REGULAR ARTICLE

Additional peaks in the cluster size distribution of amphiphile + water systems: a clue for shape/phase transition or statistical uncertainty

Reza Behjatmanesh-Ardakani

Abstract In some works on the lattice Monte Carlo simulation of amphiphilic systems additional peaks in the cluster size distribution has been interpreted as a clue for the phase or shape transition of micellar aggregates. On the other hand, some other works showed that the additional peaks are a result of finite size of the lattice box. In this paper using calculating energy-auto-correlation function and statistical error in correlated data, it is shown that how these apparently contradictory results are the same. To do this, we have simulated a pure system containing amphiphile and water molecules. A simple model of potential containing the main feature for these systems (the hydrophobicity of surfactant molecules) that cause the aggregates to be formed is considered to avoid any synthetic results due to additional non-real parameters. To relax the initial configuration faster, configurational bias Monte Carlo move is used in addition to reptation move. Periodic boundary condition and self-avoiding walks are used as former published works in this field. It is shown that the additional peaks is a result of the statistical errors for averaged cluster size distribution and can not be interpreted as a clue for shape or phase transition.

1 Introduction

Nowadays simulation techniques are extensively used to investigate the molecular behavior of different systems [1–3]. Among different simulation techniques introduced in the literature, only coarse-grained bead type model has been used for micelle formation and related phenomena. Because of

R. Behjatmanesh-Ardakani (✉)
Department of Chemistry, School of Science,
Payame-Noor University of Ardakan,
Ardakan, Yazd, Iran
e-mail: reza_b_m_a@yahoo.com

the long time scales involved in self-assembling systems, the computational requirements for equilibrating more realistic models are too demanding.

Using these simple models, it has been discussed that how micelles are formed and what are the effective parameters on them [4,5]. Furthermore, simulation can be used to study the phase behavior of surfactant systems [6–8] and to test the results of different theories in this field [9]. Although one can use different simulation methods such as lattice and off-lattice Monte Carlo [10–12], Molecular Dynamics [13], Dissipative Particle Dynamics [14] and Brownian Dynamics [15], but one point is common: Are the results reliable?

Bernardes and his coworkers have discussed the reliability condition for a lattice Monte Carlo simulation for a pure micelles system [16]. They have shown that four factors must be checked for any simulation of lattice Monte Carlo containing finite size effect, meta-stability, control of relaxation time, and correlations between snapshots considered for ensemble averaging. They have checked finite size effect through investigating the behavior of the monomer concentration versus total concentration at different lattice sizes, meta-stability by monitoring the convergence of energy for two different initial configurations (random or high energy configuration and ordered or low energy configuration), relaxation time by relaxing the total energy of the lattice and correlation by an approximate manner without calculating any auto-correlation function. On the other hand, Nelson et al., for the first time, introduced the finite size effect on the distribution of cluster sizes in the lattice model. They investigated the finite size effect on the behavior of cluster size distribution for the pure system of H_2T_2 in a face-centered cubic lattice. They have shown that at small lattices the cluster size distribution can be biased by the presence of additional peaks. These peaks disappeared at larger lattices [17]. This is in contradictory with some other simulation works of

oil–water–surfactant [18–20] or pure ones [21] that interpret the additional peaks as a shape or phase transition.

In the present work, we have conducted much more new runs with the same simulation details as our previous published work in this journal [21]. Taking two head beads (H_2) and four connected tail beads (T_4) to the heads, we have checked the correlation between snapshots taken for ensemble averaging by the energy auto-correlation function. In addition, statistical uncertainty of each data on the cluster size distribution has been calculated using standard methods has been defined in the literature.

This paper is organized as follows. We first describe the details of the model used in this research. Then we present the results of the new runs and compare them to our previous results published in this journal.

2 Model

The details of the model and the computation method are similar to our previous paper with some changes for the calculations of energy auto-correlation function. The data analysis is done after each simulation run. The simulations have been conducted on simple cubic lattice ($50 \times 50 \times 50$) with periodic boundary conditions in all of the directions and excluded volume to mimic the box of the simulation to the bulk of the solution. A chain containing four tail beads connecting to the two head sites (similar to the water beads) is used as a non-ionic asymmetric amphiphile (H_2T_4). A water molecule occupies only one site. In a random initial configuration, amphiphile molecules are placed in the lattice by random, and then the remainder of the lattice is filled by the water. In an order initial configuration, amphiphile molecules are placed in the lattice by order, and then the remainder of the lattice is filled by the water. So, in all of the configurations there is not any vacancy. A bead in the lattice can interact with six sites, in other words, the nearest sites are considered for interaction. So, diagonal sites that are far from the cutoff range are not considered. A bead in a chain is not considered to interact with the beads connected directly to it in the same chain.

The model, similar to the Nelson's model, considers only repulsive interactions between unlike sites without any attractive terms as follows,

$$E = \varepsilon/k_B T (n_{TH} + n_{TW}), \tag{1}$$

where $\varepsilon/k_B T$ is dimensionless interaction parameter that can be related to Flory–Huggins parameter χ through $\chi = z\varepsilon/k_B T$ [22]. T, H, and W are symbols represent tail, head, and water sites, respectively. One can consider different models for self-assembling of amphiphile molecules but the potential model considered here is the simplest one.

Because in some cases reptation is not an ergodic [17] move, we use reptation in conjunction with configurational

bias Monte Carlo (CBMC) defined by Siepmann and Frenkel [23]. In reptation move, one end unit of a chain is selected randomly, and is moved to its nearest neighboring site. Then, the second unit, which is attached to that end unit, is moved to the site formerly occupied by the end unit. The third unit is moved to the site formerly occupied by the second, and so on. A reptation move should not cause a chain to be severed. A move is chosen such that the self avoidance and mutual avoidance conditions are not violated. The probability of acceptance of a reptation move is calculated according to the standard Metropolis algorithm [24]. The Metropolis probability is given by:

$$P = \min(1, \exp(-\beta \Delta E)), \tag{2}$$

where ΔE is the difference in the total internal energy between the trial and old configurations. If a move is not accepted, the old configuration is maintained and a new move is tried.

Configurational bias Monte Carlo move is used to relax the system faster. This move is based on the following algorithm:

1. One chain is chosen at random for removal.
2. The Rosenbluth weight (W) for this chain is then calculated by counting the number of solvent sites ($z_s(i)$) around the next bead to be removed. For the chain that is being removed this factor is represented by W_{old}, which is defined as:

$$W_{old} = \prod_{i=m-1}^{1} \frac{z_s(i)}{z}, \tag{3}$$

where m is the chain length, and z is the coordination number of the lattice which is equal to 6.
The product in Eq. 3 represents the total number of possible configurations of deleted chain. For deletion of the chain, the end segment is deleted, but it is not considered in the above product because, after it has been deleted we do not have any segment left to calculate its choices.
3. A lattice site is chosen at random for growing the removed chain.
4. The chain is grown. After setting each site in its place on the lattice, the possible ways that the next site can be placed are calculated. By this method, the Rosenbluth weight (W) for this chain is calculated:

$$W_{new} = \prod_{i=1}^{n-1} \frac{z_s(i)}{z}. \tag{4}$$

If $z_s(i)$ becomes zero for any of the sites, then the system is returned to its old state and the configuration is not altered. Detailed balance is maintained by using the ratio of the two

Rosenbluth weights in the acceptance criterion:

$$P_{acc} = \min\left\{1, \frac{W_{new}}{W_{old}} \exp(-\beta \Delta E)\right\}. \quad (5)$$

We used reptation and the CBMC move with equal probability for both surfactant and oil molecules.

During the simulation, we calculate monomer number density, X_1, cluster number density with an aggregation number n, X_n, and the distribution of monomers in different clusters, nX_n, that is referred in the literature to as the cluster size distribution. It must be noted that these definition (number density) are different from our previous definition for concentration (mole fraction). For above parameters the following definitions are used:

$$X_1 = \frac{\text{number of monomers}(N_1)}{\text{total sites of the lattice}(L^3)}, \quad (6)$$

$$X_n = \frac{\text{number of aggregates containing n amphiphiles}(N_n)}{\text{total sites of the lattice}(L^3)}, \quad (7)$$

where, for our previous definition of mole fraction, the denominators in above equations were total number of molecules in the box of simulation. In some cases, the total concentration of surfactant is represented by the volume fraction (v_s). The volume fraction for surfactant molecules is defined as:

$$v\_s = \frac{\text{total sites of amphiphile molecules}}{\text{total sites of the lattice}(L^3)}. \quad (8)$$

To study the effect of correlation between the snapshots taken for the ensemble averaging, we have calculated the energy auto-correlation function. The energy auto-correlation function is defined as

$$A(\tau) = \frac{\langle E(t+\tau)E(t)\rangle - \langle E(t)\rangle^2}{\langle E^2(t)\rangle - \langle E(t)\rangle^2}, \quad (9)$$

where $A(\tau)$ is the correlation between snapshots at time t and $t + \tau$. It is clear that the maximum of correlation takes place at $t = 0$ and the minimum of it is zero at snapshots that are independent of each other. $E(t)$ is the total energy of the lattice at time t and $E(t + \tau)$ is the total energy for time $t + \tau$; $\langle \cdots \rangle$ refers to the averaging over snapshots. It must be noted that to get a smooth curve, very large snapshots should be taken, i.e., very long run should be done. Error estimates on averages of correlated data have been done using the blocking method of Flyvbjerg and Petersen [25]. In this method in each blocking period the data set transforms into new set with half of the length of the old set. New data in new set are calculated by averaging two near data in the old set. It has been shown that under this blocking transformation, the variance is invariant [25], they have shown that the following
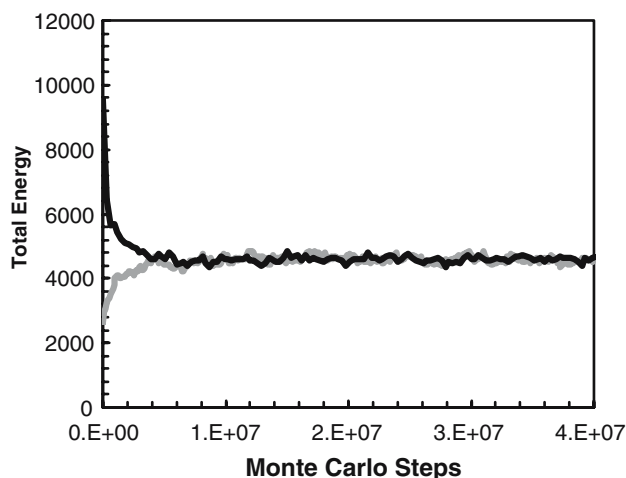
Fig. 1 Convergence of the total energy for two different initial configurations (highly random and highly order initial configurations) at volume fraction of 0.04 for the mixture of H_2T_4 and water. This is suggestive that there is no strong metastability with a timescale longer that the simulations considered here

relation is true for the variance in the block transformation,

$$\sigma^2(\bar{x}) \geq \left\langle \frac{c_0}{n-1} \right\rangle, \quad (10)$$

where $\bar{x}$ is average of n measurements of some fluctuating quantity, and c_0 is defined as follows,

$$c_0 = \frac{1}{n}\sum_{k=1}^{n}(x_k - \bar{x})(x_k - \bar{x}) \quad (11)$$

starting with a data set $x_1, x_2, \ldots, x_n$, $c_0/(n-1)$ is computed, and used as estimate of $\langle c_0/(n-1)\rangle$. Then, the blocking transformation is used to produce a new data set of $x_1', x_2', \ldots x_{n'}'$. As the old set, $c_0'/(n'-1)$ is computed, and used as estimate of $\langle c_0'/(n'-1)\rangle$. This process is repeated until $n' = 2$. (Please refer to the Ref. [25] for more information)

3 Results and discussion

The results are for the mixture of water and nonionic lattice surfactant (H_2T_4) with concentration of $v\_s = 0.04$. The convergence of the total energy for two different initial configurations shows that there is not any meta-stability in the system. This convergence has been shown in Fig. 1. Figure 2 shows monomer number density (X_1) versus total volume fraction of surfactant (v_s) for different lattice sizes to investigate any size effect. According to Bernardes et al. [16], there is not any size effects for the lattice of size of 50.

We have shown in our previous papers that the cluster size distribution for pure and mixed systems of surfactant has just one micelle peak [9,21,26,27]. In the canonical ensemble the
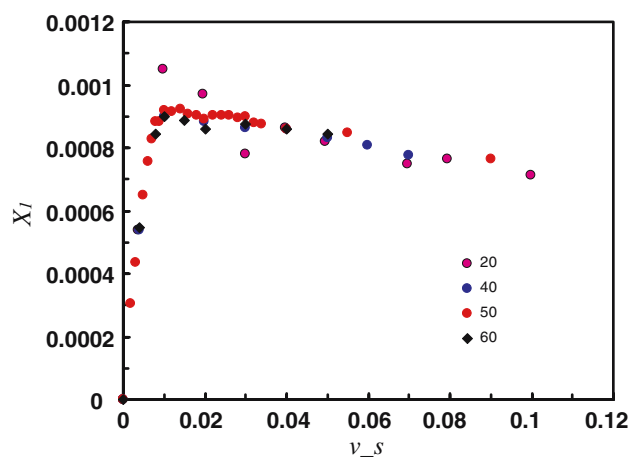
Fig. 2 Monomer number density (X_1) versus total volume fraction (v_s) for H_2T_4 at $\frac{\varepsilon}{k_B T} = 0.7$ for different lattice sizes. This figure shows that for lattice sizes of 50 there is not any size effect

Fig. 3 Auto-correlation function, $A(\tau)$, as a function of Monte Carlo steps. The best equation fitted to this figure is $A(\tau) = 0.75\,e^{-0.037x}$. After $2\tau = 2 \times 27 \times 50,000 = 2.7 \times 10^6$ Monte Carlo steps, there is no any correlation between snapshots. For getting correct cluster size distribution, Monte Carlo steps between snapshots taken for ensemble averaging must be at least 2.7×10^6 steps

equilibrium state is fulfilled when Helmholtz free energy of the system is minimized. This minimum condition provides the equilibrium population distribution of the aggregates in solution. To derive the distribution, the method of undetermined multipliers containing the constraint of $N = \sum_{n=1}^{\infty} n N_n$ can be used in which N is the total surfactant molecules and N_n represents number of aggregates with aggregation number n. So, the equilibrium condition leads to the following equations:
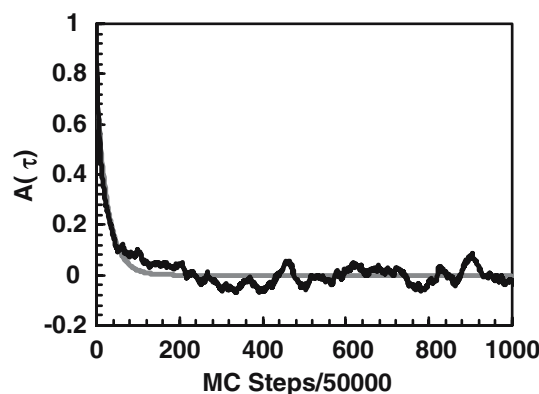
$$\mu_n = n\mu_1 \tag{12}$$

chemical potential μ_n can be related to the activity through $\mu_n = \mu_n^\circ + k_B T \ln a_n$. This equation can be written as $a_n = a_1 \exp(-\beta(\mu_n^\circ - n\mu_1^\circ))$. If we suppose that inter-aggregate interactions are absent, the above equation can be written as $x_n = x_1 \exp(-\beta(\mu_n^\circ - n\mu_1^\circ))$. The generic form assigned to $(\mu_n^\circ - n\mu_1^\circ)$ is the following (at least for nonionic surfactant)[22]:

$$\mu_n^\circ - n\mu_1^\circ = \alpha n^{\frac{4}{3}} - \beta n + \gamma n^{\frac{2}{3}}, \tag{13}$$

where α, β and γ are positive parameters that depend on temperature and the molecular structure of surfactant. Ben-Naim and Stillinger [28], Goldstein [29] and, Ruckenstein and Nagarajan [30] have considered different form of above equation and proved that in the cluster size distribution there is only one peak for micellar aggregates at the concentration above critical micelle concentration.

Besides theoretical works, some simulations have been devoted to this important problem. Nelson et al. [11,17] have shown that finite size effect of the system can bias the cluster size distribution to have multiple peaks [11,17]. On the other hand, some other simulation works in the mixed system of oil–water–surfactant have interpreted the additional peaks in the cluster size distribution (peaks beyond the micelles peak)

as phase or shape transitions [18–20]. In this paper, we show that the additional peak in the cluster size distribution for H_2T_4 at volume fraction of 0.04 in our previous paper is a result of statistical uncertainty in ensemble averaging.

At the first glance, it seems that the correlation between snapshots taken for ensemble averaging may have bias effects on the cluster size distribution. So, firstly we show the effect of correlation on the cluster size distribution without any statistical error for the averages, and then it is shown that the additional peaks in cluster size distribution are in the range of statistical uncertainty and can not be accepted as true peaks.

To show the effects of correlation, simulations were run in a lattice size of 50 (which does not show any size effect) one time without the correlation between snapshots and the other time with significant correlation between snapshots in ensemble averaging. To do this, the number of snapshots taken for the ensemble averaging, is considered on the basis of energy auto-correlation function of Fig. 3. Figure 4 shows the final configuration of the box of the simulation for the system at volume fraction of 0.04. This figure shows the aggregates are not spherical, i.e., they undergoe shape transitions. Figs. 5 and 6 show the results of cluster size distribution for the above system with and without correlation, respectively. From these results, the reader may conclude that this is the correlation biases the cluster size distribution.

On the other hand, Figs. 7 and 8 show the cluster size distribution for the system of Figs. 5 and 6 with considering statistical errors calculated by blocking method. As is clear in these figures, the additional peak for Fig. 7 is in the range of the statistical uncertainty.

In a pure system (mixture of a surfactant and water almost in the literature known as a pure system), Nelson et al. have fitted the cluster size distribution to simple theories which assume the micelles are spherical and cylindrical. The

Fig. 4 Final configuration of the box of simulation. As is clear at volume fraction of 0.04 of H_2T_4, some aggregates have cylindrical shapes

aggregation distribution is split into two components $C_N = C_N^S + C_N^C$ where C_N^S is the volume fraction of surfactants in spherical micelles of aggregation number N and C_N^C is the volume fraction of surfactants in cylindrical micelles of aggregation number N. The main peak in the distribution is fitted to a Gaussian distribution for spherical micelles:

$$C_N^S = C_M \exp\left(\frac{-(N-M)^2}{2\sigma^2}\right), \tag{14}$$

where, M is the most probable aggregation number for the spherical micelles with peak height C_M and σ is the width of the Gaussian distribution. The long tail in the distribution is fitted to an exponential distribution for cylindrical micelles of the form

$$C_N^C = \beta \exp(-\alpha N) f(N), \tag{15}$$

where $f(N)$ is a matching function of the form

$$f(N) = \begin{cases} 0 & N \le M \\ \dfrac{C_M - C_N^S}{C_M} & N > M \end{cases} \tag{16}$$

The matching function has a value of unity at large aggregation numbers and smoothly takes the cylindrical micelle distribution to zero at an aggregation number of M, at values lower than M there can be no cylindrical micelles by definition. They concluded that as the surfactant concentration is increased the distribution becomes broader, with the excess
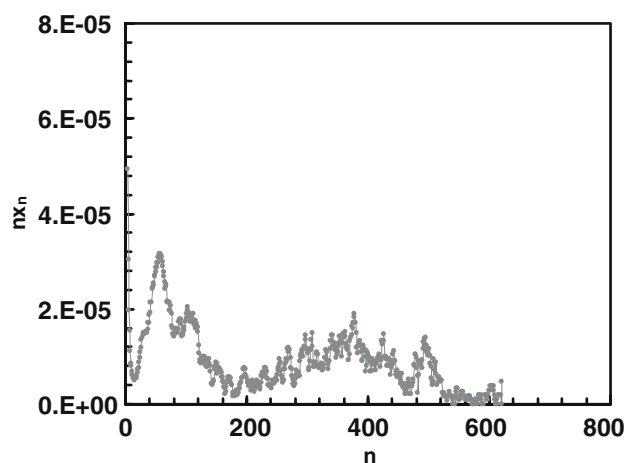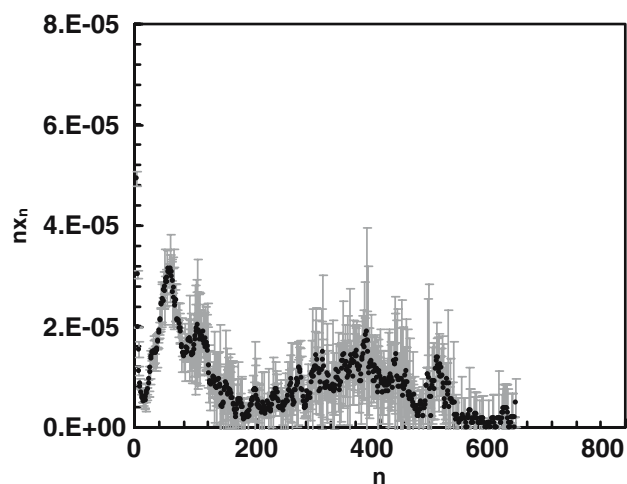

Fig. 5 Cluster size distribution for H_2T_4 with correlation between snapshots. In this figure, number of snapshots is equal to 1,024, and the total Monte Carlo steps after relaxation is equal to 3×10^8. So, the Monte Carlo steps between two snapshots is approximately equal to 3×10^5 which is less than 2τ and causes the correlation between snapshots
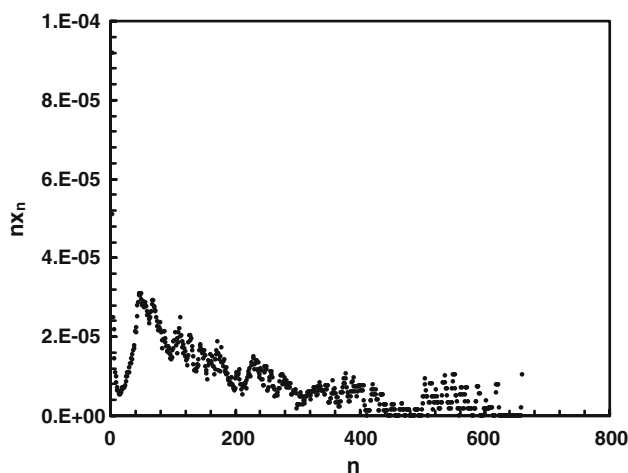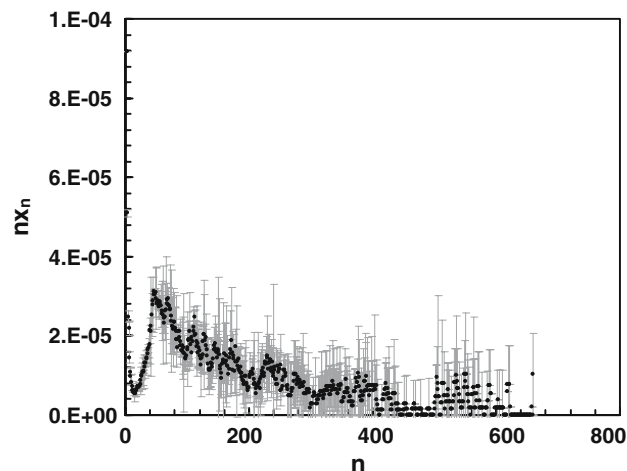

Fig. 6 Cluster size distribution for H_2T_4 without any correlation between snapshots. In this figure, number of snapshots is equal to 512, and the total Monte Carlo steps after relaxation is equal to 1.9×10^9. So, the steps between two snapshots is equal to 4.45×10^6 which is greater than 2τ and omit any correlation between snapshots

surfactant moving into the long cylindrical tail of the distribution.

We have repeated the above theory used by Nelson et al. for our system. Figure 9 shows the cluster size distribution and the functions have been used to fit our results. This figure indicates that a shape transition only changes the distribution to be broader and do not make any additional peak in the cluster size distribution. This result is similar to the result of Nelson et al., and is a confirmation to the error effects in the cluster size distribution.

Fig. 7 Cluster size distribution with the errors calculated by blocking method for H_2T_4 with the conditions as Fig. 5
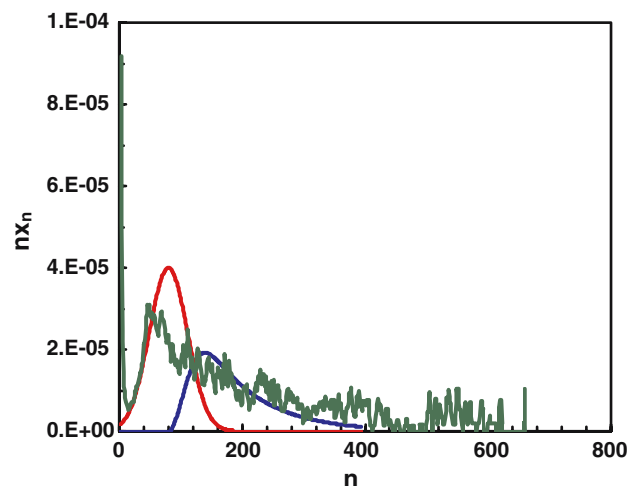

Fig. 8 Cluster size distribution with the errors calculated by blocking method for H_2T_4 with the conditions as Fig. 6

Fig. 9 Fitting the cluster size distribution without any correlation (see Fig. 6) with simple theory used in the work of Nelson et al. Red line is the Gaussian function with the formula $C_N^s = 4 \times 10^{-5} \times e^{-\frac{1}{2}\left(\frac{N-80}{31}\right)^2}$, and blue line is a function with the formula $C_N^c = 1.2 \times 10^{-4} \times \left(\frac{4\times10^{-5} - C_N^s}{4\times10^{-5}}\right) \times e^{-\frac{N}{84}}$

4 Concluding remarks

In this paper we showed that the additional peaks are results of statistical uncertainty for the averaged cluster sizes. Also it was shown that the reduction of correlation can reduce the statistical error for ensemble averaging. Simulations showed that the shape transition changes the shape of the cluster size distribution to be broader, and the presence of additional peak in the cluster size distribution can not be interpreted as a clue for the shape or phase transition.

References

1. Frenkel D, Smit B (1996) Understanding molecular simulation: from algorithms to applications. Academic, London
2. Allen MP, Tildesley DJ (1987) Computer simulation of liquids. Clarendon, Oxford
3. Binder K (1992) In: Marcel JB (Ed) Computational modeling of polymers. Dekker, New York
4. Doye JPK, Sear RP, Frenkel D (1998) J Chem Phys 108:2134
5. Bohbot Y, Ben-Shaul A, Granek R, Gelbart WM (1995) J Chem Phys 103:8764
6. Mackie AD, Onur K, Panagiotopoulos AZ (1996) J Chem Phys 104:3718
7. Larson RG, Scriven LE, Davis T (1985) J Chem Phys 83:2411
8. Larson RG (1989) J Chem Phys 91:2479
9. Behjatmanesh-Ardakani R (2005) J Chem Phys 122:204903
10. Bhattacharya A, Mahanti SD (2001) J Phys C 13:L861
11. Nelson PH, Rutledge GC, Hatton TA (1997) J Chem Phys 107:10777
12. Bhattacharya A, Mahanti SD, Chakrabati A (1998) J Chem Phys 108:10281
13. Bourov GK, Bhattacharya A (2003) J Chem Phys 119:9219
14. Groot RD, Warren PB (1997) J Chem Phys 107:4423
15. Bourov GK, Bhattacharya A (2005) J Chem Phys 122:044702
16. Bernardes AT, Henriques VB, Bisch PM (1994) J Chem Phys 101:645
17. Nelson PH (1998) Simulation of self-assembled polymer and surfactant systems. Ph.D. Thesis Massachusetts Institute of Technology
18. Talsania SK, Rodriguez-Guadarrama LA, Mohanty KK, Rajagopalan R (1998) Langmuir 14:2684
19. Lisal M, Hall CK, Gubbins KE, Panagiotopoulos AZ (2002) J Chem Phys 116:1171
20. Talsania SK, Wang Y, Rajagopalan R, Mohanty KK (1997) J Colloid Interfac Sci 190:92

21. Gharibi H, Behjatmanesh-Ardakani R, Hashemianzadeh SM, Mousavi-Khoshdel SM, Javadian S, Sohrabi B (2006) Theor Chem Acc 115:1
22. Tanford C (1973) The hydrophobic effects. Wiley, New York
23. Siepmann JI, Frenkel D (1992) Mol Phys 75:59
24. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) J Chem Phys 21:1087
25. Flyvbjerg H, Petersen HG (1989) J Chem Phys 91:461
26. Gharibi H, Behjatmanesh-Ardakani R, Hashemianzadeh M, Mousavi-Khoshdel M (2006) J Phys Chem B 110:13547
27. Gharibi H, Behjatmanesh-Ardakani R, Hashemianzadeh SM, Sohrabi B, Javadian S (2006) Bull Chem Soc Jpn 79:1355
28. Ben-Naim A, Stillinger FH (1980) J Phys Chem 84:2872
29. Goldstein RE (1986) J Chem Phys 84:3367
30. Ruckenstein E, Nagarajan R (1975) J Phys Chem 79:2622