The Role of Bonded Terms in Free Energy Simulations. 2. Calculation of Their Influence on Free Energy Differences of Solvation

Stefan Boresch[†,‡,§] and Martin Karplus*[,†,‡]

Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138, Laboratoire de Chimie Biophysique, Institut Le Bel, Université Louis Pasteur, 67000 Strasbourg, France, and Institut für Theoretische Chemie, University of Vienna, Währingerstrasse 17, A-1090 Vienna, Austria

Received: March 25, 1998

Calculations of the free energy difference of solvation are used to study the contributions arising from alchemical changes of bond stretching and angle bending energy terms in the force field. The results illustrate the theoretical analysis of such terms given in the companion paper (Boresch, S.; Karplus, M. The Role of Bonded Terms in Free Energy Simulations: 1. Theoretical Analysis. *J. Phys. Chem. A* **1998**, *103*, 103[10]). Three model systems are investigated: (a) two one-dimensional harmonic oscillators interacting with a third particle that represents the solvent, (b) the aqueous solvation of two diatomic molecules, and (c) the aqueous solvation of ethane and methanol. In each case, the computations are carried out with both a single topology and a dual topology methodology. A comparison of free energy components of the single and double free energy differences obtained in the calculations makes it possible to identify the three contributions that the theoretical analysis showed were involved, i.e., *vibrational*, *pmf-type*, and *Jacobian factor* terms.The verification of the theoretical analysis by illustrative examples provides the basis for addressing the question of whether the so-called *self-terms* can make significant contributions to double free energy differences. This is accomplished by identifying the effect of coupling of the three contributions from bonded energy terms on a double free energy difference. For the model systems studied, coupling and, hence, self-terms are found to be of little importance. The analysis resolves the ambiguities concerning this issue in the literature.

1. Introduction

Although bonded energy terms (bond stretching and angle bending terms) have often been discussed in the literature and cited as problematic, no systematic investigation concerning their role in free energy simulations has been published. Computational procedures for evaluating them explicitly or implicitly vary considerably, and a number of practical problems related to bonded terms have been encountered in free energy simulations. Bonded energy terms have been omitted from the free energy formalism by some.[1,2] Others included these terms and reported severe convergence problems in the calculations.[3,4] Moreover, there are conflicting views concerning the importance of bonded energy terms in the literature.[3−9] This is reflected in the discussion regarding the relevance of the so-called *self-term*[5] or *intraperturbed group*[6,9] contribution, since this quantity is expected to be dominated by bonded energy terms.[5,6]

A detailed theoretical analysis of the role of bonded energy terms in free energy simulations was given in the companion paper.[10] It was shown that their contribution depends foremost on the simulation methodology, i.e., single or dual topology.[11] Many apparent contradictions in the literature[3−9] were identified as resulting from an incomplete understanding of these two commonly used approaches. In single topology simulations, free energy differences resulting from an alchemical change in a bond (angle) term were demonstrated to arise from the following three physical effects: (i) *Vibrational* contributions from changes in the force constants, (ii) *Jacobian factor* contributions from changes in the equilibrium bond lengths and bond angles in the absence of nonbonded interactions, and (iii) *pmf-type* contributions which result from the change in nonbonded interaction (e.g., solute−solvent interactions) if the equilibrium (average) geometry of a molecule is altered. In dual topology simulations one has to distinguish between the use of an ideal gas molecule end state and an ideal gas atom end state. In the former case, which appears to be the more practical dual topology approach, vibrational and Jacobian factor contributions are omitted. Consequently, single free energy differences obtained in a dual topology method with an ideal gas molecule end state will differ from those obtained in single topology methods or dual topology methods with an ideal gas atom end state. Nevertheless, correct double free energy differences are obtained in all cases. All three contributions are obtained in dual topology methods when an ideal gas atom end state is used; however, the pmf-type free energy contribution is projected on nonbonded free energy components.

The separation of bonded free energy contributions into vibrational, Jacobian factor and pmf-type contributions is based on the rigid rotator harmonic oscillator (RRHO) approximation. Even though it is not exact, it is often accurate and always useful for obtaining an understanding of the underlying physical processes, as well as of the differences between the three methods (single vs the two dual topology methods). Since free energy simulation methods do not rely on the RRHO approximation, all three methods include the coupling contributions that arise from anharmonic terms and the violation of the rigid rotator assumption.

* Corresponding author.
† Department of Chemistry.
‡ Laboratoire de Chimie Biophysique.
§ Institut für Theoretische Chemie.

10.1021/jp981629f CCC: $18.00
Published on Web 12/12/1998

The degree of coupling between the three contributions is directly related to the question of self-terms. A self-term[5] or intragroup perturbed interaction[6] arises from the change in the intramolecular or intragroup energy terms of the part of the system that is alchemically mutated, i.e., ΔU_{intra} in the notation of the companion paper[10], as opposed to contributions resulting from the change of interaction caused by the mutation with the (unaltered) part of the system. If the RRHO approximation were exact, vibrational and Jacobian factor contributions of corresponding (parallel) parts of a thermodynamic cycle would cancel; i.e., no contributions from bonded energy terms to the self-term would be obtained. If, on the other hand, coupling is significant, no such cancelation can be expected and self-term contributions will result. Computations for model systems are required to determine the size of the coupling and, hence, of self-term contributions from bonded energy terms.

In this paper we illustrate and complement the theoretical analysis given in the companion paper[10] by applying it to calculations of the free energy differences of solvation for selected model systems. They are (a) two one-dimensional harmonic oscillators interacting with a third particle that represents the solvent, (b) the aqueous solvation of two diatomic molecules, and (c) the aqueous solvation of ethane and methanol. To ensure the reliability of the results, each simulation is carried out at least twice unless, as is the case for the one-dimensional system, an exact reference result can be obtained. The paper is organized as follows. We first describe the three systems that we examine, the details of the simulations, and the rationale for choosing them (section 2). In particular, the properties of the hybrid potential energy functions used are presented, and differences between single and dual topology methodologies are discussed (cf. also sections 2b and 2e of the companion paper[10]). Section 3 contains the results of the calculations. On the basis of the theoretical analysis[10] combined with the insights from actual calculations, we present our conclusions regarding the practical aspects of the computation of bonded free energy contributions, the interpretation of the bonded terms and the related question of the importance of self-terms (section 4).

2. Model Systems and Details of Simulations

2a. Free Energy Difference of "Solvation" of a One-Dimensional Model System. The first system is concerned with the "solvation" free energy difference between two one-dimensional harmonic oscillator solutes interacting with a third particle that represents the solvent (see Scheme 1). Three sets of computations aimed at investigating the different types of contributions to bond free energy components and, in particular, coupling between vibrational and pmf-type contributions were carried out. The model is taken from a recent publication by Severance et al.,[12] who used it to verify the proposed algorithm to incorporate flexible bond terms in the exponential formula (EF) approach of free energy difference simulations. We selected this model system to examine the effect of alchemical changes in the calculation of the free energy difference of solvation between two one-dimensional diatomic molecules, the simplest system discussed in section 2c of the companion paper.[10]

Scheme 1

B1−B2 B3

The distance between particles B1 and B3 is kept fixed at 5 Å. Particle B2 is bound to B1 by a harmonic oscillator term and in "solution" interacts through a Lennard-Jones potential with B3; particle B1 does not interact with B3. The two solutes

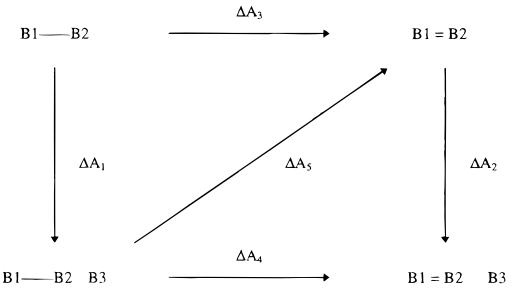

Figure 1. Thermodynamic cycle for the one-dimensional model system. ΔA_3 and ΔA_4 correspond to the alchemical mutations changing molecule B1−B2 into B1═B2 in the "gas phase" and in "solution". ΔA_1 and ΔA_2 correspond to the chemical paths along which transfer free energies between the "gas phase" and "solvated phase" would be measured. The path ΔA_5 is included to illustrate a case where bonded and nonbonded interactions are changed simultaneously as a result of an alchemical mutation.

i and f represent harmonic oscillators, one with $K_i = 260$ kcal/(mol Å^2) and $r_{0,i} = 1.526$ Å and the other with $K_f = 80$ kcal/(mol Å^2) and $r_{0,f} = 0.3$ Å. The van der Waals interaction between B2 and B3 is described by $\epsilon = -0.175$ kcal/mol and $\sigma = 3.905$ Å in both states (i and f). No charges are present.

Figure 1 depicts the paths that were used in the first set of simulations. The thermodynamic cycle in Figure 1 is analogous to that in Figure 1 of the companion paper[10] with S1 and S2 replaced by i and f. Transformation of the "solute" B1−B2 from i to f in the absence of interactions with B3 represents the gas phase (ΔA_3); ΔA_4 is the free energy difference between i and f in the presence of the "solvent" B3. The bond term (K, r_0) changes along these alchemical paths. The chemical paths (ΔA_1 and ΔA_2 of Figure 1) correspond to transferring i and f from the gas phase into "solution", which is accomplished by turning on the van der Waals interaction between particles B2 and B3 for system i and f, respectively. We also calculated the change $(i)_{solv}$ to $(f)_{gas}$ (ΔA_5 of Figure 1) in which all three terms describing the system (force constant K, equilibrium bond length r_0, and van der Waals parameters ϵ, σ) are changed in a concerted fashion; i.e., they were modified simultaneously in going from the initial to the final state.

As pointed out by Severance et al.,[12] the configurational partition functions for this system and, thus, the free energies can be obtained by numerical integration. Numerical integrations were carried out using the NIntegrate[] function of Mathematica.[13] All free energy differences depicted in Figure 1 were also calculated with the PERT module of CHARMM,[14] which in this case was used to set up the single topology, as well as the dual topology calculations, in which bond energy terms where not scaled (ideal gas molecule end state). Slow convergence resulting in high statistical uncertainties is observed when simulation methods are used to attempt dual topology simulations in which the bond energy terms are scaled to a limiting value of λ as described in section 2d of the companion paper.[10] To avoid these numerical problems, Mathematica[13] is employed to evaluate $\langle \partial U/\partial\lambda \rangle_\lambda$, $\langle \partial U_{vdW}/\partial\lambda \rangle_\lambda$, and $\langle \partial U_{bond}/\partial\lambda \rangle_\lambda$ directly by numerical integration of the respective integrals for five values of λ (0.1, 0.3, ..., 0.9). The van der Waals free energy component is obtained in the standard manner by numerical integration of the data-points with the trapezoidal rule. The values for $\langle \partial U_{bond}/\partial\lambda \rangle_\lambda$ are fitted to $C_o + C_1/\lambda$ ($C_o + C_1/(1 - \lambda)$) as described in section 2d of the companion paper;[10] this function was integrated from ϵ to 1 (0 to $1 - \epsilon$), where ϵ is given by eq 40 of the companion paper.[10]

Atoms B1 and B3 are not allowed to move during the simulations. In the single topology calculation, one B2 atom is

defined, and the energy terms are changed as a function of the coupling parameter λ (i: $\lambda = 0$, f: $\lambda = 1$). In dual topology, two B2 atoms ($B2^i$ corresponding to the initial state and $B2^f$ corresponding to the final state) are present simultaneously. Thus, there are four atoms in the dual topology calculations and only three atoms in the single topology calculations. In principle, the dual topology system could have been set up by defining two molecules (four atoms) on top of each other. However, in realistic applications the full molecule (such as a protein) is not duplicated, and it is best to duplicate only the part that changes to keep the two molecules aligned as much as possible. To make the difference between single and dual topology clear, we give expressions for the hybrid potential function used to calculate ΔA_4. The single topology hybrid potential energy function is

$$U_{s.t.}(\lambda) = (1-\lambda)\{U_{i,bond}(B1,B2) + U_{i,LJ}(B2,B3)\} + \lambda\{U_{f,bond}(B1,B2) + U_{f,LJ}(B2,B3)\} \quad (1a)$$

and the two different dual topology hybrid functions are

$$U_{d.t.1}(\lambda) = U_{i,bond}(B1,B2^i) + U_{f,bond}(B1,B2^f) + (1-\lambda)U_{i,LJ}(B2^i,B3) + \lambda U_{f,LJ}(B2^f,B3) \quad (1b)$$

and

$$U_{d.t.2}(\lambda) = (1-\lambda)U_{i,bond}(B1,B2^f) + \lambda U_{f,bond}(B1,B2^f) + (1-\lambda)U_{i,LJ}(B2^i,B3) + \lambda U_{f,LJ}(B2^f,B3) \quad (1c)$$

The equations indicate which atoms are involved in each energy term; the subscripts and superscripts i and f stand for initial (harmonic oscillator i in solution) and final state (harmonic oscillator f in solution), respectively. Aside from the difference in the number of atoms, the bond energy term is treated differently in the three hybrid functions. In $U_{s.t.}(\lambda)$ the bond term is changed alchemically, i.e., it is a function of λ. In $U_{d.t.1}(\lambda)$, the bond term does not depend on λ so that the end states correspond to ideal gas molecules as in $U_{s.t.}(\lambda)$; in_$U_{d.t.2}(\lambda)$ the bonded terms are scaled, corresponding to ideal gas atom end states (see section 2b of ref 10). For the chemical transfer processes from the gas phase into "solution" (ΔA_1, ΔA_2), there is no difference between single topology and the two dual topology methods as only the nonbonded interaction between B2 and B3 of i and f, respectively, are scaled as a function of the coupling parameter in all three hybrid potentials.

The free energy simulations are carried out with a slow-growth, thermodynamic integration protocol; in all cases, the system is equilibrated for 10000 steps (a time step of 1 fs was used) at $\lambda = 0$, then the coupling parameter is changed from 0 to 1 over the following 100 000 simulation steps. As the results agree excellently with the reference results from numerical integration of the partition function, only one simulation is carried out for each path. The velocity of particle(s) B2 ($B2^i$ and $B2^f$ in dual topology) are kept at an average temperature of 300 K by a Nosé thermostat,[15] implemented in CHARMM c24b1 by M. Watanabe. Separate thermostats for each of the two atoms ($B2^i$ and $B2^f$) are used in dual topology.

Additional calculations (sets 2 and 3) were carried out to probe how much the presence of the nonbonded interaction alters the harmonic oscillator term. In a second set of calculations the free energy difference for changing the equilibrium bond length in the presence of the nonbonded interaction between B2 and B3 was computed for force constants of varying strength, as well as for a rigid system. The comparison of the results obtained using a flexible bond with several values of the force constant to those for a rigid system addresses the question whether these two approaches (flexible vs constrained bond terms) lead to different results. These calculation were carried out with both the single and dual topology (ideal gas molecule end state) methods, and the free energy difference was also calculated using Mathematica.[13] Further, dual topology calculations in which the bond terms were scaled (ideal gas atom end state) were carried out using Mathematica.[13] In each run the bond length was changed from 1.526 to 0.3 Å; the force constant K and the van der Waals parameters ($\epsilon = -0.175$ kcal/mol, $\sigma = 3.905$ Å) remain the same. The computations are repeated with four force constants K = 1,000, 260, 80, 10, kcal/(mol Å^2) to illustrate the range of effects; clearly the value of 10 kcal/(mol Å^2) is too small to be realistic for a covalent bond though it might correspond to the effective force constant for a hydrogen bond.

The calculation with a rigid bond removes the last degree of freedom, i.e., the free energy difference reduces to

$$\Delta A = U(r_f) - U(r_i) \quad (2)$$

where U is the potential energy function. The distance (r_f or r_i) between particles B1 and B2 (at $\lambda = 0$ or 1) fully determines the system as the B1−B3 distance is fixed. There are, however, two choices for r_i and r_f. One possibility is to use the equilibrium bond length r_o as the reference value for the constraint; in this case r_f(B1−B2) = 0.3 Å and r_i(B1−B2) = 1.526 Å. An alternative choice for r_f and r_i is to use the values for r_f and r_i which give the minimum energy for the corresponding flexible system. Free energy differences are calculated according to eq 2 for both choices of r_i and r_f and compared to the result for the corresponding flexible system.

Finally, we test the case where the equilibrium bond length (1.526 Å) of the harmonic oscillator term, as well as the nonbonded interaction ($\epsilon = -0.175$ kcal/mol, $\sigma = 3.905$ Å) between particles B2 and B3, are left unchanged while the force constant is reduced to one-half of its original value (third set of computations). Two initial values of the force constant, K = 250 and K = 17 kcal/(mol Å^2) are used. The same methods as in the other cases (single topology and dual topology with ideal gas molecule end state using simulation, dual topology with ideal gas atom end state using Mathematica, and direct computation of ΔA using Mathematica[13]) are employed.

2b. Free Energy Difference of Solvation between Two Hypothetical Diatomic Molecules. To verify the conclusions derived from the simplified one-dimensional system, we study an analogous three-dimensional diatomic molecule in water. As the solute moves in three dimensions, it is possible to demonstrate the role of Jacobian factors, which do not arise in the one-dimensional case. The same system was used to study the local path-dependence of free energy components; these results will be presented in detail elsewhere (Boresch, S.; Karplus, M. To be published). The bond length of the hypothetical diatomic molecule was changed from 1 to 2 Å. All other parameters (charges and van der Waals parameters of the solute and, in the case of a flexible solute, the force constant for the bond term) remained unchanged. The parameters of the solute were force constant, K = 300 kcal/(mol Å^2); van der Waals parameters, $\epsilon = -0.08$ kcal/mol, $\sigma = 2.06$ Å, and charges, $q = \pm 0.5e$. The charges on the solute do not interact. The mass of the solute atoms was set to 15 amu.

The alchemical transformation was realized in two ways: calculations were (i) run with the PERT module of CHARMM (single topology approach) and (ii) the BLOCK module of

TABLE 1: Description of Simulations Carried out for the Diatomic Model System in Water

	no. of runs[a]	protocol	no. λ-values	details of protocol[b]
PERT (s.t.)[c]				
gas phase[d]	10 fw/10 bw	windowing/TI	20	$\lambda = 0.025$: 14000/2000 $\lambda = 0.075$: 2000/2000 $\lambda = 0.975$: 2000/2000
solution 1[e]	1 fw/1 bw	windowing/TI	20	as for gas phase
solution 2[f]	1 fw/1 bw	windowing/TI	20	as for gas phase
BLOCK (d.t.)[g]				
solution 1[h]	1 fw	windowing/TI	9	$\lambda = 0.1$: 10000/5000 $\lambda = 0.2$: 5000/5000 $\lambda = 0.9$: 5000/5000
solution 2[i]	1 fw	windowing/TI in 5 steps	3 per step	For each step: $\lambda = 0.1667$: 4000/4000 $\lambda = 0.5$: 4000/4000 $\lambda = 0.8333$: 4000/4000

[a] Results reported in Table 4 are the average of the runs listed here. The forward (fw) direction is defined as 1 → 2 Å and the backward (bw) direction as 2 → 1 Å. [b] The notation ($\lambda = x$:m/n) means the following: At a given value of $\lambda = x$ (fw or bw run) m simulation steps were used for equilibration and discarded; during the following n steps of dynamics, the ensemble average $\langle \partial U/\partial\lambda \rangle_\lambda$ was accumulated. [c] The time-step in all PERT simulations was 1 fs. [d] Flexible system. Degrees of freedom corresponding to overall rotation and translation of the system were removed, so the Jacobian factor is not included in the results.[18] [e] Flexible system. The harmonic oscillator with a bond length of 1 Å length was placed in the center of a periodic box filled with 125 CHARMM modified TIP3 water molecules[17] (box length of 15.5516 Å). Waters overlapping with the solute were deleted; in all simulations 122 water molecules were present. The Nosé constant temperature molecular dynamics algorithm[15] was used to keep the average temperature at 300 K. [f] Rigid system. The bond of the solute was kept rigid by SHAKE.[16] The constraint correction was calculated according to eq 29 of ref 10. All other details of the simulations are the same as for the flexible system. [g] The time-step in all BLOCK simulations was 2 fs. [h] Two diatomic molecules, one with 1 Å, the other with 2 Å bond length, were defined on top of each other and centered along the x-axis, cf. Figure 2c. The solute was fixed during the molecular dynamics simulations. A Berendsen thermostat was used for temperature control. As in the PERT simulations, 122 water molecules were present. [i] The total change in bond length was broken up into five increments: 1−1.2 Å, 1.2−1.4 Å, etc. The individual free energy differences were added to obtain the final result.

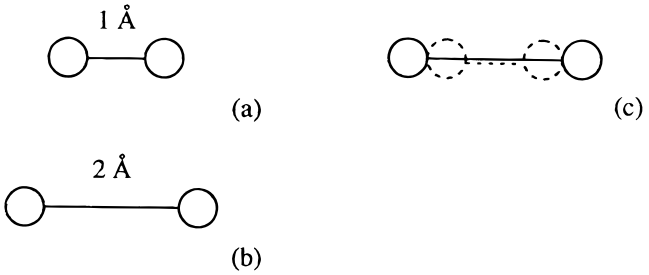

Figure 2. The hypothetical diatomic model system. (a) Initial state, (b) final state, and (c) the setup used in the dual topology simulations.

CHARMM (dual topology approach). The system is illustrated in Figure 2, which shows initial (a) and final state (b), as well as how the simulation was set up in the dual topology method c. In single topology, a flexible, as well as a rigid, solute was studied. Because of problems with keeping two flexible solutes aligned with each other, the solutes used in the dual topology calculations were not allowed to move, i.e., they were rigid. Thermodynamic integration was used in all calculations. Numerical integration of the $\langle \partial U/\partial\lambda \rangle_\lambda$ ensemble averages was carried out with the trapezoidal rule. An atom-based truncation scheme with a shifted electrostatic potential ($r_c = 7.5$ Å) was used; Lennard-Jones interactions were switched off between 6.5 and 7.5 Å. A dielectric constant of 1 was used. In the solution calculations periodic boundary conditions with the minimum image criterion were applied; they were carried out at constant volume. SHAKE was always applied to the water molecules[16] which were described by the CHARMM modified TIP3P model.[17] Additional details of the simulations are summarized in Table 1.

To interpret the results from these simulations as a free energy difference of solvation, a gas phase calculation is required, in principle. In the dual topology approach without scaling (see section 2b of the companion paper[10]), the bond terms do not change, so there is no need for a gas phase calculation. In the single topology method the only free energy contribution in the gas phase is a Jacobian factor arising from the bond length change (see section 2c of the companion paper[10]), and the Jacobian factor was calculated analytically.[18] However, we did a gas phase simulation to estimate the accuracy and precision of the free energy protocol for the flexible solute. It was set up so that the Jacobian factor free energy contribution is not included by removing the overall rotational and translational degrees of freedom from the simulation; see also ref 18.

2c. Free Energy Difference of Solvation between Ethane and Methanol. The free energy difference of solvation between ethane and methanol has become a benchmark for free energy simulation formalisms, as well as for the correctness of the implementation of computer programs.[19] Numerous simulations for this system have been published (e.g., refs 20−22). Many of these, using a variety of force fields and simulation protocols that might be considered too short by present standards, gave results (6.9 ± 0.1 kcal/mol[20,21]) in excellent agreement with the experimental value (6.9 kcal/mol[23]). It was noted recently, however, that the quantitative free energy difference of solvation is more sensitive to the details of the parametrization than thought earlier; in particular, the partial charges of the alkyl group(s) appear to play an important role.[24]

Simulations of the transformation were performed with an early version of the CHARMM 22 all-atom parameter set.[25] Both single and dual topology approaches were employed. Since the hybrid solutes used in the two cases are not the same, we discuss them separately.

Dual Topology Calculations. The BLOCK module of CHARMM[14] was employed for the simulations. When using this module, atoms that are not of the same type *and/or* do not have the same partial charge in the portion of the system that is transformed are defined separately. The parameters and partial charges used in the calculations are listed in Tables a−e in Supporting Information. Only the three hydrogens of the common methyl group in methanol and ethane are the same in the two end states. Ideal gas molecule end states were used so that the bonding parameters (bond stretching, bond angle, and

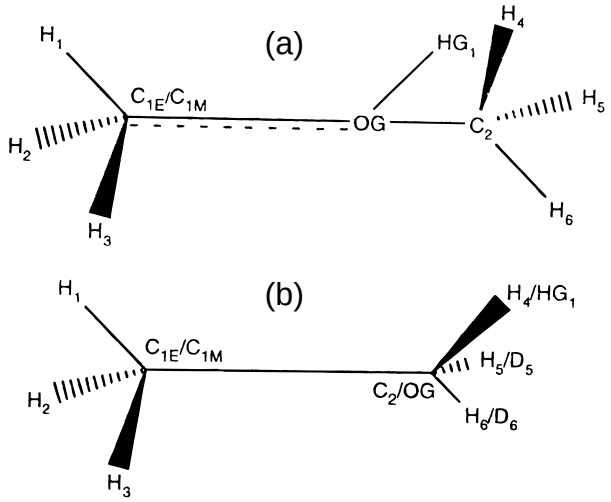

Figure 3. Ethane/methanol simulation. (a) The hybrid solute used in the dual topology simulations. (b) The hybrid solute used in the single topology simulations.

Urey−Bradley terms) were not scaled as a function of λ. Figure 3a depicts the hybrid solute used in the dual topology calculations. The simulation system is divided into three regions, which we denote as *environment* (waters and H1, H2, and H3), *reactant* (ethane part: C1E, C2, H4, H5, and H6) and *product* (methanol part: C1M, OG, and HG1) region. Reactant and product portions do not interact with each other during the simulation. The two C1 carbons (C1M and C1E) were assigned the same coordinates and were not allowed to move during the simulation. This may affect the vibrational degrees of freedom, but if C1E and C1M were free to move, there could be unphysical coupling between the two halves of the system. For example, a displacement of the ethane part that is caused by interaction with the solvent could affect the methanol part, which should not feel this particular interaction with solvent, and vice versa. However, the effect is expected to be small. The hydrogens H1, H2, and H3, which belong to the environment part, are "divalent"; i.e., each of them has bonds to C1E and to C1M. Since the bond parameters are not scaled (only simulations with an ideal gas molecule reference state were carried out), these hydrogens would have effective bond strengths equal to twice the normal value if the default parameters (cf. Table c in Supporting Information) were used. To avoid this, the force constant was divided by two for the bonds between H1, H2, H3, and C1E, C1M, respectively, and the same was done for angles of the type H1−C1E−H2, H1−C1M−H2, etc. There is an ambiguity regarding the H1−C1M-OG, H1−C1E−C2, etc., angles as the hydrogens should experience simultaneously a bond angle potential corresponding to methanol and ethane; furthermore, OG and C2 should experience the bond angle term for methanol and ethane, respectively. Since this is not possible (at least within BLOCK), the default parameters were used. A useful consequence of these artificially strong angle terms is that the C1E−C2 and C1M−OG bonds (and, therefore, reactant and product part of the hybrid) are restrained to stay close to each other throughout the simulation. The dihedral angle and 1−4 nonbonded intramolecular terms were scaled by the coupling parameter.

The simulations follow the alchemical paths in the thermodynamic cycle used to compute a free energy difference of solvation[26] (see Figure 1 of the companion paper[10]); i.e., $\Delta\Delta A_{solv} = \Delta A_4 - \Delta A_3$, where ΔA_3 and ΔA_4 are the free energy difference between ethane and methanol in the gas phase and in water, respectively. Since there are intramolecular dihedral angle and 1−4 nonbonded interactions which change when the properties of the hybrid solute are changed from ethane to methanol, the gas phase (ΔA_3) contributes to $\Delta\Delta A_{solv}$. The technical details of the simulations are summarized in Table 2a.

Single Topology Calculations. The PERT module of CHARMM was used for the single topology calculation; Figure 3b depicts the hybrid molecule employed. Since the total number of atoms must be conserved (see section 2b of the companion paper[10]), dummy atoms are introduced; i.e., two hydrogen atoms of the second methyl group in ethane (H5 and H6) were mutated into dummy atoms (D5 and D6) in methanol. Dummy atoms did not have nonbonded interactions with the rest of the system; i.e., their van der Waals parameters and charges were set to zero. The force constants of the dihedral terms to dummy atoms were also set to zero so that only harmonic energy terms connected the dummy atoms to the rest of the system (see section 2b of the companion paper[10]). The remaining bonded energy terms were maintained; i.e., the respective bond, bond angle, and Urey−Bradley parameters were equal to the values of their C−H counterparts. Two parametrizations for the bond length to the dummy atoms were used: (i) The bond length was the same as in the corresponding original bond (dummy atom type D1, calculations PERT1) and (ii) the bond length was one-fifth of this value (dummy atom type D2, calculations PERT2). The second model is expected to reduce end-point problems.[6,27,28] All intrasolute bonded and nonbonded interactions were included in the free energy to evaluate the importance of these terms. Free energy simulations were made also for a system in which the bond lengths of the hybrid molecule were constrained to their equilibrium value by SHAKE[16] (PERT2C calculations). The relationship between the single topology calculations carried out is illustrated by the thermodynamic cycle depicted in Figure 4. The initial state (ethane) is the same in the PERT1, PERT2 and PERT2C calculations as no dummy atoms are necessary for its representation. The parameters, which are based on an early version of the new CHARMM all-atom force field,[25] are listed in Tables a−e in Supporting Information.

In all cases a gas phase calculation was required to complete the thermodynamic cycle (Figure 4). The degrees of freedom corresponding to overall rotation and translation of the molecule were not removed in the gas phase calculations so that the full Jacobian factor was included there, as well as in solution.[18] In the case of a flexible molecule, the gas phase free energy differences can be estimated without MD simulations by use of normal mode calculations. This provides a useful check on the accuracy of the gas phase simulations. The two end point structures ($\lambda = 0$ and $\lambda = 1$) were minimized, and a normal coordinate analysis was performed. The resulting normal-mode frequencies were used to estimate the (classical) vibrational free energies of ethane and methanol (including the dummy atoms). If the corresponding particles in the initial and final state have the same mass, the contribution of the kinetic energy cancels out of the vibrational free energy difference; alternatively, the configurational contribution to the free energy difference can be calculated with the techniques developed in ref 29. Since in the present case masses were not changed, the entire vibrational free energy is used in calculating $\Delta\Delta A$. The difference in energy between the minimized structures of the initial and final state, i.e., ethane and methanol, has to be added to this vibrational free energy difference, as well as a term accounting for the change in moment of inertia, $\Delta A_{m.o.i} = -kT \ln(|\mathbf{I}_f|/[\mathbf{I}_i|)^{1/2}$, where $|\mathbf{I}|$ is the determinant of the moment of inertia tensor.[18] For a

TABLE 2: Summary of Dual and Single Topology Simulations for the Ethane/Methanol System[a]

(a) Single Topology

	no. of runs[b]	protocol	no. of λ-values used	details of protocol pertinent to free energy simulations[c]
gas phase	3 fw/3 bw	windowing/TI	10	$\lambda = 0.05$: 20 000/10 000 $\lambda = 0.15$: 10 000/10 000 $\lambda = 0.95$: 10 000/10 000
solution[d]	2 fw	windowing/TI	16	$\lambda = 0.05$: 5 000/60 000 $\lambda = 0.15$: 5000/60 000 $\lambda = 0.25$: 5 000/40 000 $\lambda = 0.75$: 5 000/40 000 $\lambda = 0.8125$: 10 000/40 000 $\lambda = 0.8375$: 10 000/40 000 $\lambda = 0.9875$: 10 000/40 000

(b) Single Topology[e]

	no. of runs[f]	protocol	details of protocol[g]	thermostat
gas phase				
PERT1	10 fw/10 bw	slow growth	10 000/100 000	multiple Nosé[h]
PERT2		slow growth	20 000/200 000	multiple Nosé[h]
PERT2C		slow growth	20 000/200 000	Langevin[i]
solution[j]				
PERT1	1 fw/1 bw	stepwise slow growth[k]	10 000/20 000 (per step)[l]	Nosé (ref 15)
PERT2	1 fw/1 bw	stepwise slow growth[k]	10 000/20 000 (per step)[l]	Nosé (ref 15)
PERT2C[m]	1 fw/1 bw	stepwise slow growth[k]	10 000/20 000 (per step)[l]	Nosé (ref 15)

[a] All calculations were performed with a slightly modified form of the BLOCK module in the CHARMM program[14] in which bond, bond angle and Urey−Bradley terms were not scaled as a function of λ. Numerical integration of the $\langle \partial U/\partial \lambda \rangle_\lambda$ ensemble averages was carried out with the trapezoidal rule. The mass of all solute hydrogens was set to 10 amu which improves convergence of the gas phase simulations. The temperature of the system was kept at an average value of 300 K with a Nosé thermostat.[15] An atom-based truncation scheme with a shifted electrostatic potential ($r_c = 7.5$ Å) was used; Lennard-Jones interactions were switched off between 6.5 and 7.5 Å. Nonbonded interactions were computed including 1−4 pairs; a dielectric constant $\epsilon = 1$ was used. [b] Results reported in Table 5 are the average of the number of runs listed here. The forward (fw) direction is defined as ethane → methanol; the backward (bw) direction as methanol → ethane. [c] The notation ($\lambda = x$:m/n) means the following: At a given value of λ (fw or bw run) m steps of simulation were used for equilibration and discarded; during the following n steps of dynamics, configurations were saved every 5 steps and used to calculate $\langle \partial U/\partial \lambda \rangle_\lambda$ in a postprocessing step. [d] The hybrid solute was placed in a preequilibrated box of 125 CHARMM modified TIP3P waters[17] (boxlength = 15.56 Å). Waters overlapping with the solute were deleted; in all simulations, 122 water molecules were present. SHAKE[16] was applied to the TIP3P water molecules. All simulations were carried out at constant volume. [e] All calculations were carried out with the PERT module of CHARMM.[14] The mass of all solute hydrogens was set to 10 amu. The temperature of the system was kept at an average value of 300 K; the thermostats used are listed in the table. An atom-based truncation scheme with a shifted electrostatic potential ($r_c = 7.5$ Å) was used; Lennard-Jones interactions were switched off between 6.5 and 7.5 Å. 1-4 nonbonded interactions were included; a dielectric constant $\epsilon = 1$ was used. [f] Results reported in Tables 6a−c are the average of the runs listed here. The forward (fw) direction is defined as ethane → methanol; the backward direction (bw) as methanol → ethane. [g] The notation m/n means that the system was equilibrated for m steps at the initial value of λ; the alchemical mutation from initial to final state was carried out over the following n steps using the slow-growth method. $\langle \partial U/\partial \lambda \rangle_\lambda$ was calculated on the fly and summed up to give ΔA for the step. [h] A separate Nosé heat bath was coupled to each atom of the hybrid solute.[15] [i] Temperature was controlled by Langevin dynamics; a friction coefficient of 50 ps^{-1} was used. [j] The hybrid solute was placed in a pre-equilibrated box of 125 CHARMM modified TIP3P waters[17] (box length = 15.56 Å). Waters overlapping with the solute were deleted; in all simulations 122 water molecules were present. SHAKE was applied to the water molecules.[16] All simulations were carried out at constant volume. [k] The interval $0 < \lambda < 1$ was broken into 10 steps: $0 < \lambda < 0.1$, 0.1, $\lambda < 0.2$, etc. For each of these steps, the system was (re)equilibrated for 10000 steps at the initial λ value; then λ was changed from initial to final value using the slow-growth protocol.[l] At the respective first step of a fw/bw run ($\lambda = 0$ or 1), 20000 instead of 10000 steps were used for equilibration. [m] In the PERT2C simulations constraints were applied to all bond terms of the solute. The resulting constraint correction to the free energy was calculated according to eq 29 of the companion paper.[10]

more detailed discussion, see section 3b of ref 18, in particular eqs 14 and 15.

In both the gas phase and solution simulations, a slow-growth protocol was employed. Slow-growth simulations have been criticized because, in principle, they do not sample equilibrium configurations at any instantaneous value of λ.[30] However, it has been shown that the results from a slow-growth calculation are upper and lower bounds to the free energy difference, which justifies the use of this protocol if the results are interpreted accordingly.[31] Furthermore, in solution the range $0 < \lambda < 1$ was broken up into 10 steps. While a slow-growth protocol was used to calculate the free energy difference for each of these steps, the system was allowed to reequilibrate between steps. The overall results from the forward and backward run are lower and upper bounds to the free energy difference, respectively. Moreover, the upper and lower bounds for each of the subintervals was also estimated from the simulations.[31] Additional technical details are summarized in Table 2b.

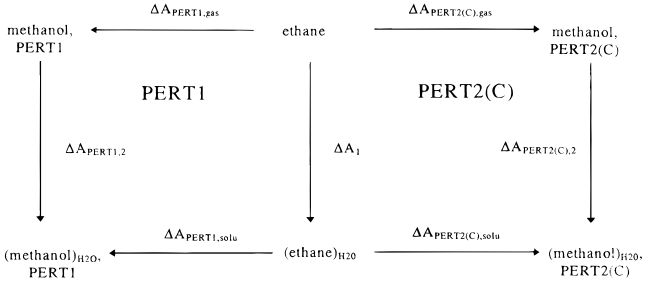

Figure 4. Thermodynamic cycles illustrating the single topology calculations for the free energy difference of solvation between ethane and methanol. The initial state (ethane) is the same for PERT1 and PERT2; the differences between the two protocols is discussed in the text. The PERT2 calculations were carried out for a flexible solute (PERT2), as well as for a solute where bond lengths were constrained to their parameter value (PERT2C); the symbol PERT2(C) indicates that this part of the diagram applies to both PERT2 and PERT2C.

TABLE 3: (a) Results for the One-Dimensional Atom, Diatomic Model System along the Paths Depicted in Figure 1.[a] (b) Free Energy Changes for a Change in Bond Length in the One-Dimensional Atom, Diatomic Model System as a Function of the Force Constant of the Bond.[r] (c) Free Energy Changes as a Function of the Force Constant Calculated with the Different Methods[s]

(a) Results for the System along the Paths Depicted in Figure 1

	single topology			dual topology (d.t.1)[b]			dual topology (d.t.2)[c]			exact[d]
	total	bond	vdW	total	bond	vdW	total	bond	vdW	total
ΔA_1	1.40		1.40	1.40		1.40	1.40		1.40	1.40
ΔA_2	−0.12		−0.12	−0.12		−0.12	−0.12		−0.12	−0.15
$\Delta\Delta A = \Delta A_2 - \Delta A_1$	−1.52		−1.52	−1.52		−1.52	−1.52		−1.52	−1.55
ΔA_3	−0.35	−0.35		0.00		0.00	−0.35	−0.35		−0.35[e]
ΔA_4	−1.92	−1.92		−1.56		−1.56	−1.87	−0.32	−1.55	−1.90[e]
$\Delta\Delta A = \Delta A_4 - \Delta A_3$	−1.57	−1.57		−1.56		−1.56	−1.52	0.03	−1.55	−1.55
$\Delta A_5 = -\Delta A_1 + \Delta A_3$ $= \Delta A_4 - \Delta A_2$[f]	−1.75	−1.40	−0.35	−1.41		−1.41	−1.72	−0.32	−1.40	−1.75[e]

(b) Free Energy Changes for a Change in Bond Length

				$\Delta A_{d.t.2}$[j]					
K	ΔA[g]	$\Delta A_{s.t.}$[h]	$\Delta A_{d.t.1}$[i]	total	bond	vdW	ΔA_{cons}[k]	ΔA_{cons}[l]	r_{oi}/r_{of} in Å[m]
1000.0	−1.58	−1.59	−1.58	−1.56	0.02	−1.58	−1.59	−1.58	1.522/0.300
260.0	−1.56	−1.56	−1.56	1.52	0.03	−1.55	−1.59	−1.54	1.513/0.300
80.0	−1.49	−1.49	−1.49	−1.42	0.04	−1.46	−1.59	−1.45	1.487/0.301
10.0	−1.16	−1.14	−1.13	−1.03	−0.13	−0.90	−1.59	−0.99	1.352/0.305

(c) Free Energy Changes as a Function of the Force Constant

			$\Delta A_{d.t.2}$[p]			
K	$\Delta A_{s.t.}$[n]	$\Delta A_{d.t.1}$[o]	total	bond	vdW	ΔA[q]
250 → 125	−0.24	−0.03	−0.25	−0.21	−0.04	−0.24
17 → 8.5	−0.36	−0.14	−0.40	−0.16	−0.23	−0.36

[a] All values in kcal/mol. K (force constant) in kcal/(mol Å^2). [b] Dual topology simulations in which bond terms were not scaled by coupling parameter λ. [c] Dual topology simulations in which bond terms were scaled by λ up to a limiting value as described in section 2d of the companion paper[10]; computations carried out using Mathematica[13] as described in section 2a. [d] Obtained by evaluating the configuration integral numerically using Mathematica.[13] [e] Compare only with single topology or d.t.2 simulations as d.t.1 does not include vibrational free energy contributions (see text). [f] See text (section 3a). [g] Calculated by direct numerical integration of the partition function of initial and final state using Mathematica.[13] [h] Simulation results using single topology, $\Delta A = \Delta A_{bond}$. [i] Simulation results obtained using dual topology without scaling of bond term, $\Delta A = \Delta A_{vdW}$. [j] Simulation results obtained using dual topology; bond terms were scaled to limiting value of λ as described in section 2d of the companion paper[10]; computations were carried out using Mathematica[13] as described in section 2a. [k] Analytical result for a constrained bond using the parameter value of the bond length to calculate the nonbonded interaction; in this case $\Delta A = U(r_{B1-B2} = 0.3$ Å$) - U(r_{B1-B2} = 1.526$ Å$)$. [l] Analytical result for a constrained bond assuming that the bond length corresponds to the minimum energy at the respective end point (cf. text). [m] Bond length corresponding to the minimum energy for the flexible system in the initial and final state, respectively. [n] Single topology simulation result; $\Delta A = \Delta A_{bond}$. [o] Dual topology simulation result in which the bond term was not scaled by λ; $\Delta A = \Delta A_{vdW}$. [p] Dual topology simulation result in which the bond term was scaled to a limiting value of λ as described in section 2d of ref 10 computations were carried out using Mathematica[13] as described in section 2a. Free energy components are listed explicitly. [q] Calculated by direct numerical integration of the partition function of initial and final state using Mathematica.[13] [r] The bond length r is changed from 1.526 to 0.3 Å. The case of a flexible bond term is compared with two ways of calculating the change in free energy for a rigid (=constrained) bond. [s] All free energy differences are in kcal/mol.

3. Results

3a. One-Dimensional Model System. The results corresponding to the paths depicted in Figure 1 are listed in Table 3a. Additional calculations aimed specifically at identifying coupling between nonbonded and bonded interactions are summarized in Tables 3b and c. For all transformations (Tables 3a−c), the free energy differences obtained from molecular dynamics simulations using either a single or dual topology methodology without scaling of the bond term (ideal gas molecule end state) are compared to those obtained from the numerical integration of the partition function. Furthermore, dual topology calculations with scaling of the bond term as described in section 2d of the companion paper[10] and section 2a were carried out using Mathematica.[13] No error bars are given for the simulations, as only one simulation was carried out for each case and method. However, the good agreement between the results obtained with the various methods (numerical integration of the partition function, single and dual topology computations) is indicative of the high accuracy and precision of the results.

The results summarized in Table 3a correspond to the thermodynamic cycle depicted in Figure 1 which can be interpreted as defining a free energy difference of solvation, $\Delta\Delta A_{solv} = \Delta A_4 - \Delta A_3 = \Delta A_2 - \Delta A_1$ (see section 2a). To compare the results of the concerted path ΔA_5 with alchemical (ΔA_3 and ΔA_4) and chemical (ΔA_1 and ΔA_2) paths, one has to add or subtract a contribution (e.g., $\Delta A_5 = -\Delta A_1 + \Delta A_3$). The free energy difference of "solvation" computed along chemical and alchemical paths agrees for all three computational methods (single topology, dual topology with (d.t.2) and without scaling of the bond term (d.t.1). However, there are systematic differences for the single free energy differences in the thermodynamic cycle, in particular for the alchemical paths ΔA_3 and ΔA_4, as well as for the concerted path ΔA_5.

In the single topology calculations, the free energy differences consist solely of a van der Waals contribution along the chemical paths and a bond contribution along the alchemical paths. Since only one component of the potential energy function is changed along each of these paths (the van der Waals interaction along the chemical paths and the bond term along the alchemical paths), no decomposition is required. The free energy difference for the additional path (ΔA_5), in which all parameters describing the system (force constant, equilibrium bond length, and van

der Waals parameters σ and ϵ) were changed in a linear concerted process, consists of both bond and van der Waals components, which are different as they must be, from the components found for analogous changes in the potential energy along the chemical and alchemical paths. The free energy differences for the chemical paths (ΔA_1 and ΔA_2), as well as the nonbonded free energy component on which it is projected, is the same in all three methods (s.t., d.t.1, and d.t.2). This is in accord with the considerations of section 2e of the companion paper[10] and reflects the fact that along the chemical paths the three methods study identical processes (the transfer of a molecule from the gas phase into solution). For the alchemical paths, two cases have to be distinguished in dual topology. If the bond potential energy term is not scaled (d.t.1), one finds $\Delta A_{3,s.t.} \neq \Delta A_{3,d.t.1}$, $\Delta A_{4,s.t.} \neq \Delta A_{4,d.t1.}$, and, similarly, $\Delta A_{5,s.t.} \neq \Delta A_{5,d.t.1}$. Only van der Waals components are obtained; the alchemical gas phase free energy difference (ΔA_3) is zero, and $\Delta A_{4,d.t.1} = \Delta\Delta A_{solv.}$. If the bond term is scaled (d.t.2); i.e., the initial bond is broken and the final bond is formed, there is a bonded contribution along the alchemical paths, and the total free energy differences (ΔA_3 and ΔA_4) are the same as in the single topology calculation. However, the respective bond and van der Waals components for ΔA_4 and ΔA_5 obtained in the s.t. and d.t.2 computations differ.

The reason for the difference in the results obtained for the alchemical paths is that different initial and final states are employed in the three methods. In the single topology calculations (for alchemical paths), molecule B1−B2^i is changed into B1−B2^f, with a corresponding, nonzero free energy difference $\Delta A_{3,s.t.}$ in the gas phase. Both $\Delta A_{3,s.t.}$ and $\Delta A_{4,s.t.}$ contain a vibrational contribution due to the change in force constant. In the dual topology calculations, where both molecules are present at all times, the results depend on the treatment of the bond term. If the bonds are not scaled (d.t.1), the hybrid molecule B1−B2^i/B1−B2^f remains the same in the gas phase and so $\Delta A_{3,d.t.1} = 0$. Similarly, in the calculation of the free energy difference in solution, the intramolecular energy terms are not changed; therefore, the free energy difference between the two molecules themselves is not included. In the d.t.1 method, the free energy difference caused by the change in intramolecular energy terms (equilibrium bond length and force constant) are omitted. This contrasts with the d.t.2 method, where the bond terms are scaled to a limiting value of the coupling parameter (see section 2d of the companion paper[10]) so that the vibrational contribution to the free energy difference is included. The overall results for ΔA_3, ΔA_4, and ΔA_5 agree with those of single topology method, in contrast to those from d.t.1, but different free energy components are obtained.

The comparison of the free energy components obtained for ΔA_3 and ΔA_4 by single topology and d.t.2 calculations makes clear the physical origin of the bond free energy components found in single topology calculations; see also the theoretical discussion in section 2c of ref 10. The gas phase free energy difference ΔA_3 is due to a change in vibrational frequency. Since the force constant of the bond term is different, there is a vibrational bond contribution. As this system is restricted to one dimension, there is no Jacobian factor contribution due to the change in bond length.[18] Both single topology and d.t.2 give the same free energy component for ΔA_3. The free energy difference in solution ΔA_4 contains the same vibrational contribution as the gas phase, but since the nonbonded interaction with the third particle changes when the bond length between B1 and B2 is altered, there is also a potential-of-mean-force-type contribution. This is evident from a comparison of s.t. and d.t.2 free energy components: While $\Delta A_{4,s.t.}$ is equal to the bond free energy component, $\Delta A_{4,d.t.2}$ consists of a bond component of about equal magnitude to that obtained in $\Delta A_{3,d.t.2}$ plus a nonbonded component. The pmf-type free energy contribution appears as a nonbonded free energy component in dual topology. The same conclusion can also be deduced from a comparison of the single topology and the d.t.1 result. As discussed, $\Delta A_{4,d.t.1}$ does not contain the (vibrational) free energy difference between B1−B2^i and B1−B2^f. From Table 3a we find $\Delta A_{4,d.t.1} = -1.56$ kcal/mol, which originates from the change in nonbonded interactions of the solute B1−B2 with the third particle B3. Upon subtracting $\Delta A_{4,d.t.1}$ from $\Delta A_{4,s.t.}$, one finds a difference of −0.36 kcal/mol, which must be caused by vibrational contributions. This is practically identical to the bond component obtained with d.t.2 (−0.32 kcal/mol). An analogous analysis can be made for the components of $\Delta A_{5,d.t.2}$, where there is a vibrational bond contribution of −0.35 kcal/mol.

The free energy differences of solvation $\Delta\Delta A_{solv}$ calculated with the three different methods (s.t., d.t.1, and d.t.2) are in excellent agreement. Since single topology and dual topology with scaling of bonded terms (d.t.2) include all contributions from the change in the bond term, this agreement for $\Delta\Delta A_{solv}$ is expected for simulations that have converged (s.t.) or computations that avoid convergence problems (d.t.2). However, the individual free energy differences in d.t.1 omit the vibrational contribution. Thus, it is important to confirm that no contribution to the physically meaningful quantity $\Delta\Delta A_{solv}$ is omitted for this case (see also section 2e of the companion paper[10]). Along the chemical paths, any influence of the nonbonded terms on the bonded terms (e.g., change in equilibrium geometry) shows up indirectly in the nonbonded free energy components. The same is true in dual topology simulations that do not scale bonded terms by the coupling parameter. The Boltzmann distribution used in the solution calculation ($\Delta A_{4,d.t.1}$) includes solvent induced changes in equilibrium geometry or vibrational behavior and, therefore, any net contribution to the free energy difference of solvation. This makes clear that, despite the different realizations of the alchemical paths (ΔA_3 and ΔA_4), *none* of the three approaches omits a contribution that is relevant to $\Delta\Delta A_{solv}$. Since the free energy is a state function, this result is to be expected. Nevertheless, we emphasize this finding and its confirmation by the results of the model calculations. It makes clear that one is free to choose the methodology that leads to the best behavior of the simulation and provides the most insight into the meaning of the results.

For the thermodynamic cycle depicted in Figure 1, $\Delta\Delta A_{solv}$ is caused almost exclusively by the change in nonbonded interactions between the solute B1−B2^i/B1−B2^f and the solvent B3. For the chosen interaction parameters, any effects of coupling between the bond term of the solute and the nonbonded interactions are smaller the precision of the calculated results.

To determine whether nonbonded interactions can have a measurable effect on the vibrational contribution, we consider first the results of calculations in which the bond length was changed from 1.526 to 0.3 Å as before, but neither the force constant nor the van der Waals parameters were altered (Table 3b). This free energy difference was calculated for four different force constants (K = 1,000, 260, 80, and 10 kcal/(mol Å^2)), ranging from an atypically strong force constant (K = 1000 kcal/(mol Å^2)) to one that is atypically weak (K = 10 kcal/(mol Å^2)) for a bond (or bond angle) term. In the absence of external forces, the free energy difference is zero as the force constant is not changed (eq 19 of the companion paper[10]). If

there were no influence of the nonbonded interaction on the harmonic oscillator term (vibrational degree of freedom), all four simulations would give the same result. For the weakest force constant ($K = 10$ kcal/(mol Å^2)) compared with the strongest force constant ($K = 1000$ kcal/(mol Å^2)) the difference is 0.42 kcal/mol, about one-third of the total free energy change. Thus, for the low force constant case the effect is important. However, for more realistic force constants, the differences in the results are quite small and would be impossible to discern in a typical free energy simulation with error bars of several tenths of a kcal/mole and more. The results for the two force constants used in the simulations summarized in Table 3a (260 and 80 kcal/(mol Å^2)) differ by 0.07 kcal/mol, a negligible amount.

Table 3b also includes the free energy differences for these four processes calculated under the assumption of a rigid bond. As described in the Methods section, two choices for the reference bond length are compared. The first approach, which uses the parameter values of the bonds as the target value for the constraint, gives the same result (−1.59 kcal/mol) in all four cases, as it should. The error, relative to a flexible bond, is significant only for the weakest bond ($K = 10$ kcal/(mole Å^2)). Use of reference values for the bond length that correspond to the minimum energy geometry of the flexible system improves the results, although the free energy difference for the weakest bond is still in error by a significant amount. This indicates that in most cases the influence of nonbonded interaction on the bond term will be small. Nevertheless, it is important to be aware that such an effect exists. Moreover, the van der Waals parameters chosen (section 2a) result in a repulsion of approximately 2.5 kT between particles 2 and 3 for the 1.526 Å bond length (the interaction is negligible for the 0.3 Å bond length). When the calculations were repeated with a stronger nonbonded interaction (approximately 10 kT between particles 2 and 3 for the 1.526 Å bond length), the effects of coupling between nonbonded (pmf-type) and vibrational contributions to the free energy difference increase significantly (results not shown).

Table 3b also includes the bond lengths corresponding to the minimum energy in initial and final state, respectively, which were used to compute the (free) energy difference according to eq 2. These can serve as a measure how strongly the nonbonded interaction influences the average geometry of the bond. The deviations from the equilibrium geometry are noticeable, even for "normal" force constants, such as 80 and 260 kcal/(mole Å^2), and the fairly weak van der Waals interaction chosen. This large effect is due in part to the one-dimensional arrangement of the model system, in which neither particle 2 nor particle 3 can avoid the unfavorable interaction.

The free energy differences for a change in force constant in the presence of nonbonded interactions are summarized in Table 3c. As the bond length parameter is not changed, there can only be a vibrational contribution and a coupling due to the van der Waals interaction between B2 and B3. The vibrational contribution of −0.21 kcal/(mol Å^2) is the same since the force constant changes by a factor of 2 in both cases and can be calculated analytically according to eq 19 of the companion paper.[10] Since the d.t.1 calculations omit vibrational free energy contributions, the d.t.1 results in Table 3c differ by exactly this value from the single topology and d.t.2 results.

In addition to the Mathematica results shown in Table 3c, the d.t.2 calculations (dual topology with scaling of bond terms to a limiting value as described in section 2d of the companion paper[10]) were also attempted with simulations. The results are generally less accurate than those with simulations based on single topology or d.t.1. A combination of reasons is responsible for this. First, we observed difficulties in achieving correct thermal equilibration for the two parts of the system although separate thermostats were used. Related problems were noticed by Wang and Hermans,[28] who coupled individual parts of a system to a Langevin heat bath to improve convergence of results. Also, the results were rather noisy so that the fits introduced errors in the overall result. Finally, in all systems studied (Tables 3a−c) there appear to be small systematic deviations between the d.t.2 Mathematica results and the analytical free energy differences. The d.t.2 algorithm makes use of a criterion for when a bond is broken (eq 40 of the companion paper[10]). At this point ($\lambda = \epsilon$), the "ideal gas" particles should not experience any nonbonded interactions; nevertheless, at $\lambda = \epsilon$ a weak van der Waals interaction can still be present. This could be corrected for by using a pair of coupling parameters; To describe the formation of a bond, the endpoints $\{\lambda_{vdW} = 0, \lambda_{bond} = \epsilon\} \rightarrow \{\lambda_{vdW} = 1, \lambda_{bond} = 1\}$ should be used.

The coupling effect when the force constant is altered originates from the (small) change in average bond length of the solute induced by nonbonded interaction with B3. This deviation becomes larger as the force constant becomes weaker, which is clearly reflected in the results (Table 3c). This indicates that the coupling should be included in the free energy methodology for alchemical changes in bond and bond angle terms with (very) low force constants in the presence of strong nonbonded interactions with the rest of the system. Significant coupling effects between intramolecular bond(ed) and intermolecular energy terms are evident in spectroscopic and theoretical studies.[32] This indicates that the form of the potential (e.g., representing bonds as harmonic rather than Morse oscillators) used in typical MM force fields[14,33,34] is not appropriate for including such effects.[32]

The calculations reported here, particular the results for the thermodynamic cycle depicted in Figure 1, make clear that single and dual topology methods utilize different end states to represent the alchemical paths usually followed in free energy simulations. In particular for dual topology simulations where bonded energy terms do not depend on the coupling parameter, different alchemical free energy differences are obtained for individual paths. This is clear from Table 3c where the results obtained with s.t. and d.t.2 agree well; $\Delta A_{d.t.1}$, on the other hand, differs systematically by the free energy difference for reducing the force constant by one-half in the absence of nonbonded interactions (0.21 kcal/mol). If this missing vibrational free energy contribution is added, the same results as in the s.t. and d.t.2 case are obtained. However, the results combined with the theoretical considerations of section 2e of the companion paper[10] demonstrate that all three methods correctly include coupling between vibrational and Jacobian factor contributions on one hand and nonbonded (e.g., solute−solvent) interactions on the other hand. Again, this is reflected in Table 3c, which shows that in the absence of nonbonded interactions, the free energy difference is a vibrational contribution resulting from the reduction of the force constant to one-half its original value. If the model calculation were expanded into a thermodynamic cycle, i.e., the computation of the "free energy difference of solvation" with the gas phase corresponding to the absence of nonbonded interactions, a gas phase free energy difference of 0.21 kcal/mol would be obtained in s.t. and d.t.2 (cf. footnote *s* in Table 3c) compared to zero for d.t.1 (since vibrational contributions are not included in this methodology). Therefore,

TABLE 4: Results of Free Energy Simulations for the Diatomic Molecule in Water[a]

single topology	ΔA_{gas}[b]	ΔA_{solu}	ΔA_{bond}[c]	ΔA_{vdW}	ΔA_{elec}	$\Delta\Delta A$[d]
run 1	−0.89(0.02)	−6.7(0.2)	−6.7(0.2)			−5.8(0.2)
run 2	−0.89(0.02)	−6.9(0.2)	−6.9(0.2)			−6.0(0.2)
dual topology[e]						
run 1		−5.8	0.0	3.4	−9.2	−5.8
run 2		−6.0	0.0	3.2	−9.2	−6.0

[a] See section 2b for the methods used. The bond length is changed from 1 to 2 Å; all other interaction parameters remain the same. Free energy differences are in kcal/mol. [b] ΔA_{gas} is the same for run 1 and run 2, and it is the sum of the values obtained separately for the Jacobian factor and the vibrational contribution. Theoretically, the only contribution is from the change in Jacobian factor, which was calculated analytically (eq 32) of the companion paper[10]; i.e., $\Delta A_J = -2k_BT \ln 2 = -0.83$ kcal/mol. Since the force constant of the solute is not changed, no vibrational contribution is expected; however, in actual simulations, a value of -0.06 ± 0.02 kcal/mol was found in quite good agreement with the theoretical result (cf. text). [c] Bond contribution for run 1, constraint correction for run 2. [d] For the single topology calculations, $\Delta\Delta A = \Delta A_{solu} - \Delta A_{gas}$, whereas for the dual topology case $\Delta\Delta A = \Delta A_{solu}$. [e] No gas phase calculation was carried out for the dual topology simulations (see text).

TABLE 5: Results of the Dual Topology Simulations (BLOCK) for the Free Energy Difference of Solvation between Ethane and Methanol[a]

	$\Delta A_{BLOCK,gas}$	$\Delta A_{BLOCK,solu}$	$\Delta\Delta A_{BLOCK,solv}$	$\Delta\Delta A^*_{BLOCK,solv}$[b]
total	5.3(0.1)	−3.4(0.4)	−8.7(0.5)	−8.5(0.4)
contributions				
dihedral	−0.4(0.1)	−0.5(0.1)	−0.1(0.2)	
VdW	0.1(0.0)	−2.2(0.2)	−2.3(0.2)	−2.2(0.2)
electrostatic	5.6(0.0)	−0.7(0.2)	−6.3(0.2)	−6.3(0.2)

[a] All free energy differences correspond to the transformation from ethane to methanol and are given in kcal/mol. The numbers in parentheses are the standard deviations obtained from averaging over the simulations. Since an ideal gas molecule end state was used in the dual topology simulations, there are no bond and bond angle free energy components and vibrational and Jacobian factor contributions are missing in both the gas phase and the solution calculation. [b] Free energy difference of solvation obtained with the neglect of the gas phase and intrasolute interactions (self-terms) in the solution calculation (cf. text).

identical double free energy differences would result for the three methods, demonstrating numerically that all three are suited for the study of double free energy differences defined by a thermodynamic cycle. In particular, coupling contributions are correctly included by all methods. The hypothetical double free energy difference that one can deduce from the results listed in Table 3c as just discussed is essentially such a coupling term; i.e., the vibrational contribution is the same in both cases; the van der Waals parameters themselves remain unchanged; and yet different values are obtained for the high and for the low values of the force constant. Although the vibrational contribution is omitted in the d.t.1 calculations, the same overall result (double free energy difference) is obtained; i.e., the coupling contribution is taken into account.

3b. Diatomic Model System. Table 4 lists the results obtained from simulations of a bond length change for a diatomic molecule in water with a charge of $\pm 0.5e$ on each atom. As was the case for the one-dimensional model systems (section 3a), we are interested in the *physical* origin of the computed free energy difference. Since this system is unrestricted; i.e., it can move in all three spatial dimensions, the Jacobian factor contributions, which were absent in the one-dimensional model systems, have to be taken into account. The details of the simulations are described in Table 1 and section 2b. The single topology (PERT) simulations labeled run 1 and run 2 differ only in the treatment of the bond term of the solute: In run 1 it is a flexible harmonic oscillator term; in run 2 it is constrained. In both dual topology simulations, a rigid solute is used (cf. section 2b and Table 1), but in run 1 the bond length is changed in one step whereas there are intermediate steps in run 2. In accord with the analysis of section 2c and e of the companion paper,[10] the free energy change can be seen to be projected on the bond term in the single topology calculation; i.e., there is no van der Waals or electrostatic contribution, and on the nonbonded energy terms (van der Waals and electrostatic) in the dual topology calculation (i.e., there is no bond contribution). The meaning of the free energy components listed in Table 4 is discussed in detail below. The single and dual topology results, with the latter determined by two different protocols, are in good agreement. This indicates that the error in these calculations is low.

As for the one-dimensional model (section 3a) and as described in the Theory section of the companion paper,[10] several contributions to the apparent bond free energy component (which for this simple model is equal to the full free energy difference) have to be distinguished in the single topology result. No vibrational contribution to the free energy difference is expected, as only the bond length (and not the force constant) is changed. Gas phase simulations using the relatively short protocol described in Table 1 yield $\Delta A_{gas} = -0.06 \pm 0.02$ kcal/mol, which is included in the double free energy difference $\Delta\Delta A$ in Table 4. This vibrational gas phase contribution, which is close to the theoretical value of zero, indicates that the protocol used is sufficient to describe the change in bond free energy (vibrational contribution) with a computational error of less than 0.1 kcal/mol. It is likely that the same accuracy is obtained in the solution simulation. The overall error for the single topology solution calculations was found to be 0.2 kcal/mol (Table 4) on the basis of standard deviation of the average results of the forward and backward runs. Since this is a three-dimensional system and the diatomic molecule can rotate freely in solution and in the gas phase, there is a Jacobian factor contribution due to the change in bond length[18,28] of $-2k_BT \ln 2 = -0.83$ kcal/mol. In the gas phase result listed in Table 5 this analytical result is added to the calculated value of -0.06 kcal/mol (see above) because overall translation and rotation are removed in the calculation. In the solution calculations, the Jacobian factor contribution is automatically accounted for in the free energy simulation.[18]

The major component of the free energy difference in solution ΔA_{solu} as obtained by the single topology methodology, is a pmf-type contribution. It reflects the change in nonbonded interactions as a result of the change in the size and, hence, the dipole moment of the solute. Since none of the nonbonded

parameters are altered, no explicit nonbonded free energy components are obtained in the single topology calculation. Table 4 contains single topology results from a calculation using a flexible bond (run 1) and one where the bond of the solute is constrained (run 2). The difference between the two results (0.2 kcal/mol) could reflect coupling between the vibrational degree of freedom and the solute−solvent interactions (cf. the analogous results for the one-dimensional case in Table 3b). However, as the statistical error of the calculation is of the same magnitude, i.e., 0.2 kcal/mol, it is not certain that this is the origin of the difference.

In both dual topology simulations, the solute is held fixed. This eliminates the need for a gas phase simulation, and there is no Jacobian factor contribution in the solution simulations; in fact, ΔA_{solu} equals the double free energy difference of solvation $\Delta\Delta A$. The dual topology calculations make use of a path in which molecule 1 is desolvated and molecule 2 is solvated simultaneously and reflect how the solute−solvent interaction changes as the bond length of the solute is increased or decreased. There is a stabilizing electrostatic interaction (−9.2 kcal/mol) with water due to the increase in bond length (and dipole moment) which is partially offset by a destabilizing van der Waals contribution (3.3 kcal/mol) due to the larger size of the solute. These two opposing contributions are analogous to those observed in the free energy difference of solvation between Br^- and Cl^-.[35]

The above results demonstrate that the solvation free energy difference of approximately −6 kcal/mol is the result of the increase in the dipole moment of the solute when the bond length changes from 1 to 2 Å. Comparison of the single and dual topology methods is useful for a better understanding of the physical origin of the free energy change. Since the charges of the solute did not interact, there is no intramolecular nonbonded contribution. The change of bond length in the gas phase results in a Jacobian factor contribution, provided the molecule is free (see section 2c of the companion paper[10]). In solution, the single topology methodology leads to a large pmf-type contribution. Its physical origin is made explicit by comparison with the dual topology calculations, which show that the change in bond length alters both the van der Waals and electrostatic interaction with the solvent.

3c. Ethane to Methanol. To further illustrate the analysis of contributions of bonded terms to free energy changes, we consider the solvent effect on the ethane to methanol transformation, which is a realistic system that has been simulated as a test case by several other groups.[20−22] We first describe separately the dual and single topology results and then compare the components obtained with the two methodologies to obtain a clearer understanding of their significance. The average free energy difference of solvation $\Delta\Delta A_{solv}$ between ethane and methanol found with the various protocols discussed below (Tables 5−7) is −8.7 ± 0.2 kcal/mol. This result is rather far from the experimental value (−6.9 kcal/mol),[23] as well as from most calculations reported to date.[20−22] Since elaborate protocols with small statistical error were used in the present study, the difference cannot arise from the lack of convergence of the simulations. It is possible, though not likely, that the relatively small size of the water box (122 water molecules) with a relatively short cutoff radius plays a role. However, it has been shown recently that the result of the ethane to methanol free energy simulation depends strongly on small changes in the point charges used for methyl carbons and hydrogens.[24] It was found that a free energy difference of solvation between ethane and methanol of −7.9 kcal/mol is obtained for charges fitted to

TABLE 6: Results of the Single Topology Simulations ((a) PERT1,[a] (b) PERT2,[g] and (c) PERT2C[b]) for the Free Energy Difference of Solvation in Transforming Ethane into Methanol

(a) PERT1

	$\Delta A_{PERT1,gas}$[c]	$\Delta A_{PERT1,solu}$	$\Delta\Delta A_{PERT1,solv}$
total	5.7(0.1)	−2.8(−0.9)	−8.5(1.0)
contributions			
bond	1.8(0.0)	2.1(0.0)	0.3(0.0)
angle	0.7(0.0)	0.6(0.0)	−0.1(0.0)
Urey−Bradley	−0.8(0.0)	−0.8(0.0)	0.0(0.0)
dihedral	−0.5(0.1)	−0.5(0.0)	0.0(0.1)
VdW	0.1(0.0)	−2.4(0.5)	−2.5(0.5)
electrostatic	4.4(0.0)	−1.8(0.4)	−6.2(0.4)
estimate of contributions			
to bonded components[d]			
vibrational	0.35	0.35	0.0
Jacobian factor	0.26	0.26	0.0

(b) PERT2[g]

	$\Delta A_{PERT2,gas}$[e]	$\Delta A_{PERT2,solu}$	$\Delta\Delta A_{PERT2,solv}$
total	9.5(0.1)	0.6(0.1)	−8.8(0.2)
contributions			
bond	7.4(0.0)	6.6(0.2)	−0.8(0.2)
angle	0.6(0.0)	0.6(0.0)	0.0(0.0)
Urey−Bradley	−1.7(0.0)	−1.7(0.0)	0.0(0.0)
dihedral	−0.4(0.0)	−0.4(0.0)	0.0(0.0)
VdW	−0.2(0.0)	−1.6(0.1)	−1.4(0.1)
electrostatic	3.7(0.0)	−2.9(0.1)	−6.6(0.1)
estimate of contributions			
to bonded components[d]			
vibrational	0.35	0.35	0.0
Jacobian factor	4.10	4.10	0.0

(c) PERT2C[b]

	$\Delta A_{PERT2,gas}$	$\Delta A_{PERT2,solu}$	$\Delta\Delta A_{PERT2,Csolv}$
total	9.1(0.1)	0.2(0.1)	−8.9(0.2)
contributions			
bond (=constraint[f])	8.5(0.0)	7.6(0.1)	−0.9(0.2)
angle	0.7(0.0)	0.6(0.0)	−0.1(0.0)
Urey−Bradley	−3.1(0.0)	−3.1(0.0)	0.0(0.)
dihedral	−0.4(0.0)	−0.4(0.1)	0.0(0.1)
VdW	−0.2(0.0)	−1.7(0.1)	−1.5(0.1)
electrostatic	3.7(0.0)	−2.9(0.0)	−6.6(0.1)
estimate of contributions			
to bonded components[d]			
vibrational			
Jacobian factor	4.10	4.10	0.0

[a] Calculation in which the bond lengths to dummy atoms were unchanged dummy atom type D1. All free energy differences are in kcal/mol. The values listed are the average of at least two simulations (cf. the description of the simulations, section 2c), and the values in parentheses are the corresponding standard deviation. [b] Calculations with reduced bond lengths to dummy atoms, dummy atom type D2. Bond lengths of the solute were constrained using SHAKE. All free energy differences are in kcal/mol. The values listed are the average of at least two simulations (cf. the description of the simulations, section 2c), and the numbers in parentheses are the corresponding standard deviation. [c] Using normal mode analysis, a gas phase free energy difference of 5.8 kcal/mol is obtained (this includes the contribution due to the change in moment in inertia). [d] These contributions to the bond(ed) free energy components listed above were estimated with the analytical formulas developed in section 2c of the companion paper.[10] Only contributions from changes in equilibrium bond lengths and bond stretching force constants are included; the neglect of contributions from bond angle and Urey−Bradley terms is discussed in the text. [e] Using normal mode analysis, a gas phase free energy difference of 9.6 kcal/mol is obtained (this includes a contribution due to the change in the moment of inertia). [f] Obtained as constraint correction using eq 29 of the companion paper.[10] [g] Calculation with reduced bond lengths to dummy atoms, dummy atom type D2. All free energy differences are in kcal/mol. The values listed are the average of at least two simulations (cf. the description of the simulations, section 2c), and the values in parentheses are the corresponding standard deviations.

TABLE 7: Comparison of Free Energy Differences of Solvation between Ethane and Method Obtained with the Various Protocols[a]

	contributions				
	$\Delta\Delta A_{\mathrm{BLOCK}}$	$\Delta\Delta A^*_{\mathrm{BLOCK}}$	$\Delta\Delta A_{\mathrm{PERT1}}$	$\Delta\Delta A_{\mathrm{PERT2}}$	$\Delta\Delta A_{\mathrm{PERT2C}}$
bond			0.3(0.0)	−0.8(0.2)	−0.9(0.1)[c]
angle			−0.1(0.0)	0.0(0.0)	−0.1(0.0)
Urey−Bradley			0.0(0.0)	0.0(0.0)	0.0(0.0)
dihedral	−0.1(0.2)		0.0(0.1)	0.0(0.0)	0.0(0.1)
VdW	−2.3(0.2)	−2.2(0.2)	−2.5(0.5)	−1.4(0.1)	−1.5(0.1)
electrostatic	−6.3(0.2)	−6.3(0.2)	−6.2(0.4)	−6.6(0.1)	−6.6(0.0)
total	−8.7(0.5)	−8.5(0.4)	−8.5(1.0)	−8.8(0.2)	−8.9(0.1)

[a] All free energy differences are in kcal/mol. The numbers in parentheses are the corresponding standard deviations. As all entries refer to free energy differences of solvation, the subscript solv used in Tables 5 and 6a−c was dropped. [b] See footnote *b* in Table 5. [c] Obtained as constraint correction using eq 29 of the companion paper.[10]

the electrostatic potential surface (EPS) from ab initio calculations on the 6-31G* level, compared to −12.2 kcal/mol (!) for charges derived from a Mulliken population analysis of the same quantum mechanical calculations. It was also pointed out that some of the early calculations may correspond to unconverged results, which fortuitously were in better agreement with experiment.[24] The Mulliken population analysis results in a negative partial charge on the methyl carbon (−0.471) and positive partial charges on the methyl hydrogens (0.157); the opposite result (positive partial charges on the carbon (0.027), negative partial charges on the hydrogens (−0.009)) is obtained with the EPS method. The parameters used in the present study (Supporting Information) fall between the two extremes compared in Ref 24, and so the value for the solvation free energy change seems reasonable.

The very similar results obtained with different simulation methodologies and protocols (which are presented and discussed individually below) provide strong evidence for the fact that -8.7 ± 0.2 kcal/mol is the correct free energy difference of solvation between ethane and methanol for the parameters and system size (water box) used here. Consequently, the detailed analysis of the results is expected to be meaningful, even if they do not reproduce the exact experimental value.

Dual Topology Simulations. Table 5 contains the results of the dual topology simulations that determine the free energy difference of solvation between ethane and methanol. A dual topology methodology employing an ideal gas molecule end state was used (see section 2c). As discussed in section 2b and d of the companion paper,[10] this implies that vibrational and Jacobian factor contributions arising from changes of the bonded energy terms (bond stretching, bond angle bending and Urey−Bradley terms) are omitted in both the gas phase and in the solution calculations. The free energy components obtained from a decomposition of the respective overall free energy difference are included in Table 6. The first three columns in Table 5 list the gas phase free energy difference, the result of the solvation calculation, and the double free energy difference of solvation $\Delta\Delta A_{\mathrm{solv}} = \Delta A_{\mathrm{solu}} - \Delta A_{\mathrm{gas}}$, respectively. The last column contains the double free energy of solvation calculated under the assumption that the self-terms in ΔA_{solu} and ΔA_{gas}, which in this case consist of dihedral and intramolecular nonbonded interactions, cancel from the thermodynamic cycle; we refer to this result as $\Delta\Delta A^*$. It is obtained from the calculation that was used to determine ΔA_{solu}. The results for $\Delta\Delta A$ and $\Delta\Delta A^*$ listed in Table 5 are essentially identical. The dihedral contribution of −0.1 kcal/mol to $\Delta\Delta A$, which is not included in $\Delta\Delta A^*$, is small and statistically not significant: it has the highest relative error (compared to its absolute magnitude), both in the gas phase and in solution.

Since bond and bond angle terms are not included in the calculated free energy (i.e., these terms were not scaled as a function of λ; see section 2d of the companion paper), neither step of the thermodynamic cycle (gas phase or solution) yields the true free energy difference between ethane and methanol. However, the correct solvation free energy difference is obtained since the same reference state is used for both ΔA_{gas} and ΔA_{solu}; i.e., the intrasolute energy terms are treated in the same manner in the gas phase and in solution. This behavior is analogous to that of the one-dimensional model systems discussed in section 3a.

The gas phase free energy difference is large and favors ethane relative to methanol. As can be seen from the component analysis, the source of this difference is the intramolecular electrostatic interaction. It arises from the repulsive interaction between the charges (0.09*e*) on the hydrogens in the methyl groups and the large charge (0.43*e*) on the hydroxyl hydrogen of methanol versus the smaller repulsion between the two sets of methyl hydrogens in ethane. Since the dual topology results do not contain any bond or bond angle contributions, the self-term consists of the electrostatic interactions plus the dihedral and van der Waals interactions. Both of the latter are relatively small in the gas phase. Although the electrostatic self-term is large in the gas phase, it essentially cancels from the double free energy difference as reflected by the good agreement between $\Delta\Delta A$ and $\Delta\Delta A^*$. Further, comparing ΔA_{solv} with ΔA_{gas}, we see that the electrostatic destabilization of methanol relative to ethane in the gas phase is more than compensated by the solute−solvent electrostatic interactions in solution. In addition, the van der Waals contribution in solution also favors methanol since it is slightly smaller than ethanol (cf. the parameters listed in Supporting Information).

The decomposition made here, which provides insights into the nature of the contributions to the free energy, depends on the consistent use of component analysis; i.e., the same local path is used for ΔA_{gas} and ΔA_{solu} (see Ref 36).

Single Topology Results. As described in the Methods section, several different single topology calculations are performed to make clear the role of dummy atoms. Table 6a contains the results for ethane to methanol(D1), $(\Delta)\Delta A_{\mathrm{PERT1}}$; Table 6b lists the analogous results for ethane to methanol(D2), $(\Delta)\Delta A_{\mathrm{PERT2}}$; and Table 6c gives the results for ethane to methanol(D2) simulations, in which all bond lengths of the solute were constrained by the use of SHAKE,[61] $(\Delta)\Delta A_{\mathrm{PERT2C}}$. The three columns of each table contain the gas phase free energy difference between ethane and methanol ΔA_{gas}, the corresponding free energy difference in solution ΔA_{solu}, and the free energy difference of solvation $\Delta\Delta A_{\mathrm{solv}} = \Delta A_{\mathrm{solu}} - \Delta A_{\mathrm{gas}}$, respectively. Both the overall (double) free energy differences, as well as the free energy components segregated according to type of interaction energy (bond, bond angle, etc.), are listed. In addition, the vibrational and Jacobian factor contributions due to all changes in bond stretching terms are reported. These last two

quantities are not additional free energy components; they are a part of the bonded free energy components. As derived in section 2c of the companion paper,[10] bonded free energy components, i.e., free energy components due to bond stretching, angle bending, and Urey−Bradley terms of the potential energy function, originate from a number of physical effects: vibrational, Jacobian factor, and pmf-type contributions plus coupling among them. On the basis of analytical formulas of the companion paper, it is straightforward to calculate the vibrational and Jacobian factor contributions of the bond stretching terms, and these are listed in Tables 6a−c. However, the simultaneous use of bond angle and Urey−Bradley energy terms for the angle bending degrees of freedom makes it difficult to do the same for these terms (cf. section 2c of the companion paper). Since a number of bond angle parameters change during the alchemical mutation of ethane into methanol, the calculation would be rather complex. While Jacobian factor contributions may be safely neglected as all equilibrium angles are within a very narrow range of ±4°, the force constants differ considerably between the initial and final states. However, in all instances when a weaker bond angle force constant is replaced by a stronger one, there is a corresponding loss of a Urey−Bradley term (e.g., ∠CT3−CT3−HA to ∠CT3−OH1−H or ∠HA−CT3−CT3 to ∠HA−CT3−OH1). Inserting the actual values (Table d in Supporting Information) into the first-order approximation to obtain an effective angle bending force constant $K_{\theta,\mathrm{eff}}$ for the bond angle and Urey−Bradley interaction energy, i.e., $K_{\theta,\mathrm{eff}} = K_\theta + K_{\mathrm{UB}}$ (section 2c of the companion paper[10]), one can deduce that the resulting vibrational contribution, which is proportional to $\ln(K^{\mathrm{f}}_{\theta,\mathrm{eff}}/K^{\mathrm{i}}_{\theta,\mathrm{eff}})$, is small because of these two canceling effects. In other cases, the force constants involved remain unchanged. The improved approximation outlined in section 2c of the companion paper[10] is quite involved in the case of such coupled degrees of freedom, as in a methyl group or the hydroxyl group plus two dummy atoms. Considering that the contributions are expected to be small and that the error in the component analysis due to statistical fluctuations may well be nonnegligible, we list only the unambiguous partial contributions from bond stretching degrees of freedom. Pmf-type contributions can be obtained only as the difference between the bonded free energy components from a decomposition and the sum of vibrational plus Jacobian factor contributions. Because of the problems in reliably determining the latter for this system, we estimate pmf-type contributions by comparing free energy components among the various simulation methodologies (dual vs single topology, PERT1 vs PERT2); this is described in detail below. The results of such comparisons are not exact, and so we do not include the values in Tables 6a−c.

The difference between methanol(D1), $(\Delta)\Delta A_{\mathrm{PERT1}}$, and methanol(D2), $(\Delta)\Delta A_{\mathrm{PERT2}}$, is the equilibrium bond length of the hydroxyl oxygen to dummy atom bonds, as well as the corresponding Urey−Bradley parameter. The total free energy differences in the gas phase and in solution are different in the two calculations since the endpoints of the simulations are different, although the (double) free energy difference of solvation agrees (within error bars), as it must. The differences between the PERT1 and PERT2 results (both gas phase and solution) are primarily caused by the Jacobian factors from the different bond length to dummy atoms. Following ref 18 and ignoring any vibrational and Jacobian factor contributions from bond angle/Urey−Bradley terms due to the complications just discussed, these can be estimated to be $-4k_{\mathrm{B}}T\ln 5 = -3.8$ kcal/mol, which coincides with the difference between $\Delta A_{\mathrm{PERT1,gas}} = 5.7$ kcal/mol and $\Delta A_{\mathrm{PERT2,gas}} = 9.5$ kcal/mol (see the respective first column of Tables 6a and b) and is quite close to the −3.4 kcal/mol difference between $\Delta A_{\mathrm{PERT1,solu}} = -2.8$ kcal/mol and $\Delta A_{\mathrm{PERT2,solu}} = 0.6$ kcal/mol (see second column of Tables 6a and b). Also, the omitted contributions from the coupled change in Urey−Bradley 1−3 distances are expected to be small (cf. discussion in previous paragraph).

In addition to the different Jacobian factor contribution to the free energy difference ($\Delta A_{\mathrm{PERT1,gas}}$ and $\Delta A_{\mathrm{PERT2,gas}}$), there is also a different pmf contribution that depends on whether the bonds to the dummy atoms remain the same (PERT1) or are shrunk/grown (PERT2). This affects the free energy components rather than the total free energy difference; i.e., bond, van der Waals, and electrostatic free energy components differ significantly between PERT1 and PERT2. One can distinguish between two pmf-type free energy contribution, one due to the intramolecular nonbonded interactions and the other due to solute−solvent interactions. In the gas phase only the former is present. It originates from the change of the intramolecular 1−4 interactions when the geometry of the system (bond lengths and bond angles) is altered in going from ethane to methanol. Combined with the Jacobian factor contribution, this explains the differences in bond, Urey−Bradley (which are coupled to the equilibrium bond lengths), van der Waals, and electrostatic free energy components for $\Delta A_{\mathrm{PERT1,gas}}$ and $\Delta A_{\mathrm{PERT2,gas}}$. The dihedral and bond angle components, which are treated identically in the two protocols, agree within the error bars of the calculation. In solution, there are additional pmf-type contributions from the change in solute−solvent interaction as a result of changes in the solute geometry (bond lengths and bond angles).

A careful comparison of the free energy components listed in Tables 6a and b makes it possible to distinguish pmf contributions caused by the intrasolute 1−4 nonbonded interactions from those caused by the solvent and to determine whether they arise from physical changes in bond and bond angle terms due to the transmutation of ethane into methanol or from the treatment of the unphysical bonds to dummy atoms. Such an analysis assumes that vibrational and Jacobian factor contributions cancel in the thermodynamic cycle, which is the case unless there is significant coupling with the nonbonded interactions. In the simple model systems studied so far (section 3a and b), coupling between vibrational and Jacobian factor contributions on the one hand and pmf-type contributions on the other hand was observed only in the case of untypically low force constants. For the present case, coupling cannot be excluded rigorously, but it is expected to be small relative to the term of interest. Looking at the bonded free energy components (bond, bond angle, and Urey−Bradley) of the free energy difference of solvation (third column in Tables 6a−c), we can immediately discard the Urey−Bradley component as a source of coupling as it is zero in all three cases. In the case of constrained bond terms (PERT2C) there are neither vibrational contributions nor coupling between Jacobian factor and pmf-type contributions. Since the bond component in the PERT2 result (−0.8 ± 0.2 kcal/mol) is the same (within error bars) as the bond or, more exactly, constraint component in the PERT2C result (−0.9 ± 0.1 kcal/mol), it can be identified as a pmf-type contribution; i.e., it is highly unlikely that any coupling contributions are contained in the PERT2 bond component (where, in principle, they cannot be ruled out). Finally, the angle component of the free energy difference of solvation is small in all three cases (≤0.1 kcal/mol), so that any resulting coupling will be negligible. Thus, coupling contributions, if any, will be neglected in the following analysis.

The gas phase by definition contains only intra-solute pmf-type contributions. In the PERT1 calculation there is no pmf-type contributions from the bonds to dummy atoms as their bond length is not changed. The formal bond, bond angle (and Urey−Bradley) contributions obtained in the gas phase mainly reflect the potential of mean force of the 1−4 intrasolute interactions, as well as (small) vibrational (0.35 kcal/mol) and Jacobian factor (0.26 kcal/mol) contributions. The two values from Table 6a for the vibrational and Jacobian factor contributions, respectively, include only the contributions for the bond stretching terms for the reasons outlined above. When one subtracts the gas phase free energy components from the solvation data; i.e., when one calculates the free energy of solvation, one is left with the potential of mean force contributions from the solvent. As seen in Table 6a we find 0.3 kcal/mol as the bond contribution and −0.1 kcal/mol as the bond angle contribution. These values are relatively small, but the bond and bond angle changes to which they correspond are small as well (≤0.15 Å difference in bond lengths and ≤4° for bond angles). In other cases larger effects could result.

Repeating this procedure for the PERT2 calculation, we find a pmf-type contribution due to the interaction with the solvent of −0.8 kcal/mol, which shows up as the formal bond contribution to $\Delta\Delta A_{\mathrm{PERT2}}$. The interpretation of the bond components of $\Delta\Delta A_{\mathrm{solv}}$ as being pmf-type contributions is supported by the observation that the differences in $\Delta\Delta A_{\mathrm{bond}}$ for the two systems (PERT1 and PERT2) are accompanied by a reciprocal change in the formal nonbonded contributions, mainly, $\Delta\Delta A_{\mathrm{vdW}}$. Thus, the bond components of $\Delta\Delta \mathrm{A}_{\mathrm{solv}}$ of PERT1 and PERT2 (third column, second row in Tables 6a and b, respectively), which correspond to the solvent induced potential of mean force contributions, differ by approximately 1 kcal/mol. Since the only difference between the PERT1 and the PERT2 simulations is the equilibrium bond length of the bonds to the two the dummy atoms, this difference of 1 kcal/mol is seen to be caused solely by the two bonds involving the dummy atoms. This demonstrates the influence which the treatment of the unphysical dummy atoms has on the free energy components.

Protocols in which the bond lengths to dummy atoms are very short (<0.5 Å vs normal bond length of >1.0 Å) have been suggested as a tool to improve the convergence of free energy simulations (cf. e.g., refs 28, 37, and 38. The overall result for the free energy of solvation obtained for the two parametrizations of the bonds to the dummy atoms (PERT1 vs PERT2) agrees within 0.3 kcal/mol in the present case. However, the uncertainty of the result is significantly larger for the PERT1 calculations relative to PERT2. The free energy components indicate where the errors originate. Comparing Table 6a with Tables 6b and c, we see that the main contribution comes from the solute−solvent van der Waals and electrostatic interactions. Most of the error can be traced to the interval $0.6 < \lambda < 1.0$; i.e., the methanol end of the simulation where the difference in the bond length to the dummy atoms is most important (data not shown). A similar analysis in the dual topology case (results not shown) also identifies the range $0.6 < \lambda < 1.0$ as the main source of error.

The results for the system with constrained bonds $(\Delta)\Delta A_{\mathrm{PERT2C}}$ are summarized in Table 6c. Bond free energy components are replaced by the corresponding constraint correction.[28] The free energy difference of solvation agrees well with the results for the flexible systems. As already pointed out, this confirms that for the ethane/methanol system the harmonic bond terms do not couple to nonbonded solute−solvent interactions. The overall gas phase and solution free energy differences deviate by 0.4 kcal/mol for the system with and without constraints (cf. Tables 6b and 6c). To interpret this result, we write the (gas phase) free energy of a polyatomic molecule within the rigid rotor, harmonic oscillator (RRHO) approximation as[18,29]

$$\mathrm{A} = -k_{\mathrm{B}}T\ln \frac{(2\pi k_{\mathrm{B}}T)^{3N/2-3}}{|F_{\mathrm{S}}|^{1/2}} \prod_i \mathrm{J}_{i\mathrm{S}} \qquad (3)$$

Introducing bond length constraints does not affect the Jacobian factors $\prod_i J_{is}$ in the above equation;[18] the force matrix F_S, however, is changed as degrees of freedom are removed. In other words, the constrained system lacks the vibrational bond contribution present in the flexible system due to the changes in the force constants between ethane and methanol (see parameters in Table 2). This contribution from the change in force constants of the harmonic bond stretch terms (computed under the assumption that the force constants which are different between ethane and methanol describe independent harmonic oscillators) has been listed in Tables 6a and b. The value of 0.35 kcal/mol (obtained using the force constants listed in Table c in Supporting Information) is in good agreement with the difference of 0.4 kcal/mol between the results for the flexible (Table 6b) and rigid systems (Table 6c). In addition, the bond/constraint and Urey−Bradley free components in the gas phase and in solution are not the same as those for the flexible and rigid calculations. In particular, there is a large difference for the Urey−Bradley free energy component. As one sees from Tables 6b and c, its absolute value increases from −1.7 kcal/mol for the flexible system (PERT2) to −3.1 kcal/mol for the rigid system (PERT2C), both in the gas phase and in solution. This more negative Urey−Bradley free energy component is almost exactly compensated by a more positive bond free energy component, 8.5 instead of 7.4 kcal/mol in the gas phase and 7.6 instead of 6.6 kcal/mol in solution, respectively. This behavior reflects the strong coupling between equilibrium bond lengths and Urey−Bradley free energy contributions (cf. the discussion of angle bending degrees of freedom in section 2c of the companion paper.[10])

Comparison of Single and Dual Topology Results. A summary of the results for the ethane−methanol system is presented in Table 7. The average free energy difference of solvation between ethane and methanol found with the various protocols is −8.7 ± 0.2 kcal/mol (average over all $\Delta\Delta A_{\mathrm{solv}}$ values listed in Table 7). The narrow range of the results (all within the estimated standard deviation) provides numerical confirmation that none of the approaches for calculating $\Delta\Delta A_{\mathrm{solv}}$ omits significant contributions. In particular, the two dual topology results ($\Delta\Delta A_{\mathrm{BLOCK}}$ and $\Delta\Delta A_{\mathrm{BLOCK*}}$) do not omit any relevant contributions from the double free energy of solvation, although they exclude vibrational and Jacobian factor contributions in the individual steps of the thermodynamic cycle. The components of $\Delta\Delta A_{\mathrm{BLOCK}}$ (dual topology) and $\Delta\Delta A_{\mathrm{PERT1}}$ (single topology) are nearly the same; the small differences in nonbonded contributions are compensated by the presence of bonded terms in the single topology result. This is also in accord with the identification of these bond and bond angle components as pmf-type contributions. All of the bonded contributions in $\Delta\Delta A_{\mathrm{PERT1}}$ are small in this case. They could be larger in other systems, but the identification as pmf contributions would still hold.

Comparisons of the gas phase and the solvation calculations in dual topology (Table 5) and single topology (Tables 6a−c) serve to increase our understanding, though they have to be

made with care because the simulations involve different initial and final states (see sections 2b and e of the companion paper). In contrast to the $\Delta\Delta A_{solv}$ values in Table 7, the individual gas phase (ΔA_{gas}) and solution results (ΔA_{solu}) of the BLOCK and PERT1 calculations in Tables 5 and 6a show large differences. In the gas phase results, the larger bonded free energy components of the single topology simulations originate in the vibrational and Jacobian factor contributions (cf. Table 6a) plus the intramolecular potential of mean force type contributions. The former two are omitted in the dual topology calculations, while the latter one is projected on the nonbonded components. If the dual topology methodology in which bonded terms are scaled to a limiting value of the force constant (see section 2d of the companion paper) were used, these bonded contributions would be accounted for as well. Corresponding considerations apply to the results of the solvent calculations. The individual gas phase and solution results of the BLOCK and PERT2 calculations differ by large pmf and (unphysical) Jacobian factor contributions that are included in PERT2 but not included in BLOCK.

4. Discussion

A theoretical analysis, given in the companion paper,[10] and the model calculations presented here have been used to examine the role of bond stretching and bond angle energy terms in free energy simulations. In the companion paper[10] a number of results were obtained which we summarize briefly: (i) Practical approaches to compute free energy differences due to changes in bonded terms were compared. It was shown that in thermodynamic integration, changes in bond stretching and bond angle terms are straightforward to incorporate in the standard methodology. By contrast, there are practical difficulties in the exponential formula if flexible bond terms are used (cf. ref 6), although a method for overcoming this limitation has been proposed.[12] Overall, thermodynamic integration appears to be more flexible and easier to use in this regard. Three methodologies for calculating the bonded terms in thermodynamic integration were considered. They are single topology, dual topology with an ideal gas molecule end state and dual topology with an ideal gas atom end state. (ii) Any method that attempts to create or remove a bond(ed) term in a straightforward manner, i.e. by changing its force constant from or to a zero value, is likely to lead to erratic results since the theoretical expressions diverge at the limits; see also Sun et al.[9] for a related discussion that does not provide a full solution to the problem. The complication that arises can be avoided in two ways. One uses ideal gas molecules as end states and the other uses a physically meaningful cutoff. In single topology simulations ideal gas molecule end states seem to have been used in all practical applications, although reasons for their use were never given explicitly. In a number of dual topology calculations, which used ideal gas atom end states, convergence problems have arisen.[3−5] The theoretical analysis demonstrated that the consistent use of an ideal gas molecule end state is possible in dual topology and eliminates the convergence problem. In addition, an approach that avoids the singularity caused by a bond with a zero force constant by a cutoff was introduced (section 2d of the compansion paper[10]). (iii) The free energy contributions resulting from changes in bonded energy terms were shown to consist of vibrational, Jacobian factor, and pmf-type contributions, as well as coupling among these terms. It was demonstrated that these contributions, although physically meaningful, appear as different free energy components depending on the simulation methodology used; i.e., they appear as bonded free energy contributions in single topology and as nonbonded or a mixture of bonded and nonbonded contributions in the two dual topology approaches. (iv) The three methodologies were shown to lead to identical double free energy differences defined by a thermodynamic cycle (e.g., a free energy difference of solvation), as they must. However, quite different single free energy differences can be obtained as different end states are involved. Thus, it is necessary to use a consistent methodology in the two parts of any thermodynamic cycle and to be careful in the description of the physical interpretation of the results.

The model calculations presented in this paper support and amplify the conclusions of the companion paper.[10] The thermodynamic integration approach was used throughout because of complications intrinsic to the exponential formula.[6,10,12] The solvation free energies of three systems were examined, using both single and dual topology approaches. They are (i) two one-dimensional harmonic oscillators interacting with a third particle that represents the solvent (an ideal gas atom as well as molecule end state was used in the dual topology calculations), (ii) the aqueous solvation of two diatomic molecules, and (iii) the aqueous solvation of ethane and methanol. In all cases identical (double) free energy differences of solvation were obtained within statistical error bars (see, in particular, Tables 3, 4, and 7 and section 3), confirming the theoretical considerations. The explicit verification of this agreement is important as it allows one to choose the methodology that is best suited for a given problem. In contrast to the overall results (the free energy differences of solvation, which is independent of the method used), differences were found, as expected,[10] in the results for the respective single free energy calculations and free energy components (see Tables 3−6 and section 3). A critical comparison of the free energy components obtained with different simulation methodologies that follow different simulation paths, led to a clear understanding of the physical origin of the various contributions to the overall results, as summarized below.

The simulations provided representative examples of each of the physical contributions described in the companion paper,[10] i.e., vibrational, Jacobian factor, pmf-type, and coupling contributions. The role of vibrational contributions due to a change in the force constant is clearly visible in the one-dimensional model system (section 3a). In the gas phase, it is the only contribution to the thermodynamic cycle depicted in Figure 1, and the same result (ΔA_3 in Table 3a) is obtained with single topology and dual topology using an ideal gas atom end state (d.t.2). It is not obtained in the dual topology simulations using an ideal gas molecule end state (d.t.1 in Table 3a) as this approach omits vibrational and Jacobian factor contributions. In the corresponding solvent calculation (ΔA_4 in Table 3a) the free energy difference consists of vibrational and pmf-type contributions. Here the use of dual topology using an ideal gas atom end state made it possible to separate the two terms into bond and van der Waals contributions, respectively (d.t.2 entries in Table 3a), whereas in single topology both appear as part of the bond free energy component. For the ethane to methanol calculation, it could be shown that the vibrational contributions due to the changes in bond stretching force constants are responsible for the difference of 0.4 kcal/mol between the single free energy differences (gas phase and solution) obtained with the PERT2 (flexible bond stretching terms) and the PERT2C (constrained bond lengths) protocol (Tables 6b and 6c).

The role of Jacobian factors was analyzed in detail in ref 18. The present results provide additional examples. In the computations of the free energy difference of solvation of two diatomic solutes (section 3b), the difference of 0.8 kcal/mol

between the solution results using single and dual topology (ΔA_{solu} in Table 4) is caused by the automatic inclusion of the Jacobian factor in the single topology methodology used in CHARMM[14] and its omission in the dual topology methodology using an ideal gas molecule reference state. (In addition, in this specific calculation there cannot be a Jacobian factor contribution as the solute was held fixed in the dual topology simulations.) The second example is the difference of 3.8 kcal/mol between the PERT1 and PERT2(C) gas phase (ΔA_{gas}) and solution free energy differences (ΔA_{solu}) for the ethane to methanol system (Tables 6a−c). This large value was explained by computing analytically the Jacobian factor contribution due the change of bond lengths to the dummy atoms (see section 3c).

The meaning of pmf-type contributions in single topology is evident from a comparison of single and dual topology free energy components in all three model systems. Particularly striking are the results for the diatomic molecules in water (section 3b, Table 4). After subtracting the Jacobian factor contribution of −0.8 kcal/mol (cf. previous paragraph), the free energy difference of solvation consists of a pmf-type bond contribution of −5.9 kcal/mol with the single topology methodology. In the dual topology simulations, the identical result is obtained, but it consists solely of van der Waals and electrostatic contributions. This makes clear that pmf-type contributions describe the change in nonbonded interactions as a result of a change in molecular geometry.

The presence and magnitude of coupling between the three contributions just discussed is especially interesting as it is directly related to the importance of *self-term contributions*, to which we return shortly. In the calculations on the one-dimensional model system summarized in Tables 3b and c, clear instances of coupling could be discerned. It becomes important when the force constant or equilibrium bond length of a weak bond is changed in the presence of strong nonbonded interactions. On the other hand, when typical values of the force constant were chosen (e.g., in the ethane to methanol system), no coupling was observed.

Bond and bond angle terms are part of the intramolecular term of the hybrid potential energy function (ΔU_{intra} defined in the Introduction of ref 10) and are commonly associated with *self-term* contributions to the free energy difference;[5] i.e., the free energy components that result from ΔU_{intra} make up the self-term. Contradictory results are described in the literature concerning the importance of self-terms. Prevost et al.[5] and Pearlman and Kollman[6] found sizable contributions from self-terms or intraperturbed group contributions. In contrast, refs 7−9 report little or no contributions. Attempts by Nilsson and co-workers[3,4] to resolve this issue failed because of convergence problems in some of their computations. To be clear about what is involved, it is essential to understand the differences between single and dual topology methods, as well as the origin of the problems encountered during breaking or forming of a bond term. Both aspects were analyzed theoretically in the companion paper[10] and illustrated by the model simulations. In the following paragraphs, we consider the results obtained in refs 3−9 and show how they can be understood by taking account of the simulation methodology (single or dual topology) used in the calculation. The classification of a contribution as a *self-term* is based, as it should be, on the *physical* significance of the bonded free energy components, which in turn depend on the simulation path.[35,36,39−44]

In single topology calculations formal bond (angle) components are obtained, so the question regarding the importance of self-terms can be addressed directly. Clearly, vibrational and Jacobian factor contributions obtained in such calculations are self-terms. On the other hand, a pmf-type contribution reflects the change in nonbonded interactions due to a change in the equilibrium geometry of (a part of) the system. Thus, the so-called "overlooked bond-stretching contribution" of Pearlman and Kollman,[6] which is a pmf-type contribution, should *not* be interpreted as a self-term (or intra-group perturbed contribution). This is supported by a comparison of the components of the free energy difference of solvation between ethane and methanol listed in Table 7. The bond component between the PERT1 and the PERT2(C) results differ by more than 1 kcal/mol. However, it was demonstrated in Section 3c that this bond contribution is nonbonded in origin, i.e., a pmf-type contribution. This can be deduced directly from Table 7 by observing that any change in bonded free energy components is compensated by a reciprocal change in the nonbonded components. Furthermore, in the BLOCK, i.e., dual topology, results in Table 7, the bond and bond angle free energy components obtained in a single topology calculation are nonbonded components. Thus, while it is imperative to include the pmf-type contribution to obtain the correct free energy difference, it is part of the nonbonded free energy contributions and, therefore, is not a self-term. The only pmf-type contributions that do involve self-terms arise from coupling with vibrational and Jacobian factor contributions because they reflect changes in the intramolecular contributions (vibrational and Jacobian factor) due to the interaction with different environments (e.g., gas phase and solution). However, in the model calculations, coupling was found to be important only for very weak bonded terms (low force constants). The influence of solvent on the equilibrium geometry of the solutes appears to be too small in most cases to lead to significant differences in the Jacobian factor contributions. Correspondingly, the force constants typically used in molecular mechanics force fields are too strong to be affected to a significant degree by nonbonded interactions. Somewhat larger effects might be found if anharmonicity (e.g., via a Morse potential) were included in bond stretching. Within the precision of the calculations, vibrational and Jacobian factor contributions canceled from the thermodynamic cycles of the two realistic systems studied here (diatomic molecule in water, ethane to methanol, section 3b and c). This is in line with the work of Harris and Loew[7] and Rao et al.[8] who have suggested that self-terms are not important. In fact, the present results indicate that selected terms of the energy function might be omitted from the free energy formalism *unless* single free energy differences are required. However, we feel that it is best not to omit any such terms because including them does not introduce any difficulties in the calculations, does not lead to detrimental effects on the precision of the calculations (in contrast to ref 6) and avoids problems that could arise from their neglect.

In the studies by Prevost et al.[5] and Nilsson and co-workers[3,4] a dual topology approach was used. It seems likely from the present analysis that the significant bond and bond angle components obtained in these studies are due to inaccuracies arising from the making and breaking of bonds. In a correct implementation of the dual topology method, bond and bond angle terms must be preserved or limiting values of the coupling parameters have to be used (section 2d of the companion paper[10]). The former excludes vibrational and Jacobian factor contributions from the total free energy difference. As mentioned earlier, different single free energy differences result, but this is irrelevant for thermodynamic cycles. This is reflected in the d.t.1 results for the one-dimensional systems (Tables 3a−c) and

the dual topology results for the diatomic molecule (Table 4) and the ethane to methanol system (Tables 5 and 7). Contributions to self-terms that do not cancel from a double free energy difference arise only if coupling between these terms and nonbonded interactions is important; these enter the dual topology formalism indirectly through the Boltzmann factor.[10] Within a dual topology framework that uses an ideal gas molecule reference state, i.e., bonded terms are not scaled, self-term contributions to double free energy differences due to bond (angle) can only be discerned by a comparison of results from separate calculations using flexible and constrained bond (angle) terms. This was not done here since equivalent information is available from a comparison of single and dual topology results.

The present analysis also makes clear that it is essential to avoid a simplistic comparison of results and conclusions obtained with single and dual topology methods. A pmf-style "bond" contribution cannot be obtained in this form in a dual topology simulation as attempted in refs 3 and 4 in their "FULL" calculations. We stress that the results of the present study show that the "DONT" results of refs 3 and 4, in which bond and bond angle terms were excluded from the free energy formalism, are, in fact, correct and do *not* omit any contribution to the double free energies of interest.

Dihedral angles and intramolecular nonbonded terms can also make contributions to self-terms. It has been suggested that dihedral angles might be strongly coupled with nonbonded terms since their force constants are weak compared to bond stretching and bond angle terms.[45] There are no problems, in principle, to including dihedral angles explicitly in both dual and single topology simulations (cf. the ethane/methanol calculations, sections 2c and 3c) and thus to ensure that no contribution is omitted from the free energy difference of interest. While the dihedral angle contribution to the free energy difference of solvation cancels for the ethane to methanol system (Table 7), this need not be true in general. However, it should be noted that the only real complication expected from dihedral angles is the presence of multiple conformational substates that will not be sampled adequately in most straightforward computer simulations. Special techniques have been developed to deal with this type of problem (see refs 46 and 47, Kuczera and Karplus, unpublished results). Nonbonded energy terms can also contribute to self-terms. For the alchemical mutation of neopentane into methane, Pearlman and Kollman reported a contribution of 0.6−0.9 kcal/mol,[6] although in later work on solvation free energy differences of fluorocarbons and alkanes from the same group[48,49] these terms were again omitted. For the ethane/methanol model systems studied here, we did not find significant contributions from intramolecular nonbonded terms.

In the two papers presented here, the role of bond (angle) terms in free energy difference simulations was analyzed using the example of solvation processes. Since there are no fundamental differences between solvation calculations and other applications of free energy simulations, such as determinations of protein stability[5,50,51] or ligand binding,[52,53] the results of this study remain valid. For example, in the calculation of the binding free energy difference between two ligands L1 and L2, the two alchemical mutations involve transformation of ligand L1 into L2 in solution (this step is identical to the solvation calculation considered here) and bound to the receptor. In the latter case, the receptor "replaces" the solvent. The choice of simulation methodology (single or dual topology) determines the correct treatment of changes in bond and bond angle parameters, as well as the interpretation of results. A recent study that favored single topology because of faster convergence also found that for certain transformations a dual topology method is required.[11] Equivalent results can be obtained with the two types of methods provided that the respective "peculiarities" that have been described here are taken into account. For the simple solutes used here, the definition of the self-term was straightforward and corresponded to the intrasolute free energy contributions. For bigger systems the definition (aside from the differences arising from the choice of a single vs a dual topology method) may be less clear (e.g., if in a protein one side chain is altered alchemically, one has to decide how to consider the contributions from energy terms involving the backbone atoms of the mutated residue). Their parameters remain unchanged, yet they are part of the mutated residue. A good example with explicit definitions can be found in ref 5. In addition to solvent pmf-type contributions, which we regard as nonbonded in origin, self-term contributions were found to be negligible for the model systems studied. A protein may provide a less isotropic environment than aqueous solution; it could induce distortions from the equilibrium geometry (strain) or conformational changes, which in turn would lead to nonvanishing self-term contributions. This possibility should be kept in mind and a generalization of the quantitative results reported here should be made with care, although the principles are general.

The current study provides a framework for the calculation and interpretation of bonded terms by single and/or dual topology free energy simulations on larger systems. In combination with other recent improvements in methodology, particularly those concerned with the van der Waals endpoint problem,[54−56] the correct treatment of long-range electrostatic interactions,[57,58] and the understanding of the role of Jacobian factors,[18] as well as the availability of more computer time, this makes possible the performance of converged free energy simulations for realistic systems of chemical and biological interest. The results of such simulations and their analysis will, we believe, find a useful place, as a complement to experiment, in the elucidation of the properties of complex systems.

Acknowledgment. We thank Tom Simonson, Georgios Archontis, and Arnaud Blondel for helpful discussions. This work was supported by a grant from the National Institute of Health.

Supporting Information Available: Tables of the parameters and charges of the CHARMM22 all-atom parameter set used in the ethane/methanol free energy calculations (3 pages). Ordering information is given on any current masthead page.

References and Notes

(1) Bash, P. A.; Singh, U. C.; Langridge, R.; Kollman, P. A. *Science* **1987**, *236*, 564−568.
(2) Rao, S. N.; Singh, U. C.; Bash, P. A.; Kollman, P. *Nature* **1987**, *328*, 551.
(3) Elofsson, A.; Nilsson, L. *Mol. Simul.* **1993**, *10*, 255−276.
(4) Eriksson, M. A. L.; Nilsson, L. *J. Mol. Biol.* **1995**, *253*, 453−472.
(5) Prevost, M.; Wodak, S. J.; Tidor, B.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 10880−10884.
(6) Pearlman, D. A.; Kollman, P. A. *J. Chem. Phys.* **1991**, *94*, 4532−4546.
(7) Harris, D.; Loew, G. *J. Comput. Chem.* **1996**, *17*, 273−288.
(8) Rao, B. G.; Kim, E. E.; Murcko, M. A. *J. Comput. Aid. Mol. Des.* **1996**, *10*, 23−30.
(9) Sun, Y.-C.; Veenstra, D. L.; Kollman, P. A. *Protein Eng.* **1996**, *9*, 273−281.
(10) Boresch, S.; Karplus, M. *J. Phys. Chem.* **1998**, *103*, 103.
(11) Pearlman, D. A. *J. Phys. Chem.* **1994**, *98*, 1487−1493.
(12) Severance. D. L.; Essex, J. W.; Jorgensen, W. L. *J. Comput. Chem.* **1995**, *16*, 311−327.
(13) Wolfram, S. *Mathematica: A System for Doing Mathematics by Computer*, version 2.2.4; Wolfram Research, Inc: Champaign, IL, 1995.

(14) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(15) Nosé, S. *Mol. Phys.* **1984**, *52*, 255−268.
(16) Ryckaert, J. P.; Cicotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(17) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935. Reiher, III Ph.D. Thesis, Harvard University, 1985. Neria, E.; Fischer, S.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 1902−1921. See also the tables in Supporting Information.
(18) Boresch, S.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 5145−5154.
(19) Beveridge, D. L.; DiCapua, F. M. *Annu. Rev. Biophys. Chem.* **1989**, *18*, 431−492.
(20) Jorgensen, W. L.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050−3054.
(21) Fleischman, S. H.; Brooks, C. L. III *J. Chem. Phys.* **1987**, *87*, 3029−3037.
(22) Singh, U. C.; Brown, F. K.; Bash, P. A.: Kollman, P. A. *J. Am. Chem. Soc.* **1987**, *109*, 1607−1614.
(23) Ben-Naim, A.; Marcus, Y. *J. Chem. Phys.* **1984**, *81*, 2016−2027.
(24) Carlson, H. A.; Nguyen, T. B.; Orozco, M.; Jorgensen, W. L. *J. Comput. Chem.* **1993**, *14*, 1240−1249.
(25) MacKerell, A. D. Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L. Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Workiewicz-Kuczera, J.; Yinb, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(26) Tembe, B. L.; McCammon, J. A. *Comput. Chem.* **1984**, *8*, 281−283.
(27) Hermans, J.; Yun, R. H.; Anderson, A. G. *J. Comput. Chem.* **1992**, *13*, 429−442.
(28) Wang, L.; Hermans, J. *J. Chem. Phys.* **1994**, *100*, 9129−9139.
(29) Herschbach, D. R.; Johnston, H. S.; Rapp, D. *J. Chem. Phys.* **1959**, *31*, 1652.
(30) Wood, R. H. *J. Phys. Chem.* **1991**, *95*, 4838−4842.
(31) Hunter, J. E., III; Reinhardt, W. P.; Davis, T. F. *J. Chem. Phys.* **1993**, *99*, 6856−6864.
(32) de Souza, L. E. S.; Guerin, C. B. E.; Ben-Amotz, D.; Szleifer, I. *J. Chem. Phys.* **1993**, *99*, 9954−9961.
(33) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(34) Jorgensen, W. L.; Swenson, C. *J. Am. Chem. Soc.* **1985**, *106*, 765−784.
(35) Boresch, S.; Archontis, G.; Karplus, M. *Proteins* **1994**, *20*, 25−33.
(36) Boresch, S.; Karplus, M. *J. Mol. Biol.* **1995**, *254*, 801−807.
(37) Straatsma, T. P.; Zacharias, M.; McCammon, J. A. In *Computer Simulations in Biological Systems*; van Gunsteren, W. F. Weiner, P. K., Wilkinson, A. J., Eds.; Escom Science: Netherlands, 1994.
(38) Straatsma, T. P.; Zacharias, M.; McCammon, J. A. *Chem. Phys. Lett.* **1992**, *196*, 297−302.
(39) Mark, A. E.; van Gunsteren, W. F. *J. Mol. Biol.* **1994**, *240*, 167−176.
(40) Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1994**, *100*, 577−585.
(41) Van Gunsteren, W. F.; Beutler, T. C.; Fraternali, F.; King, P. M.; Mark, A. E.; Smith, P. E. In *Computer Simulations in Biological Systems*; van Gunsteren, W. F. Weiner, P. K., Wilkinson, A. J., Eds.; Escom Science: Netherlands, 1994.
(42) Simonson, T.; Brünger, A. T. *Biochemistry* **1992**, *31*, 8661−8674.
(43) Brady, G. P.; Sharp, K. A. *J. Mol. Biol.* **1995**, *254*, 77−85.
(44) Archontis, G.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 11246−11260.
(45) Van Gunsteren, W. F. In *Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications*; Van Gunsteren, W. F., Weiner, P. K., Eds.; Escom Science: Netherlands, 1989; pp 27−59.
(46) Straatsma, T.; McCammon, J. A. *J. Chem. Phys.* **1989**, *90*, 3300−3304.
(47) Hodel, A.; Rice, L. M.; Simonson, T.; Fox, R. O.; Brünger, A. T. *Protein Sci.* **1995**, *4*, 636−654.
(48) Sun, Y.; Spellmeyer, D.; Pearlman, D. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1992**, *114*, 6798−6801.
(49) Gough, C. A.; Pearlman, D. A.; Kollman, P. A. *J. Chem. Phys.* **1993**, *99*, 9103−9110.
(50) Dang, L. X.; Merz, K. M.; Kollman, P. A. *J. Am. Chem. Soc.* **1989**, *111*, 8505−8508.
(51) Tidor, B.; Karplus, M. *Biochemistry* **1991**, *30*, 3217−3228.
(52) Lau, F. T. K.; Karplus, M. *J. Mol. Biol.* **1994**, 1049−1066.
(53) Miyamoto, S.; Kollman, P. A. *Proteins* **1993**, *16*, 226−245.
(54) Simonson, T. *Mol. Phys.* **1993**, *80*, 441−447.
(55) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1994**, *222*, 529−539.
(56) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1994**, *100*, 9025.
(57) Simonson, T.; Archontis, G.; Karplus, M. *J. Phys. Chem.* **1997**, *41*, 8347−8360.
(58) Hummer, G.; Pratt, L. R.; Garcia, A. E. *J. Phys. Chem.* **1996**, *100*, 1206−1215.