Multi-body effects on molecular recognition regulated by translational motion of solvent particles

松尾, 美香

A Thesis Submitted to Kyushu University for the Degree of Doctor of Science

Multi-body effects on molecular recognition regulated by translational motion of solvent particles

Mika Matsuo

Department of Chemistry,
Graduate School of Science, Kyushu University, Japan

January 2024

Abstract

Theoretical studies were carried out to investigate effective interactions between a guest and a host in molecular recognition phenomena in liquids. When the degrees of freedom of solvent particles are ignored, the recognition, namely the association between a host and a guest, could reduce the system entropy. However, we now address the idea of entropic attraction driven by the translational motion of solvent particles existing in a system. From the viewpoint of the van der Waals picture, we consider our approach essential, especially in molecular recognition phenomena in nonpolar solvents. Thermodynamic experiments for molecular recognition suggest the contribution of entropy gain. I think that the number of experiments cannot be negligible.

First, a theoretical analysis of molecular recognition by model cyclic molecules, such as cyclodextrins, was carried out using integral equation theory. Because we are dealing with a nonspherical solute system here, such as molecular recognition, a theory that can be described on a 3D grid we adopted. Some approximations were made, the accuracies of which have not yet been examined in the case of nonspherical solutes. Thus, an assessment studies were carried out to determine an adequate theory for molecular recognition, using a dimer of two spheres in contact as the solute. Three approximations were employed: PY, HNC, and MHNC closures. The theory's accuracy was assessed by comparing it with exact results given by the grant canonical Monte Carlo (GCMC) simulations. The comparison indicated that the MHNC results were in almost full agreement with the GCMC results, more so than with PY and HNC. Analyses of the triplet distribution function were also performed. In terms of the triplet distribution function, the MHNC theory did not necessarily have an improvement in the triplet correlation.

In a further study, calculations were carried out to determine the potential of mean force (PMF) between a cyclic host molecule and a guest molecule that fits into the host cavity. The MHNC theory was adopted because the GCMC calculation was hard for the calculation for multicomponent systems, and the accuracy of the MHNC was noted to be the best in the above test calculations. I observed entropic stabilization at the recognition conformation and a free energy barrier in a multicomponent system. Compared with the PMF for the one-component system, that for a multicomponent system was less oscillatory, and the free energy barrier was significantly reduced. This difference is caused by the interference of the density wave of solvents, caused by the solute molecules. On the other hand, the free energy stabilization in the multicomponent systems was sufficiently large. The stability of multicomponent systems is, however, slightly lower than that of one-component systems. These results indicated that, in a multicomponent system, the free energy barrier can be overcome with less energy, and stability almost identical to that in one-component systems can be achieved.

Contents

Chapter 1

General introduction

The aim of this study is to describe molecular associations between macromolecules in a solution, such as protein–protein associations. Such molecular recognition phenomena have been discussed in biophysics and chemical physics [1,2]. The phenomena are now discussed from a unique aspect, based on a liquid theory, the so-called "van der Waals picture [3–8]."

In this study, we shed light on the granularity of the solvent in the association process in liquid phases. Most researchers, especially in the field of biochemistry, have assumed that the dominant driving force for association is direct interactions between associated molecules [9–12]. If the solvent is considered an inert background, then it could be a reasonable premise. In other words, the direct interactions between host and guest molecules could lead to the stabilization of complex formation and compete with the conformational entropy loss in the assumption. However, the above assumption is often in conflict with thermodynamic experiments. For example, it has been observed, experimentally, that both enthalpy and entropy changes in molecular recognition of the HIV protease are positive, that is, entropy driven [13, 14]. Thus, these examples contradict the assumption, meaning that in molecular recognition studies in the liquid phase it is important to understand the total system from a thermodynamic aspect. The entropy factor implies, inter alia, the need to observe the translational motion of solvent particles.

The structure of this chapter is as follows. In Section 1.1, the current interpretations of molecular recognition and the concepts of this study are explained. Some experimental results are also shown in this section. In Section 1.2, the properties of

liquids are explained based on the van der Waals equation of state [7]. The thermodynamic theory reveals the importance of the repulsive part of the interaction between molecules. It gives us the van der Waals picture of liquid structure, which explains that "the structure of liquid is determined primarily by the repulsive part of the two-body potential [3–5]." Therefore, attention will here be given to the repulsive part of the direct interaction between particles, and an explanation is provided for the idea of entropy-driven attraction or ordering in a solution phase in Section 1.3. In Section 1.4, the aim of this study is presented in detail.

1.1 Molecular recognition in liquids

The goal of this study was to clarify the molecular recognition mechanism. Molecular recognition refers to a process in which a host molecule recognizes a specific guest molecule to form a complex. Molecular recognition is usually found in a liquid phase that consists of solvent, host, and guest molecules. It is well known that the molecular recognition phenomenon plays important roles in various biological systems, such as enzyme–substrate reactions [2]. Many studies have focused only on host and guest molecules. However, molecular recognition in the liquid phase must be understood to understand the whole system [9–12]. Thermodynamics is the most powerful tool for discussing molecular recognition in a liquid system.

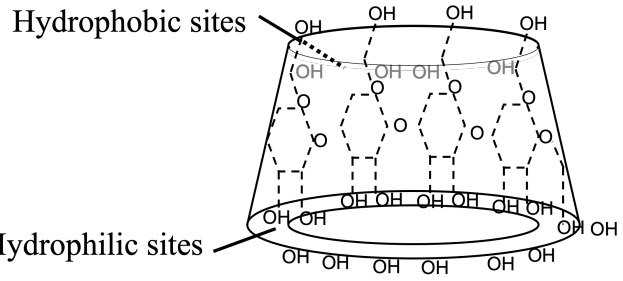
Here, we introduce some experimental results for molecular recognition between HIV protease and indinavir. The free-energy differences between the initial separation and the final recognized states were obtained experimentally. The differences were divided into enthalpy and entropy contributions (Table 1.1 [13, 14]). In the case of indinavir here, the enthalpy change was positive, and the molecular recognition

Table 1.1: Binding affinity between indinavir and HIV protease [13, 14]

	ΔG(kcal/mol)	ΔH(kcal/mol)	$-T\Delta S$ (kcal/mol)
MDR-HM	-9.3	8.4	-17.7
MDR-QM	-10.4	6.4	-16.8
V82A/I8A	-10.8	3.7	-14.5
M46I/I54V	-12.2	3.5	-15.7
L10I/L90M	-11.8	3.0	-14.8

Table 1.2: Chemical Structure and Properties of Cyclodextrins (CDs) [19–22]

	α-CD	β-CD	γ-CD
Molecular Weight	972	1135	1297
No. of Glucose Units	5	7	8
Cavity Diameter(nm)	0.47	0.60	0.75
Height of Torus(nm)	0.79	0.79	0.79



was driven by entropy gain. This example is not an exception. Entropy-driven molecular recognition has been reported experimentally; therefore, discussion of molecular recognition only in terms of direct attraction between host and guest molecules would be inappropriate. Furthermore, most of the explanations that are based only on direct attraction are contradicted by thermodynamic experimental results if they exist. In particular, the entropy factor points to the importance of observing the translational motion of solvent particles [1, 15–18].

In artificial systems, molecular recognition has also been applied to the construction of many molecular systems. To elucidate the molecular recognition mechanism, the present study focused on molecular recognition by cyclic molecules such as cyclodextrins (CDs). CDs are a family of macrocyclic oligosaccharides, the most common of which are composed of 6(α), 7(β), or 8(γ) α-1,4-linked D-glucopyranose units. The properties of typical CDs are listed in Table 1.2 [19–22]. CDs form complexes in aqueous solutions by the recognition of guest molecules with a size that fits the CD cavity [23]. We can apply the recognition by the CDs to construct various unique molecular systems, such as self-healing materials and slide-ring gels [24, 25]. Therefore, understanding molecular recognition by cyclic molecules provides fundamental knowledge of molecular recognition as well as methods for designing new molecular systems using the function.

The formation of the inclusion complex occurs despite a loss of the conformational

entropy of the guest and host molecules. The current interpretation is as follows: enthalpy gain from direct attraction between CD and guest molecules, such as van der Waals interaction and hydrophobic interactions, competes with the conformational entropy loss [26]. However, this picture has only the CD and guest molecules and cannot explain the mechanism of entropy-driven molecular recognition shown above. By contrast, if we adopt the granularity of solvent molecules, we can explain it due to the increase in translational entropy of the solvent molecules with the formation of the host–guest complex. It is a natural explanation that this entropy gain leads to an effective interaction between the host–guest molecules and drives molecular recognition. The stability of molecular recognition was investigated here based on the translational motion of the solvent particles. In the next section, the van der Waals picture of liquid structure is presented with the aim of clarifying the role of solvents in the association process.

1.2 Roles of repulsive force in liquids

Finding an ideal model system for liquids is more difficult than for gases or solids. Clearly, an ideal gas is an excellent ideal system for gases [27, 28]. The atoms in a solid retain their individual sites; basically, they do not exchange positions. When the atoms in a solid are connected by harmonic oscillators, the total partition function can be expressed as the product of the partition functions for each mode [27,28]. Thus, the system can be expressed as an ideal system because of the factorization. On the other hand, in a liquid, the particles exchange positions with each other, although the interactions are strong. Thus many particles interact with each other at any time. In other words, multibody interactions are important in a liquid system, thus making it difficult to find an adequate ideal system. Difficulties are encountered in the analysis of liquids. One of the earliest approaches to express the liquid phase was proposed by van der Waals [7]. Therein, the basic idea was that both the gaseous state and the liquid state could be described by the same equation of state, incorporating the effects of intermolecular forces into the equation of state for an ideal gas. The effect of repulsion reduces the effective volume, given the finite size of the molecule (Nb), and the effect of intermolecular attraction is simply treated as a

certain amount of cohesive energy ($-Na/V^2$). Under this assumption, van der Waals offered his equation of state:

$$\left(P + \frac{Na}{V^2}\right)(V - Nb) = kT, \tag{1.1}$$

where P is pressure, V is volume, N is the number of molecules, k is the Boltzmann constant, and T is absolute temperature. This showed that gas–liquid transition is caused by intermolecular attraction, which must be below the critical temperature. The repulsive part, namely the individual particle volume, is essential to express the liquid phase. The equation of state can then be rewritten as

$$P = \frac{\rho kT}{1 - b\rho} - a\rho^2, \tag{1.2}$$

where $\rho(= N/V)$ is the number density of the particle, and the first term means the particle pressure. The first term is the exact pressure of the 1D hard-sphere fluid (or hard-rod fluid) given by statistical mechanics. Surprisingly, van der Waals reported the first term in 1873 without statistical mechanics. If the first term was expressed using the polynomial $P(\rho)$, then the liquid–gas phase transition could not be expressed in his theory. This description indicates that the hard-sphere system is important as a reference system to analyze liquid behaviors. Indeed, there are many perturbation theories with a hard-sphere reference fluid in liquid studies [27].

On the other hand, interparticle repulsion is one of the most important factors in crystallization. In colloidal dispersion systems [29–31], and in the Wigner crystal of electrons [32, 33], those ordering phenomena are driven by the repulsive interaction. In these systems, the repulsive interaction steeply varies around the center of particle mass, and the particle excludes other particles. In other words, the particle has individual volume. The particle with the individual volume can be idealized as a hard sphere, for which the pair potential (shown in Fig.1.1(a)) is given by

$$u_{HS}(r) = \begin{cases} \infty & r \leq \sigma \\ 0 & r > \sigma, \end{cases} \tag{1.3}$$

where r is the distance between centers of particles and σ is the diameter of particles. This simple model is adequate for studying phenomena driven by the repulsion

between particles. Computer simulations provide one of the most exact results for the hard-sphere model, and are sometimes referred to as "numerical experiments" because of the sampling. Such calculations have shown that the hard-sphere system shows the fluid–crystal transition despite the absence of attractive forces between particles [34–38]. The absence of attraction also means only a single fluid phase, namely the absence of the gas–liquid transition.

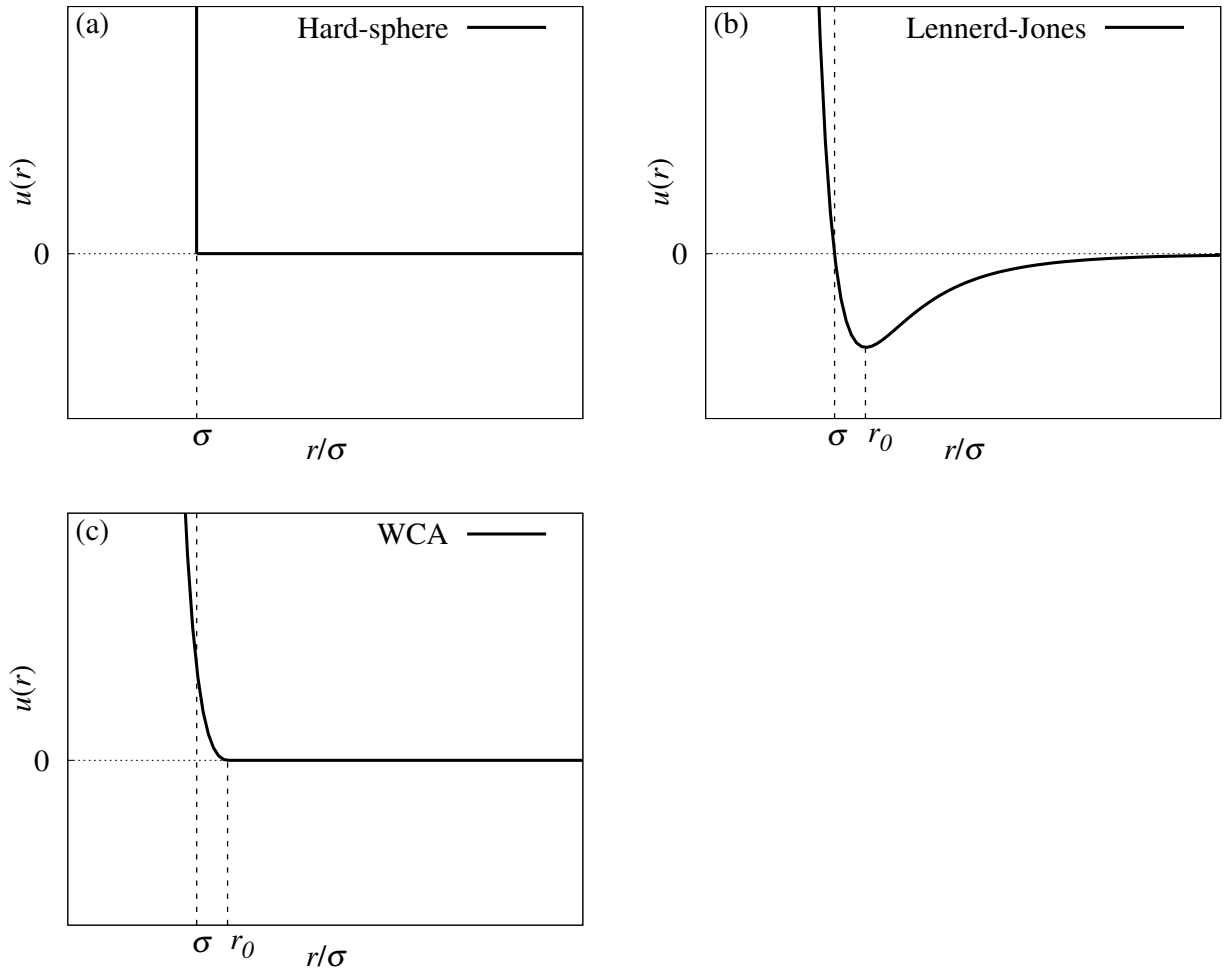


Figure 1.1: Pair interactions of (a) hard–sphere, (b) Lennerd–Jones and (c) WCA models [3–5].

Emphasizing the repulsive interaction between particles explains many properties of a liquid, as shown in Eq. (1.1). This is referred to as the van der Waals picture. For example, the liquid structure near the triple point can be obtained only using the repulsive part of the intermolecular interaction, such as the Lennard-Jones interaction, for the same temperature and the number density. It means the repulsive forces

between molecules primarily determine liquid structure. The effects of the smaller and more slowly varying attractive interaction part can be treated using a mean field approximation [3–5].

Comparing the structure of a liquid of simple molecules with that of a fluid of hard spheres is now addressed. The structure of liquids is commonly described using a radial distribution function $g(r)$ (see Section 2.1 for details). For convenience of comparison, the Lennard-Jones fluid is considered. The Lennard-Jones pair potential (shown in Fig.1.1(b)) is

$$u_{LJ}(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right], \tag{1.4}$$

where ϵ is the depth of the potential well at the minimum in $u(r)$. The properties of this fluid are known from the results of computer simulations. Then, together with suitably chosen values of the energy and length scale parameters ϵ and σ, it provides an accurate model for real atomic liquids such as argon. Fig. 1.2 is a radial distribution function $g(r)$ for the Lennard-Jones fluid at temperature $T = 0.88\epsilon/k$ and density $\rho\sigma^3 = 0.85$, which is a thermodynamic state close to the triple point. It is compared with $g(r)$ of a nonattracting system. Comparison of $g(r)$ have already been done by Chandler et al. [3–5] but here I use the results of my retest. The radial distribution function was obtained by computer simulation, by a method described in Section 2.3.1. Here, the WCA potential (shown in Fig.1.1(c)) is

$$u_{WCA}(r) = \begin{cases} u_{LJ}(r) + \epsilon & r \le r_0 \\ 0 & r > r_0, \end{cases} \tag{1.5}$$

where $r_0 = 2^{\frac{1}{6}}$ is the distance at which the potential reaches its minimum value [3–5]. The WCA potential has the same repulsive forces as the Lennard-Jones potential and no attractive forces. The figure shows that the WCA model is an excellent simpritication for the Lennard-Jones model. When the density is sufficiently high that neighboring particles are extremely close to one another, the change in energy associated with any local displacement will clearly be dominated by the interparticle repulsive forces. The attractive forces are much weaker and tend to cancel one another, leaving only an average uniform background energy. Notably, the radial distribution function of the fluid of hard spheres is almost identical to $g(r)$ of the other fluids (except

for $r = \sigma$, with impulsive potential variation). Thus, it can be seen that $g(r)$ for the Lennard-Jones liquid (a common model for argon) is accurately fitted by the radial distribution function produced by the repulsive forces or the associated hard-sphere fluid.

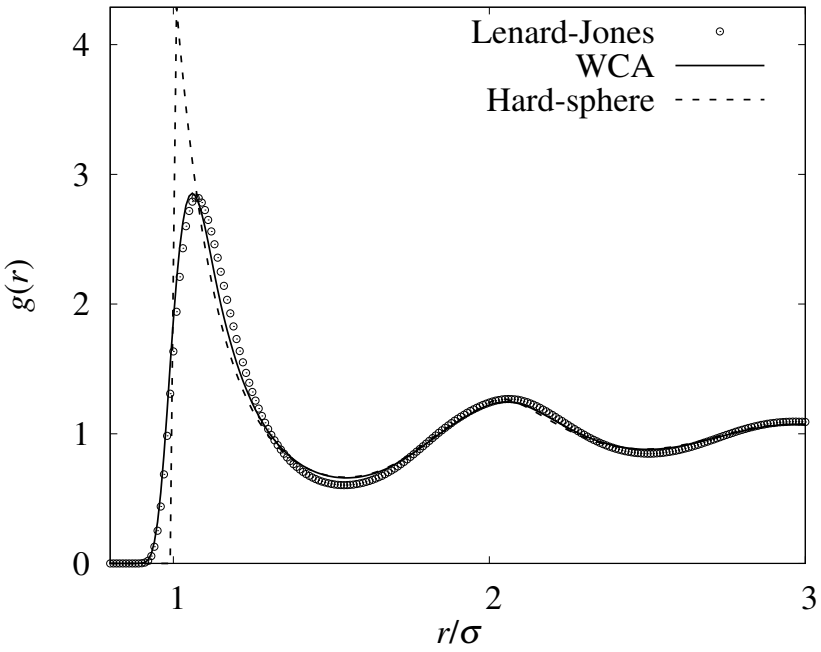


Figure 1.2: The radial distribution function, $g(r)$, for the Lennard-Jones liquid at a state near the triple point with $\rho\sigma^3 = 0.85$ and $kT/\epsilon = 0.88$. It is compared with $g(r)$ for WCA and hard-sphere fluid at the same temperature and density (My data).

Once the radial distribution function $g(r)$ is known, the effective interaction $W(r)$ can be obtained as

$$W(r) = -kT \ln g(r). \tag{1.6}$$

The effective interaction $W(r)$ calculated using $g(r)$ in Fig.1.2 is shown in Fig.1.3 [28]. The figure shows that the results from the three models are in excellent agreement. Thus, I confirmed the assertion by Chandler et al. that the structure of a liquid is determined by repulsion. This implies that repulsion even determines affinity in liquids, which I also proved by comparing effective interactions.

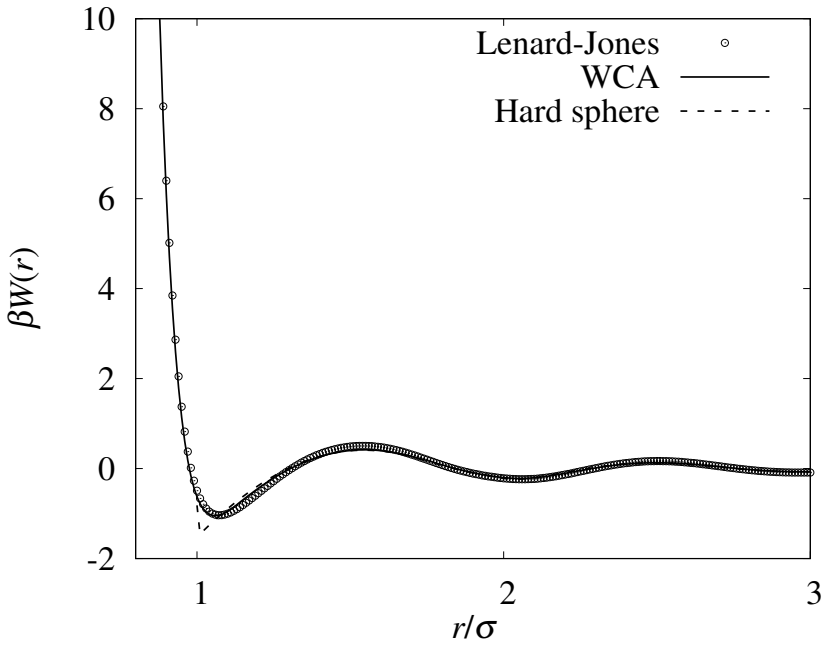


Figure 1.3: The effective interaction, $W(r)$, for the Lennard-Jones liquid at a state near the triple point with $\rho\sigma^3 = 0.85$ and $kT/\epsilon = 0.88$. It is compared with $W(r)$ for WCA and hard-sphere fluid at the same temperature and density (My data).

1.3 Entropic interaction between macromolecules in liquids

As noted, the structure of liquids is mostly determined by the strong repulsion between molecules, namely by the molecular sizes and shapes. The simplest and earliest approach is the Asakura–Oosawa (AO) theory [39–41], which has been applied to the effective interaction between macromolecules immersed in a dilute polymer solution. An explanation of the AO theory is now given. Herein, the solvent granularity is ignored, and the effective interaction arising from the translational motions of polymers is estimated only from the configurational volume of the polymers with van't Hoff's equation. Thus, the polymer–polymer interaction is ignored. Here, the polymers are the depletants. However, we cannot ignore the solvent granularity in discussing the molecular recognition phenomena in a solution phase because of the high packing fraction of depletants, namely solvent molecules. Therefore, we introduce the results calculated using more precise statistical mechanics theories. The solvent–solvent interaction affects the activation free energy for the association re-

action. The above results are the effective interaction between spherical solutes. In Section 1.3.3, the results for nonspherical solutes are introduced.

1.3.1 Asakura Oosawa theory

As shown in Fig.1.4(a), a large spherical particle (diameter σ_L) is immersed in a sea of small spherical particles (diameter σ_S). The small particles are excluded by the large particle. The excluded volume is spherical, and the diameter is $\sigma_S + \sigma_L$ (colored gray). When two large spheres contact each other, these excluded-volume regions overlap (the overlapped space is red shaded in Fig. 1.4.(b)) and the total volume accessible to the small spheres increases by this amount. Thus, the entropy of small spheres increases, resulting in an attractive interaction being induced between the large spheres between the large spheres.

In the AO theory, this effective interaction between two large spheres separated by a distance r is written as

$$W(r) = -\rho kT \Delta V_{ex}, \tag{1.7}$$

where ρ is the number density of small spheres and ΔV_{ex} is the overlapped volume. This formula was derived by Asakura and Oosawa using statistical mechanics with a simple model [39, 40]. In the model, the interaction between small particles is ignored. Therefore, the small particles behave as an ideal gas where the pressure P_{id} is ρkT. This is the most distinctive feature in the statistical theories for entropic interaction, and we can rewrite the above formula as

$$W(r) = P_{id} \Delta V_{ex}. \tag{1.8}$$

This formula suggests that the effective interaction $W(r)$ means that the quasistatic work increases as the excluded volume for the ideal gas increases, namely PV–work for the ideal gas.

1.3.2 Effects of liquid structure [42]

This study includes investigation of the ordering phenomena in a solution phase, such as molecular recognition. The AO theory can be applied to discuss the stabilization of the large molecules' association. The AO theory models the small sphere as an

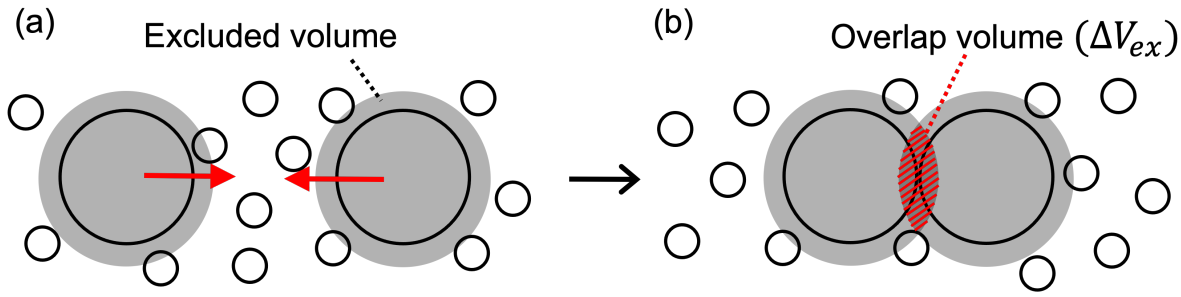


Figure 1.4: Contact between two large particles (diameter σ_L) in small particles (diameter σ_S). Spheres in gray (diameter $\sigma_S + \sigma_L$) are the excluded region for the small particles. Upon contact, the two excluded regions overlap (the overlapped space is red shaded).

ideal gas. In earlier studies carried out by Asakura and Oosawa, the small particles meant a dilute polymer solution, and the solvent particles were an inert background [39, 40]. Therefore, the ideal gas model, more precisely the ideal dilute solution model, was adequate, and the equation of state for an ideal gas, that is, van't Hoff's law for osmotic pressure, provide fruitful discussion. However, this ideal model for the solvent could be problematic in studies of the ordering phenomena in a solution phase.

It appears that the repulsive interaction between solvent particles is essential in the quantitative comparison and in the discussion of the free energy barrier. In the original study carried out by Asakura and Oosawa, the solvent particles permeate the polymers and macromolecules. Thus, we do not need to take into account the exclusion of the solvent particles. However, we need to consider the exclusion the solvent particles in the molecular recognition phenomena. When the inert solvent is replaced with a solvent consisted of granular particles, the system packing fraction becomes high similar to that of the triple point. Taking into account the short-range repulsive interaction between solvent particles, the effective interaction includes the effects caused by the liquid structure. Many studies have been examined the effective interaction between spherical solutes in liquids. According to the literature [3–5], the interaction has a free energy barrier caused by the liquid structure. The association stabilities change is caused by the pressure increase due to multibody collision [15, 43–46].

As noted in Section 1.2, in which the van der Waals picture is discussed, an appro-

priate model for the small sphere is a fluid of hard spheres [3–5]. The effective forces between the macromolecules immersed in a high-density hard-sphere solvent have been calculated using statistical mechanics theories, such as the integral equation theory (IET) [15, 44], the density functional theory (DFT) for a classical fluid [45], and computer simulations [43].

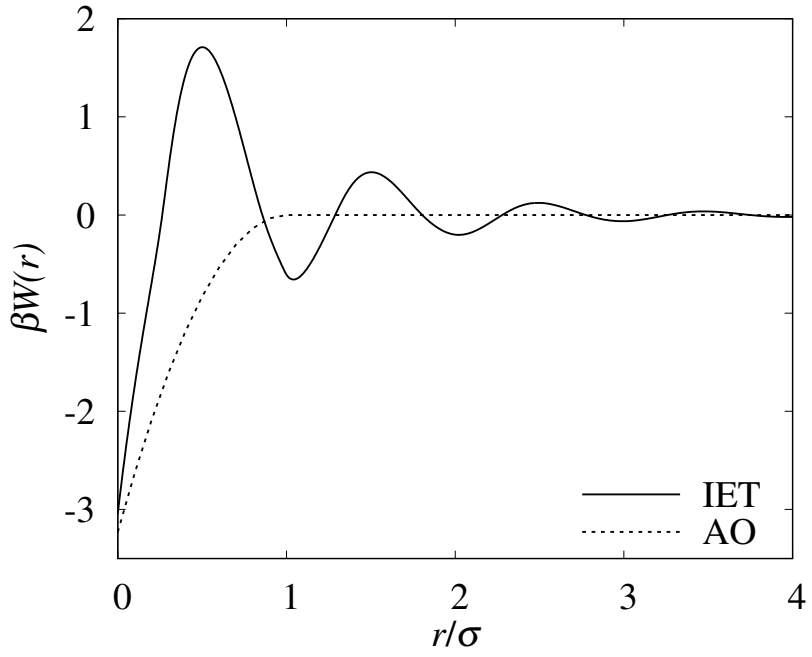


Figure 1.5: The effective forces induced between large particles (diameter $\sigma_L = 5\sigma_S$) immersed in small particles (diameter σ_S, packing fraction 0.38). Surface separation is given by r. The results calculated by AO theory and IET are compared.

Fig.1.5 shows the effective interaction between two large spheres immersed in small spheres as observed by the AO theory and IET. The results show the importance of interparticle correlations, that is, solvent structure, and bring new features to solvent-mediated interactions. The AO interaction $W(r)$ is always negative between $0 < r < \sigma_S$, and the absolute value of $W(r)$ increases monotonically with decreasing r in $0 < r$, and $W(r) = 0$ in $r > \sigma_S$.

By contrast, the interaction observed by IET is oscillatory and the period is the solvent particle diameter. Thus, the oscillation reaches beyond 1σ. This means that we may find some free energy barriers in the association process. The highest barrier is located outside the attractive well near the large particle. This is caused by

solvent particles with high density on the surface of large particles. For example, when a narrow slit – like region is formed between the large particles, as illustrated in Fig.1.6, the effective attraction between the large particles becomes stronger due to the vacuum formation. Thermodynamic work for the vacuum formation in the channel is increased until the slit size = σ_S. If the slit size becomes larger than σ_S, the small particles enter the slit. It causes the effective repulsion between the large particles by the solvent particles which enter the slit. Therefore, the free energy barrier is induced when large particles are located, as shown in Fig.1.6. When discussing the free energy barrier in the molecular recognition process, the repulsive interaction between solvent particles is essential.

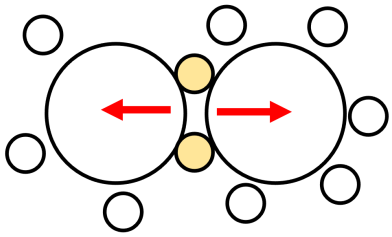


Figure 1.6: Picture illustrating the free energy barrier at $0 < r < \sigma_S$.

1.3.3 Application to molecular recognition phenomena

The discussion that now follows addresses a host molecule and a guest molecule in a sea of small particles with a hard-body model. Under certain concentrations of small particle, the recognition affinity can be estimated using the AO model [47,48], with the overlap of the excluded volumes. This idea was applied to the molecular recognition affinity by cyclic oligomers of glucose, known as cyclodextrins (CDs).

The CD molecules form inclusion complexes with the guest molecules in an aqueous solution. According to the literature [23, 26], the experimental results can be summarized as follows. It is clear that when the outer diameter of the guest molecule is larger than the inner diameter of the CD molecule, a complex cannot be formed. On the other hand, when the outer diameter of the guest is too small, the recognition is also very weak. This is a kind of "lock-and-key" relationship [2]. The best fitting, when the size of the guest molecule is equal to the guest cavity, gives stable recognition. As a result, high selectivity exists between the sizes of the outer and inner

diameters.

Returning to the AO theory, the overlap of the excluded volumes is a key amount under certain depletant concentrations. In Fig.1.7, the excluded volumes produced by the cyclic molecule and guest molecule are shown in gray. The above experimental conclusion can be explained by the key amount. The best fitting gives the maximum affinity. Therefore, using the AO theory, the recognition stability can be addressed qualitatively.

By contrast, employing adequate statistical mechanics theory with hard-body models seems essential in the quantitative comparison and the discussion of the free energy barriers. The above simple discussion that includes the overlap volume shows the importance of the idea of entropic attraction. However, the van der Waals picture and the results for the association between two large hard spheres immersed in small hard spheres indicate the need for the IET, DFT, and molecular simulations in discussion of the features of effective interaction, such as the free energy barriers [15–18]. Other calculation results reveal that the liquid structure affects the effective interaction between a key molecule and a lock molecule immersed in a high-density solvent.

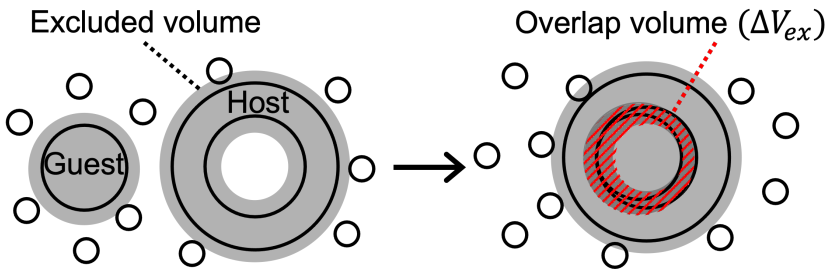


Figure 1.7: Description of molecular recognition by cyclic molecules in liquids using the AO model.

1.4 Aim of and motivation for this study

The aim of this study is to clarify the mechanism of molecular recognition in liquids by cyclic molecules such as CDs. Specifically, the purpose is to clarify the effects arising from the translational motion of solvent particles. The entropic effects are the most important factor in the affinity between molecules in a liquid; it is an aspect of the van der Waals picture [3–5]. Thus, a study of the entropic effects on effective interactions or molecular recognition affinity is essential.

For this purpose, the solvent granularity is important, and all the solvent particles must be included in the theoretical model. It is known that CD molecules form inclusion complexes with specific molecules in aqueous solution [23]. In the mainstream interpretation of the affinity between the guest and the host, the origin of attraction is the van der Waals interactions between the inner surface of the CD and the guest molecules. Moreover, the association of two solute molecules is considered to be an entropy loss, and hydrophobic effects correct the effective interaction [26]. However, the thermodynamic experiments for molecular recognition in an aqueous solution contradict the validity of the conventional narrative for molecular recognition because the association enthalpy is positive in most cases [13, 14]. The stability in the recognition arises from the entropy gain in the experiments. Interestingly, the formation of CD complexes is observed experimentally, even in organic solvents where there are no hydrophobic interactions. Unfortunately, the thermodynamic data were not found. However, these experimental results provide strong motivation for the study based on the van der Waals picture.

This thesis comprises two main chapters. The first is an assessment of the approximation of the integral equation theories for a fluid consisting of hard-body particles. (An assessment for a nonspherical dimer solute has not been carried out.) Here, the most adequate approximation is discussed. The second part addresses calculations for the recognition system using the best integral equation theory. Here, the importance of solvent granularity in the regulation of molecular recognition is discussed.

Understanding the recognition mechanism in a liquid provides not only fundamental knowledge of molecular recognition but also guidelines for the design and control

of new molecular systems that utilize the molecular recognition function. Recently, various interesting related systems have been proposed. For example, the formation of complexes with linear polymers by CDs has attracted attention. Such complexes are a molecular necklace with many CDs threaded through the polymer chain. It is expected that this will be useful to construct unique nanostructures by utilizing the degrees of freedom.

Chapter 2

Calculation

2.1 Distribution function [27, 28, 42]

For a system at equilibrium, integration of a reduced distribution function over the remaining momenta yields an equilibrium particle density, where $\rho^{(n)}(\mathbf{r}_1, \ldots, \mathbf{r}_n)$ specifies the probability that a set of n molecules will be found in a configuration $\mathbf{r}_1, \ldots, \mathbf{r}_n$. In canonical ensembles, the n-particle density is defined as:

$$\rho^{(n)}(\mathbf{r}_1, \ldots, \mathbf{r}_n) = \frac{N!}{(N-n)!} \frac{\int \ldots \int e^{-\beta U_N} d\mathbf{r}_{n+1} \ldots d\mathbf{r}_N}{Z_N} \tag{2.1}$$

, and the configuration integral Z_N is:

$$Z_N = \int \ldots \int e^{-\beta U_N} d\mathbf{r}_1 \ldots d\mathbf{r}_N \tag{2.2}$$

, where $\beta = (k_B T)^{-1}$. U_N is the total potential energy, written as a sum of a pair potential $u(r_{ij})$:

$$U_N = \sum_{i<j}^{N} u(r_{ij}) \ , \quad r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|. \tag{2.3}$$

In a homogeneous system, the n-particle distribution function $g^{(n)}(\mathbf{r}_1, \ldots, \mathbf{r}_n)$ is defined in terms of the corresponding particle density by:

$$\rho^{(n)}(\mathbf{r}_1, \ldots, \mathbf{r}_n) = \rho^n g^{(n)}(\mathbf{r}_1, \ldots, \mathbf{r}_n). \tag{2.4}$$

. This function is dimensionless and measures the extent of deviation of the structure of the fluid from complete randomness. In a liquid of spherically symmetric

molecules, $g^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$ depends only on the relative distance $r_{12} = |\mathbf{r}_2 - \mathbf{r}_1|$; it is the usually called radial distribution function and is written simply as $g(r)$. We can write:

$$\rho^{(2)}(r) = \rho^2 g(r). \tag{2.5}$$

The radial distribution function plays a central role in the theory of liquids. The first reason for this statement is that the radial distribution function is directly measurable by radiation-scattering experiments. The second reason is that knowledge of radial distribution functions is often sufficient to calculate the equation of state and other thermodynamic properties of the system.

N is very large; therefore, Eqs(2.1) and (2.2) give:

$$g(r) = \frac{V^2 \int \dots \int e^{-\beta U_N} d\mathbf{r}_3 \dots d\mathbf{r}_N}{\int \dots \int e^{-\beta U_N} d\mathbf{r}_1 \dots d\mathbf{r}_N}. \tag{2.6}$$

By contrast, when particle 1 is fixed, the force acting on particle 2 is $-dU_N/dr$, and its average is:

$$K(r) = \frac{-\int \dots \int e^{-\beta U_N} d\mathbf{r}_3 \dots d\mathbf{r}_N}{\int \dots \int e^{-\beta U_N} d\mathbf{r}_1 \dots d\mathbf{r}_N}. \tag{2.7}$$

Thus, the relationship is as follows:

$$K(r) = \frac{1}{\beta} \frac{d \ln g(r)}{dr}. \tag{2.8}$$

If the potential of mean force (PMF) is written as:

$$W(r) = -\int_{\infty}^{r} K(r) dr, \tag{2.9}$$

we can write:

$$g(r) = e^{-\beta W(r)}. \tag{2.10}$$

2.2 Integral equation theory [27, 42]

We now introduce the total correlation function: $h(r_{12}) = g(r_{12}) - 1$. This function is a measure of the total influence of molecule 1 on molecule 2 at a distance r_{12}. Ornstein and Zernike proposed that $h(r_{12})$ be written as the sum of the direct and indirect

effects. The direct effect is given by the direct correlation function $c(r_{12})$. The defining equation of the direct correlation function is called the Ornstein–Zernike (OZ) equation:

$$h(r_{12}) = c(r_{12}) + \rho \int c(r_{13})h(r_{23})d\mathbf{r}_3. \tag{2.11}$$

If we repeatedly replace $h(r)$ by [$c(r)$+integral] within the integral of Eq.(2.11), we get:

$$\begin{aligned} h(r_{12}) = c(r_{12}) + \rho \int c(r_{13})c(r_{23})d\mathbf{r}_3 \\ + \rho^2 \int\int c(r_{13})c(r_{34})c(r_{42})d\mathbf{r}_3 d\mathbf{r}_4 + \cdots . \end{aligned} \tag{2.12}$$

The total correlation $h(r_{12})$ is decomposed into a direct correlation of 1 and 2, through $c(r_{12})$, and indirectly through all possible chains of direct correlation within the fluid.

If $c(r)$ were given in terms of $h(r)$, for example, its substitution into the OZ equation would give a closed integral equation for $h(r)$. This can be done in the following approximate way.

2.2.1 The hypernetted chain (HNC) and Percuss Yevick (PY) approximations

In Boltzmann form, the radial distribution function $g(r_{12})$ may be written in terms of the PMF, thus:

$$\ln g(r_{12}) = -W(r_{12})/kT. \tag{2.13}$$

The PMF may be split up into two components, much as in the OZ equation:

$$W(r_{12}) = u(r_{12}) + \Phi(r_{12}), \tag{2.14}$$

where $u(r_{12})$ is the direct pair potential, and $\Phi(r_{12})$ is the indirect pair potential, which represents the mean effect of a third particle averaged over all possible positions.

The expression for the direct correlation in the HNC approximation is given as:

$$\begin{aligned} c_{\mathrm{HNC}}(r_{12}) &= h(r_{12}) + \frac{\Phi(r_{12})}{kT} \\ &= h(r_{12}) - \ln g(r_{12}) - \frac{u(r_{12})}{kT}. \end{aligned} \tag{2.15}$$

When inserted in the OZ equation, this expression provides a closed equation for the total correlation and radial distribution function.

Similarly, in the PY approximation

$$c_{\mathrm{PY}}(r_{12}) = \exp\left(-\frac{W(r_{12})}{kT}\right) - \exp\left(-\frac{\Phi(r_{12})}{kT}\right)$$
$$= g(r_{12})\left[1 - \exp\left(\frac{u(r_{12})}{kT}\right)\right]. \tag{2.16}$$

2.2.2 Bridge function

Now, with extension of the Eq.(2.15), we have the exact relation:

$$c(r_{12}) = h(r_{12}) - \ln g(r_{12}) - \frac{u(r_{12})}{kT} + b(r_{12}), \tag{2.17}$$

where $b(r)$ is called a bridge function. When $b(r) = 0$, Eq.(2.17) reduces to the HNC approximation. Various bridge functions were proposed to improve the theory for the correlation functions. In this study, three bridge functions, namely Verlet [49], Duh-Henderson [50], and MHNC [44, 51, 52], were used. Verlet proposed the following functional form:

$$b(r_{12}) = -0.5\frac{\gamma^2(r_{12})}{1 + 0.8\gamma(r_{12})}, \tag{2.18}$$

where $\gamma(r) = h(r) - c(r)$ [49]. The distribution function, $g(r)$, obtained using this form was shown to be quite accurate. However, Duh and Henderson argued that the form is not suitable for size-asymmetric mixtures, and they proposed the following form:

$$\begin{aligned} b(r_{12}) &= -0.5\frac{\gamma_{12}^2(r)}{1 + 0.8\gamma(r_{12})} \quad (\gamma > 0) \\ &= -0.5\gamma_{12}^2(r) \qquad\qquad (\gamma < 0). \end{aligned} \tag{2.19}$$

Eq.(3.7) and its first and second derivatives with respect to $\gamma(r)$ are also continuous at $\gamma(r) = 0$ [50]. However, Eq.(3.8) for $\gamma < 0$ gives poor results when the size asymmetry increases. Kinoshita proposed the following form, named the modified HNC (MHNC) bridge function [44, 51, 52]:

$$\begin{aligned} b(r_{12}) &= -0.5\frac{\gamma^2(r_{12})}{1 + 0.8\gamma(r_{12})} \quad (\gamma > 0) \\ &= -0.5\frac{\gamma^2(r_{12})}{1 - 0.8\gamma(r_{12})} \quad (\gamma < 0). \end{aligned} \tag{2.20}$$

2.3 Computer simulation [53, 54]

The adequacy of the approximations involved in the integral equation method can be checked by comparison with experimental results. In a comparison with experimental data from X-ray scattering, it may be difficult to distinguish whether a disagreement is caused by the approximation or by the choice of intermolecular potential. Clearly, it would be great if essentially exact results could be obtained for a given model system without having to rely on approximate theories. Computer simulation allows doing just that. We can compare the calculated properties of a model system with those of an experimental system. If they disagree, our model is inadequate, that is, estimation of the intermolecular interactions must be improved. It is also possible to compare the simulation results of a given model system with the predictions of an approximate theory applied to the same model. In this case, the computer simulation plays the role of the "experiment" to test the theory. Computer experiments have now become standard practice to test the theoretical results.

2.3.1 Monte Carlo method

Monte Carlo (MC) is a computer simulation method. First, we recall that the average of a function of the coordinates $F(\mathbf{r}_1, \ldots, \mathbf{r}_n)$ is given by:

$$\langle F \rangle = \frac{\int \ldots \int F(\mathbf{r}_1, \ldots, \mathbf{r}_n) e^{-\beta U_N} d\mathbf{r}_1 \ldots d\mathbf{r}_N}{\int \ldots \int e^{-\beta U_N} d\mathbf{r}_1 \ldots d\mathbf{r}_N}. \tag{2.21}$$

In numerical calculations, the integral is always replaced by a sum over a set of suitably chosen points. In the MC method, the points are chosen at random. In the present context, one would choose some large but manageable number is of configurations choosing them with a probability. Let M be the number of points and the resulting estimate of $\langle F \rangle$ is:

$$\langle F \rangle = \sum_{i=1}^{M} F_j e^{-\beta U_i} \; / \; \sum_{i=1}^{M} e^{-\beta U_i}. \tag{2.22}$$

In the region of liquid densities, however, this method produces unsatisfactory results because the overwhelming majority of the points would correspond to very low

statistical weights, which makes erroneous results unavoidable. Then importance sampling is usually preferred. The basic principle of this method is to allow for the statistical weight in choosing the points in such a way that the frequency of each configuration j becomes proportional to $e^{\beta U_i}$. Eq.(2.22) then reads:

$$\langle F \rangle = \frac{1}{M} \sum_{i=1}^{M} F_i. \tag{2.23}$$

The Metropolis method is one of the most popular ways of ensuring that the configuration space is sampled with the weight [55]. The transition probabilities must satisfy the normalization conditions:

$$\sum_j p_{ij} = 1. \tag{2.24}$$

Let us assume that any state j may be reached from any state i in a finite number of steps m (ergodicity condition). Then Markov's theorem states that the limits

$$\lim_{m\to\infty} p_{ij}^{(m)} = p_i \tag{2.25}$$

exist for all j and are independent of j:

$$\sum_i p_i = 1, \tag{2.26}$$

and

$$p_i = \sum_j p_j p_{ij}. \tag{2.27}$$

In fact, the limits p_i are uniquely determined by Eqs(2.25) and (2.26). However, we wish to invert the usual process; that is, we inquire what set of transition probabilities p_{ij} will converge to the previously known set of p_i:

$$p_i = e^{-\beta U_i} \; / \; \sum_i e^{-\beta U_i}. \tag{2.28}$$

Substitution of Eq.(2.28) into Eq.(2.27) gives, with use of Eq.(2.24):

$$p_{ij} e^{-\beta U_i} = p_{ji} e^{-\beta U_j}. \tag{2.29}$$

Eq.(2.29) evidently does not determine a unique set p_{ij}. Metropolis et al. proposed the transition probabilities as follows:

$$p_{ij} = \begin{cases} \alpha_{ij} & U_j < U_i \\ \alpha_{ij} e^{-\beta(U_j - U_i)} & U_j > U_i \end{cases}. \tag{2.30}$$

$$p_{ii} = 1 - \sum_{j \neq i} p_{ij}, \tag{2.31}$$

where α must satisfy the following:

$$\begin{cases} \alpha_{ij} & = \alpha_{ji} \\ \sum\limits_j \alpha_{ij} & = 1. \end{cases} \tag{2.32}$$

Chapter 3

Spatial distribution obtained by GCMC and IET with bridge functions [56]

3.1 Introduction

To predict precise distribution functions around a hard–body solute, such as a contact solvent dimer immersed in a hard–sphere fluid, we must prepare some statistical mechanics theories and assess them. In the context of the study of entropic interaction, the theories have been examined with hard-body models. Because the repulsive part of the direct interaction has an essential role in the formation of a microscopic liquid structure [3–5], the progress of the statistical theories is expected to understand the effective interaction in high-density fluids. This viewpoint has been called the van der Waals picture. In the picture, a hard-sphere system is the most fundamental model in the studies of a dense fluid. Although the precise prediction becomes more difficult as the packing fraction becomes higher, there are some precise approaches, such as density functional theories (DFT) [45, 57–59] and integral equation theories [42, 60, 61].

I carried out the theoretical calculation using some integral equation theories for a liquid to discuss the accuracy of the solvent density profiles around a contact solvent dimer. I solved the Ornstein–Zernike (OZ) equation with a closure relation [42] and examined various closure relations. In the preceding studies, the Percus–Yevick (PY) closure, the hypernetted–chain (HNC) closure, etc., have been used as the traditional closures. The accuracy of the results depends on the closure relation, namely the

approximation. It is known that the spatial distribution functions calculated using the traditional closure, such as the PY and the HNC closures, are reasonable qualitatively. However, we can find differences between them. It is known that the PY closure is better than the HNC closure for the one–component hard–sphere fluid. By contrast, the HNC closure has been adopted in the calculations of hard-sphere mixtures because it is better than the PY closure when the size ratio is not 1. My test calculation also supported the above conclusion.

The MHNC closure proposed by Kinoshita was examined in the preceding studies [44,51,52,62]. In that study, a hard sphere solute was immersed in a hard–sphere fluid whose packing fraction was 0.38, and various size solutes were examined. The calculated spatial distribution functions show that the results calculated by the MHNC-OZ theory are much more quantitative than those calculated by the PY–OZ and the HNC–OZ theories. Furthermore, the MHNC closure is accurate even with significant size asymmetry. The MHNC closure is the most expected approximation for the hard-body particle systems.

In my studies, a contact dimer of solvent particles is examined as a nonspherical solute particle. Molecular simulation studies can be applied to obtain the spatial distribution functions around the nonspherical solute particle. However, a study on the accuracy of distribution functions in the vicinity of a nonspherical solute is unusual because of the computational cost. Moreover, I will obtain the correlation function for the muticomponent solvent system with a model cyclic host molecule like cyclodextrin in the next chapter. The computational cost of molecular simulation for the systems is very high. The concentration dependence studies must be hopeless.

As mentioned above, the MHNC closure is an adequate closure for a spherical solute in a solvent hard-sphere fluid. The study on the accuracy of distribution functions in the vicinity of a nonspherical solute in a muticomponent solvent is unusual because of the expensive computational cost. However, it is possible to assess the closures using molecular simulations if more simple nonspherical solute and the one-component solvent are adopted. The assessment must give us the prospect of the closure choice.

In this study, the simulations are carried out to compare the results obtained using

the 3D integral equation theory. We adopt the contact dimer of solvent hard–sphere as a nonspherical solute. The spatial distribution function around a nonspherical solute for the coordinate (x, y, z) in the calculation using the integral equation theory must be obtained. If we take symmetry into account, we can reduce the computational cost. Therefore, I adopt the contact dimer of the solvent hard-sphere as the nonspherical solute. The grand canonical ensemble Monte Carlo (GCMC) simulations were performed to verify the 3D–MHNC–OZ theory in the present study because the integral equation theories have been formulated under the grand canonical ensemble

3.2 Model and Methods

3.2.1 Model

A fluid is composed of hard spheres. The diameter used was σ_V (V denotes the solvent particles). The scaled number density of the fluid $\rho{\sigma_V}^3$ was at 0.7315. This density of the fluid corresponds to the packing fraction $\eta = 0.383$. This value is the packing fraction of pure water at standard temperature and pressure. We immersed a solute into the hard-sphere fluid. The solute was a contact dimer that consisted of two hard spheres fixed at coordinates $(-0.5\sigma_V, 0, 0)$ and $(0.5\sigma_V, 0, 0)$. Each diameter of two hard spheres was $\sigma_U = \sigma_V$ (U denotes solute particles).

3.2.2 Integral Equation Theory [27, 42]

In the present study, we solved the OZ equation with a closure equation to obtain the spatial distribution $g_{UV}(x, y, z)$ of solvent particles around a solute. The OZ equation for the bulk solvent was

$$h_{VV}(r) = c_{VV}(r) + \rho_V \int c_{VV}(r) h_{VV}(|\mathbf{r}' - \mathbf{r}|) d\mathbf{r}', \tag{3.1}$$

where ρ is the number density, h is the total correlation function, and c is the direct correlation function. The system is spherically symmetric and r is the distance between the centers of solvent particles. At first, we solved this equation with a closure equation.

The solute particle has a nonspherical shape. Therefore, the spatial distribution functions between the solute and the solvent particles are $g_{UV}(x, y, z) = h_{UV}(x, y, z) + 1$. To solve the equations, we prepared a 3D grid covering enough volume. Then, the OZ equation is given in the discrete form as follows,

$$h_{UV}(x, y, z) = c_{UV}(x, y, z) + \rho_V \int c_{UV}(x, y, z) h_{VV}(|\mathbf{r}' - \mathbf{r}|) dx' dy' dz', \tag{3.2}$$

where $\mathbf{r} = (x, y, z)$ and $\mathbf{r}' = (x', y', z')$. Eq.(4.2) is written in the wavenumber $\mathbf{k}$ space as follows:

$$\hat{\gamma}_{UV}(k_x, k_y, k_z) = \rho_V \hat{c}_{UV}(k_x, k_y, k_z) \hat{h}_{VV}(|\mathbf{k}|), \tag{3.3}$$

where $\gamma = h - c$, the symbol "^" indicates the Fourier transform. The vector $\mathbf{k} = (k_x, k_y, k_z)$.

We examined some closures. The closure equation is written as

$$c_{ij}(\mathbf{r}) = \exp[-\beta u_{ij}(\mathbf{r})] \exp[\gamma_{ij}(\mathbf{r}) + b_{ij}(\mathbf{r})] - \gamma_{ij}(\mathbf{r}) - 1, \tag{3.4}$$

where $\beta = (k_B T)^{-1}$, $k_B T$ is Boltzmann constant times the absolute temperature. The functions u and b are the potential and the bridge function, respectively. In the calculation of bulk solvent, $i = j = V$ and $\mathbf{r}$ is replaced with r. By contrast, $i = U$, $j = V$ and $\mathbf{r} = (x, y, z)$ in the calculation of the solute–solvent correlation functions.

Eq.(4.4) includes the bridge function. If we had the perfect bridge function, the closure would also be perfect. However, the exact form of the bridge function was yet unknown. In the case of the HNC approximation,

$$b_{ij}(\mathbf{r}) = 0. \tag{3.5}$$

In the present study, we examined the MHNC closure proposed by Kinoshita [44,51, 62]. The bridge function is as follows.

$$\begin{aligned} b_{ij}(\mathbf{r}) &= -0.5 \frac{\gamma_{ij}^2(\mathbf{r})}{1 + 0.8\gamma_{ij}(\mathbf{r})} \quad (\gamma_{ij} > 0) \\ &= -0.5 \frac{\gamma_{ij}^2(\mathbf{r})}{1 - 0.8\gamma_{ij}(\mathbf{r})} \quad (\gamma_{ij} < 0). \end{aligned} \tag{3.6}$$

We also examined the following two bridge functions. The bridge function proposed by Verlet is as follows [49].

$$b_{ij}(\mathbf{r}) = -0.5 \frac{\gamma_{ij}^2(\mathbf{r})}{1 + 0.8\gamma_{ij}(\mathbf{r})}. \tag{3.7}$$

Another bridge function, proposed by Duh and Henderson, is as follows [50].

$$\begin{aligned} b_{ij}(\mathbf{r}) &= -0.5\frac{\gamma_{ij}^2(\mathbf{r})}{1+0.8\gamma_{ij}(\mathbf{r})} \quad (\gamma_{ij} > 0) \\ &= -0.5\gamma_{ij}^2(\mathbf{r}) \qquad\qquad (\gamma_{ij} < 0). \end{aligned} \tag{3.8}$$

We also examined the PY closure:

$$c_{ij}(\mathbf{r}) = \exp[-\beta u_{ij}(\mathbf{r})][\gamma_{ij}(\mathbf{r}) + 1] - \gamma_{ij}(\mathbf{r}) - 1. \tag{3.9}$$

It is known that the PY closure is excellent for the one-component hard-sphere fluid. The spatial distribution of solvents $g_{UV}(x, y, z)$ is obtained using the calculated $\gamma_{UV}(x, y, z)$ and $c_{UV}(x, y, z)$ as follows.

$$g_{UV}(x, y, z) = \gamma_{UV}(x, y, z) + c_{UV}(x, y, z) + 1. \tag{3.10}$$

The numerical procedures were as follows: (a) $h_{VV}(r)$ was calculated using Eq.(4.1) and one of the closures, (b) $h_{VV}(x, y, z)$ was prepared using $h_{VV}(r)$ and transformed to $\hat{h}_{VV}(k_x, k_y, k_z)$ using the 3D fast Fourier transform (3D-FFT), (c) $u_{UV}(x, y, z)$ was calculated at each 3D grid point and $\gamma_{UV}(x, y, z)$ was initialized to zero, (d) $c_{UV}(x, y, z)$ was calculated using Eq.(4.2) and the same closure adopted in step (a), (e) $c_{UV}(x, y, z)$ was transformed to $\hat{c}_{UV}(x, y, z)$ using the 3D-FFT, (f) $\hat{\gamma}_{UV}(x, y, z)$ was calculated using Eq.(4.3), (g) $\hat{\gamma}_{UV}(x, y, z)$ was transformed to $\gamma_{UV}(x, y, z)$ using the inversed 3D-FFT and steps (d)–(g) were repeated until the difference between the input and output functions become smaller than the given value, (h) $g_{UV}(x, y, z)$ were calculated using Eq.(3.10). The grid spacing $(\Delta x, \Delta y, \Delta z)$ was $0.02\sigma_V$ and the grid resolution (N_x, N_y, N_z) was 512.

3.2.3 MC simulation [53, 54]

We fixed the basic cell size and adopted the periodic boundary condition. In the present study, three types of MC simulations were carried out. We examined the canonical MC (CMC) simulations and the GCMC simulation. The CMC has a problem in comparison with the results calculated by the integral equations. The integral equation theory is formulated using the grand canonical ensemble, and the solvent density ρ_V is determined at the reservoir. By contrast, the solvent number density in

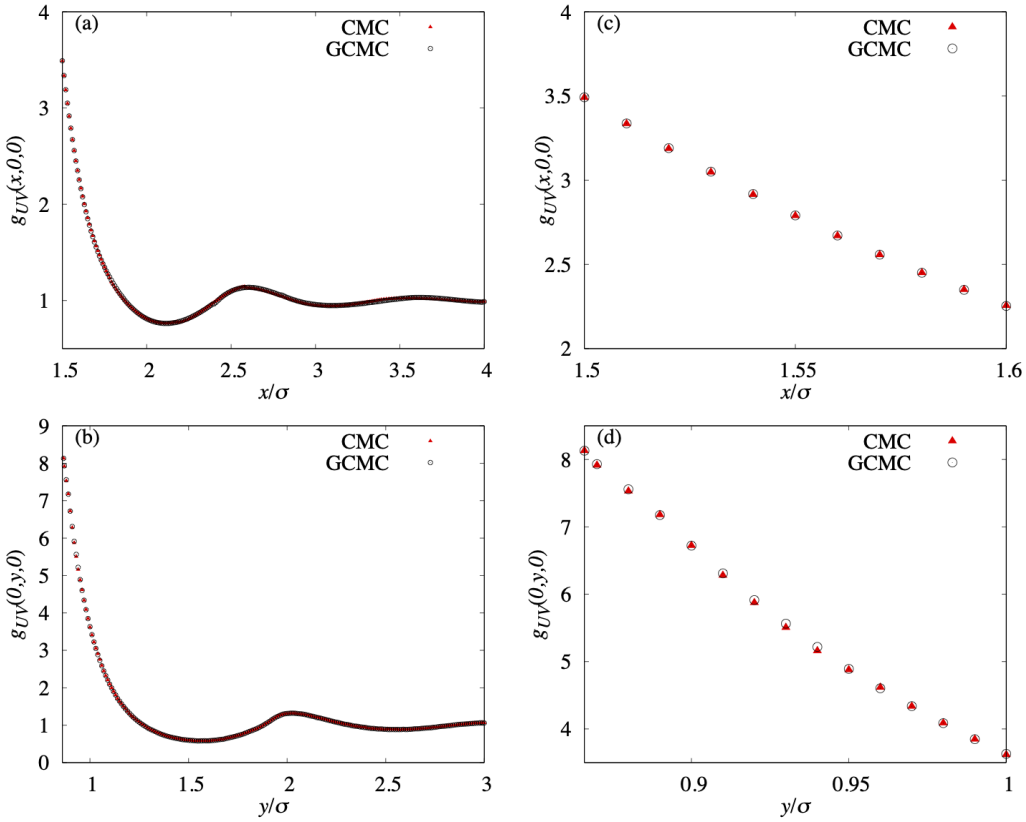


Figure 3.1: The spatial distribution (a) along the x-axis $g_{UV}(x, 0, 0)$ and (b) along the y-axis $g_{UV}(0, y, 0)$. The results of CMC and GCMC are compared. (c) and (d) are magnifications near the contact of (a) and (b), respectively.

the basic cell for the CMC deviates due to the insertion of the solute particle into the fluid. Then, we must obtain the number of solvent particles and the volume of the basic cell. However, this, in general, is not easy, but Schmidt and Skinner proposed a recipe [63]. Therefore, we adopted the recipe and adjusted the cell volume by using the following rule,

$$V = \frac{N}{\rho_V} + \Delta V_{ex}, \tag{3.11}$$

where N is total number of particles (solvent and solute) and ΔV_{ex} is the difference between the excluded volumes of solute and solvent particle. If the solute particle is spherical, $\Delta V = \frac{\pi}{6}[(\sigma_U + \sigma_V)^3 - (\sigma_V + \sigma_V)^3]$, where σ_U and σ_V are the solute and the solvent diameters, respectively. This recipe has been adopted in the case of a spherical solute particle, and gave satisfactory results [62–64], despite the fact that in the present study the solute shape was not spherical. We call this approach CMC1.

In addition to the methods noted in the main manuscript, another CMC simulation

was carried out. This simulation does not take into account the adjustment for the excluded volume. We added two hard-spheres of solute to the total number of particles and determined the cell volume using the set total number density. This method was called CMC2.

We also carried out the GCMC [65–68]. Although the calculation cost of the GCMC is relatively high, this choice is most adequate because the integral equation theory is formulated using the grand canonical ensemble.

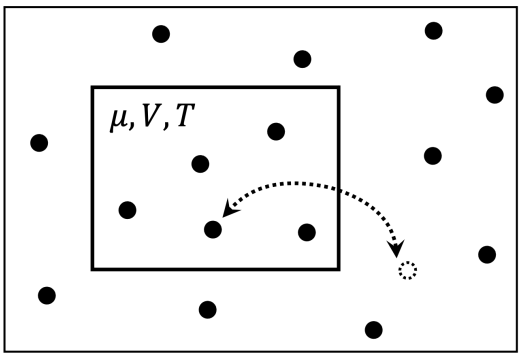


Figure 3.2: A system with fixed μ, V and T that exchanges energy and particles with a reservoir.

We considered a one-component system with fixed volume(V), temperature(T), and chemical potential(μ) which exchanges energy and particles with the reservoir (Fig.3.2). $\mathbf{x}_i$ denotes a given configuration i of the system, and $U(\mathbf{x}_i)$ is the corresponding total potential energy. The probability P_i of configuration $\mathbf{x}_i$ in the grand canonical ensemble is given by the following:

$$P_i = \frac{1}{\Xi}\frac{1}{N!\Lambda^{3N}}\exp[\beta N\mu - \beta U(\mathbf{x}_i)]. \tag{3.12}$$

Ξ is the grand canonical partition function

$$\Xi = \sum_{N=0}^{\infty}\frac{1}{N!\Lambda^{3N}}\int\dots\int\exp[\beta N\mu - \beta U(\mathbf{x}_i)]d\mathbf{x}_1\dots d\mathbf{x}_N, \tag{3.13}$$

where $\Lambda = h/(2\pi mkT)^{1/2}$. h and m are Plank's constant and the mass of the particle, respectively. The chemical potential for an ideal gas is

$$\mu_{id} = kT\left[\ln\Lambda^3 + \ln\frac{\bar{N}}{V}\right], \tag{3.14}$$

where $\bar{N}$ is the average number of particles. Transforming the integrals to dimensionless particle coordinates, $d\tau = V^{-1}d\mathbf{x}$, and substituting the chemical potential for an ideal gas produces

$$P_i = \frac{1}{\Xi}\frac{1}{N!}\exp[\beta N\mu_{ex} - \beta U(\tau_i) + N\ln\bar{N}]. \tag{3.15}$$

In Eq.(3.15), μ_{ex} is the excess chemical potential of fluid relative to a perfect gas with the same particle mass, density, and temperature. Before the GCMC simulation, we used Widom's insertion method in a canonical ensemble for determining the excess chemical potential as a function of the bulk density.

The grid cell volume was $\Delta V = \Delta x\Delta y\Delta z$, where Δx, Δy, and Δz denote the grid spacings in Cartesian coordinates and $\Delta x = \Delta y = \Delta z = 0.01\sigma_V$. The cell size was set equal to $L = 12.36\sigma_V$ and the number of particles N was 1372 in the CMC simulation. The number of particles varies in the GCMC simulation. The spatial distribution function $g(x, y, z)$ was calculated as follows:

$$g_{US}(x, y, z) = \Delta N(x, y, z)/\rho\Delta V, \tag{3.16}$$

where $\Delta N(x, y, z)$ is the number of solvent particles in the grid cell. We used the Metropolis algorithm and performed 10^7 MC steps for equilibration of hard-sphere fluids and over 10^{10} MC steps for the collection of ensemble averages. The contact value was estimated by extrapolation [62].

We compared the distribution functions around the dimer obtained using CMC1, CMC2 and GCMC in Fig.3.1. The agreement between three results was good. When the GCMC sampling is hard, CMC could be used. Because the results of CMCs and GCMC agreed well, we can conclude that the size of the simulation box was large enough and the sampling number was sufficient for obtaining the distribution functions. The GCMC results can thus be recognized as exact results.

3.3 Results and Discussion

3.3.1 Calculation for a spherical solute

First, we calculated the distribution functions around a spherical solute to test the feasibility of the numerical solution of the integral equation theory on a 3D discretized

grid with a finite number of grid points. For a spherical system, the results of the 3D integral equation theory and the radial-symmetric integral equation theory must agree. The distribution function of solvents along the x-axis through the center of the particle is shown in Fig.3.3. The plot agrees with the result calculated using the radial-symmetric integral equation theory. The deviations between the two distribution functions are small enough. We confirmed that the 3D integral equation theory accurately reproduced the correct density distribution.

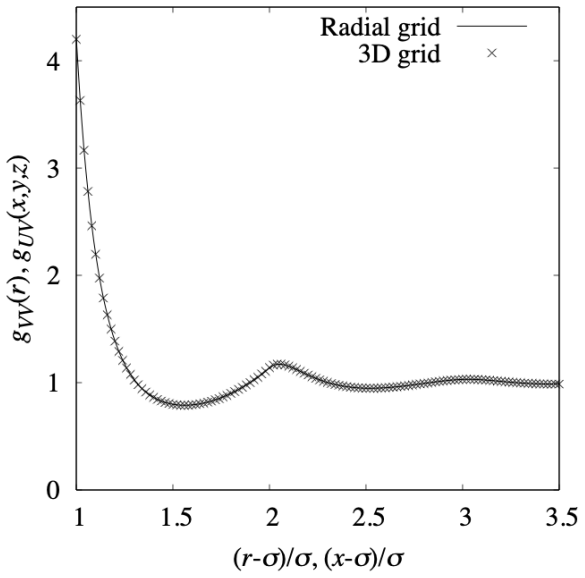


Figure 3.3: The distribution function of solvents around a spherical solute obtained from the integral equation theory on a radial grid with spacing $0.01\sigma_V$(solid line) and a 3D grid with spacing $0.02\sigma_V$(crocces). The HNC closure was used.

3.3.2 Spatial distribution of solvent around a contact dimer

The spatial distribution $g_{UV}(x, y, z)$ around a nonspherical solute, a contact dimer was obtained by GCMC simulation. Fig.3.4(a) shows the spatial distribution around the dimer $g_{UV}(x, y, 0)$. We compared this distribution function with the distribution functions obtained using the 3D integral equation theories. In Fig.3.4(b), we show the result obtained by PY closure. Similar map has obtained in the previous studies [69, 70]. These color maps for the solvent distribution are similar to each other. This means that the PY closure is qualitatively reasonable. The same applied to the spatial

distributions $g_{UV}(x, y, 0)$ calculated using other closures.

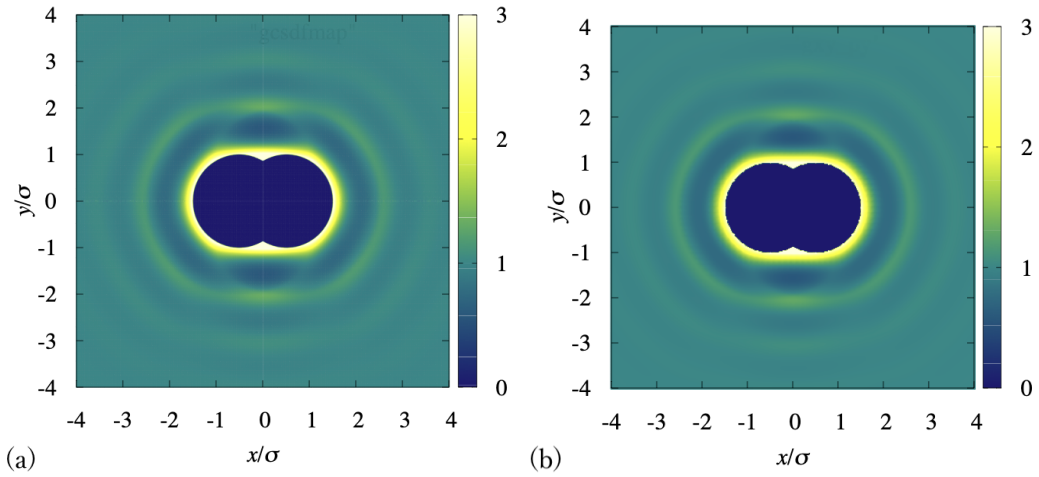


Figure 3.4: The spatial distribution around dimer in the xy plane $g_{UV}(x, y, 0)$ calculated by (a) GCMC simulation and (b) PY closure.

We also plotted the spatial distribution along the x-axis $g_{UV}(x, 0, 0)$ and that along the y-axis $g_{UV}(0, y, 0)$ to quantitatively compare the results obtained by MC and closures. as shown in Fig.3.5. Two specific sections are chosen, namely the most convex and the most concave sections, and the distribution functions, namely $g_{UV}(x, 0, 0)$ and $g_{UV}(0, y, 0)$ are plotted.

In Fig.3.5(a), (c), we show the solvent distribution around the most convex surface of the solute, namely $g_{UV}(x, 0, 0)$. The curvature of the solute particle is the same as that of the solvent particles. We can find the difference near the contact distance. The result obtained by the MHNC approximation almost agrees with that obtained by the GCMC simulation. By contrast, the PY approximation underestimates, and the HNC approximation overestimates the values near the solute surface. This agreement and these differences were also observed when the solute was a spherical one with a size of solvent particle [62]. These results were reasonable because the curvature of the most convex surface of the solute was the equal to that of the solvent particles.

By contrast, the behaviors of $g(0, y, 0)$ (Fig.3.5(b),(d)) differ from that of $g(x, 0, 0)$. The deviations from the GCMC simulation result become significant. The contact value of the HNC approximation was about 1.5 the GCMC result, and the contact value of the PY approximation was about half the GCMC result. The MHNC result maintained a small deviation from that for GCMC. The ratio of the contact value

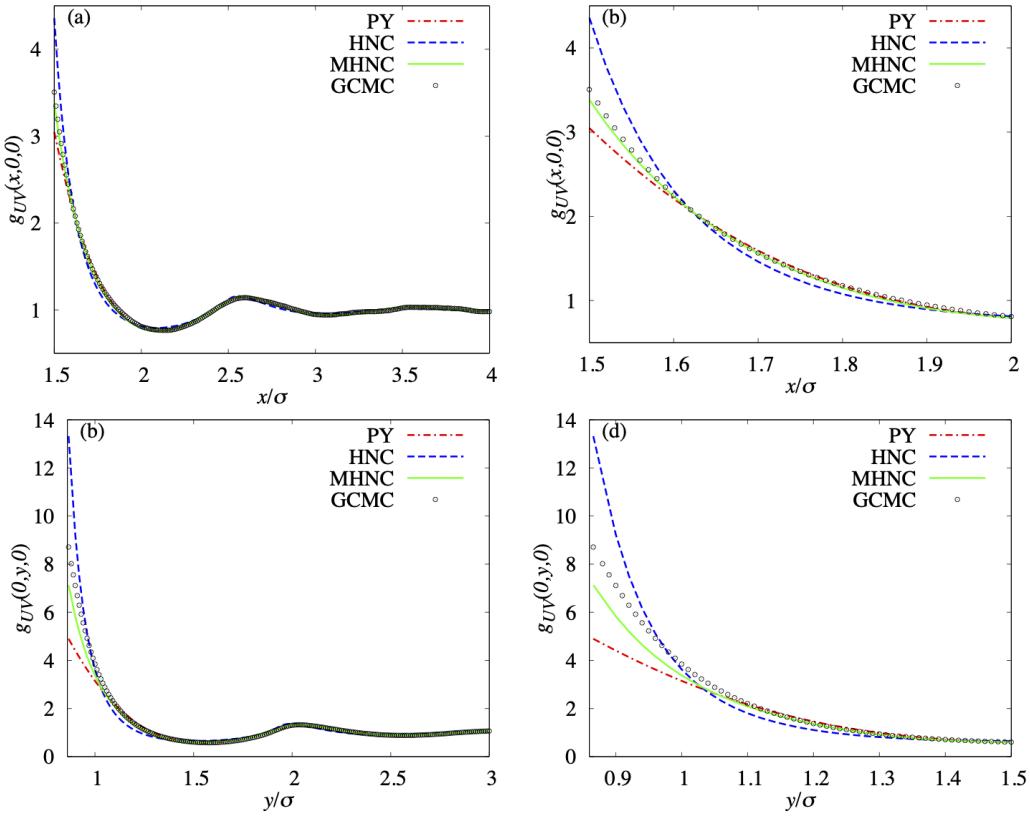


Figure 3.5: The spatial distribution (a) along the x-axis $g_{UV}(x,0,0)$ and (b) along the y-axis $g_{UV}(0,y,0)$. The results of integral equation theory with the PY, HNC, and MHNC closures are compared with the GCMC simulation. (c) and (d) are magnifications near the contact of (a) and (b), respectively.

was about 0.85. These results indicates that the MHNC closure provides a good approximation even near the concave surface.

We also analyzed other bridge functions, namely, the Verlet bridge function (Eq.(3.7)) and the Duh–Henderson bridge function (Eq.(3.8)). The comparisons with the MHNC bridge function (Eq.4.5) are plotted in Fig.3.6. The plots for the three bridge functions cannot be distinguished. As also investigated in the previous papers [51, 62], when a spherical solute was immersed in the one-component hard-sphere fluid, these bridge functions provided adequate spatial distribution functions. Therefore, we discuss that the superiority of the MHNC closure is slight in the three bridge functions. The superiority of the MHNC closure becomes significant when the value of $\gamma_{ij}(r)$ is negative and the absolute value $|\gamma_{ij}(r)|$ is large enough. In the case of a one-component solvent system, as mentioned in our previous paper [62], the function

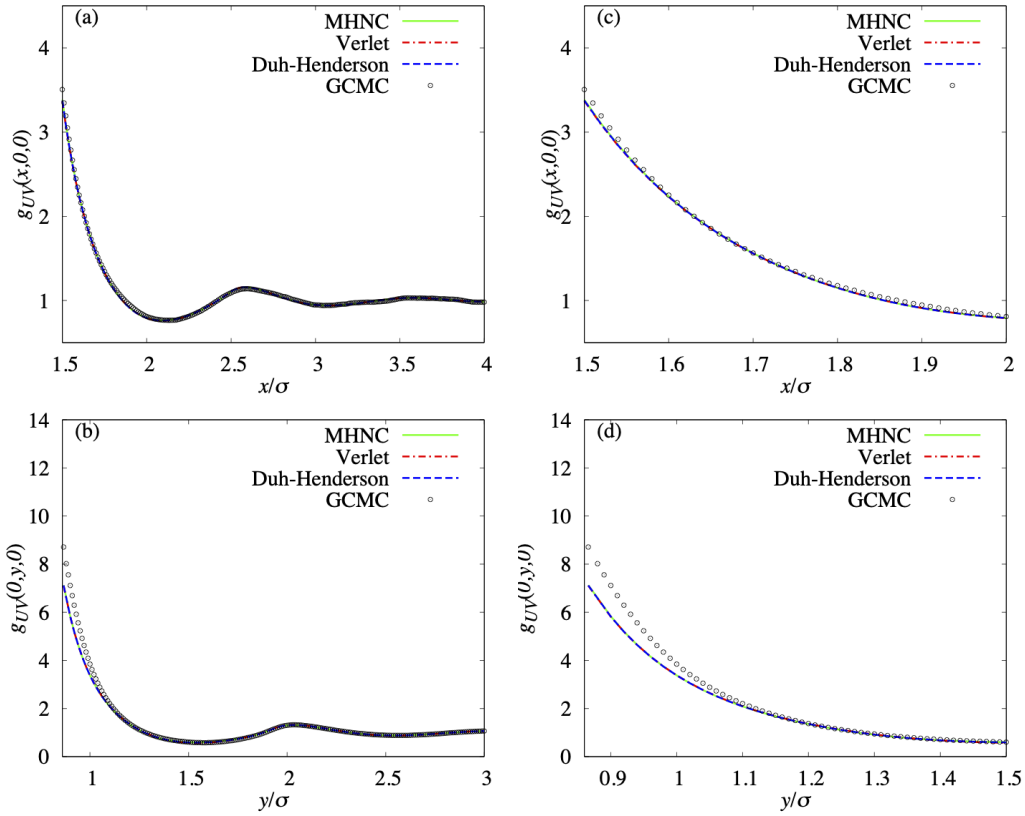


Figure 3.6: The spatial distribution (a) along the x-axis $g_{UV}(x,0,0)$ and (b) along the y-axis $g_{UV}(0,y,0)$. The results of integral equation theory with the MHNC, Verlet and Duh–Henderson bridge functions are compared with the GCMC simulation. (c) and (d) are magnifications near the contact of (a) and (b), respectively.

does not have a large negative value. This is the reason for the small differences between the results calculated using three bridge functions. In the present study, this conclusion was maintained even when the surface of the solute was concave, as in the case of $g(0,y,0)$ (See Fig.3.6(b), (d)). The accuracy of the approximations with the bridge functions depends on the surface curvature of the solute. Fig.3.5 suggests that the accuracy near the convex surface of the solute is better than that near the concave surface. The spatial distribution functions $g(0,y,0)$ obtained using the approximation with the bridge functions are slightly smaller than that obtained using GCMC.

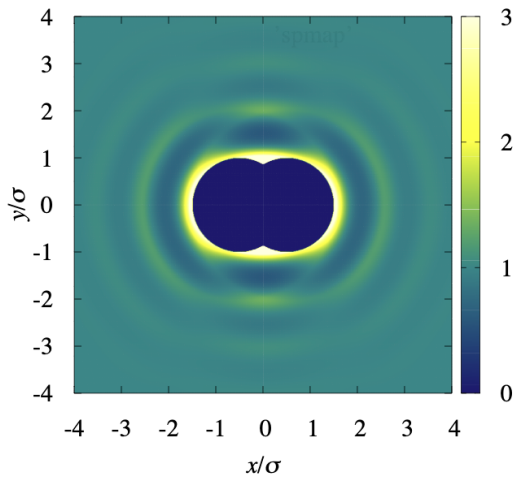


Figure 3.7: The spatial distribution obtained by using the superposition approximation (SA) around the dimer in the xy plane $g(x, y, 0)$.

3.3.3 Triplet distribution functions

The validity of the closure can also be evaluated in terms of the three-body correlation. The triplet distribution functions are discussed here. The triplet distribution functions $g^{(3)}(\mathbf{r_1}, \mathbf{r_2}, \mathbf{r_3})$ mean the reduced probability of finding three particles at positions $\mathbf{r_1}$, $\mathbf{r_2}$, and $\mathbf{r_3}$ [71]. Then $g^{(3)}$ can be written as a triple product of the pair distribution functions $g^{(2)}$, which is called superposition approximation (SA) [72]:

$$g^{(3)}_{\mathrm{SA}}(\mathbf{r_1}, \mathbf{r_2}, \mathbf{r_3}) = g(\mathbf{r_1}, \mathbf{r_2})g(\mathbf{r_2}, \mathbf{r_3})g(\mathbf{r_3}, \mathbf{r_1}). \tag{3.17}$$

When two particles are fixed at $\mathbf{r_1}$ and $\mathbf{r_2}$, a reduced probability of finding a particle at position $\mathbf{r_3}$ is as follows using Eq.(3.17):

$$\begin{aligned} g_{\mathrm{SA}}(r_{12}; \mathbf{r_3}) &= g^{(3)}_{\mathrm{SA}}(\mathbf{r_1}, \mathbf{r_2}, \mathbf{r_3})/g(\mathbf{r_1}, \mathbf{r_2}) \\ &= g(\mathbf{r_2}, \mathbf{r_3})g(\mathbf{r_3}, \mathbf{r_1}), \end{aligned} \tag{3.18}$$

where

$$r_{12} \equiv |\mathbf{r_1} - \mathbf{r_2}|. \tag{3.19}$$

Choosing the third Cartesian coordinate $\mathbf{r_3} = (x_3, y_3, z_3)$ in Eq.(3.18) and setting $\mathbf{r_1} = (-0.5\sigma_V, 0, 0)$ and $\mathbf{r_2} = (0.5\sigma_V, 0, 0)$, we can calculate the spatial distribution function $g(x_3, y_3, z_3)$ around a contact dimer. The spatial distribution in xy plane calculated from Eq.(3.18) is shown in Fig.3.7. Here, $g(\mathbf{r_2}, \mathbf{r_3})$ and $g(\mathbf{r_3}, \mathbf{r_1})$ are obtained

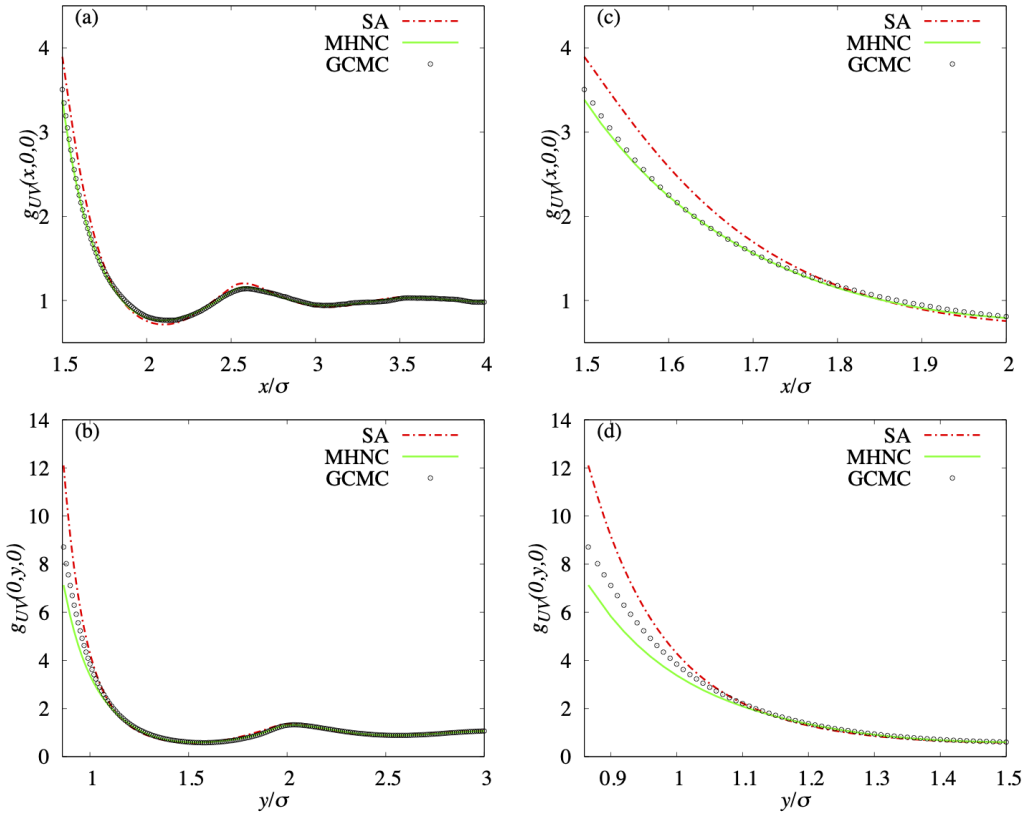


Figure 3.8: The spatial distribution (a) along the x-axis $g_{UV}(x,0,0)$ and (b) along the y-axis $g_{UV}(0,y,0)$. (c) and (d) are magnifications near the contact of (a) and (b), respectively. Here, $g(r)$ obtained by GCMC is adopted in SA calculation.

using the radial distribution functions $g_{VV}(r)$ calculated by GCMC simulation. We can find the density wave of the hard-sphere fluid around a dimer as shown in Fig.3.7. The shape of density waves is in good agreement with that in Fig3.4. This means that for simple solute models, the shape of the distribution of solvents can be estimated more or less accurately from SA.

Fig.3.8 shows the spatial distributions calculated using the MHNC approximation and the GCMC simulation with the results obtained using Eq.(3.18), namely SA. The features of the functions $g_{UV}(x,0,0)$ are similar. The position of the peaks and the minimums are almost the same. The numerical agreement between the GCMC and MHNC results is excellent (the contact value $g(1.5\sigma,0,0) = 3.8935$(SA) and 3.5062(GCMC), 3.3838(MHNC) in Fig.3.5,3.8). By contrast, the difference between the SA results and the exact results is the largest but it is smaller than the differences for PY or HNC ($g(1.5\sigma,0,0) = 3.0473$(PY) and 4.3583(HNC) in Fig.3.5).

The same behavior is obtained in $g_{UV}(0, y, 0)$ as $g_{UV}(x, 0, 0)$, and the differences become more significant. The first peaks of functions $g_{UV}(0, y, 0)$ are higher than those of the $g_{UV}(x, 0, 0)$ because the excluded volume at the contact position increases. The behaviors of three functions $g_{UV}(0, y, 0)$ for SA, MHNC, and GCMC are similar to each other again (the contact value $g(0, \sqrt{3}/2\sigma, 0) = 12.1041$(SA) and 8.7112(GCMC), 7.1283(MHNC) in Fig.3.5,3.8). The spatial distributions calculated using the SA with $g(r)$ obtained by GCMC (or the MHNC approximation) are more accurate than those obtained using the HNC or the PY approximations ($g(0, \sqrt{3}/2\sigma, 0) = 4.8957$(PY) and 13.3187(HNC) in Fig.3.5). This suggests that when discussing the qualitative behavior of the distribution functions, it is not bad for SA to use the accurate $g(r)$ despite the simplicity. By contrast, the differences between the SA results and the exact results show the existence of multiple body correlations.

Here, we discuss the approximations based on the three-body effect. The SA takes into account only three independent pair correlations and ignores the effect of any of these pair correlations. Therefore, the SA results differ from the exact results when the third particle is in the neighborhood of the pair. For an assessment of the true three-body effect, we consider the ratio $g^{(3)}$ to the value of the SA(Eq.(3.18)) [70, 73, 74]:

$$\Gamma(\mathbf{r_1}, \mathbf{r_2}, \mathbf{r_3}) = \frac{g^{(3)}(\mathbf{r_1}, \mathbf{r_2}, \mathbf{r_3})}{g(\mathbf{r_1}, \mathbf{r_2})g(\mathbf{r_2}, \mathbf{r_3})g(\mathbf{r_3}, \mathbf{r_1})}, \tag{3.20}$$

which can be written as

$$\Gamma(r_{12}; \mathbf{r_3}) = \frac{g^{(3)}(r_{12}; \mathbf{r_3})}{g(\mathbf{r_2}, \mathbf{r_3})g(\mathbf{r_3}, \mathbf{r_1})}. \tag{3.21}$$

If the SA were exact, the ratio Γ should be unity. However, the calculated results are not unity except the SA because the SA becomes worse as the density of fluid increases.

Here, a contact dimer (particles 1 and 2) and a solvent particle 3 are considered. We define the ratio $\Gamma(\theta) = \Gamma(\mathbf{r_3})$ as a function of the angle θ enclosed by $\mathbf{r_1} - \mathbf{r_2}$ and $\mathbf{r_3} - \mathbf{r_2}$ (see Fig.3.9). The results $\Gamma(\theta)$ at $|\mathbf{r_3} - \mathbf{r_2}| = \sigma$ calculated using the HNC, PY, and MHNC closures are shown in Fig.3.10(a). The result of GCMC is also shown in

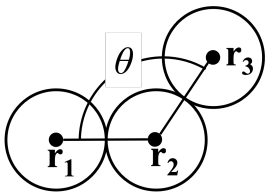


Figure 3.9: Description of θ

the figure. First, we compare the SA and GCMC results. The exact result (GCMC) has two minimums around 60° and 180°, and there is a peak around 120°. The first minimum at 60° is explained by the overestimation of the SA. In other words, the numerator is smaller than the denominator in Eq.(3.21), which is the definition of Γ. The reference value, namely the denominator, is obtained by the SA in Eq.(3.21). In fact, particle 3 is stable at 60° due to the reduction of excluded volume for the three particles. However, the exclusion effect for particle 3 by particle 2 is not included in the SA. Therefore, the peak of the solvent spatial distribution estimated by the SA becomes larger than the exact value at 60°. This explained the minimum of the $\Gamma(\theta)$ at 60°[1]. In addition, it seems that the value smoothly approaches unity at 180° because of the reductions of the three-body effect.

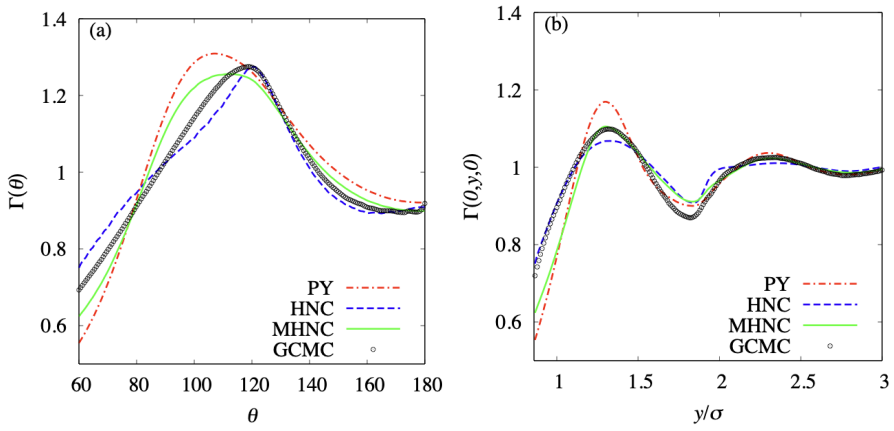


Figure 3.10: The ratio (a) $\Gamma(\theta)$ and (b) $\Gamma(0, y, 0)$ obtained by the GCMC method and the integral equation theory with the PY, HNC and MHNC closures.

The GCMC result have the peak around 120° because of the influence of the solvent located at 60°. A high peak for the reduced spatial distribution appears at 60°,

[1]This minimum can also be explained in terms of the triplet overlap of the excluded volume. In the SA, the overlap of the excluded volume is double–counted.

and the peak value is much larger than unity. As the probability of the existence of the solvent particle at 60° increases, the fourth particle at 120° is stabilized because of the reduction of the excluded volume caused by the adsorption. In the SA, this three-body effect is not included. Therefore, the probability of particle existence obtained using the GCMC at 120° becomes larger than that estimated by the SA, and the peak appears around 120° in Fig.3.10(a).

Fig.3.10(a) also shows $\Gamma(\theta)$ for the various closures, namely the PY, the HNC, and the MHNC closures. The functions $\Gamma(\theta)$ maintain the same characteristics: two minimums at 60° and 180° and one peak at 120°. In contrast to the qualitative validity of these approximate three-body correlations, quantitative deviations are obtained. The differences between the PY result and the exact result (the GCMC simulation) is the largest. The HNC results have the opposite sign of deviation to the MHNC results, but both differences are small. The exact result is also close to these two plots. The HNC result is very accurate in the plots $\Gamma(\theta)$. However, it does not mean that the HNC closure is very accurate in the calculation of the spatial distribution. The deviation of the spatial distribution function calculated using the HNC closure is larger than that calculated using the MHNC closure because of the large deviation of the two-body correlation from the exact solution.

Fig.3.10(b) has the plots $\Gamma(\mathbf{r_3}) = \Gamma(x, y, z)$ as a function of the coordinate (x, y, z). $\Gamma(0, y, 0)$ at $y = \sqrt{3}/2$ is equivalent to $\Gamma(\theta)$ at $\theta = 60°$. The plots $\Gamma(0, y, 0)$ oscillate, and the qualitative features of the plots in the Fig.3.10(b) are similar. There are two deep minimums around $y = \sqrt{3}/2$ and 1.8, with a distinct peak around 1.3. However, the difference between the PY result and the exact (GCMC) results is the largest. The accuracies of the MHNC and the HNC approximations depend on the value y. The HNC approximation is the most accurate around $y = \sqrt{3}/2$, while the MHNC becomes the most precise around the first peak at $y = 1.8$. However, the difference is not quantitively large.

Here, we discuss the detail of the differences of $\Gamma(0, y, 0)$. The MHNC results evaluated well the value of $\Gamma(0, y, 0)$ near the first and second peaks. These peak positions correspond to the bottoms (or minimums) of the distribution function $g(0, y, 0)$, where SA underestimates. We examined the $g(r)$ between two hard spheres im-

mersed in a one–component hard–sphere fluid using the MHNC approximation in the previous study [62]. The MHNC results showed very accurate first minimums in the $g(r)$. The accuracy was much better than that of the HNC results. It can be concluded that these are common features in the present study. Therefore, the evaluation of the distribution functions using the MHNC approximation is very accurate, even when discussing the value around the minimum. By contrast, the difference of the HNC results for $\Gamma(0, y, 0)$ around the first minimum is smaller than that of the MHNC result, although the differences of the HNC results for $g(r)$ are much larger than those of the MHNC results. The HNC results overestimate both the numerator and the denominator of Eq.(3.21). By contrast, in the MHNC results, only the denominator is very accurate in the calculation Eq.(3.21).

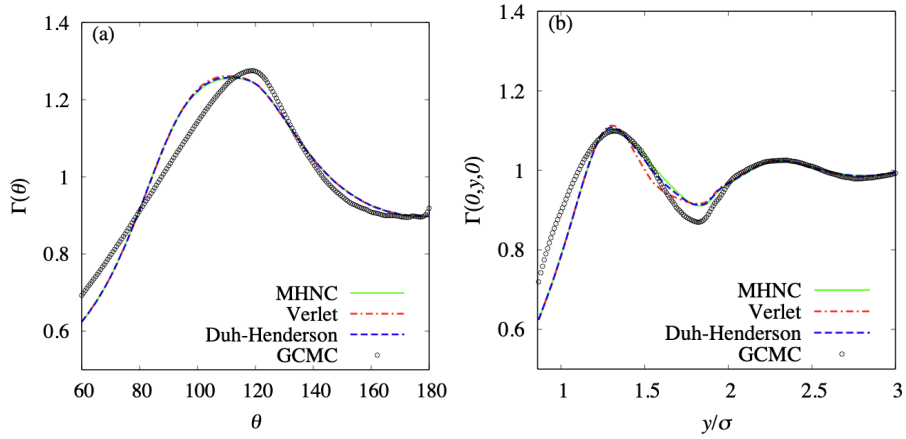


Figure 3.11: The ratio $\Gamma(\theta)$ and $\Gamma(0, y, 0)$ obtained by the GCMC method and the integral equation theory with the MHNC, Verlet and Duh–Henderson bridge functions.

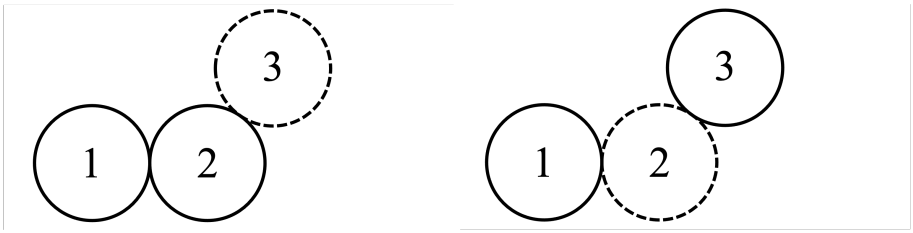
We also examined $\Gamma(\theta)$ and $\Gamma(0, y, 0)$ using three bridge functions: the MHNC, the Verlet, and the Duh–Henderson bridge functions. Fig.3.11 shows the comparison with GCMC simulation. The three bridge functions have virtually the same results. We can find differences in $\Gamma(0, y, 0)$ around $y = 1.5\sigma_V$. The difference between the Verlet bridge result and the exact result is the largest, but it can be ignored. Therefore, it is consistent with the agreement in the spatial distribution functions (see Fig.3.6).

The HNC approximation seems to overestimate the spatial distribution function, especially the peak of the function, due to the ignorance of the bridge function (see Eq.(3.5).) However, in the preceding study for the correlation functions of the hard-

sphere solvent around a hard sphere, the radial distribution function calculated by the MHNC approximation is very accurate [62]. In the study, the exact radial distribution function was calculated using the canonical MC simulation with a correction. The conclusion can be believed because we checked the validity of the correction using the GCMC simulation. Furthermore, it is known that the pair correlation function also agrees well with MC if the bridge function is constructed to satisfy the thermodynamic consistency [43, 75]. Although the thermodynamic consistency in the case of the HNC approximation is insufficient, the results given by the MHNC approximation are automatically almost completely satisfied [76]. Therefore, we expected that the bridge function of the MHNC approximation would be an adequate improvement and that the three-body correlation is also better than that calculated using the HNC approximation.

However, the superiority of three closures with a bridge function, namely MHNC, Verlet, and Duh–Henderson closures, in the spatial distribution functions is not caused by the superiority of the three-body correlations. Here, we can also confirm the validity of the closure in terms of the three-body correlation obtained in the present study. Although the incorporation of the Verlet, Duh–Henderson, and Kinoshita bridge functions improved the pair correlation function, it did not necessarily lead to better results for the triplet distribution function under the present calculation conditions. Therefore, it was interesting to compare the results from the MC simulation and the OZ equations coupled with PY, HNC, and the three different closures in terms of the triplet distribution function. I can conclude that the calculated results do not always imply that the MHNC approximation gives better results than the HNC approximation for the three-body correlation.

Here, I mention the calculation procedures for $\Gamma(\theta)$. There are two procedures. Configuration 1 (see Fig.3.12(a)) was adopted in the above calculation because it is easy to obtain the function in the molecular simulation. However, the $\Gamma(\theta)$ can also be obtained by using configuration 2 (see Fig3.12(b)). In the method, a pair of two separate particles, 1 and 3, was fixed at r_1, and r_3 was immersed in the solvent particles. Thus, the spatial distribution function around them $g^{(3)}(r_{13}; r_2)$ to obtain $\Gamma(\theta)$ can be calculated. In other words, $\Gamma(\theta)$ was calculated using the spatial distribution



(a) Configuration 1
: Particle 1 and 2 are fixed

(b) Configuration 2
: Particle 1 and 3 are fixed

Figure 3.12: Description of the two configurations for the $\Gamma(\theta)$ calculation: (a)$\Gamma(\theta) = \Gamma(r_{12}; \mathbf{r_3})$ and (b)$\Gamma(\theta) = \Gamma(r_{13}; \mathbf{r_2})$. The solid spheres are fixed and the broken spheres are obtained for distribution functions.

function at $\mathbf{r}_2$ when particles 1 and 3 were located at variable θ. Then, $\Gamma(\theta)$ could be written as

$$\Gamma(\theta) = \Gamma(r_{13}; \mathbf{r_2}) = \frac{g^{(3)}(r_{13}; \mathbf{r_2})}{g(\mathbf{r_1}, \mathbf{r_2})g(\mathbf{r_2}, \mathbf{r_3})}. \tag{3.22}$$

Here, $g^{(3)}(r_{13}; r_2)$ was not the spatial distribution around a contact dimer solute. However, we could compare the $\Gamma(\theta)$ between the two configurations (2 vs. 1). In Fig.3.13, we show the $\Gamma(\theta)$ results when configuration 2 is adopted. The discrepancies between the exact GCMC result and the results for the various closures are much larger than those obtained using configuration 1 (Fig.3.10).

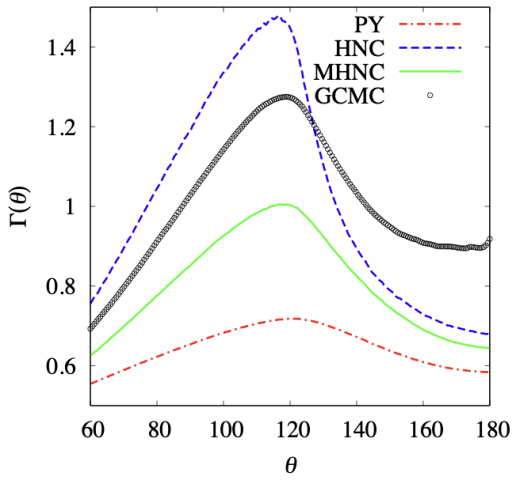


Figure 3.13: The ratio $\Gamma(\theta)$ calculated by the integral equation theory with the PY, HNC, and MHNC closures at configuration 2 and the GCMC method.

Here, the difference between the calculated results using the integral equation the-

ories and the exact result using GCMC in Fig.3.13 is discussed. For the exact results, the plots for configuration 2 and configuration 1 were the same as mentioned above. The closure dependence (See Fig.3.10(a)) appeared for configuration 1. However, the dependence was much smaller than that for configuration 2 (See Fig.3.13). For configuration 2, the value calculated using MHNC was about 0.65 when $\theta = 180°$. At $\theta = 180°$, because the three particles were aligned on a straight line, we expected the three-body correlation effect to be less apparent. This is because particle 3 is not strongly affected by particle 1, as particle 3 is located just behind particle 2 (See Fig.3.12). In fact, in the case of configuration 1, all plots converge to a value of about 0.9 at $\theta = 180°$ (See Fig.3.10(a)). That is, in the case of configuration 1, especially when $\theta = 180°$, any approximation is close to the exact result. These results contrast with those of configuration 2. The poor approximation for the three-body correlation by closures is emphasized using configuration 2.

I discuss the reason for the difference between configurations 1 and 2 as follows. The three-body correlation becomes weak in a dilute system, and Γ goes to 1 [69]. By contrast, the multi-body correlation becomes significant in a high-packing fluid because the calculation of the multi–body correlation becomes more significant as the local concentration increases. Then, the value Γ is expected to worsen as the local density increases. The local density depends on the configuration. As the overlap of the excluded volume increases, the configuration of the three particles becomes stable in the present packing fraction, and the local density of the third particle becomes high at the location. Let us think of this location. Therefore, the large value of the excluded volume overlap for the third particle, namely particle 2 for configuration 2, should correlate with the accuracy of the multi-body correlation.

For example, in configuration 2, particle 2 contacts particles 1 and 3. Then, particle 2 has two overlaps of excluded volume with particles 1 and 3 at any angle. Because of the large amount of overlap, it seems that the local density of the third particle for configuration 2 is high, and the deviation of the three-body correlation value from the exact value becomes large in the entire range of $\theta = 60°$ to $180°$ except the cross point near $\theta = 130°$ for the HNC closure.

Here, I will discuss the validity of the above argument regarding the relationship

between local density and three-body correlations. In configuration 1, the excluded volume of particle 3 overlaps with both those of particles 1 and 2 near $\theta = 60°$. Therefore, at $60° < \theta < 120°$, the differences between the results of the exact calculation and those calculated using the integral equation theories appear. However, when $120° < \theta < 180°$, the excluded volume of particle 3 overlaps only that of particle 2. Since the overlap is less than that for $60° < \theta < 120°$, the results of the integral equation theory roughly agree with the exact results. The above results coincide with the argument that the three-body correlation tends to be inaccurate in configurations where the local density of particles is high. We can expect that the deviation between the exact results and the results for configuration 1 becomes smaller than those for configuration 2.

3.4 Conclusion

The reduced spatial density profile of hard spheres in the vicinity of a hard-sphere dimer was calculated using the 3D–OZ equation with closure and GCMC simulations. Not only the PY and HNC closures but also the closures with bridge functions proposed by Verlet, Duh–Henderson, and Kinoshita were examined to assess the approximations. An assessment using the spherical solute has been published. However, a nonspherical solute has not been examined in the assessment of the integral equation theories. There was one more advantage in this assessment study. I adopted the GCMC simulation, not the canonical MC simulation because the integral equations were formulated under the grand canonical ensemble. This advantage was not clear In the case of the present study because the solute particle was small. However, the advantage must become clear when the solute is a macromolecule.

I compared the results obtained from the integral equation theories with the exact result obtained using the GCMC. The results using the three closures taking account of the bridge function, such as the MHNC closure, were much better than others. The results given by the SA were not so bad, although the calculation cost is reasonable. Surprisingly, if we used the precise radial distribution function $g(r)$, the reduced density profile obtained by the SA was better than the results calculated using the PY and HNC closures. However, the advantage of the SA is not so valuable

because the 3D–OZ calculation is not so expensive. By contrast, the closures with bridge functions have an advantage in the calculation of the reduced density profiles $g_{\mathrm{UV}}(x, y, z)$. As discussed above, the advantage of the bridge function proposed by Kinoshita should appear in the other studies, the concentration dependence studies for the multi-component solvent system.

I also compared the three–body correlations obtained using the integral equation theory with the exact results. The comparisons with the GCMC results showed that the inferiorities of SA and PY approximations towards the NHC and other closures in accuracy were clear. On the other hand, the superiority of the three closures, taking into account the bridge function over the results using the HNC closure, was not clear under the present calculation conditions. I can conclude the above results as follows. The HNC approximation overestimates the pair correlation function due to the ignorance of the bridge function. By contrast, it seems that the closures with the bridge functions are slightly worse than the HNC approximation for the three–body correlation, although the results of the closures with bridge functions are almost completely accurate. Because the HNC closure was insufficient for thermodynamic consistency, the slight worsening caused by the bridge function in the closures with the bridge functions was surprising. I think that there is a cancellation due to the deviations in the denominator and the numerator of Eq.(3.21) in the case of the HNC closure.

Chapter 4

Entropic molecular recognition in a hard-sphere mixture

4.1 Introduction

In the study of depletion interaction, one of the simplest and most popular systems has been the model of two large hard spheres immersed in a small hard-sphere fluid. The effective interaction between two large hard spheres has been calculated: that is, when the large spheres are in contact, the excluded volumes for the small spheres overlap, and the configuration volume for small spheres increases. The configurational entropy for the small spheres increases due to the increase in the configurational volume. Therefore, the free energy of the system decreases due to the association of large spheres. The interpretation described above was given by Asakura and Oosawa in 1954 [39–41] (see section 1.3).

In the case of nonspherical solutes, the shape dependence of the depletion interaction is very interesting, as it gives the guest selectivity by a host molecule. Kinoshita studied the depletion interaction between a hard body with a hemispherical cavity and a large hard sphere using the 3D-HNC-OZ theory [15–18]. In that study, the selectivity of the diameter of the large sphere was shown. This lock-and-key relation can be explained by the Asakura–Oosawa (AO) theory. In the AO theory, as the overlap of the excluded volumes increases, the entropy gain increases. Here, the overlap reaches maximum when the large sphere fits into the cavity exactly. Therefore, the entropy gain becomes maximum. This selectivity could play an important role in the

molecular recognition phenomenon in biological and synthetic systems.

In the discussion for selectivity, Asakura and Oosawa's idea has an advantage, as mentioned above, and presents a problem in the calculation of the effective interaction shape. For example, there are problems in the estimations of the barrier in the recognition process. The remarkable simplification of the theory is in the modeling of small spheres as an ideal gas. The interaction between small particles is ignored in the AO theory. Therefore, the small spheres do not construct any liquid structure near the solute molecules even when the packing fraction is large, such as 0.4. The effective attraction increases monotonically as the large particles approach each other in the simplification. However, when the fluid of small particles is dense, we can expect that the effective interaction is oscillatory due to the inhomogeneity of the number density of small particles, namely the liquid structure. In particular, the high density of small particles on the surface of large particles leads to a free-energy barrier to overcome before reaching the contact. Kinoshita et al. used integral equation theories to clarify the problem in the shape of the effective interaction [15–18].

Some studies show the depletion interaction between spherical solutes immersed in a hard-sphere mixture [77, 78]. The effective interaction depends on the mixing ratio, the size ratio, and the packing fraction. For example, as the number density of smaller particles increases, the stabilization free energy at contact increases, and the free-energy barrier becomes higher. Roth and Kinoshita reported a reduction in the free-energy barrier between two large spheres in multicomponent systems of smaller spheres [79]. The oscillatory structure of depletion interaction in a multicomponent system is less than that in a one-component system, and the free-energy barrier for the association process is significantly reduced in a multicomponent system. They adopted two large spheres as solute molecules. It seems that the reduction is caused by the interference of the density waves caused by the various-sized particles. This idea of barrier reduction can be applied to the molecular recognition phenomenon.

The PMFs were calculated between a cyclic model (cyclodextrin-like) molecule and a spherical molecule in a multicomponent mixture using the 3D–MHNC–OZ theory. Chapter 2 shows that the MHNC theory is also a very accurate approximation in nonspherical systems with hard-body particles. Thus, the MHNC theory is a

suitable approximation for the present analysis. The results of the PMF obtained by 3D-HNC-OZ and AO theory are also shown for comparison.

4.2 Model and Method

4.2.1 Model

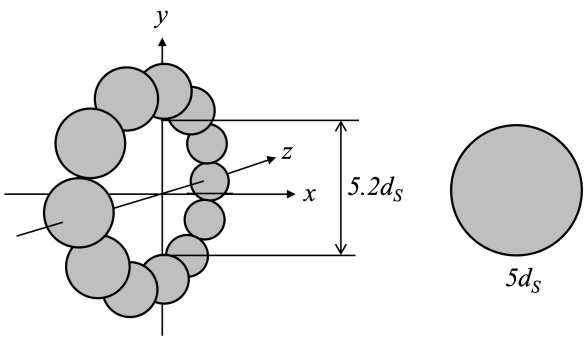


Figure 4.1: Model solutes considered. Solute 1 is a large hard sphere with diameter $5d_S$. Solute 2 is a ring cluster of hard spheres having a cavity with the diameter $5.2d_S$. The diameter of each hard sphere is $1.0d_S$. The coordinate system is chosen such that the origin is at the center of the ring.

Two solutes are immersed in a solvent. In the analysis of occurring behavior purely entropic in origin, a hard-core potential can be adopted for the system. Solute 1 is a large hard-sphere mixture acting as a guest molecule and solute 2 hused hard spheres forming a cyclic host molecule, (Fig4.1). The solvent is a multicomponent mixture of hard spheres with the diameter λd_S in the range of $1 < \lambda < 5$. The total packing fraction of the mixture is kept at 0.380 [77, 80]. Ambient water has this packing fraction. A total of nine systems were examined, and their specifications are summarized in Table4.1. The amplitude of effective interaction between large spheres in a mixture is damped the most when the two components have almost the same packing fraction [61, 79, 80]. According to this result, the packing fractions of the components were set at the same value.

4.2.2 Integral equation theory [27, 43]

The computer simulations can hardly be applied to the analysis of the molecular recognition in a mixture of small and medium spheres. When the size asymmetry of

Table 4.1: The nine systems considered. In system 9, for example, the packing fractions of the small spheres with the diameters d_S, $2d_S$, $3d_S$, and $4d_S$ are 0.095, 0.095, 0.095, and 0.095, respectively. The total packing fraction is kept at 0.380 in all the systems.

System	ds	$2ds$	$3ds$	$4ds$
1	0.380	⋯	⋯	⋯
2	⋯	0.380	⋯	⋯
3	⋯	⋯	0.380	⋯
4	⋯	⋯	⋯	0.380
5	0.190	0.190	⋯	⋯
6	0.190	⋯	0.190	⋯
7	0.190	⋯	⋯	0.190
8	0.127	0.127	0.126	⋯
9	0.095	0.095	0.095	0.095

the mixture is high, as in the present study, the basic cell required is unacceptably large. Moreover, when the guest concentration is diluted, obtaining a sufficiently accurate ensemble average becomes a difficult task. In contrast, the integral equation theory allows calculations in the infinite bulk system with very little computational effort.

The OZ equation is solved together with a closure equation [27, 43]. In this study, the correlation functions were obtained in the dilution limit of a solute particle immersed in a solvent. The OZ equation between solute (U) and solvent (V) is expressed by:

$$h_{UV}(r) = c_{UV}(r) + \rho_V \int c_{UV}(r) h_{VV}(|\mathbf{r}' - \mathbf{r}|) d\mathbf{r}', \tag{4.1}$$

where ρ is the number density, h is the total correlation function, and c is the direct correlation function. The system is spherically symmetric and r is the distance between the centers of solvent particles. For a nonspherical solute, the OZ equation is given on a 3D grid as follows:

$$h_{UV}(x, y, z) = c_{UV}(x, y, z) + \rho_V \int c_{UV}(x, y, z) h_{VV}(|\mathbf{r}' - \mathbf{r}|) dx' dy' dz'. \tag{4.2}$$

The vectors are $\mathbf{r} = (x, y, z)$ and $\mathbf{r}' = (x', y', z')$. Eq.4.2 in the momentum $\mathbf{k}$ space is written as:

$$\hat{\gamma}_{UV}(k_x, k_y, k_z) = \rho_V \hat{c}_{UV}(k_x, k_y, k_z) \hat{h}_{VV}(|\mathbf{k}|), \tag{4.3}$$

where $\gamma = h - c$, the symbol "^" indicates the Fourier transform. The vector $\mathbf{k} = (k_x, k_y, k_z)$. $\hat{h}_{VV}(|\mathbf{k}|)$ is calculated using the integral equation theory for solvent.

The closure relation equation is as follows:

$$c_{UV}(\mathbf{r}) = \exp[-\beta u_{UV}(\mathbf{r})] \exp[\gamma_{UV}(\mathbf{r}) + b_{UV}(\mathbf{r})] - \gamma_{UV}(\mathbf{r}) - 1, \tag{4.4}$$

where $\beta = (k_B T)^{-1}$, $k_B T$ is the Boltzmann constant times the absolute temperature. The functions u and b are the potential and the bridge function, respectively. For a spherical system, that is, coupled with Eq.(4.1), $\mathbf{r}$ is replaced with r. For a non-spherical system, that is, coupled with Eq.(4.2), $\mathbf{r} = (x, y, z)$. In this study, the HNC closures and modified closures with semiempirical bridge functions were used. The results from the HNC theory are not very accurate because the bridge function was neglected. The MHNC closure includes the bridge function proposed by Kinoshita [44, 51, 52, 62], as follows:

$$\begin{aligned} b_{UV}(\mathbf{r}) &= -0.5 \frac{\gamma_{UV}^2(\mathbf{r})}{1 + 0.8\gamma_{UV}(\mathbf{r})} \quad (\gamma > 0) \\ &= -0.5 \frac{\gamma_{UV}^2(\mathbf{r})}{1 - 0.8\gamma_{UV}(\mathbf{r})} \quad (\gamma < 0). \end{aligned} \tag{4.5}$$

The reliability of the MHNC closure has been demonstrated for mixtures with high size asymmetry. Furthermore, it has been verified that the MHNC closure provides more accurate spatial distribution functions than the PY and HNC closures. Thus, the MHNC theory is best suited for the analysis in this work.

First, the solute 1–solvent correlation functions are calculated using Eq.(4.1) and one of the closures ($U = 1$). Second, the solute 2–solvent correlation functions are calculated using Eq.(4.2) and the same closure ($U = 2$). The PMF between solute 1 and 2 is obtained from:

$$\beta W_{12}(x, y, z) = \beta u_{12}(x, y, z) - \gamma_{12}(x, y, z) - b_{12}(x, y, z), \tag{4.6}$$

where $\gamma_{12}(x, y, z)$ is calculated using inverting $\hat{\gamma}_{12}(k_x, k_y, k_z)$ given by:

$$\hat{\gamma}_{12}(k_x, k_y, k_z) = \rho_V \hat{c}_{2V}(k_x, k_y, k_z) \hat{h}_{1V}(|\mathbf{k}|). \tag{4.7}$$

4.3 Results and Discussion

4.3.1 Size selectivity

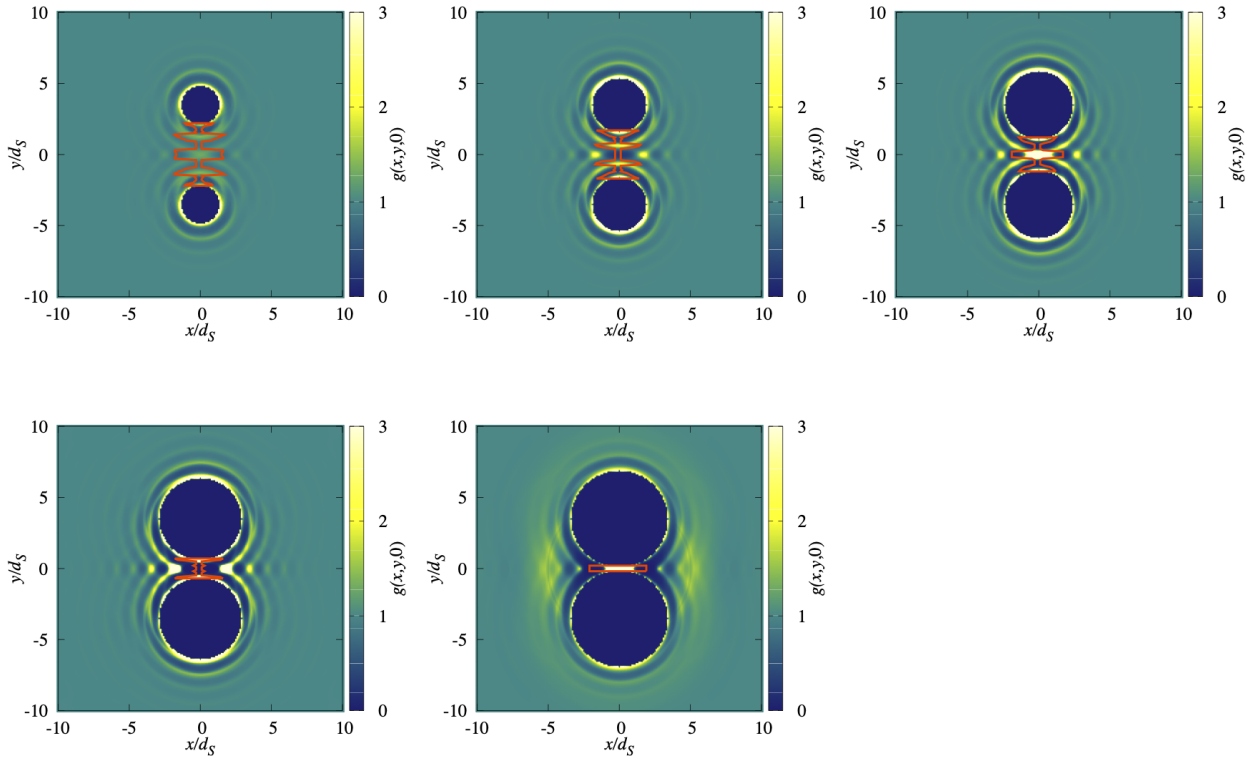


Figure 4.2: The spatial distribution of medium spheres around solute 2 in the xy plane $g_{2M}(x, y, 0)$ calculated by MHNC theory. The region enclosed by the red line is the range from the peak position at $x = 0$ to the minima along the x-axis, and medium spheres in this region were decided to be inclusions.

In this study, a large sphere with a diameter of $5d_S$ was adopted as a guest molecule for the host (Solute 2). According to the lock-and-key relationship, it must be an adequate combination of the host and guest because the guest size is almost the same as the cavity size of the host molecule. The size dependence of the recognition stability is discussed here based on the AO theory. The overlap of excluded volumes of the host and guest becomes maximum at the recognition site when the size fitting is the best, such as the $5d_S$ guest. Thus, it could be estimated that the $5d_S$ guest was very stable at the recognition site of the ring-like host molecule, which had a $5.2d_S$ cavity. Therefore, the $5d_S$ guest was adopted as the guest. The validity was also confirmed

Table 4.2: The number and volume of the existence region of included medium spheres with diameter $d_S, 2d_S, 3d_S, 4d_S, 5d_S$.

	d_S	$2d_S$	$3d_S$	$4d_S$	$5d_S$
n	2.09×10^{-2}	1.38×10^{-2}	1.11×10^{-2}	1.11×10^{-2}	5.41×10^{-2}
v/d_S^3	17.69	10.14	5.04	2.03	0.09

using the OZ–MHNC theory. The spatial distribution of medium spheres (guests) around a host molecule is presented in Fig.4.2. Five sizes of candidate medium spheres (i.e., $1d_S, 2d_S, 3d_S, 4d_S, 5d_S$) were examined. For all guest candidates, the high peaks are found at the recognition site. The host is immersed in a binary mixture. The number density of the medium sphere is $\rho_M = 0.1 \times 10^{-2}$, and the total packing fraction is kept at 0.380. The highest peak was found for the 5ds guest in the spatial distribution function $g_{2M}(x, y, z)$ for medium spheres around the host (where 2 denotes solute 2). The amount of guest molecules around the recognition site was estimated by integration of the spatial distribution in the surrounded volume v (red curves in Fig.4.2). v is the area from the peak position at $x = 0$ to the minima along the x-axis. The amount of recognized guest molecules can be calculated as:

$$n = \rho_M \int_v g_{2M}(x, y, z) dv. \tag{4.8}$$

The amount for the $5d_S$ guest was much larger than others (see Table 4.2), although the volume of the recognition site was the smallest. The selectivity of the $5d_S$ guest was also the highest.

4.3.2 The PMF between host and guest molecules immersed in a one-component system

Figs4.3–4.6 show the PMFs between host and guest molecules immersed in four one-component hard-sphere fluids (systems 1–4). Three different theories, AO, HNC, and MHNC, were adopted in the calculation. The PMF gives interesting information on the stability caused by the recognition and the free-energy barrier in the recognition process. All three theories give the most stable state at the recognition site $(x, y, z) = (0, 0, 0)$. The PMFs have negative values, and the maximum absolute value of each PMF plot is at $x = 0$.

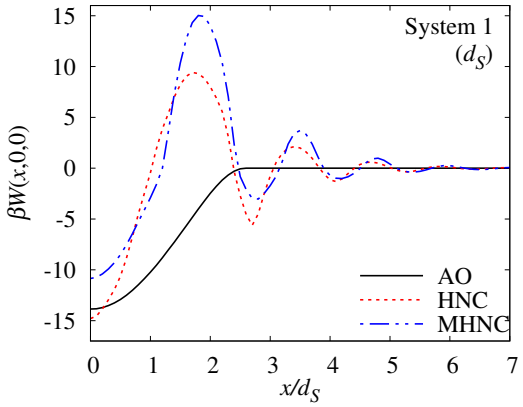


Figure 4.3: The PMF in system 1 between solutes 1 and 2 along the trajectory of solute 1: $x > 0$, $y = 0$, and $z = 0$. The results were obtained by three theories: namely, AO, HNC, and MHNC.

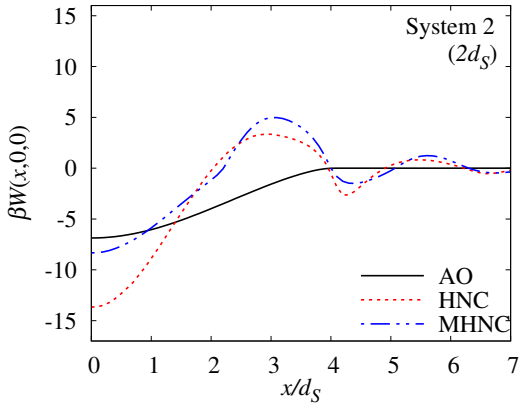


Figure 4.4: The PMF in system 2 between solutes 1 and 2 obtained by three theories: namely, AO, HNC, and MHNC.

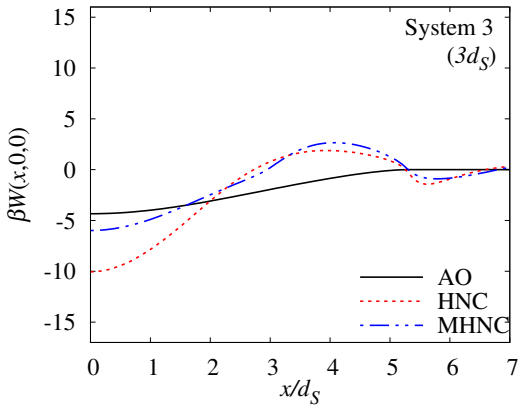


Figure 4.5: The PMF in system 3 between solutes 1 and 2 obtained by three theories: namely, AO, HNC, and MHNC.

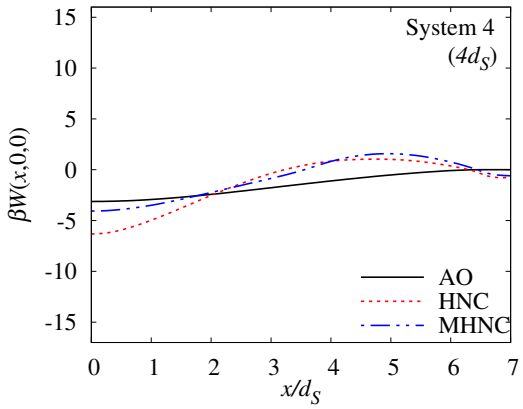


Figure 4.6: The PMF in system 4 between solutes 1 and 2 obtained by three theories: namely, AO, HNC, and MHNC.

The HNC approximation gave the deepest well in the three theoretical results at the most stable location. In the case of the spherical particle system, it is known that the HNC approximation overestimates the peak value in the distribution function [42, 62]. The study in the previous chapter also concluded that the HNC approximation overestimated peak values near the concave surface [56]. The peak in the distribution function corresponded to the minimum in PMF because of the relation $W_{ij}(x, y, z) = -kT \ln g_{ij}(x, y, z)$. Therefore, the present results for the PMFs at $x = 0$ agree well with the literature.

The results shown in the previous chapter suggest the exact value to be located

between the HNC and the MHNC results near the concave surface, and the MHNC result was much closer to the exact value than the HNC result. In the present study, we can expect the MHNC results to be close to the exact value at the recognition site $(x, y, z) = (0, 0, 0)$.

Next, we focused on the PMF oscillation behaviors. Two integral equation theories give the PMFs with oscillation. By contrast, the AO theory cannot predict the free-energy barrier in the recognition process because of the lack of correlation between the solvent particles. Thus, the oscillation behaviors were caused by the correlation between the solvent particles. The correlation is caused by the short-range repulsive interaction between solvent particles when the solvent packing fraction is high. The short-range repulsive interaction is important for the discussion of the free-energy barrier in a high packing-fraction system, such as a liquid.

The first peaks, namely the first barriers, are the highest (see Figs 4.3–4.6.) When the location of the first barrier is in the vicinity of a hard wall, we can expect that the location is $x = \alpha d_S$, where α is constant [77]. However, the location of the first barrier was about the solvent size $+d_S$. For example, the highest barrier was located around $1.8d_S$ for system 1 (the solvent particle size was d_S). In the case of system 4 (the solvent particle size was $4d_S$), the location was about $5d_S$. In the case of the location of the first barrier in the vicinity of a hard wall, the first minimum of the PMF is sharp [77]. By contrast, the present free-energy well around $x = 0$ is broad and gradually increases. It seems that the breadth gives the first barrier position around $x =$ solvent size $+d_S$.

As the solvent diameter increases, the well at $(x, y, z) = (0, 0, 0)$ becomes shallower, and the peaks become lower. In the case of system 4 (i.e., the largest solvent particle), the PMFs were the flattest among the four pure solvent systems (i.e., systems 1–4). These systems had the same packing fraction. Then, the number density reduced as the solvent particle increased. In the framework of the AO theory, the recognition stability is proportional to the number density if the excluded volumes overlap and the temperatures are kept. The dependence of the excluded volume overlap on the solvent size appeared weaker than the number density change. Moreover, the strong localization of a particle caused strong exclusion around the particle.

Then, the first well shallowing causes a decrease in barrier height. Therefore, the PMFs became flatter with the solvent size increasing. A similar discussion was conducted for the PMFs between two hard spheres immersed in a binary hard sphere mixture [77, 80].

4.3.3 The PMF between host and guest molecules immersed in mixture

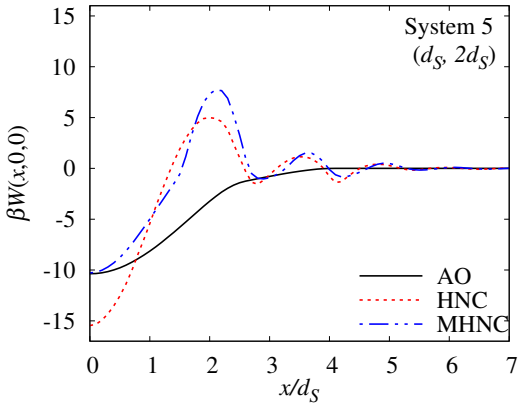


Figure 4.7: The PMF in system 5 between solutes 1 and 2 obtained by three theories: namely, AO, HNC, and MHNC.

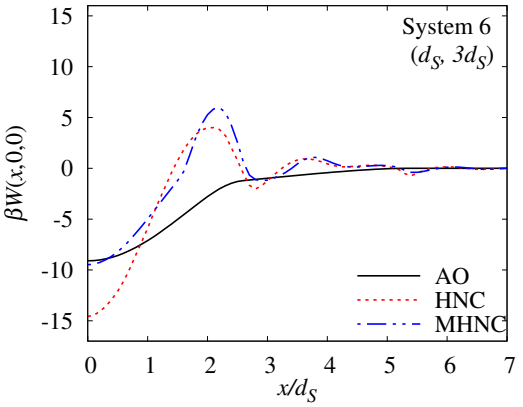


Figure 4.8: The PMF in system 6 between solutes 1 and 2 obtained by three theories: namely, AO, HNC, and MHNC.

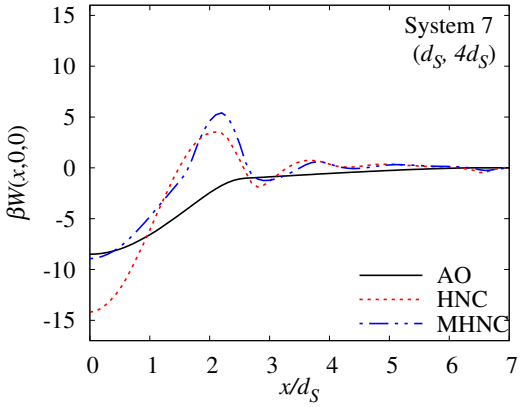


Figure 4.9: The PMF in system 7 between solutes 1 and 2 obtained by three theories: namely, AO, HNC, and MHNC.

The PMF between solute 1 (guest) and solute 2 (host) immersed in a hard-sphere mixture was also investigated. Five mixtures were examined. The AO potential in

the m-component mixture was obtained as:

$$W = \frac{1}{\beta} \sum_{i=1,m} \rho_i \Delta V_i, \tag{4.9}$$

where ρ_i is the number density of the solvent i, and ΔV_i is the overlap of the excluded volume for solvent i. The PMFs for a binary mixture are shown in Figs4.7–4.9. The binary mixtures were prepared as systems 5–7. From the results of the AO theory, the stability at $x = 0$ is lower with increasing diameter of the second component. This phenomenon can be explained based on the number density of the solvent particles. The total number density decreases due to the constant total packing-fraction condition. As discussed in the previous section, this decrease causes the decrease in stability at $x = 0$. These behaviors were also observed in the results obtained by the HNC and MHNC theories.

The results obtained by two integral equation theories are also discussed here. The depth obtained by the MHNC and AO theories were almost the same, whereas that obtained by the HNC theory became larger than others. According to Chapter 3 [56], it seems that the wells obtained by the HNC theory were too deep. The results obtained by the MHNC theory were expected to be almost reasonable. Indeed, there was surprising agreement on depth between the AO and MHNC theories. However, this agreement seemed to be an accidental coincidence because the AO theory does not take into account the correlation between the solvent particles. The lack of correlation caused the absence of a barrier.

Because the first barrier is discussed here, we focus on the results of the MHNC theory. The heights of the barriers for systems 5–7 were smaller than the barrier for system 1. However, the height for these binary solvent systems weakly depended on the diameter of the second component. Moreover, the effect of the second component on the period of PMF oscillation was also weak. The PMF period in systems 5–7 was close to d_S (i.e., the diameter of the first component). Because the number density of the first component was much larger than that of the second component, the effect of the first component on the PMF oscillation period was dominant. As a result of this effect, the heights of the barriers for systems 5–7 were almost the same.

The PMFs in systems 8 and 9 are shown in Figs4.10 and 4.11, respectively. Sys-

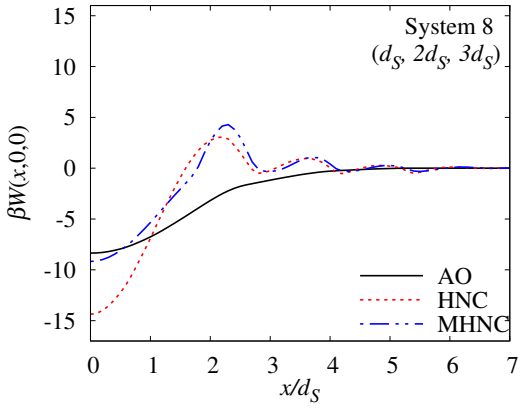


Figure 4.10: The PMF in system 8 between solutes 1 and 2 obtained by three theories: namely, AO, HNC, and MHNC.

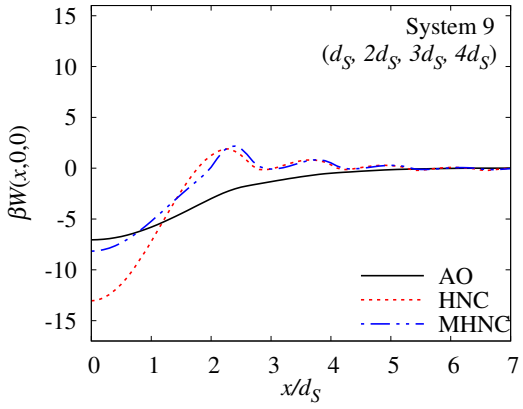


Figure 4.11: The PMF in system 9 between solutes 1 and 2 obtained by three theories: namely, AO, HNC, and MHNC.

tems 8 and 9 have three- and four-component solvent particles, respectively. The free-energy barriers in these multicomponent systems were lower than those in the binary mixtures and are much lower than system 1. As the range of the sphere diameter (i.e., d_S–$3d_S$ and d_S–$4d_S$ in systems 8 and 9, respectively) increases, the barrier becomes lower.

The first well depth and the first barrier height in the two most different systems were compared. The PMF results for systems 1 and 9 are shown in Figs 4.12 and 4.13. System 1 is the simplest case, and system 9 has the most complex solvent mixture. First, the wells' depth (i.e., the recognition stabilizations) are discussed. In the case of the AO theory, the reduction in total number density caused the reduction in stability in system 9. The value for system 9 was almost half (48%) that of system 1.

By contrast, the results for the MHNC theory were different. The reduction was much smaller than that obtained by the AO theory. In the MHNC theory, the stability in system 9 is $-8.16kT$, which is smaller than that in system 1 ($-10.85kT$) only by 25%. This difference is caused by the correlation between solvent particles. In the case of the AO theory, the correlation is ignored, and the reduction in the total number density directory affects the recognition stability. The packing around the host, guest, and surrounding solvents becomes important if correlations between the solvent particles exist. These results are nontrivial phenomena caused by a multi-

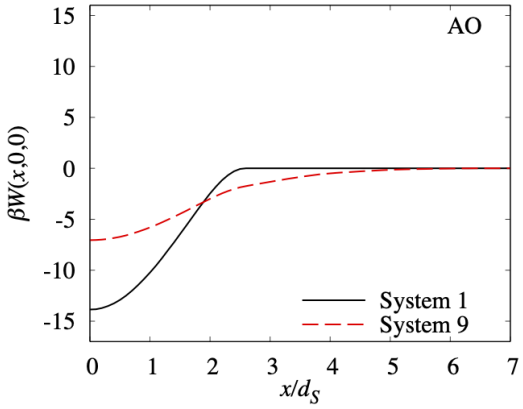


Figure 4.12: The PMF in system 1 and 9 between solutes 1 and 2 obtained by the AO theory.

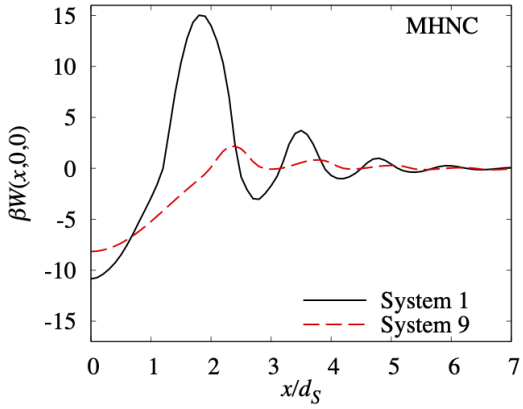


Figure 4.13: The PMF in system 1 and 9 between solutes 1 and 2 obtained by the MHNC theory.

body effect.They must be discussed based on the partial molar volume change in the recognition process. The weak dependence on the first well depth should be useful in the appreciation. Therefore, in our next study, the partial molar volumes will be calculated and the effects will be revealed using the integral equation theory.

Second, the barrier difference is discussed based on the MHNC results. The first free-energy barrier in system 9 was $2.16kT$, a value that is about 48% of that for system 1 ($15.02kT$). This dramatic reduction in the highest barrier in the recognition process is practically useful because solvent mixing is not usually difficult.

The above results for the first barrier on the solvent mixing effect are also non-trivial. Here, we rationalize the solvent mixing effect on the barrier. The spatial distributions of the PMF for solute 1 (guest) around solute 2 (host) on the xy plane are drawn in Figs4.14 and 4.15. The contact surface of solute 2 is surrounded by a negative value layer of the PMF, but the negative value layer is also surrounded by a positive value. These layers repeat again and again, and the amplitude is decreasing. Thus, solute 2 (host) is surrounded by a decreasing oscillation. When we replace the PMF with the spatial distribution of solute 1 (guest), the layers are the decreasing density wave of solute 1, which is generated by solute 2. We can find the interference in both dimensional maps. In these figures, it looks like the interference is caused by the density wave from the upper part of solute 2 and that from the lower part of solute 2. However, the interference is caused by the density waves from all spheres

forming solute 2.

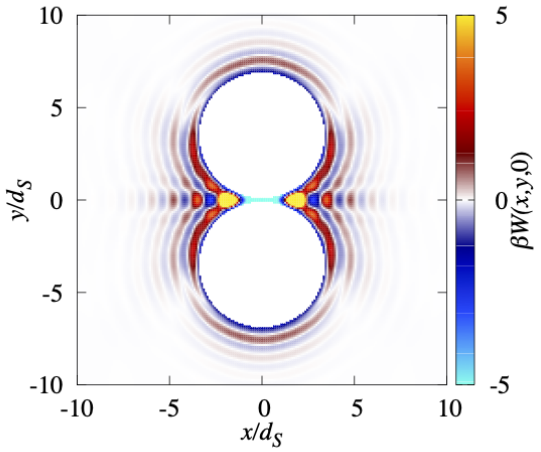


Figure 4.14: PMF on the *xy* plane in system 1 obtained by the MHNC theory.

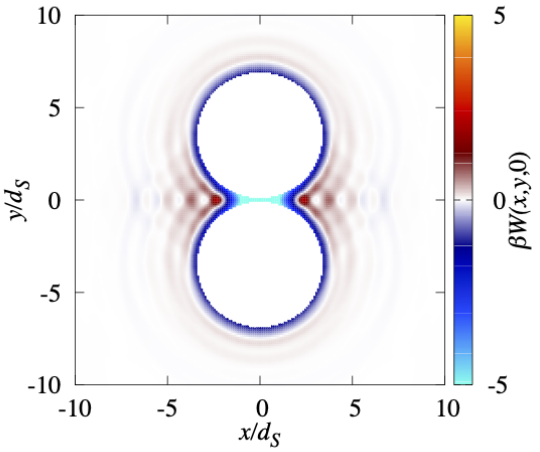


Figure 4.15: PMF on the *xy* plane in system 9 obtained by the MHNC theory.

The interference feature in the PMF wave (or the density wave) is a common feature in Figs4.14 and 4.15. However, the decreasing features differ. The decrease for system 1 (see Fig.4.14) is slower than for system 9 (see Fig.4.15). The density distribution in the solvent mixture has multiple components. For example, the density distribution of solute 1 for system 9 has four components, and the oscillation periods are $1d_S$, $2d_S$, $3d_S$, and $4d_S$. They also interfere with each other, and the starting position of each wave is different (see Fig.4.16). As a result, the PMF wave of solute 2 cannot have a high peak at the first barrier position due to the destructive interference. Therefore, the first barrier for system 9 becomes much lower than

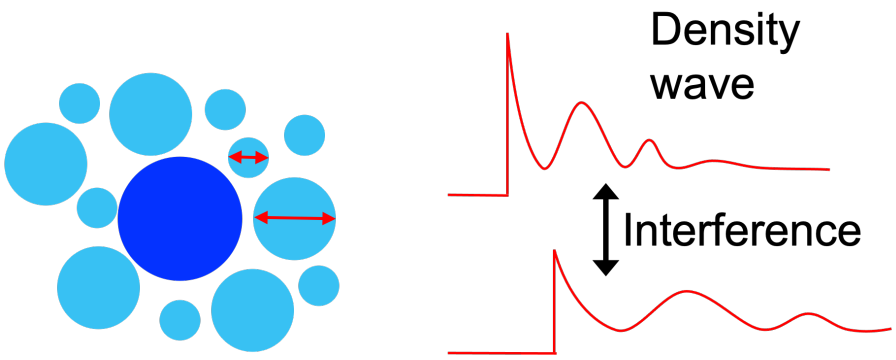


Figure 4.16: Picture illustrating density wave interference in a mixture.

that for system 1. A similar decrease in the oscillation by the interference caused by the multicomponent system is found in the PMF between two spherical solutes immersed in a mixture solvent [80]. In the present study, the interference of waves and the decreasing behavior of the amplitude can be found in these two-dimensional maps.

4.4 Conclusion

The PMFs between a cyclic model molecule and a large sphere in a multicomponent mixture of smaller spheres were investigated using the AO, OZ–HNC, and OZ–MHNC theories. The AO theory is the simplest theory to describe the depletion interaction. The interaction between small spheres is ignored in the AO theory; therefore, the predicted depletion interactions do not oscillate. In contrast, the real entropic force is oscillatory and possesses the free-energy barrier, which plays an important role. Nine systems were prepared and the stabilities of the recognition and the free-energy barrier in the recognition process were obtained.

The PMFs in the mixtures differed significantly from those in a one-component system. The oscillatory structure of the depletion interaction in a multicomponent system was lower than in a one-component system, and the free-energy barrier for the association process in a multicomponent system was reduced. The strongest reduction of the free-energy barrier was observed in a multicomponent mixture of smaller spheres with varying diameters, with the same packing-fraction value of each component. This feature was similar to the previous study of spherical solutes [80]. When the viscosities of the solvents are similar to each other, this could mean a reduction in the time scale in which the guest molecules reach a recognition site of host molecules. Furthermore, the PMF in the multicomponent mixture has a short-range attraction in the cavity, which induces sufficient stability for molecular recognition. In the case of spherical solute particles, the stabilities were reduced when the barrier heights were reduced. By contrast, in the present study, the stabilities were maintained when the barrier heights were reduced. These results showed that the free-energy barrier could be controlled while the recognition maintained sufficient stability by choosing an appropriate multicomponent mixture.

Chapter 5

Summary

In this thesis, an analysis of molecular recognition was described using 3D integral equation theories combined with grand canonical MC (GCMC) simulations. In the analysis, hard-body particle models were adopted to discuss the effect of the translational motion of solvent molecules on the effective interaction. This study is summarized as follows.

In Chapter 2, simulation studies were performed to assess the integral equation theories for molecular recognition because molecular recognition studies need concentration dependence, and molecular simulations for concentration dependence are usually difficult to perform. The spatial distribution function around a contact dimer was calculated using 3D OZ equations coupled with closures in which PY and HNC approximations were used or in which the MHNC bridge functions proposed by Kinoshita were incorporated. The results were compared with those from GCMC simulations because the grand canonical ensemble was adopted in the formulation of integral equation theories for liquid. The spatial distribution functions obtained by the 3D–MHNC–OZ theory were much more accurate than those obtained by other theories. The advantage of the 3D–MHNC–OZ theory was maintained even when the solute particle also had concave surfaces. Here, it should be stressed that although the PY approximation was believed to be more adequate in the case of a hard-particle fluid, the MHNC approximation was found to be much better than the former. Therefore, the 3D–MHNC–OZ theory is a powerful tool in systems that have strong concavity, such as molecular recognition, when the hard-body particle model was adopted. Analyses on the triplet distribution function were also per-

formed. It was found that in terms of the triplet distribution function, the use of the 3D–MHNC–OZ theory does not necessarily lead to improvement. This means that the 3D–MHNC–OZ theory shows high accuracy because of some cancellations. However, the details have not been clarified.

In Chapter 3, based on the results from the previous chapter, the 3D–MHNC–OZ theory was adopted for the molecular recognition study in a multicomponent mixture. The PMFs between a cyclic host molecule and a spherical guest molecule were calculated. The PMF in the mixture, in particular the first peak, was significantly different from that in a one-component system of small spheres. The PMF in a one-component hard-sphere solvent has an oscillatory structure. By contrast, the oscillation structure almost disappeared in a multicomponent system. This difference is related to the reduction of the first barrier height in the recognition process. Moreover, the stability with recognition in a multicomponent system of small spheres whose diameters and the mixing ratios were suitably chosen was large enough. These results suggested that the cosolvent can regulate the stability and the recognition barriers. The interference of the density waves can explain the disappearance of the oscillation structure in the effective interaction.

Acknowledgement

I would like to express my sincere gratitude to Prof. Ryo Akiyama for his support and warm encouragement during the course of my study. His advice and comments were an enormous help to me. Special thanks to Prof. Ayumi Suematsu of Nishinippon Institute of Technology for constructive comments and suggestions on my study.

I would also like to thank Prof. Masahiro Kinoshita of Kyoto University for the software program used in the numerical preparation to obtain the distribution functions. I would like to express my appreciation to Prof. Yuka Nakamura of Nigata University for their valuable comments and discussion on my study. I would also like to thank my colleagues of "Biophysics and Chemical Physics and Biophysics Group" at Kyushu University for their friendships and support. Finally, I am deeply grateful to my family for their warm support and encouragement.

This work was supported by JST, the establishment of university fellowships towards the creation of science technology innovation, Grant Number JPMJFS2132. The computation was performed using the Research Center for Computational Science, Okazaki, Japan and the Research Institute for Information Technology, Kyushu University.

Bibliography

[1] M. Kinoshita, Biophys. Rev. **5**, 283 (2013).

[2] B. Alberts *et al.*, *Molecular Biology of the Cell*, 7th ed. (Garland science, 2017).

[3] D. Chandler and J. D. Weeks, Phys. Rev. Lett. **25**, 149 (1970).

[4] D. Chandler, J. D. Weeks, and H. C. Andersen, Science **220**, 787 (1983).

[5] J. D. Weeks and D. Chandler, J. Chem. Phys. **54**, 5237 (1971).

[6] B. Widom, Science **157**, 375 (1967).

[7] J. D. van der Waals, *On the Continuity of Gaseous and Liquid States* (Dover Publications, New York, 2004).

[8] H. C. Longuet-Higgins and B. Widom, Mol. Phys. **8**, 549 (1964).

[9] M. Nishino, M. Hirota, and Y. Umezawa, *The CH/π Interaction : Evidence, Nature, and Consequences*(*Wiley − VCH*, 1998).

[10] M. Watanabe, M. Nishiyama, T. Yamamoto, and Y. Koie, Tetrahidron **56**, 741 (2000).

[11] Y. Nakagawa, K. Irie, R. C. Yanagita, H. Ohigashi, and K. Tsuda, J. Am. Chem. Soc. **127**, 5746 (2005).

[12] Y. Umezawa and M. Nishio, Biopolymers **79**, 248 (2005).

[13] H. Ohtaka, A. Schon, and E. Freire, Biochemistry **42**, 13659 (2003).

[14] J. Kardos, K. Yamamoto, K. Hasegawa, H. Naiki, and Y. Goto, J. Biol. Chem. **279**, 55308 (2004).

[15] M. Kinoshita, J. Chem. Phys. **116**, 3493 (2002).

[16] M. Kinoshita and T. Oguni, Chem. Phys. Lett. **351**, 79 (2002).

[17] K. Amano and M. Kinoshita, Chem. Phys. Lett. **488**, 1 (2010).

[18] K. Amano and M. Kinoshita, Chem. Phys. Lett. **504**, 221 (2011).

[19] M. L. Bender and M. Komiyama, *Reactivity and Structure Concepts in Organic Chemistry* (Springer, Berlin, Germany, 1978).

[20] J. Szejtli, *Cyclodextrins and Their Inclusion Complexes* (Akadémiai Kiadó: Budapest, Hungary, 1982).

[21] J. Szejtli, T. Osa, and Eds, *Cyclodextrins* (Pergamon: Oxford, U.K., 1996).

[22] C. J. Easton and S. F. Lincoln, *Modified Cyclodextrins: Scaffolds and Templates for Supramolecular Chemistry* (Imperial College Press: London, U.K., 1999).

[23] T. Loftsson and M. Brewster, J. Pharm. Sci. **85**, 1017 (1996).

[24] T. Kakuta *et al.*, Adv. Mater. **25**, 2849 (2013).

[25] J. M. Belitsky, A.Nelson, and J. F. Stoddart, Org. Biomol. Chem. **4**, 250 (2006).

[26] J. Szetjli, Chem. Rev. **98**, 1743 (1998).

[27] D. A. Mcquarrie, *Statistical Mechanics* (Harper and Row, New York, 1976).

[28] D. Chandler, *Introduction to Modern Statistical* (Oxford University Press, Oxford, UK, 1987).

[29] W.Luck, M. Klier, and H. Wesslau, Ber. Bunsenges. Phys. Chem. **67**, 75 (1963).

[30] S. Hachisu, Y. Kobayashi, and A. Kose, J. Colloid Interface Sci. **42**, 342 (1973).

[31] S. Hachisu and Y. Kobayashi, J. Colloid Interface Sci. **46**, 470 (1974).

[32] E. Winger, Phys. Rev. **46**, 11 (1934).

[33] E. Winger, Trans. Faraday. Soc. **34**, 678 (1938).

[34] B. J. Alder, H. G. Hoover, and D. A. Young, J. Chem. Phys. **49**, 3988 (1968).

[35] M. Wadati and M. Toda, J. Phys. Soc. Japan. **32**, 1147 (1972).

[36] J. G. Kirkwood, J. Chem. Phys. **7**, 919 (1939).

[37] B. J. Alder and T. Wainwright, J. Chem. Phys. **127**, 459 (1959).

[38] B. J. Alder and T. Wainwright, Phys. Rev. **127**, 359 (1962).

[39] S. Asakura and F. Oosawa, J. Chem. Phys. **22**, 1255 (1954).

[40] S. Asakura and F. Oosawa, J. Polym. Sci. **33**, 183 (1958).

[41] A. Vrij, Pure and applied chemistry **48**, 471 (1976).

[42] J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, 2nd ed. (Academic Press, Lodon, 1986).

[43] J. P. Hansen and G. Zerah, Phys. Lett. A **108**, 277 (1985).

[44] M. Kinoshita, Chem. Phys. Lett. **353**, 259 (2002).

[45] R. Roth, B. Götzelmann, and S. Dietrich, Phys. Rev. Lett. **83**, 448 (1999).

[46] O. G. Berg, Biopolymers **30**, 1027 (1990).

[47] D. Hall and A. P. Minton, Biochim. Biophys. Acta **1649**, 127 (2003).

[48] R. Akiyama, Y. Karino, H. Obama, and A. Yoshifuku, PCCP **12**, 3096 (2010).

[49] L. Verlet, Mol. Phys. **41**, 183 (1980).

[50] D. M. Duh and D. Henderson, J. Chem. Phys. **104**, 6742 (1996).

[51] M. Kinoshita, J. Chem. Phys. **118**, 8969 (2003).

[52] M. Kinoshita and T. Hayashi, J. Molec. Liq. **247**, 403 (2017).

[53] D. Frenkel and B.Smit, *Underatanding Molecular Simulation From Algorithms to Applications* (Academic Press, San Diego, 1996).

[54] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford Science, Oxford, 1987).

[55] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[56] M. Matsuo, Y. Nakamura, M. Kinoshita, and R. Akiyama, 2023, arXiv:2311.05893.

[57] Y. Rosenfeld, Phys. Rev. Lett. **63**, 980 (1989).

[58] R. Roth, R. Evans, A. Lang, and G. Kahl, J. Phys: Condens. Matter **14**, 12063 (2002).

[59] Y. Yang-Xin and W. Jianzhong, J. Chem. Phys. **117**, 10156 (2002).

[60] P. Attard and G. N. Patey, J. Chem. Phys. **92**, 4970 (1990).

[61] M. Kinoshita, S. Iba, K. Kuwamoto, and M. Harada, J. Chem. Phys. **105**, 7177 (1996).

[62] Y. Nakamura, S. Arai, M. Kinoshita, A. Yoshimori, and R. Akiyama, J. Chem. Phys. **151**, 044506 (2019).

[63] J. R. Schmidt and J. L. Skinner, J. Chem. Phys. **119**, 8062 (2003).

[64] R. O. Sokolovskii, M. Thachuk, and G. N. Patey, J. Chem. Phys. **125**, 204502 (2006).

[65] D. Frenkel and A. A. Louis, Phys. Rev. Lett. **68**, 3363 (1992).

[66] D. J. Adams, Mol. Phys. **28**, 1241 (1974).

[67] D. J. Adams, Mol. Phys. **29**, 307 (1975).

[68] I. Nezbeda and J. Kolafa, Mol. Simulat. **5**, 391 (1991).

[69] Y. Kubota and R. Akiyama, J. Phys. Soc. Jpn. **81**, SA017 (2012).

[70] Y. Uehara, Y. T. Lee, T. Ree, and T. H. Ree, J. Chem. Phys. **70**, 1884 (1979).

[71] B. J. Alder, Phys. Rev. Lett. **12**, 317 (1964).

[72] J. G. Kirkwood, J. Chem. Phys. **3**, 300 (1935).

[73] P. Attard and G. Stell, Chem. Phys. Lett. **189**, 128 (1992).

[74] B. Bildstein and G. Kahl, J. Chem. Phys. **100**, 5882 (1994).

[75] F. J. Rogers and D. A. Young, Phys. Prev. A **30**, 999 (1984).

[76] T. Hayashi, H.Oshima, Y.Harano, and M.Kinoshita, J. Phys: Condens. Matter **28**, 344003 (2016).

[77] R. Akiyama, Y. Karino, Y. Hagiwara, and M. Kinoshita, J. Phys. Soc. Jpn. **75**, 064804 (2006).

[78] Y. Karino and R. Akiyama, Chem. Phys. Lett. **478**, 180 (2009).

[79] R. Roth, Y. Harano, and M. Kinoshita, Phys. Rev. Lett. **97**, 078101 (2006).

[80] R. Roth and M. Kinoshita, J. Chem. Phys. **125**, 084910 (2006).