

Cumulant expansion of the free energy: Application to free energy derivatives and component analysis

G. Archontis and M. Karplus

Citation: *The Journal of Chemical Physics* **105**, 11246 (1996); doi: 10.1063/1.472921

View online: <http://dx.doi.org/10.1063/1.472921>

View Table of Contents: <http://scitation.aip.org/content/aip/journal/jcp/105/24?ver=pdfcov>

Published by the [AIP Publishing](#)

Articles you may be interested in

Hydration of the tetramethylammonium ion: From water structure to the free energy of hydration

AIP Conf. Proc. **492**, 202 (1999); 10.1063/1.1301529

Treatment of electrostatic interactions in computer simulations and calculation of thermodynamic properties such as free energies and pressures

AIP Conf. Proc. **492**, 84 (1999); 10.1063/1.1301522

A thermodynamic distance criterion of optimality for the calculation of free energy changes from computer simulations

J. Chem. Phys. **105**, 10072 (1996); 10.1063/1.472836

Curvature elasticity and thermodynamic stability of electrically charged membranes

J. Chem. Phys. **105**, 8314 (1996); 10.1063/1.472686

Free energy simulations: Correcting for electrostatic cutoffs by use of the Poisson equation

J. Chem. Phys. **104**, 7645 (1996); 10.1063/1.471472



Cumulant expansion of the free energy: Application to free energy derivatives and component analysis

G. Archontis and M. Karplus^{a)}

Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138

(Received 22 April 1996; accepted 26 August 1996)

The free-energy change in a system represented by an empirical energy function, when calculated by thermodynamic integration from the initial to the final state, can be expressed as a sum of “free-energy components,” that correspond to contributions from different energy terms or different parts of the system. Although the overall free energy is path independent, the values of the components depend on the integration path, which determines the physical processes involved in the change. A cumulant expansion is used to express the total free-energy change as a sum of terms that contain the cumulants of a single energy term plus a coupling term involving cumulants of more than one energy term. The dependence of the free-energy components on the thermodynamic integration path is shown to arise from the different partitionings of the latter term. A concerted linear integration path, along which the total energy function changes from the initial to the final state as a linear function of an external parameter λ , as well as piecewise-linear paths, along which individual energy terms are changed linearly, are examined. Along the concerted linear path the partitioning of the coupling term preserves permutation symmetry with respect to energy components to all orders. This contrasts to piecewise-linear paths, where the total coupling is projected onto one component. To illustrate the analysis, calculations are made for the transformation $\text{Cl}^{(0)} \rightarrow \text{Br}^{(-)}$ (a neutral particle with the van der Waals parameters of $\text{Cl}^{(-)}$ to a charged particle with the van der Waals parameters of $\text{Br}^{(-)}$) in aqueous solution along various paths. Terms through third order in the cumulant expansion for the change in the solvation free energy are evaluated by use of the reference interaction site model with the hypernetted chain closure, and the values obtained along the linear and piecewise-linear paths are compared and analyzed. The utility of cumulants for the prediction of the effect of specific changes in the energy terms, and for interpreting free-energy components is discussed. In most cases a first-order approximation is not accurate enough to be useful, but results through third order are accurate to 5% or better. In the Appendix a general expression for the free-energy derivatives of a Hamiltonian with a general functional dependence on a parameter λ is derived. © 1996 American Institute of Physics. [S0021-9606(96)50145-5]

I. INTRODUCTION

Free-energy changes in processes of chemical and biological interest, such as the solvation of atomic and molecular solutes in water or organic solvents, protein-ligand associations, or the introduction of single or multiple mutations in proteins, are the focus of much experimental and theoretical research. In many calculations of free-energy differences a given chemical entity (e.g., methanol) is transformed into another (e.g., ethane); we refer to such a transformation as “alchemical”^{1,2} because it is the computational analog of the transmutation of the elements, such as changing lead into gold. Since an alchemical transformation can be done reversibly on a computer, it is possible to calculate the free energy change involved.

Free-energy differences in solutions and in macromolecules can be calculated by several methods, including integral equation theories,^{3–5} Poisson–Boltzmann calculations,^{6,7} Monte Carlo,^{8,9} and molecular dynamics simulations.^{10,11} One approach uses thermodynamic integration¹² in which the

change in the Hamiltonian of the system is expressed as a function of a coupling parameter that connects the initial and the final state. If the change in potential energy in the Hamiltonian consists of a sum of terms, as it does in many empirical energy functions for complex systems, the total free-energy change can be decomposed into a sum of terms, called “free-energy components,” each of which contains explicitly a single interaction energy term.¹⁰ These free-energy components are often associated with the contribution of a specific interaction (e.g., van der Waals vs electrostatic), or a specific part of the system (e.g., mutated protein vs solvent), and so provide useful insights into the physical origin of the free-energy change. This is important because only the overall free-energy change can be measured experimentally. The component analysis has been applied to the effect of mutations on protein stability^{13,14} and ligand binding.^{14,15} For example, the validity of the hydrogen-bond inventory model has been examined with a free-energy simulation of binding of the tyrosine substrate to the Tyrosyl-tRNA synthetase mutant Tyr96→Phe96.¹⁴ More recently, free-energy derivatives, which correspond to low order approximations to the components, have been used to estimate the effect on

^{a)}Also at: Laboratoire de Chimie Biophysique, Institut Le Bel, Université Louis Pasteur, 1 rue Blaise Pascal, 67000 Strasbourg, France.

the binding affinity of localized changes in ligand molecules.^{16,17}

The free energy components obtained from thermodynamic integration are path dependent, although the total free-energy change is not. It has been shown that this path dependence has a local and a global aspect.^{18,19} The global aspect is associated with changes in the components when physically different paths are examined; e.g., the effect of a mutation in the folded and denatured protein are obtained along one path and the difference in the potential of mean force for the folding reaction of the mutant and wild-type protein along another.¹⁸ The local aspect of the path dependence is concerned with the nature of the coupling terms that arise in the free-energy components due to the Boltzmann factor involved in their evaluation. In Ref. 18 a detailed examination was made of the concerted linear path, along which the total energy change from the initial to the final system is switched on in the Hamiltonian as a linear function of an external parameter λ . The expansion of the Boltzmann factor was used to show that this concerted linear path (also referred to as the alchemical path^{1,18,10}), leads to a symmetric partitioning of the coupling terms through second order.

In this paper we investigate the origin of dependence of the free-energy components on the thermodynamic integration path. The simple case of two energy terms that change between the initial and final state is described; extension to any number of terms is straightforward. We utilize a cumulant expansion of the total free-energy change and extend to infinite order the results through second order given previously.¹⁸ By means of this cumulant expansion the total free-energy change is expressed as a sum of terms that contain the cumulants of each energy term, plus a coupling term that includes cumulants of both energy terms. The dependence of the free-energy components on the thermodynamic integration path arises from the different partitioning of the coupling term into components along different paths. Along the concerted linear path the partitioning of the coupling term preserves the permutation symmetry with respect to energy components to all orders. This contrasts to piecewise-linear paths, where only one component changes over each interval and the coupling term is projected onto that component in the interval. It is shown also that along the concerted linear path the free-energy components only change sign when the initial and final state are inverted. Other paths do not have these desirable properties.

To illustrate the conclusions from the cumulant expansion we consider the transformation $\text{Cl}^{(0)} \rightarrow \text{Br}^{(-)}$ in aqueous solution. This transformation is introduced because it involves altering two terms in the Hamiltonian; that is, the van der Waals interaction and the Coulombic interaction between the solute particle and the solvent. Along the concerted linear path the functional form for the total solute-solvent interaction energy is varied from the neutral $\text{Cl}^{(-)}$ -like Lennard-Jones particle to the $\text{Br}^{(-)}$ ion by means of linear coupling of both energy terms to an external parameter. The results are compared with piecewise-linear paths, along which the van der Waals and electrostatic interactions are altered one at a time. The thermodynamic cycle associated with these trans-

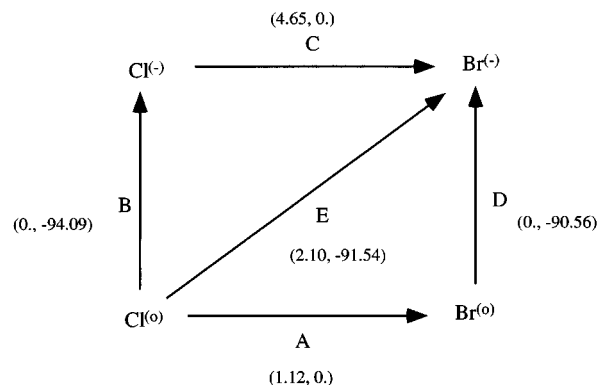


FIG. 1. Thermodynamic cycle for the transformation $\text{Cl}^{(0)} \rightarrow \text{Br}^{(-)}$ in water. The free energy differences between the various species were calculated with the xRISM-HNC methodology (Refs. 20 and 21). Along paths A through D the total free energy change is projected on one free energy component. Path E corresponds to a simultaneous linear switching of both van der Waals and electrostatic interactions [the alchemical path (Refs. 1, 2 and 10)]. The free energy components along this path are calculated in Eq. (5). In the pair of numbers in parentheses for each path, the first corresponds to the van der Waals and the second to the electrostatic component.

formations is shown in Fig. 1. We evaluate the terms through third order in the cumulant expansion for the change in solvation free energy and compare the values obtained along the linear and piecewise-linear paths, by means of the extended reference interaction site model (xRISM) theory.^{4,20,21} We also discuss the utility of cumulants for the prediction of the effect of specific changes in the energy terms, and for interpreting free-energy components. In the Appendix we derive a general expression for the free-energy derivatives of a Hamiltonian with a general functional dependence on a parameter λ .

In Sec. II A we outline the thermodynamic integration and exponential formalism for calculating the change in the free energy, and introduce the free-energy components that contribute to the total free-energy change. In Sec. II B we derive a cumulant-expansion for the total free-energy change and for the free-energy components. We compare the concerted linear (alchemical) path with piecewise-linear paths and use this comparison to analyze the dependence of the free-energy decomposition on the path. We also consider the case of a path with a general functional dependence of the Hamiltonian on an external parameter λ . The cumulant expansion is based on the selection of a given state as the reference state and expresses the perturbation Hamiltonian ΔH relative to that state. In Sec. III we demonstrate the symmetry of the free-energy components along the concerted linear path for an interchange of the initial and final states as the reference. In Sec. IV we illustrate the theoretical conclusions with a numerical example that makes use of the alchemical transformation of a $\text{Cl}^{(0)}$ neutral particle with van der Waals parameters of $\text{Cl}^{(-)}$, to a charged particle with van der Waals parameters of $\text{Br}^{(-)}$, in aqueous solution. We utilize the extended reference interaction site model (xRISM) theory,^{4,20,21} coupled with the hypernetted-chain (HNC) approximation,²² to evaluate the terms in the cumulant expansion of the free-energy change along the various portions

of the corresponding thermodynamic cycle (Fig. 1). The first-order terms are calculated analytically, and the rest by numerical evaluation of the second- and third-order free-energy derivatives. The n th order free-energy derivative coincides with the n th order term in the cumulant expansion of the free-energy change along a linear path, as shown in the Appendix. The partitioning of the coupling term between the van der Waals and electrostatic components in the transformation is presented. An analysis is made of the use of low order derivatives to estimate the effect of a change in the properties of the system. The conclusions are summarized in Sec. V. A cumulant expansion for evaluating free-energy derivatives to infinite order is derived in the Appendix for a general dependence of the Hamiltonian on λ .

II. THEORY

In this section we analyze the origin of the dependence on the coupling parameters in the Hamiltonian of the decomposition of the free energy into contributions from different free-energy components. We first summarize the thermodynamic integration and exponential formalisms for calculating the free-energy change in Sec. II A, and review the meaning of free-energy components obtained in the thermodynamic integration formalism. In Sec. II B we utilize the cumulant-expansion theorem,²³ to show that the total free-energy change can be expressed as a sum of cumulants containing a single energy term, and cumulants involving coupling between different energy terms. We establish the connection between the various terms in the cumulant expansion and the free-energy components of the thermodynamic integration formalism along piecewise-linear paths, as well as the concerted linear path. By so doing, we extend to infinite order the results obtained in Ref. 18 through second order. We also treat the case in which the Hamiltonian has a general functional dependence on an external parameter λ .

A. Overview of the free-energy formalism

In the canonical ensemble the partition function $Z(N, V, T)$ for a classical system of N indistinguishable atoms, that occupy volume V at temperature T , is given by

$$Z(N, V, T) = \frac{1}{h^{3N} N!} \int d\mathbf{r} d\mathbf{p} \exp[-\beta H(\mathbf{r}, \mathbf{p})], \quad (1)$$

where \mathbf{r} are the particle coordinates and \mathbf{p} the conjugate momenta, H is the Hamiltonian of the system, h is Planck's constant, $\beta = 1/(k_B T)$ with k_B Boltzmann's constant, and the factor $N!$ takes into account the indistinguishability of the particles. The Helmholtz free energy of the system is given in terms of the partition function by the formula

$$A(N, V, T) = -\frac{1}{\beta} \ln Z(N, V, T). \quad (2)$$

A transformation of a system from an initial to a final state is expressed mathematically by a change in the Hamiltonian; e.g., the initial state is described by the Hamiltonian H_0 and the final state by the Hamiltonian $H_0 + \Delta H$. Most generally, the change in the Hamiltonian, ΔH , can corre-

spond to a sum over several intramolecular and/or intermolecular potential energy terms, as well as the kinetic energy. The transformation is accompanied by a change ΔA in the free energy A of the system. When a Cartesian system of coordinates is used and the potential energy does not depend on the particle momenta, the partition function factorizes into a kinetic and a configurational part. The kinetic contribution to the free-energy change can be calculated analytically by performing the integration over the momentum variables in Eq. (2) for the initial and final state. For this reason we focus below on the contribution to the free-energy change due to the configurational part of the partition function.

An expression for the configurational contribution to the free energy change ΔA can be obtained by the method of thermodynamic integration.¹² In this formulation the transformation from the initial to the final state is accomplished gradually and reversibly, as in an experimental measurement of the free-energy change. The changes in the energy terms ΔH_κ are expressed parametrically in terms of a set of variables $\{\lambda_1, \lambda_2, \dots\}$. A typical Hamiltonian has the form

$$H(\lambda_1, \lambda_2, \dots) = H_0 + \sum_\kappa \Delta H_\kappa(\lambda_\kappa). \quad (3)$$

In Eq. (3) it is assumed for simplicity that each component $\Delta H_\kappa(\lambda_\kappa)$ depends on a single parameter λ_κ . The variables $\{\lambda_\kappa\}$ change from a set of initial values $\{\lambda_\kappa^{\text{in}} = 0\}$ to a set of final values $\{\lambda_\kappa^{\text{f}} = 1\}$. The functional dependence $\Delta H_\kappa(\lambda_\kappa)$ is arbitrary, with the only constraint that it satisfy the boundary conditions $\Delta H_\kappa(\lambda_\kappa = 0) = 0$ and $\Delta H_\kappa(\lambda_\kappa = 1) = \Delta H_\kappa$. The set of values $\{\lambda_\kappa\}$ defines the integration path C , along which the exact free-energy change ΔA is calculated by the line integral

$$\begin{aligned} \Delta A &= -\frac{1}{\beta} \int_C d \ln Z(\lambda_1, \lambda_2, \dots) \\ &= \int_C d\lambda_1 \left\langle \frac{\partial \Delta H_1(\lambda_1)}{\partial \lambda_1} \right\rangle_{\lambda_1, \lambda_2, \dots} \\ &\quad + \int_C d\lambda_2 \left\langle \frac{\partial \Delta H_2(\lambda_2)}{\partial \lambda_2} \right\rangle_{\lambda_1, \lambda_2, \dots} + \dots \end{aligned} \quad (4)$$

Of particular interest is the concerted linear (alchemical^{1,18,10}) path, along which the Hamiltonian components vary as $\Delta H_\kappa(\lambda) = \lambda \Delta H_\kappa$. In this case the thermodynamic integration expression, Eq. (4), simplifies to

$$\Delta A = \int_0^1 d\lambda \langle \Delta H_1 \rangle_\lambda + \int_0^1 d\lambda \langle \Delta H_2 \rangle_\lambda + \dots \quad (5)$$

Equation (5) consists of a sum of terms, each containing the average of a single Hamiltonian component ΔH_κ . The appearance of the components ΔH_κ as independent terms in Eq. (5) has motivated the identification of each term of the above sum with the corresponding free-energy component.^{1,10} The thermodynamic integration formula reduces to an expression similar to Eq. (5) for the case of piecewise-linear paths, along which only some Hamiltonian components vary (linearly) and the rest remain constant. The

total free-energy change ΔA obtained along any of these piecewise-linear paths is path independent if the initial and final states are the same, because the free energy is a state function. However, the values obtained for the individual free-energy components are not necessarily the same along different paths, because the averages $\langle \Delta H_\kappa \rangle$ depend on the total Hamiltonian through its presence in the Boltzmann factor of the partition function, Eq. (1), and thus on the value of all the parameters $\{\lambda_\kappa\}$.^{11,15,19,24}

An alternative procedure for calculating free-energy changes is based on a formula for the free-energy difference between two systems with Hamiltonians H and $H + \Delta H$, first derived by Zwanzig;²⁵ it is

$$\Delta A = -\frac{1}{\beta} \ln \left\langle \exp \left[-\beta \sum_\kappa \Delta H_\kappa \right] \right\rangle_0, \quad (6)$$

where “0” denotes the initial state, as above. Although Eq. (6) is often referred to as the “thermodynamic perturbation” method, we prefer the term “exponential formula” because Eq. (6) is exact and does not correspond to a perturbation approximation. In the exponential formula the energy change is evaluated from an ensemble average over the initial (unperturbed) state. Unlike the thermodynamic integration formula Eq. (5), the exponential formula Eq. (6) cannot be decomposed into terms that each contain a single energy component ΔH_κ , except when these terms depend on different coordinates so that the integral for the partition function of the system factorizes²⁶ (e.g., the kinetic and potential energy components of the Hamiltonian). When the Hamiltonian terms depend on a common set of coordinates, a decomposition is possible only in the limit of a very small perturbation. The exponential formula for ΔA then reduces to a first-order perturbation expression²²

$$\Delta A \approx \sum_\kappa \langle \Delta H_\kappa \rangle_0. \quad (7)$$

Equation (7) is obtained as the first term in an expansion of ΔA in cumulants of the energy components ΔH_κ , as we show below. Outside the limit of a small perturbation, Eq. (7) is satisfied only if the contributions of the higher order cumulants vanish identically. For small ΔH_κ , the exponential formula approximation, Eq. (7), is equivalent to thermodynamic integration, Eq. (4). This can be verified by noting that the latter, in the limit of an infinitesimal perturbation from a reference state $\{\lambda_1, \lambda_2, \dots\}$, becomes

$$\Delta A \approx \sum_\kappa \left\langle d\lambda_\kappa \frac{\partial H(\lambda_\kappa)}{\partial \lambda_\kappa} \right\rangle_{\lambda_1, \lambda_2, \dots}, \quad (8)$$

which is the same as Eq. (7) but for the reference state $\{\lambda_1, \lambda_2, \dots\}$.

B. Dependence of the free-energy components on the simulation path

The exponential formula for ΔA [Eq. (6)] expresses the total free-energy change (a state function) in terms of an ensemble average over the initial state. Thus it is a path-

independent expression. In the following we show that this expression can be used to investigate the meaning of free-energy components along different linear paths. For simplicity we consider the case in which only two energy components change. The conclusions from this case can be generalized in a straightforward manner to systems with any number of energy terms that change in going from the initial to the final state.

An example that involves two energy components is the solvation of a charged monoatomic solute in a polar solvent. In this case, both the dispersion (van der Waals) and Coulombic interactions between the solute molecule and the solvent are changed from zero to their full value in solution. A related case, shown in Fig. 1 is the thermodynamic cycle for the transformation (mutation) of an uncharged chloridelike van der Waals sphere to the bromide ion, $(\text{Cl}^{(0)} \rightarrow \text{Br}^{(-)})$ in aqueous solution. This process involves a change in the van der Waals interaction and the switching on of the electrostatic interactions between the solute and solvent. It was considered in a previous analysis of the meaning of the decomposition of the total free-energy change into contributions from individual components.^{18,19} Reference 18 demonstrated that the decomposition in Eq. (5) leads to free-energy components, with the first-order coupling term between them partitioned in a symmetric manner. In Ref. 19 it was shown that different paths project the free-energy change onto different free-energy components, in a manner that is consistent and useful for interpreting the type of change introduced into the Hamiltonian. In what follows we use a cumulant expansion to evaluate the coupling terms to infinite order and show explicitly that the symmetry is preserved along the linear path. The formulation is an extension of the expansion used by Zwanzig in his original paper on the exponential formula.²⁵

For the case of two terms that change in the Hamiltonian, the total free-energy change is equal to

$$\Delta A = -\frac{1}{\beta} \ln \langle \exp[-\beta(\Delta H_1 + \Delta H_2)] \rangle_0, \quad (9)$$

by Eq. (6). In the transformation $\text{Cl}^{(0)} \rightarrow \text{Br}^{(-)}$ we denote with “0” the reference state $\text{Cl}^{(0)}$ and with “1” the reference state $\text{Br}^{(-)}$; ΔH_1 denotes the difference in the solute-solvent van der Waals energy term between the $\text{Cl}^{(0)}$ and $\text{Br}^{(-)}$ solutes, and ΔH_2 is the corresponding difference in the Coulomb term, which is the total bromide ion-water Coulombic interaction. The corresponding free energy components are denoted by ΔA_1 and ΔA_2 . The total free energy change along any step I of the thermodynamic cycle of Fig. 1 is denoted with ΔA_I .

To decompose the free energy ΔA in Eq. (9) into a sum of terms, we use the cumulant-expansion theorem²³

$$\ln \langle \exp[-\beta V] \rangle_0 = \langle \exp[-\beta V] - 1 \rangle_{0,c}, \quad (10)$$

where $\langle \dots \rangle_{0,c}$ is the cumulant evaluated at reference state “0”. The right-hand side of Eq. (10) contains cumulants of the form $\langle V^n \rangle_{0,c}$, which result from writing the exponential as a

power series. These cumulants can be expressed in terms of averages of powers of V ; the cumulants up to third order are given by²³

$$\langle V \rangle_{0,c} = \langle V \rangle_0,$$

$$\langle V^2 \rangle_{0,c} = \langle V^2 \rangle_0 - \langle V \rangle_0^2,$$

$$\langle V^3 \rangle_{0,c} = \langle V^3 \rangle_0 - 3\langle V^2 \rangle_0 \langle V \rangle_0 + 2\langle V \rangle_0^3.$$

Application of this theorem to Eq. (9) yields

$$\begin{aligned} \Delta A &= -\frac{1}{\beta} \langle \exp[-\beta(\Delta H_1 + \Delta H_2)] - 1 \rangle_{0,c} \\ &= \sum_{n=1}^{\infty} \frac{(-\beta)^{n-1}}{n!} \langle \Delta H_1^n \rangle_{0,c} + \sum_{n=1}^{\infty} \frac{(-\beta)^{n-1}}{n!} \\ &\quad \times \langle \Delta H_2^n \rangle_{0,c} + \sum_{n=2}^{\infty} (-\beta)^{n-1} \sum_{m=1}^{n-1} \frac{1}{m!(n-m)!} \\ &\quad \times \langle \Delta H_1^m \Delta H_2^{n-m} \rangle_{0,c} \\ &= \langle \Delta H_1 \rangle_0 - \frac{\beta}{2} [\langle \Delta H_1^2 \rangle_0 - \langle \Delta H_1 \rangle_0^2] + \frac{\beta^2}{3!} [\langle \Delta H_1^3 \rangle_0 \\ &\quad - 3\langle \Delta H_1^2 \rangle_0 \langle \Delta H_1 \rangle_0 + 2\langle \Delta H_1 \rangle_0^3] + \dots \\ &\quad - \frac{\beta}{2} [\langle \Delta H_1 \Delta H_2 \rangle_0 - \langle \Delta H_1 \rangle_0 \langle \Delta H_2 \rangle_0] \\ &\quad + \frac{\beta^2}{2!} \{ \langle (\Delta H_1^2 - \langle \Delta H_1^2 \rangle_0) \Delta H_2 \rangle_0 - 2\langle \Delta H_1 \rangle_0 \\ &\quad \times [\langle \Delta H_1 \Delta H_2 \rangle_0 - \langle \Delta H_1 \rangle_0 \langle \Delta H_2 \rangle_0] \} + \dots \\ &\quad + \{1 \leftrightarrow 2\}, \end{aligned} \quad (11)$$

where $\{1 \leftrightarrow 2\}$ represents terms obtained by permuting indices 1 and 2. Thus, the free-energy change ΔA can be decomposed into a sum of terms that involve cumulants of the separate energy components, ΔH_1 and ΔH_2 , plus a coupling term that contains cumulants of both components. The coupling term first appears in second order, in agreement with Eq. (7). Equation (11) corresponds to an expansion in β , or the inverse temperature; at high temperatures ($\beta \rightarrow 0$) only the first term, which does not have an explicit dependence on β , is significant. From the cumulant-expansion theorem, the first two sums on the r.h.s. of Eq. (11) satisfy the relations

$$\begin{aligned} &\sum_{n=1}^{\infty} \frac{(-\beta)^{n-1}}{n!} \langle \Delta H_1^n \rangle_{0,c} \\ &= -\frac{1}{\beta} \ln \langle \exp[-\beta \Delta H_1] \rangle_{0,c} \\ &= \int_0^1 d\lambda_1 \langle \Delta H_1 \rangle_{\lambda_1,0} = \Delta A_1(1,0) = \Delta A_A. \end{aligned} \quad (12)$$

$$\begin{aligned} &\sum_{n=1}^{\infty} \frac{(-\beta)^{n-1}}{n!} \langle \Delta H_2^n \rangle_{0,c} \\ &= -\frac{1}{\beta} \ln \langle \exp[-\beta \Delta H_2] \rangle_{0,c} \\ &= \int_0^1 d\lambda_2 \langle \Delta H_2 \rangle_{0,\lambda_2} = \Delta A_2(0,1) = \Delta A_B. \end{aligned} \quad (13)$$

These are the exact expressions for introducing ΔH_1 or ΔH_2 , starting with H_0 as the initial state. In the case $\text{Cl}^{(0)} \rightarrow \text{Br}^{(-)}$, Eq. (12) corresponds to the free energy change ΔA_A in the transformation $\text{Cl}^{(0)} \rightarrow \text{Br}^{(0)}$ (path A in Fig. 1), and Eq. (13) to the free energy change ΔA_B in the transformation $\text{Cl}^{(0)} \rightarrow \text{Cl}^{(-)}$ (path B). As indicated in Eqs. (12) and (13), the infinite sums of cumulants of ΔH_1 and ΔH_2 are equal to the free-energy components ΔA_1 and ΔA_2 obtained from the thermodynamic integration formula, evaluated along paths A and B, respectively.

To make clear the origin of the changes in the partitioning of the free energy along alternative paths, we write (see Fig. 1)

$$\begin{aligned} \Delta A &= \Delta A_B + \Delta A_C = (\Delta A_A + \Delta A_B) + (\Delta A_C - \Delta A_A) \\ &= \Delta A_A + \Delta A_D = (\Delta A_A + \Delta A_B) + (\Delta A_D - \Delta A_B). \end{aligned} \quad (14)$$

A comparison of Eqs. (11)–(14) shows that the terms included in the second parentheses of Eq. (14) are equal to the coupling term of Eq. (11); i.e.,

$$\begin{aligned} \Delta A_C - \Delta A_A &= \Delta A_D - \Delta A_B \\ &= \sum_{n=2}^{\infty} (-\beta)^{n-1} \sum_{m=1}^{n-1} \frac{1}{m!(n-m)!} \\ &\quad \times \langle \Delta H_1^m \Delta H_2^{n-m} \rangle_{0,c}. \end{aligned} \quad (15)$$

Consequently, the free energy change ΔA_C is equal to the sum of the free energy change ΔA_A and the full coupling term of Eq. (11), while ΔA_D is equal to the sum of the free-energy change ΔA_B and the full coupling term. From Fig. 1, ΔA_C is equal to the van der Waals free-energy component evaluated along path C in the presence of the full electrostatic interaction. Correspondingly, along path D the electrostatic free energy component is calculated for the system with the final state van der Waals interaction. Thus, for a mutation from a reference state to a final system that involves alteration of two components in the Hamiltonian, the path along which only one energy component is changed (e.g., path A), and the second component is kept at the reference value (e.g., 0), leads to a free energy component that does not have any contribution from the coupling term of Eq. (11). A path along which one component changes after the second component has been switched on (e.g., path C) leads to a free-energy component that includes the full coupling term.

For the linear path E, which has been referred to as the “alchemical path”,¹⁸ because the entire Hamiltonian is changed simultaneously and linearly as one chemical species is “transmuted” into another, we now consider how the

three terms of Eq. (11) are partitioned in the free-energy components ΔA_1 and ΔA_2 , when evaluated by Eq. (5). We write ΔA_1 in the form

$$\begin{aligned}\Delta A_1 &= \int_0^1 d\lambda \langle \Delta H_1 \rangle_\lambda \\ &= \int_0^1 d\lambda \frac{\int d\mathbf{r} e^{-\beta H_0} e^{-\beta \lambda (\Delta H_1 + \Delta H_2)} \Delta H_1}{\int d\mathbf{r} e^{-\beta H_0} e^{-\beta \lambda (\Delta H_1 + \Delta H_2)}} \\ &= \int_0^1 d\lambda \frac{\int d\mathbf{r} e^{-\beta H_0} e^{-\beta \lambda (\Delta H_1 + \Delta H_2)} \Delta H_1}{\int d\mathbf{r} e^{-\beta H_0} e^{-\beta \lambda (\Delta H_1 + \Delta H_2)}} \frac{\int d\mathbf{r} e^{-\beta H_0}}{\int d\mathbf{r} e^{-\beta H_0}} \\ &= \int_0^1 d\lambda \frac{\langle e^{-\beta \lambda (\Delta H_1 + \Delta H_2)} \Delta H_1 \rangle_0}{\langle e^{-\beta \lambda (\Delta H_1 + \Delta H_2)} \rangle_0}.\end{aligned}\quad (16)$$

Multiplication and division by the factor $\int d\mathbf{r} \exp(-\beta H_0)$ in Eq. (16) introduces averages evaluated at 0 (i.e., for the reference system) instead of averages evaluated at λ . To compare the component ΔA_1 , obtained by Eq. (16), with the cumulant expansion expression, Eq. (11), we introduce the auxiliary function

$$f(\kappa) = \ln \langle e^{-\beta \lambda (\kappa \Delta H_1 + \Delta H_2)} \rangle_0. \quad (17)$$

This function satisfies the relation

$$-\frac{1}{\beta \lambda} \frac{df(\kappa)}{d\kappa} \bigg|_{\kappa=1} = \frac{\langle e^{-\beta \lambda (\Delta H_1 + \Delta H_2)} \Delta H_1 \rangle_0}{\langle e^{-\beta \lambda (\Delta H_1 + \Delta H_2)} \rangle_0}. \quad (18)$$

From Eqs. (16) and (17), the component ΔA_1 evaluated along path E can be expressed in terms of $f(\kappa)$ as

$$\Delta A_1 = \int_0^1 d\lambda \left(-\frac{1}{\beta \lambda} \frac{df(\kappa)}{d\kappa} \bigg|_{\kappa=1} \right). \quad (19)$$

The function $f(\kappa)$ can be expanded in terms of cumulants by use of the cumulant expansion theorem

$$\begin{aligned}f(\kappa) &= \sum_{n=1}^{\infty} (-\beta \lambda)^n \sum_{m=0}^n \frac{\kappa^m}{m!(n-m)!} \\ &\quad \times \langle \Delta H_1^m \Delta H_2^{n-m} \rangle_{0,c}.\end{aligned}\quad (20)$$

Differentiation of $f(\kappa)$ with respect to κ yields

$$\begin{aligned}-\frac{1}{\beta \lambda} \frac{df(\kappa)}{d\kappa} \bigg|_{\kappa=1} &= \sum_{n=1}^{\infty} (-\beta \lambda)^{n-1} \sum_{m=1}^n \frac{1}{(m-1)!(n-m)!} \\ &\quad \times \langle \Delta H_1^m \Delta H_2^{n-m} \rangle_{0,c}.\end{aligned}\quad (21)$$

To calculate the free-energy component ΔA_1 , we need to integrate Eq. (21) with respect to λ , in accord with Eq. (19). Performing the integration, we find

$$\begin{aligned}\Delta A_1 &= \sum_{n=1}^{\infty} (-\beta)^{n-1} \sum_{m=1}^n \frac{m/n}{m!(n-m)!} \\ &\quad \times \langle \Delta H_1^m \Delta H_2^{n-m} \rangle_{0,c}.\end{aligned}\quad (22)$$

Separating out the term $m=n$, we have

$$\begin{aligned}\Delta A_1 &= \sum_{n=1}^{\infty} \frac{(-\beta)^{n-1}}{n!} \langle \Delta H_1^n \rangle_{0,c} + \sum_{n=2}^{\infty} (-\beta)^{n-1} \\ &\quad \times \sum_{m=1}^{n-1} \frac{m/n}{m!(n-m)!} \langle \Delta H_1^m \Delta H_2^{n-m} \rangle_{0,c}.\end{aligned}\quad (23)$$

The first term on the right-hand side of Eq. (23) is the total cumulant-expansion contribution of ΔH_1 to ΔA and corresponds to the free energy component ΔA_1 along path A, as in Eqs. (12) and (13). Thus, the component ΔA_1 evaluated along the linear path (E in Fig. 1) includes the full ΔA_1 term evaluated along path A, plus an additional contribution that is part of the coupling term in the total cumulant expansion Eq. (11). Similarly, for ΔA_2 (along path E) we find

$$\begin{aligned}\Delta A_2 &= \sum_{n=1}^{\infty} \frac{(-\beta)^{n-1}}{n!} \langle \Delta H_2^n \rangle_{0,c} + \sum_{n=2}^{\infty} (-\beta)^{n-1} \\ &\quad \times \sum_{m=1}^{n-1} \frac{m/n}{m!(n-m)!} \langle \Delta H_1^{n-m} \Delta H_2^m \rangle_{0,c}.\end{aligned}\quad (24)$$

The component ΔA_2 evaluated along path E includes a term equal to the free energy component contribution ΔA_2 evaluated along path B, plus an additional contribution from the coupling term of Eq. (11). Comparing Eqs. (23) and (24), we see that the coupling term is divided symmetrically along the concerted linear path; i.e., the coupling term in Eq. (24) is obtained from that in Eq. (23) by interchanging ΔH_1 and ΔH_2 . We expand Eqs. (23) and (24) to order $O(\beta^2)$ and obtain

$$\begin{aligned}\Delta A_1: & -\frac{\beta}{2!} \langle \Delta H_1 \Delta H_2 \rangle + \frac{\beta^2}{3!} \{2 \langle \Delta H_1^2 \Delta H_2 \rangle_{0,c} \\ & + \langle \Delta H_1 \Delta H_2^2 \rangle_{0,c}\} + \dots,\end{aligned}\quad (25)$$

$$\begin{aligned}\Delta A_2: & -\frac{\beta}{2!} \langle \Delta H_1 \Delta H_2 \rangle + \frac{\beta^2}{3!} \{ \langle \Delta H_1^2 \Delta H_2 \rangle_{0,c} \\ & + 2 \langle \Delta H_1 \Delta H_2^2 \rangle_{0,c} \} + \dots.\end{aligned}\quad (26)$$

The first term is partitioned symmetrically between ΔA_1 and ΔA_2 , as was already demonstrated in Ref. 18. In the higher order contributions the coupling terms are partitioned in a way that preserves the symmetry with respect to the interchange of components 1 and 2.

From the above analysis, the exact free-energy change in a transformation that involves two energy components which are altered can be expressed as a sum of three terms: Two terms correspond to the free-energy changes in processes where only one energy component is varied, while the other remains at the reference value, and the third term is a coupling term between the two energy components, evaluated at the reference state. This coupling term is partitioned between the two free-energy components in a path-dependent manner. Along the concerted linear path (E) the partitioning is symmetric with respect to the interchange of components 1 and 2. Along piecewise linear paths (A+D or B+C) the coupling term is projected onto the free-energy component that changes along the second step (D or C in the two examples

in Fig. 1), whereas the free energy evaluated along the first part of the path (A or B in Fig. 1) includes only a single energy component.

Although we do not use it here, the above procedure can be employed to derive a cumulant-expansion formula for the free-energy components in the case of a more general dependence of the Hamiltonian on the parameter λ . The resulting relation for component ΔA_1 is

$$\Delta A_1 = \sum_{n=1}^{\infty} \frac{(-\beta)^{n-1}}{(n-1)!} \left\langle \int_0^1 d\lambda \frac{\partial \Delta H_1(\lambda)}{\partial \lambda} \Delta H^{n-1}(\lambda) \right\rangle_{0,c} \quad (27)$$

For a Hamiltonian with two changing components and the functional dependence $\Delta H(\lambda) = f(\lambda)\Delta H_1 + g(\lambda)\Delta H_2$, Eq. (27) becomes

$$\Delta A_1 = \sum_{n=1}^{\infty} (-\beta)^{n-1} \sum_{m=0}^{n-1} \left\{ \left[\int_0^1 d\lambda f'(\lambda) f^m(\lambda) g^{n-m-1}(\lambda) \right] \times \frac{1}{m!(n-m-1)!} \langle \Delta H_1^{m+1} \Delta H_2^{n-m-1} \rangle_{0,c} \right\}. \quad (28)$$

The diagonal contribution (i.e., that involving only ΔH_1), is given by the term $m = n - 1$ of Eq. (28).

III. SYMMETRY OF THE LINEAR (ALCHEMICAL) PATH

In the previous section we showed that the concerted linear path led to a symmetric partitioning of the coupling term to all orders in the perturbation. We now show that inversion of the direction along the path changes only the sign of the free-energy components. This contrasts with what happens for piecewise-linear paths, where the magnitudes of the components depend on the choice of reference state.

According to Eq. (11), the total free energy change ΔA splits naturally into two terms that include cumulants of a single Hamiltonian component and a coupling term. The first two terms are calculated along paths A and B (see Fig. 1), and the coupling term is equal to the difference $\Delta A_C - \Delta A_A = \Delta A_D - \Delta A_B$. However, this definition of free energy components is associated with the choice of $\text{Cl}^{(0)}$ as the reference state. If $\text{Br}^{(-)}$ is chosen as the reference state (1), the analogous cumulant-expansion expression for the free energy change (in the opposite direction) is

$$\begin{aligned} -\Delta A &= -\frac{1}{\beta} \langle \exp[\beta(\Delta H_1 + \Delta H_2)] - 1 \rangle_{1,c} \\ &= -\sum_{n=1}^{\infty} \frac{\beta^{n-1}}{n!} \langle \Delta H_1^n \rangle_{1,c} - \sum_{n=1}^{\infty} \frac{\beta^{n-1}}{n!} \langle \Delta H_2^n \rangle_{1,c} \\ &\quad - \sum_{n=2}^{\infty} \beta^{n-1} \sum_{m=1}^{n-1} \frac{1}{m!(n-m)!} \langle \Delta H_1^m \Delta H_2^{n-m} \rangle_{1,c}. \end{aligned} \quad (29)$$

The first two terms are equal to the free energy differences $-\Delta A_C$ and $-\Delta A_D$ respectively, and the last is the coupling term. Consequently, if $\text{Br}^{(-)}$ is chosen as the reference state, paths C and D would, by the above argument, give the van

der Waals and electrostatic components respectively, of the mutation $\text{Br}^{(-)} \rightarrow \text{Cl}^{(0)}$. However, the resulting contributions are clearly different from these obtained with $\text{Cl}^{(0)}$ as the reference state. If we denote the coupling term of Eq. (11) by K^{for} and the corresponding term of Eq. (29) by K^{rev} , it is easily shown that $K^{\text{for}} = K^{\text{rev}}$; that is

$$\begin{aligned} -\Delta A &= \Delta A_{-C} + \Delta A_{-D} + K^{\text{rev}} \\ &= -\Delta A_A - K^{\text{for}} - \Delta A_B - K^{\text{for}} + K^{\text{rev}} \\ &= -\Delta A - K^{\text{for}} + K^{\text{rev}} \end{aligned} \quad (30)$$

Following the same reasoning for the choice of the free energy components and using the above analysis, we see that if we choose paths A+D (the forward mutation) or -D-A (the reverse mutation) and perform a free-energy decomposition we find the following results for the free energy components:

$$\text{Path A+D: } \Delta A_1 = \Delta A_A, \quad \Delta A_2 = \Delta A_B + K^{\text{for}},$$

$$\text{Path -D-A: } \Delta A_1 = \Delta A_{-C} + K^{\text{for}}, \quad \Delta A_2 = \Delta A_{-D}.$$

This shows that along the path A+D, ΔA_1 is equal to the pure van der Waals component (for this direction), but ΔA_2 differs from the pure electrostatic component by an amount equal to the coupling term K^{for} . Along the opposite direction, ΔA_2 coincides with the electrostatic component (for the second direction), and ΔA_1 differs by K^{for} .

Along the linear path E both components ΔA_1 and ΔA_2 include parts of the coupling term K^{for} . In the forward (E) and backward (-E) directions of path E component 1 satisfies the relation $\Delta A_1^{(\text{E})} = -\Delta A_1^{(-\text{E})}$. The same relation holds for component 2. To demonstrate this we consider the forward direction and write for $\Delta A_1^{(\text{E})}$

$$\begin{aligned} \Delta A_1^{(\text{E})} &= \int_0^1 d\lambda \frac{\int d\mathbf{r} e^{-\beta H_0 - \beta(\Delta H_1 + \Delta H_2) + \beta(1-\lambda)(\Delta H_1 + \Delta H_2)} \Delta H_1}{\int d\mathbf{r} e^{-\beta H_0 - \beta(\Delta H_1 + \Delta H_2) + \beta(1-\lambda)(\Delta H_1 + \Delta H_2)}} \\ &= \int_0^1 d\lambda \frac{\langle e^{\beta(1-\lambda)(\Delta H_1 + \Delta H_2)} \Delta H_1 \rangle_1}{\langle e^{\beta(1-\lambda)(\Delta H_1 + \Delta H_2)} \rangle_1}. \end{aligned} \quad (31)$$

This expression can be written in terms of cumulants by use of an auxiliary function $f(\kappa) = \ln \langle \exp[\beta(1-\lambda)(\kappa \Delta H_1 + \Delta H_2)] \rangle_1$. The final result is

$$\Delta A_1^{(\text{E})} = \sum_{n=1}^{\infty} \beta^{n-1} \sum_{m=0}^n \frac{m/n}{m!(n-m)!} \langle \Delta H_1^m \Delta H_2^{n-m} \rangle_{1,c}. \quad (32)$$

In the opposite direction, ΔA_1 is expressed as

$$\begin{aligned}\Delta A_1^{(-E)} &= \int_0^1 d\lambda \frac{\int d\mathbf{r} e^{-\beta H_0 - \beta(\Delta H_1 + \Delta H_2) + \beta\lambda(\Delta H_1 + \Delta H_2)} (-\Delta H_1)}{\int d\mathbf{r} e^{-\beta H_0 - \beta(\Delta H_1 + \Delta H_2) + \beta\lambda(\Delta H_1 + \Delta H_2)}} \\ &= - \int_0^1 d\lambda \frac{\langle e^{\beta\lambda(\Delta H_1 + \Delta H_2)} \Delta H_1 \rangle_1}{\langle e^{\beta\lambda(\Delta H_1 + \Delta H_2)} \rangle_1}.\end{aligned}\quad (33)$$

Equation (33) can be used in conjunction with the auxiliary function $f(\kappa) = \ln\langle \exp[\beta\lambda(\kappa\Delta H_1 + \Delta H_2)] \rangle_1$, to express $\Delta A_1^{(-E)}$ in terms of cumulants

$$\begin{aligned}\Delta A_1^{(-E)} &= - \sum_{n=1}^{\infty} \beta^{n-1} \sum_{m=0}^n \frac{m/n}{m!(n-m)!} \\ &\quad \times \langle \Delta H_1^m \Delta H_2^{n-m} \rangle_{1,c} = -\Delta A_1^{(E)}.\end{aligned}\quad (34)$$

Thus the free-energy components evaluated along the linear path E change sign if the direction of mutation is reversed. By associating the van der Waals and electrostatic free energy components with the values calculated along path E, we obtain a symmetric partition of the coupling term K^{for} between the two components. Further, the result does not depend on whether the initial or the final state is used as the reference state, since switching the reference state and the direction of the mutation only reverses the sign of the components, and does not affect their magnitudes. This supports the choice of the concerted linear path for the decomposition of the free-energy change into components.

IV. NUMERICAL EXAMPLE

In this section, the total free-energy change, as well as the corresponding free-energy components, which have been expressed as an infinite sum of cumulants [Eqs. (11), (23), and (24)] are used to study a model system and to determine the importance of the various contributions. The cumulants are used to determine the accuracy of approximating the total free-energy change as the sum of the first few terms. The partitioning into free-energy components of the coupling between energy terms in the cumulant expansion is examined and insights into the importance of the various physical processes involved are obtained. Cumulants through third order of the solvation free-energy change in the alchemical transformation $\text{Cl}^{(0)} \rightarrow \text{Br}^{(-)}$ in aqueous solution (see Fig. 1) are evaluated. The transformation involves a change in the van der Waals interaction and the electrostatic interaction between the monoatomic solute and the aqueous solvent. A concerted linear (alchemical) path and piecewise-linear paths are used to evaluate the free-energy change. The cumulant expansion, Eqs. (11), (23), and (24), are used to calculate total free-energy changes and the free-energy components.

The change in solvation free-energy in this alchemical transformation is calculated by the extended reference interaction site model (xRISM) methodology^{4,20,21} with the hypernetted-chain (HNC) closure,²² as in Ref. 19. The xRISM method, although approximate, permits calculations without statistical errors, which is important for determining accurate free energy derivatives. The solvent is represented by a modified three-site TIP3P model,²⁷ with bondlength

TABLE I. Potential parameters used in the xRISM calculations.^a

Atom	σ (Å)	ϵ (kcal/mol)	q (e)
H_w^b	0.40	0.046	0.417
O_w	3.15	0.152	-0.834
$\text{Cl}^{(0)}(\text{Cl}^{(-)})$	4.446	0.107	0.0(-1.0)
$\text{Br}^{(0)}(\text{Br}^{(-)})$	4.623	0.09	0.0(-1.0)

^aThe standard combination rules are used, $\sigma_{ab} = (\sigma_a + \sigma_b)/2$, $\epsilon_{ab} = \sqrt{\epsilon_a \epsilon_b}$.

^bModified to avoid a Coulombic singularity (Reference 28).

0.9572 Å and angle 104.52°; nonzero Lennard-Jones parameters were used for the water hydrogens to avoid a Coulombic singularity in the RISM-HNC calculations.²⁸ The solute-solvent interaction is a sum of the van der Waals interactions between the spherical solute and the solvent atoms, represented by a Lennard-Jones potential, and the solute-solvent Coulombic interactions for the charged spheres. The parameters used are listed in Table I; they are the same as in Ref. 19. The water density was set to 0.997 g/cm³ and the absolute temperature of the solution to 298.15 K.

The four corners of the thermodynamic cycle shown in Fig. 1 correspond to four different monoatomic solutes. The excess chemical potential of solvation for each solute was obtained from the corresponding solute-solvent radial distribution and pair-correlation functions from the formula³

$$\Delta\mu = \frac{\rho_v}{\beta} \sum_{v=1}^{N_v} \int d\mathbf{r} \left[\frac{1}{2} h_{uv}^2(r) - \frac{1}{2} h_{uv}(r) c_{uv}(r) - c_{uv}(r) \right], \quad (35)$$

where v labels the solvent sites, u is the monoatomic solute, $h_{uv} = g_{uv} - 1$ with g_{uv} the solute-solvent radial distribution function (between the solute and solvent site v), c_{uv} is the corresponding pair-correlation function, and ρ_v is the solvent density. The changes in free energy of solvation along paths A–E were obtained as the differences between the excess chemical potential of solvation of the corresponding monoatomic solutes; no thermodynamic integration was required. The free-energy components along path E were calculated by thermodynamic integration [Eq. (5)], using the solute-solvent radial distribution and pair-correlation functions for a series of intermediate systems; these functions were evaluated with the RISM method.^{20,21} The intermediate systems were monoatomic solutes with a solute-solvent vacuum interaction that changed linearly from the chloride–water to the bromide–water potential. The switching parameter λ was changed from 0.0 to 1.0 with a step of 0.01; i.e., a total of 101 hybrid solutes were considered. Similar calculations, with a hybrid potential and a step of 0.01 were performed along steps A–D, to obtain energy derivatives.

The first-order cumulants in the expression for the free-energy change along paths A–E are the averages $\langle \Delta H \rangle_\lambda$, where ΔH refers to the difference in van der Waals interaction energy between the initial and final state (paths A and C), the electrostatic difference (paths B and D), or their sum (path E). These averages were evaluated analytically from the relation

TABLE II. Free-energy derivatives along the mutation $\text{Cl}^{(0)} \rightarrow \text{Br}^{(-)}$.^a

Derivatives	Paths					
	A	B	C	D	E	
					vdW	elec
First ^b	1.23	-7.61	5.56	-1.17	1.23	-7.61
Second ^b	-0.28	-138.48	-2.34	-135.0	-0.10	-138.31
Third ^b	0.10	-75.78	1.98	-67.92	3.95	-68.0
First ^c	1.01	-194.73	3.91	-186.83	3.91	-186.83
Second ^c	-0.19	-219.77	-1.17	-210.24	5.31	-203.80
Third ^c	0.05	-35.18	0.68	-34.10	4.10	-20.2

^aAll quantities in kcal/mol. All values at 298.15 °C.^bFor each path, the derivatives are evaluated at the tail of the arrow (see Fig. 1 and text).^cFor each path, the derivatives are evaluated at the head of the arrow (see Fig. 1 and text). The signs at both arrow ends correspond to the direction from the tail to the head.

$$\langle \Delta H \rangle_\lambda = \rho_v \sum_{v=1}^{N_v} \int d\mathbf{r} g_{uv}(r; \lambda) \Delta H_{uv}(r), \quad (36)$$

where the radial distribution functions $g_{uv}(r; \lambda)$ are obtained from the RISM calculations. Higher order cumulants involve more than one pair of solute-solvent sites, and cannot be expressed in terms of solute-solvent radial pair distribution functions. Thus, they cannot be obtained directly from RISM calculations. However, they can be evaluated from the derivatives of the free-energy change with respect to the coupling parameter λ of the hybrid Hamiltonian. The derivatives are determined numerically from RISM calculations at a series of λ values. For a linear functional dependence of the Hamiltonian on λ the free-energy derivatives are proportional to the cumulants [see Eq. (A3)]. In particular, the first-order cumulant is equal to the first-order derivative $d\Delta A/d\lambda$.

The first-order cumulants obtained from Eq. (36) were compared to the numerical derivatives of the excess chemical potential of solvation at the reference state; the difference was less than 0.001 in all cases. A similar comparison was made between the value of the total free-energy change along any path, obtained from the difference in the excess chemical potential of solvation of the initial and final states and from the thermodynamic integration formula; the difference between the values from the two methods was less than 0.001 for all paths. The second-order derivatives were obtained by taking the numerical difference of the first derivatives between points that differ by 0.02 in λ . At the end points ($\lambda=0$ and $\lambda=1$) the derivatives were obtained by additional RISM calculations for states with $\lambda < 0$ and $\lambda > 1$ [outside the interval (0,1) we performed calculations for $0 \geq \lambda \geq -0.1$ and $1 \leq \lambda \leq 1.1$]. The third-order derivatives were calculated at the end points of each path by taking the numerical difference of the second-order derivatives between points that differ by 0.02 in λ . The results for the van der Waals and electrostatic free-energy component term for each path are given in Fig. 1. The values correspond to the infinite-order result in the cumulant expansion Eq. (14). The free-energy derivatives, up to third order, are listed in Table II. For each path the listed derivatives are determined at the

reference state value corresponding to the beginning or the end of the path (see Fig. 1); e.g., for path C the derivatives are evaluated at the states $\text{Cl}^{(-)}$ and $\text{Br}^{(-)}$.

The free-energy derivative values of Table II were used to calculate the cumulants of the energy components, up to third order. As an example, the second order cross-term at the reference state $\text{Cl}^{(0)}$ (i.e., 0) was obtained with the expression [see Eq. (A3)]

$$\begin{aligned} & -\beta \langle \Delta H_1 \Delta H_2 \rangle_{0,c} \\ &= \frac{1}{2} [-\beta \langle (\Delta H_1 + \Delta H_2)^2 \rangle_{0,c} + \beta \langle \Delta H_1^2 \rangle_{0,c} + \beta \langle \Delta H_2^2 \rangle_{0,c}] \\ &= \frac{1}{2} \left[\lim_{\lambda \rightarrow 0} \frac{d^2}{d\lambda^2} \Delta A(\lambda, \lambda) - \lim_{\lambda \rightarrow 0} \frac{d^2}{d\lambda^2} \Delta A(\lambda, 0) \right. \\ & \quad \left. - \lim_{\lambda \rightarrow 0} \frac{d^2}{d\lambda^2} \Delta A(0, \lambda) \right], \end{aligned} \quad (37)$$

and the first third-order cross-term at the same reference state $\text{Cl}^{(0)}$ (Table III) was calculated with the expression

$$\begin{aligned} & \beta^2 (2 \langle \Delta H_1^2 \Delta H_2 \rangle_{0,c} + \langle \Delta H_1 \Delta H_2^2 \rangle_{0,c}) \\ &= \beta^2 [\langle \Delta H_1^3 \rangle_{0,c} + (2 \langle \Delta H_1^2 \Delta H_2 \rangle_{0,c} \\ & \quad + \langle \Delta H_1 \Delta H_2^2 \rangle_{0,c})] - \beta^2 \langle \Delta H_1^3 \rangle_{0,c} \\ &= \lim_{\lambda \rightarrow 0} \frac{d^3}{d\lambda^3} \Delta A_1(\lambda, \lambda) - \lim_{\lambda \rightarrow 0} \frac{d^3}{d\lambda^3} \Delta A(\lambda, 0). \end{aligned} \quad (38)$$

The other third-order cross-term in Table III was obtained by the difference $\lim_{\lambda \rightarrow 0} (d^3/d\lambda^3) \Delta A_2(\lambda, \lambda) - \lim_{\lambda \rightarrow 0} (d^3/d\lambda^3) \Delta A(0, \lambda)$. The resulting values for the beginning points of the paths of Fig. 1 are listed in Table III, and the corresponding values for the end points in Table IV. The cumulants are multiplied by the coefficients that appear in the free-energy expansion, Eq. (14). The sums of the first two and three cumulants, which correspond to the two- and three-term approximation to the total free-energy change, are also in-

TABLE III. Cumulant contributions to the free-energy changes for the mutation $\text{Cl}^{(0)} \rightarrow \text{Br}^{(-)}$.^a The reference state 0 is at the beginning of each path.^b

Cumulant	Paths				
	A	B	C	D	E
$\langle \Delta H_1 \rangle_0$	1.23		5.56		1.23
$\langle \Delta H_2 \rangle_0$		-7.61		-7.47	-7.61
$-\beta/2 \langle \Delta H_1^2 \rangle_{0,c}$	-0.14		-1.17		-0.14
$-\beta/2 \langle \Delta H_2^2 \rangle_{0,c}$		-69.24		-67.51	-69.24
$-\beta \langle \Delta H_1 \Delta H_2 \rangle_{0,c}$					0.17
$\beta^2/3! \langle \Delta H_1^3 \rangle_{0,c}$	0.02		0.33		0.02
$\beta^2/3! \langle \Delta H_2^3 \rangle_{0,c}$		-12.63		-11.32	-12.63
$\beta^2/3! (2 \langle \Delta H_1^2 \Delta H_2 \rangle_{0,c} + \langle \Delta H_1 \Delta H_2^2 \rangle_{0,c})$					0.64
$\beta^2/3! (\langle \Delta H_1^2 \Delta H_2 \rangle_{0,c} + 2 \langle \Delta H_1 \Delta H_2^2 \rangle_{0,c})$					1.30
Two-term sum ^c	1.09	-76.85	4.39	-74.98	-75.58
Three-term sum ^d	1.11	-89.48	4.72	-86.30	-86.24
Total (infinite order sum) ^e	1.11	-94.10	4.64	-90.60	-89.46

^aAll quantities in kcal/mol. All values at 298.15 °C.^bFor all quantities the signs correspond to the direction from the tail to the head of the arrow (see Fig. 1), as in Table II.^cSum of first two terms [cumulants up to order $O(\beta)$] in cumulant expansion, Eq. (14).^dAs in c, but for cumulants up to $O(\beta^2)$.^eChange in free energy in going from the initial (tail) to the final (head) state.

cluded. The approximate values can be compared with the total free-energy changes, listed in the last rows of Tables III and IV.

We consider first the cumulants evaluated at the beginning points of the paths of Fig. 1 (i.e., at the tail of each arrow). The values are listed in Table III. Along path A, which corresponds to the van der Waals contribution for the uncharged spheres, the cumulants above second order are very small; i.e., the two-term approximation is very close to the total change. For path C, which is the van der Waals contribution for the charged spheres, higher order terms become significant; inclusion of the third-order term yields an

approximation that is within 2% of the total value. For both cases (paths A and C), even the first-order term (simple perturbation theory) gives the right order of magnitude. For paths B and D, which correspond to the introduction of a charge on the two spheres, the situation is very different. The first-order term contributes less than 10% of the total, and the third-order term $\beta^2 \langle \Delta H_2^3 \rangle_{0,c}$ is larger than the first order term. Comparing the result to infinite order with the value through third order, one sees that the remaining error is still 5%. An interesting point is that the linear response (Gaussian approximation) result, which corresponds to the summation through second order,^{29,30} is significantly in error. Linear re-

TABLE IV. Cumulant contributions to the free-energy changes for the mutation $\text{Cl}^{(0)} \rightarrow \text{Br}^{(-)}$.^a The reference state 1 is at the end of each path.^b

Cumulant	Paths				
	A	B	C	D	E
$\langle \Delta H_1 \rangle_1$	1.01		3.91		3.91
$\langle \Delta H_2 \rangle_1$		-194.73		-186.83	-186.83
$-\beta/2 \langle \Delta H_1^2 \rangle_{1,c}$	-0.10		-0.59		-0.59
$-\beta/2 \langle \Delta H_2^2 \rangle_{1,c}$		-109.89		-105.12	-105.12
$-\beta \langle \Delta H_1 \Delta H_2 \rangle_{1,c}$					6.46
$\beta^2/3! \langle \Delta H_1^3 \rangle_{1,c}$	0.01		0.11		0.11
$\beta^2/3! \langle \Delta H_2^3 \rangle_{1,c}$		-5.86		-5.68	-5.68
$\beta^2/3! (2 \langle \Delta H_1^2 \Delta H_2 \rangle_{1,c} + \langle \Delta H_1 \Delta H_2^2 \rangle_{1,c})$					0.57
$\beta^2/3! (\langle \Delta H_1^2 \Delta H_2 \rangle_{1,c} + 2 \langle \Delta H_1 \Delta H_2^2 \rangle_{1,c})$					2.32
Two-term sum ^c	-1.11	84.84	-4.50	81.71	83.67
Three-term sum ^d	-1.12	90.70	-4.61	87.39	86.35
Total (infinite order sum) ^e	-1.11	94.10	-4.64	90.60	89.46

^aAll quantities in kcal/mol. All values at 298.15 °C.^bFor the cumulants the signs correspond to the direction from the tail to the head of the arrow (see Fig. 1), as in Table II.^cSum of first two terms [cumulants up to order $O(\beta)$] in cumulant expansion, Eq. (14). The signs correspond to the free-energy changes along the direction from the head to the tail of the arrows of Fig. 1.^dAs in c, but for cumulants up to order $O(\beta^2)$.^eChange in free energy in going from the final (head) to the initial (tail) state.

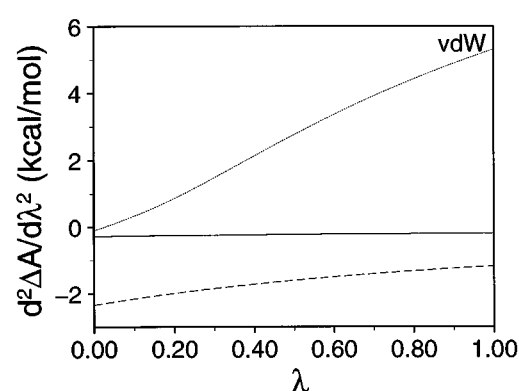
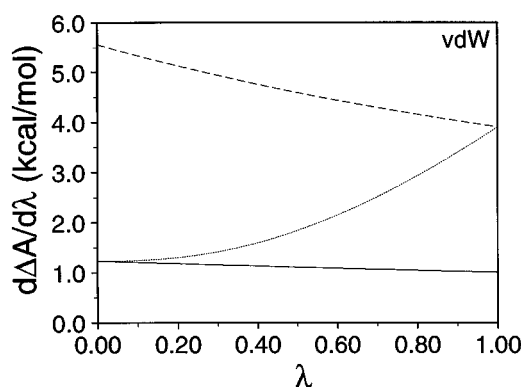
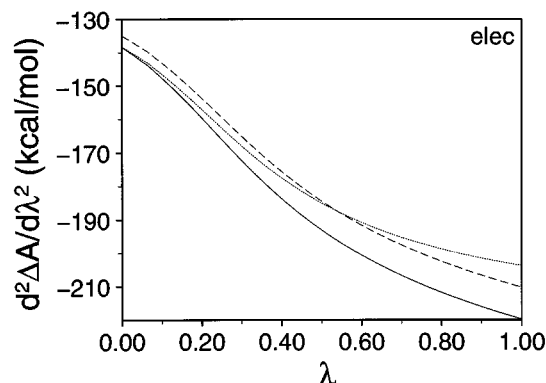
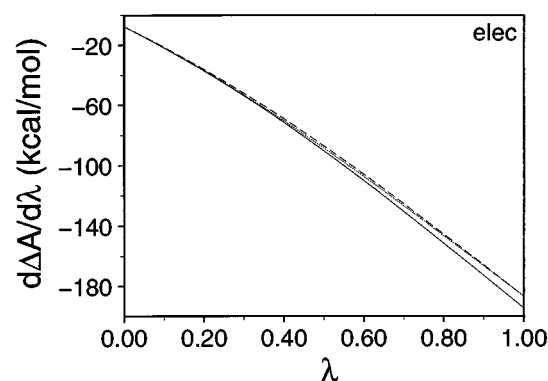


FIG. 2. First derivatives of the solvation free-energy change along the paths of Fig. 1. The derivatives are evaluated as the averages $\langle \Delta H \rangle_\lambda$. **vdw**: path A (—); path C (---); path E (···). **elec**: path B (—); path D (---); path E (···).

FIG. 3. Second derivatives of the solvation free-energy change along the paths of Fig. 1. The derivatives are evaluated numerically from the first derivatives $\langle \Delta H \rangle_\lambda$. **vdw**: path A (—); path C (---); path E (···). **elec**: path B (—); path D (---); path E (···).

sponse is a good approximation for a charge as large as ~ 0.4 (see below and Fig. 4). This is in accord with other calculations of electrostatic free-energy changes.^{29–31}

The significance of the higher order contributions can be illustrated by examining the variation of the free-energy derivatives with respect to λ along the various paths. A comparison of the derivatives at the beginning and end points of the paths is provided in Table I. In Fig. 2 we plot the first order free-energy derivatives, evaluated as the averages $\langle \Delta H \rangle_\lambda$ from the radial distribution functions of the hybrid solutes. The van der Waals derivatives along paths A and C, and the electrostatic derivatives along B, D, and E diminish with λ . The deviation of the first derivatives from their values at $\lambda=0$ is caused by the contribution of the higher-order cumulants. Along paths A through D these higher-order terms include cumulants of component ΔH_1 or ΔH_2 . The second order free-energy derivatives are plotted in Fig. 3. The electrostatic second derivatives diminish along λ due to the higher-order contributions, whereas the van der Waals derivatives increase (see also Table II). This implies that the “average” electrostatic second derivative along paths B, D is smaller than the initial value, whereas the average van der Waals derivative along A, D, is larger than the initial value. Thus, if the total free energy change along a path is approximated by a Taylor expansion around the beginning of the

path ($\lambda=0$), truncated at second order, the electrostatic free energy is underestimated and the van der Waals is overestimated. This is shown in Fig. 4, where we plot for each path the total free-energy change as a function of λ , together with the approximation by a second- and third-order Taylor expansion at the beginning of the path. For linear scaling, the n th order free-energy derivative is proportional to the n th order cumulant, as is shown in the Appendix [see Eqs. (A2) and (A3)]. Thus, the Taylor expansion is equivalent to the cumulant expansion, Eq. (11), with the cumulants evaluated at the point where the Taylor expansion is performed.

The cumulant values at the end points of the various paths of Fig. 1 are included in Table IV. For paths A and C the two- and three-term contributions are very close to the total sum. They are in closer agreement with the total free-energy changes than the approximate values obtained by the cumulants at the beginning points (see Table III). For paths B and D, the first-order term is larger by a factor of 2 (in absolute value) than the total free-energy change. This is in contrast to the situation observed in the beginning points of the “electrostatic” paths (A, D), and the concerted linear path E, where the average first electrostatic derivative is evaluated with the radial distribution function of the uncharged spheres, and gives a result that is much smaller than the total free-energy change. The second derivative corrects

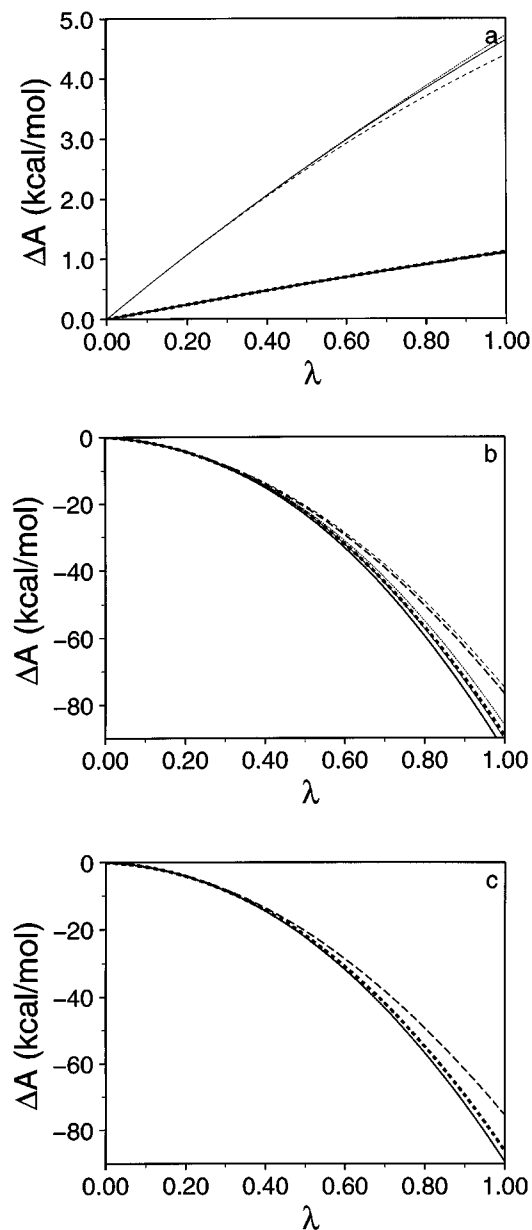


FIG. 4. Comparison between the total solvation free-energy change along the paths of Fig. 1 and a Taylor expansion at the beginning of each path of the free-energy change, truncated at second and third order. The Taylor expansion is equivalent to the cumulant expansion, Eq. (11), truncated at order $O(\beta)$ and $O(\beta^2)$, respectively, with the cumulants evaluated at the beginning of each path (see text). (a) Path A: total (—); second order (---); third order (- - -). Path C: total (—); second order (- - -); third order (···). (b) Path B: total (—); second order (---); third order (- - -). Path D: total (—); second order (---); third order (···). (c) path E: total (—); second order (---); third order (- - -).

the overestimation of the total free-energy change by the first derivative. Inclusion of the third derivative brings the result in very close agreement with the total value. As for the case of paths A and C, the agreement is slightly improved with respect to the results of Table III.

The results from this numerical example can be used to assess the accuracy of approximating the total free-energy change by the few first cumulants. This assessment is of

interest, because it has been proposed that improved ligands can be designed based on the first-order approximation to the free energy (first-order perturbation theory); i.e., by use of the first free-energy derivative with respect to terms in the Hamiltonian that describe the interaction between the ligand and the receptor.^{16,17} This approximation has the great advantage that a single simulation for the unperturbed system can be used to evaluate all of the derivatives. Such a study has been made by Gerber and van Gunsteren¹⁶ for the binding of trimethoprin to dihydrofolate reductase. The method was tested by comparing the measured binding of related compounds with that estimated from the perturbation results. In employing the latter, Gerber and van Gunsteren used not only the individual first derivatives, which often do not correspond to physically realizable changes, but also the results from several derivatives introduced at the same time. Because the first-order approximation was used, no coupling is present, since, as we have shown, the coupling terms appear only in second order. Comparison of the first-order estimates with the measured total free-energy differences for the related compounds revealed little correlation between the two.¹⁶

This lack of correlation is readily explained by the results of the numerical example presented here. We compare the first-derivative results with those obtained from a full calculation. Analysis of the transformation $\text{Cl}^{(0)} \rightarrow \text{Br}^{(-)}$ is of interest in this regard because it involves both van der Waals and electrostatic terms, which are the interactions that contribute to ligand binding. With the exception of path A, which corresponds to a small change in the solute-solvent van der Waals interaction, the values obtained from the first derivatives are very different from the total free-energy changes. In path A the solvent distribution around the van der Waals solute changes to a much smaller degree than in the other paths, along which the solute is charged. This indicates that the first derivative might be a satisfactory approximation to the total free-energy change if the environment does not change appreciably during the perturbation (e.g., for a small mutation of a ligand in a nonflexible active site). For the electrostatic paths (B and D) a calculation at the uncharged initial state underestimates the total free energy by more than 90%. However, the sum of the diagonal contribution of the derivatives through the second order term gives a result (80% of total) that could be sufficiently accurate for the purpose of ligand design. Further investigation of this possibility appears worthwhile. The effect of the higher-order terms is demonstrated by the large difference between the first derivatives evaluated at the beginning and the end of the various paths.

We now consider the various contributions along the concerted linear path E, making use of Tables III and IV, and the figures. In contrast to the piecewise linear paths (A+C, B+D), path E contains contributions from terms that include two components and first appear in second order (see Tables III and IV). Considering first the values of Table III, we see that the second-order coupling term evaluated at $\text{Cl}^{(0)}$ is very small (0.17 kcal/mol), but that the third-order terms are larger. By subtracting the free-energy differences along paths

C and A, or D and B [Eq. (15)], it follows that the total coupling term is equal to 3.52 kcal/mol. This is significantly larger than the value through third order (2.11 kcal/mol). Along the linear path E this term is split into contributions of 0.98 and 2.54 kcal/mol to ΔA_1 and ΔA_2 , respectively. That the portion of the coupling term associated with the electrostatic component is larger in magnitude than that associated with the van der Waals component arises from the contributions beyond second order and the fact that the electrostatic perturbation is larger than the van der Waals perturbation. The third-order contribution to the van der Waals free-energy component is equal to 0.64 kcal/mol, whereas the corresponding electrostatic term is 1.30 kcal/mol.

Along path E the free-energy derivatives include also cumulants that contain products of the two energy components [see Eq. (11)]. These terms cause the van der Waals derivative $\langle \Delta H_1 \rangle_\lambda$ along path E to vary much faster than along paths A or C, as is shown in Fig. 2. At the endpoint of path E, the second-order cross-term is much larger (6.46 kcal/mol) than the total coupling term (see Table IV). This is in accord with the large values of the first electrostatic derivatives at the end points of paths B, D, and E. The value of the coupling term through third order is equal to 3.57 kcal/mol, i.e., very close to the total value (3.52 kcal/mol).

To examine the partitioning of the coupling term into the van der Waals and electrostatic components as a function of λ , we consider the thermodynamic cycles with corners (0,0), $(\lambda,0)$, $(0,\lambda)$, and (λ,λ) for all intermediate λ values. The total coupling term as a function of λ results from the double differences along parallel, partially completed (from 0 to λ) paths, and the partitioning is determined by evaluating the free energy components along the diagonal path $(0,0) \rightarrow (\lambda,\lambda)$ and subtracting the total free-energy changes along the corresponding piecewise-linear paths $[(0,0) \rightarrow (\lambda,0)$ and $(0,0) \rightarrow (0,\lambda)]$. The resulting partitionings, along with the total coupling term, are plotted in Fig. 5(a). The difference between the electrostatic and van der Waals partitioning of the coupling term grows with λ . In Fig. 5(b) we plot the difference between the electrostatic second-order free-energy derivatives evaluated along paths E and A, and the corresponding van der Waals difference for the derivatives along E and B. These double differences are equal to

$$\begin{aligned} \text{vdW: } & -\beta \langle \Delta H_1 \Delta H_2 \rangle_{0,c} + \beta^2 \lambda [2 \langle \Delta H_1^2 \Delta H_2 \rangle_{0,c} \\ & + \langle \Delta H_1 \Delta H_2^2 \rangle_{0,c}] + \dots, \\ \text{elec: } & -\beta \langle \Delta H_1 \Delta H_2 \rangle_{0,c} + \beta^2 \lambda [\langle \Delta H_1^2 \Delta H_2 \rangle_{0,c} \\ & + 2 \langle \Delta H_1 \Delta H_2^2 \rangle_{0,c}] + \dots. \end{aligned}$$

Thus, at $\lambda=0$ the differences converge to the cross-correlation between the two components, that is equal to 0.17 (Table III). The electrostatic difference grows faster [see Fig. 5(b)] due to the higher magnitude of the electrostatic cumulants.

The coupling term is projected on the van der Waals contribution along the path B+C, and on the electrostatic contribution along the path A+D. The difference between the van der Waals contribution for the change $\text{Cl}^{(0)}$ to $\text{Br}^{(0)}$

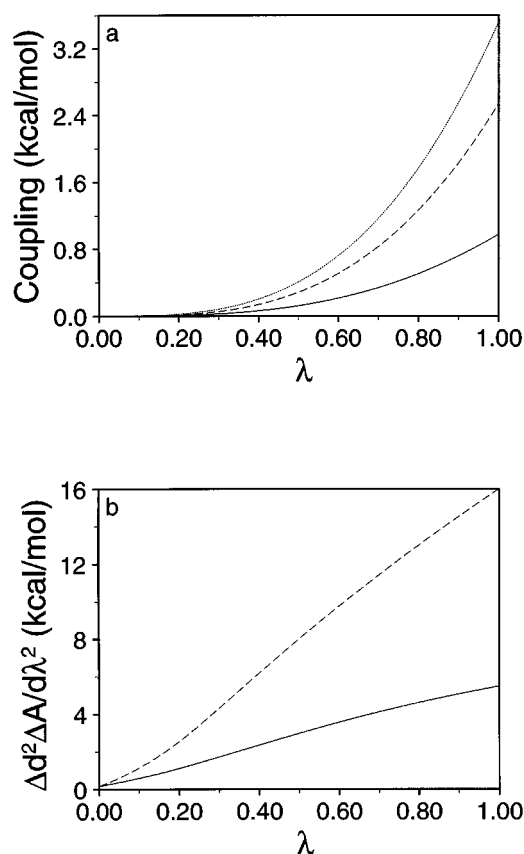


FIG. 5. Partition of the coupling term in the free-energy components along path E. vdw (—); elec (---); total (···). (b) Second free-energy derivative, evaluated from the numerical derivative of the first derivative $\langle \Delta H \rangle_\lambda$. The differences between the values along path E (van der Waals derivative) and A (—) or path E (electrostatic derivative) and B (---) are plotted. At $\lambda=0$ both differences converge to the value $-\beta \langle \Delta H_1 \Delta H_2 \rangle_{0,c}$ (see the text).

(1.12 kcal/mol), and for the change $\text{Cl}^{(-)}$ to $\text{Br}^{(-)}$ (4.65 kcal/mol), is exactly the coupling term between the electrostatic and van der Waals components in the cumulant expansion, Eq. (11). If the piecewise paths B+C are used, the coupling term does not appear explicitly. Instead it is entirely subsumed in the $\langle \Delta H_1^2 \rangle_{0,c}$ contributions along path C, which are significantly larger than the corresponding contributions along path A (see Table III). A complementary description of this difference in terms of the radial distribution function for the neutral and charged solutes was given in Ref. 19.

V. CONCLUSIONS

Free-energy simulations that use thermodynamic integration permit the decomposition of the total free-energy change into components.¹⁰ As has been pointed out, the free-energy components depend on the integration path.^{11,24,32,33} We have employed an infinite-order cumulant expansion of the free-energy change for two components to evaluate the dependence of the free-energy components along piecewise-linear paths and the concerted linear (alchemical) path. In this way the total free-energy change can be expressed as a sum of terms that contain the cumulants of a single energy component, and a coupling term between pairs of compo-

nents. Along piecewise-linear paths the coupling term is projected on one free-energy component. The concerted linear path on the other hand provides a symmetric partitioning of this coupling term into free-energy components. This had been demonstrated to second order in earlier work.¹⁸ The extension to infinite order provides the theory necessary for a full treatment of free-energy components. The cumulant expansion also permits us to compare the first-order change in the free energy involving a single parameter with higher-order, including infinite-order, results. It is shown that for some perturbations (e.g., change in the charge), the first-order approximation is very poor but the value obtained by including terms through third order is accurate to 5%.

ACKNOWLEDGMENT

This work was supported in part by a grant from the National Institutes of Health.

APPENDIX: FREE-ENERGY DERIVATIVES

In this Appendix we use the cumulant-expansion method to derive an expression for the n th order derivative of the free-energy change $\Delta A^{(n)}(\lambda) \equiv d^{(n)}A(\lambda)/d\lambda^n$ with respect to the parameter λ , for the case of a general functional dependence of the Hamiltonian on a single variable λ . Smith and van Gunsteren evaluated derivatives of the free-energy change $\Delta A(\lambda)$ with respect to λ .³¹ They derived a formula for the n th derivative $\Delta A^{(n)}(\lambda)$ for the special case of a linear dependence $\Delta H(\lambda) = \lambda \Delta H$, and provided expressions for the up to 5th order derivatives in the case of a general functional dependence of $\Delta H(\lambda)$. They used higher order free-energy derivatives, evaluated from a single simulation at the reference state, and a Taylor expansion of $\Delta A(\lambda)$, to estimate the total free-energy change from a single simulation. The expression derived in this Appendix generalizes the formulas obtained in Ref. 31 to all orders.

To derive the general formula we employ a Taylor expansion of the free energy for a change relative to the reference value λ ,

$$\Delta\Delta A(\lambda) \equiv \Delta A(\lambda + \delta\lambda) - \Delta A(\lambda) = \sum_{n=1}^{\infty} \frac{1}{n!} \Delta A^{(n)}(\lambda) (\delta\lambda)^n. \quad (\text{A1})$$

For linear scaling the change in the Hamiltonian around λ is equal to $\delta\lambda \Delta H$. From Eq. (A1) and the cumulant-expansion for $\Delta\Delta A(\lambda)$ at the reference state λ and for linear scaling

$$\Delta\Delta A(\lambda) = \sum_{n=1}^{\infty} \frac{(-\beta)^{n-1}}{n!} (\delta\lambda)^n \langle \Delta H^n \rangle_{\lambda,c}, \quad (\text{A2})$$

it follows that the n th derivative of ΔA for the case of linear scaling is equal to

$$\Delta A^{(n)}(\lambda) = (-\beta)^{n-1} \langle \Delta H^n \rangle_{\lambda,c}. \quad (\text{A3})$$

This expression has been derived by Smith and van Gunsteren.³¹ To derive a formula for the n th free-energy derivative of the free energy for the case of a general functional dependence $\Delta H(\lambda)$ we use the Taylor expansion:

$$\Delta H(\lambda + \delta\lambda) = \sum_{m=0}^{\infty} \frac{1}{m!} \Delta H^{(m)}(\lambda) (\delta\lambda)^m, \quad (\text{A4})$$

and substitute in the cumulant expansion for $\Delta\Delta A(\lambda)$ to deduce

$$\begin{aligned} \Delta\Delta A(\lambda) &= \sum_{n=1}^{\infty} \frac{(-\beta)^{n-1}}{n!} \langle (\Delta H(\lambda + \delta\lambda) - \Delta H(\lambda))^n \rangle_{\lambda,c} \\ &= \sum_{n=1}^{\infty} \frac{(-\beta)^{n-1}}{n!} \left\langle \left[\sum_{m=1}^{\infty} \frac{\Delta H^{(m)}(\lambda)}{m!} (\delta\lambda)^m \right]^n \right\rangle_{\lambda,c} \\ &= \frac{1}{\beta} - \sum_{n=0}^{\infty} \frac{1}{n!} \left[\sum_{k_1}^{\infty} \cdots \sum_{k_m}^{\infty} \cdots \delta \left(\sum_{i=1}^{\infty} i k_i - n \right) \right. \\ &\quad \times (-\beta)^{-1 + \sum_{i=1}^{\infty} k_i} \frac{n!}{\prod_{i=1}^{\infty} k_i! \prod_{i=1}^{\infty} (i!)^{k_i}} \\ &\quad \left. \times \left\langle \prod_i \Delta H^{(i)}(\lambda)^{k_i} \right\rangle_{\lambda,c} \right] (\delta\lambda)^n. \end{aligned} \quad (\text{A5})$$

The term inside the last brackets [...] in Eq. (43) is equal to the n th derivative of ΔA . The function $\delta(\sum_{i=1}^{\infty} i k_i - n)$ is a Kronecker delta that constrains the sum of terms $i k_i$ to be equal to the order n of the n th derivative. Employing Eq. (43) we find for the three first derivatives:

$$\begin{aligned} \Delta A^{(1)}(\lambda) &= \langle \Delta H^{(1)}(\lambda) \rangle_{\lambda,c}, \\ \Delta A^{(2)}(\lambda) &= -\beta \langle \Delta H^{(1)}(\lambda)^2 \rangle_{\lambda,c} + \langle \Delta H^{(2)}(\lambda) \rangle_{\lambda,c}, \\ \Delta A^{(3)}(\lambda) &= \beta^2 \langle \Delta H^{(1)}(\lambda)^3 \rangle_{\lambda,c} - 3\beta \langle \Delta H^{(1)}(\lambda) \Delta H^{(2)}(\lambda) \rangle_{\lambda,c} \\ &\quad + \langle \Delta H^{(3)}(\lambda) \rangle_{\lambda,c}. \end{aligned}$$

These formulas coincide with the expressions given in the Appendix a of Smith and van Gunsteren.³¹

¹J. Gao, K. Kuczera, B. Tidor, and M. Karplus, *Science* **244**, 1069 (1989).

²S. Boresch and M. Karplus, *J. Chem. Phys.* **105**, 5145 (1996).

³S. J. Singer and D. Chandler, *Mol. Phys.* **55**, 621 (1985).

⁴F. Hirata and P. J. Rossky, *Chem. Phys. Lett.* **83**, 329 (1981).

⁵B. M. Pettitt, M. Karplus, and P. J. Rossky, *J. Phys. Chem.* **90**, 6335 (1986).

⁶A. S. Yang and B. Honig, *J. Mol. Biol.* **252**, 351 (1995).

⁷A. D. MacKerell, Jr., M. S. Sommer, and M. Karplus, *J. Mol. Biol.* **247**, 774 (1995).

⁸G. Kaminski, E. M. Duffy, T. Matsui, and W. J. Jorgensen, *J. Phys. Chem.* **98**, 13077 (1994).

⁹J. Forsman and B. Jonsson, *J. Chem. Phys.* **101**, 5116 (1994).

¹⁰B. Tidor and M. Karplus, *Biochemistry* **30**, 3217 (1991).

¹¹P. E. Smith and W. F. van Gunsteren, *J. Phys. Chem.* **98**, 13735 (1994).

¹²J. G. Kirkwood, *J. Chem. Phys.* **3**, 300 (1935).

¹³L. X. Dang, K. M. Merz, and P. A. Kollman, *J. Am. Chem. Soc.* **111**, 8505 (1989).

¹⁴F. T. K. Lau and M. Karplus, *J. Mol. Biol.* **236**, 1049 (1994).

¹⁵T. Simonson and A. Brunger, *Biochemistry* **31**, 8661 (1992).

¹⁶P. Gerber, A. Mark, and W. F. van Gunsteren, *J. Comp. Aid. Mol. Des.* **7**, 305 (1993).

¹⁷Y.-P. Pang and P. A. Kollman, *Perspective in Drug Discovery and Design*, (ESCOM Science Publishers, 1995), Vol. 3.

¹⁸S. Boresch and M. Karplus, *J. Mol. Biol.* **254**, 801 (1995).

- ¹⁹S. Boresch, G. Archontis, and M. Karplus, *Proteins: Struct., Funct. Gen.* **20**, 25 (1994).
- ²⁰H. C. Andersen and D. Chandler, *J. Chem. Phys.* **57**, 1918 (1972).
- ²¹D. Chandler and H. C. Andersen, *J. Chem. Phys.* **57**, 1930 (1972).
- ²²D. A. McQuarrie, *Statistical Mechanics* (Harper and Row, New York, 1976).
- ²³S. K. Ma, *Statistical Mechanics* (World Scientific, Singapore, 1985).
- ²⁴A. E. Mark and W. F. van Gunsteren, *J. Mol. Biol.* **240**, 167 (1994).
- ²⁵R. W. Zwanzig, *J. Chem. Phys.* **22**, 1420 (1954).
- ²⁶P. K. Weiner W. van Gunsteren and A. J. Wilkinson, *Computer Simulation of Biomolecular Systems* (ESCOM Science Publishers, 1993), Vol. 2.
- ²⁷W. L. Jorgensen, *J. Am. Chem. Soc.* **103**, 335 (1981).
- ²⁸B. M. Pettitt and P. J. Rossky, *J. Chem. Phys.* **77**, 1451 (1982).
- ²⁹T. Simonson and D. Perahia, *J. Am. Chem. Soc.* **117**, 7987 (1995).
- ³⁰R. M. Levy, M. Belhadj, and D. Kitchen, *J. Chem. Phys.* **95**, 3627 (1991).
- ³¹P. E. Smith and W. F. van Gunsteren, *J. Chem. Phys.* **100**, 577 (1994).
- ³²S. Yun-yu, A. E. Mark, W. Cun-xin, H. Fuhua, H. J. C. Berendsen, and W. F. van Gunsteren, *Protein Eng.* **6**, 289 (1993).
- ³³G. P. Brady and K. A. Sharp, *J. Mol. Biol.* **254**, 77 (1995).