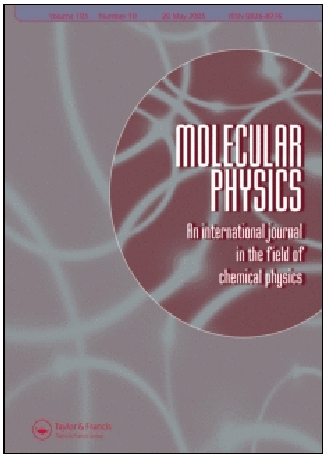
This article was downloaded by:[New York University]
[New York University]
On: 16 July 2007
Access Details: [subscription number 769426389]
Publisher: Taylor & Francis
Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Molecular Physics

An International Journal in the Field of Chemical Physics

Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713395160

Test of the overlap ratio method on the calculation of the aqueous hydration free energy difference between acetone and dimethylamine

Mihaly Mezei [a]
[a] Department of Chemistry, Hunter College of the CUNY. New York, NY. U.S.A.

Online Publication Date: 01 September 1988
To cite this Article: Mezei, Mihaly , (1988) 'Test of the overlap ratio method on the calculation of the aqueous hydration free energy difference between acetone and dimethylamine', Molecular Physics, 65:1, 219 - 223
To link to this article: DOI: 10.1080/00268978800100971
URL: http://dx.doi.org/10.1080/00268978800100971

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

© Taylor and Francis 2007

MOLECULAR PHYSICS, 1988, VOL. 65, NO. 1, 219–223

Test of the overlap ratio method on the calculation of the aqueous hydration free energy difference between acetone and dimethylamine

by MIHALY MEZEI
Department of Chemistry and Center for
Study in Gene Structure and Function,
Hunter College of the CUNY, New York, NY 10021, U.S.A.

(Received 16 March 1988; accepted 25 March 1988)

The overlap ratio method, used previously for relatively simple systems only, has been tested successfully in an aqueous system: the calculation of the hydration free energy difference between acetone and dimethylamine. The results are compared with previous calculations on the same system using the perturbation method and the finite difference thermodynamic integration method.

1. Introduction

The methods for the calculation of free energy from computer simulations, a computationally exacting task, have been periodically reviewed in the recent years [1–7]. Most methods used for complex systems fall into one of the following three classes: thermodynamic integration, the so-called perturbation method and the probability ratio method. A detailed comparison between the perturbation method and thermodynamic integration using three different paths has just been completed [8] and in a different study the probability ratio method with adaptive umbrella sampling has also been tested on the same system [9]. However, the overlap ratio method [10, 11], that should be comparable with the perturbation method (both have to sample 'adequately' the configuration space between two fixed coupling parameter values) has been neglected in recent free energy simulations on mostly aqueous systems. The purpose of this study is to examine the performance of the overlap ratio method on the calculation of the aqueous hydration free energy difference between acetone and dimethyl amine and compare it with the earlier calculations using both thermodynamic integration and the perturbation method on the same problem. (This change involves significant reorganization of the solvent around the solute during the transition between the two states with an extremum in the free energy profile.)

2. Background

The overlap ratio method calculates the free energy difference between two systems, characterized by the energy functions $E_0(X^N)$ and $E_1(X^N)$. It was introduced by Jacucci and Quirke [10], based on a formula given by Bennett [11]. This formula connects the distributions of the energy difference ΔE values and the free energy difference between the two systems:

$$\Delta A = \Delta E - kT \ln \left[P_1(\Delta E)/P_0(\Delta E)\right] \tag{1}$$

where $P_1(\Delta E)$ and $P_0(\Delta E)$ are the distributions of the energy differences in the two ensembles corresponding to the two systems between which the free energy difference is to be calculated. ΔE is an energy difference for which both $P_1(\Delta E)$ and $P_0(\Delta E)$ are nonzero. $P_1(\Delta E)$ and $P_0(\Delta E)$ can be calculated with minimum extra computational effort either during the simulation or from the simulation history and thus the free energy difference can be calculated from equation (1), as long as there is a ΔE range where the two distributions are both significantly different from zero, that is, if the two distributions overlap. This overlap criterion is easy to verify, making the application of the method rather easy. Also, the free energy difference can be calculated at several different ΔE value, giving a consistency check.

To take advantage of the several different estimates that equation (1) offers a weighted average over the overlap region was calculated as the estimate of the free energy difference:

$$\langle \Delta A \rangle = \sum_i P_1(\Delta E_i) P_0(\Delta E_i) \Delta A_i \Big/ \sum_i P_1(\Delta E_i) P_0(\Delta A_i), \tag{2}$$

where the summation is over the ΔE gridpoints where both P_1 (ΔE_i) and $P_0(\Delta E_i)$ are nonzero.

The ΔE value at which the two distributions cross coincides with ΔA since at that point the logarithmic terms vanishes in equation (1). This crossing point also is a good estimate for ΔA since that is the point where both distributions are sampled to an equal extent. The coincidence has an interesting consequence: the two distributions cannot cross more than once since two different crossing points would imply two different values for ΔA. This is a nontrivial restriction on the distributions and is probably responsible for the unimodality of the energy distributions calculated in various simulations.

3. Calculations

The simulations used the Metropolis algorithm [12], modified by incorporating the force bias procedure [13] and preferential sampling [14] for convergence acceleration. The energy calculations were based on the OPLS functions developed by Jorgensen and co-workers (using the parameters developed for small amides, the system closest to ours for which parameters are available) for the solute–water interactions [15, 16] and the TIP4P water–water potential [17, 18]. Further details of the simulation setup are described in [8].

The calculation described in this paper follows a path that transforms linearly both the coordinates of the transformed molecule and the potential coefficients, resulting in a non-linear path (the dependence of $E(\lambda)$ is not linear in λ) as described in detail in [8].

4. Results and discussion

The free energy difference between several intermediate states was computed by the overlap ratio method. The λ parameters defining the intermediate states were chosen to allow direct comparisons with the earlier calculations using the perturbation method and the finite difference thermodynamic integration method as well

Table 1. Comparison of the free energy differences computed with the overlap ratio method, perturbation method and finite difference thermodynamic integration.

λ_-(*a*):	*N*(*b*)	0·0	0·3	0·0	0·7	0·8	0·8	0·9	0·9	0·95	0·0
λ_+:		0·3	0·7	0·7	1·0	1·0	0·9	1·0	0·95	1·0	1·0
TI(*c*)	1	14·9	5·48	20·3	−11·3	−9·92	−3·31	−6·61	−2·76	−3·85	9·08
PM	1	14·4	4·77	19·1	−11·3	−9·54	−2·22	−7·11	−2·09	−5·44	8·83
PM/CK(*d*)	1					−9·33		−9·20			
OR	1	15·6	4·39	16·8	−9·33	—	−2·22	−7·20	−2·18	−5·19	10·71
OR/CK	1			20·0		−7·74		−7·36			
OR	2	15·5	4·39	18·4	−10·4	−10·4	−2·51	−7·36	−2·38	−5·77	9·50(*e*)
OR/CK	2			19·9		−9·87		−8·16			
OR	4			19·7	−11·0						8·70(*f*)

(*a*) λ_-, λ_+ are the end point coupling parameter values; (*b*) *N* is the length of the run in millions of attempted moves; (*c*) the TI result for the interval was obtained by integrating the polynomial of the gaussian quadrature; (*d*) the data labelled by /CK gives the sum of the values calculated on the subdivision of the coupling parameter interval into two parts; (*e*) the value for the whole interval at 1000 K and 2000 K were computed as the sum over the [0·0, 0·3], [0·3, 0·7] and [0.7, 1·0] intervals; (*f*) the value for the whole interval at 4000 K were computed as the sum over the [0·0, 0·7] and [0.7, 1·0] intervals; (*g*) all energies are in kJ/mol^{-1}; (*h*) $\lambda = 0$: acetone, $\lambda = 1$: dimethylamine.

as self-consistency checks within the method itself. Particular attention was paid to the $\lambda = 1$ end of the coupling parameter space where the creation of the oxygen begins since this region is the most problematic.

Table 1 summarizes the computed free energy differences at various stages of the calculation using the perturbation method, the finite difference thermodynamic integration and the overlap ratio method computed between intermediate states and between the two systems. The overlap ratio and perturbation method results have been repeated on several subdivision of the coupling parameter intervals—the sums of the subinterval results are labelled with CK. Good agreement was found both with the previously calculated free energy differences over the various intervals considered as well as over the whole interval and with the values calculated on the subintervals with the overlap ratio method as a self test.

The overlap ratio method can also be compared with the perturbation method with respect to the coupling parameter interval width over which they are able to calculate the free energy reliably and the length of runs required to achieve 'converged' results. The overlap ratio method works well on larger coupling parameter intervals—it worked even on the [0·0, 0·7] interval—but required proportionally longer runs to converge. Thus the use of the overlap ratio method can provide the convenience of having to set up (and equilibrate) fewer runs but will incur about the same computational cost as the perturbation method. The comparison over the 'critical' [0·9, 1·0] coupling parameter interval where the subinterval test was performed on both methods show that the overlap ratio method results are likely to be more reliable than the perturbation method results since the self test for the perturbation method is off by 2·1 kJ mol^{-1} (although the values appeared 'converged' [8]). This finding is an indication of the difficulty of calculating reliable exponential averages—a necessary ingredient of the perturbation method.

Table 2 collects information on the overlap of the distributions: the maxima of the two distribution computed, $P_{\max 0}$ and $P_{\max 1}$, to be compared with the value at

Table 2. Characterization of the overlapping distributions.

$\lambda_-(a)$	$N(b)$	0·0	0·3	0·0	0·7	0·8	0·8	0·9	0·9	0·95
λ_+:		0·3	0·7	0·7	1·0	1·0	0·9	1·0	0·95	1·0
$P_{\max 0}(c)$	1	0·06	0·06	0·02	0·09		0·22	0·17	0·24	0·22
$P_{\max 1}$		0·07	0·06	0·02	0·01		0·16	0·41	0·22	0·08
$P(X)(d)$		0·02	0·03	0·02	0·01		0·16	0·41	0·21	0·06
$\Delta^w E(e)$		25·9	30·9	45·2	32·6		14·2	20·1	13·4	24·3
$\Delta^e A(f)$		0·63	1·38	0·75	0·38		1·63	0·75	0·46	0·46
$\Delta^x A(g)$		0·29	0·29	0·46	0·75		0·25	0·88	0·21	0·21
$P_{\max 0}$	2	0·06	0·06	0·02	0·08	0·10	0·22	0·15	0·24	0·21
$P_{\max 1}$		0·06	0·06	0·03	0·01	0·01	0·15	0·37	0·21	0·10
$P(X)$		0·02	0·03	0·01	0·01	0·01	0·15	0·31	0·21	0·04
$\Delta^w E$		28·5	31·1	45·2	41·8	29·3	15·1	24·3	14·2	24·3
$\Delta^e A$		0·17	0·50	0·63	0·38	0·29	2·26	0·63	0·38	0·84
$\Delta^x A$		0·04	0·21	1·28	0·64	0·16	0·25	0·38	0·13	0·79
$P_{\max 0}$	4			0·03	0·08					
$P_{\max 1}$				0·03	0·01					
$P(X)$				0·01	0·01					
$\Delta^w E$				49·4	41·8					
$\Delta^e A$				0·96	0·42					
$\Delta^x A$				1·20	0·84					

(*a*) λ_-, λ_+ are the end point coupling parameter values; (*b*) N is the length of the run in millions of attempted moves; (*c*) $P_{\max 0}$, $P_{\max 1}$ are the maxima of distributions $P_0(\Delta E)$ and $P_1(\Delta E)$, respectively; (*d*) $P(X)$ is the value of the probability distributions $P_0(\Delta E)$ and $P_1(\Delta E)$ where the two distributions cross; (*e*) $\Delta^w E$ is the energy range over which the two distributions have both nonzero value; (*f*) $\Delta^e A$ is the range of free energy estimates over a 4·0 kJ mol^{-1} wide interval around the crossing of the two distributions; (*g*) $\Delta^x A$ is the difference between the free energy value calculated by equation (2) and the value of ΔE at which the two distributions cross; (*h*) $\lambda = 0$: acetone, $\lambda = 1$: dimethylamine

the crossing, $P(X)$, the width of the overlap, $\Delta^w E$, and the fluctuation of the free energy estimate in a 4 kJ mol^{-1} interval around the crossing, $\Delta^e A$. It is surprising that even distributions belonging to largely different coupling parameter values overlap significantly although in these cases the distributions turn out to be rather flat. The performance of the consistency check, however, is disappointing. In some instances (like for the [0·0, 0·7] interval) it predicted significantly smaller error than the actual error, estimated either from comparison with results using the perturbation method and the thermodynamic integration or with overlap ratio method results performed on two subintervals. This finding underlines the difficulty of calculating reliable error estimates from a single free energy calculation. The deviation of the calculated free energy from the crossing point of the two distributions, $\Delta^x A$, is also shown in table 2. The deviations are consistently small, as expected.

In summary, the overlap ratio method was found to be capable of calculating solvation free energy differences in aqueous systems. In comparison with the perturbation method, the overlap ratio method was shown to be able to work over wider coupling parameter intervals than the perturbation method, although the run length required to achieve comparable precision also increased. At least in one case it gave more reliable results than the perturbation method. The advantage of the overlap ratio method over the perturbation method is the simplification due to the fewer calculations required and a somewhat increased confidence in the result rather than in computer time.

This work was supported under an RCMI grant #SRC5G12RR0307 from NIH to Hunter College. Computing resources were provided by the City University of New York, University Computing Center. Illuminating discussions with Drs. D. Tildesley and N. Quirke at the CECAM workshop on free energy calculations at Orsay are gratefully acknowledged.

References

[1] Barker, J. A., and Henderson, D., 1976, *Rev. mod. Phys.*, **48,** 587.

[2] Quirke, N., 1980, *Proceedings of the NATO Summer School on Superionic Conductors*, Odense, Denmark (Plenum Press).

[3] Shing, K. S., and Gubbins, K. E., 1983, *Advances in Chemistry Series*, Vol. 204, *Molecular Based Study of Fluids*, edited by J. M. Haile and G. A. Mansoori (American Chemical Society).

[4] Levesque, J. J., Weiss, J. J., and Hansen, J. P., 1984, *Monte Carlo Methods in Statistical Physics*, Vol. II, edited by K. Binder (Springer-Verlag).

[5] Pohorille, A., and Pratt, L. R., 1986, *Meth. Enzymol. Biomembranes*, **127,** 64.

[6] Mezei, M., and Beveridge, D. L., 1986, *Ann. N.Y. Acad. Sci.*, **482,** 1.

[7] Frenkel, D., 1986, *Molecular-Dynamics Simulation on Statistical Mechanical Systems* (Soc. Italiana di Fisica).

[8] Mezei, M., 1987, *J. chem. Phys.*, **86,** 7084.

[9] Mezel, M. (in preparation).

[10] Jacucci, G., and Quirke, N., 1980, *Molec. Phys.*, **40,** 1005.

[11] Bennet, C. H., 1976, *J. Comput. Phys.*, **22,** 245.

[12] Metropolis, N. A., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E., 1953, *J. chem. Phys.*, **21,** 1087.

[13] Rao, M., Pangali, C. S., and Berne, B. J., 1979, *Molec. Phys.*, **37,** 1779.

[14] Owicki, J. C., and Scheraga, H. A., 1979, *Chem. Phys. Lett.*, **47,** 600. Owicki, J. C., 1987, *Computer Modeling of Matter*, edited by P. G. Lykos (American Chemical Society).

[15] Jorgensen, W. L., and Swenson, C. J., 1985, *J. Am. chem. Soc.*, **107,** 569.

[16] Jorgensen, W. L., and Swenson, C. J., 1985, *J. Am. chem. Soc.*, **107,** 1489.

[17] Jorgensen, W. L., Chandrashekar, J., Madura, J. D., Impey, R., and Klein, M. L., 1983, *J. chem. Phys.*, **79,** 926.

[18] Jorgensen, W. L., and Madura, J. D., 1985, *Molec. Phys.*, **56,** 1381.