

Title	An efficient method for analyzing conformational properties of a polymer in solvent
Author(s)	Amano, Ken-ich; Oshima, Hiraku; Kinoshita, Masahiro
Citation	Chemical Physics Letters (2011), 504(1-3): 7-12
Issue Date	2011-02-28
URL	<a href="http://hdl.handle.net/2433/139483">http://hdl.handle.net/2433/139483</a>
Right	© 2011 Elsevier B.V.; This is not the published version. Please cite only the published version. この論文は出版社版ではありません。引用の際には出版社版をご確認ご利用ください。
Type	Journal Article
Textversion	author

# **An efficient method for analyzing conformational properties of a polymer in solvent**

Ken-ich Amano<sup>1</sup>, Hiraku Oshima<sup>2</sup>, and Masahiro Kinoshita<sup>2</sup>

<sup>1</sup>*Graduate School of Energy Science, Kyoto University, Uji, Kyoto 611-0011, Japan*

<sup>2</sup>*Institute of Advanced Energy, Kyoto University, Uji, Kyoto 611-0011, Japan*

Author to whom correspondence should be addressed: Masahiro Kinoshita.

*E-mail address:* kinoshit@iae.kyoto-u.ac.jp.

## **Abstract**

We propose an efficient method for investigating conformational properties of a polymer in solvent. The method is a combination of a Monte Carlo (MC) simulation applied to the polymer alone and a statistical-thermodynamic approach for incorporating solvent effects. To illustrate it, we analyze conformations of a simple polymer chain stabilized in a hard-sphere solvent. The generation of polymer conformations is performed using the self-avoiding random walk on a cubic lattice. We argue that by introducing the generalized-ensemble techniques to the MC simulation part, the method can be applied to studies on protein conformations in aqueous solution under any thermodynamic condition.

## 1. Introduction

Analysis on the conformational properties of polymers including biopolymers such as proteins is a challenging subject in polymer science, biophysics, and biochemistry. However, a difficult point is that solvent has enormous effects on the conformational properties. The most important solvent is doubtlessly water or aqueous solution. Though the continuum model has frequently been used to incorporate solvent effects, it is not capable of accounting for the solvophobicity which is essential in determining the conformational properties. As argued in our earlier works [1,2] and in later paragraphs of the present letter, a molecular model is necessitated for solvent to elucidate the solvent effects. In usual computer simulations, a polymer and many solvent molecules are simultaneously treated. However, the number of solvent molecules required becomes progressively larger as the polymer size increases. It is important to carefully check if the number is sufficiently large, but such a check is often skipped, which lowers the reliability of the results obtained.

Here we propose an efficient method for incorporating the solvent effects to the full extent through the solvation free energy (SFE) of a polymer in a fixed conformation. The SFE is calculated using a hybrid of the integral equation theory [3] (a statistical-mechanical theory for fluids) and the morphometric approach [4]. Since the integral equation theory is employed, the number of solvent molecules considered is infinitely large. In the method, a Monte Carlo (MC) simulation applied to the polymer alone is combined with the hybrid. The calculation of the SFE is finished quite rapidly even for a large polymer, and most of the time is consumed in generating the polymer conformations using the MC algorithm. Hereafter, the method is referred to as “combined method”.

When a polymer takes a more compact conformation in solvent, the excluded volume generated by the polymer (i.e., the volume of the space which the centers of solvent molecules cannot enter) decreases, leading to increases in the total volume available to the translational displacement of the coexisting solvent molecules, in the number of accessible translational configurations of the solvent, and in the configurational entropy of the solvent [1,2]. The conformational properties of a polymer are largely influenced by this entropic excluded-volume effect. In the conventional concept, only the solvent in the close vicinity of a polymer surface is considered when the solvent effects are discussed: The solvent near the surface is entropically unstable in comparison with the bulk solvent. We remark that the solvent-entropy effect emphasized in the present letter has the following characteristics in comparison with the conventionally argued one [1,2]: It reaches a far larger length scale; it is much more insensitive to the polymer-solvent interaction potential [5]; it is considerably larger [5]; and it can be taken into

account only by a molecular model for solvent. (The conventionally argued effect is also incorporated in our theoretical treatment as discussed in the later section describing the morphometric approach; see Eq. (4).)

When the entropic excluded-volume effect plays an essential role or the solvation entropy is the key quantity, due to its insensitivity to the polymer-solvent interaction potential, a polymer can be modeled as a set of fused hard spheres with the result that the morphometric approach [4] becomes particularly powerful. In the combined method for analyzing the conformational properties of a polymer, any model can be employed for solvent in which a homo- or hetero-polymer or a protein is immersed. The solvent can be under any thermodynamic condition. We have shown that protein folding [5,6], receptor-ligand binding [7], and heat [8,9], cold [10,11], and pressure [12] denaturing of proteins can be elucidated by our theories wherein the entropic excluded-volume effect is treated as the dominant factor [4].

In the present letter, our combined method is illustrated for a simple polymer chain immersed in a hard-sphere solvent. The polymer chain is modeled as connected hard spheres. The conformational properties are then determined purely by the competition of the two entropic components, the conformational entropy of the polymer chain and the configurational entropy of the solvent. The generation of polymer-chain conformations is made using the self-avoiding random walk on a cubic lattice. The SFE of the polymer chain with each conformation is calculated using a hybrid of the integral equation theory for spherical particles [3] and the morphometric approach [4].

Even for the simple polymer chain, its conformational properties in solvent are not fully understood. This is because the thorough incorporation of solvent effects is a rather difficult task. It is physically more insightful to consider somewhat simplified model systems rather than realistic systems in which as many factors as possible are incorporated. In such computer simulations previously performed, a polymer is constrained within a spherical space and the solvent effects are accounted for simply by this constraint [13], or the solvent particles are placed on sites of a lattice and allowed to move only discretely [14-16]. With these limitations, the entropic excluded-volume effect cannot suitably be taken into account. As described above, the computer simulations which explicitly incorporate a sufficiently many solvent particles suffer from a large computational burden [17,18]. In the illustration of our combined method, we calculate the conformations of the simple polymer chain stabilized and the conformational entropies in vacuum and in solvent with low and high densities.

Last, we argue that by introducing a generalized-ensemble technique [19,20] to the MC simulation part, the combined method can be applied to studies on protein conformations in aqueous solution under any thermodynamic condition. Heat, cold, and pressure denaturing of

proteins and the prediction of the native structure can be investigated using the combined method.

## 2. Method

### 2.1. Models of polymer chain and solvent

Small hard spheres with diameter  $d_S=2.8 \text{ \AA}$  form the solvent. Two reduced solvent densities,  $\rho_S d_S^3=0.35$  and  $0.70$  ( $\rho_S$  is the solvent number density), are tested. The polymer chain is modeled as connected hard spheres with diameter  $d_U=2.5d_S$ . Hereafter, the hard spheres are referred to as “unit spheres”. The chain has neither branching nor side chains. The chain is immersed in the solvent at infinite dilution. Since rigid-body models are employed, all the accessible system configurations share the same energy and the system behavior is purely entropic in origin. The number of unit spheres  $L$  tested is in the range 4-20. The conformations of the chain are generated using the self-avoiding random walk on a cubic lattice.

### 2.2. Incorporation of solvent effects through solvation free energy

We consider a system comprising a polymer and solvent molecules. In the Boltzmann factor,  $\exp\{-E/(k_B T)\}$  ( $k_B$  is the Boltzmann constant and  $T$  is the absolute temperature),  $E$  is the instant value of the potential energy for a system configuration. Mitsutake *et al.* [19] has shown that  $E$  can be replaced by  $E_C+\Delta\mu$  ( $E_C$  is the intramolecular energy and  $\Delta\mu$  is the SFE; these are defined for the polymer in a fixed conformation) when the instant value of a quantity which is of interest depends on the polymer conformation alone. In our model system,  $E_C=0$  (and  $\Delta\mu/(k_B T)= -\Delta S/k_B$  where  $\Delta S$  is the solvation entropy). Therefore, the existing probability of conformation  $i$  of the polymer,  $P_i$ , is expressed by

$$P_i = \exp\{-\Delta\mu_i/(k_B T)\} / \sum_{i=1}^N \exp\{-\Delta\mu_i/(k_B T)\}, \quad (1)$$

where  $\Delta\mu_i$  is the SFE of the polymer with conformation  $i$  and  $N$  is the total number of accessible conformations of the polymer. We note that Eq. (1) is formally exact [19].

For  $L=4-12$ , we consider all the accessible conformations and  $P_i$  is calculated using Eq. (1). For  $L=13-20$ , however, we make use of an MC simulation where a sufficiently large number of accessible conformations are generated: Firstly, a unit sphere is placed on the original point of the

cubic lattice; the second unit sphere is randomly placed at an adjacent site; the third unit sphere is randomly placed at a vacant, adjacent site. In this manner, the unit spheres are suitably placed one after another.  $P_i$  for  $L=13-20$  is then given by

$$P_i = \exp\{-\Delta\mu_i/(k_B T)\} / [(N/n) \sum_{i=1}^n \exp\{-\Delta\mu_i/(k_B T)\}], \quad (2)$$

where  $n$  is the number of conformations actually sampled ( $n < N$ ). The exact values of  $N$  are available in literature [21] and in the present study  $n$  is set at  $10^7$ .

The conformational entropy of the polymer chain, which is denoted by  $S_C$ , is a measure of the number of accessible conformations of the polymer.  $S_C$  in solvent is obtained from

$$S_C/k_B = (1/m) \sum_{j=1}^m \{-(N/n) \sum_{i=1}^n P_i \ln P_i\}, \quad (3)$$

where  $m$  is the number of trial runs. We employ the following parameter setting:  $m=1$  and  $n=N$  for  $L=4-12$  and  $m \geq 200$  and  $n=10^7$  for  $L=13-20$ . It has been verified that with these values the standard error of  $S_C/k_B$  is always smaller than 0.1.

The MC simulation is applied to the polymer chain alone for its conformational sampling, and the SFE is theoretically calculated as an ensemble-averaged quantity for the solvent configurations in equilibrium with the chain in a fixed conformation. With the aid of the morphometric approach [4], the calculation of the SFE is finished only in less than 1 millisecond on our workstation: Most of the time is consumed in the conformational sampling for the polymer chain. The method described above is referred to as “combined method”.

### 2.3. Calculation of solvation free energy

The solvation free energy (SFE) can be calculated using the three-dimensional integral equation theory (3D-IET) [6,7,22-24] coupled with the hypernetted-chain (HNC) approximation [3]. However, the calculation requires a large amount of computer memory and a long computation time. This problem can be overcome by adopting the morphometric approach [4]. The idea of the approach is to express a solvation quantity such as the SFE  $\Delta\mu$  by the linear combination of only four geometric measures of a solute molecule:

$$\Delta\mu/(k_B T) = C_1 V_{\text{ex}} + C_2 A + C_3 X + C_4 Y. \quad (4)$$

Here,  $V_{\text{ex}}$  is the excluded volume,  $A$  is the solvent-accessible surface area, and  $X$  and  $Y$  are the integrated mean and Gaussian curvatures of the accessible surface, respectively. Though in usual cases  $C_1V_{\text{ex}}$  is the principal term, the other three terms also influence  $\Delta\mu$ . In the approach, the solute shape enters  $\Delta\mu$  only via the four geometric measures. Therefore, the four coefficients ( $C_1$ – $C_4$ ) can be determined in simple geometries. They are calculated from the values of  $\Delta\mu$  for hard-sphere solutes with various diameters ( $d_B$ :  $0 \leq d_B \leq 15d_S$ ) immersed in the model solvent. The integral equation theory for spherical particles [3] with the HNC approximation [3] is employed in the calculation. The four coefficients are determined by the least square fitting applied to the following equation for hard-sphere solutes:

$$\Delta\mu/(k_B T) = C_1(4\pi R^3/3) + C_2(4\pi R^2) + C_3(4\pi R) + C_4(4\pi), \quad R = (d_B + d_S)/2, \quad (5)$$

where  $\Delta\mu/(k_B T)$  is a function of  $d_B$ . The fitting is achieved almost perfectly. The four coefficients thus determined are  $C_1 = 0.0392 \text{ \AA}^{-3}$ ,  $C_2 = -0.0137 \text{ \AA}^{-2}$ ,  $4\pi C_3 = 0.0512 \text{ \AA}^{-1}$ ,  $4\pi C_4 = 0.0289$  for  $\rho_S d_S^3 = 0.35$ ; and  $C_1 = 0.2323 \text{ \AA}^{-3}$ ,  $C_2 = -0.1468 \text{ \AA}^{-2}$ ,  $4\pi C_3 = 1.2601 \text{ \AA}^{-1}$ ,  $4\pi C_4 = -0.2930$  for  $\rho_S d_S^3 = 0.70$ .

The method thus constructed is referred to as “a hybrid of the integral equation theory and the morphometric approach”. The high reliability of the hybrid method has already been demonstrated [1,2,4]: For a model protein immersed in a simple fluid, the results from the 3D-IET can be reproduced with sufficient accuracy by the morphometric approach where the four coefficients are determined in the manner described above.

#### 2.4. Entropic excluded-volume effect

We explain the entropic excluded-volume effect using the simple polymer chain with 4 unit spheres immersed in the hard-sphere solvent (see Fig. 1). A unit sphere generates an excluded volume for the solvent particles. The accessible conformations can be characterized by the number of overlaps of the excluded volumes generated by nonadjacent unit spheres (for  $L=1-3$ , the number of overlaps is zero). For example, in Fig. 1(a) there are no overlaps while in Fig. 1(b) there is one overlap. Therefore, the total excluded volume in Fig. 1(b) is smaller than that in Fig. 1(a): The absolute value of the SFE becomes smaller for the polymer conformation in Fig. 1(b). A polymer conformation with a smaller total excluded volume is more favored in solvent. (In the simple polymer chain we consider, smaller  $V_{\text{ex}}$  also leads to smaller  $A$ .)

### 3. Results and Discussion

#### 3.1. Conformations of polymer chain stabilized in solvent

Figure 2(a) shows the relation between the polymer-chain conformations and their existing probabilities for  $L=20$  ( $L$  is the number of unit spheres). The accessible conformations are characterized by the number of overlaps of the excluded volumes generated by nonadjacent unit spheres. The relation is plotted for vacuum and for solvents with the low-density ( $\rho_S d_S^3=0.35$ ) and high density ( $\rho_S d_S^3=0.70$ ). As the number of overlaps increases, the corresponding conformations become more compact with a smaller total excluded volume. Figures 2(b)-(d) show the conformations with the highest existing probability in vacuum and in the low-density and high-density solvents, respectively. As observed in the figures, the conformations stabilized in solvent are considerably more compact than those in vacuum and the compactness is enhanced as the solvent density increases. The solvent effects thus clarified are qualitatively the same as those uncovered by computer simulations [17,18].

#### 3.2. Conformational entropy of polymer chain in solvent

Figure 3 shows  $S_C/k_B$  of the simple polymer chains with 4-20 unit spheres in vacuum and in the low-density and high-density solvents.  $L=1-3$ , for which the number of overlaps of the excluded volumes is zero, are not considered. For our model system,  $S_C/k_B$  in vacuum is given by  $\ln N$  where  $N$  is the total number of accessible conformations of the polymer.  $S_C/k_B$  in solvent is smaller than that in vacuum, and this effect becomes larger as the solvent density increases. The same can be said for the slope. Thus, the solvent effects, which are purely entropic in origin, become larger as the solvent density or  $L$  increases. The number of accessible conformations of a polymer chain is reduced by the constraint due to the solvent effects, and the reduction becomes larger as the solvent density or  $L$  increases. These results are physically reasonable.

In the plot for the high-density solvent,  $S_C/k_B$  does not always increase as the number of unit spheres  $L$  increases. For example, the polymer chains with  $L=8, 12, 16,$  and  $18$  possess smaller conformational entropies than those with  $L=7, 11, 15,$  and  $17$ , respectively. This initially surprising behavior can be understood in the following way. When the solvent density is high, the trend that a conformation with a larger number of overlaps of the excluded volumes is more favored becomes quite strong. In the most compact conformations, the polymer chains with  $L=8, 12, 16,$  and  $18$  can gain larger numbers of overlaps than those with  $L=7, 11, 15,$  and  $17$ , respectively. For instance, the chain with  $L=8$  has five overlaps in the most compact conformations, while that with



$L=7$  has only three overlaps. The chain with  $L=8$  is forced to take the most compact conformations more strongly, leading to smaller  $S_C/k_B$ . As observed in Fig. 3,  $S_C/k_B$  notably increases when  $L=8$ , 12, and 18 change to  $L=9$ , 13, and 19, respectively. This behavior can also be understood as follows. The polymer chains with  $L=8$  and 9, for instance, share the same number of overlaps in the most compact conformations. In such a case, the chain with  $L=9$ , for which the number of possible conformations is larger, is less constrained to take the most compact conformations, leading to considerably larger  $S_C/k_B$ . This subtle effect can thus be captured by our combined method.

### 3.3. Beyond self-avoiding random walk on a cubic lattice

With the self-avoiding random walk, the conformation of the simple polymer chain is varied *discretely* because the centers of unit spheres are placed on cubic-lattice sites. In a realistic model, however, the polymer is to change its conformation *continuously*. The continuous exploration of the conformational space is thus necessitated. Such exploration can be performed widely and effectively by the generalized-ensemble techniques [19,20].

To see whether our combined method can treat the realistic model or not, we consider a simple polymer chain with  $L=4$  whose conformation is varied as illustrated in Fig. 4(a): The unit sphere drawn in the broken circle continuously moves from the position with  $\theta=0^\circ$  to that with  $\theta=90^\circ$ . We calculate the SFE of the polymer chain as a function of  $\theta$  via the two routes: the 3D-IET [6,7,22-24]; and the hybrid of the integral equation theory for spherical particles [3] and the morphometric approach [4]. We compare the two results for the low-density solvent in Fig. 4(b) and for the high-density solvent in Fig. 4(c).  $\Delta F$  represents “the SFE at  $\theta=\theta$ ” minus “the SFE at  $\theta=0$ ”. As observed in Fig. 4(b), the two results are almost indistinguishable for the low-density solvent. When the solvent density is raised (see Fig. 4(c)), a discrepancy can be appreciated between the two results. However, they share almost the same qualitative behavior. The agreement is satisfactory even in a quantitative sense. In particular,  $\Delta F$  at  $\theta=90^\circ$  is accurately calculated by the hybrid method. Considering that in the hybrid method the computation time required is shorter by three to four orders of magnitude, we can conclude that it acts as a powerful tool even for a more realistic model of the polymer chain.

## 4. Perspective

### 4.1. Extension to studies on protein conformations in aqueous solution

In principle, our combined method can be extended to studies on protein conformations in aqueous solution under any thermodynamic condition. The energetic components as well as the entropic components are to be taken into account. The SFE is expressed as

$$\Delta\mu = \Delta E - T\Delta S, \quad (6)$$

where  $\Delta\mu$ ,  $\Delta E$ , and  $\Delta S$  are the SFE, solvation energy, and solvation entropy, respectively. “ $E_C + \Delta\mu$ ” ( $E_C$  is the protein intramolecular energy) then equals “ $E_C + \Delta E - T\Delta S$ ”. The quantities,  $\Delta\mu$ ,  $\Delta E$ ,  $\Delta S$ , and  $E_C$ , are defined for a protein in a fixed conformation.  $E_C$  comprises the torsion energy  $\tau$  and Coulomb plus Lennard-Jones (LJ) terms  $E_{CLJ}$ . As argued in our earlier work [25], “ $E_{CLJ} + \Delta E$ ” can be replaced by the so-called total dehydration penalty (TDP)  $\Lambda$ . Thus,

$$E_C + \Delta\mu = \tau + E_{CLJ} + \Delta E - T\Delta S = \tau + \Lambda - T\Delta S. \quad (7)$$

The existing probability of conformation  $i$  of the protein,  $P_i$ , is then written as

$$P_i = \exp\{-(\tau_i + \Lambda_i - T\Delta S_i)/(k_B T)\} / \sum_{i=1}^n \exp\{-(\tau_i + \Lambda_i - T\Delta S_i)/(k_B T)\}, \quad (8)$$

where  $n$  is the number of conformations actually sampled and the subscript  $i$  represents that the value is for conformation  $i$ .

In Ref. [25], the structures are given beforehand and their torsion energies take physically reasonable values: The difference between any two structures in terms of the torsion energy can be neglected. With the free-energy function,  $\Lambda_i - T\Delta S_i$ , we have been exceptionally successful in discriminating the native fold from a number of misfolded decoys for significantly many proteins [25]: The free-energy function has been shown to be better than any other physics-based or knowledge-based potential function in terms of the performance. In the present case, however, the structures are generated during the simulation and those with unreasonably high torsion energy should be avoided. This is why the torsion energy is explicitly incorporated in Eq. (8).

The solvation entropy  $\Delta S_i$  is calculated using a hybrid of the angle-dependent integral equation theory for molecular fluids [26-30] applied to a multipolar model of water [26,27] and

the morphometric approach [4]. The protein can be modeled as a set of fused hard spheres in calculating  $\Delta S_i$ . The TDP is judiciously calculated in the following simple manner [25]. A fully extended structure possesses the maximum number of hydrogen bonds with water molecules and no intramolecular hydrogen bonds. “ $E_{\text{CLJ}} + \Delta E$ ”, when the fully extended structure is chosen as the standard one, corresponds to the TDP occurring upon the transition to a more compact structure. Compared to the fully extended structure with  $A=0$ , in a more compact structure some donors and acceptors (e.g., N and O, respectively) are buried in the interior after the break of hydrogen bonds with water molecules (e.g.,  $\text{NH}\cdots\text{W}$ ,  $\text{CO}\cdots\text{W}$ ; W denotes a water molecule). There is no problem if the intramolecular hydrogen bonds ( $\text{CO}\cdots\text{HN}$ , etc.) are formed. However, such hydrogen bonds are not always formed, leading to the dehydration penalty. When a donor and an acceptor are buried in the interior after the break of hydrogen bonds with water molecules, if they form an intramolecular hydrogen bond, we impose no penalty. On the other hand, when a donor or an acceptor is buried with no intramolecular hydrogen bonds formed, we impose an energetic penalty. We note that  $A_i$  takes a positive value. More details are described in our earlier publication [25].

It is crucially important to widely explore the conformational space of a protein lest the simulation should be trapped in a state with a local minimum of the function. The wide exploration can effectively be achieved by employing the generalized-ensemble techniques such as the replica-exchange MC algorithm [19,20]. In any case, the number of conformations for which  $\Delta S_i$  and  $A_i$  are calculated can be huge, but the calculation of these quantities is finished only in  $\sim 0.1$  sec per conformation on our workstation.

#### *4.2. Analyses on heat-, cold-, and pressure-denatured structures of proteins*

In our earlier studies on heat [8,9], cold [10,11], and pressure [12] denaturing of a protein, a denatured state (i.e., a set of unfolded structures) was constructed as the input data on the basis of experimentally available information, and the free energy of the protein-aqueous solution system was analyzed for the two cases where the protein takes the denatured state and the native structure, respectively. The free-energy change upon structural transition from one of the two cases to the other comprised the changes in the conformational entropy, intramolecular protein energy, and SFE. Thus, the number of accessible protein conformations and the water effects were separately treated though they are strongly coupled in a real system. Significantly many original results were successfully obtained, but a more complete analysis can efficiently be performed in the following way: By combining the hybrid of the angle-dependent integral equation theory [26-30] and the morphometric approach [4] with a generalized-ensemble technique [19,20], the set of denatured structures stabilized in aqueous solution at a low or high temperature or at an elevated pressure is

obtained as the output data by accounting for the strong coupling of the number of accessible protein conformations and the water effects.

#### 4.3. Prediction of the native structure of a protein

Prediction of the native structure of a protein from its amino-acid sequence is one of the most challenging problems in modern science. In a standard method for predicting the native structure of a protein, the all-atom Lennard-Jones and Coulomb potentials are employed for the whole system comprising a protein and water molecules. However, such a full-scale computer simulation suffers unacceptably heavy computational burden. What is worse, the results obtained are strongly dependent on the force parameters used [31]. It is experimentally known that most of the proteins are not foldable in pure water. This is due to the strong water-protein electrostatic attractive interactions. With salts such as NaCl added at sufficiently high concentrations, such interactions are screened and protein folding is facilitated. Nevertheless, in the standard method, the simulations of protein folding are usually performed in pure water. In the light of such a status, we should make an effort to develop another method based on a completely different viewpoint. Our combined method is expected to provide such a method.

The structures of a protein stabilized in aqueous solution under the physiological condition can be specified in our combined method. The problem of uncertain force-field parameters mentioned above is prudently avoided in our methods for calculating  $\Delta S_i$  and  $A_i$ . The salt effects are implicitly taken into consideration through the evaluation of  $A_i$  (i.e., with the reduced TDP [25]). It is worthwhile to test our combined method for predicting the native structure of a protein. As discussed above, the full-scale computer simulation is not always superior to a method using *logically* simplified models.

### 5. Concluding Remarks

We have proposed an efficient method for investigating conformational properties of polymers including biopolymers such as proteins in solvent. The method is a combination of a Monte Carlo (MC) simulation applied to the polymer alone and a hybrid of the integral equation theory [3] and the morphometric approach [4] for incorporating solvent effects through the solvation free energy (SFE). The calculation of the SFE is finished quite rapidly even for a large polymer, and most of the time is consumed in generating the polymer conformations using the MC algorithm. To illustrate the combined method, we have analyzed conformations of a simple

polymer chain stabilized in a hard-sphere solvent. The generation of polymer conformations is made using the self-avoiding random walk on a cubic lattice. The entropic excluded-volume effect has been shown to be crucially important. This result is consistent with the previous suggestions based on the improvement of the reference interaction site model (RISM) theories [32,33].

A most interesting target of the combined method is an analysis on protein conformations in aqueous solution under any thermodynamic condition. Since the solvent is water, the integral equation theory used is the angle-dependent version [26-30] applied to a multipolar model of water [26,27]. Solvent effects are theoretically incorporated through the solvation entropy and total dehydration penalty (TDP) (see Eq. (8)). The generalized-ensemble technique such as the replica-exchange MC algorithm [19,20] is employed for exploring the conformational space widely and effectively. There is no need to apply the MC algorithm to the solvent. The calculation of the solvation entropy and TDP can be finished only in  $\sim 0.1$  sec on our workstation. Another great advantage is that the number of replicas required is largely decreased in the combined method [19]. Thus, the computational burden is expected to be reduced to a drastic extent. We intend to revisit heat, cold, and pressure denaturations of proteins after establishing the method of estimating the TDP under any thermodynamic condition. Further, it is worthwhile to test our combined method for predicting the native structure of a protein. Works in these directions are in progress in our group.

## **Acknowledgment**

We thank T. Yoshidome for checking the results presented in Fig. 4 which can be done using the computer program for the morphometric approach developed by R. Roth, Y. Harano, and M. Kinoshita [4]. This work was supported by Grants-in-Aid for Scientific Research on Innovative Areas (No. 20118004) and that on (B) (No. 22300100) from the Ministry of Education, Culture, Sports, Science and Technology of Japan, by the Grand Challenges in Next-Generation Integrated Nanoscience, MEXT, Japan, by Kyoto University Global Center of Excellence (GCOE) of Energy Science, and by Kyoto University Pioneering Research Unit for Next Generation.

## **References**

- [1] M. Kinoshita, *Front. Biosci.* 14 (2009) 3419.
- [2] M. Kinoshita, *Int. J. Mol. Sci.* 10 (2009) 1064.

- [3] J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, 3rd ed. (Academic, London, 2006).
- [4] R. Roth, Y. Harano, and M. Kinoshita, *Phys. Rev. Lett.* 97 (2006) 078101.
- [5] T. Yoshidome, M. Kinoshita, S. Hirota, N. Baden, and M. Terazima, *J. Chem. Phys.* 128 (2008) 225104.
- [6] Y. Harano and M. Kinoshita, *Biophys. J.* 89 (2005) 2701.
- [7] M. Kinoshita, *J. Chem. Phys.* 116 (2002) 3493.
- [8] K. Amano, T. Yoshidome, Y. Harano, K. Oda, and M. Kinoshita, *Chem. Phys. Lett.* 474 (2009) 190. In this article, the entropic excluded-volume effect is treated as the dominant factor.
- [9] K. Oda, R. Kodama, T. Yoshidome, M. Yamanaka, Y. Sambongi, and M. Kinoshita, *J. Chem. Phys.*, in press.
- [10] T. Yoshidome and M. Kinoshita, *Phys. Rev. E* 79 (2009) 030905(R).
- [11] H. Oshima, T. Yoshidome, K. Amano, and M. Kinoshita, *J. Chem. Phys.* 131 (2009) 205102. When a protein takes a more compact conformation, for example, the energy lowering occurs due to the formation of intramolecular hydrogen bonds and van der Waals attractive interactions among the protein atoms. However, the dehydration penalty is always accompanied. The penalty comprises the break of protein-water hydrogen bonds and the loss of protein-water van der Waals attractive interactions. In this article, the energy lowering and the dehydration penalty are assumed to be compensating.
- [12] Y. Harano, T. Yoshidome, and M. Kinoshita, *J. Chem. Phys.* 129 (2008) 145103. In this article, the entropic excluded-volume effect is treated as the dominant factor.
- [13] M. Luo and J. Xu, *Eur. Polym. J.* 33 (1997) 1395.
- [14] M. Dijkstra and D. Frenkel, *Phys. Rev. Lett.* 72 (1994) 298.
- [15] G. Luna-Bàrcenas, G. E. Bennett, I. C. Sanchez, and K. P. Johnston, *J. Chem. Phys.* 104 (1996) 9971.
- [16] M. Gaudreault and J. Viñals, *Phys. Rev. E* 80 (2009) 021916.
- [17] R. Chang and A. Yethiraj, *J. Chem. Phys.* 114 (2001) 7688.
- [18] J. M. Polson and M. J. Zuckermann, *J. Chem. Phys.* 116 (2002) 7244.
- [19] A. Mitsutake, M. Kinoshita, Y. Okamoto, and F. Hirata, *J. Phys. Chem. B* 108 (2004) 19002.
- [20] A. Mitsutake, Y. Sugita, and Y. Okamoto, *Biopolymers (Peptide Science)* 60 (2001) 96.
- [21] N. Clisby, R. Liang, and G. Slade, *J. Phys. A: Math. Theor.* 40 (2007) 10973.
- [22] M. Ikeguchi and J. Doi, *J. Chem. Phys.* 103 (1995) 5011.
- [23] K. Amano and M. Kinoshita, *Chem. Phys. Lett.* 488 (2010) 1.
- [24] K. Amano, T. Yoshidome, M. Iwaki, M. Suzuki, and M. Kinoshita, *J. Chem. Phys.* 133

- (2010) 045103.
- [25] T. Yoshidome, K. Oda, Y. Harano, R. Roth, Y. Sugita, M. Ikeguchi, and M. Kinoshita, *Proteins* 77 (2009) 950.
- [26] P. G. Kusalik and G. N. Patey, *J. Chem. Phys.* 88 (1988) 7715.
- [27] P. G. Kusalik and G. N. Patey, *Mol. Phys.* 65 (1988) 1105.
- [28] M. Kinoshita and M. Harada, *Mol. Phys.* 81 (1994) 1473.
- [29] M. Kinoshita and D. R. Bérard, *J. Comput. Phys.* 124 (1996) 230.
- [30] M. Kinoshita, *J. Chem. Phys.* 128 (2008) 024507.
- [31] T. Yoda, Y. Sugita, and Y. Okamoto, *Chem. Phys. Lett.* 386 (2004) 460.
- [32] D. S. Palmer, V. P. Sergiievskiy, F. Jensen, and M. V. Fedorov, *J. Chem. Phys.* 133 (2010) 044104.
- [33] D. S. Palmer, A. I. Frolov, E. L. Ratkova, and M. V. Fedorov, *J. Phys.: Condens. Matter* 22 (2010) 492101.

## Figure Captions

Fig. 1. A polymer chain with 4 unit spheres immersed in solvent. Open circles: solvent particles. Closed circles: unit spheres of the polymer chain. White line: backbone. Space drawn in black and gray: Excluded space which the centers of the solvent particles cannot enter.

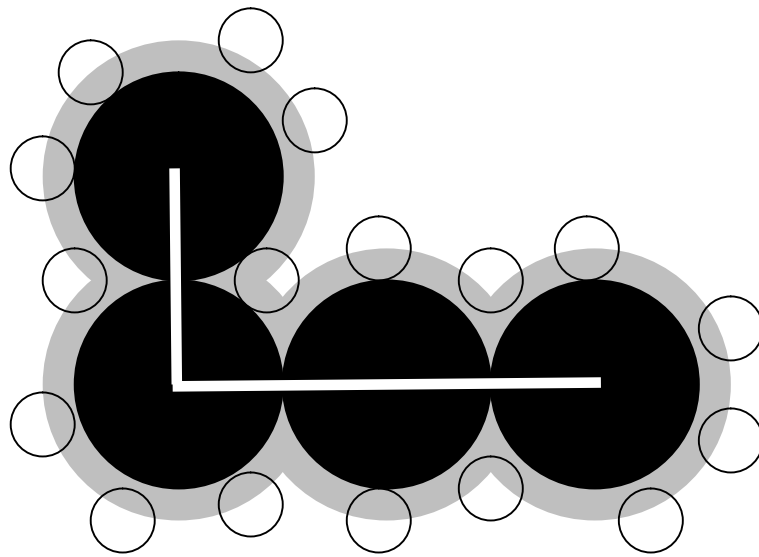
Fig. 2. (a) Relation between polymer-chain conformations and their existing probabilities. The simple polymer chain with 20 unit spheres is treated. The conformations are characterized by the number of overlaps of the excluded volumes generated by nonadjacent unit spheres. The relation is plotted for vacuum (open circles) and for solvents with the low-density ( $\rho_S d_S^3=0.35$ ; closed triangles) and high density ( $\rho_S d_S^3=0.70$ ; closed circles). (b)-(d) Conformations with the highest existing probability in vacuum (b) and in the low-density (c) and high-density (d) solvents.

Fig. 3. Conformational entropy normalized by the Boltzmann constant,  $S_C/k_B$ , plotted against the number of unit spheres  $L$ . The simple polymer chains with 4-20 unit spheres in vacuum (open circles) and in the low-density ( $\rho_S d_S^3=0.35$ ; closed triangles) and high-density ( $\rho_S d_S^3=0.70$ ; closed circles) solvents are considered.

Fig. 4. (a) Illustration of conformational variation of a simple polymer chain with 4 unit spheres. The unit sphere drawn in the broken circle continuously moves from the position with  $\theta=0^\circ$  to that with  $\theta=90^\circ$ . (b), (c) Solvation free energy (SFE) of the polymer chain as a function of  $\theta$ .  $\Delta F$  represents “SFE at  $\theta=\theta$ ” minus “SFE at  $\theta=0$ ”. The calculation is made using the three-dimensional integral equation theory (closed circles) or the hybrid of the integral equation theory for spherical particles and the morphometric approach (solid curves). The solvent density  $\rho_S d_S^3$  is 0.35 in (b) and 0.70 in (c).



(a)



(b)

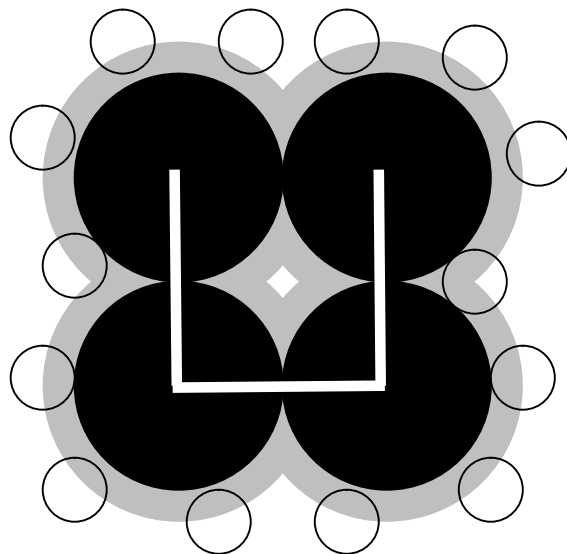
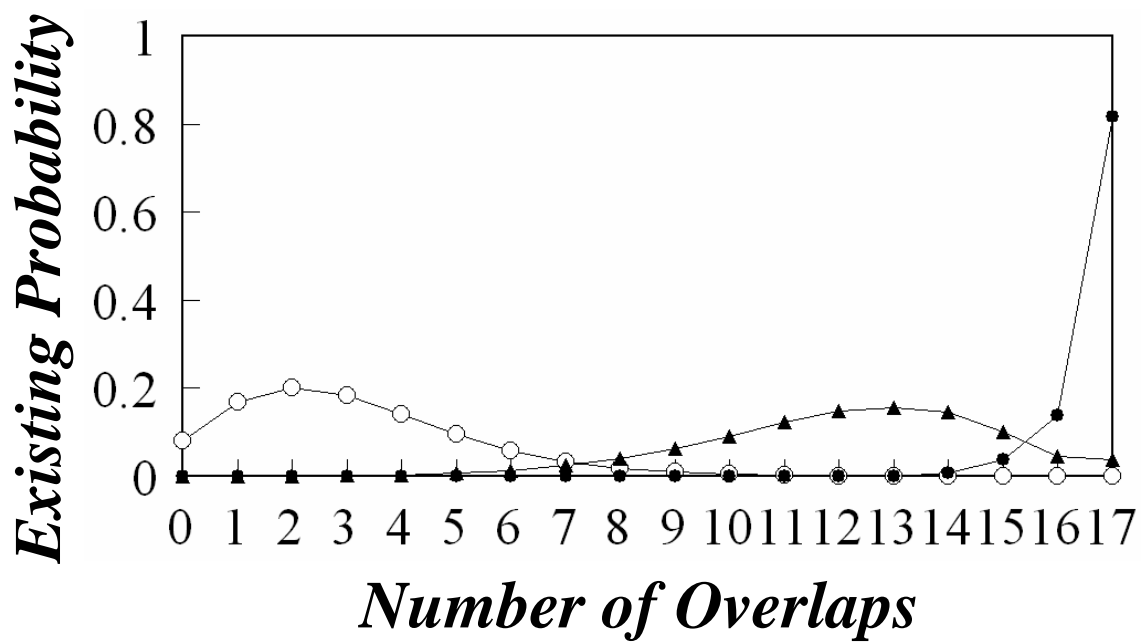
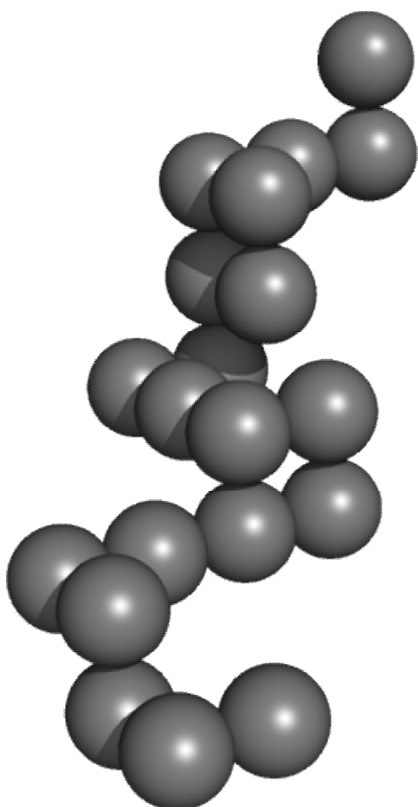


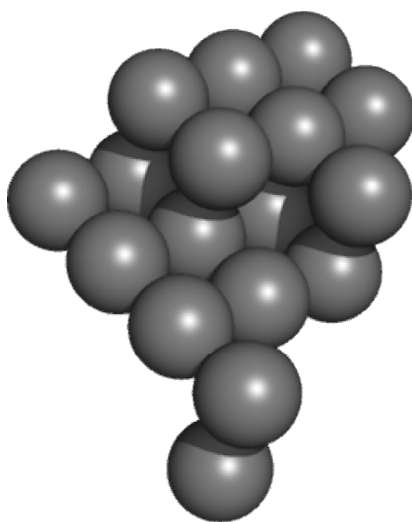
Fig. 1



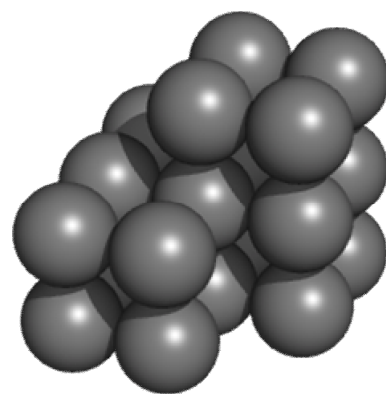
(a)



(b)



(c)



(d)

Fig. 2

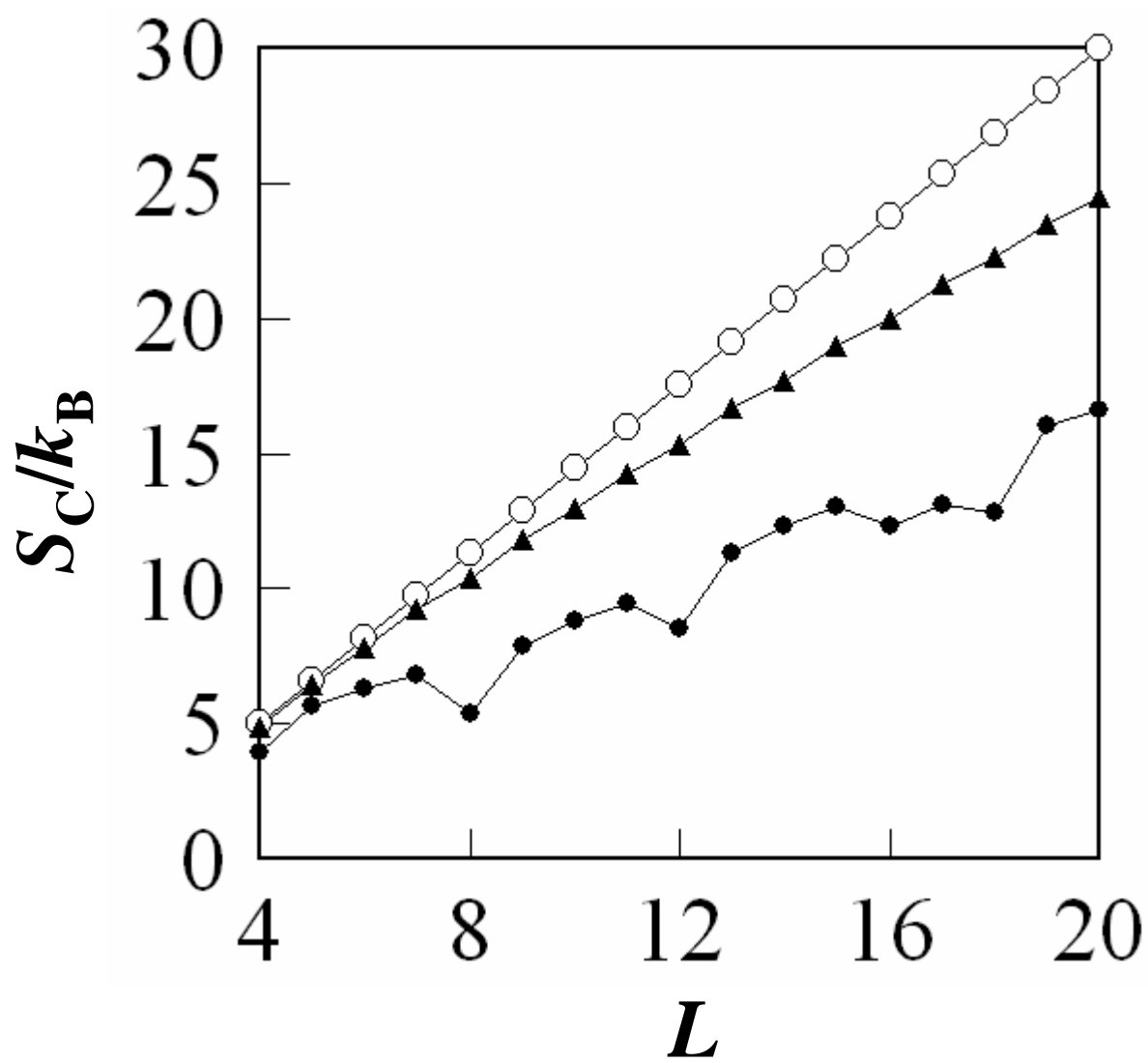


Fig. 3

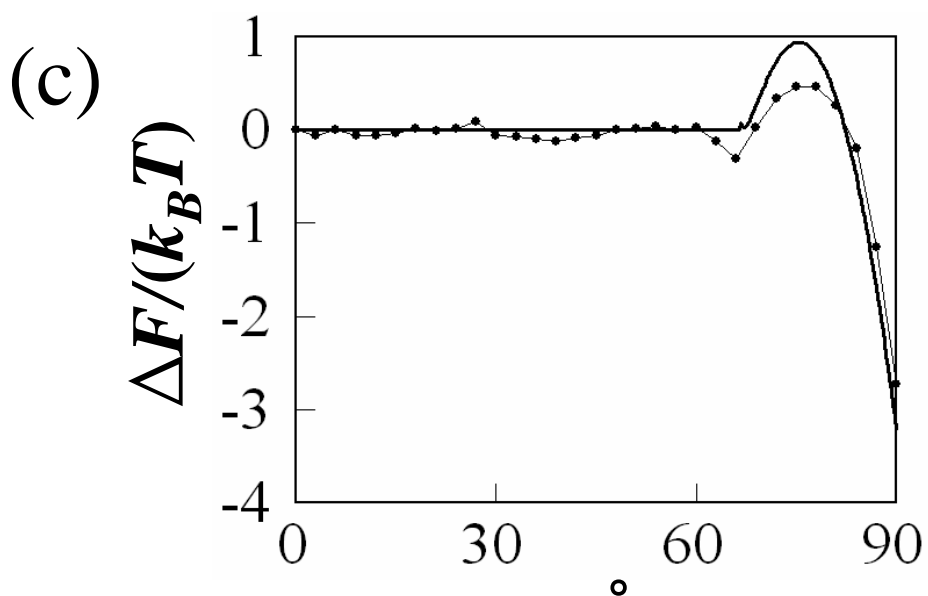
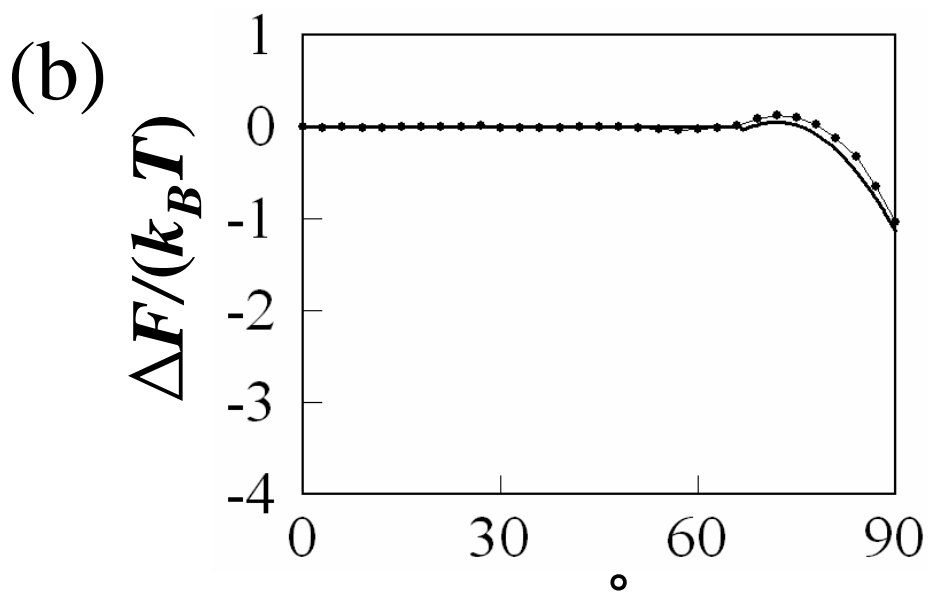
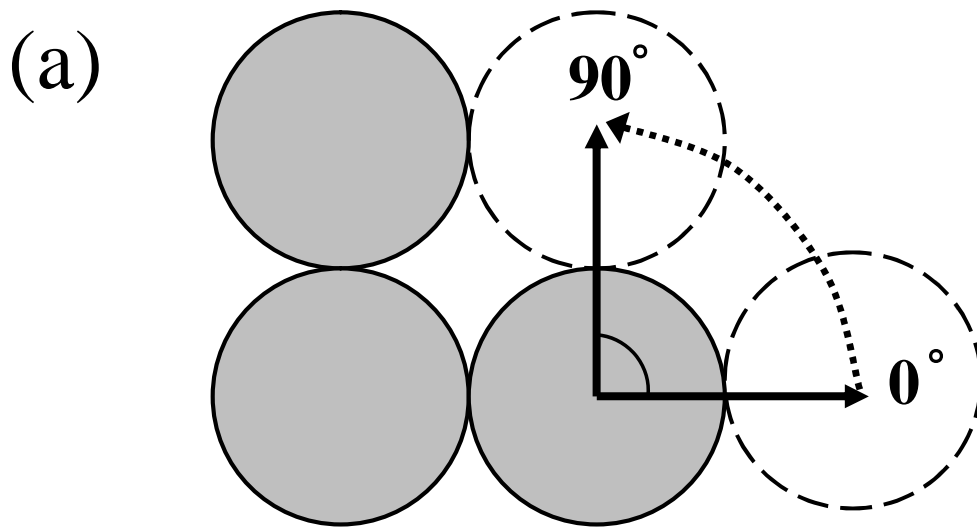


Fig. 4