

# **Pair correlation functions of simple solutes in a Lennard-Jones solvent\***

**Arieh Ben-Naim\*\* and Vincent K. Shen**

**Physical and Chemical Properties Division  
National Institute of Standards and Technology  
100 Bureau Drive MS 8380  
Gaithersburg, Maryland 20899-8380  
U.S.A.**

## Abstract:

We have calculated the pair correlation functions for several binary mixtures composed of simple solutes in a Lennard-Jones solvent. In particular, we have studied the solute-solute pair correlation functions and their dependence on the total density, the solvent Lennard-Jones parameters and on the solute-solute energy parameter. All the results were obtained from solving the Percus-Yevick equations, as well as from Monte Carlo simulations. The relevance of these results to the problem of hydrophobic interactions is also discussed.

\*\* Present address:

Department of Physical Chemistry  
The Hebrew University of Jerusalem  
Givat Ram, Jerusalem  
Israel 91904

\* Contribution of the National Institute of Standards and Technology, not subject to U.S. copyright.

## 1. Introduction

In the course of studying the underlying molecular origins of the hydrophobic interaction<sup>1,2</sup>, one of the authors (ABN) examined the potential of mean force (*PMF*) in a one dimensional system composed of a hard-rod solute in two types of solvents. The first was a simple solvent consisting of particles interacting via a square-well potential, and the second was a water-like solvent consisting of particles interacting via hydrogen-bond-like potential functions.<sup>3-5</sup>

The two solvents showed very different behavior in their pure states; one behaved “normally” while the second exhibited water-like behavior. Interestingly, the solubility and the *PMF* of a hard-rod (*HR*) solute in the two solvents could be made very similar in magnitude by adjusting the strength of the solvent-solvent interaction. It was therefore suspected that the strength of the so-called hydrophobic interaction (i.e., the *PMF* between two simple solutes in water) is not a result of the peculiarities of the structure of water but a result of the strength of the solvent-solvent interaction.

To test this conjecture, we have carried out a series of calculations of the pair correlation function (*PCF*) of a simple solute in a Lennard-Jones (*LJ*) solvent. We have examined the dependence of the *PCF* on both the strength of the intermolecular interaction as well as the size of the solvent particles. In the following section, we describe very briefly the methods used to calculate the *PCF*. A sample of results is presented in section 3, and some conclusions on the relevance of these results to the problem of hydrophobic interactions are discussed in section 4.

## 2. Methods of calculating the pair correlation function

### 2.1 Numerical solution to the Percus-Yevick equations

All calculations were done for binary mixtures. The solvent *B* consisted of particles interacting through pair potentials of the form

$$U_{BB}(R) = 4\varepsilon_{BB} \left[ \left( \frac{\sigma_{BB}}{R} \right)^{12} - \left( \frac{\sigma_{BB}}{R} \right)^6 \right] \quad (2.1)$$

where  $\varepsilon$  represents the strength of the inter particle attraction and  $\sigma$  represents the particle diameter.

The solute  $A$  was treated as either hard sphere particles with interaction

$$U_{AA}(R) = \begin{cases} \infty & R \leq \sigma_{BB} \\ 0 & R > \sigma_{BB} \end{cases} \quad (2.2)$$

or  $LJ$  particles with interactions as in (2.1). Interaction parameters involving unlike species were obtained using conventional Lorentz-Berthelot mixing rules <sup>6</sup>

$$\sigma_{AB} = \sigma_{BA} = (\sigma_{AA} + \sigma_{BB}) / 2 \quad (2.3)$$

$$\varepsilon_{AB} = \varepsilon_{BA} = (\varepsilon_{AA} \varepsilon_{BB})^{1/2} \quad (2.4)$$

The Percus Yevick ( $PY$ ) equation for pure spherical particles has the form <sup>7-10</sup>

$$\begin{aligned} y(\bar{R}_1, \bar{R}_2) = 1 + \rho \int_V y(\bar{R}_1, \bar{R}_2) f(\bar{R}_1, \bar{R}_3) \\ \times [y(\bar{R}_2, \bar{R}_3) f(\bar{R}_2, \bar{R}_3) + y(\bar{R}_2, \bar{R}_3) - 1] dR_3 \end{aligned} \quad (2.5)$$

where  $f(R)$  is the Mayer function defined as

$$f(R) = \exp[-\beta U(R)] - 1 \quad (2.6)$$

and  $y(R)$  is defined as

$$y(R) = g(R) \exp[\beta U(R)] \quad (2.7)$$

A simpler and more useful form of this equation is obtained by transforming to bipolar coordinates

$$u = |\bar{R}_1 - \bar{R}_3|, \quad v = |\bar{R}_2 - \bar{R}_3|, \quad R = |\bar{R}_1 - \bar{R}_2| \quad (2.8)$$

In bipolar coordinates, the element of volume is expressed as

$$dR_3 = 2\pi uv \, du \, dv / R \quad (2.9)$$

and thus (2.5) can be transformed into

$$y(R) = 1 + 2\pi\rho R^{-1} \int_0^\infty y(u) f(u) u \, du \int_{|\bar{R}-u|}^{R+u} [y(v) f(v) + y(v) - 1] v \, dv \quad (2.10)$$

For numerical purposes, it is convenient to transform again the *PY* equation (2.10) by defining the function<sup>1,2,7,8</sup>

$$z(R) = y(R)R \quad (2.11)$$

Hence, an integral equation for  $z(R)$  can be obtained

$$z(R) = R + 2\pi\rho \int_0^\infty z(u)f(u)du \int_{|R-u|}^{R+u} [z(v)f(v) + z(v) - v]dv \quad (2.12)$$

For a binary mixture, equation (2.12) is generalized to

$$z_{\alpha\beta}(R) = R + \sum_{\gamma=A,B} 2\pi\rho_\gamma \int_0^\infty z_{\alpha\gamma}(u)f_{\alpha\gamma}(u)du \int_{|R-u|}^{R+u} [z_{\gamma\beta}(v)f_{\gamma\beta}(v) + z_{\gamma\beta}(v) - v]dv \quad (2.13)$$

where  $\alpha\beta = AA, AB$ , or  $BB$ . The numerical procedure used to solve the multi-component *PY* equation is similar to that used to solve the pure-fluid variant. One starts with

$$z_{\alpha\beta}(R) = R \quad (2.14)$$

for all four functions  $z_{\alpha\beta}(R)$  and proceeds to solve the four integral equations (2.13) by iteration. [For more details, see references (1) and (2)]. Once the functions  $z_{\alpha\beta}(R)$  are obtained, the pair correlation functions can be calculated using equations (2.7) and (2.11).

## 2.2 Monte Carlo simulations

Pair correlation functions were also calculated using Monte Carlo (*MC*) simulations in the canonical ensemble, that is at fixed  $N$  (number of particles),  $V$  (system volume), and  $T$  (temperature). For convenience, the quadratically shifted-force version of the Lennard-Jones interaction was used in the simulations.<sup>11</sup> The cutoff distance  $R_c$  was taken to be  $2.5\sigma_{\alpha\beta}$  (where  $\alpha\beta = AA, BB$ , or  $AB$ ). In all cases, a minimum of 300 particles was used. Since the simulations were performed in the canonical ensemble, only random trial displacements were employed. Cluster moves were not employed in the mixtures involving large size asymmetry. Instead, multiple independent simulations (a minimum of five in the case of large size asymmetry, and a minimum of three in other cases) were performed. The reported *PCFs* represent the mean of the independent runs. The standard deviation gives an indication of how well configuration space is explored. Each

simulation consisted of an equilibration period of at least 500,000,000 trial moves and a subsequent production period of at least 5 billion moves. Maximum trial displacements for each species were adjusted to yield an acceptance rate of 0.40. In addition, systems composed of 500 and 1000 particles were investigated to verify that system size effects were not important. Within the uncertainties, which are taken as the standard deviation of at least three independent runs, system size effects were negligible.

### **3. Results from the Percus Yevick (PY) equations and the Monte Carlo simulations**

In this section, we present a sample of the results obtained from (a) solution of the *PY* equations and (b) by Monte Carlo simulations. We have carried out four sets of calculations for a mixture composition of  $x_A = 0.1$  (where  $x_A$  is the solute mole fraction) and temperature  $T = 1$ . These are discussed in the following subsections:

#### **3.1 Dependence on the solvent size parameter**

The first series consists of results obtained for a fixed solute diameter  $\sigma_{AA} = 1$  while varying the solvent diameter  $\sigma_{BB}$ . The other parameters for these calculations are

$$\varepsilon_{AA} / k_B T = \varepsilon_{BB} / k_B T = 1, \text{ and } \eta = 0.4 \quad (3.1)$$

where  $\eta$  is the packing fraction, defined as

$$\eta = \frac{\pi}{6} (\rho_A \sigma_{AA}^3 + \rho_B \sigma_{BB}^3) \quad (3.2)$$

where  $\rho_i$  is the number density of species  $i$ .

Figures 1 and 2 show the *PCF*'s calculated by *PY* equations for  $\sigma_{BB} = 1$  to 5 and  $\sigma_{BB} = 6$  to 10, respectively. Figure 3 shows  $g_{AA}(R)$  in the region near the first maxima. As can be seen from these figures, the height of the first peak of  $g_{AA}$  increases with solvent size  $\sigma_{BB}$  while the heights of the first peaks of either  $g_{AB}$  and  $g_{BB}$  remain relatively constant. In Figure 4, the heights of the first and second peaks of the *PCF*'s predicted by theory and simulation are compared. The predictions of the two methods are in good agreement.

### 3.2 Dependence on packing fraction $\eta$

In this series of calculations, the packing fraction  $\eta$  is varied while fixing the following parameter values

$$\begin{aligned}\sigma_{AA} &= 1, & \sigma_{BB} &= 6 \\ \varepsilon_{AA} / k_B T &= \varepsilon_{BB} / k_B T = 1\end{aligned}\quad (3.3)$$

We vary  $\eta$  between 0.01 and 0.5. Figures 5 and 6 show the results computed by using the *PY* equations. In both the pure *LJ* liquid and *LJ* mixtures,<sup>1,2,12</sup> it is well known that the first peak increases with increasing  $\eta$ . The interesting finding in our calculations is that the height of the first peak in  $g_{AB}$  almost does not change, while that of  $g_{BB}$  changes only slightly, with increasing  $\eta$ . In contrast, the height of the first peak in  $g_{AA}$  changes quite dramatically with  $\eta$ . Figure 6 shows the details of  $g_{AA}$  around the first maxima. Figure 7 compares the locations of the first and second peaks of the *PCF* obtained by the two methods.

### 3.3 Dependence on $\varepsilon_{AA}$

Figure 8 shows a series of results with the following fixed parameters

$$\begin{aligned}\sigma_{AA} &= 1, & \sigma_{BB} &= 6 \\ \varepsilon_{BB} / k_B T &= 1, & \eta &= 0.4,\end{aligned}\quad (3.4)$$

while varying  $\varepsilon_{AA}$ . Note that in this case where  $\varepsilon_{AA}=0$ , the solute is treated as a hard sphere. Again, as expected, the solvent *PCF*,  $g_{BB}$ , does not change significantly, indicating that the solvent is not affected by the solute-solute interactions. On the other hand,  $g_{AA}$  changes considerably with  $\varepsilon_{AA}$ . A comparison of the peak heights predicted by the *PY* equations and *MC* simulation is provided in Figure 9.

### 3.4 Dependence on $\varepsilon_{BB}$

In the last series of calculations, the influence of the strength of solvent-solvent interaction is investigated. Other relevant parameter values were held fixed

$$\sigma_{AA} = \sigma_{BB} = 1,$$

$$\begin{aligned}\varepsilon_{AA}/k_B T &= 0, \\ \eta &= 0.4\end{aligned}\tag{3.5}$$

Notice that this corresponds physically to a hard-sphere solute in a LJ solvent. In the case where  $\varepsilon_{BB} = 0$ , the solvent is also a hard-sphere. The predictions of the *PY* equations are shown in Figure 10. Several important aspects of these results should be noted. First, the function  $g_{AB}$  is relatively insensitive to the strength of the solvent-solvent interaction. On the other hand, the function  $g_{BB}$  changes as expected for a normal liquid<sup>10</sup>, i.e., all the peaks increase with  $\varepsilon_{BB}$ . These results are similar to the results obtained in one dimensional systems.<sup>3,3,14</sup>

The most interesting finding is the behavior of  $g_{AA}$ . With increasing  $\varepsilon_{BB}$ , the height of the first peak of  $g_{AA}$  increases even more than the height of the first peak of  $g_{BB}$ . More important is the *range* of the solute-solute correlation, which grows with  $\varepsilon_{BB}$ .

Figure 11 shows the comparison with the *MC* results. We suspect that the disagreement between the two sets of results in this case is due to a phase separation. To clarify this point, we calculated the finite Kirkwood-Buff integrals defined by

$$G_{ij}(R_M) = \int_0^{R_M} [g_{ij}(R) - 1] 4\pi R^2 dR \tag{3.6}$$

Figure 12 shows that for small values of  $\varepsilon_{BB}$ , the function  $G_{ij}(R_M)$  behaves “normally”, i.e., as expected.<sup>10</sup> However, for larger values of  $\varepsilon_{BB}$  the functions  $G_{ij}(R_M)$  do not seem to converge as  $R_M$  increases. We also calculated the parameter

$$\Delta_{AB} = G_{AA} + G_{BB} - 2G_{AB} \tag{3.7}$$

Although we do not have the Kirkwood-Buff integrals (i.e.,  $G_{ij}(R_M \rightarrow \infty)$ ) the values of  $G_{ij}$  at  $R_M = 10$  as well as  $\Delta_{AB}$ , as reported in Table 1 indicates that something grows “wild” as we increase  $\varepsilon_{BB}$ . In Figure 13, we show a phase-separated configuration for  $\varepsilon_{BB} = 1.1$ .

#### 4. Discussion

All the results reported in the previous sections in terms of the *PCF* may be reinterpreted in terms of the potential of mean force (*PMF*). The relationship between the two quantities is

$$W_{ij}(R) = -k_B T \ln g_{ij}(R) \quad (4.1)$$

Since we are interested in the solute-solute *PMF*, and since the direct solute-solute interaction is kept constant in each of the series of results, we can reinterpret the results in terms of the solvent induced interaction, i.e., we write

$$W_{AA} = U_{AA} + \delta A_{AA} \quad (4.2)$$

where  $U_{AA}$  is the direct solute-solute interaction and  $\delta A_{AA}$  is the solvent induced part of the *PMF*. The quantity  $\delta A_{AA}$  was studied extensively in connection with the problem of hydrophobic interactions.<sup>1,2</sup>

There are several conclusions that can be drawn from the results of this work. The first is “technical”; that is the good agreement between the *PY* and the *MC* results proves once again the reliability of the *PY* equations. Even in the case where the results disagree, we can trace the origin of the discrepancy, and learn something about the applicability of the *PY* approximation in regions near phase separation.

The second conclusion is similar to the one drawn in a previous article (Ref. 3). As in reference 3, we found that in a one-dimensional system of mixtures that the correlation can be made large, or the solute-solute *PMF* can be made deeper by varying the solvent-solvent interactions. These findings strongly indicate that the solute-solute *PMF* for non-polar solutes in real water is basically a result of the *strength* of the solvent-solvent interactions. In real water, these interactions are due to hydrogen-bonding. However, it seems that hydrogen-bonding *per se*, or the peculiarities of the structure of water, are not necessary for the manifestation of strong solute-solute correlation, as is currently believed to be the case for so-called hydrophobic interactions.



As was also pointed out in reference 3, the extended *range* of the solute-solute *PCF* in a solvent with strong intermolecular interactions (or at low temperatures) is also not a phenomenon relevant to the hydrophobic interaction. In our case, it is probably due to the fact that the mixture becomes less symmetrical and less ideal, and perhaps approaching the region of phase separation.

**Table 1**  
**The values of  $G_{ij}$  ( $R_M = 10$ ) and  $\Delta_{AB}$  for the series of calculations in section 3.4**

	$\epsilon_{BB} = 0$	$\epsilon_{BB} = 0.5$	$\epsilon_{BB} = 1.0$	$\epsilon_{BB} = 1.5$
$G_{AA}$	-1.09	10.36	414.8	1843
$G_{AB}$	-1.09	-3.27	-84.17	-443
$G_{BB}$	-1.09	-0.67	15.53	105
$\Delta_{AB}$	0	16.23	598.7	2834

## References

1. A. Ben-Naim, *Water and Aqueous Solutions*, Plenum Press, New York, (1974)
2. A. Ben-Naim, *Water and Aqueous Solutions*, World Scientific, Singapore (In press 2009)
3. A. Ben-Naim, *J. Chem. Phys.* (submitted 2008)
4. A. Ben-Naim, *J. Chem. Phys.* **128**, 024505 (2008)
5. A. Ben-Naim, *J. Chem. Phys.* **128**, 024506 (2008)
6. M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press, New York, (1989)
7. J.K. Percus, and G.J. Yevick, *Phys. Rev.*, **110**, 1 (1958)
8. E.W. Grundke, and D.Henderson, *Mol. Phys.* **24**, 269, (1972)
9. A.A. Broyles, *J. Chem. Phys.* **33**, 456 (1960) and **35**, 493(1961)
10. G. J.Throop and R. J. Bearman, *J. Chem. Phys.* **42**, 2838 (1965) and **44**, 1423 (1966)
11. S. D. Stoddard and J. Ford, *Phys. Rev. A* **8**, 1504 (1973).
12. A. Ben-Naim, *Molecular Theory of Solutions*, Oxford Press Univ. Press, Oxford (2006)
13. A. B. Kolomeisky and B. Widom, *Faraday Disc.* **112**, 81 (1999)
14. B. Widom, *Polish J. Chem.* **75**, 507 (2001)

## Figure captions

Figure 1: The pair correlation functions (*PCF*) for a mixture of a solute  $A$  in a solvent  $B$  predicted by the *PY* equations. The parameters for this series are provided in equations 3.1.

Figure 2: Same as in Figure 1 but for higher values of the solvent diameters  $\sigma_{BB}$ .

Figure 3: The same as in Figure 1 and 2 but a close up view near the first peak of the *PCF*.

Figure 4: Comparison of the locations of the first and second peaks of  $g_{AA}$  obtained by solution of the Percus Yevick equations (diamonds) and by Monte-Carlo simulations (circles). Error bars represent the standard deviation of at least three independent runs.

Figure 5: Dependence of the *PCFs* on the packing fraction. Parameters for this series are provided in equations 3.3.

Figure 6: As in Figure 5, but a close up view near the first peak.

Figure 7: Comparison of the results from *PY* (diamonds) and *MC* (circles) for the first and second peak of  $g_{AA}$  and their dependence on  $\eta$ . Error bars represent the standard deviation of at least three independent runs.

Figure 8: Dependence of the *PCF* on the solute-solute energy parameter  $\varepsilon_{AA}$ . The other parameters for this series are provided in equations 3.4.

Figure 9: Comparison of the results from the *PY* (diamonds) and the *MC* (circles) for the first and second peaks of  $g_{AA}$  and the dependence on  $\varepsilon_{AA}$ . Error bars represent the standard deviation of at least three independent runs.

Figure 10: Dependence of the *PCF* on the solvent-solvent energy parameters  $\varepsilon_{BB}$ . The other parameters for this series are provided in equations 3.5.

Figure 11: Comparison of the results from the *PY* (diamonds) and *MC* (circles) for the first and second peaks of  $g_{AA}$  and their dependence on  $\varepsilon_{BB}$ . For clarity, the upper panel corresponds to the height of the first peak in  $g_{AA}$ , and the lower panel corresponds to the height of the second peak. Error bars represent the standard deviation of at least three independent runs.

Figure 12: The finite Kirkwood-Buff integrals for the same series as in Figure 10. The four columns correspond to the four values of  $\epsilon_{BB}$  (0,0.5,1.0,1.5) as in Figure 10.

Figure 13: Phase separation for a hard-sphere solute in a LJ solvent ( $\epsilon_{BB} = 1.1$ ) observed in MC simulation. The total number of particles used in this simulation was 1000. (a) Both solvent and solute are shown. (b) For clarity, the solvent particles have been removed. Note also that periodic boundaries are applied.

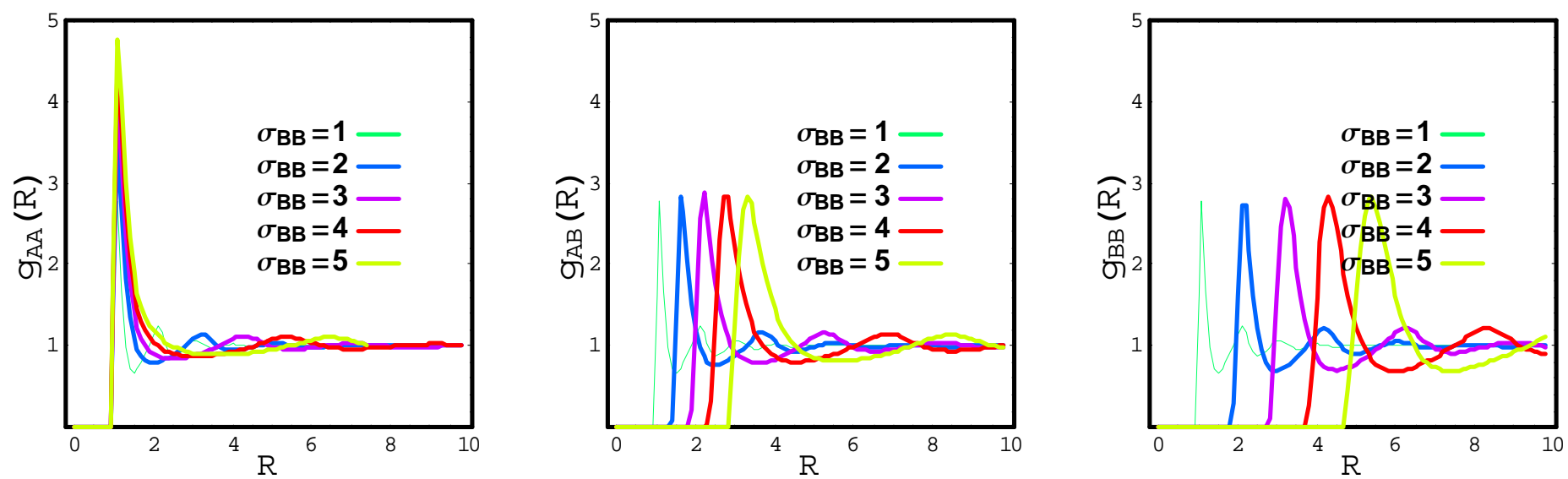


Figure 1

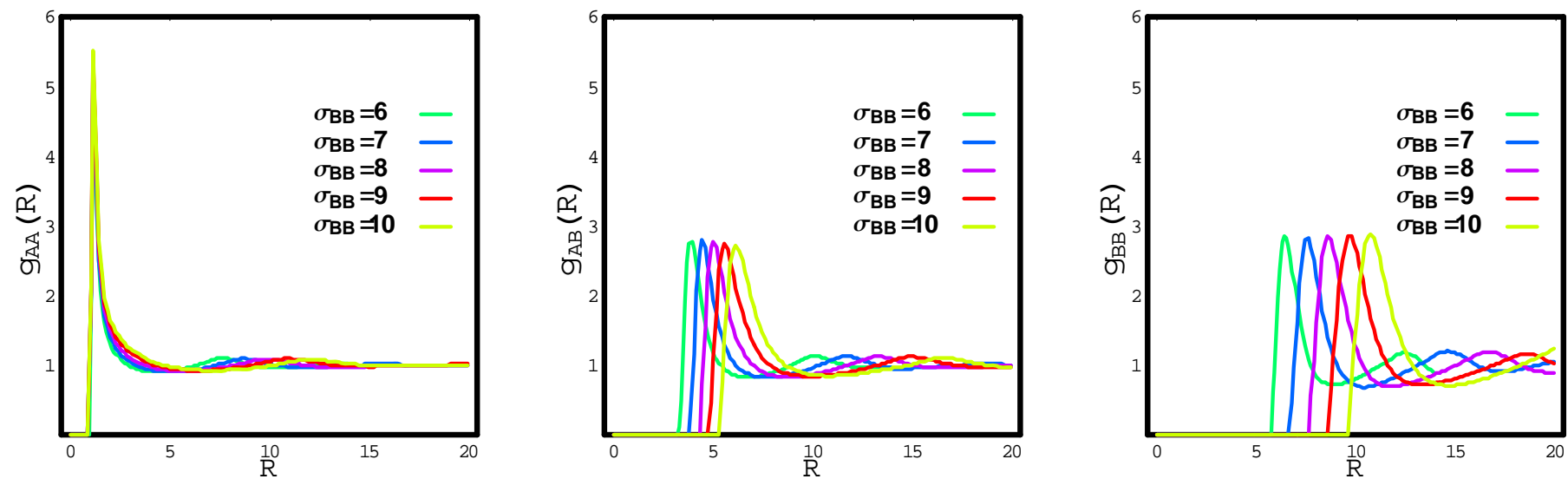


Figure 2

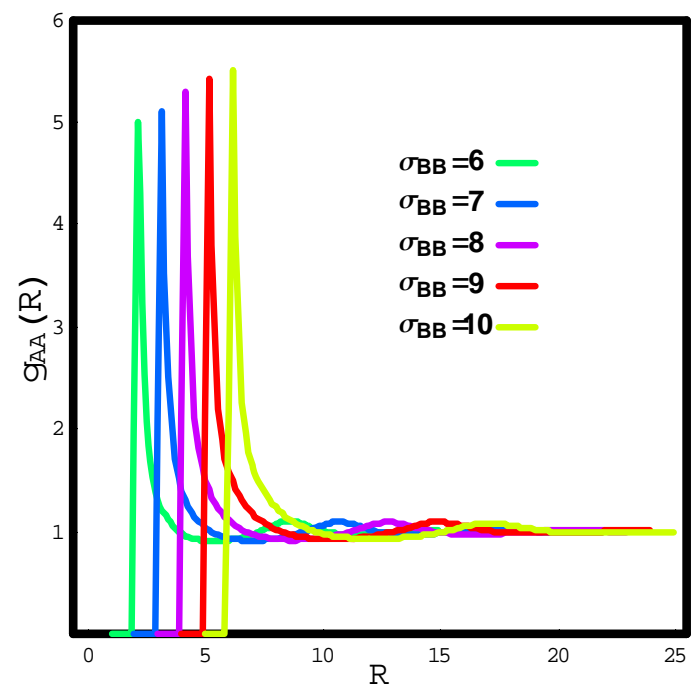
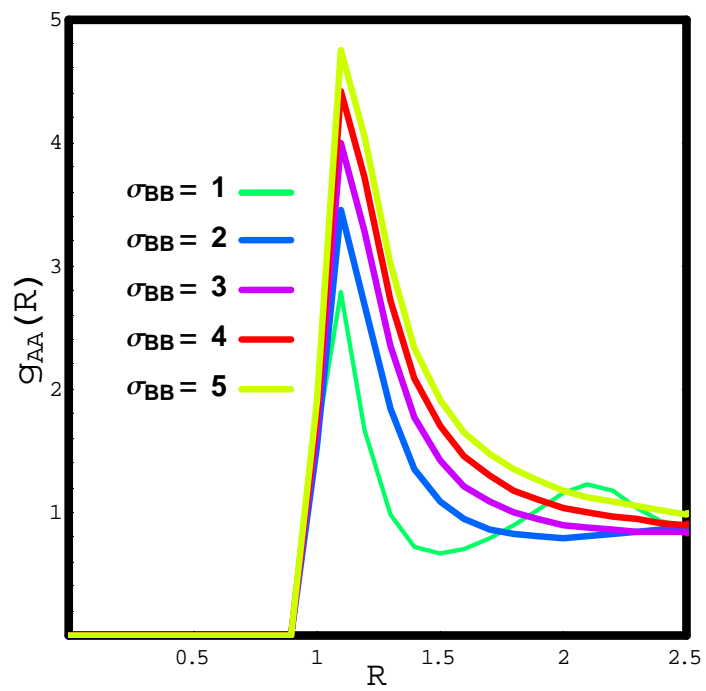


Figure 3

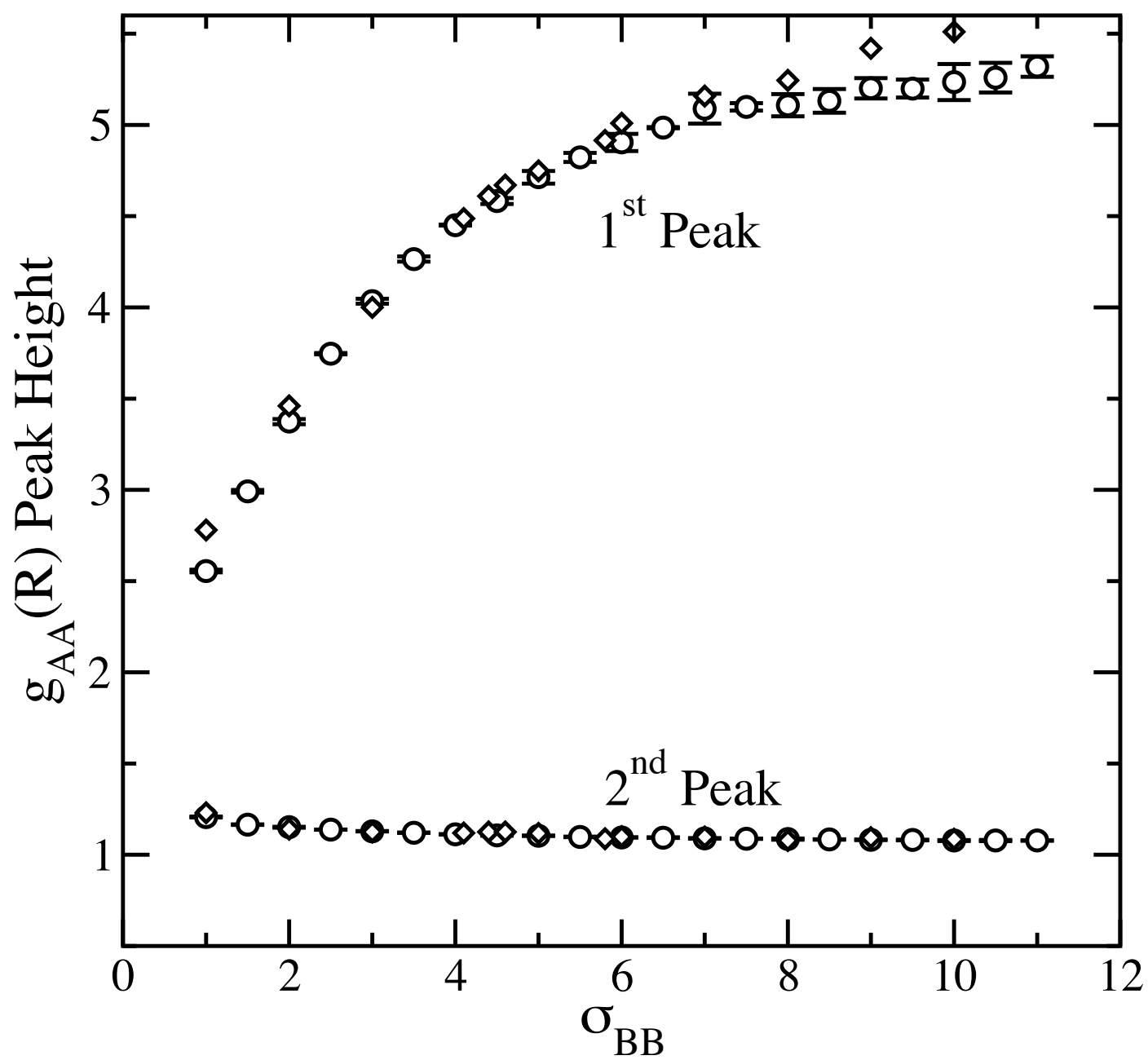


Figure 4



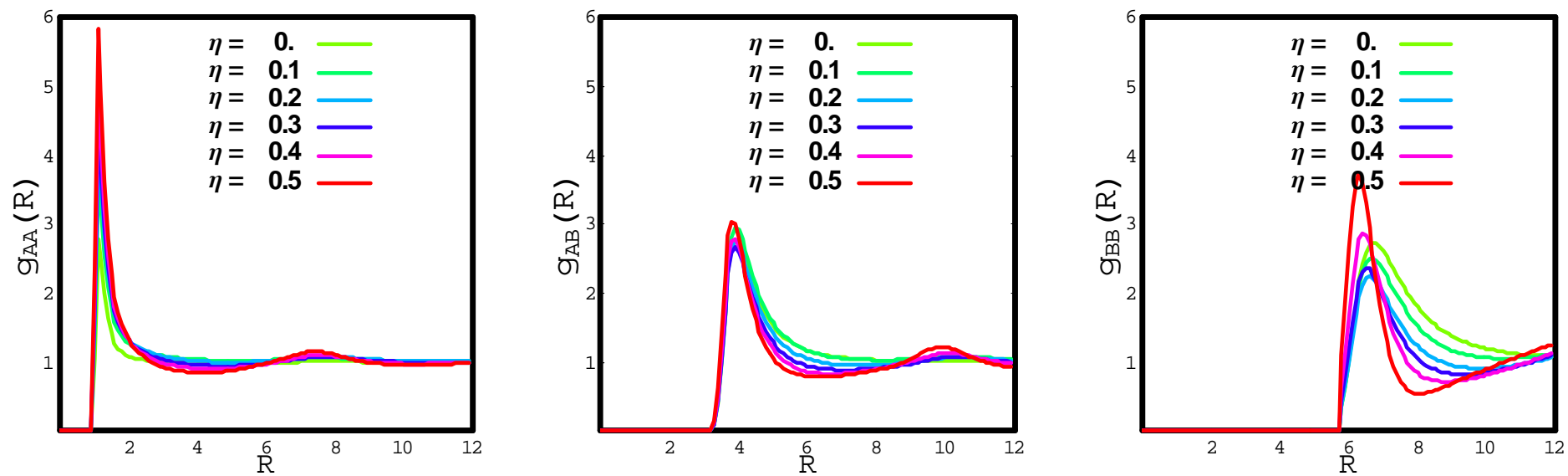


Figure 5

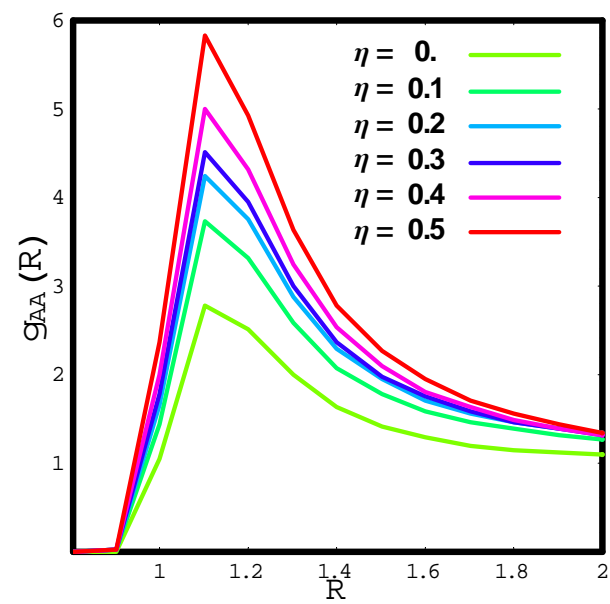


Figure 6

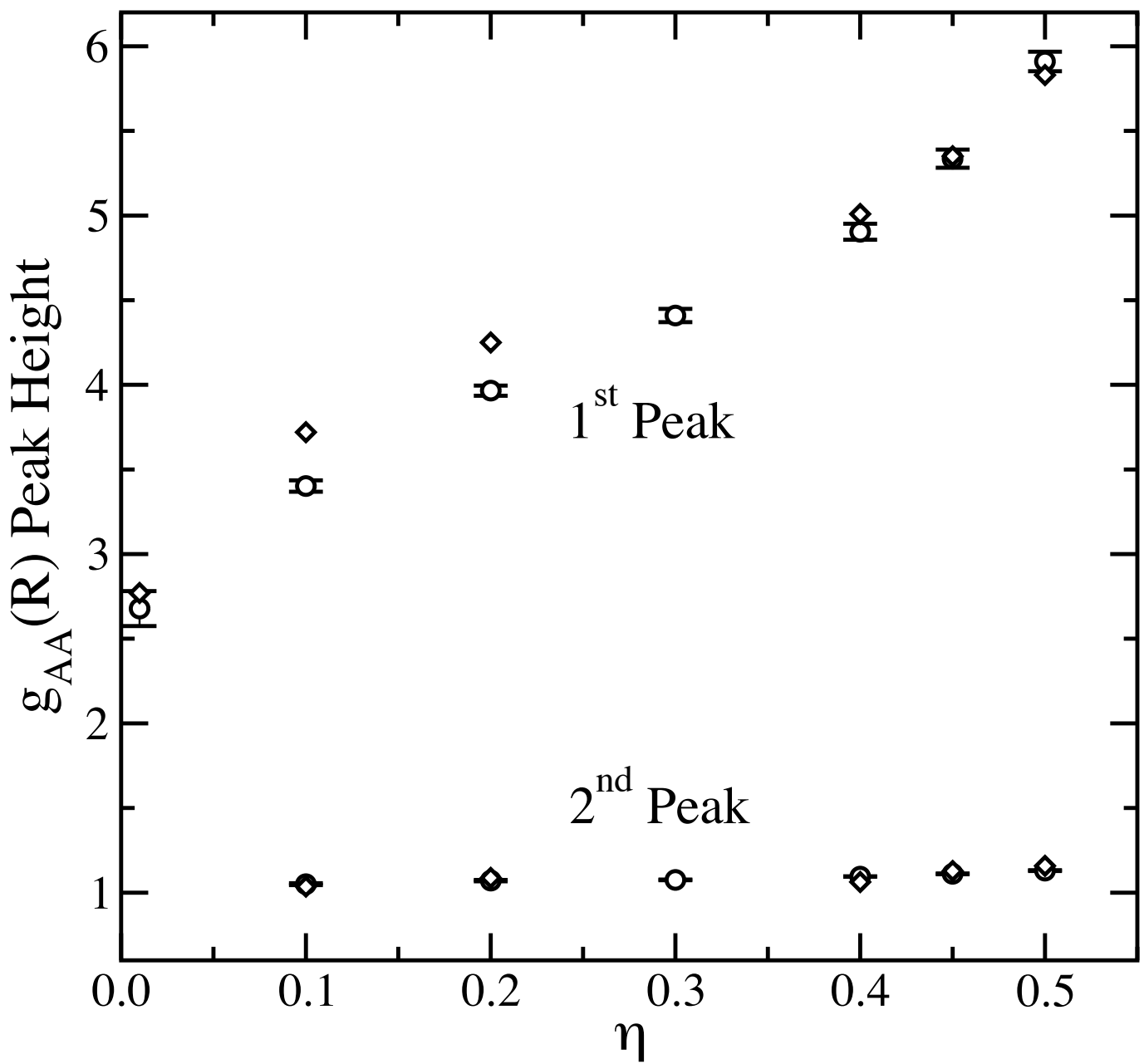


Figure 7

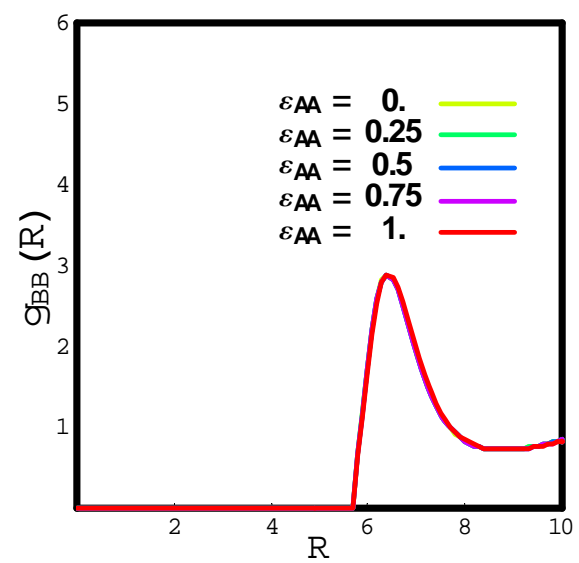
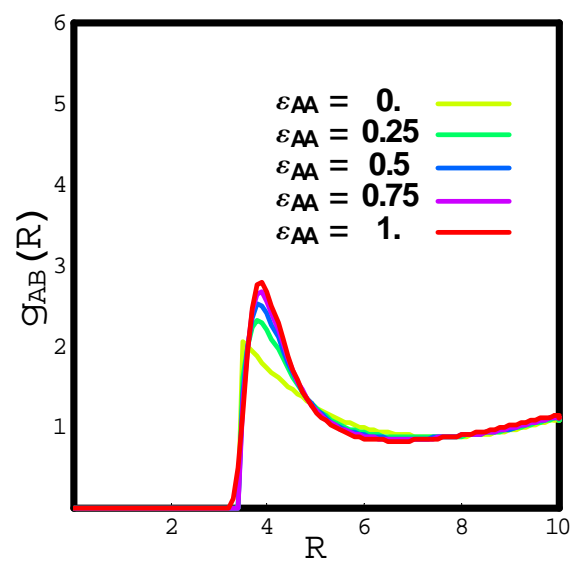
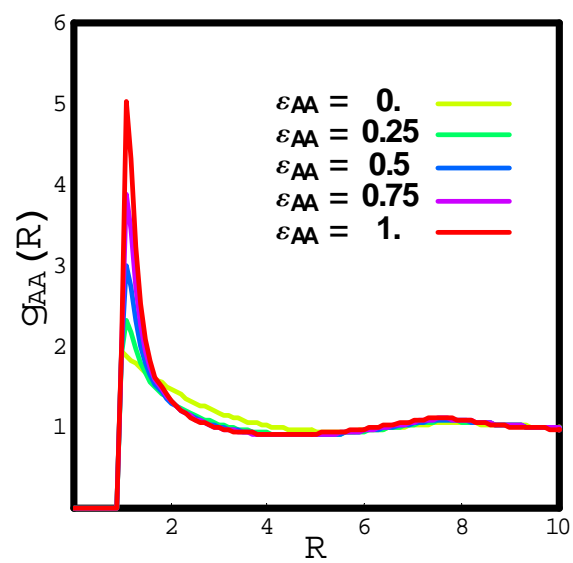


Figure 8

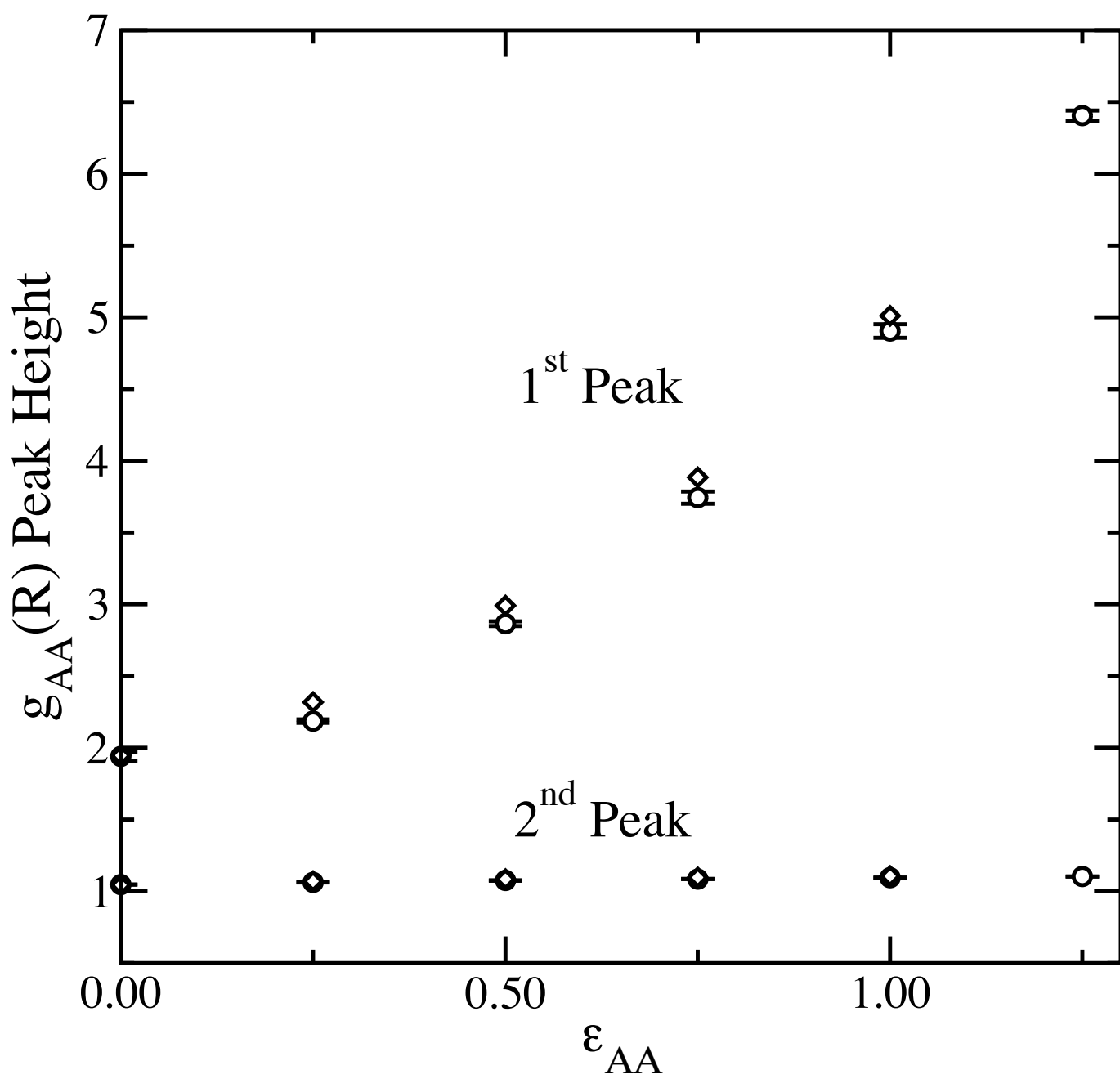


Figure 9

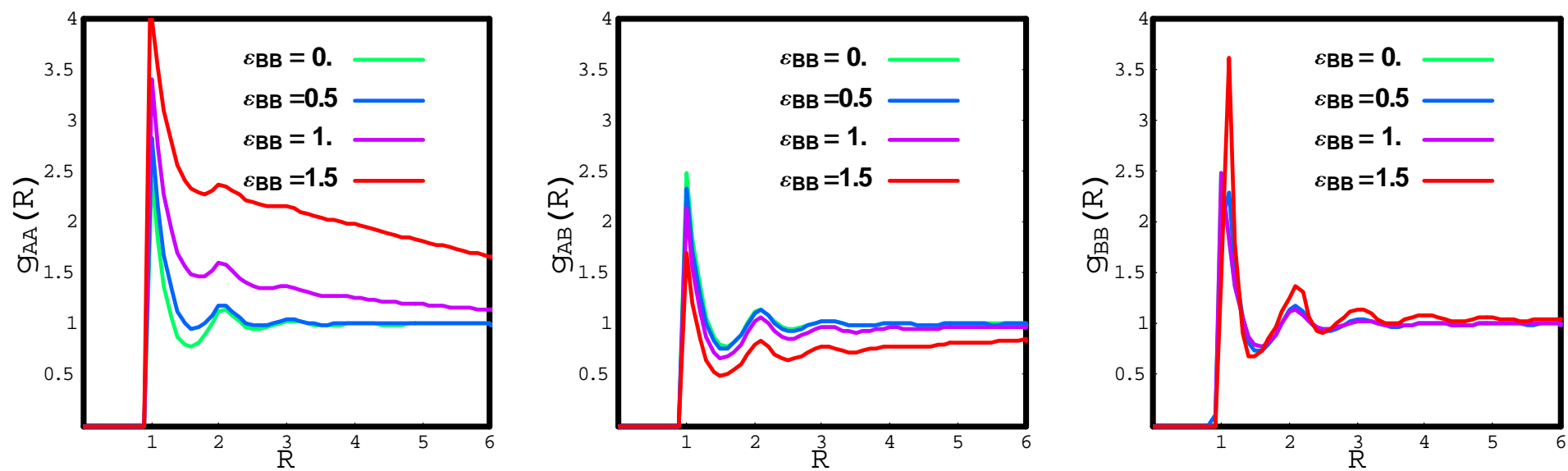


Figure 10

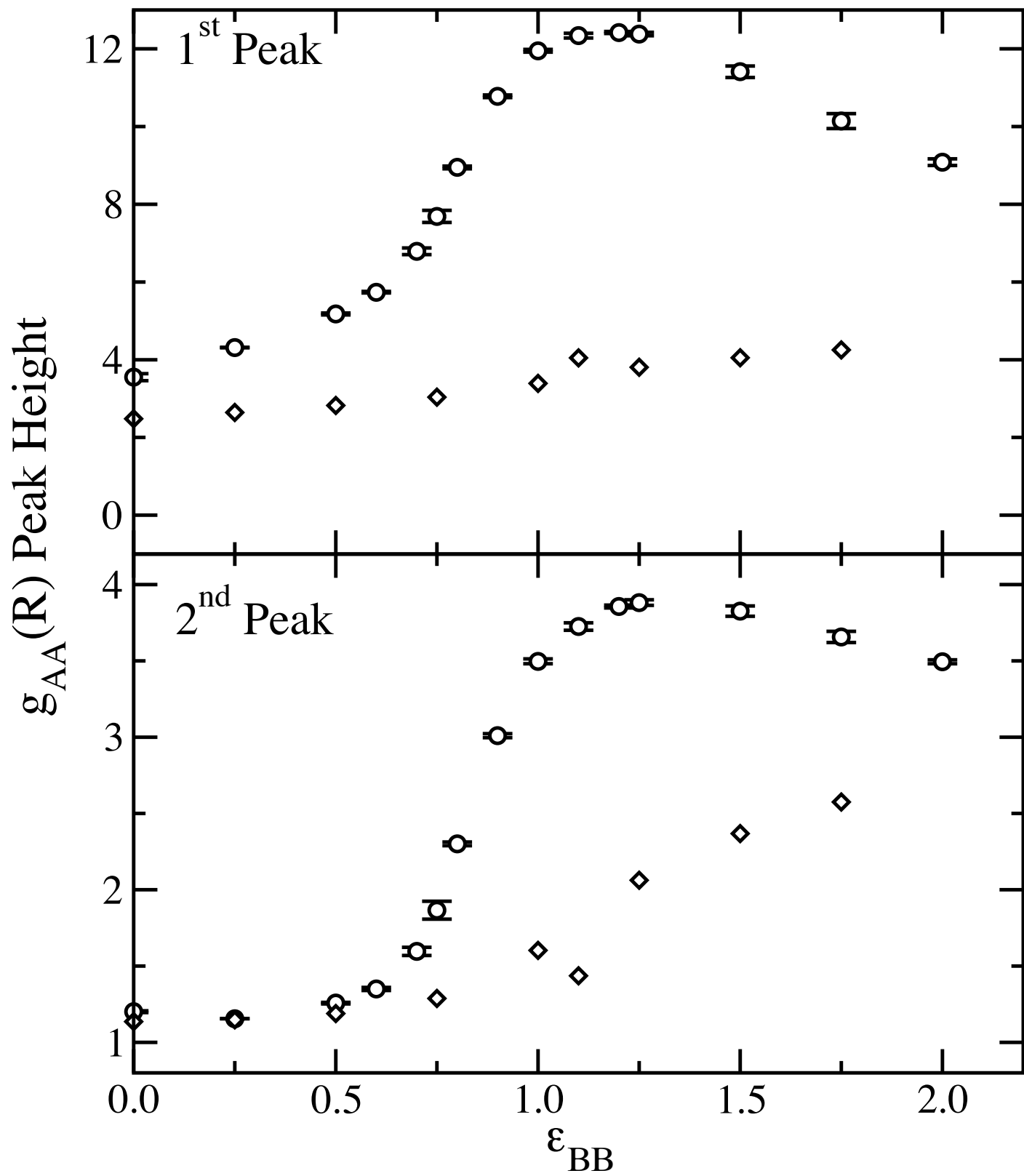


Figure 11

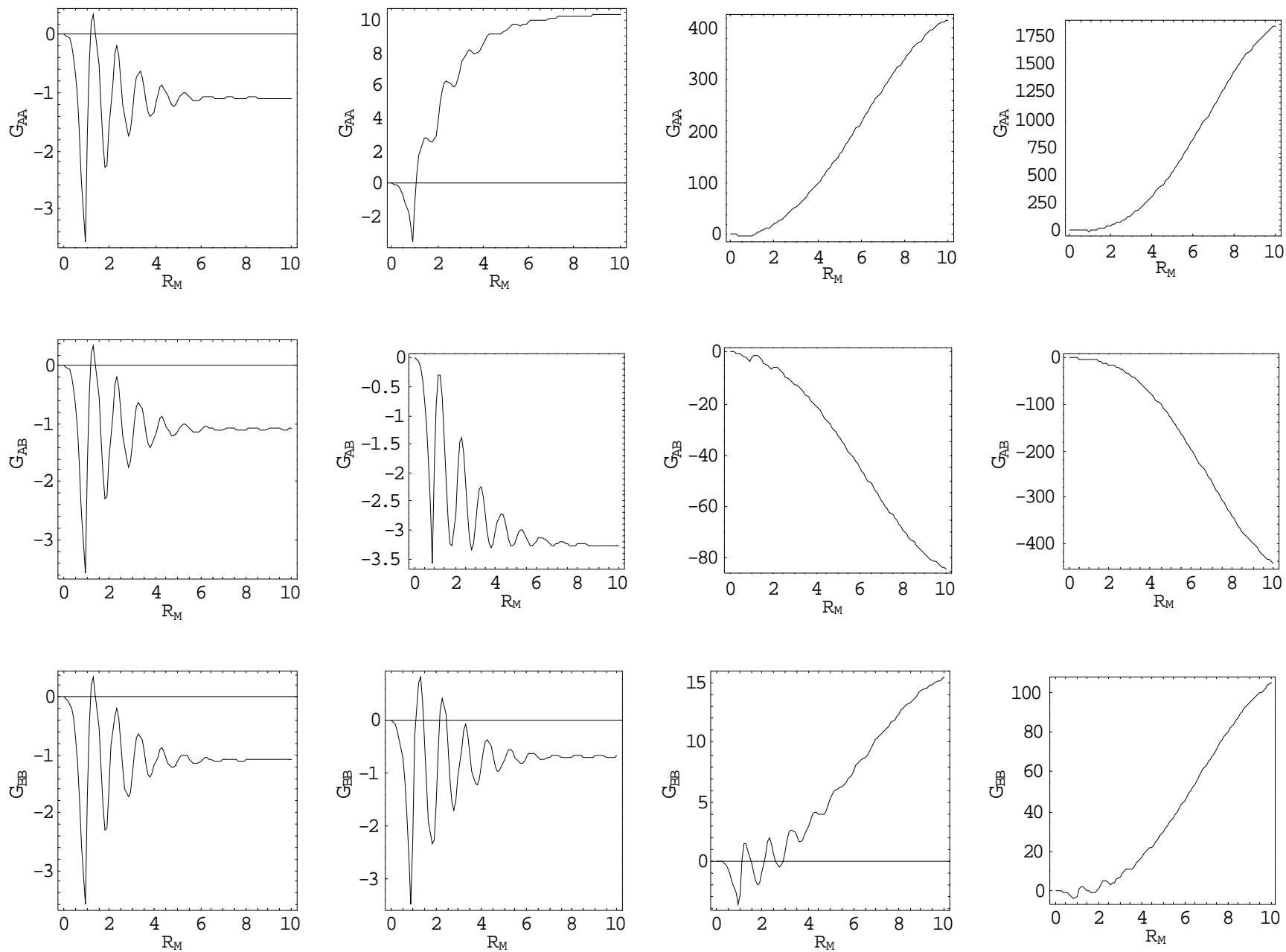


Figure 12



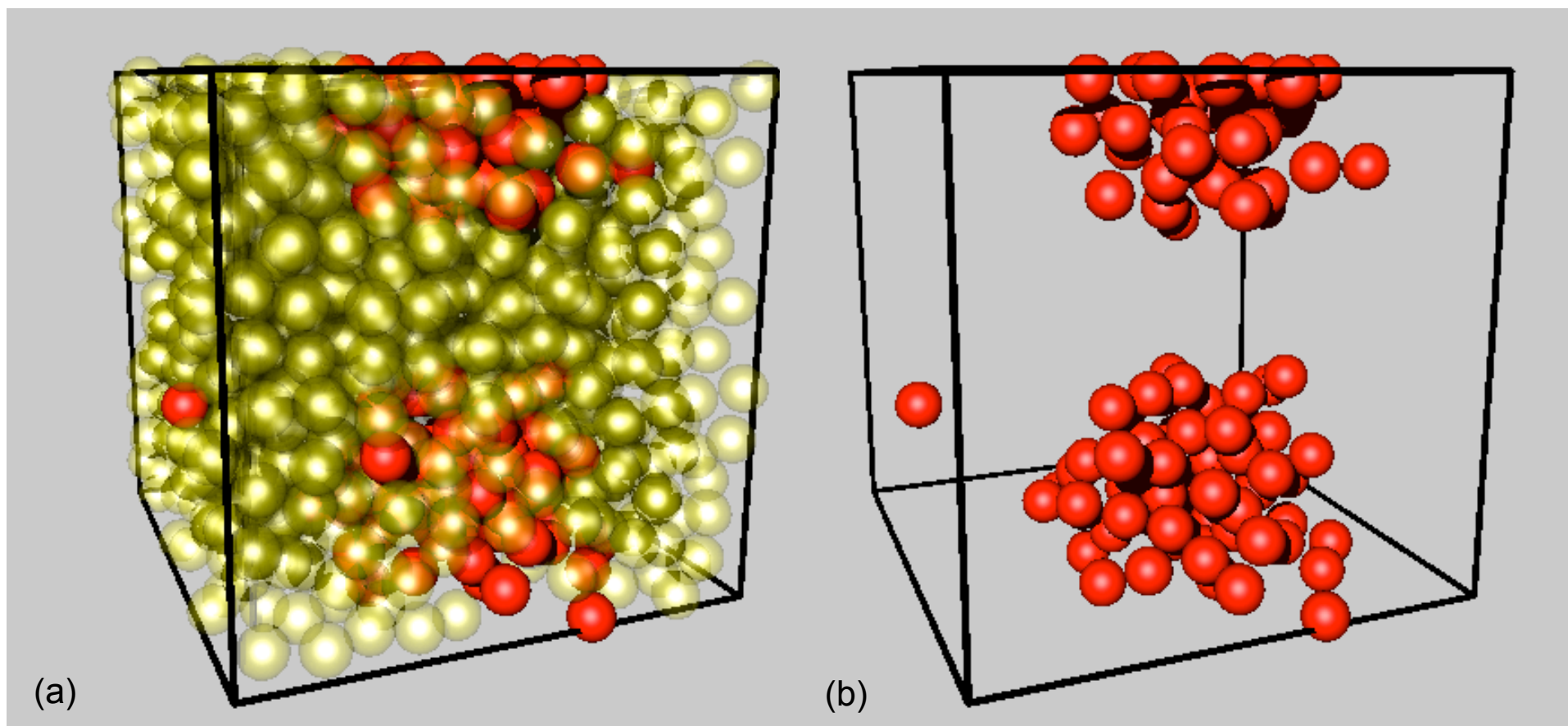


Figure 13.