Solvation structure and stability of peptides in aqueous solutions analyzed by the reference interaction site model theory

Masahiro Kinoshita, Yuko Okamoto, and Fumio Hirata

Citation: J. Chem. Phys. **107**, 1586 (1997); doi: 10.1063/1.474511
View online: https://doi.org/10.1063/1.474511
View Table of Contents: http://aip.scitation.org/toc/jcp/107/5
Published by the American Institute of Physics

Solvation structure and stability of peptides in aqueous solutions analyzed by the reference interaction site model theory

Masahiro Kinoshita
Advanced Energy Utilization Division, Institute of Advanced Energy, Kyoto University, Uji, Kyoto 611, Japan

Yuko Okamoto and Fumio Hirata
Department of Theoretical Studies, Institute for Molecular Science, Okazaki, Aichi 444, Japan

(Received 25 February 1997; accepted 29 April 1997)

We report results of numerical analyses on solvation structure and conformational stability of a dipeptide and Met-enkephalin in the extended simple point charge (SPC/E) model water. The reference interaction site model (RISM) theory is fully solved using our robust, highly efficient algorithm. It is shown that water structure near the peptides and the hydration free energy are greatly dependent on the peptide conformations. Stability of Met-enkephalin is examined in terms of the total energy defined as the sum of the conformational energy and the hydration free energy of the peptide. We test several different conformations including that with the minimum energy in gas phase, which takes rather compact form due to an intramolecular hydrogen bond. It is shown that a fully extended conformation has the highest stability in water. Our results are in qualitative accord with the recent nuclear magnetic resonance (NMR) experiments which suggest fully extended conformations with large fluctuations for the solution structure of the peptide. A conformation which is similar to that obtained from the NMR experiments in miceller solutions, is much less stable when it is put in water. Thus, the peptide conformations are greatly sensitive to microscopic solvent environment, and any native treatment of the solvent such as the continuum model will end in failure. © *1997 American Institute of Physics.* [S0021-9606(97)51629-1]

I. INTRODUCTION

The prediction of tertiary structures of proteins from their primary structures is one of the most challenging problems in biophysics and physical chemistry. The problem amounts to finding the lowest-energy conformation of a protein out of a huge number of possible conformations. Recently, promising simulation methods such as the simulated annealing[1] and the multicanonical algorithm[2] were developed to avoid getting trapped in a local minimum of the energy surface. Their usefulness was demonstrated for problems of peptide conformation prediction in gas phase[3–7] where the energy function is simply the conformational energy. These methods are now being extended to small proteins. However, another essential problem still remains unresolved: The incorporation of the effects due to the solvent (water molecules, anions, and cations) in the energy function.

The reference interaction site model (RISM) theory[8–10] provides a reliable approach of accounting for the solvent effects and potentially allows us to analyze a protein–solvent system on an atomic level. Pettitt and his co-workers[11,12] applied the RISM theory to the calculation of the free-energy surface of di- and tri-peptides, but they used the superposition approximation in which the entire free energy of a peptide is expressed as the sum of the potential of mean forces between pairs of atoms. The work of Kitao *et al.*[13] was the first one that employed the full RISM theory for a free-energy analysis for melittin (a small protein) in water. As one would have expected, however, a huge amount of computational effort was required to solve the basic equations. This problem appears to be a major stumbling block in elaborate studies based on the statistical-mechanical treatment for taking account of the solvent effects on the protein conformations. Fortunately, we recently developed an algorithm[14] for solving the full RISM equations for the system of a solute molecule with many atoms in water which is orders of magnitude faster than a conventional one.

Our ultimate goal is to combine the fast solution algorithm for the RISM theory with the powerful conformational sampling methods[1,2] mentioned above so that the lowest-energy conformation of a protein can be found in an aqueous solution (the energy function is the sum of the conformational energy and the solvation free energy). As an essential step in this direction, we consider in the present article a dipeptide (NH_2–$CHCH_3$–CO–NH–$CHCH_3$–COOH) and Met-enkephalin (the numbers of the atomic sites are 23 and 75, respectively) of some different conformations in the extended simple point charge (SPC/E) model water[15] and apply the full RISM theory to analyses of water structure (density structure) near the peptide and calculations of the hydration free energy. It is assumed in the analyses that these peptides are not ionized, but cases of zwitterions with zero net charges are also considered. It is shown that water structure and the hydration free energy are greatly dependent on the peptide conformations. The relation between water structure and the hydration free energy is analyzed and discussed in detail.

The conformations of Met-enkephalin in an aqueous solution (buffered to pH=3.87 using CH_3COONa at an ionic concentration of 0.05 M) determined from the recent nuclear

 0021-9606/97/107(5)/1586/14/$10.00

magnetic resonance (NMR) experiments[16] are fully extended and quite different from the lowest-energy conformation in gas phase which was already obtained using various simulation methods[5,17,18]. We test several different conformations including the lowest-energy conformation in gas phase and a conformation which is similar to those determined from the NMR experiments. We show that the latter is the most stable in water with the lowest total energy (the total energy is defined as the sum of the conformational energy and the hydration free energy) in both of the unionized and zwitterion cases. Although the effects due to the presence of CH_3COONa is unknown and need to be investigated in further studies, our results are very encouraging, implying that solvent plays essential roles and that the RISM theory can be a reliable tool for taking account of the solvent effects.

II. THEORETICAL MODEL

In the present article the subscripts ''v'' and ''s'' denote ''water'' and ''solute'', respectively. It is assumed that the solute molecules (peptides) are present at infinite dilution. The calculation process is then split into two steps where bulk water (step 1) and water near a solute molecule (step 2) are treated, respectively. The site–site intermolecular total correlation functions calculated in step 1 are used as input variables for step 2. The calculation in step 1 is performed using the RISM theory improved by Perkyns and Pettitt,[19,20] which assures the dielectric consistency. We consider step 2 hereafter.

It is assumed that the solute molecule and a water molecule has m and three interaction sites, respectively. The site–site Ornstein–Zernike (SSOZ) equation is expressed as

$$\tilde{\boldsymbol{\eta}}_{sv}=\tilde{\mathbf{w}}_{ss}\tilde{\mathbf{c}}_{sv}\tilde{\mathbf{H}}_{vv}-\tilde{\mathbf{c}}_{sv}, \tag{1a}$$

$$\tilde{\boldsymbol{\eta}}_{sv}=\tilde{\mathbf{h}}_{sv}-\tilde{\mathbf{c}}_{sv}, \tag{1b}$$

$$\tilde{\mathbf{H}}_{vv}=\tilde{\mathbf{w}}_{vv}+\boldsymbol{\rho}_v\tilde{\mathbf{h}}_{vv}, \tag{2}$$

where $\tilde{\mathbf{H}}_{vv}$, $\tilde{\boldsymbol{\eta}}_{sv}$, and $\tilde{\mathbf{w}}_{ss}$, for example, are 3×3, $m\times3$, and $m\times m$ matrices, respectively. $\boldsymbol{\rho}_v$ is the matrix of number density of water molecules in the bulk, $\mathbf{h}$ is the matrix of site–site intermolecular total correlation functions, $\mathbf{c}$ is the matrix of site–site intermolecular direct correlation functions, $\mathbf{w}$ is the intramolecular correlation matrix, and ''~'' represents Fourier transforms. $\tilde{\mathbf{H}}_{vv}$ is dependent on properties of the bulk water alone and is part of the input data for step 2. More detailed information is given in Ref. 14.

The closure equation employed is of the hypernetted-chain (IINC) type given by

$$c_{AB}(r)=\exp\{-u_{AB}(r)/(k_BT)+\eta_{AB}(r)\}-\eta_{AB}(r)-1, \quad A=1,\ldots,m;\quad B=\mathrm{H,O}, \tag{3}$$

where $u_{AH}(u_{AO})$ is the pair potential between site A of the solute molecule and the water–hydrogen (oxygen) and k_B is the Boltzmann constant. For instance, $c_{AH}(c_{AO})$ is the site–site direct correlation function between site A of the solute–molecule and the water–hydrogen (oxygen).

The full pair distribution function $g(12)$ defined for molecular fluids is dependent on the distance between centers of molecules 1 and 2 and on the orientations of these two molecules. Averaging $g(12)$ over all orientations of molecules 1 and 2 with fixing atomic site A in molecule 1 and atomic site B in molecule 2 yields the site–site pair distribution function $g_{AB}(r)$ (r is the distance between the two atomic sites, A and B). We discuss the structure of water near the peptides in terms of $g_{AB}(r)$. $\rho_B g_{AB}(r)$ in this case can be regarded as the orientationally averaged density profile of atom B of the water–molecule near atom A of the peptide.

The hydration free energy for the solute molecule $\Delta\mu_s$ is calculated from[10,13,21]

$$\Delta\mu_s/(k_BT)=4\pi\int_0^\infty F(r)dr, \tag{4a}$$

$$F(r)=\sum_{A=1}^{m}\sum_{B=\mathrm{H,O}}\rho_B r^2[\{h_{AB}(r)\}^2/2-c_{AB}(r)-h_{AB}(r)c_{AB}(r)/2], \tag{4b}$$

$$\rho_{\mathrm{H}}=2\rho_v,\quad \rho_{\mathrm{O}}=\rho_v. \tag{4c}$$

The site–site correlation functions $h_{AB}(r)$ and $c_{AB}(r)$ are calculated by solving Eqs. (1) and (3). We consider two cases: Case (a) where the full values of the site–charges of the solute molecule are used and case (b) where all the site–charges are set to zero. The values of the solvation free energies in cases (a) and (b) are denoted by $\Delta\mu_{sa}$ and $\Delta\mu_{sb}$, respectively.

The model of a water–molecule is the SPC/E model.[15] The temperature is set at 298.15 K. $u_{AB}(r)$ has the form

$$u_{AB}(r)=q_Aq_B/r+4\epsilon_{AB}\{(\sigma_{AB}/r)^{12}-(\sigma_{AB}/r)^6\}, \quad A=1,\ldots,m;\quad B=\mathrm{H,O}, \tag{5}$$

where q_A is the partial charge on site A of the solute–molecule and the standard combination rule

$$\epsilon_{AB}=(\epsilon_A\epsilon_B)^{1/2},\quad \sigma_{AB}=(\sigma_A+\sigma_B)/2, \tag{6}$$

is employed for calculating the Lennard-Jones potential parameters. The potential-energy functions and parameters are adopted from KONF90 (Ref. 22) which is based on ECEPP/2 (Refs. 23–25). The values of q_A and σ_A used for the dipeptide chosen (Ala–Ala) are given in Table I (the electronic charge is -1). Those for some representative sites of Met-enkephalin (Tyr–Gly–Gly–Phe–Met) are given in Table II. The carbonyl carbons have large, positive site–charges, and oxygens (in particular, the two oxygens at the C-terminus of the zwitterions) and nitrogens have large, negative site–charges. For the SPC/E water, we have $q_{\mathrm{H}}=0.4238$, $q_{\mathrm{O}}=-0.8476$, $\epsilon_{\mathrm{H}}=0.046$ kcal/mol, $\epsilon_{\mathrm{O}}=0.156$ kcal/mol, $\sigma_{\mathrm{H}}=0.040$ nm, and $\sigma_{\mathrm{O}}=0.316$ nm. The σ-value of the water–hydrogen is exceptionally small. The dimensionless number density of water $\rho_v d^3$ ($d=0.28$ nm) is 0.7317.

TABLE I. Values of q_A and σ_A used for the dipeptide. The electronic charge is -1. The last four rows are for the zwitterion (1 H Ala[1], 3 H Ala[1], and 23 H Ala[1] form NH_3^+). The q-values and σ-values of water are $q_H=0.4238$, $q_O=-0.8476$, $\sigma_H=0.040$ nm, and $\sigma_O=0.316$ nm.

A	$q_A(-)$	σ_A(nm)
1 H Ala[1]	0.176	0.239
2 N Ala[1]	−0.356	0.313
5 HB1 Ala[1]	0.040	0.260
8 CA Ala[1]	0.064	0.367
9 HA Ala[1]	0.020	0.260
10 C Ala[1]	0.450	0.333
11 O Ala[1]	−0.384	0.278
12 N Ala[2]	−0.356	0.313
14 CB Ala[2]	−0.090	0.367
13 H Ala[2]	0.176	0.239
20 C Ala[2]	0.450	0.333
21 O Ala[2]	−0.384	0.278
22 O Ala[2]	−0.380	0.289
23 H Ala[2]	0.204	0.252
1 H Ala[1]	0.285	0.239
21 O Ala[2]	−0.532	0.278
22 O Ala[2]	−0.532	0.278
23 H Ala[1]	0.285	0.239

III. NUMERICAL METHOD

A sufficiently long-range r_L is divided into N mesh points ($r_i=i\delta r$, $i=0,1,\ldots,N-1$; $\delta r=r_L/N$) and all the functions are represented by their values on these points. The long-range Coulomb potentials are handled in a special manner so that r_L can be minimized. In the present analysis, δr and N are set at 0.02 d ($d=0.28$ nm) and 512, respectively. The details of the algorithm for solving the full RISM equations are described in Ref. 14. The algorithm is a hybrid of the Newton–Raphson and Picard methods. The Jacobian matrix is read from a file as part of the input data. We have found that the same matrix can be used for a considerably large set of different conformations of the solute molecule. Since the matrix is part of the input data, it is completely independent of the initial guess, and sufficient stability is assured even with a crude initial guess. The algorithm is capable of treating a molecular solute with many atomic sites with minor computational effort on an interactive workstation.

TABLE II. Values of q_A and σ_A used for Met-enkephalin. The electronic charge is -1. The last four rows are for the zwitterion (1 H Tyr[1], 3 H Tyr[1], and 75 H Tyr[1] form NH_3^+).

A	$q_A(-)$	σ_A(nm)
1 H Tyr[1]	0.176	0.239
2 N Tyr[1]	−0.356	0.313
4 CB Tyr[1]	−0.040	0.367
5 HB1 Tyr[1]	0.025	0.260
10 CE1 Tyr[1]	−0.060	0.330
11 HE1 Tyr[1]	0.030	0.261
13 OH Tyr[1]	−0.330	0.289
14 HH Tyr[1]	0.165	0.252
55 C Phe[4]	0.450	0.333
56 O Phe[4]	−0.384	0.278
57 N Met[5]	−0.356	0.313
58 H Met[5]	0.176	0.239
62 CG Met[5]	−0.120	0.367
66 CE Met[5]	−0.190	0.367
72 C Met[5]	0.450	0.333
73 O Met[5]	−0.384	0.278
74 O Met[5]	−0.380	0.289
75 H Met[5]	0.204	0.252
1 H Tyr[1]	0.285	0.239
73 O Met[5]	−0.532	0.278
74 O Met[5]	−0.532	0.278
75 H Tyr[1]	0.285	0.239

The convergence criterion for the iterative calculation is set so that the hydration free energy can be calculated with the accuracy ± 0.3 kcal/mol (0.6 kcal/mol$\sim k_BT$). We note that this is a severe criterion particularly for Met-enkephalin. Nevertheless, with a crude initial guess convergence is achieved in ~ 1 and 10 min for the dipeptide and Met-enkephalin, respectively, on our workstation (IBM RS6000/3CT; 64MB). When the converged solution for another conformation is available, convergence is achieved in a few tens of seconds and several minutes for the dipeptide and Met-enkephalin, respectively. We emphasize that the full RISM equations are rigorously solved with no approximate treatment to accelerate convergence.

IV. RESULTS AND DISCUSSION

The unionized peptides are first treated, and then the zwitterions are considered in a later section. We have tested several different conformations of Met-enkephalin (Tyr–Gly–Gly–Phe–Met), but in the present article we describe four of them illustrated in Fig. 1. Conformation 1 is the lowest-energy conformation in gas phase determined by the multicanonical algorithm[5] and has hydrogen bonding between 14 HH Tyr[1] and 36 O Gly[3]. In conformation 2 the five carbonyl oxygens are not far apart, and in particular 29 O Gly[2] and 56 O Phe[4] are close to each other. Conformation 3 is a conformation we have obtained from the backbone dihedral angles given in Ref. 16. These angles were determined from NMR experiments for Met-enkephalin in an aqueous solution with the presence of 50 mM sodium dodecyl sulfate (SDS) (the critical micellar concentration is 8.3 mM). In a strict sense conformation 3 is different from the one shown in Fig. 3 of Ref. 16 because the aromatic side chains of Tyr[1] and Phe[4] are not close together in our case (since Ref. 16 gives only the backbone dihedral angles, it is difficult to obtain appropriate side-chain orientations). However, conformation 3 maintains the main feature that all five of the carbonyl oxygens are on the same side as seen in Fig. 1(d). We have prepared conformation 4 such that it is fully extended as implied by the NMR results.[16] Conformational energies for the four conformations are -12.0, 12.2, -2.5, and 0.8 kcal/mol, respectively. Conformation 2 has the highest energy among the four conformations.

Here, we define the total energy as the sum of the conformational energy and the hydration free energy. The total energies for conformations 2, 3, and 4 relative to the total energy for conformation 1 are compared in Table III. In column b of the table, all the site–charges of Met-enkephalin

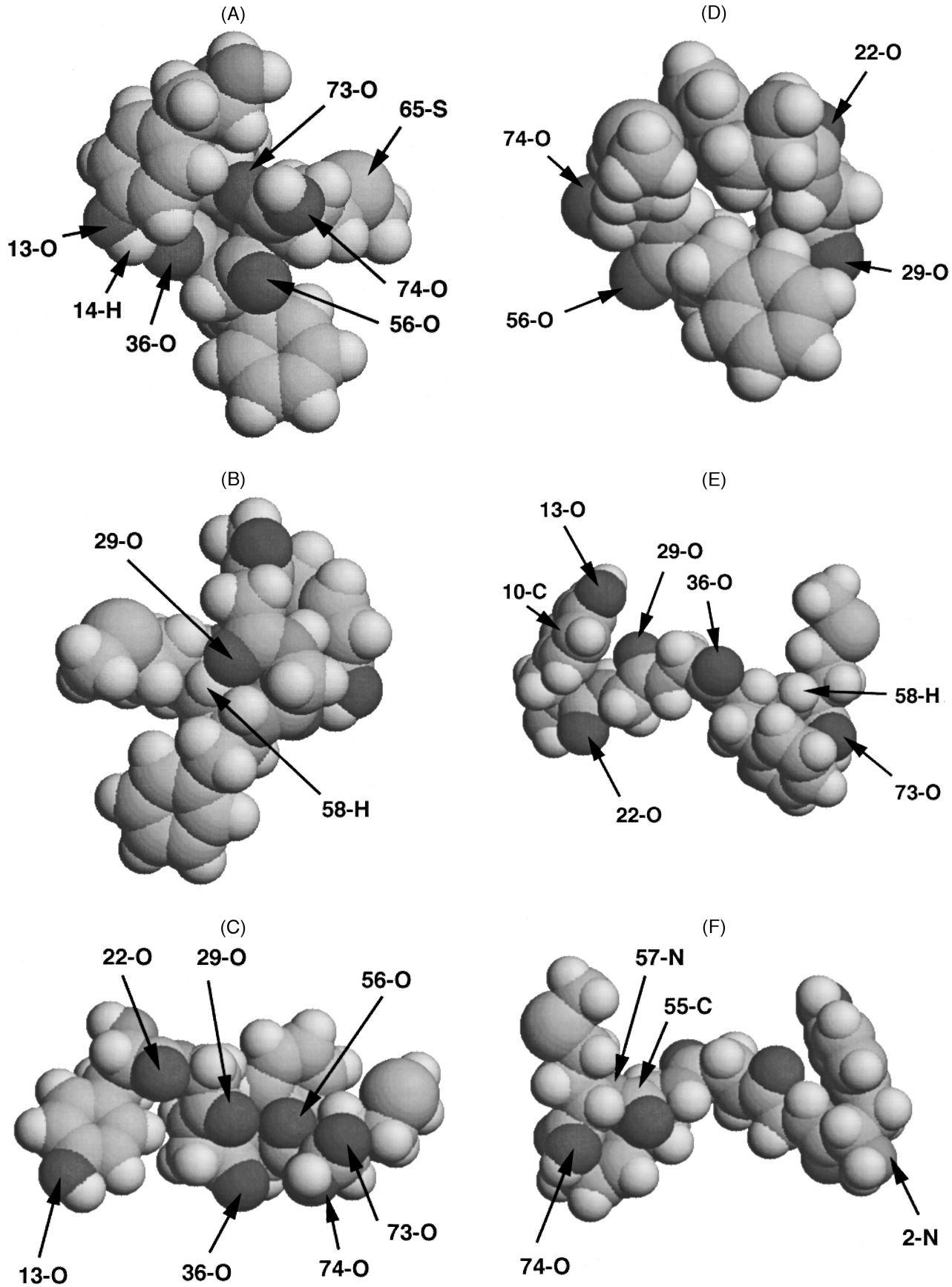

FIG. 1. Four different conformations of Met-enkephalin considered. C, H, N, O, and S denote carbon, hydrogen, nitrogen, oxygen, and sulfer atoms, respectively. ''10-C'', ''13-O'', ''14-H'', ''29-O'', ''36-O'', ''55-C'', ''57-N'', and ''74-O'', for example, represent ''10 CE1 Tyr^{1}'', ''13 OH Tyr^{1}'', ''14 HH Tyr^{1}'', ''29 O Gly^{2}'', ''36 O Gly^{3}'', ''55 C Phe^{4}''. ''57 N Met^{5}'', and ''74 O Met^{5}'', respectively. This figure was prepared by RasMol. (a) Conformation 1. (b) Conformation 1 viewed from another angle. (c) Conformation 2. (d) Conformation 3. (e) Conformation 4. (f) Conformation 4 viewed from another angle.

TABLE III. Total energies for conformations 2, 3, and 4 of Met-enkephalin relative to the total energy for conformation 1 (kcal/mol). Column a: In the case where the full values of the site–charges are used. Column b: In the case where all the site–charges are set to zero. Column c: In gas phase [i.e., conformational energies relative to the conformational energy for conformation 1 (kcal/mol)]. Column d: In the case where the solvent is the repulsive potential system explained in the text. The changes (kcal/mol) are also given with arrows.

Conformation	a		b		c		d
1	0.0		0.0		0.0		0.0
2	5.4	-11.2 $\leftarrow$	16.6	-7.6 $\leftarrow$	24.2	-1.8 $\rightarrow$	22.4
3	15.4	-6.2 $\leftarrow$	21.6	$+12.1$ $\leftarrow$	9.5	$+5.9$ $\rightarrow$	15.4
4	-7.2	-4.6 $\leftarrow$	-2.6	-15.4 $\leftarrow$	12.8	-3.2 $\rightarrow$	9.6

are set to zero. In column d. we consider a simple, repulsive potential system as the solvent. The particles of this solvent interact through

$$u(r)=4\epsilon(\sigma/r)^{12}, \tag{7}$$

where $\epsilon=0.156$ kcal/mol and $\sigma=0.28$ nm. The dimensionless number density in the bulk is 0.7317. The interaction between the particle and an atomic site of the peptide is also expressed as the form of Eq. (7) with the combination rule of Eq. (6). First, we discuss the results in columns a, b, and c of the table. Conformation 4, a fully extended conformation, is considerably more stable in water than conformation 1, which is in qualitative accord with the experimental observations[16] (data obtained by the NMR technique). This indicates that solvent plays essential roles in determining the conformation of a peptide and that the RISM theory is a promising tool for taking account of the solvent effects. The decrease in the hydration free energy caused by the presence of water is very large in conformation 2. Nevertheless, due to the highest conformational energy, this conformation is less stable in water than conformation 1. It is interesting to note that conformation 3 is the most unstable in water without the presence of SDS. We then consider the results in column d. In the repulsive potential system, conformation 1 is still the most stable with the lowest total energy. Water is thus clearly distinguished from the simple solvent, even when all the site–charges of the peptide are set to zero (we note that conformation 4 is more stable than conformation 1 even in column b).

In the succeeding sections, water structure near the peptides is analyzed and the relation between the structure and the hydration free energy is discussed in detail. The dipeptide (Ala–Ala), a peptide which is smaller and simpler than Met-enkephalin, is also considered.

A. Spherical particles

Before considering the peptides, we treat spherical particles (isolated atoms) which are present at infinite dilution in water. Eight different types of atoms are treated. Those of types 1–8 are characterized by the partial charges and Lennard-Jones parameters of 1 H Ala[1], 5 HB 1 Ala[1], 10 C Ala[1], 12 N Ala[2], 14 CB Ala[2], 21 O Ala[2], 22 O Ala[2], and 23 H Ala[2] of the dipeptide, respectively. Those of types 9 and 10 have the partial charges and Lennard-Jones parameters of 22 O Ala[2] and 23 H Ala[1] of the zwitterion, respectively. The hydration free energies calculated for the eight types of atoms are given in Table IV. [Some of the site–site pair distribution functions $g_{AB}(r)$ will be shown in Figs. 3–6 of Sec. IV B.] The first peaks of $g_{AB}(r)$ (B is a water–hydrogen: $B=$H) for $A=$12 N Ala[1], 21 O Ala[2], and 22 O Ala[2] are very sharp, and the first-peak values are 2.11, 3.91, and 3.73, respectively. Thus, each of these atoms is strongly bonded with water–hydrogens, giving rise to a large, negative hydration free energy. This is ascribed to the large, negative partial charges of these atoms and strong, electrostatic attractive interactions with water–hydrogens. The first-peak value of $g_{AB}(r)$ (B is a water–oxygen: $B=$O) for $A=$10 C Ala[1] (this atom has a large, positive partial charge) is 2.14, leading to a negative hydration free energy. We note that 23 H Ala[2] has a relatively large, positive site–charge [the first-peak value of $g_{AO}(r)$ is 2.02] but the hydration free energy for this atom is positive. This is because the core diameter (σ-value) of 23 H Ala[2] is not small (in fact, it is much larger than that for water–hydrogens) and the electrostatic attractive interaction between this atom and a water–oxygen is not sufficiently strong. As expected, the hydration free energies for atoms of types 9 and 10 are negative. In particular, the first-peak value of $g_{AH}(r)$ for $A=$23 H Ala[1] is 8.07, resulting in an extremely large, negative hydration free energy.

TABLE IV. Hydration free energies (kcal/mol) calculated for eight types of imaginary spherical particles, isolated atoms. The last two (types 9 and 10) are for isolated atoms related to the zwitterion case.

Type	Atom	$\Delta\mu_s$(kcal/mol)
1	1 H Ala[1]	2.0
2	5 HB1 Ala[1]	4.4
3	10 C Ala[1]	-5.1
4	12 N Ala[2]	-11.5
5	14 CB Ala[2]	6.9
6	21 O Ala[2]	-18.1
7	22 O Ala[2]	-16.9
8	23 H Ala[2]	1.7
9	22 O Ala[2]	-39.8
10	1 H Ala[1]	-2.0

However, the structure of water near an atom of the dipeptide (or Met-enkephalin) and the contribution from the atom to the hydration free energy are far more complicated than those discussed above, because they are greatly dependent on the neighboring atoms.

B. Dipeptide

To analyze the effects of the neighboring atoms on water structure around a peptide atom, we consider two different conformations of the dipeptide (Ala–Ala) illustrated in Fig. 2 (actually, we have tested several different conformations, but we describe two of them). The conformational energies for conformations 1 and 2 are 4.6 and 38.2 kcal/mol, respec-

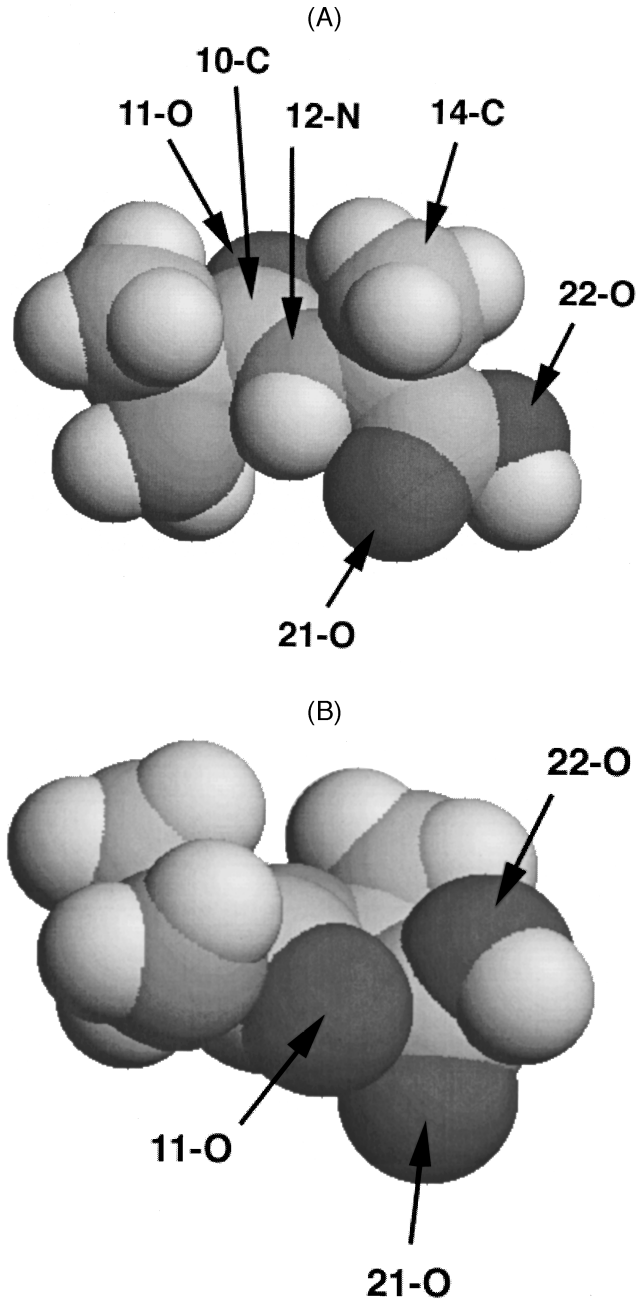

FIG. 2. Two different conformations of the dipeptide considered. ''10-C'' and ''14-C'', for example, represent ''10 C Ala[1]'' and ''14 CB Ala[2]'', respectively. This figure was prepared by RasMol. (a) Conformation 1. (b) Conformation 2.

tively. Conformation 1 is the lowest-energy conformation in gas phase determined by the Monte Carlo simulated annealing for the present study. Conformation 2 has a much higher conformational energy.

The first-peak values of some representative site–site pair distribution functions $g_{AB}(r)$ in the two conformations are compared in Table V. Figures 3–6 show $g_{AB}(r)$ for A = 11 O Ala[1], 12 N Ala[2], 10 C Ala[1], and 14 CB Ala[2], respectively. Here, the solute atom A is either an atom of the dipeptide in conformation 1 or a spherical particle (isolated atom) which has the same site–charge and Lennard-Jones potential

TABLE V. First-peak values of some representative site–site radial distribution functions $g_{AB}(r)$. A and B denote atomic sites in the dipeptide and a water–molecule, respectively.

A	B	Conf. 1	Conf. 2
2 N Ala[1]	H	0.43	0.51
11 O Ala[1]	H	1.84	2.00
12 N Ala[2]	H	0.22	0.15
21 O Ala[2]	H	1.78	1.82
22 O Ala[2]	H	0.85	0.88
1 H Ala[1]	O	1.19	1.11
3 H Ala[1]	O	1.25	1.16
10 C Ala[1]	O	0.90	0.80
20 C Ala[2]	O	1.20	1.23
23 H Ala[2]	O	1.39	1.37

parameters. Table VI gives the contributions from some representative atoms to the hydration free energy for conformations 1 and 2 in case a, where the full values of the site–charges are used. The first-peak values of $g_{AB}(r)$ for oxygens, nitrogens, and carbonyl carbons are much lower than those for the isolated atoms treated in Sec. IV A as observed in Table V and Figs. 3–5. Also, the contributions from these atoms to the hydration free energy (Table VI) are more or less shifted in more ''hydrophobic'' directions. We discuss these results in more detail below.

For A = 11 O Ala[1] (Fig. 3), $g_{AB}(r)$ in the dipeptide and isolated-atom cases possess sharp first peaks at the same positions. The position of the first peak of $g_{AH}(r)$ is about half of that of $g_{AO}(r)$. Since the core diameter of 11 O Ala[1] does not significantly differ from that of a water–oxygen, Fig. 3 indicates that water–hydrogens are rather strongly bonded with 11 O Ala[1] (i.e., the bond formation, 11 O Ala[1]–H–O). 11 O Ala[1] is covalently bonded with one carbon atom having a certain core diameter. Moreover, this atom has a positive site–charge. Hence, water–hydrogens can form bonding with 11 O Ala[1] in more limited orientations than with the isolated atom, resulting in lower first-peak values of $g_{AB}(r)$. Also, the contribution from 11 O Ala[1] to the hydration free energy is −5.8 and −9.7 kcal/mol in conformations 1 and 2, respectively, rather than −18.1 kcal/mol for the isolated atom.

For A = 12 N Ala[2] (Fig. 4), $g_{AB}(r)$ in the dipeptide and isolated-atom cases possess first peaks at the same positions, though the peak values in the dipeptide case are much lower. 12 N Ala[2] is covalently bonded with and lies among one hydrogen atom (13 H Ala[2]) and two carbon atoms (10 C Ala[1] and 18 CA Ala[2]) having certain core diameters and positive site–charges. Hence, water–hydrogens can form bonding with 12 N Ala[2] only in very limited orientations, which leads to much lower first-peak values of $g_{AB}(r)$ than in the isolated-atom case. However, even in the dipeptide case the bond formation. 12 N Ala[2]–H–O, is present [the core diameter of 12 N Ala[2] does not significantly differ from that of a water–oxygen, and the position of the first peak of $g_{AH}(r)$ is about half of that of $g_{AO}(r)$]. Compared with the hydration free energy for the isolated atom, the negative contribution

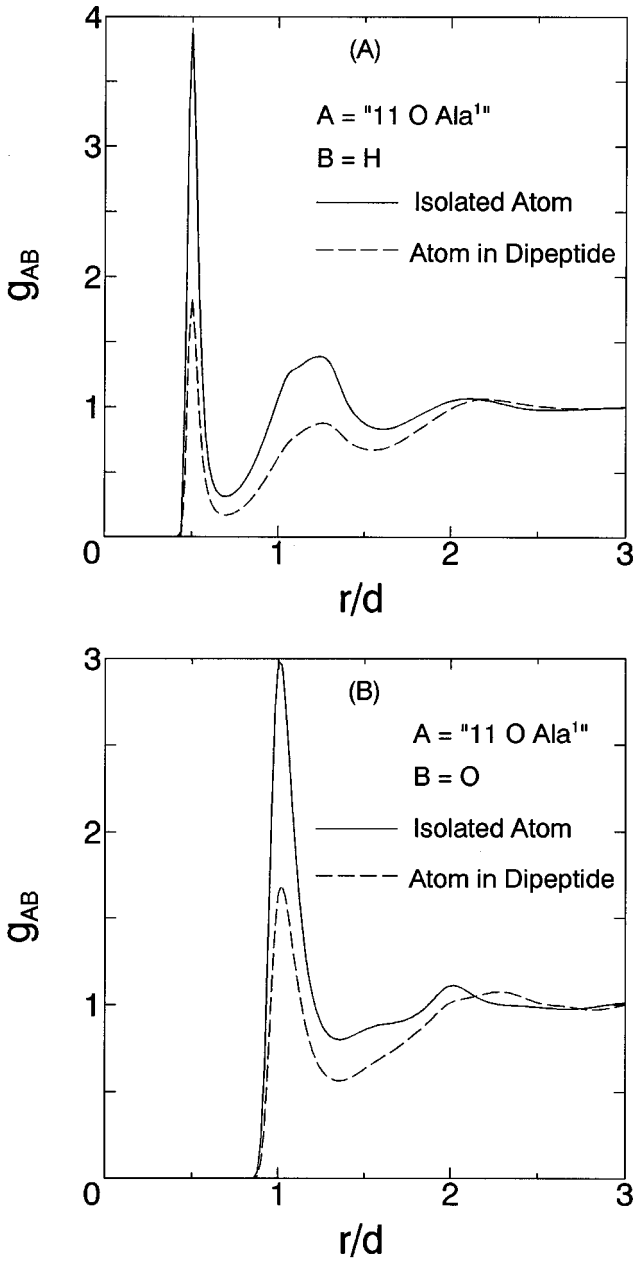

FIG. 3. Site–site pair distribution functions $g_{AB}(r)$ for $A=$''11 O Ala^1'' in two cases ($d=0.28$ nm). Atom A is isolated for one of the curves and an atom of the dipeptide (conformation 1) for the other. (a) B is a water–hydrogen. (b) B is a water–oxygen.

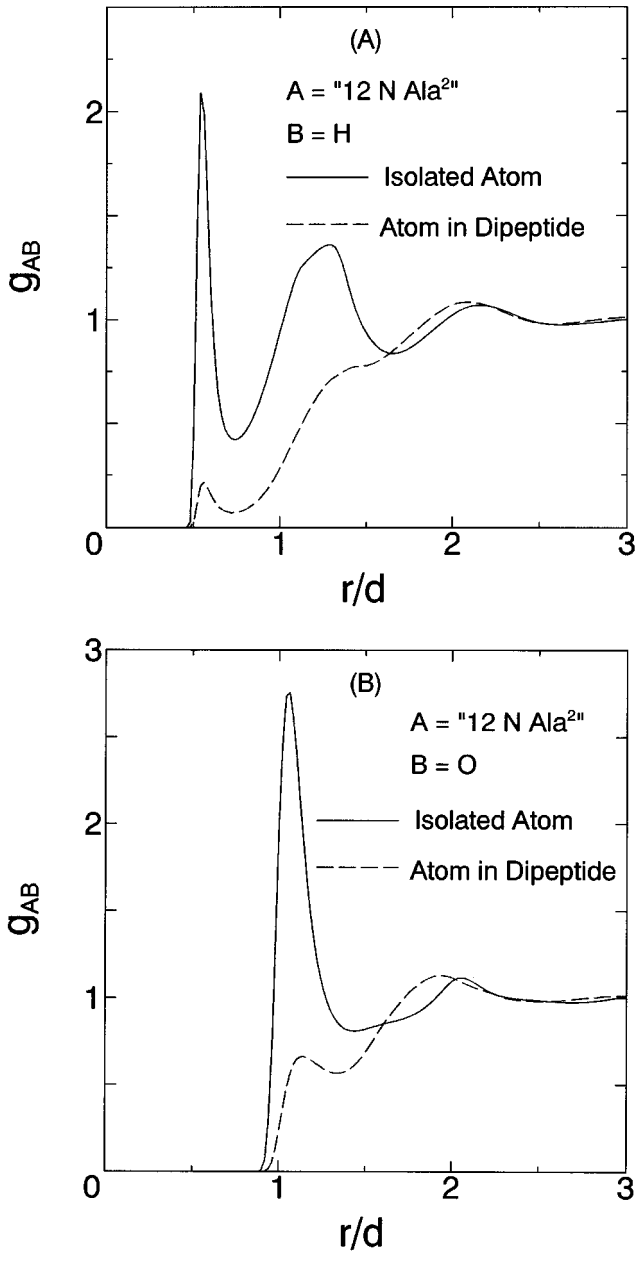

FIG. 4. Site–site pair distribution functions $g_{AB}(r)$ for $A=$''12 N Ala^2'' in two cases ($d=0.28$ nm). Atom A is isolated for one of the curves and an atom of the dipeptide (conformation 1) for the other. (a) B is a water–hydrogen. (b) B is a water–oxygen.

from 12 N Ala^2 to the hydration free energy is considerably less.

10 C Ala^1 is covalently bonded with and lies among 8 CA Ala^1 (this atom has a positive site–charge but it is very small), 11 O Ala^1, and 12 N Ala^2. The latter two atoms have negative site–charges. Moreover, water–oxygens have a large core diameter and hence they cannot approach 10 C Ala^1 close enough, which is reflected in the much lower value and the farther position of the first peak of $g_{AO}(r)$ than in the isolated-atom case (Fig. 5). While water–oxygens with large core diameters cannot approach even positively charged atoms very close when they lie among other atoms, water–hydrogens come to the close vicinity of negatively charged atoms by the rotation of water molecules. Water–hydrogens are attracted to the negatively charged atoms adjacent to 10 C Ala^1, which gives rise to enhancement of the repulsive electrostatic interaction with 10 C Ala^1. The contribution from 10 C Ala^1 to the hydration free energy is then positive and significantly large as seen in Table VI.

The contributions from carbons with negative site–charges such as 14 CB Ala^2 are often negative (Table VI), but the total contribution from the methyl group is positive and significantly large. As observed in Fig. 6, the first-peak positions of $g_{AB}(r)$ for $A=14$ CB Ala^1 in the dipeptide case are significantly farther than in the isolated-atom case. This

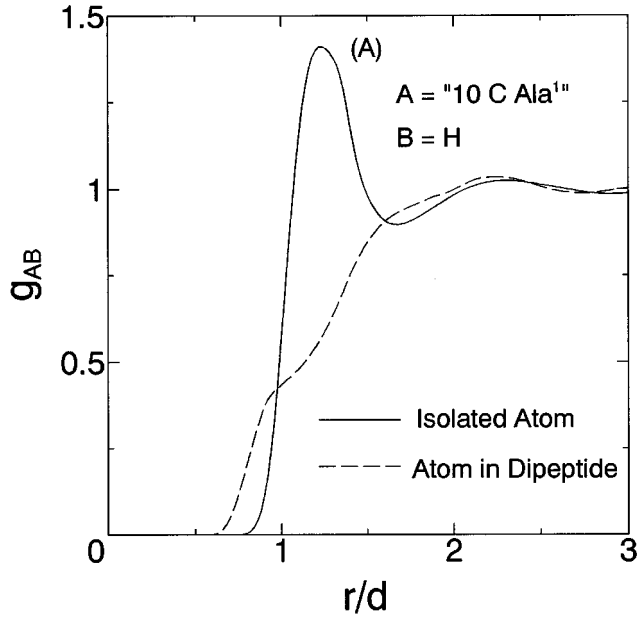

FIG. 5. Site–site pair distribution functions $g_{AB}(r)$ for $A=$''10 C Ala^1'' in two cases ($d=0.28$ nm). Atom A is isolated for one of the curves and an atom of the dipeptide (conformation 1) for the other. (a) B is a water–hydrogen. (b) B is a water–oxygen.

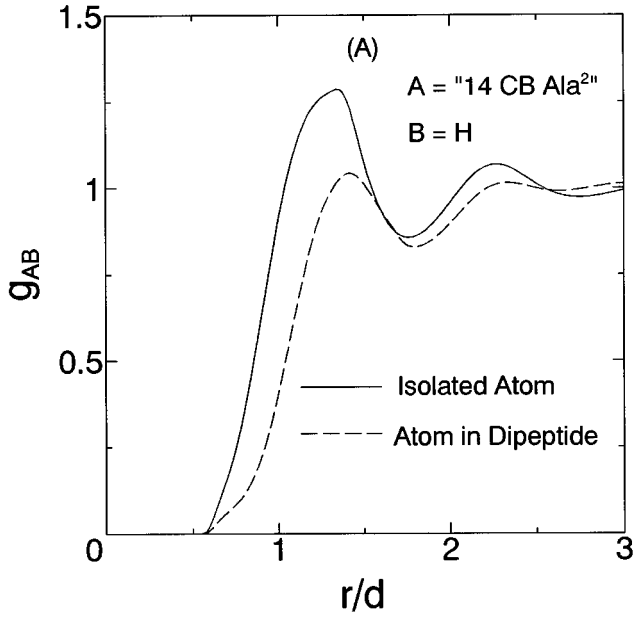

FIG. 6. Site–site pair distribution functions $g_{AB}(r)$ for $A=$''14 CB Ala^2'' in two cases ($d=0.28$ nm). Atom A is isolated for one of the curves and an atom of the dipeptide (conformation 1) for the other. (a) B is a water–hydrogen. (b) B is a water–oxygen.

is presumably because water–hydrogens are somewhat repelled from the three hydrogens of the methyl group (having positive site–charges) covalently bonded with 14 CB Ala^2.

We note that the sum of the hydration free energies of the 23 isolated atoms is -12.4 kcal/mol, which is much lower (i.e., more ''hydrophilic'') than the value of the dipeptide. Pettitt and his co-workers[11,12] applied the RISM theory to the calculation of the hydration free energy of small peptides, but they used the superposition approximation in which the entire free energy of a peptide is expressed as the sum of the potential of mean forces between pairs of isolated atoms. Our results suggest that this is a poor approximation.

Finally, the hydration free energies calculated are summarized in Table VII. The contributions from the five portions of the dipeptide [NH_2, $CHCH_3(1)$, CONH, $CHCH_3(2)$, and COOH] to the hydration free energies are also given in the table. We note that $\Delta\mu_{sb}$ and $|\Delta\mu_{sa}-\Delta\mu_{sb}|$ are measures of contributions from the hydrophobic and electrostatic interactions with water, respectively.

In general, oxygens and nitrogens have relatively large, negative contributions to the hydration free energy due to the formation of hydrogen bonding, and this is particularly true for carbonyl oxygens. The result for conformation 2 is characterized by the formation of stronger hydrogen bonding between carbonyl oxygen and water–hydrogens than for conformation 1 (Tables V and VI). As observed in Fig. 2(b),

TABLE VI. Contributions of some representative atoms to the hydration free energy (kcal/mol) in case a (the full values of the site–charges are used) for conformations 1 and 2 of the dipeptide.

Atom	Conf. 1	Conf. 2
1 H Ala[1]	3.9	4.4
2 N Ala[1]	−2.1	−2.9
5 HB1 Ala[1]	3.8	4.1
10 C Ala[1]	8.9	12.6
11 O Ala[1]	−5.8	−9.7
12 N Ala[2]	−5.7	−5.8
13 H Ala[2]	5.4	4.0
14 CB Ala[2]	−2.2	−1.2
18 CA Ala[2]	2.0	4.1
19 HA Ala[2]	3.9	3.9
20 C Ala[2]	8.8	8.8
21 O Ala[2]	−5.7	−8.5
22 O Ala[2]	−3.9	−6.1
23 H Ala[2]	5.5	7.1

11 O Ala[1] and 21 O Ala[2] in conformation 2 are close to each other and they are sufficiently far apart from the methyl group, the most hydrophobic portion of the dipeptide. Relatively strong hydrogen bonding is then formed, and conformation 2 has a much lower hydration free energy than conformation 1 (Table VII).

C. Met-enkephalin

Figures 7–10 show the site–site pair distribution functions $g_{AB}(r)$ for $A=$36 O Gly[3], 57 N Met[5], 55 C Phe[4], and 10 CE1 Tyr[1], respectively, for Met-enkephalin in conformation 4. We note that $g_{AB}(r)$ shown in Figs. 7–9 are qualitatively the same as those in Figs. 3–5, respectively. For example, Figs. 7 and 8 indicate that water–hydrogens form rather strong bonding with 36 O Gly[3] and 57 N Met[5], though the bond formation occurs in limited orientations particularly for 57 N Met[5]. The water density near 10 CE1 Tyr[1], a typical hydrophobic atom, is rather low despite that this atom is relatively well exposed to water (Fig. 10).

The first-peak values of some representative site–site pair distribution functions $g_{AB}(r)$ are given in Table VIII. Table IX gives the contributions from some representative atoms to the hydration free energy in case a for the four conformations. 13 OH Tyr[1] is bonded with and lies between one carbon atom and one hydrogen atom, so the first-peak value of $g_{AH}(r)$ is not high (Table VIII). On the other hand, 14 HH Tyr[1] is bonded only with one oxygen atom, so $g_{AO}(r)$ has a higher first-peak value. However, 13 OH Tyr[1] and 14 HH Tyr[1] have negative and positive contributions to the hydration free energy, respectively (Table IX). This is because water–oxygens with large core diameters cannot approach even positively charged atoms very close, but water–hydrogens come to the close vicinity of negatively charged atoms by the rotation of water molecules. Due to the hydrogen bonding between 14 HH Tyr[1] and 36 O Gly[3] in confor-

TABLE VII. Hydration free energies (kcal/mol) for the five portions of the dipeptide and total hydration free energy (kcal/mol). ''$\Delta\mu_{sa}-\Delta\mu_{sb}$'' for conformations 1 and 2 are −9.8 and −14.0 kcal/mol, respectively. For CONH and COOH, the difference between the two values in cases *a* and *b* is larger in conformation 2.

Conformation	NH_2	$CHCH_3(1)$	CONH	$CHCH_3(2)$	COOH	Total
1[a]	6.0	16.4	2.8	17.7	4.7	47.5
2[a]	5.6	15.0	1.1	19.1	1.3	42.0
1[b]	7.3	15.1	8.8	15.5	10.6	57.3
2[b]	7.4	13.9	9.0	16.9	8.8	56.0

[a]The full values of the site–charges are used.
[b]All the site–charges are set to zero.

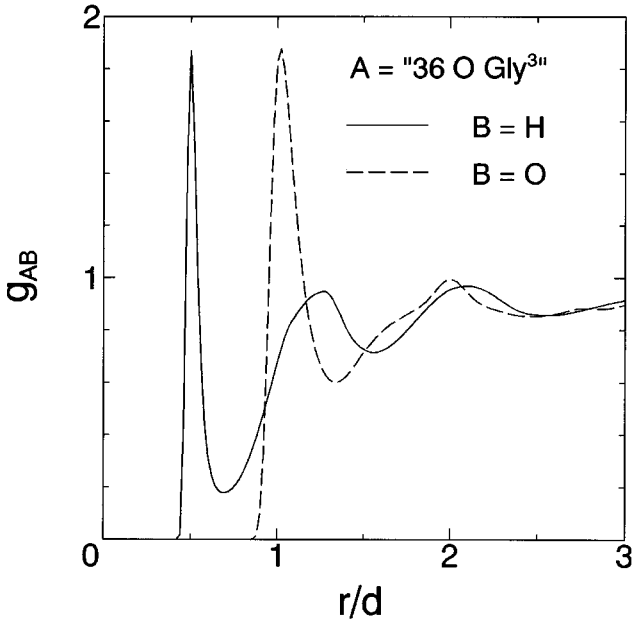

FIG. 7. Site–site pair distribution functions $g_{AB}(r)$ for $A=$''36 O Gly[3]'' of Met-enkephalin in conformation 4 ($d=0.28$ nm).

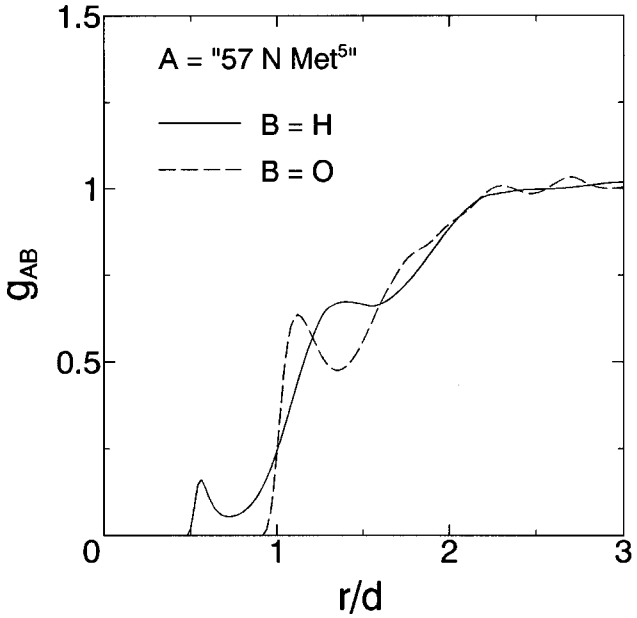

FIG. 8. Site–site pair distribution functions $g_{AB}(r)$ for $A=$''57 N Met[5]'' of Met-enkephalin in conformation 4 ($d=0.28$ nm).

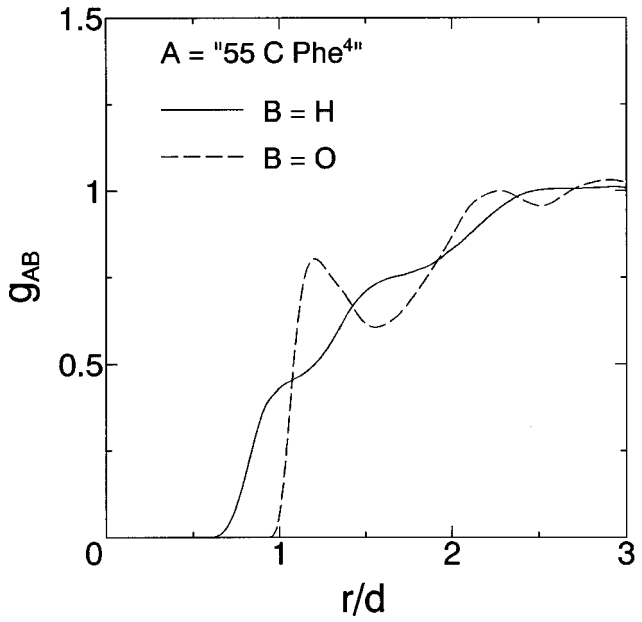

FIG. 9. Site–site pair distribution functions $g_{AB}(r)$ for $A=$''55 C Phe^4'' of Met-enkephalin in conformation 4 ($d=0.28$ nm).

mation 1, $g_{AO}(r)$ for $A=14$ HH Tyr^1 and $g_{AH}(r)$ for $A=36$ O Gly^3 have the lowest first-peaks among the five conformations (Table VIII). In conformation 1, 29 O Gly^2 and 73 O Met^5 are not well exposed to water as observed in Figs. 1(a) and 1(b), leading to lower first-peak values of $g_{AH}(r)$.

The hydration free energies calculated are summarized in Table X. (The solvation free energies for the simple, repulsive potential system tested above are 223.9, 222.1, 229.8, and 220.7 kcal/mol in conformations 1, 2, 3, and 4, respectively.) The contributions from the five residues of Met-enkephalin (Tyr^1, Gly^2, Gly^3, Phe^4, and Met^5) to the hydration free energies are also given in the table. The hydration free energies for the six backbone portions, NH_2, CONH(1), CONH(2), CONH(3), CONH(4), and COOH, are compared in Table XI. All the five carbonyl oxygens are well exposed to water in conformations 2, 3, and 4. However, the result for conformation 2 can be distinguished from those for the other four conformations due to the largest electrostatic interactions with water molecules and the resultant decrease in the hydration free energy as shown in Tables X and XI. In conformation 2, the five carbonyl oxygens are not far apart, and in particular 29 O Gly^2 and 56 O Phe^4 are close to each other. Conformation 2 of Met-enkephalin is similar to conformation 2 of the dipeptide in this respect. A relatively

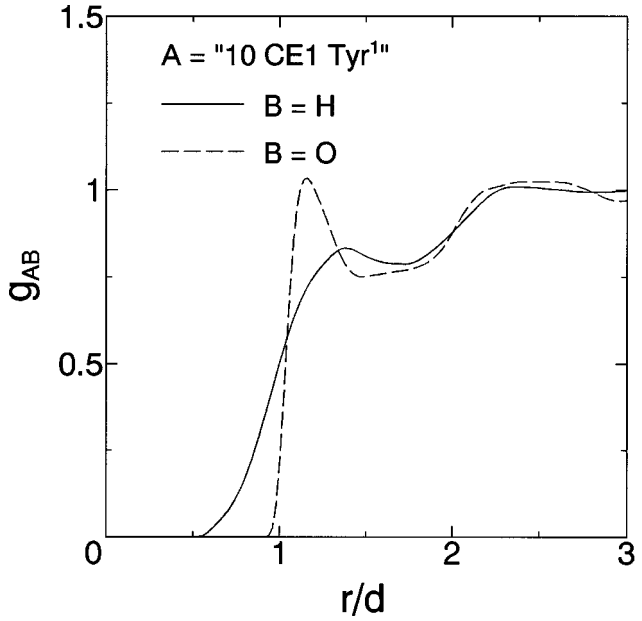

FIG. 10. Site–site pair distribution functions $g_{AB}(r)$ for $A=$''10 CE1 Tyr^1'' of Met-enkephalin in conformation 4 ($d=0.28$ nm).

TABLE VIII. First-peak values of some representative site–site radial distribution functions $g_{AB}(r)$. A and B denote atomic sites in Met-enkephalin and a water molecule, respectively.

A	B	Conf. 1	Conf. 2	Conf. 3	Conf. 4
2 N Tyr^1	H	0.54	0.44	0.52	0.51
13 OH Tyr^1	H	0.84	1.03	1.04	1.01
22 O Tyr^1	H	1.83	1.84	1.82	1.82
23 N Gly^2	H	0.12	0.18	0.18	0.23
29 O Gly^2	H	1.45	1.91	1.75	1.81
30 N Gly^3	H	0.18	0.11	0.16	0.14
36 O Gly^3	H	1.55	1.73	1.80	1.87
37 N Phe^4	H	0.14	0.19	0.17	0.17
56 O Phe^4	H	1.84	1.73	1.87	1.85
57 N Met^5	H	0.16	0.15	0.21	0.16
73 O Met^5	H	1.31	1.67	1.90	1.77
74 O Met^5	H	0.85	0.91	0.82	0.87
1 H Tyr^1	O	1.58	1.05	1.40	1.32
3 H Tyr^1	O	1.09	1.31	1.09	1.12
14 HH Tyr^1	O	1.24	1.31	1.46	1.44
55 C Phe^4	O	0.93	1.00	1.06	0.80
72 C Met^5	O	1.05	1.23	1.32	1.06
75 H Met^5	O	1.58	1.62	1.55	1.46

TABLE IX. Contributions of some representative atoms to the hydration free energy (kcal/mol) in case a (the full values of the site–charges are used) for conformations 1, 2, 3, and 4 of Met-enkephalin.

Atom	Conf.1	Conf.2	Conf.3	Conf.4
1 H Tyr^1	4.5	2.5	4.0	3.7
2 N Tyr^1	−2.1	1.2	−1.2	−1.2
4 CB Tyr^1	−0.2	−0.2	1.4	0.5
5 HB1 Tyr^1	4.8	4.0	4.7	4.2
10 CE1 Tyr^1	2.7	2.1	4.5	2.3
11 HE1 Tyr^1	4.0	3.4	3.5	3.8
13 OH Tyr^1	−3.4	−2.6	−2.6	−3.1
14 HH Tyr^1	5.3	4.3	4.7	4.6
29 O Gly^2	−4.1	−7.6	−5.4	−4.8
36 O Gly^3	−7.5	−7.1	−6.5	−5.5
37 N Phe^4	−8.6	−4.9	−2.5	−4.2
55 C Phe^4	12.3	10.2	13.5	11.6
56 O Phe^4	−6.8	−7.9	−6.1	−6.0
57 N Met^5	−4.9	−5.9	−5.3	−6.2
58 H Met^5	4.0	6.4	2.6	5.8
72 C Met^5	12.7	12.0	8.8	9.2
73 O Met^5	−5.5	−8.6	−3.1	−5.8
74 O Met^5	−4.5	−6.3	−3.9	−4.0
75 H Met^5	6.5	7.8	5.7	6.1

TABLE X. Hydration free energies (kcal/mol) for the five residues of Met-enkephalin and total hydration free energy (kcal/mol). "$\Delta\mu_{sa}-\Delta\mu_{sb}$" for conformations 1, 2, 3, and 4 are −19.6, −30.8, −25.8, and −24.2 kcal/mol, respectively.

Conformation	Tyr[1]	Gly[2]	Gly[3]	Phe[4]	Met[5]	Total
1[a]	60.2	17.4	18.8	50.9	49.4	196.8
2[a]	52.9	15.2	14.5	50.1	45.3	178.0
3[a]	61.3	17.1	17.2	56.8	50.4	202.7
4[a]	52.9	14.7	14.2	51.1	44.0	176.8
1[b]	62.0	20.3	22.8	54.2	57.1	216.4
2[b]	57.9	20.2	19.9	54.9	55.9	208.8
3[b]	65.4	21.9	21.8	60.7	58.7	228.5
4[b]	56.6	19.3	18.3	54.2	52.7	201.0

[a]The full values of the site–charges are used.
[b]All the site–charges are set to zero.

strong hydrogen bonding is formed between 29 O Gly[2] and water–hydrogens as shown in Tables VIII and IX. The hydrogen bonding for 56 O Phe[4] is weaker due to the presence of the phenyl group in Phe[4] shown in Fig. 1(c). Overall, the electrostatic interactions with water–molecules are the largest in conformation 2. Carbonyl oxygens (the most hydrophilic atoms) are less exposed to water in conformation 1 than in the others, and this is reflected in the changes given between columns a and b in Table III.

We emphasize that the hydration free energy is significantly dependent on the peptide conformations even in cases where all the site–charges are set to zero (Tables IX and X). It is obvious that water molecules cannot closely approach an atom which is not well exposed to water, which is verified by Figs. 1(b), 1(e), and 11. 58 H Met[5] in conformation 1 is less exposed to water than that in conformation 4. Hence, $g_{AB}(r)$ for $A=58$ H Met[5] in these two conformations are greatly different. However, matters are more complicated than this: Even when atom A is well exposed to water in all the conformations considered, $g_{AB}(r)$ and the contribution from this atom to the hydration free energy vary significantly depending on details of the surroundings (i.e., the peptide conformations). Last, it is interesting to note that the solvation free energy for the simple, repulsive potential system is much less dependent on the peptide conformations than in the water case.

TABLE XI. Hydration free energies (kcal/mol) for the six backbone portions of Met-enkephalin. Except for CONH(1), the difference between the two values in cases a and b is the highest in conformation 2.

Conformation	NH_2	CONH(1)	CONH(2)	CONH(3)	CONH(4)	COOH
1[a]	6.8	5.5	9.8	6.1	4.5	9.2
2[a]	7.0	5.2	4.4	5.2	2.7	4.9
3[a]	7.7	7.7	7.8	7.6	4.7	7.5
4[a]	6.6	5.3	6.1	5.6	5.2	5.4
1[b]	8.1	10.0	14.6	12.0	10.1	14.5
2[b]	8.5	11.6	12.0	12.2	10.2	12.3
3[b]	8.8	13.2	14.1	14.4	11.3	13.5
4[b]	7.8	10.5	11.8	11.5	11.5	11.9

[a]The full values of the site–charges are used.
[b]All the site–charges are set to zero.

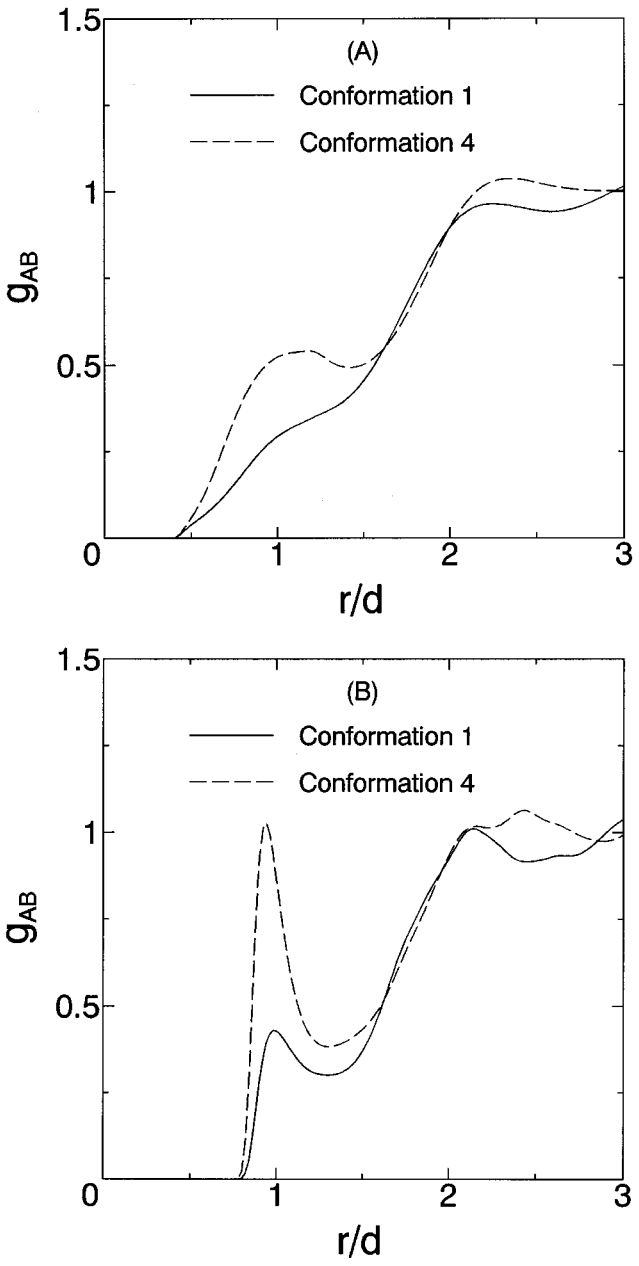

FIG. 11. Site–site pair distribution functions $g_{AB}(r)$ for $A=$"58 H Met[5]" in conformations 1 and 4. All the site–charges of Met-enkephalin are set to zero ($d=0.28$ nm). (a) B is a water–hydrogen. (b) B is a water–oxygen.

D. Zwitterions

In the aqueous solution buffered to pH=3.87 used in the NMR experiments,[16] Met-enkephalin should be present as a zwitterion with zero net charge. However, NH_3^+ and COO^- can be screened by counterions, CH_3COO^- and Na^+, respectively. Hence, it is probable that the chemical form of Met-enkephalin in the aqueous solution is in effect an intermediate of the unionized molecule and the zwitterion.

The zwitterions with zero net charges are now considered. The hydration free energies for the dipeptide are summarized in Table XII with the contributions from the five portions. Table XIII gives the hydration free energies for Met-enkephalin and the contributions from the five residues.

TABLE XII. Hydration free energies (kcal/mol) for the five portions of the dipeptide (zwitterion) and total hydration free energy (kcal/mol). This table should be compared with Table VII.

Conformation	NH_3^+	$CHCH_3(1)$	CONH	$CHCH_3(2)$	COO^-	Total
1[a]	0.7	17.4	1.3	18.2	−28.8	8.7
2[a]	2.1	16.1	−1.6	19.6	−36.2	−0.1

[a]The full values of the site–charges are used.

The hydration free energies for the six backbone portions, NH_3^+, CONH(1), CONH(2), CONH(3), CONH(4), and COO^-, are given in Table XIV. For the dipeptide, the sum of the hydration free energies for NH_3^+ and COO^- in the zwitterion case is about 40 kcal/mol lower than that for NH_2 and COOH in the unionized case. For Met-enkephalin, the former is 38–56 kcal/mol lower than the latter. As expected, the total hydration free energies decrease greatly by the ionization. The degree of the decrease for the dipeptide is larger in conformation 2 than in conformation 1. This is because the three oxygens are close together in conformation 2 [Fig. 2(b)].

Figures 12 and 13 show $g_{AB}(r)$ for $A=74$ O Met[5] at the C-terminus and 1 H Tyr[1] at the N-terminus, respectively, for Met-enkephalin in conformation 4. These figures imply the formation of strong hydrogen bonding between 74 O Met[5] and water–hydrogens and between 1 H Tyr[1] and water–oxygens in the zwitterion case. The former bonding is particularly strong, which results in a very large, negative hydration free energy.

Since 73 O Met[5] is not well exposed to water in conformation 1 of Met-enkephalin [Fig. 1(a)], the decrease in the hydration free energy by the ionization is less than in the other three conformations (Table XIV). In contrast, the ionization leads to the largest decrease in the hydration free energy in conformation 2 where oxygens in CONH(1), CONH(2), CONH(3), CONH(4), and COO^-, are not far apart [Fig. 1(c)]. The total energies for conformations 2, 3, and 4 relative to the total energy for conformation 1 are compared in Table XV. The conformational energy for conformation 2 is very high due to Coulombic repulsions among the like-charged atoms which are close to one another. Hence, despite the lowest hydration free energy, the total energy for conformation 2 is considerably higher than that for conformation 4. The most stable and unstable conformations are conformations 4 and 3, respectively. Thus, the qualitative aspects of the conclusions are not significantly altered by the ionization.

TABLE XIII. Hydration free energies (kcal/mol) for the five residues of Met-enkephalin (zwitterion) and total hydration free energy (kcal/mol). This table should be compared with Table X.

Conformation	Tyr[1]	Gly[2]	Gly[3]	Phe[4]	Met[5]	Total
1[a]	58.1	17.7	18.2	50.5	13.4	157.9
2[a]	39.0	15.3	12.9	49.9	−1.8	115.4
3[a]	53.6	18.1	18.3	56.7	11.1	157.7
4[a]	42.8	15.9	14.7	51.9	1.1	126.5

[a]The full values of the site–charges are used.

TABLE XIV. Hydration free energies (kcal/mol) for the six backbone portions of Met-enkephalin (zwitterion). This table should be compared with Table XI.

Conformation	NH_3^+	CONH(1)	CONH(2)	CONH(3)	CONH(4)	COO^-
1[a]	1.9	7.7	8.7	5.1	2.6	−25.5
2[a]	−4.9	4.3	2.2	2.3	0.3	−40.1
3[a]	0.0	9.6	8.0	7.7	2.5	−31.4
4[a]	−3.2	6.9	6.1	4.8	3.2	−36.3

[a]The full values of the site–charges are used.

Last, we comment on the reported shortcomings of the RISM-HNC theory. The theory tends to give too large values of hydration free energies for nonpolar solutes.[10,26] Also, the theory often violates the stoichiometry of the coordination numbers of water around mono- and multivalent ions calculated from the ion–oxygen and ion–hydrogen pair distribution functions.[27,28] However, these shortcomings do not raise any serious problem as long as we are concerned only with the relative values between two different peptide conformations. Besides, there is no comparable theoretical alternative which allows us to analyze the solvation structure of peptides at the same microscopic level.

V. CONCLUSION

The full RISM equations have been solved for a dipeptide and Met-enkephalin in the SPC/E water[15] using our robust, highly efficient algorithm. Some different conformations of these peptides have been considered. The site–site pair distribution functions and hydration free energies calcu-

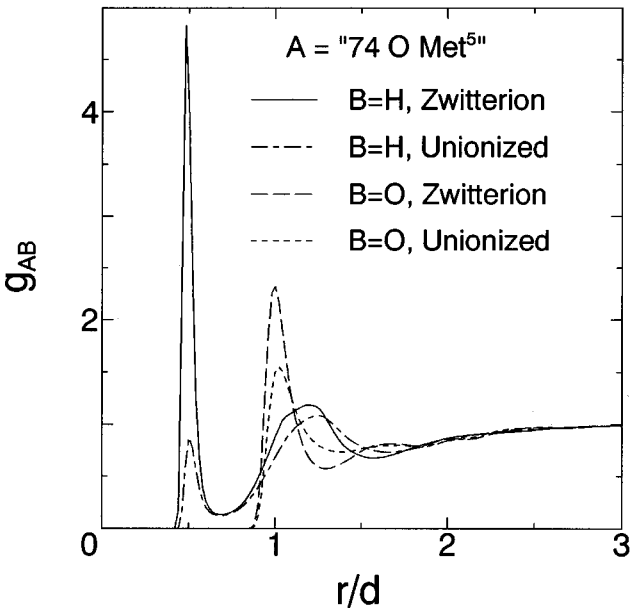

FIG. 12. Site–site pair distribution functions $g_{AB}(r)$ for $A=$ ''74 O Met[5]'' of Met-enkephalin (zwitterion) in conformation 4 ($d=0.28$ nm). The contribution of this atom to the hydration free energy is −31.0 kcal/mol. The functions in the unionized case are also shown for comparison.

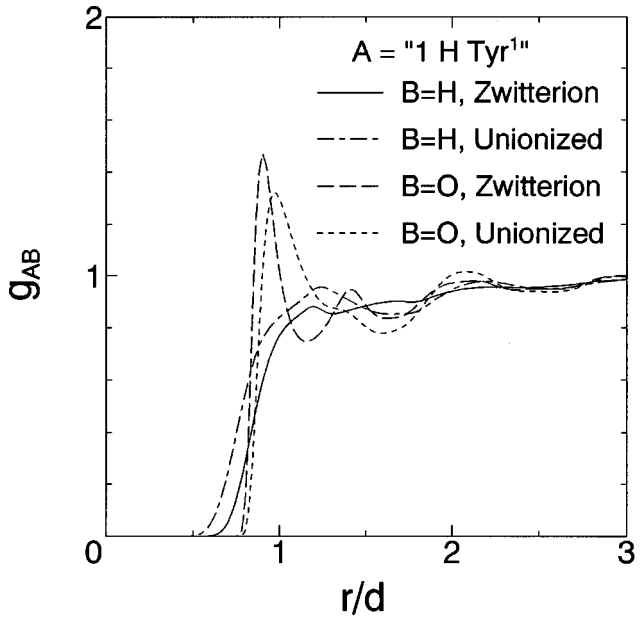

FIG. 13. Site–site pair distribution functions $g_{AB}(r)$ for $A=$''1 H Tyr^1'' of Met-enkephalin (zwitterion) in conformation 4 ($d=0.28$ nm). The contribution of this atom to the hydration free energy is -2.9 kcal/mol. The functions in the unionized case are also shown for comparison.

lated as major output data have been analyzed in detail. It has been shown that solvent plays essential roles in determining the conformation of the peptide and that the RISM theory is a promising tool for taking account of the solvent effects.

The most stable conformation of the peptide in water is the one that has the lowest total energy. The total energy is determined not only from the conformational energy but from the interactions with water molecules which are greately dependent on the peptide conformations. The conformations of Met-enkephalin determined in NMR experiments[16] are quite different from the lowest-energy conformation in gas phase.[5] We have tested four different conformations including the lowest-energy conformation (conformation 1) and a conformation which is similar to the experimentally determined conformations (conformation 4). It has been shown that conformation 4 is the most stable in water with the lowest total energy in both of the unionized and zwitterion cases. (Actually, we have tested more conformations than we described in the present article, and still conformation 4 is the most stable.) It is interesting that a conformation which is similar to that obtained from the NMR experiments in miceller solutions, is the least stable when it is put in water. Although the effects due to 0.05 M CH_3COONa in the solution used in the NMR experiments are unknown and need to be investigated in further studies, our results are quite encouraging. When the SPC/E water is replaced by a simple, repulsive potential system, the lowest-energy conformation in gas phase is still the most stable among the four conformations, because the solvation free energy is much less dependent on the conformations than in the water case. Water is clearly distinguished from the simple solvent even when all the site–charges of the peptide are set to zero.

The site–site pair distribution function $g_{AB}(r)$ for atom A (B is a water–hydrogen or oxygen) and the contribution from this atom to the hydration free energy is greatly dependent on the neighboring atoms. An atom with a large (negative or positive) site–charge is covalently bonded with oppositely charged atoms with certain core diameters in most cases. Consequently, the first-peak values of $g_{AB}(r)$ for carbonyl carbons, oxygens, and nitrogens are much lower than those for the isolated atoms (imaginary spherical particles) treated in Sec. IV A. Compared with the hydration free energies of the isolated atoms, the contributions from these atoms of the peptide to the hydration free energy are considerably shifted in more hydrophobic directions. The superposition approximation, in which the entire free energy of a peptide is expressed as the sum of the potential of mean forces between pairs of isolated atoms, is a poor approximation.

The carbonyl oxygen is covalently bonded only with one atom (carbonyl carbon), and it often forms very strong bonding with water–hydrogens. In general, oxygens and nitrogens often have relatively large, negative contributions to the hydration free energy due to the formation of hydrogen bonding. This is particularly true for the two oxygens at the C-terminus of zwitterions. When more than two carbonyl oxygens are close together, well exposed to water, and at least one of them is sufficiently far apart from a hydrophobic portion, strong hydrogen bonding is formed between the carbonyl oxygen and water–hydrogens.

The hydration free energy for a portion of the peptide is also greatly dependent on the neighboring portions. For example, the value for CONH of the unionized dipeptide is in the range from 1–3 kcal/mol, while that of Met-enkephalin (unionized) is more variable, ranging from 2–10 kcal/mol. The value for COOH of the dipeptide (1–5 kcal/mol) is smaller than that of Met-enkephalin (4–10 kcal/mol). (We repeat that we have tested more conformations than we describe in the present article.) This is because Met-enkephalin has larger hydrophobic portions (e.g., the phenyl group) and they are often close to CONH or COOH.

In the course of the present study, we have noticed the

TABLE XV. Total energies for conformations 2, 3, and 4 of Met-enkephalin (zwitterion) relative to the total energy for conformation 1 (kcal/mol). Column a: In the case where the full values of the site–charges are used. Column c: In gas phase [(i.e., conformational energies relative to the conformational energy for conformation 1 (kcal/mol)]. The changes (kcal/mol) are also given with arrows. This table should be compared with Table III.

Conformation	a		c
1	0.0		0.0
2	−2.4	$\overset{-42.5}{\leftarrow}$	40.1
3	15.3	$\overset{-0.2}{\leftarrow}$	15.5
4	−10.9	$\overset{-31.4}{\leftarrow}$	20.5

following. When a hydrophobic atom gets very close to a hydrophilic atom, the hydration free energy for the former decreases while that for the latter increases. However, the sum of the two hydration free energies tends to increase (i.e., shift in a more hydrophobic direction). Though this could be an artifact of the RISM theory and needs to be investigated further, we are inclined to think that for a larger peptide the contribution from the hydrophobic interaction with water becomes larger. In fact, the values of $(\Delta\mu_{sb}-\Delta\mu_{sa})/\Delta\mu_{sb}$ ($\Delta\mu_{sb}$ and $|\Delta\mu_{sa}-\Delta\mu_{sb}|$ are measures of contributions from the hydrophobic and electrostatic interactions with water, respectively) are around 0.2 for the unionized dipeptide (0.17 and 0.25 in conformations 1 and 2, respectively), but they are only ~0.1 for Met-enkephalin (unionized) (0.09, 0.15, 0.11, and 0.12 in the four conformations, respectively). In the zwitterion case, the values of $(\Delta\mu_{sb}-\Delta\mu_{sa})/\Delta\mu_{sb}$ are around 0.9 for the dipeptide while they are ~0.35 for Met-enkephalin.

We are now combining the solution of the full RISM equations with powerful conformational sampling methods[1,2] to find the lowest-energy conformation of a peptide in water. This can be done with moderate computational effort on a workstation because the algorithm used for solving the RISM equations is robust and extremely fast. The extension to a small protein or inclusion of ions in water are also under way.

ACKNOWLEDGMENT

The present study was supported by the Joint Studies Program (1996) of the Institute for Molecular Science (IMS), ''Drastic Reduction in Computation Time for Hydration Free Energy of Protein Using the RISM Theory''.

[1] S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi, Science **220**, 671 (1983).
[2] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991); Phys. Rev. Lett. **68**, 9 (1992).
[3] S. R. Wilson, W. Cui, J. W. Moskowitz, and K. E. Schmidt, Tetrahedron Lett. **29**, 4373 (1988).
[4] H. Kawai, T. Kikuchi, and Y. Okamoto, Protein Engng. **3**, 85 (1989).
[5] U. H. E. Hansmann and Y. Okamoto, J. Comput. Chem. **14**, 1333 (1993).
[6] M. H. Hao and H. A. Scheraga, J. Phys. Chem. **98**, 4940 (1994).
[7] Y. Okamoto, Proteins: Struc. Funct. Genet. **19**, 14 (1994).
[8] D. Chandler and H. C. Andersen, J. Chem. Phys. **57**, 1930 (1972).
[9] F. Hirata and P. J. Rossky, Chem. Phys. Lett. **83**, 329 (1981); F. Hirata and R. M. Levy, Int. J. Quantum Chem. **15**, 179 (1988).
[10] M. Kinoshita and F. Hirata, J. Chem. Phys. **106**, 5202 (1997).
[11] B. M. Pettitt and M. Karplus, J. Am. Chem. Soc. **107**, 1166 (1985); B. M. Pettitt and M. Karplus, Chem. Phys. Lett. **121**, 194 (1985).
[12] G. L. Ramé, W. F. Lau, and B. M. Pettitt, Int. J. Peptide Protein Res. **35**, 315 (1990).
[13] A. Kitao, F. Hirata, and N. Go, J. Phys. Chem. **97**, 10233 (1993).
[14] M. Kinoshita, Y. Okamoto, F. Hirata, J. Comput. Chem. (in press).
[15] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, J. Phys. Chem. **91**, 6269 (1987).
[16] W. H. Graham, E. S. Carter II, and R. P. Hickes, Biopolymers **32**, 1755 (1992).
[17] Z. Li and H. A. Scheraga, Proc. Natl. Acad. Sci. USA. **84**, 6611 (1987).
[18] Y. Okamoto, T. Kikuchi, and H. Kawai, Chem. Lett. **1992**, 1275.
[19] J. Perkyns and B. M. Pettitt, Chem. Phys. Lett. **190**, 626 (1992).
[20] J. Perkyns and B. M. Pettitt, J. Chem. Phys. **97**, 7656 (1992).
[21] H. A. Yu, B. Roux, and M. Karplus, J. Chem. Phys. **92**, 5020 (1990).
[22] H. Kawai, Y. Okamoto, M. Fukugita, T. Nakazawa, and T. Kikuchi, Chem. Lett. **1991**, 213; Y. Okamoto, M. Fukugita, T. Nakazawa, and H. Kawai, Protein Engng. **4**, 639 (1991).
[23] F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga, J. Phys. Chem. **79**, 2361 (1975).
[24] G. Némethy, M. S. Pottle, and H. A. Scheraga, J. Phys. Chem. **87**, 1883 (1983).
[25] M. J. Sipple, G. Némethy, and H. A. Scheraga, J. Phys. Chem. **88**, 6231 (1984).
[26] T. Ichiye and D. Chandler, J. Phys. Chem. **92**, 5257 (1988).
[27] R. A. Kuharski and D. Chandler, J. Phys. Chem. **91**, 2978 (1987).
[28] T. Ichiye and A. D. J. Haymet, J. Chem. Phys. **89**, 4315 (1988).