

## COMMUNICATION

## On the Decomposition of Free Energies

G. Patrick Brady<sup>1</sup>, Attila Szabo<sup>2</sup> and Kim A. Sharp<sup>1\*</sup>

<sup>1</sup>Dept. of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia PA 19104-6059, USA

<sup>2</sup>Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Disease, National Institutes of Health, Bethesda MD 20892-0520, USA

The decomposition of free energies and entropies into components has recently been discussed within the framework of the free energy perturbation (FEP) and thermodynamic integration (TI) methods. In FEP, the cumulant expansion of the excess free energy contains coupling terms in second and higher orders. It is shown here that this expansion can be expressed in terms of temperature derivatives of the mean energy, suggesting a natural decomposition of the free energy into components corresponding to each term in the Hamiltonian. This result is derived in such a way that it establishes the equivalence to a particular form of component analysis based on TI in which all terms in the interaction energy are turned on simultaneously using  $1/kT$  as the coupling parameter.

© 1996 Academic Press Limited

**Keywords:** free energy components; entropy decomposition; enthalpy decomposition; free energy separability

\*Corresponding author

The decomposition of the total free energy of a system into components has been a subject of much interest and discussion (Bash *et al.*, 1987; Boresch *et al.*, 1994; Brady & Sharp, 1995; Gao *et al.*, 1989; Mark & van Gunsteren, 1994; Shi *et al.*, 1993; Smith & van Gunsteren, 1994; Yu & Karplus, 1988). Suppose that the total potential energy is expressed as a sum of terms  $U = \sum U_j$ . The mean configurational energy can then be written as:

$$E^{\text{conf}} = \langle U \rangle = \langle \sum_j U_j \rangle = \sum_j \langle U_j \rangle = \sum_j E_j \quad (1)$$

where  $\langle \dots \rangle$  denotes a canonical ensemble average  $\langle (\dots) \rangle = \int (\dots) e^{-\beta U(q)} dq / \int e^{-\beta U(q)} dq$  where  $\beta = 1/kT$ ,  $T$  is the absolute temperature, and  $k$  is Boltzmann's constant. Equation (1) shows that the mean energy can be decomposed into specific components,  $E_j$ , arising from the corresponding terms in the potential function representing specific types of interactions. It should be emphasized that, even for the mean energy, the above decomposition does not imply that each component is independent. Since the ensemble average involves the total potential, each  $\langle U_i \rangle$  clearly has an implicit dependence on all the other terms in the potential function. Can one find an analogous decomposition of the free energy?

The excess free energy  $A^{\text{ex}}$  (i.e. the free energy relative to an ideal gas) is:

$$\beta A^{\text{ex}} = \ln \frac{\int dq}{\int e^{-\beta U(q)} dq} = \ln \left( \frac{Q(0)}{Q(\beta)} \right) = \ln \frac{\int e^{\beta U(q)} e^{-\beta U(q)} dq}{\int e^{-\beta U(q)} dq} \quad (2a)$$

$$= \ln \langle e^{\beta U(q)} \rangle \quad (2b)$$

where  $Q(0) = \int dq = V^N$  is the partition function of  $N$  independent particles confined to a volume  $V$ . The cumulant expansion of the excess free energy, obtained by expanding equation (2b) in powers of  $\beta$  is (Zwanzig, 1954):

$$A^{\text{ex}} = \langle U \rangle + \frac{\beta}{2!} \langle \Delta U^2 \rangle + \frac{\beta^2}{3!} \langle \Delta U^3 \rangle + \frac{\beta^3}{4!} (\langle \Delta U^4 \rangle - \langle \Delta U^2 \rangle^2) + \dots \quad (3)$$

where  $\Delta U = U - \langle U \rangle$ . When  $U = \sum U_j$  the second-order term contains averages of the type  $\langle U_i U_j \rangle$  so it appears that the excess free energy is inherently non-separable. However, by direct algebraic manipulation one finds that the second and third moment terms in the cumulant expansion are related to temperature derivatives of the mean energy (Brady & Sharp, 1995), namely:

$$\langle \Delta U^2 \rangle = -\frac{\partial \langle U \rangle}{\partial \beta} \quad (4a)$$

Abbreviations used: TI, thermodynamic integration; FEP, free energy perturbation; TDE, temperature derivative expansion.

$$\langle \Delta U^3 \rangle = \frac{\partial^2 \langle U \rangle}{\partial \beta^2} \quad (4b)$$

If this were true to all orders, then one could express  $A^{\text{ex}}$  solely in terms of  $\langle U \rangle$  and its temperature derivatives. Since:

$$\frac{\partial^n \langle U \rangle}{\partial \beta^n} = \frac{\partial^n \sum_j \langle U_j \rangle}{\partial \beta^n} = \sum_j \frac{\partial^n \langle U_j \rangle}{\partial \beta^n} \quad (5)$$

this would lead to a natural decomposition of the excess free energy in terms of the average energy of each component and its temperature derivatives.

We now prove the above result to all orders, in such a way that we establish the connection with thermodynamic integration (TI). Since the mean energy is given by

$$\langle U \rangle_\beta = - \frac{\partial \ln Q(\beta)}{\partial \beta},$$

it immediately follows that:

$$\int_0^\beta \langle U \rangle_{\beta'} d\beta' = \ln \left( \frac{Q(0)}{Q(\beta)} \right) = \beta A^{\text{ex}} \quad (6)$$

Expanding  $\langle U \rangle_{\beta'}$  in a Taylor series about  $\beta$  gives:

$$\langle U \rangle_{\beta'} = \sum_{n=0}^{\infty} \frac{(\beta' - \beta)^n}{n!} \frac{\partial^n \langle U \rangle_\beta}{\partial \beta^n} \quad (7)$$

Substituting this expansion into equation (6) and integrating with respect to  $\beta'$ , we have:

$$\begin{aligned} A^{\text{ex}} &= \sum_{n=0}^{\infty} \frac{(-\beta)^n}{(n+1)!} \frac{\partial^n \langle U \rangle_\beta}{\partial \beta^n} \\ &= \langle U \rangle_\beta - \frac{1}{2!} \frac{\partial \langle U \rangle_\beta}{\partial \beta} + \frac{1}{3!} \frac{\partial^2 \langle U \rangle_\beta}{\partial \beta^2} - \frac{1}{4!} \frac{\partial^3 \langle U \rangle_\beta}{\partial \beta^3} \dots \quad (8) \end{aligned}$$

which formally expresses the free energy in terms of  $\langle U \rangle$  and its temperature derivatives. When  $U = \sum U_j$  it follows from equations (5) and (8) that the excess free energy is separable into terms corresponding to each term in the potential energy in the same sense that the mean energy is:

$$A^{\text{ex}} = \sum_j \left( \sum_{n=0}^{\infty} \frac{(-\beta)^n}{(n+1)!} \frac{\partial^n \langle U_j \rangle_\beta}{\partial \beta^n} \right) = \sum_j A_j^{\text{ex}} \quad (9)$$

Moreover, since equation (8) was obtained using equation (6) as a starting point, it immediately follows that:

$$\beta A_j^{\text{ex}} = \int_0^\beta \langle U_j \rangle_{\beta'} d\beta' \quad (10)$$

thus establishing the equivalence between the decomposition of the free energy based on the temperature derivative expansion (equation (8)) and component analysis using a particular form of thermodynamic integration (equations (6) and (10)). This form of TI turns on all components of the

potential simultaneously using  $\beta$  as a linear coupling parameter.

Inserting  $U = \sum U_j$  into the cumulant expansion (equation (3)) gives rise to terms that represent explicit coupling between free energy components, namely  $\langle \Delta U_i \Delta U_j \rangle$  in second order,  $\langle \Delta U_i^2 \Delta U_j \rangle$ ,  $\langle \Delta U_i \Delta U_j^2 \rangle$  and  $\langle \Delta U_i \Delta U_j \Delta U_k \rangle$  in third order, and so on. Even though no explicit coupling terms appear in the equivalent temperature derivative expansion (TDE), equation (8), these must of course be present. The decomposition in equation (9) is, in effect, a prescription for partitioning all coupling terms. For example direct algebraic manipulation shows that the term  $\langle \Delta U_i \Delta U_j \rangle$  is partitioned equally between  $A_i^{\text{ex}}$  and  $A_j^{\text{ex}}$ ,  $\langle \Delta U_i^2 \Delta U_j \rangle$  is partitioned 2:1 between the  $i$ th and  $j$ th excess free energy components, while  $\langle \Delta U_i \Delta U_j \Delta U_k \rangle$  is partitioned equally among the  $i$ th,  $j$ th and  $k$ th components. In the context of TI, the identical partitioning has previously been found to second order (Boresch & Karplus, 1995) and to all orders by a cumulant expansion (G. Archontis & M. Karplus, unpublished results), as is to be expected, since equations (9) and (10) are equivalent.

It should be emphasized that the decomposition of the excess free energy into components based on the TDE, although equivalent to a form of TI, is not the same as that in conventional TI. In conventional TI the integration parameter  $\lambda$  specifies the extent to which the potential function resembles the initial ( $\lambda = 0$ ) state,  $U^i$ . One defines  $U(\lambda) = U^i + \lambda(U^f - U^i) = U^i + \lambda \delta U$ , so that  $\lambda = 1$  corresponds to the final state. If the perturbation is applied uniformly to all the components of the Hamiltonian that are changing,  $\delta U$ , then the decomposition of the difference in free energy between the initial and final states,  $\Delta A$ , is given by (Boresch *et al.*, 1994; Boresch & Karplus, 1995; Brooks, 1990)

$$\begin{aligned} \Delta A &= \int_0^1 \langle \delta U \rangle_\lambda d\lambda = \int_0^1 \left\langle \sum_j \delta U_j \right\rangle_\lambda d\lambda \\ &= \sum_j \int_0^1 \langle \delta U_j \rangle_\lambda d\lambda = \sum_j \Delta A_j \quad (11) \end{aligned}$$

In the TI component decomposition given by equation (11), only contributions corresponding to Hamiltonian components that change between the two states, i.e. that are part of  $\delta U$ , appear in the free energy change  $\Delta A$ . To illustrate the difference between the TDE and conventional TI decompositions, consider the simple example clearly discussed by Boresch *et al.* (1994), namely the free energy difference between  $\text{Cl}^-$  and  $\text{Br}^-$  in water. If one calculates the free energy change using TI by changing only the ion radius, from equation (11) one would necessarily attribute all the free energy change to the water-ion van der Waals interaction, since this is the only component that is changing. On the other hand, if the free energy difference is obtained using a cycle whereby  $\text{Cl}^-$  is discharged,

its radius changed to that of  $\text{Br}^-$ , and the ion recharged, one would conclude that the free energy difference is predominantly due to electrostatic interactions (Boresch *et al.*, 1994; Boresch & Karplus, 1995). This latter decomposition accords more with our physical intuition based on the familiar Born-type model of ion solvation (Hirata *et al.*, 1989; Rashin & Honig, 1985; Roux *et al.*, 1990). Thus, application of equation (11), although permitting one to decompose free energy contributions, does not remove the necessity for choosing the most informative path. In the context of the TDE decomposition (or equivalently, temperature, TI) one would obtain the free energy difference by subtracting the excess free energy of the  $\text{Cl}^-$  water system from that of  $\text{Br}^-$  water. We note that this TDE decomposition also implicitly specified the cycle in which  $\text{Cl}^- + \text{water}$  is transformed into  $\text{Br}^- + \text{water}$  via an ideal gas of  $N_{\text{water}} + 1$  particles. Using TDE the total free energy change is decomposed into components resulting from changes in the water-water, water-ion electrostatic, and water-ion van der Waals interactions.

A more detailed analysis of the theoretical and computational aspects of TDE decomposition will be presented elsewhere.

---

## Acknowledgements

We thank Bhupinder Madan, Xiaoxi Chen and Martin Karplus for helpful discussion, the E. R. Johnson Foundation, NIH (GM54105) and NSF (MCB95-06900) for financial support.

## References

Bash, P., Singh, C., Brown, F., Langridge, R. & Kollman, P. (1987). Calculation of the relative change in binding free energy of a protein-inhibitor complex. *Science*, **235**, 574–576.

- Boresch, S. & Karplus, M. (1995). The meaning of component analysis: decomposition of the free energy in terms of specific interactions. *J. Mol. Biol.* **254**, 801–807.
- Boresch, S., Archontis, G. & Karplus, M. (1994). Free-energy simulations—the meaning of the individual contributions from a component analysis. *Proteins: Struct. Funct. Genet.* **20**, 25–33.
- Brady, G. P. & Sharp, K. A. (1995). Decomposition of interaction free energies in proteins and other complex systems. *J. Mol. Biol.* **254**, 77–85.
- Brooks, C. L. (1990). Molecular simulations of protein structure, dynamics and thermodynamics. In *Computer Modelling of Fluids, Polymers and Solids* (Catlow, C., Parker, S. & Allen, M., eds), pp. 289–334, Kluwer Academic Publishers, Dordrecht.
- Gao, J., Kuczera, K., Tidor, B. & Karplus, M. (1989). Hidden thermodynamics of mutant proteins: a molecular dynamics analysis. *Science*, **244**, 1069–1072.
- Hirata, F., Redfern, P. & Levy, R. (1989). Viewing the Born model of ion hydration through a microscope. *Int. J. Quant. Chem.* **15**, 179–190.
- Mark, A. E. & van Gunsteren, W. F. (1994). Decomposition of the free energy of a system in terms of specific interactions. *J. Mol. Biol.* **240**, 167–176.
- Rashin, A. A. & Honig, B. (1985). Reevaluation of the Born model of ion hydration. *J. Phys. Chem.* **89**, 5588–5593.
- Roux, B., Yu, H. A. & Karplus, M. (1990). Molecular basis for the Born model of ion solvation. *J. Phys. Chem.* **94**, 4683–4688.
- Shi, Y., Mark, A. E., Wang, C., Huang, F., Berendsen, H. & van Gunsteren, W. (1993). Can the stability of protein mutants be predicted by free energy calculations? *Protein Eng.* **6**, 289–295.
- Smith, P. E. & van Gunsteren, W. F. (1994). When are free energy components meaningful? *J. Phys. Chem.* **98**, 13735–13740.
- Yu, H. A. & Karplus, M. (1988). A thermodynamic analysis of solvation. *J. Chem. Phys.* **89**, 2366–2379.
- Zwanzig, R. W. (1954). High-temperature equation of state by a perturbation method. I Nonpolar gases. *J. Chem. Phys.* **22**, 1420–1426.

*Edited by B. Honig*

*(Received 30 May 1996; accepted 5 August 1996)*