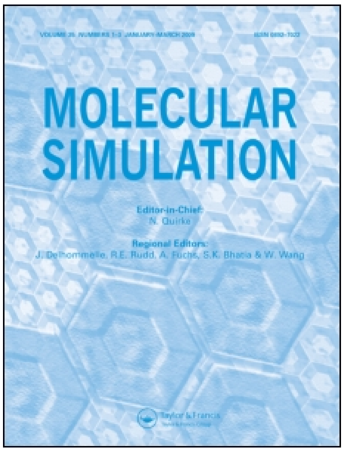

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Structural organization of a chain molecule with specific charge distribution: A molecular dynamics study

A. Baumketner[a]; H. Shimizu[b]; Y. Hiwatari[a]
[a] Department of Computational Sciences, Faculty of Science, Kanazawa University, Kanazawa, Japan [b] Department of Physics, Shinshu University, Matsumoto, Japan

To cite this Article Baumketner, A. , Shimizu, H. and Hiwatari, Y.(2002) 'Structural organization of a chain molecule with specific charge distribution: A molecular dynamics study', Molecular Simulation, 28: 4, 359 — 375
To link to this Article: DOI: 10.1080/08927020290018750
URL: http://dx.doi.org/10.1080/08927020290018750

Molecular Simulation, 2002 VOL. 28 (4), pp. 359–375

STRUCTURAL ORGANIZATION OF A CHAIN MOLECULE WITH SPECIFIC CHARGE DISTRIBUTION: A MOLECULAR DYNAMICS STUDY

A. BAUMKETNER[a,*], H. SHIMIZU[b] and Y. HIWATARI[a]

[a]*Department of Computational Sciences, Faculty of Science, Kanazawa University, Kakuma, Kanazawa 920-1192, Japan;* [b]*Department of Physics, Shinshu University, 3-1-1 Asahi, Matsumoto 390-8621, Japan*

We report the results of molecular dynamics simulations of a charged bead-monomer chain molecule with charge distribution adopted from immunoglobulin-binding domain B1 of protein-g. The beads of the model are connected by invariable bonds and interact with each other *via* the Coulomb potential. To study the low-temperature conformational space of the designed model we use standard canonical, microcanonical and multicanonical molecular dynamics simulations. We find that at low temperature $T = T_c$ the chain undergoes a continuous freezing transition into one of many low-energy conformations. Below T_c the molecule is a compact globule composed of an inner core, containing mostly charged monomers, and an outer corona, consisting of all the rest neutral units. All frozen conformations have almost equal potential energy but differ in structure. The potential energy surface of the model does not posses a pronounced ground-state minimum—an essential feature of protein-like heteropolymers.

Keywords: Molecular dynamics; Multicanonical ensemble; Long-range interaction; Structural transition; Off-lattice heteropolymer

INTRODUCTION

That the long-range Coulomb interaction plays an important role in biologically active molecules, and especially proteins, is a generally recognized fact [1,2]. It has long been known that the Coulomb force substantially contributes to both

---

*Corresponding author. Permanent address: Institute for Condensed Matter Physics, 1 Svientsitsky Str., Lviv 79011, Ukraine. E-mail: andrij@icmp.lviv.ua

ISSN 0892-7022 © 2002 Taylor & Francis Ltd
DOI: 10.1080/08927020290018750

stabilization of native folds [3] and folding process itself [1]. The recent interest to the electrostatic force in the context of the protein folding problem has been motivated by the appreciation that the charge of the residues capable of carrying charged side chains in compact globular states may not be fully screened because of the lack of the solvent in inner parts of the protein. This goes in contrast to the situation encountered in extended coil conformations, where solvent molecules effectively screen protein charges thereby diminishing the effect of the Coulomb force [4,5]. Given the long-range nature of the latter, one may expect the electrostatic potential to play a central role in the protein folding process, along with the hydrophobic and hydrogen bond interactions [2].

The primary structure of a protein is believed to fully determine its three-dimensional native conformation. The non-random character [4] of the primary sequence of amino acids implies that electrical charges in protein molecules are fixed at certain specified sites along the chain. As was shown both experimentally [6] and by computer simulations [7], varying the amount of charge on some specific position of the sequence leads to qualitative structural changes of the whole molecule, namely formation of salt bridges. Thus, the distribution of charged monomers along the backbone emerges as another important aspect of the protein's potential energy field that can substantially influence the way in which it reaches its final native fold. As far as the charge distribution is concerned a question of particular interest is whether having a "correct" (adapted from a real protein) charge sequence is a sufficient physical prerequisite for a model polymer to have protein-like properties, e.g. pronounced potential energy minimum located at a unique ground state [8]. That the long-range interaction may lead to the uniqueness of the ground state in chain heteropolymers was demonstrated in Ref. [9], where by means of exact enumerations of short heteropolymeric chains it was shown that polyampholytes have only one, or few, lowest energy conformations in contrast to polymers with short-range potential between monomers, whose ground states are highly degenerate. Unique ground states were also found in several other works on polyampholytes [10,11]. Despite the fact that heteropolymeric chains with long-range interaction between monomers have been a subject of many theoretical studies so far (see for example Ref. [12] for a recent review), the implications of the specificity of charge distribution, adapted from real protein prototypes, to the structural properties of polymer models have not yet been studied. To partially bridge this gap in our understanding of the statistics of charged heteropolymers is the main motivation of the present work.

In this paper we study structural, thermodynamic and dynamic properties of a chain molecule consisting of spherical beads which interact *via* electrostatic potential and which are connected by invariable chemical bonds; the molecule's charge sequence was adopted from the binding domain B1 of protein-g, 2gb1 in the PDB classification. In addition to the Coulomb force, to mimic the excluded volume effect, all the monomers are subject to pairwise soft-core potential. For our model, we performed molecular dynamics simulations in microcanonical (*NVE*), canonical (*NVT*), and multicanonical ensembles (MMD) systematically in a wide range of temperature. At high temperatures the model is found preferably in extended coil conformations. Its dynamics is governed primarily by soft-core collisions and the Coulomb force plays a minor role. As temperature declines the molecule transforms into a liquid-like globule. Upon further decrease of temperature the molecule undergoes a glass transition at $T{=}T_{\mathrm{c}}$ into one of the multitude of low-energy conformations. The transition manifests itself as a maximum of the specific heat at constant volume C_v. Particular final frozen conformation occupied by the chain at $T < T_{\mathrm{c}}$ strongly depends on the rate of the corresponding cooling process. All frozen globules are composed of an inner core, containing mostly charged monomers, and an outer corona, formed by all the rest neutral monomers. The motions of the charged and neutral monomers, governed by the Coulomb and soft-core potential, respectively, take place at distinct time scales and follow different dynamic patterns in frozen globules. While the charged monomers can only vibrate around their equilibrium positions the neutral monomers experience slow propagating-type motions restricted only by bond connectivity and soft-core repulsion.

Availability of a large array of compact conformations obtained in the MMD run enabled us to calculate the potential energy surface as a function of two coordinates—distances from the target conformation to two arbitrarily chosen conformations. The potential energy surface exhibits no distinct minimum appropriate for designed protein sequences but has, instead, broad regions of minima of comparable depth, i.e. the system behaves much like random heteropolymers [13].

The plan of the paper is as follows. In “Model and computational methods” section, we give a short outline of our model and describe the computational methods of multicanonical molecular dynamics and other techniques we used in this study. In “Results and discussion” section, we give the results of the numerical simulations carried out for the model. Finally, we summarize our work in “Concluding remarks” section.

MODEL AND COMPUTATIONAL METHODS

To study the specificity of the charge distributions found in proteins we constructed an off-lattice heteropolymer model by using the immunoglobulin-binding domain B1 of protein-g, 2gbl, as the prototype system. The original protein contains 56 amino acid residues, six of which are positively charged $+e$, 10 are negatively charged $-e$, and the rest 40 monomers are neutral (non-charged); by e we denote here the elementary electric charge. Each amino acid residue is represented in the model by a bead of the same for all monomers diameter σ. Each monomer is connected to an adjacent bead by a virtual bond of fixed length which we also set σ. The resultant sequence of monomers is shown in Fig. 4. For the interaction between charged monomers we adopt the Coulomb potential:

$$V_{\mathrm{c}}(r_{ij}) = \frac{q_i q_j}{|\vec{r}_i - \vec{r}_j|}, \tag{1}$$

where q_i and $\vec{r}_i$ are charge and position of monomer i, respectively. All monomers, i.e. charged and neutral, arc subject to the repulsive soft-core interactions:

$$V_{\mathrm{sc}}(r_{ij}) = 4\epsilon_{\mathrm{sc}}\left(\frac{\sigma}{|\vec{r}_i - \vec{r}_j|}\right)^{12}, \tag{2}$$

where parameter ϵ_{sc} describes the potential strength. Covalent bond interaction between adjacent along the chain units is realized through the constant bond-length constraints:

$$\sigma = |\vec{r}_i - \vec{r}_j| = \text{constant}. \tag{3}$$

The original protein-g is a globally charged molecule with the net charge $-4e$. In solutions this net charge is compensated for by the presence of counterions which make the whole system electrically neutral. In the present study we do not introduce counterions but we cancel the total charge of the molecule by reassigning fractional charges $-0.6e$ to all negative monomers, instead. The positive and neutral monomers were assigned charges $q_i = +e$ and $q_i = 0$, respectively.

The parameters of the soft-core potential were taken from Ref. [10]: $\epsilon_{\mathrm{sc}} = 2.5 \times 10^{-3}$ a.u., $\sigma = 3.4$ a.u., $m_i = m = 137$ a.u. Throughout the paper, including the formulas given above, we use the atomic system of units. In addition, temperature will be given in units of $e^2/\sigma k_{\mathrm{b}}$ where k_{b} is the Boltzmann constant.

The main computational tools in the present study are molecular dynamics methods. Their principal advantage over Monte Carlo methods is that they allow one to assess both structural properties and microscopic dynamics of the system. By employing an off-lattice model we avoid potential difficulties related to inadequate description of closely packed native conformations, which are encountered in simulations of lattice models [14]. The well-known difficulty faced by standard MD and MC techniques in simulations of systems with complicated potential energy functions, particularly polymers, is the multiple minima problem. Because of the presence of bond-length constraints, potential energy surface of model polymers has many local minima separated by high energy barriers. At low temperatures the events of crossing the barriers become increasingly rare and the system is very likely to get trapped in one of the local minima. This results in very long relaxation times at low T which may far exceed computational times manageable by the present-day computers. When the escape time from a local minimum becomes comparable to some characteristic time of the problem in study, for example folding time for the protein folding problem, the system is called a glass. A method specifically designed to overcome the slow relaxation problem is the multicanonical ensemble sampling, proposed by Berg and Neuhaus [15,16]. Its efficiency was demonstrated in simulations of a first order transition in model spin systems where use of the multicanonical technique reduced the total computational time by at least two orders of magnitude as compared to the traditional *NVT* sampling. The multicanonical scheme was adapted to the molecular dynamics method by Nakajima *et al.* [17,18] in their MD studies of a model double-well system and a small five-residue peptide. It was shown, for the case of the peptide, that the multicanonical MD algorithm was able to sample much broader conformational space than the conventional *NVT* scheme. In the present paper, we use a combined technique of the multicanonical MD algorithm with a bond constraints method, devised by Shimizu *et al.* [10] for a study of di-block polyampholytes. We confine ourselves here to noting that the bond-length constraint in our simulations was realized by the NIMM algorithm [19] and for all other details, regarding the temperature thermostat and the procedure to smooth energy histograms employed in the present work we refer to Ref. [10].

RESULTS AND DISCUSSION

To establish the results quoted in Introduction we performed several canonical and multicanonical simulations. The multicanonical simulation, consisting of an initial *NVT* run and several subsequent MMD iterations, was carried out at $T = 0.02$. We recall that the temperature is measured in reduced units $e^2/\sigma k_b$. The initial canonical run contained 20×10^6 time steps; its main purpose was to determine a first estimate of the density of states $n(E)$ which was refined in later iterations according to the multicanonical recursion procedure [16]. A total of 10 multicanonical iterations of the same length as that of the initial run were performed. After the 10th MMD iteration we found that the broadening of the energy distribution function $P_\mu^{10}(E)$ in the direction of low energies became exceptionally slow. In order to explore in more detail the region of low energies, where the most interesting physics is expected to occur, we performed a second MMD simulation at the temperature of 0.003. This consisted of 30 runs: one initial *NVT* run and 29 iterations. Like in the first MMD simulation, all the runs of the second MMD simulation were 20×10^6 time-step-long. An important point we would like to make here about the second MMD run is that the convergence of energy distributions, $P_\mu^k(E)$, towards a function of flat shape was unexpectedly poor. The energy histograms, instead of getting more and more broad and flat with each successive iteration, behaved completely conversely. They exhibited sharp peaked structures with one maximum or two maxima, centred at distinct energies. In order to obtain more or less broad and reasonably flat energy histogram we had to run as many as 29 iterations. Jumping ahead of ourselves, let us say that the statistics of our model at $T \sim 0.003$ is dominated by glassy states. Exactly the same behaviour of the multicanonical algorithm as that found by our study has recently been reported by Bhattacharya and Sethna [20] in the MC study of glassy Lennard–Jones liquids. It was argued that it is the entropy barriers, which trap simulations in multicanonical ensemble in contrast to the potential energy barriers trapping conventional MD simulations. The remedy to the situation that the authors came up with in [20] was dynamic update of the potential energy function during the simulation. This is consistent with the results of the present study since the multicanonical potential energy update is equivalent to frequent recalculation of the weight function; i.e. performing many MMD iterations.

In Fig. 1 we display the energy distribution functions computed in canonical runs at selected temperatures and those obtained from the MMD simulations by the reweighting procedure [21]. The two multicanonical energy distributions, $P_\mu(E)$s, are sufficiently flat. The reweighted canonical distribution functions generated with their help are in good agreement with the $P_\beta(E)$s of the proper canonical simulations, as seen from Fig. 1. Specific heats calculated from canonical and reweighted multicanonical histograms did not differ by more than 1%. We consider the agreement between C_vs calculated by the two independent methods as enough evidence for sufficient reliability of our simulations. The total energy range covered by the two MMD simulations is from -1.74 to -0.5 a.u. which roughly corresponds to the temperature interval 0.002–0.06.

In addition to the described above multicanonical simulations, in this study we also tested the performance of another widely used for the task of tackling slow relaxation phenomena computational tool—the simulated annealing (SA) method [22]. To assess the efficiency of the SA routine as applied to our model we designed two different cooling processes: (a) simulated annealing with

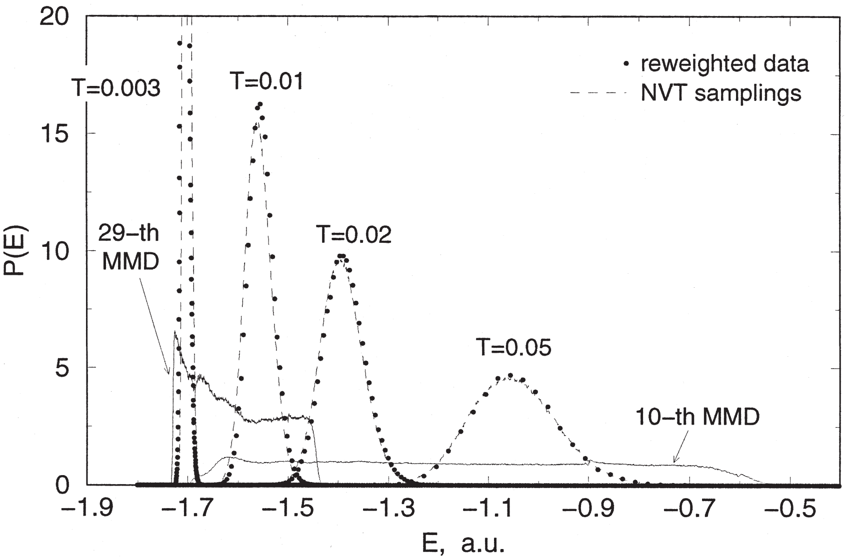

FIGURE 1 Potential energy probability distributions calculated by means of conventional canonical MD and MMD methods at selected temperatures. Arrows show the data from two MMD runs: the 10th MMD iteration at $T = 0.02$ and the 29th iteration at $T = 0.003$.

linear cooling protocol (LNSA) $T_n = T_{n-1} - \Delta T$, and (b) simulated annealing with logarithmic cooling protocol (LGSA) $T_n = T_{n-1}\alpha$. The linear simulated annealing method was used to cool the system down from $T \sim 0.08$ to $T \sim 0.002$ in 5×10^6 time-step-long simulations with the temperature interval $\Delta T = 0.002$. We selected the length of the quoted above equilibration period t_{ec} so as to make it five times longer than our estimate of the system's relaxation time at $T \sim 0.05$; the latter being 1×10^6 time steps. To further enhance the relaxation processes taking place in the system during the cooling the particle velocities were set anew with random directions, not just rescaled, at each successive temperature. For the case of the logarithmic simulated annealing protocol we used 20 independent replicas, each starting from a different random conformation. Initially, the temperature-rate constant α for all simulated replicas was chosen to be 0.85 and time, over which the system was let to equilibrate at each new temperature, was set $t_{\mathrm{ec}} = 5 \times 10^6$ time steps. In the course of the cooling process, by monitoring the average energy of all 20 replicas at every T_n we found that the energies begun to deviate from each other notably at T about 0.01. The deviation of the potential energy may be interpreted as a sign of the onset of the glass transition in the system, since the relaxation time at $T \sim 0.01$ grows rather long, longer than 5×10^6 steps (cf. with 1×10^5 steps at $T = 0.05$). To better assess the divergence of the relaxation time as temperature declines, we launched another series of simulations for the same number of replicas from $T = 0.012$ to $T = 0.002$ with $\alpha = 0.9$ and $t_{\mathrm{ec}} = 20 \times 10^6$ time steps. This time, the deviation of the energies started at $T \sim 0.0086$ signifying relaxation times longer than 20×10^6 time steps at lower temperatures. As one can see the gain in the temperature region of the reliable canonical ensemble sampling, 0.0086–0.012, achieved by additional simulations is negligibly small as compared with the four-times increased computational costs. Such behaviour of the canonical sampling was to be expected from an NP-hard computational problem. With the present limitation of our computational power being about 50×10^6 time steps, the temperature boundary of canonical-ensemble simulation of the studied system can be drawn at about 0.008. In order to proceed to lower temperatures we have to resort to the multicanonical method as described above.

The time step δt used to integrate the equations of motion varied in our simulations from 150 a.u. at low temperatures, where the monomer motions are relatively slow to 80 a.u. at high temperatures, where the monomers move faster. This choice of the integration step allowed us to achieve less than 0.05% relative temperature drift in all the MD runs carried out. In addition to the described above *NVT* simulations of the cooling processes we also performed several *NVE* runs, mostly to calculate velocity autocorrelation functions. The length of these runs varied from 10^5 to 20×10^6 time steps; the total energy conservation in those runs was not worse than 0.1% of its magnitude.

Figure 2 summarizes the results of all our simulations, including the reweighted MMD data, for potential energy E, specific heat at constant volume C_v and radius of gyration R_{g}. The radius of gyration R_{g} is defined as:

$$R_{\mathrm{g}} = \frac{1}{N}\sqrt{N\sum_i \vec{r}_i{}^2 - \sum_i \vec{r}_i^2}, \tag{4}$$

where the angular brackets stand for ensemble average. Essentially, R_{g} is the mean square distance between two random monomers of the molecule and characterizes its overall size. From Fig. 2(a) one can hardly distinguish any difference between the results of all three computational methods used. They all predict almost the same potential energy for the studied heteropolymer both at high and low temperatures. The situation with the specific heat and radius of gyration is notably different. While at high temperatures the three methods agree reasonably well with each other at $T < 0.01$ the disagreement is apparent. As seen from Fig. 2(b) and (c) both C_v and R_{g} calculated in the canonical simulations of the LNSA and LGSA methods experience remarkable departures from the reweighted in the MMD simulations data. The disagreement between canonical

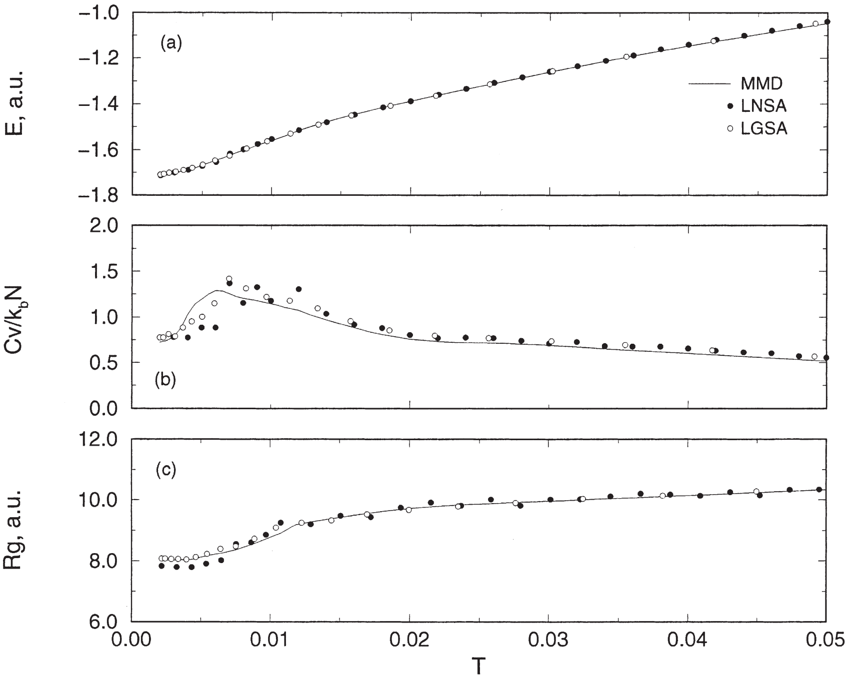

FIGURE 2 Temperature dependence of: (a) potential energy; (b) specific heat at constant volume; (c) radius of gyration as obtained from the two cooling methods described in the text. The solid line denotes the reweighted MMD data.

and multicanonical results lends further support for the conclusions as to the reliability of simulations in canonical ensemble that we made above by analyzing the system relaxation times. Clearly, at $T < 0.01$ the specific heat and gyration radius obtained in both canonical simulations neither are consisted with themselves nor with the MMD simulation. Of the two annealing techniques studied the LGSA method perform marginally better than its linear protocol counterpart as seen from Fig. 2(c) for R_g. This is partly because we used as many as 20 replicas for calculating averages in the LGSA method whereas the LNSA protocol was applied to one replica only, and partly due to the generally established fact that the logarithmic temperature schedule works better. A note concerning comparison of the performance of the logarithmic SA and MMD method is in place here. The multicanonical simulation was able to locate conformations with energy lower than -1.74 a.u. in 10 iterations. This amounts to about 120×10^6 time steps of total computational effort. Of all 20 replicas that we used for the SA method only one reached a conformation with $E < -1.74$ a.u. Thus, we estimate the effort needed for the SA method to achieve a result comparable to that of the MMD method to be of $\sim 7000 \times 10^6$ time steps. Contrasting these two numbers clearly shows that, when applied to the present model, the MMD scheme outperforms the best of our simulated annealing protocols by at least one order of magnitude. Furthermore, the MMD procedure, in contrast to the SA methods, has the advantage of providing canonical averages in the whole sampled temperature range in addition to finding low energy conformations. A conclusion that immediately follows from this comparison is that despite all the difficulties expressed for the multicanonical scheme above, the MMD algorithm still remains the method of choice in numerical studies of linear heteropolymers.

At a temperature about 0.006 the MMD-generated C_v has a distinct maximum which signals a structural transition taking place in the system. The structural transition is accompanied by a rapid compaction of the system as seen from Fig. 2(c) for R_g. The radius of gyration decreases smoothly from 10 Å at $T \sim 0.002$ to 8 Å at $T \sim 0.05$. The largest variation of R_g is observed at $T_c \sim 0.006$ while in the temperature range 0.01–0.05 it is quite flat. In Fig. 3 we display the temperature dependence of the radius of gyration distribution function $P_\beta(R_g)$ reweighted from the MMD histograms. At all temperatures $P_\beta(R_g)$ has unimodal shape. The maximum position smoothly decreases from 10 to 8 Å which implies continuous character for the underlying structural transition. It is important to emphasize the decisive role of the multicanonical sampling in obtaining $P_\beta(R_g)$. As noted above, equilibrating the system at temperatures below 0.008 by means

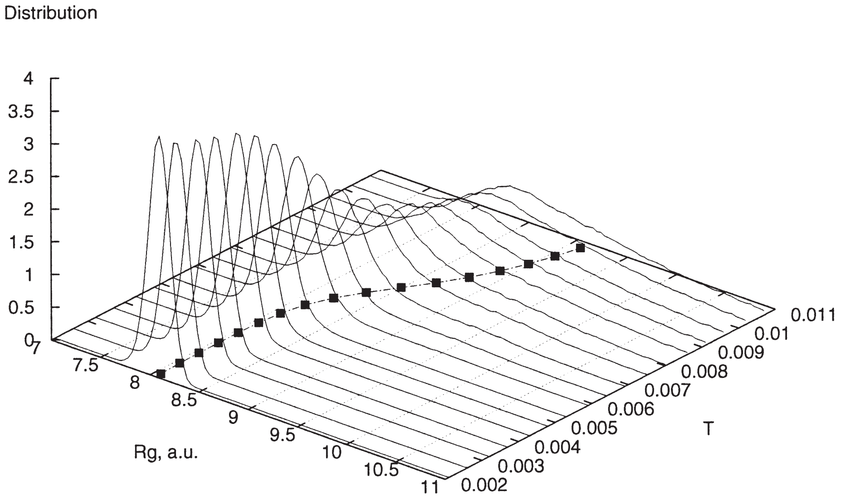

FIGURE 3 Reweighted gyration radius distribution function as obtained from the MMD simulation. By squares, we denote canonical expectations of R_g.

of canonical simulations is beyond our capacity. Thus, every finite-time canonical run at these temperatures would be nothing but a snapshot of few fluctuations. In Fig. 4 we display two example monomer configurations found in the multicanonical simulation at $T = 0.003$. At low energies corresponding to $T < T_c$, Fig. 4(a), our system is a compact globule, composed of an inner core and an outer corona. Majority units of the core are charged monomers while the corona contains all the rest neutral monomers. At high energies, Fig. 4(b), the system finds itself in extended random coil conformations. The fact that both compact and extended conformations were encountered in one MMD simulations provides further support for the utility of the multicanonical scheme. Apparently, the free energy barriers separating the conformations shown in Fig. 4 were effectively reduced during the simulation which enabled the molecule to populate broad region of the conformation space.

Now we turn to discussing the structure of low- and high-T phases. The best way of studying the system's structure is through its pair distribution functions (PDF). The PDFs of positive–negative, positive–positive and negative–negative pairs of monomers at $T = 0.02$ and $T = 0.002$ are shown in Fig. 5. The temperature 0.02 is well above transition point while that of 0.002 is below it. At $T = 0.002$ the PDF was sampled in the vicinity of one of the local energy minima. The considered polymeric chain has a pair of positive–negative and a pair of negative–negative nearest neighbours along the chain (see Fig. 4). The invariable distance σ separating the nearest neighbours gives rise to the first peaks of the PDFs in Fig. 5(a) and (c). All the rest of the PDFs' maxima and minima reflect structural preferences of the present model. At $T = 0.02$ all the PDFs have a liquid-like structure with two well-defined maxima. The positions of the maxima of the monomers of like charges fall onto the positions of the minima

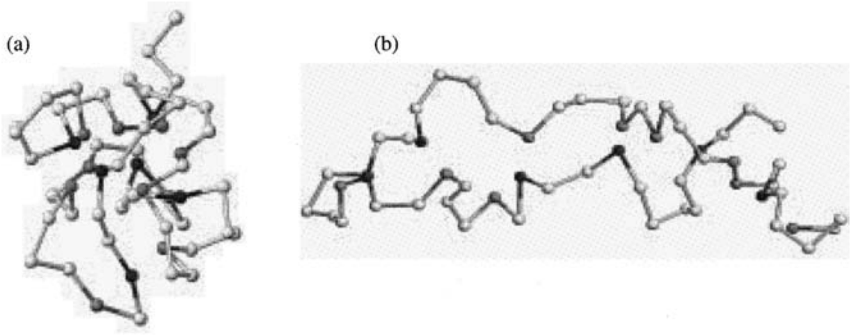

FIGURE 4 Example configurations found in the MMD run at $T = 0.003$ corresponding to: (a) the smallest gyration radius; (b) the largest gyration radius. Black spheres denote positively charged, grey—negatively charged, and white—electrically neutral monomers.

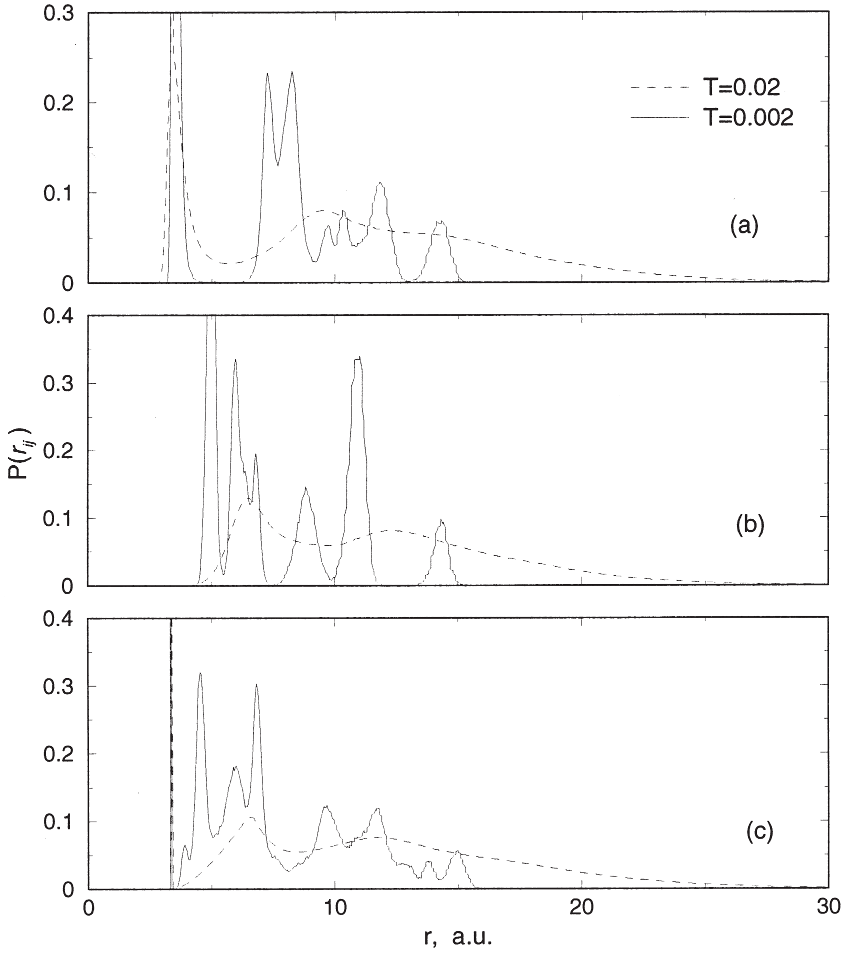

FIGURE 5 Pair distribution functions for: (a) positive–negative pairs of monomers; (b) positive–positive pairs; (c) negative–negative pairs taken at temperatures $T = 0.002$ and $T = 0.02$.

of unlike-charge monomers exactly in the same way as in liquid ionic solutions [23]. This alternating-shell structure of PDF implies screening of the long-range electrostatic interaction within the system limits. At the lower temperature $T = 0.002$ the system's PDF resembles the structure of a solid-like body in having many sharp maxima and minima. We conclude from this sharply peaked shape of PDF that at $T = 0.002$ the molecule is frozen. The freezing transition in random polyampholytes has recently been studied theoretically by Pande *et al.* [4] and Shakhnovich *et al.* [5]. The results of Ref. [4] rely on the Debye–Hückel theory of electrolytes. The validity of this theory is limited to weakly coupled systems where either the solution is dilute or the charges are sufficiently small. Quantitatively, the Debye–Hückel approximation breaks down whenever average distance between particles exceeds the screening radius. In the case of our model, the screening radius is at least twice as small as the typical intermonomer distance in semicompact, liquid conformations and even smaller than that in frozen states. In other words, our chain is a strongly coupled system and its freezing cannot be described in terms of the simple Debye–Hückel approximation. To study the freezing of this system we need to resort to numerical methods. It was predicted in Ref. [5] that freezing of polyampholyte globules may only occur at high densities and fractions of charged monomers. This conclusion is consistent with our findings.

Another important feature of the model studied in this paper is the multiplicity of frozen states at temperatures below T_c. The energy of these states is almost the same—it differs from state to state by no more than 0.1% of its magnitude. But the structure of the frozen states is remarkably different. Summarizing the above observations the potential energy surface of the system has multiple local minima, which correspond to different conformations and have almost same energy. In Fig. 6 we show the potential energy as a function of two coordinates, which measure the distance from a tagged conformation to two randomly chosen

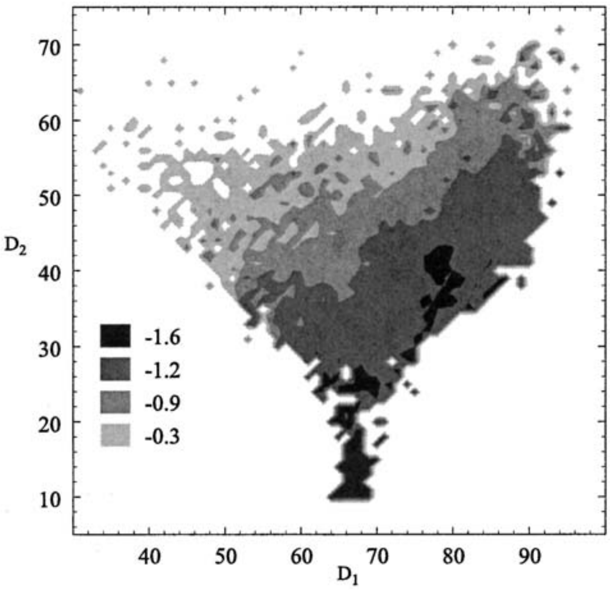

FIGURE 6 Potential energy surface in two coordinates, distances D_1 and D_2 are given in arbitrary units. For the definition of the distances see the text. Energy is given in atomic units.

compact structures. The distance between conformations k and l is defined as

$$D_{kl} = \min\left\{ \sqrt{\sum_{i=1}^{N} (\vec{r}_i^{\,k} - \vec{r}_i^{\,l})^2} \right\} \tag{5}$$

where $\vec{r}_i^{\,p}$ is the position of monomer i in conformation p. The minimum operation min{ } is performed over all relative rotations and translations of the two conformations. Figure 6 was obtained by calculating the potential energy of every conformation of 20,000 monomer configurations recorded in the second MMD simulation, and by subsequent measuring of the distance from it to two compact conformations 1 and 2, which were chosen at random. We see that the potential energy surface has no pronounced minimum corresponding to unique ground state but rather broad regions of minima of almost same depth as predicted earlier. It follows, thus, that by its potential energy properties, our model resembles more random heteropolymers than protein-like biopolymers. Clearly, fixing the charges in appropriate places along the chain is *not* enough to ensure protein-like properties of the molecule.

At the end of the paper we would like to devote a little attention to the microscopic dynamics of the model system studied. The best insight into the dynamics is provided by the simplest dynamical function possible—the velocity autocorrelation function (VAF) $Z(t)$. The velocity autocorrelation functions of charged and neutral monomers of the system are depicted in Fig. 7 at selected temperatures above the transition point, below it and in the transition region. All the VAFs were computed in 20×10^6 time-step-long microcanonical simulations. Figure 7 shows that at $T = 0.05$ the system dynamically behaves much like a gas. The monomers undergo very fast oscillations at approximately same time scale for charged and neutral monomers. The monomer dynamics at that temperature is almost entirely dominated by the stiff soft-core binary collisions. In the transition region of $T \sim 0.006$ one observes a considerable slowing down of the particle motions. Also the time scale of the motions of charged and neutral units begins to separate at this temperature. Charged units are moving much faster than their neutral counterparts. Another characteristics of $Z(t)$ in the transition temperature region is the presence of long-decaying tails in the velocity autocorrelation functions of all types. The tails decay exponentially as $\sim t^{\alpha}$, where $\alpha \sim -2.2$. We took special care to make sure that these poorly decaying tails do not arise from the rotational motion of the molecule as a whole. The angular momentum, which is an integral of motion in the *NVE* simulations performed, was zero within the numerical error of the algorithm.

As we mentioned earlier, the escape time from local potential energy minima at $T < T_{\mathrm{c}}$ is much longer than presently accessible simulation time. Equilibrating

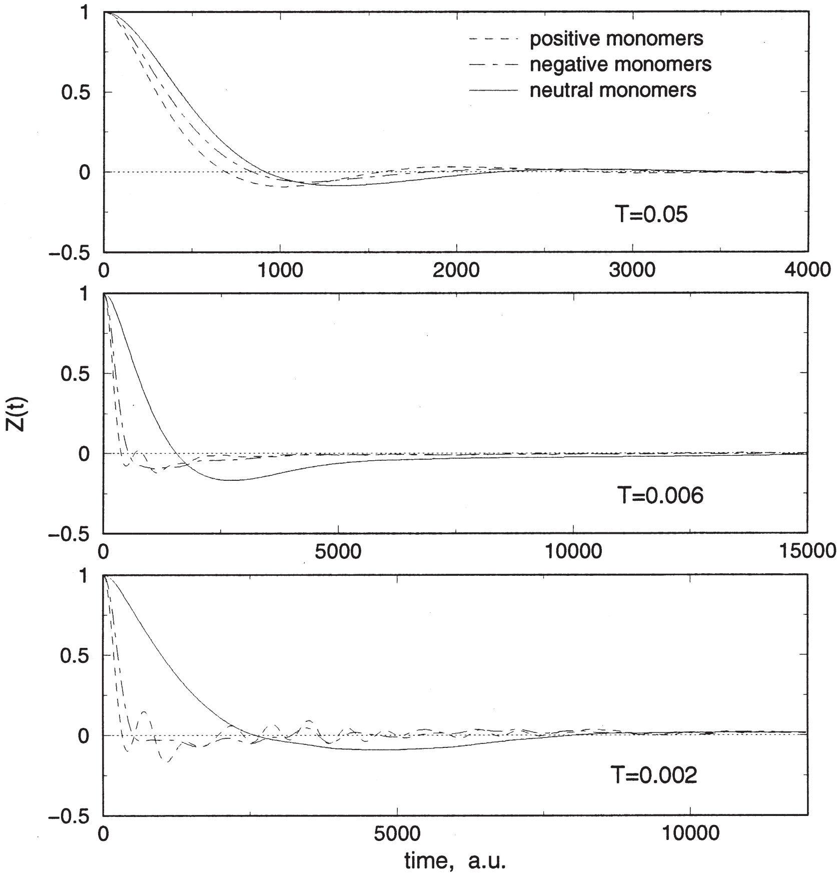

FIGURE 7 Velocity autocorrelation functions of positive, negative and neutral monomers calculated from *NVE* simulations at the temperatures of 0.05, 0.006 and 0.002.

the system at these T in the microcanonical ensemble is not possible and thus, averaged in an *NVE* run, $Z(t)$s depend on the particular minimum from which the simulation started. The difference, however, concerns only the long-time part of the velocity autocorrelation functions of the charged monomers and has no crucial relevance to our analysis. It is seen from Fig. 7 that at $T = 0.002$ the VAFs of charged monomers have pronounced oscillatory character—very similar picture to that found in solid bodies. The neutral monomers are moving significantly slower and follow diffusive rather than oscillatory patterns. This behaviour of the chain monomers can be explained as follows. The charged monomers, which are predominantly confined within the globule core, are bound together by strong electrostatic interaction. Their motions are restricted by the

Coulomb force and by the globule boundaries—a sort of cage effect. The neutral monomers, on the other side, are not restricted by the chain geometry since they are found mostly in the outer corona, surrounding the core. Their mobility is further enhanced by the absence of any attractive force exerted on them, except covalent bonds. All this results in the diffusion-like pattern of the neutral monomers velocity autocorrelation function.

CONCLUDING REMARKS

In this paper we reported the results of our molecular dynamics study of a simple polymer model, constructed on the basis of the immunoglobulin-binding domain B1 of protein-g. The numerical results show that the model undergoes a continuous freezing transition at a temperature about 0.006. The transition takes place into one of multiple low-energy states. The frozen globules consist of a core, containing mostly charged monomers, with all the rest neutral monomers surrounding it. The long-range electrostatic interaction is effectively screened within the core and the core itself has the plasma structure of alternating shells of positive and negative units. The potential energy surface of the studied model does not possess a pronounced minimum corresponding to a unique ground state.

Above the transition temperature the model heteropolymer is a random globule. The pair distribution functions of this globule are very similar to the PDFs of classical liquids near freezing. Their alternating nature of maxima and minima of charged monomers is preserved.

From the dynamical point of view, the time scale of the motions of charged monomers at the core and the motions of neutral monomers around the core is distinctly different.

Acknowledgements

This work was supported by the Research and Development Applying Advanced Computational Science and Technology, Japan Science and Technology Corporation (JST).

References

[1] Fraga, S., Parker, J.M.R. and Pocock, J.M. (1995) Computer simulations of protein structures and interactions (Springer, Berlin).
[2] Creighton, T.E. (1992) In: Creighton, T.E., ed, Protein Folding (W.H. Freeman, New York).

[3] Nilsson, O. and Tapia, O. (1992) "Electrostatic forces and the structural stability of a modelled bacteriophage T4 glutaredoxin fold: molecular dynamics simulations of polyglycine 87-mers", *J. Mol. Struct. (Theochem)* **256**, 295–305.

[4] Pande, V.S., Grosberg, A.Y., Joerg, C., Kardar, M. and Tanaka, T. (1996) "Freezing transition of compact polyampholytes", *Phys. Rev. Lett.* **77**, 3565–3568.

[5] Angerman, H.J. and Shakhnovich, E. (1999) "Freezing in polyampholyte globules: influence of the long-range nature of the interaction", *J. Chem. Phys.* **111**, 772–785.

[6] Shoemaker, K.R., Kim, P.S., Brems, D.N., Marqusee, S., York, E.J., Chaiken, I.M., Stewart, J.M. and Baldwin, R.L. (1985) "Nature of the charged-group effect on the stability of the C-peptide helix", *Proc. Natl. Acad. Sci. USA* **82**, 2349–2353.

[7] Hansmann, U.H.E. and Okamoto, Y. (1998) "Tertiary structure prediction of C-peptide of ribonuclease A by multicanonical algorithm", *J. Phys. Chem. B* **102**, 653–663.

[8] Shakhnovich, E.I. and Gutin, A.M. (1990) "Implications of thermodynamics of protein folding for evolution of primary sequences", *Nature* **346**, 773–775.

[9] Kantor, Y. and Kardar, M. (1995) "Randomly charged polymers: an exact enumeration study", *Phys. Rev. E* **52**, 835–846.

[10] Shimizu, H., Uehara, K., Yamamoto, K. and Hiwatari, Y. (1999) "Structural phase transition of di-block polyampholyte", *Mol. Sim.* **22**, 285–301.

[11] Tanaka, M. and Tanaka, T. (2000) "Clumps of randomly charged polymers: molecular dynamics simulation of condensation, crystallization, and swelling", *Phys. Rev. E* **62**, 3803–3816.

[12] Pande, V., Grosberg, A.Y. and Tanaka, T. (2000) "Heteropolymer freezing and design: towards physical models of protein folding", *Rev. Mod. Phys.* **72**, 259–314.

[13] Abkevich, V.I., Gutin, A.M. and Shakhnovich, E.I. (1994) "Free energy landscape for protein folding kinetics: intermediates, traps, and multiple pathways in theory and lattice model simulations", *J. Chem. Phys.* **101**, 6052–6062.

[14] Tanaka, M., Grosberg, A.Y., Pande, V.S. and Tanaka, T. (1997) "Molecular dynamics study of the structure organization in a strongly coupled chain of charged particles", *Phys. Rev. E* **56**, 5798–5808.

[15] Berg, B.A. and Neuhaus, T. (1992) "Multicanonical Ensemble: a new approach to simulate first-order phase transition", *Phys. Rev. Lett.* **68**, 9–12.

[16] Berg, B.A. (1999) Introduction to Multicanonical Monte Carlo Simulations, LANL e-print archive cond-mat 236.

[17] Nakajima, N., Nakamura, H. and Kidcra, A. (1997) "Multicanonical ensemble generated by molecular dynamics simulation for enhanced conformational sampling of peptides", *J. Phys. Chem. B* **101**, 817–824.

[18] Nakajima, N., Higo, J., Kidera, A. and Nakamura, H. (1997) "Flexible docking of a ligand peptide to a receptor protein by multicanonical molecular dynamics simulation", *Chem. Phys. Lett.* **278**, 297–301.

[19] Yoneya, M., Berendsen, H.J.C. and Hirasawa, K. (1994) "A non-iterative matrix method for constraint molecular dynamics simulations", *Mol. Sim.* **13**, 395–405.

[20] Bhattacharya, K.K. and Sethna, J.P. (1998) "Multicanonical methods, molecular dynamics, and Monte Carlo methods: comparison for Lennard–Jones glasses", *Phys. Rev. E* **57**, 2553–2562.

[21] Ferrenberg, A.M. and Swendsen, R.H. (1988) "New Monte Carlo technique for studying phase transitions", *Phys. Rev. Lett.* **61**, 2635–2638.

[22] Kirkpatrick, S., Gelatt, C.D. and Vecchi, J.M.P. (1983) "Optimization by simulated annealing", *Science* **220**, 671–680.

[23] Hansen, J.P. and McDonald, I.R. (1986) Theory of Simple Liquids, 2nd ed. (Academic Press, London).