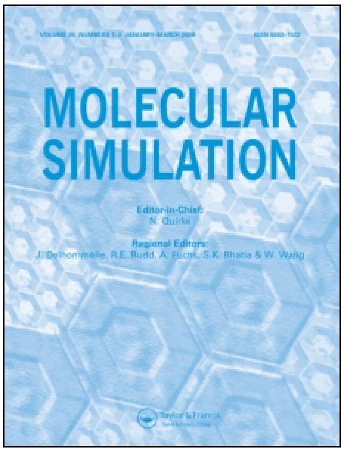
This article was downloaded by:
On: *14 January 2011*
Access details: *Access Details: Free Access*
Publisher *Taylor & Francis*
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Molecular Simulation

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

Statistical Errors in Constrained Molecular Dynamics Calculations of the Mean Force Potential

E. Guardia[a]; R. Rey[ab]; J. A. Padro[b]
[a] Departament de Fisica i Enginyeria Nuclear, Universitat Politècnica de Catalunya, Barcelona, Spain [b] Departament de Fisica Fonamental, Universitat de Barcelona, Barcelona, Spain

To cite this Article Guardia, E. , Rey, R. and Padro, J. A.(1992) 'Statistical Errors in Constrained Molecular Dynamics Calculations of the Mean Force Potential', Molecular Simulation, 9: 3, 201 — 211
To link to this Article: DOI: 10.1080/08927029208047427
URL: http://dx.doi.org/10.1080/08927029208047427

PLEASE SCROLL DOWN FOR ARTICLE

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

This article may be 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.

Molecular Simulation, 1992, Vol. 9, pp. 201–211
Reprints available directly from the publisher
Photocopying permitted by license only

© 1992 Gordon and Breach Science Publishers S.A.
Printed in Great Britain

STATISTICAL ERRORS IN CONSTRAINED MOLECULAR DYNAMICS CALCULATIONS OF THE MEAN FORCE POTENTIAL

E. GUARDIA[a], R. REY[a,b] and J.A. PADRO[b]

[a]*Departament de Física i Enginyeria Nuclear, Universitat Politècnica de Catalunya, Pau Gargallo, 5. 08028 Barcelona, Spain*
[b]*Departament de Física Fonamental, Universitat de Barcelona, Diagonal, 647. 08028 Barcelona, Spain*

(*Received January 1992, accepted February 1992*)

A procedure to estimate the statistical errors in constrained molecular dynamics calculations of the solute–solute mean force potential is presented. The concept of statistical inefficiency is used to determine the decorrelation time for independent samples of the ion–solvent mean force. A compact expression is given for the computation of the statistical error in the mean force potential. The proposed methodology allows for a systematic analysis of the convergence of the simulation data. Results of its application to Na^+-Cl^-, Cl^--Cl^- and Na^+-Na^+ ion pairs in water are reported.

KEY WORDS: Constrained molecular dynamics, mean force, mean force potential, statistical inefficiency, statistical error

Downloaded At: 19:40 14 January 2011

INTRODUCTION

The solute–solute mean force potential ($W(r)$) plays a central role in theoretical studies of both equilibrium and nonequilibrium properties of liquid solutions [1,2]. The calculation of $W(r)$ by computer simulation techniques is of increasing interest. Ordinary simulations are not well suited for calculating $W(r)$ because a homogeneous sampling at different solute–solute distances cannot be achieved. To overcome this problem, several specialized techniques have been developed. Umbrella sampling [3,4], thermodynamic perturbation [5,6] and thermodynamic integration [7] methods are based on statistical thermodynamic techniques. A completely different approach is provided by the constrained molecular dynamics (MD) technique [8]. This method is based on the calculation of the mean force acting on the solute particles due to the solvent molecules, while the solute particles are kept at fixed separations. $W(r)$ is afterwards determined by numerical integration of the solute–solute mean force ($F(r)$).

Constrained MD constitutes a conceptually simple route to $W(r)$. Its implementation in an ordinary MD program is straightforward. However, there is a widely extended opinion that it is numerically less reliable [9]. Contrary to this conventional belief, recent calculations have confirmed the reliability of the method and its computational feasibility [8, 10, 11].

In the present work, we develop a procedure that allows for a systematic analysis of the convergence of $F(r)$ and $W(r)$ during a constrained MD calculation. This procedure is applied to estimate the statistical errors in previously reported calculations of $W(r)$ for Na^+–Cl^-, Cl^-–Cl^- and Na^+–Na^+ ion pairs in water.

METHODOLOGY

Constrained MD

We consider a system made up of two solute particles (A, B) and N solvent molecules. MD simulations are performed keeping the AB separation fixed. To this end a constraint between the positions of the solute particles is added

$$\xi(\mathbf{r}_A, \mathbf{r}_B) = (\mathbf{r}_A - \mathbf{r}_B)^2 - r^2 = 0 \tag{1}$$

The force $(\Delta F(t;r))$ exerted by the solvent molecules along the solute–solute internuclear axis is computed during the constrained simulations. $\Delta F(t;r)$ can be expressed as

$$\Delta F(t;r) = \mu\left[\frac{\mathbf{F}_{AS}(t;r)}{m_A} - \frac{\mathbf{F}_{BS}(t;r)}{m_B}\right]\cdot\hat{\mathbf{r}} \tag{2}$$

where $\mathbf{F}_{AS}(t;r)$ and $\mathbf{F}_{BS}(t;r)$ are the total forces exerted by the solvent molecules on each of the two solute particles, m_A and m_B are their respective masses, $\mu = m_A m_B/(m_A + m_B)$ is the reduced mass and $\hat{\mathbf{r}}$ is a unit vector along the AB direction

$$\hat{\mathbf{r}} = \frac{\mathbf{r}_A - \mathbf{r}_B}{|\mathbf{r}_A - \mathbf{r}_B|} \tag{3}$$

The solvent contribution to $F(r)$ is calculated as the average of $\Delta F(t;r)$ over the constrained MD configurations. From Equation (2) and taking into account that $\langle \mathbf{F}_{AS}(t;r)\rangle = -\langle \mathbf{F}_{BS}(t;r)\rangle$ it follows that

$$\Delta F(r) \equiv \langle \Delta F(t;r)\rangle = \tfrac{1}{2}\langle(\mathbf{F}_{AS}(t;r) - \mathbf{F}_{BS}(t;r))\cdot\hat{\mathbf{r}}\rangle \tag{4}$$

This expression coincides with the one proposed in reference [8] for computing the mean solvent force on the solute particles.

If $F_d(r)$ is the direct solute–solute force, the total mean force between the solute particles may be written as

$$F(r) = F_d(r) + \Delta F(r) \tag{5}$$

and $W(r)$ can be calculated by integrating

$$W(r) = W(r_0) - \int_{r_0}^{r} F(r)\,dr \tag{6}$$

where $W(r_0)$ should be suitably chosen to obtain reliable $W(r)$ values at long distances.

Error Estimates

We divide the sequence of MD steps corresponding to a given solute–solute separation into n_b blocks of time length τ_b. We denote by $\langle\Delta F(t;r)\rangle_b$ the mean value of $\Delta F(t;r)$ calculated for each block. We define the variance of $\langle\Delta F(t;r)\rangle_b$ as

$$\sigma_b^2(\Delta F(r)) = \frac{1}{n_b}\sum_{b=1}^{n_b}(\langle\Delta F(t;r)\rangle_b - \Delta F(r))^2 \tag{7}$$

Downloaded At: 19:40 14 January 2011

Following reference [12], we define the statistical inefficiency as

$$s(r) = \lim_{\tau_b \to \infty} \frac{\tau_b \sigma_b^2(\Delta F(r))}{\sigma^2(\Delta F(r))} \tag{8}$$

where $\sigma^2(\Delta F(r))$ is the variance of $\Delta F(t; r)$ resulting from the whole simulation run. $s(r)$ gives the interval of time necessary to obtain statistically independent samples of the solvent mean force. The statistical error in the calculation of $\Delta F(r)$ is then given by

$$\varepsilon_{\Delta F}(r) = \left[\frac{s(r)}{\tau_{\text{run}}}\right]^{1/2} \sigma(\Delta F(r)) \tag{9}$$

where τ_{run} is the total length of the MD run. Equation (9) allows us to obtain the statistical error in $\Delta F(r)$ which is also the error in the solute–solute mean force for a given r value (see Equation (5)). Let us now turn to the calculation of the statistical error in the mean force potential.

Let us assume that we use the composite trapezoide rule [13] to carry out the numerical integration in Equation (6). Then

$$W(r) = W(r_0) - \sum_{i=0}^{n-1} \Delta r \frac{F(r_i) + F(r_{i+1})}{2} \tag{10}$$

where the interval between r and r_0 has been divided into n subintervals of width Δr and the discretized interionic separation is given by $r_i = r_0 - i\Delta r$. Equation (10) can be expressed as

$$W(r) = W(r_0) - \sum_{i=0}^{n} \zeta_i(r_i) \tag{11}$$

with

$$\begin{aligned} \zeta_0(r_0) &= (\Delta r/2) F(r_0) \\ \zeta_i(r_i) &= \Delta r F(r_i) \qquad 1 \leqslant i \leqslant n-1 \\ \zeta_n(r_n) &= (\Delta r/2) F(r) \end{aligned} \tag{12}$$

If the $F(r_i)$ forces are calculated from runs with independent initial configurations, $\zeta_i(r_i)$ are statistically independent variables and the statistical error in the mean force potential is given by [14]

$$\varepsilon_W(r) = \left[\sum_{i=0}^{n} \varepsilon_{\zeta_i}^2\right]^{1/2} \tag{13}$$

If we further assume that the $\varepsilon_{\Delta F}(r)$ errors corresponding to different interionic separations may be approximated by their mean value, i.e., $\varepsilon_{\Delta F}(r_i) \simeq \bar{\varepsilon}_{\Delta F}$, we get the compact expression

$$\varepsilon_W(r) \simeq \Delta r [n - 1/2]^{1/2} \, \bar{\varepsilon}_{\Delta F} \tag{14}$$

which indicates that as r diminishes there is a systematic increase of $\varepsilon_W(r)$ proportional to $n^{1/2}$. Equation (14) constitutes our algorithm for computing the statistical error in $W(r)$.

Downloaded At: 19:40 14 January 2011

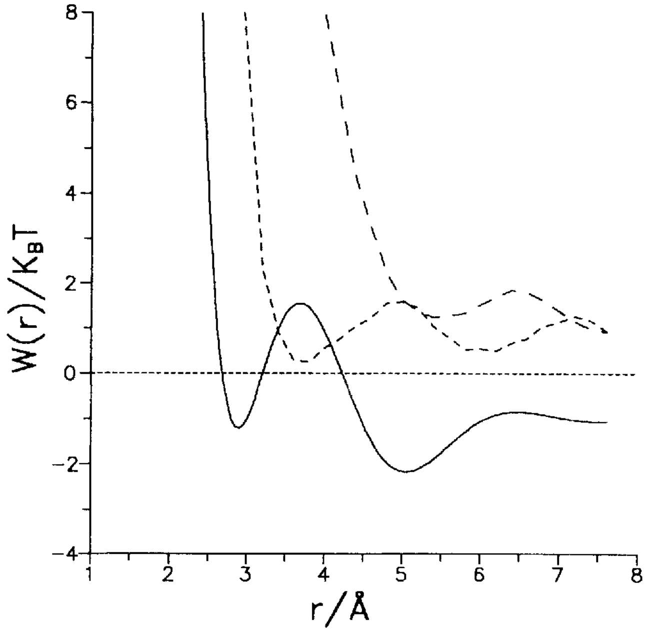


Figure 1 Ion–ion mean force potentials as a function of the interionic separation. ———, Na^+–Cl^-; — — —, Cl^-–Cl^-; - - - -, Na^+–Na^+.

APPLICATION TO IONS IN WATER

Description of the Systems

The above procedure has been applied to estimate the statistical errors in previously reported constrained MD calculations of $W(r)$ for Na^+–Cl^-, Cl^-–Cl^- and Na^+–Na^+ ion pairs in water at normal room temperature conditions [10, 11]. The calculations were performed on systems made up of one ion pair and 125 water molecules. We assumed flexible water molecules according to the interaction model proposed by Toukan and Rahman [15]. The ion–water potential functions were modelled by a Coulomb plus a Lennard-Jones term. The Huggins–Mayer potential form was used for the ion–ion interactions. The potential functions and parameters used for ion–ion and ion–water interactions were the same as in reference [16]. The Ewald summation method [17] was used for computing the long-range forces. For each ion pair a series of simulations were performed to cover an interionic separation ranging from r_0 = 7.6 Å to 2.4 Å for Na^+–Cl^-, and from r_0 = 7.6 Å to 3.0 Å for Cl^-–Cl^- and Na^+–Na^+, with Δr = 0.2 Å. The constraint on the ion–ion separation was imposed by the SHAKE technique [18]. Figure 1 displays the obtained $W(r)$ potentials. More technical details concerning their calculation together with a physical discussion of the results are given in the original papers [10, 11].

Results and Discussion

For each ion pair we have evaluated the statistical errors $\varepsilon_{\Delta F}(r)$ for three different

Downloaded At: 19:40 14 January 2011

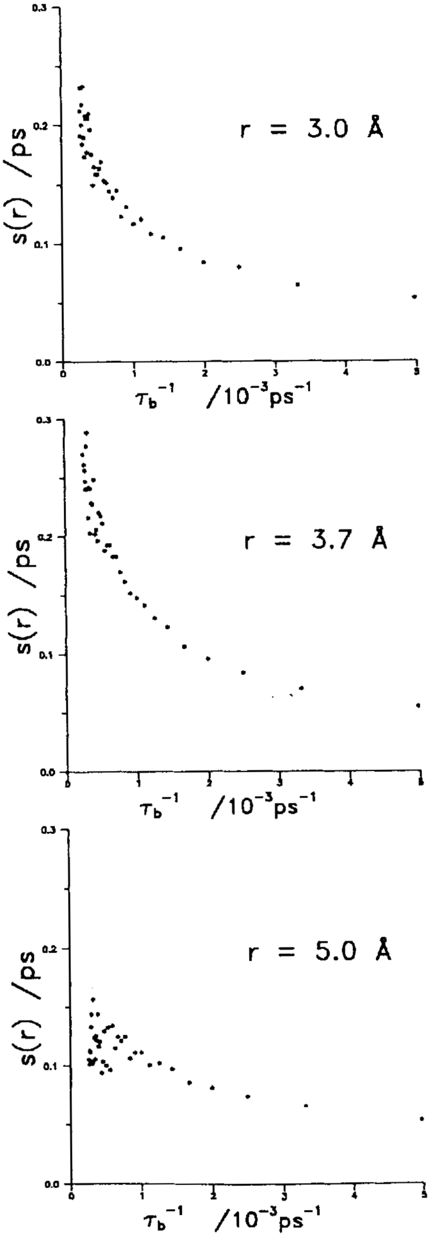


Figure 2 Statistical inefficiency of $\Delta F(r)$ for Na^+–Cl^- at different interionic separations.

Downloaded At: 19:40 14 January 2011

Table 1 Statistical inefficiency and correlation time of the ion-solvent mean force.

Ion pair	*Interionic distance* r(Å)	$s(r)$ *(ps)*	$\tau_{\Delta F}$*(ps)*
Na^+-Cl^-	3.0	0.20	0.10
	3.7	0.25	0.10
	5.0	0.12	0.04
Cl^--Cl^-	3.6	0.14	0.07
	5.4	0.13	0.06
	6.4	0.15	0.06
Na^+-Na^+	3.8	0.18	0.19
	5.0	0.22	0.13
	6.0	0.10	0.06

Downloaded At: 19:40 14 January 2011

interionic separations. The values of r considered are close to the positions of the minima and maxima of the corresponding $W(r)$ (see Figure 1). In the case of Cl^--Cl^- we also considered $r = 3.6$ Å which corresponds to a close contact ion pair configuration with an extremely low probability. In each case, $s(r)$ was obtained from a constrained run of 150 ps by applying Equation (8). Figure 2 shows the plots of $\tau_b\sigma_b^2(\Delta F(r))/\sigma(\Delta F(r))$ against $1/\tau_b$ resulting for the Na^+-Cl^- ion pair. For all the considered systems, an asymptotic behaviour is observed which indicates that $s(r)$ is a well defined quantity. The estimated values of $s(r)$ range between 0.1 ps and 0.25 ps (see Table 1). They do not show any strong dependence on the ion pair neither on the interionic separation. The lowest values are obtained for the Cl^--Cl^- ion pair, which may be related to the lower residence time of water molecules around the Cl^- ion [11]. The highest values correspond to the maximum of the Na^+-Cl^- and Na^+-Na^+ mean force potential. This can be interpreted in terms of the relative unstability of these configurations.

In early applications of the constrained MD method [8, 10,11], the decorrelation time for independent samples of the ion-solvent mean force was determined from the force-force autocorrelation function

$$C_{\Delta F}(t;r) = \frac{\langle \delta\Delta F(t;r)\,\delta\Delta F(0;r)\rangle}{\langle[\delta\Delta F(0;r)]^2\rangle} \tag{15}$$

where $\delta\Delta F(t;r) = \Delta F(t;r) - \Delta F(r)$. Figure 3 displays the $C_{\Delta F}(t;r)$ functions resulting from the data generated during the constrained MD runs for the Na^+-Cl^- ion pair.

There is not a unique prescription to determine the decorrelation time from $C_{\Delta F}(t;r)$. We have assumed that it is given by the correlation time

$$\tau_{\Delta F} = \int_0^\infty C_{\Delta F}(t;r)\,dt \tag{16}$$

The resulting $\tau_{\Delta F}$'s are close to the position of the minimum of the corresponding $C_{\Delta F}(t;r)$ functions and underestimate the $s(r)$ values by a factor of approximately two (see Table 1).

In order to monitor the convergence of the constrained MD technique, we have computed the cumulative average of $\Delta F(t;r)$ and its statistical error as a function of the length of the constrained simulations (see Figure 4). Accordingly with the reported values of $s(r)$, the convergence is slightly faster in the case of the Cl^--Cl^- ion pair. In general, there is an important diminution of the deviations with respect to the mean

Downloaded At: 19:40 14 January 2011

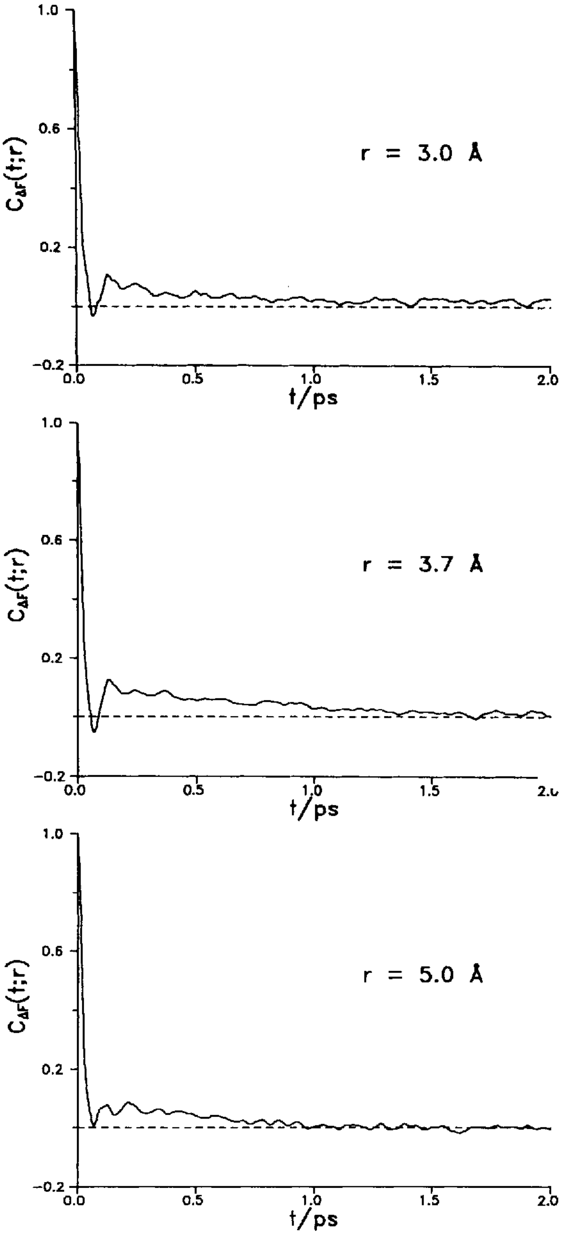


Figure 3 Ion-solvent force autocorrelation function $C_{\Delta F}(t; r)$ for Na^+–Cl^- at different interionic separations.

Downloaded At: 19:40 14 January 2011

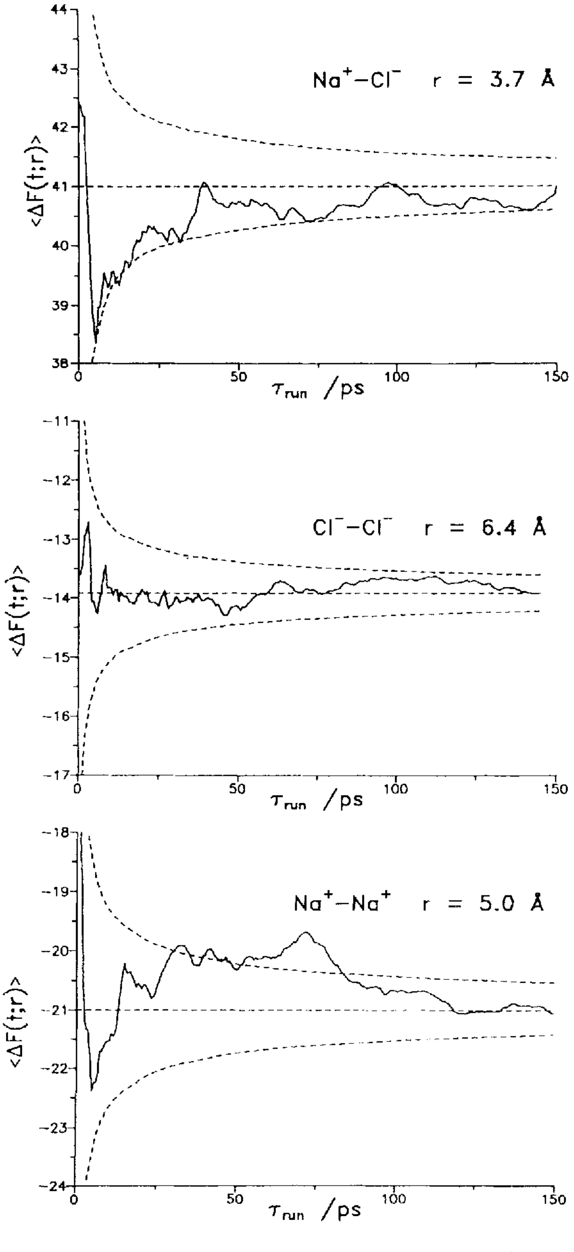


Figure 4 Cumulative average and statistical error of $\Delta F(t;r)$ (in units of $k_BT/\text{Å}$) as a function of the length of the simulation. Full line: $\langle \Delta F(t\cdot r)\rangle$, $\Delta F(r)+\varepsilon_{\Delta F}(r)$, $\Delta F(r)-\varepsilon_{\Delta F}(r)$.

Downloaded At: 19:40 14 January 2011

Table 2 Error estimates in the ion-solvent mean force in units of $k_B T/\text{Å}$.

Ion pair	*Interionic distance* r *(Å)*	$\sigma(\Delta F(r))$			$\varepsilon_{\Delta F}(r)$		
		τ_{run} = *30 ps*	*80 ps*	*150 ps*	τ_{run} = *30 ps*	*80 ps*	*150 ps*
Na^+-Cl^-	3.0	8.8	8.8	8.8	0.72	0.44	0.32
	3.7	10.3	10.3	10.6	0.97	0.58	0.43
	5.0	11.9	11.4	11.5	0.73	0.44	0.32
Cl^--Cl^-	3.6	9.0	9.1	9.1	0.62	0.38	0.28
	5.4	10.0	9.9	9.8	0.64	0.40	0.29
	6.4	9.7	10.0	9.9	0.70	0.43	0.31
Na^+-Na^+	3.8	10.9	10.5	10.7	0.81	0.50	0.36
	5.0	11.4	11.5	11.5	0.98	0.60	0.44
	6.0	11.7	11.4	11.3	0.65	0.40	0.29

value $\Delta F(r)$ as the simulation length exceeds ~ 50 ps. A further prolongation of the runs does not result in a significant improvement of the statistics. Moreover, the variance $\sigma(\Delta F(r))$ remains essentially constant for $\tau_{\text{run}} \geqslant 10$ ps. Then, the statistical error $\varepsilon_{\Delta F}(r)$ obtained from Equation (9) is proportional to $1/\tau_{\text{run}}^{1/2}$ (see Table 2). For a given ion pair, the changes of $\varepsilon_{\Delta F}(r)$ with the interionic distance are very small. This justifies the approximation used to derive Equation (14).

We have calculated $\varepsilon_W(r)$ by using equation (14). For each ion pair, $\bar{\varepsilon}_{\Delta F}$ has been taken as the arithmetic mean of the $\varepsilon_{\Delta F}(r)$ values reported in Table 2 for the different interionic separations. Figure 5 displays the obtained $\varepsilon_w(r)$ errors for Na^+-Cl^-, at different values of the simulation length. It is remarkable that the positions of the minima and maxima of $W(r)$ can be obtained even from short runs of 30 ps. In accordance with the discussion of the convergence of $\Delta F(r)$, the statistical errors in the mean force potential are only slightly reduced when the constrained MD runs are extended from 80 ps to 150 ps. The $\varepsilon_W(r)$ errors for Cl^--Cl^- are smaller than for the other ion pairs. For a given ion pair, the bigger errors correspond to the shorter distances, but $\varepsilon_W(r)$ remain bounded for all the range of interionic separations.

SUMMARY

A method to evaluate the statistical errors in constrained MD calculations of the solute–solute mean force potential has been proposed. In our opinion, the possibility of performing such a systematic analysis of the statistical errors constitutes a significant advantage of the constrained MD technique in front of other methods for the calculation of $W(r)$ by computer simulation.

The proposed procedure has been applied to the Na^+-Cl^-, Cl^--Cl^- and Na^+-Na^+ ion pairs in water. For all the analyzed cases, it has been obtained a decorrelation time which lies between 0.1 ps–0.25 ps and it is approximately two times the ion-solvent force correlation time. We have not detected any significant influence of the ion-pair neither of the ionic distance on the convergence of the results. As a general rule, constrained runs of a minimum length of 50 ps must be performed in order to obtain reliable mean force potentials. Although they have been obtained for specific systems, these conclusions can provide useful information for the calculation of $W(r)$ for ionic species in water by constrained MD techniques.

Downloaded At: 19:40 14 January 2011

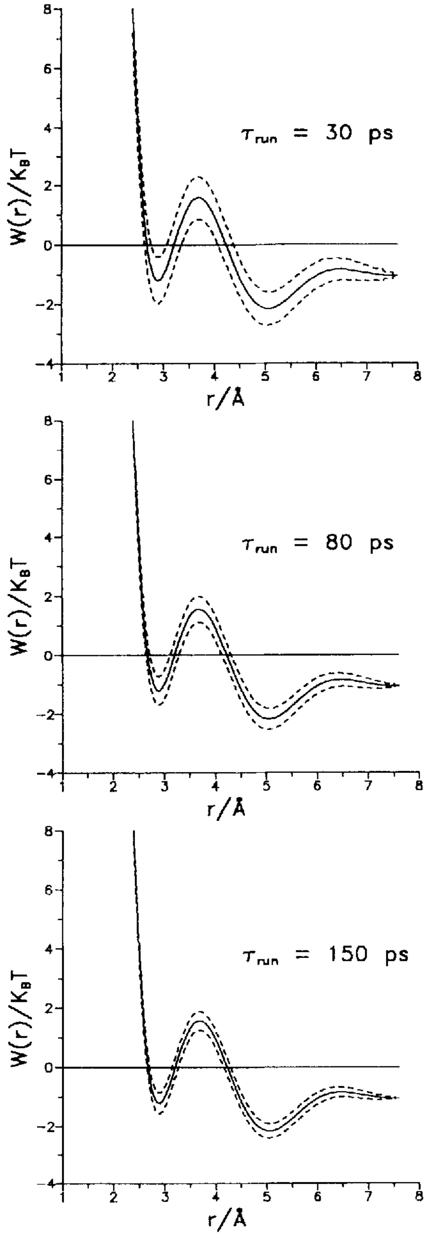


Figure 5 Statistical error in $W(r)$ for Na^+–Cl^- as a function of the length of the simulation. Full line: $W(r)$. Dashed lines: $W(r) + \varepsilon_w(r)$ and $W(r) - \varepsilon_w(r)$.

Acknowledgements

The financial support by DGICYT, Project No. PB90-0613-C03 is acknowledged.

References

[1] H.L. Friedman, *A Course in Statistical Mechanics* (Prentice Hall, Englewoods Cliffs, 1985), Ch. 10.
[2] J.T. Hynes, "The theory of reactions in solution", in *The Theory of Chemical Reaction Dynamics*, Vol. 4, M. Baer (ed.), (CRC Press, Boca Raton, 1985).
[3] J.P. Valleau and G.M. Torrie, in *Statistical Mechanics, Part A: Equilibrium Techniques*, B.J. Berne (ed.) (Plenum Press, New York, 1977), Chs 4 and 5.
[4] C. Pangali, M. Rao and B.J. Berne, "A Monte Carlo simulation of the hydrophobic interaction", *J. Chem. Phys.*, **71**, 2975 (1979).
[5] R.W. Zwanzig, "High temperature equation of state by a perturbation method. I. Non polar gases", *J. Chem. Phys.*, **22**, 245 (1954).
[6] D.J. Tobias and C.L. Brooks, "Calculation of free energy surfaces using the methods of thermodynamic perturbation theory", *Chem. Phys. Lett.*, **142**, 472 (1987).
[7] J. Van Eerden, W.J. Briels, S. Harkema and D. Feil, "Potential of mean force by thermodynamic integration: molecular dynamics simulation of decomplexation", *Chem. Phys. Lett.*, **164**, 370 (1989).
[8] G. Ciccotti, M. Ferrario, J.T. Hynes and R. Kapral, "Constrained Molecular dynamics and the mean force potential for an ion pair in a polar solvent", *Chem. Phys.*, **129**, 241 (1989).
[9] D.L. Beveridge and F.M. DiCapua, "Free energy via molecular simulation: A primer", in *Computer Simulation of Biomolecular Systems*, W.F. Van Gunsteren and P.K. Weiner (eds.) (Escom, Leiden, 1989).
[10] E. Guàrdia, R. Rey and J.A. Padró, "Potential of mean force by constrained molecular dynamics: a sodium chloride ion-pair in water", *Chem. Phys.*, **155**, 187 (1991).
[11] E. Guàrdia, R. Rey and J.A. Padró, "Na^+-Na^+ and Cl^--Cl^- ion pairs in water: mean force potentials by constrained molecular dynamics", *J. Chem. Phys.*, **95**, 2823 (1991).
[12] M.P. Allen and D.J. Tildesley, *Computer Simulation of Liquids* (Clarendon Press, Oxford, 1987), Ch. 6.
[13] G.E. Forsythe, M.A. Malcolm and C.B. Moler, *Computer Methods for Mathematical Computations* (Prentice-Hall, Englewoods Cliffs, 1977), Ch. 5.
[14] A. Papoulis, *Probability, Random Variables and Stochastic Processes* (McGraw-Hill, 1969).
[15] K. Toukan and A. Rahman, "Molecular dynamics study of atomic motions in water", *Phys. Rev.*, **B31**, 2643 (1985).
[16] B.M. Pettit and P.J. Rossky, "Alkali Halides in water: Ion–solvent and ion–ion potentials of mean force at infinite dilution", *J. Chem. Phys.*, **84**, 5836 (1986).
[17] S.W. De Leeuw, J.W. Perram and E.R. Smith, "Simulation of electrostatic systems in periodic boundary conditions. I. Lattice sums and dielectric constant", *Proc. R. Soc. Lond.*, **A373**, 27 (1980).
[18] J.P. Ryckaert, "The method of constraints in molecular dynamics. General aspects and application to chain molecules", in *Molecular Dynamics Simulation of Statistical Mechanical Systems*, G. Ciccotti and W.G. Hoover (eds.) (North-Holland, Amsterdam, 1986).

Downloaded At: 19:40 14 January 2011