration', Molecular Simulation, 14: 6, 417 — 423
To link to this Article: DOI: 10.1080/08927029508022035
URL: http://dx.doi.org/10.1080/08927029508022035

Molecular Simulation, 1995, Vol. 14, pp. 417–423
Reprints available directly from the publisher
Photocopying permitted by license only

PATH DEPENDENCE OF FREE ENERGY COMPONENTS IN THERMODYNAMIC INTEGRATION

M. ZACHARIAS[1] and T. P. STRAATSMA[2]

Department of Chemistry, University of Houston, Houston, TX 77204-5641, USA

(Received July 1994, accepted November 1994)

In this article it is shown that the decomposition of free energy differences, as evaluated by molecular simulations using thermodynamic integration, in terms of components corresponding to terms in the Hamiltonian, cannot be uniquely performed. This path dependence of the contributions to the total free energy difference is shown for the calculated difference in free energy of hydration of ethanol and ethane. The conclusion of this study is that the individual free energy difference contributions from such calculations should, in general, be interpreted with care.

KEY WORDS: Free energy, thermodynamic integration, hydration, ethanol, ethane.

1 INTRODUCTION

Thermodynamic integration has been widely used to extract free energy differences from molecular simulations [1,2]. The Hamiltonian of the molecular system is coupled to a variable λ in such a way that for $\lambda = 0$ the system is represented by a Hamiltonian $\mathscr{H}_A$, corresponding to state A, and for $\lambda = 1$ by a Hamiltonian $\mathscr{H}_B$, corresponding to state B. During a molecular dynamics simulation interconversion between state A and B is achieved by varying λ between 0 and 1. In the thermodynamic integration the derivative of the Hamiltonian vs. λ is recorded and integrated over the change in λ [1]. The free energy difference is given by,

$$\Delta G = \int_0^1 \left\langle \frac{\partial \mathscr{H}(\lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda \tag{1}$$

which, in a computer simulation experiment, is approximated by a sum over ensemble averages for the derivative of the Hamiltonian versus λ at discrete steps in λ,

$$\Delta G = \sum_i \left\langle \frac{\partial \mathscr{H}(\lambda)}{\partial \lambda} \right\rangle_{\lambda_i} \Delta \lambda_i \tag{2}$$

---

[1] Current address: Health Sciences Center, University of Colorado, Denver, CO 80262.
[2] To whom correspondence should addressed.

METHODOLOGY

The Hamiltonian in a molecular dynamics (MD) simulation consists of various bonded and nonbonded potential energy terms. In the GROMOS force field [3] used in the present study, but also in most other currently used force fields, the Hamiltonian for the bonded interaction is a sum of pairwise harmonic bond terms, three-body angle terms and four-body dihedral angle terms. The Hamiltonian representing the nonbonded interactions is a sum over pairwise van der Waals terms, represented by a Lennard-Jones (LJ) type potential, and electrostatic terms. The detailed form of the Hamiltonian used in the present study is given in the Appendix. During thermodynamic integration the ensemble average of the derivative for each term of the Hamiltonian vs. λ can be evaluated and integrated in the course of the simulation. It is, therefore, possible to formally split the recorded free energy change into different free energy components.

Thermodynamic integration has frequently been used to determine the contribution of individual amino acid residues as well as different terms in the potential energy function to the calculated free energy difference [4–8].

One needs, however, to be aware of the fact that these contributions to the free energy change are not necessarily independent of the path taken to perform the change in the system. It has been pointed out [9, 10] that only the sum of all free energy contributions is a state function, here given for the NPT ensemble,

$$\Delta G = \int_0^1 \left\langle \frac{\partial \mathcal{H}(\lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda = \int_0^1 \frac{\int \frac{\partial \mathcal{H}(\lambda)}{\partial \lambda} \exp\left(-\frac{\sum_j \mathcal{H}_j(\lambda) + pV}{k_B T} \right) d\Gamma}{\int \exp\left(-\frac{\sum_j \mathcal{H}_j(\lambda) + pV}{k_B T} \right) d\Gamma} \, d\lambda \qquad (3)$$

with $\mathcal{H} = \sum_i \mathcal{H}_j(\lambda) + pV$ representing the Hamiltonian of the system and the index j corresponding to the various interaction terms. The integrals over $d\Gamma$ represent integration over phase space. The integral in Equation 3 is complete, i.e. it satisfies the chain rule, and ΔG depends only on the initial and final state of the system. This is not true for a free energy component, indicated with index i,

$$\Delta G = \int_0^1 \left\langle \frac{\partial \mathcal{H}(\lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda = \int_0^1 \frac{\int \frac{\partial \mathcal{H}_i(\lambda)}{\partial \lambda} \exp\left(-\frac{\sum_j \mathcal{H}_j(\lambda) + pV}{k_B T} \right) d\Gamma}{\int \exp\left(-\frac{\sum_j \mathcal{H}_j(\lambda) + pV}{k_B T} \right) d\Gamma} \, d\lambda \qquad (4)$$

Here, ΔG_i depends on the functional form of the complete Hamiltonian, $\sum_j \mathcal{H}_j(\lambda) + pV$, and therefore on the path taken to perform the thermodynamic integration [9]. For example, in case of the annihilation of a charged atom, the Hamiltonian $\mathcal{H}$ in Equation 4 would depend on whether one removes charges and LJ interaction simulataneously or in two separate steps resulting in different Lennard-Jones and electrostatic free energy components.

Although in previous work on thermodynamic integration with free energy decomposition a specific path has always been indicated [4–8] it has never been thoroughly investigated to which degree the calculated free energy components change upon varying the path taken to perform the TI. Purpose of the present study is the systematic evaluation of free energy components for a small molecule transition in aqueous solution using thermodynamic integration following various paths.

3 RESULTS AND DISCUSSION

The significance of the path dependence of free energy decomposition in thermodynamic integration will be demonstrated in the present study on the calculation of the free energy difference of ethanol and ethane in aqueous solution. Free energy difference calculations on ethanol to ethane in a box containing 192 SPC/E water molecules [11] were carried out with the ARGOS program package [12] using the GROMOS parameter set [3]. The bond sprouting/desprouting technique was used to increase the statistical accuracy when annihilating the hydroxyl group of ethanol [13]. Bond lengths were kept fixed using SHAKE [14]. This results in three components of the calculated free energy difference, Lennard-Jones (LJ), electrostatic and constraint contributions. The constraint contribution corresponds to the free energy change upon changing the bond length between O1, H1 and C1, O1, respectively, in the course of the thermodynamic integration using the bond sprouting/desprouting technique. It is referred to as constraint contribution and not bond contribution since for each individual value of the control parameter λ the bond lengths are constraint to their optimal lengths. Figure 1 illustrates the various paths to transform ethanol into ethane. Paths which lead to the annihilation of the LJ interactions before decreasing the bond length between hydrogen/oxygen (H1/O1) and oxygen/carbon (O1/C1) were not included in the present study because it would result in large statistical errors [13]. Each transformation was performed using multiconfigurational thermodynamic integration (MCTI) [15]. As can be judged from Table 1, the overall free energy change comparing the different paths of transforming ethanol to ethane is relatively well conserved. Since the bond length between atoms H1, O1 and O1, C1, respectively, changed during the TI (by factor 0.5), the calculated free energy difference needs to be corrected for a moment of inertia contribution before finally comparing it with experiment [13]. This correction amounts to 3.3 kJ mol^{-1} [13]. Since there are no intramolecular contributions to the free energy change, an MCTI in vacuum to complete the thermodynamic cycle is not necessary. The calculated average free energy difference between ethanol and ethane in aqueous solution after correcting for the change in moment of inertia is 27.3 kJ mol^{-1} in good agreement with the experimental value of 28.8 kJ mol^{-1} [16]. The differences in calculated free energies of solvation following various paths were less than 3 kJ mol^{-1} indicating convergence of the MCTI calculations. The free energy decomposition, however, depends strongly on the path taken to perform the transition. In case of the direct transition from ethanol to ethane, changing all components at the same time along path 0 in Figure 1, one observes a relatively large electrostatic contribution and a small LJ contribution. In contrast, for example, when first decreasing the bond lengths and removing the partial charges and subsequently in a second step annihilating the LJ interaction, path 1a + 1c, the LJ contribution increases 2.5 times and the electrostatic and constraint contributions are

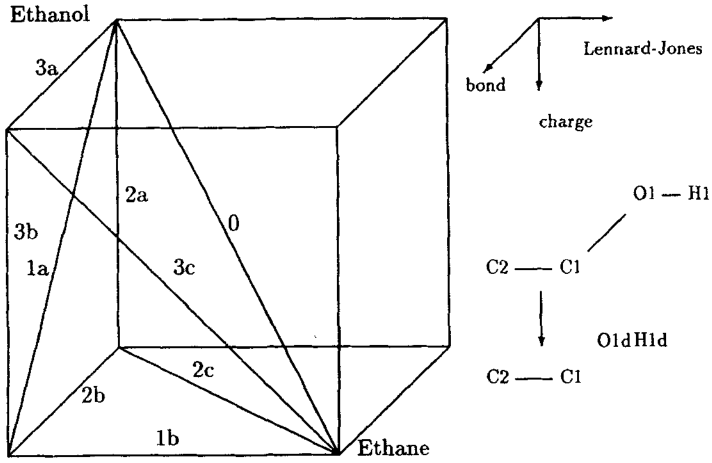

Figure 1 Ethanol to ethane cube. Figure 1 illustrates the various possibilities to transform ethanol to ethane within a thermodynamic integration calculation. Each edge of the cube corresponds to the change in one component of the interaction function modified in the course of the simulation. All paths for which the free energy change was evaluated are indicated by a number. Path 0, for example, is the "direct" path of transforming ethanol to ethane by simultaneously reducing the bondlengths (O1–C1, H1–O1), removing partial charges and annihilating the LJ interaction. Path 1a + 1b corresponds to first decreasing the bond lengths and removing the partial charges (1a) and then in a second simulation (1b) annihilating the LJ interaction. The simulated transition of ethanol to ethane is shown schematically including the reduction in bond length between atoms O1–C1 and H1–O1, respectively, and the transformation into dummy atoms (O1d, H1d). The free energies evaluated for the different paths are listed in Table 1.

Table 1 Path Dependence of Free Energy Components.

Path[a]	ΔG_{total}[b]	ΔG_{LJ}[c]	$\Delta G_{electrostatic}$[d]	$\Delta G_{constraint}$[e]
0	31.9 ± 0.6	2.3	16.5	13.1
1a + 1b	31.1 ± 0.8	5.7	14.7	10.6
2a + 2c	29.2 + 0.7	4.4	23.8	1.0
2a + 2b + 1b	29.5 ± 0.7	5.7	23.8	0.05
3a + 3c	30.9 + 0.5	5.6	2.7	22.7
3a + 3b + 1b	30.9 ± 0.6	5.7	2.5	22.7

[a] the various paths to transform ethanol to ethane are illustrated in Figure 1.

[b] total calculated free energy difference between ethanol and ethane in aqueous solution, without moment of inertia correction. ΔG is given in kJ mol^{-1}. The statistical error was evaluated using a method by Straatsma *et al.*[17]

[c] Lennard-Jones free energy contribution.

[d] electrostatic free energy contribution.

[e] constraint contribution to the calculated free energy change.

smaller than for path 0. Even more dramatic, when first reducing the bond length and then uncharging and annihilating the LJ potential, path 3a + 3c, the electrostatic contribution decreases to 1/6th whereas both the LJ and constraint contributions are two times larger compared to the direct transition, path 0. The increased constraint contribution could in part be interpreted as an effective reduction of the electrostatic dipole moment upon reducing the bond length between atoms with partial charges (O1, H1 and C1, O1). However, beside of the coupling to the electrostatic interaction the constraint contribution is also coupled to the Lennard-Jones (van der Waals) interaction of the solute atoms with solvent which is manifested in a change of the LJ contribution for path 3a + 3c compared to path 0.

The calculated free energy contributions show variations by as much as 2/3 of the total free energy change depending on the path taken to perform the TI. This result clearly illustrates the significance of the path dependence for free energy decomposition in thermodynamic integration and shows that free energy decomposition without path indication is not unique.

CONCLUSION

Although this and other studies [18, 19] demonstrate the danger of the free energy decompositions that have often been attempted, such decompositions would of course be very useful if they could be made reliably. Free energy decomposition in thermodynamic integration can be meaningful in cases where the specified path is of particular interest, for example, when calculating the potential of mean force along the line connecting to particles in solution. However, at present, the only decompositions that can generally be made in a rigorous fashion are mechanical thermodynamical functions, such as internal energy or enthalpy. The mean electrostatic potential energy of a system in a given thermodynamic state, for example, is well-defined. Whether optimal approximations can be defined for decompositions of non-mechanical thermodynamic functions and their changes is deserving of further study.

Acknowledgement

We thank J. A. McCammon, W. F.van Gunsteren, P. E. Smith and M. Karplus for stimulating discussions. M. Z. was supported in part by a postdoctoral fellowship from Deutsche Forschungsgemeinschaft. This work was supported in part grants to Professor McCammon from NSF, the Robert A. Welch Foundation, and the Grand Challenge program of the NSF Supercomputer centers.

References

[1] T.P. Straatsma and J.A. McCammon, "Computational alchemy", *Ann. Rev. Phys. Chem.*, **43**, 407–432 (1992).
[2] C.A. Reynolds, P.M. King and W.G. Richards, "Free energy calculations in molecular biophysics", *Mol. Biophys.*, **76**, 251–275 (1992).
[3] W.F. van Gunsteren and H.J.C. Berendsen, *Groningen Molecular Simulation (GRO-MOS) Library Manual*, BIOMOSB.V., Nijenborgh 16, 9747 AG Groningen, The Nether lands, 1987.
[4] A. Warshel, F. Sussman and J-K. Hwang, "Evaluation of catalytic free energies in genetically modified proteins", *J. Mol. Biol.*, **201**, 139–159 (1988).
[5] V. Daggett, F. Brown and P.A. Kollman, "Free energy component analysis: A study of the glutamic acid 165-aspartic acid 165 mutation in triose phosphate isomerase", *J. Am. Chem. Soc.*, **111**, 8247–8256 (1989).

[6] J. Gao, K. Kuczera, B. Tidor and M. Karplus, "Hidden thermodynamics of mutant proteins: A molecular dynamics study", *Science*, **244**, 1069–1072 (1989).

[7] K. Kuczera, J. Gao, B. Tidor and M. Karplus, "Free energy of sickling: A simulation analysis", *Proc. Natl. Acad. Sci. USA*, **87**, 8481–8485 (1990).

[8] F.S. Lee and A. Warshel, "A local reaction field method for fast evaluation of long range electrostatic interactions in molecular simulations", *J. Chem. Phys.*, **97**, 3100–3107 (1992).

[9] T.P. Straatsma, M. Zacharias and J.A. McCammon, "Free energy difference calculations in biomecular systems", pp. 349–367 in *Computer Simulation of Biomecular Systems Vol. 2: Theoretical and Experimental Applications*, W. F. van Gunsteren, P.K. Weiner and A.J. Wilkinson, eds., ESCOM Science Publisher, Leiden, 1994.

[10] W.F. van Gunsteren, T.C. Beutler, F. Fraternali, P.M. King, A.E. Mark and P.E. Smith, in *Computer Simulation of Biomolecular Systems Vol. 2: Theoretical and Experimental Applications*, W.F. van Gunsteren, P.K. Weiner and A.J. Wilkinson, eds., ESCOM Science Publisher, Leiden, 1994.

[11] H.J.C. Berendsen, J.P.M. Postma and T.P. Straatsma, "The missing term in effective pair potentials", *J. Phys. Chem.*, **91**, 6269–6271 (1987).

[12] T.P. Straatsma and J.A. McCammon, "ARGOS: A vectorized molecular dynamics program package", *J. Comp. Chem.*, **11**, 943–947 (1990).

[13] T.P. Straatsma, M. Zacharias and J.A. McCammon, "Holonomic constraint contributions to free energy differences from thermodynamic integration molecular dynamics simulations", *Chem. Phys. Lett.*, **196**, 297–303 (1992).

[14] J.P. Ryckaert, G. Ciccotti and H.J.C. Berendsen, "Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes", *J. Comput. Phys.*, **23**, 327–343 (1977).

[15] T.P. Straatsma and J.A. McCammon, "Multiconfiguration thermodynamic integration", *J. Chem. Phys.*, **95**, 1175–1188 (1991).

[16] S. Cabani, P. Gianni, V. Mollica and L. Lepori, "Group contributions to the thermodynamic properties of non-ionic organic solutes in dilute aqueous solution", *J. Sol. Chem.*, **10**, 563 (1981).

[17] T.P. Straatsma, H.J.C. Berendsen and A.J. Stam, "Estimation of statistical errors in molecular simulation calculations", *Mol. Phys.*, **57**, 89 (1987).

[18] P.E. Smith and W.F. van Gunsteren, "When are free energy components meaningful?", *J. Phys. Chem.*, in press.

[19] A.E. Mark and W.F. van Gunsteren, "Decomposition of the free energy of a system in terms of specific interactions", *J. Mol. Biol.*, **240**, 167–176 (1994).

[20] L. Verlet, "Computer experiments on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules", *Physical Rev.*, **159**, 98–112 (1967).

[21] H.J.C. Berendsen, J.P.M. Postma, W.F van Gunsteren, A. DiNola and J.R. Haak, "Molecular dynamics with coupling to an external bath", *J. Chem. Phys.*, **81**, 3684–3690 (1984).

APPENDIX

The Hamiltonian $\mathscr{H}$ used in the present simulation consisted of bond, angle, proper dihedral torsion, improper dihedral torsion, Lennard-Jones and electrostatic terms and had the following functional form,

$$\mathscr{H} = \sum_{\text{bonds}} \frac{1}{2} K_b (b - b_0)^2 + \sum_{\text{angles}} \frac{1}{2} K_\theta (\theta - \theta_0)^2 + \sum_{\text{proper} - \text{torsion}} K_\phi (1 + \cos(n\phi - \delta)) + \sum_{\text{improper} - \text{torsion}} \frac{1}{2} K_\xi (\xi - \xi_0)^2 + \sum_{i<j} \left(\frac{C_{12}}{r_{ij}^{12}} - \frac{C_6}{r_{ij}^6} + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right) \tag{5}$$

Here, the K_x (with $x = b, \theta, \phi, \xi$) are appropriate force constants for the different bonded terms. C_{12}, C_6 are repulsive and attractive Lennard-Jones parameters, respectively, r_{ij} is the interatomic distance, ε_0 is the dielectric constant in vacuum and q_i and q_j

are the charges on atoms i and j. The GROMOS parameter set [3] was used. The ARGOS program package [12] was used for all computer simulations. ARGOS uses the leapfrog algorithm which is a modification of the Verlet algorithm [20] to simulate Newton's equation of motion for the system consisting of an ethanol(ethane) molecule in a box of 192 water molecules. Rectangular periodic boundary conditions were employed with a pressure of 1.0×10^5 Pa. The average box size was $(1.8)^3$ nm^3. Water molecules were representred by the SPC/E model [11]. All molecular dynamics/thermodynamic integration calculations were run under constant temperature (300 K) by coupling to an external heat bath [21]. All bond lengths, including the perturbed bond, were constrained to their standard length using SHAKE [14], which allowed a time step of 2.0 fs in the MD simulations. The cutoff radius for all nonbonded interactions between solute atoms and solvent atoms was half the box size (0.9 nm).

Thermodynamic integration was performed using the multiconfigurational thermodynamic integration technique (MCTI) [15], which offers the possibility of systematically improving the calculations for each step in the control parameter λ. Ensemble average for the derivative of the Hamiltonian with respect to λ a well as other properties can be evaluated independently for each λ. The statistical error of the free energy difference for each value of λ was calculated from a correlation analysis of the MD run for each λ, [17] the statistical error of the complete MCTI was then calculated from [15]

$$E(\Delta G) = \left(\sum_i E_i^2 \Delta\lambda_i\right)^{1/2}, \tag{6}$$

The drift in the calculated free energy for the MCTI was evaluated from the average of the drift of each run at each λ [15]. Free energy difference calculations were performed in 21 evenly spaced steps of the control variable λ with 1000 equilibration MD steps and 2500 data gathering MD steps per each λ essentially as described earlier [13]. The control parameter λ was linearly coupled to the Hamiltonian. The statistical error of the free energy evaluation along each path was less than 1.0 kJ mol^{-1} ps^{-1}. The drift in the calculated free energies was less than 0.5 kJ mol^{-1} ps^{-1}. The statistical error for each path j taken to perform the ethanol to ethane transition was calculated from,

$$E(\Delta G_j) = \left(\sum_i E_i^2\right)^{1/2}, \tag{7}$$

where the sum is over the individual MCTIs along the edges of the cube (Figure 1).