Modeling of the Three-Dimensional Structure of Polypeptides in Solution Using Potential-Scaled/Hot-Solute Molecular Dynamics

Hideki Tsujishita,* Ikuo Moriguchi,‡ and Shuichi Hirono‡
*New Drug Research Laboratories, Kanebo Ltd., Osaka 534, and ‡School of Pharmaceutical Sciences, Kitasato University, Tokyo 108 Japan

ABSTRACT We present here an efficient and accurate procedure for modeling of the three-dimensional structures of polypeptides in the explicit solvent water based on molecular dynamics calculations. Using the toxic domain analog of heat-stable enterotoxin as a model peptide, we examined the utilities of two molecular dynamics techniques with the system containing the explicit solvent. One is the potential-scaled molecular dynamics that had been designed for effective conformational analyses of biomolecules with the explicit solvent water by partially scaling down the potential energies involved in the solute molecules. The other is the variation of Berendsen's weak coupling method (referred to as "hot-solute" method), in which only the solute of the system is heated to a high temperature while the solvent is kept at a normal temperature. Each method successfully increased the rate of folding of the peptides, and the most effective was a combination of the two methods. Moreover, the final structure obtained via cooling process successfully reproduced the experimentally known structure from the extended amino acid sequence using only the distance restraints representing three disulfide bonds in the peptide. Additional distance restraints derived from some of the NOE cross peaks accelerated the folding of the peptide, but gave almost the same structure as in the case without these additional restraints. Because a similar calculation without the explicit solvent could not reproduce the known structure, it is suggested that the explicit solvent water could play an important role in the modeling. The methods presented here have the potential for accurate modeling even when less experimental information was available.

INTRODUCTION

Molecular dynamics (MD) have great utility in the modeling of the three-dimensional structures of biomolecules such as proteins or nucleotides. In MD, the kinetic energy of atoms allows the system to surmount energy barriers surrounding local minima existing on the energy surface. In the recent years, MD with simulated annealing (Kirkpatrick et al., 1983) protocol has been widely used in the modeling of biomolecules, especially in the refinement of three-dimensional structures derived from x-ray or NMR experiments (Brünger et al., 1987; Brünger et al., 1989; Holak et al., 1989; Saulitis et al., 1992; Miller et al., 1992). This method includes a high temperature process to increase the ability to surmount energy barriers. However, such a technique cannot be directly applied to the system containing explicit solvent water because raising the temperature would cause unrealistic behavior of water (extremely high pressure or vaporization of water). Therefore, another efficient modeling method is needed for treating the system containing explicit water. Such a modeling method would give a more reliable three-dimensional structure because it is generally suggested that MD with explicit solvent gives more accurate results (Levitt and Sharon, 1988; van Gunsteren, 1992).

We have previously investigated the "potential-scaled molecular dynamics" method as an effective method for conformational analyses in explicit water (Tsujishita et al., 1993). In this method, the potential energies involved in the desired degrees of freedom (those of solute molecules in our case) are partially scaled down, instead of raising the temperature of the whole system. This is a kind of umbrella sampling, and it can accelerate the conformational change in the solute molecule without unfavorable influences on the behavior of the solvent. In the previous report, the method was applied to a small dipeptide analog and proved to be capable of accurately and more efficiently searching conformational space than the normal MD. In this paper, we attempt the modeling of the solution structure of a larger peptide, the full toxic domain analog of heat-stable enterotoxin STp, Mpr^5-STp(5-17), using the potential-scaled MD with explicit water. Heat-stable enterotoxins (STs, Fig. 1) are produced by enterotoxigenic *Escherichia coli* and 13 residues near the C-terminus are needed for full enterotoxic activity (Aimoto et al., 1982, 1983; Takao et al., 1983; Yoshimura et al., 1985). This toxic domain includes six cysteines that form three disulfide linkages, and the mode of the disulfide bonds has been determined (Shimonishi et al., 1987; Gariépy et al., 1987). Structures of STs in solution and in crystal have been explored in detail by x-ray crystallography (Ozaki et al., 1991b) or NMR experiments (Gariépy et al., 1987; Ohkubo et al., 1986; Gariépy et al., 1986; Ozaki et al., 1991a) and the coordinate of the three-dimensional structure in crystal is available for Mpr^5-STp(5-17) (Ozaki et al., 1991b). In addition to the potential-scaled method, we also tested another method (referred to as "hot-solute" technique hereafter) in which only the solute of the system is heated to a high temperature while the solvent is kept at a

Received for publication 29 November 1993 and in final form 28 February 1994.

Address reprint requests to Shuichi Hirono, School of Pharmaceutical Sciences, Kitasato University, 9-1, Shirokane 5 Chome, Minato-Ku, Tokyo 108 Japan. Tel.: 011-81-3-3443-7780; Fax: 011-81-3-3440-5246.

Abbreviations used: MD, molecular dynamics; NOE, nuclear Overhauser effect; ST, heat-stable enterotoxin; Mpr, β-mercaptopropionic acid; RMS, root mean square; R_g, radius of gyration.

0006-3495/94/06/1815/08 $2.00

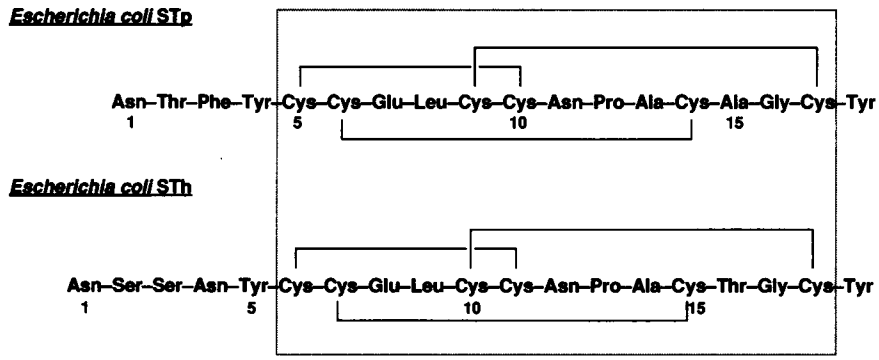

FIGURE 1 Amino acid sequences of heat-stable enterotoxins produced by *Escherichia coli*. Lines between cysteines represent disulfide linkages. The toxic domains for fully enterotoxic activity are displayed by the rectangle.

normal temperature. This method is also expected to accelerate the conformational change in the solute molecule without unexpected behavior of the solvent and to be useful as an efficient modeling tool in explicit water.

In the following sections, we will examine the efficiency of both methods compared to the normal MD. The predicted structure will be compared with the x-ray and NMR experimental data, and the accuracy of the methods will also be discussed.

COMPUTATIONAL PROCEDURES

All calculations in this paper were performed with AMBER version 4.0.1 (Pearlman et al., 1991) and its all-atom force field (Weiner et al., 1984; Weiner and Kollman. 1986). The hot-solute algorithm and the distance-restraint function used here (see below) were incorporated into the program by us.

Potential-scaled MD

The idea of potential-scaled MD was originally found in the work by Mark et al. (1991). The details of the method have been previously described (Tsujishita et al., 1993). Briefly, this method is based on the fact that both raising the temperature and decreasing the potential energy at a constant temperature have the same effect on the Boltzmann factor, which governs the population of conformers in a system. This means that a larger conformational space of the system can be searched by scaling down its potential energy as well as by raising its temperature. The potential-scaled method is quite useful particularly in MD with explicit water. Scaling down the potential only in the degrees of freedom involved in the solute molecules leads to the effective conformational search of the solute without unexpected behavior of the solvent water. In the following calculations, only the bond angle and dihedral terms in the potential energy function were reduced by 2.0 in the potential-scaled MD. This procedure of potential scaling has been proved to accelerate the conformational change of the solute without any unfavorable influences on the solvent waters (Tsujishita et al., 1993).

Hot-solute method

To maintain the temperatures of the solute and solvent at different values, we modified Berendsen's weak-coupling algorithm to control the temperature of the system (Berendsen et al., 1984). This technique couples the temperature of the system T to an external heat bath of reference temperature, T_0, by scaling the velocities of atoms using a factor λ. The factor λ is determined by the reference and current temperatures and the temperature relaxation time, τ_T, that controls the strength of the coupling. When the system is divided into two parts, A and B, and atoms i and j are involved in parts A and B, respectively, the temperatures of A and B can be kept at different values by scaling the velocities of atoms i and j using different factors, λ^A and λ^B. These factors are determined by the following equations:

$$\frac{3N^A}{2}k_B T^A(t) = E^A_{kin} = \sum_{i=1}^{N^A} \frac{1}{2} m_i \{v_i(t)\}^2, \quad (1)$$

$$\frac{3N^B}{2}k_B T^B(t) = E^B_{kin} = \sum_{j=1}^{N^B} \frac{1}{2} m_j \{v_j(t)\}^2, \quad (2)$$

$$\lambda^A = \left[1 + \frac{\Delta t}{\tau^A_T}\left(\frac{T^A_0}{T^A(t)} - 1\right)\right]^{1/2}, \quad (3)$$

$$\lambda^B = \left[1 + \frac{\Delta t}{\tau^B_T}\left(\frac{T^B_0}{T^B(t)} - 1\right)\right]^{1/2}, \quad (4)$$

where N is the number of atoms, k_B is the Boltzmann constant, E_{kin} is the kinetic energy, Δt is the integration time step, and m and v are the mass and velocity of an atom, respectively. τ_T is the temperature relaxation time that controls the strength of the coupling. The temperatures of part A (T^A) and part B (T^B) are calculated independently from the sum of the kinetic energies of the atoms in each part. The scaling factors, λ^A and λ^B, are then obtained from these temperatures and the reference temperatures, T^A_0 and T^B_0. In our case, part A corresponds to solute molecules and part B to solvent. In the following calculations that includes this protocol, we set the reference temperature for the solute to 900 K and that for the solvent to 300 K.

Distance-restraint function

Test calculations in this paper include some distance restraints (see below). As the restraint function, we adopted the function proposed by Nilges et al. (1988) described in Fig. 2. The force derived from this function is largest at $r = r_s$ and small enough when the restrained atoms are far apart. Therefore, in the early stage of the calculations (distances between the restrained atoms may be large), the forces of the restraints are negligibly small. The restrained atoms experience significant forces only when the distance between them approaches r_s during the course of the calculation, and the restraints will converge at this time. This profile of the function is expected to avoid the incorrect folding that would occur when using the ordinary square-well type function. In the following calculations, the slope of asymptote c was set to zero.

Initial structure

The initial structure of Mpr^5-STp(5-17) for the test calculations was generated in the following manner. The extended structure (all backbone dihedrals ϕ and ψ were set to 180°) of STp was first generated using the standard database of AMBER. The molecule was solvated by TIP3P water molecules (Jorgensen et al., 1983), which were added within 10 Å of the surface of the solute. This system was then minimized until the root mean square (RMS) value of the potential gradient was below 0.05 kcal mol^{-1} Å^{-1} (0.2 kJ mol^{-1} Å^{-1}). After the minimization, a normal MD calculation was

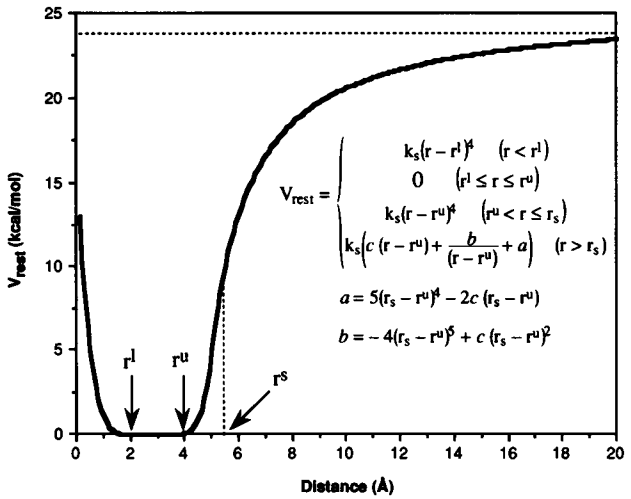

FIGURE 2 The restraint function (Nilges et al., 1988) used in our calculations. In the figure, r is the current distance, r^l and r^u are the lower and upper limits of the distance, respectively, and k_s is the force constant. The values of a and b are chosen such that the potential V_{rest} is continuous and differentiable at r_s. c is the slope of the asymptote, and the dashed line in the figure represents this asymptote.

carried out for 10 ps using an integration time step Δt = 2 fs. The temperature of the whole system was maintained at 300 K by the Berendsen's algorithm with τ_T = 0.1 ps. The bond lengths were fixed by the SHAKE algorithm (van Gunsteren and Berendsen. 1977). This MD run was performed to equilibrate the system and to reduce the maximum length of the solute. This results in saving computational time needed for the following simulations, because the required number of water molecules can be reduced. The maximum length of the solute was 33.30 Å after the MD run, whereas it was 40.38 Å before the run. The backbone dihedral angles were slightly shifted from their initial values, but no significant folding was observed. The water molecules were removed after the MD, and the solute was then solvated again by 2233 TIP3P water molecules in a rectangular box whose dimensions were 53.6x39.3x36.8 Å. The system was then minimized until the RMS value of the potential gradient was below 0.03 kcal mol^{-1} Å^{-1} (0.13 kJ mol^{-1} Å^{-1}) under the periodic boundary condition. The resulting system was used as the starting point of the following tests.

Tests of modeling

We performed four kinds of tests to evaluate the efficiency of the methods in modeling the structure of the molecule. These were A (normal MD), B (potential-scaled MD), C (hot-solute MD), and D (potential-scaled/hot-solute MD). All tests included distance restraints representing three disulfide linkages in STp. We carried out another test of the potential-scaled/hot-solute MD (test E) that included distance restraints derived from NOEs in aqueous solution observed in the NMR experiment (Ozaki et al., 1991a) in addition to the restraints for the disulfide bonds. There were 29 NOE cross peaks observed in the NMR study but we considered only six NOEs derived from the atom pairs that were separated by more than one amino acid residue in the sequence (Table 1). The purpose of this test calculation was to investigate how the efficiency and the accuracy of the methods will change when more experimental information was included in the calculation. For each disulfide bond, there were three restraints, S_i-S_j, S_i-C_j^β and C_i^β-S_j, whose target values were set to 2.02 ± 0.02, 2.99 ± 0.5 and 2.99 ± 0.5 Å, respectively. The force constant, k_s, was 0.05 kcal mol^{-1} Å^{-4} (0.2 kJ mol^{-1} Å^{-4}), and r_s, the distance at which the restraint force was maximum, was set to a value 3.0 Å longer than the upper limit of each restraint. The NOE restraints in test E used the target values shown in Table 1 and the same k_s and r_s values. The initial velocities in the whole system were chosen from a Maxwellian distribution at 300 K, and in the tests for hot-solute, the reference temperature for the solute was gradually increased from 300 to 900 K in the first 5 ps. The temperature relaxation time τ_T was set to 0.1 ps both for the solute and solvent. All of the test calculations were carried out for 50 ps using Δt = 1.0 fs, except for test E, which was performed for 30 ps. The coordinates were saved every 100 steps for the analysis. During the calculations, the pressure of the system was maintained at 1 atm using the Berendsen's algorithm under periodic boundary condition, and the bond lengths were fixed by SHAKE. We set force constants of torsional parameters for backbone amide linkages to a value 2.5 times larger than the standard values in AMBER to avoid *cis-trans* isomerizations. The nonbonded interactions beyond 9.0 Å were truncated in all of the calculations. The analysis of the effectiveness of the methods was made by a comparison of the rate of convergence of the value of radius of gyration (R_g) of STp into that in the x-ray structure.

TABLE 1 Additional distance restraints included in test E

Atom pairs*		Target values for restraints (Å)	
		Lower limit	Upper limit
Cys^6 H_α	Cys^9 H_α	2.0	4.0
Cys^6 H_α	$Asn^{11}H_N$	2.0	4.0
Cys^6 H_α	$Cys^{14}H_\beta$	2.0	5.0
Cys^9 H_α	$Asn^{11}H_N$	2.0	3.5
$Asn^{11}H_\beta$	$Ala^{13}H_N$	2.0	4.0
$Ala^{15}H_\alpha$	$Cys^{17}H_N$	2.0	4.0

* These restraints were derived from 6 out of a total of 29 NOE cross peaks observed in H_2O/D_2O (80:20) at 5 or 17°C, pH 5.6 (Ozaki et al., 1991a). These six restraints were chosen in a manner such that the restrained atoms were separated by more than one amino acid residue in the sequence.

Final structures

The final structures were obtained from the results of the potential-scaled/hot-solute MD (tests D and E) through the cooling process as follows. After each test was run, a successive calculation of the potential-scaled/hot-solute MD was done for 15 ps, increasing the force constant k_s of the restraints by multiplying its value by $10^{0.1}$ after every 500 steps. The value of k_s was 50.0 kcal mol^{-1} Å^{-4} (210 kJ mol^{-1} Å^{-4}) at the end of the run. A following MD run was performed for 10 ps, gradually cooling the heated solute from 900 to 300 K and increasing the level of the scaled potential to its original level. All distance restraints were then removed, and a 10-ps normal MD calculation was carried out, considering the disulfide linkages as explicit covalent bonds. For the last 5 ps, the coordinates were dumped every 50 steps. The saved coordinates were then averaged and minimized until the RMS of the potential gradient was below 0.05 kcal mol^{-1} Å^{-1} (0.2 kJ mol^{-1} Å^{-1}). The obtained structures were used in the following discussion. The three-dimensional structure figures in this paper were created using UCSF MidasPlus software (Ferrin et al., 1988).

Test of an implicit solvent model

We carried out an additional calculation without the explicit solvent water. This test, denoted test F, used a distance-dependent dielectric constant as a solvent model, instead of explicit water. The test was performed to evaluate the effect of the explicit solvent on the final structure. The calculation started from the same initial structure, and it was carried out under the same conditions and protocols as in test D. However, the periodic boundary condition was not applied due to the lack of the explicit solvent and a distance-dependent dielectric constant $\epsilon = 80r$ was used. The final structure was obtained in the same manner described above.

RESULTS

Effects of the hot-solute method

At first, we investigated if the hot-solute method worked correctly. Fig. 3 indicates changes in the temperatures of the

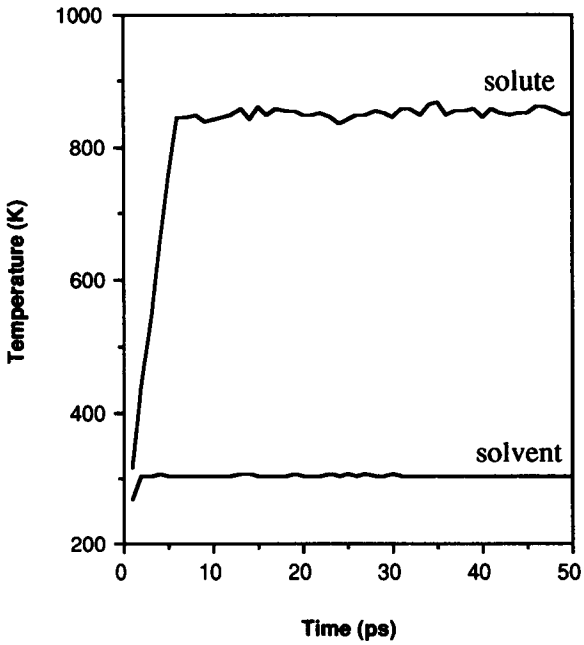

FIGURE 3 Change in the temperatures of the solute and the solvent in the system during the calculation of the potential-scaled/hot-solute MD (test D). The values were sampled every 1 ps.

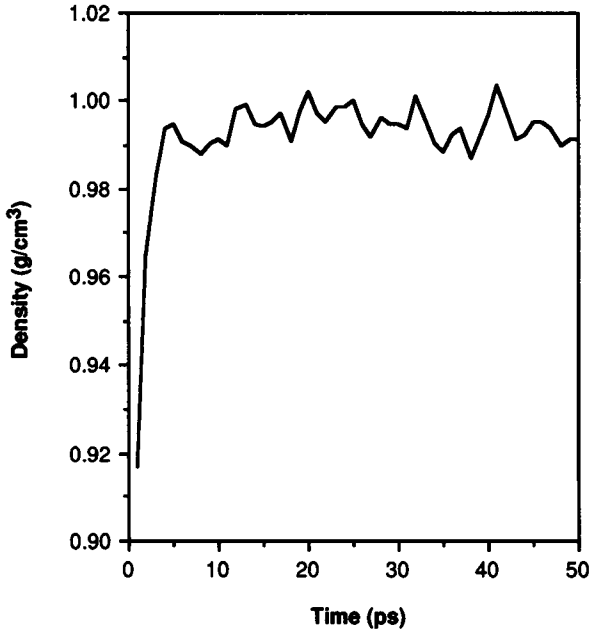

FIGURE 4 Change in the density of the system during the calculation of the potential-scaled/hot-solute MD (test D). The values were sampled every 1 ps.

solute and the solvent during the 50-ps run of the potential-scaled/hot-solute MD (test D). The temperature of the solute increased rapidly, and it was maintained at the external bath value for the solute without affecting the temperature of the solvent. The average values of the temperatures during the 10–50 ps were 852.43 ± 49.02 K for the solute and 304.02 ± 2.64 K for the solvent, and well reflected the values of the external baths. Fig. 4 displays the density of the system observed in test D as a function of time. Although part of the system was heated, the density of the entire system was maintained at the natural level for water (nearly 1 g cm^{-3}). These results indicate that the method was capable of only heating the solute without influencing the solvent's behavior.

Effects of the distance restrain function

We monitored the distances between the sulfide atoms that were constrained to examine the utility of the restrain function used here. Fig. 5 shows the change in these distances in the course of the test calculation D. This figure clearly indicates that the restrained atoms slightly felt the restraint forces during the early period of the calculation when the atoms were far apart. As they were brought close together in the course of the calculation, they were gradually captured by the restraint forces and finally the distances between them fell into the target values. These behaviors of the restraint function indicate that the function worked as expected.

Comparison of the rate of folding

Fig. 6 displays the change in radius of gyration (R_g) of STp for tests A–D. During the normal MD (test A), the change

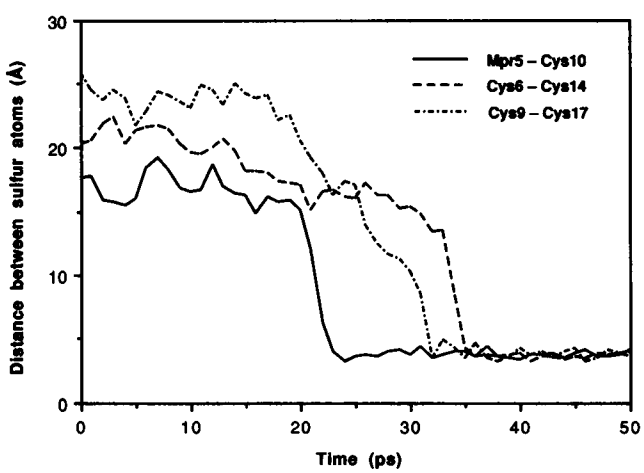

FIGURE 5 Effect of the restraint potential function in the calculation of the potential-scaled/hot-solute MD (test D). Each point represents the distance between the sulfur atoms of Mpr^5-Cys^{10} (———), Cys^6-Cys^{14} (– – –), and Cys^9-Cys^{17} (·····), which were restrained to form a disulfide bond, respectively.

in the R_g value was very small and the structure of the molecule retained nearly linear over the 50-ps calculation. The rate of folding during the potential-scaled MD (test B) was faster than that during the normal MD. This result proves that the potential-scaled MD was effective in this test system as well as in our previous case. We next examined the hot-solute method (test C). The rate of folding was greatly increased during this test. The molecule began to fold near 20 ps in the calculation, and the R_g took a value close to that in the x-ray structure (approximately 6 Å) at the end of the calculation.

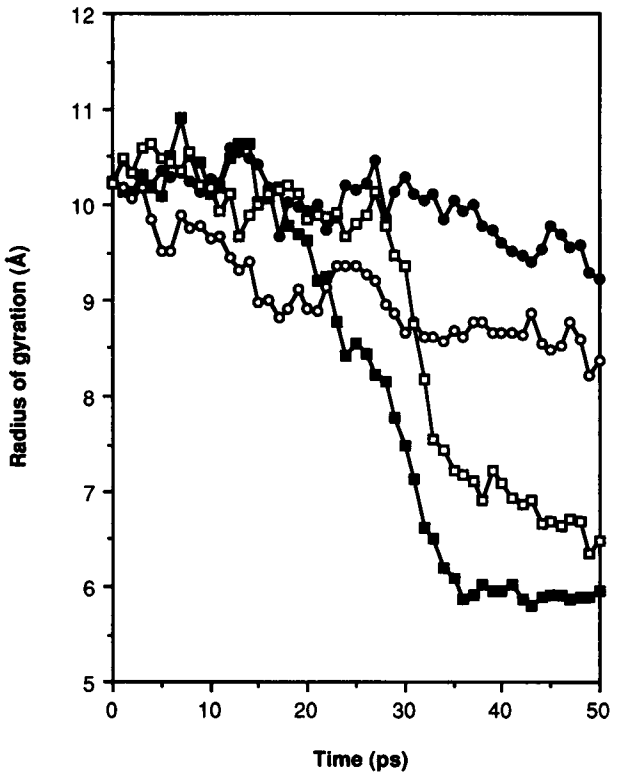

FIGURE 6 Change in the radius of gyration of Mpr[5]-STp(5-17). ●, normal MD (test A); ○, potential-scaled MD (test B); □, hot-solute MD (test C); ■, potential-scaled/hot-solute MD (test D). The radius of gyration was calculated using the heavy atoms in STp.

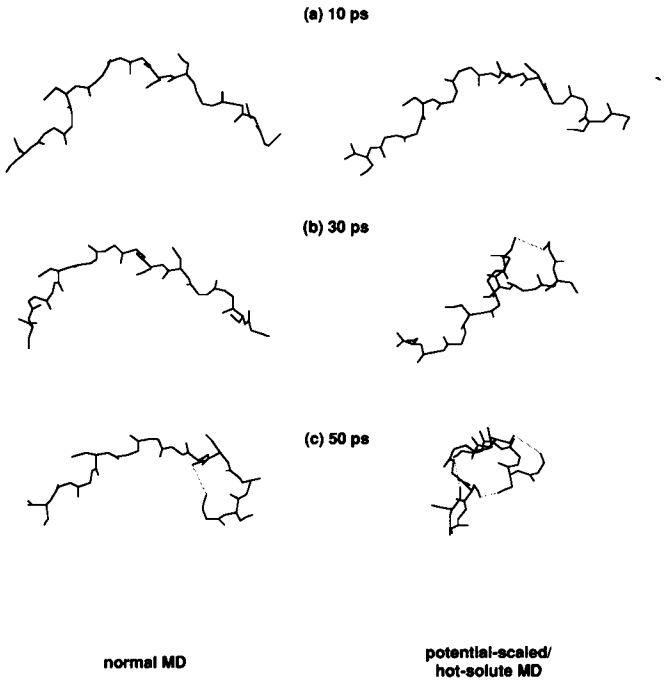

FIGURE 7 Snapshots of the structures of Mpr[5]-STp(5-17). (*a*) 10 ps, (*b*) 30 ps, and (*c*) 50 ps in the calculations of the normal MD (test A, *left*) and the potential-scaled/hot-solute MD (test D, *right*). Only the backbone atoms and the side-chain atoms of the Cys (or Mpr) residues are displayed. Dotted lines indicate the converged distance restraints.

The fastest acceleration of the folding was observed in the potential-scaled/hot-solute MD (test D). As indicated in Fig. 6, the molecule had already shifted to the globular state at 35 ps. Moreover, the molecule retained its R_g value until the end of the calculation once the folding had been completed, and no unfolding was observed, at least during this calculation period. Fig. 7 is the snapshots of the molecule at 10, 30, and 50 ps during the normal MD (test A) and the potential-scaled/hot-solute MD (test D). As suggested by the analysis of the R_g values, the molecule folded much faster in test D. In summary, these results suggested that the combination of the potential-scaled MD and the hot-solute technique could be the most effective of all the tests.

Evaluation of the predicted structure of STp

To examine the accuracy of the methods, we compared the predicted structure of STp using test D with the structure observed in the x-ray crystallography analysis (Ozaki et al., 1991b). The predicted structure superimposed on the x-ray structure is displayed in Fig. 8. The predicted structure exhibited a right-handed spiral shape, which was the same folding pattern as in the x-ray structure. There were three β-turns between Cys[6]-Cys[9], Asn[11]-Cys[14], and Cys[14]-Cys[17], which were observed in both the x-ray and NMR experiments (Ozaki et al., 1991a), and these turns were also observed in the predicted structure. The structural properties for both

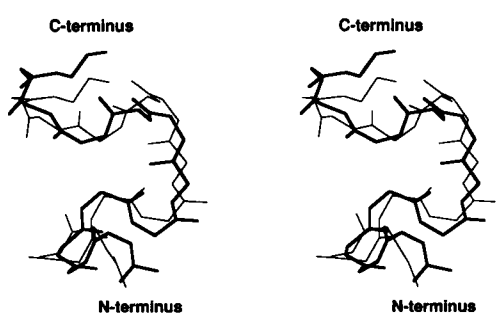

FIGURE 8 Stereo view of the resulting structure from the calculation of the potential-scaled/hot-solute MD (test D, *thick lines*) and the x-ray structure of Mpr[5]-STp(5-17) (*thin lines*). Only the backbone atoms are displayed. Structural superposition was made so that the root mean square of the deviations of the backbone atoms (N, C_α, C, O) could have the smallest value.

structures are summarized in Table 2. The R_g value was at the same level in the x-ray structure. The values of the root mean square of the deviations from the x-ray structure were 2.26 Å for the heavy atoms, 1.45 Å for the backbone atoms, and 1.44 Å for the α-carbons. These values indicate that the predicted structure was closely related to the x-ray structure. In addition, the predicted structure satisfied most of the proton-proton distances derived from NOEs observed in aqueous solution and violated only 2 of the 29 NOEs. The violated NOEs were those observed between H_N-Glu[7] and H_N-Leu[8], and between H_α-Ala[15] and H_N-Cys[17]. These NOEs belong to the loops in the N- and C-termini of STp, respec-

TABLE 2 Comparison of the structural properties obtained from the potential-scaled/hot-solute MD simulations of Mpr^5-STp(5-17) with experimental data

Property	Test D	Test E*	Test F‡	X-ray§
RMS of deviation (Å)‖				
All-atoms	2.26	2.36	3.89	—
Main chain	1.45	1.43	2.74	—
α-carbons	1.44	1.42	2.50	—
Mean deviation of main-chain dihedral (degree)‖	41.7	44.4	58.7	—
Radius of gyration (Å)	5.82	5.76	5.45	5.54
NOE violations¶	2	1	9	0

* Calculated with six additional distance restraints derived from NOEs (see text).

‡ Calculated without explicit solvent, using a distance-dependent dielectric constant (see text).

§ Estimated from the x-ray structure of Mpr^5-STp(5-17) (Ozaki et al., 1991b).

‖ Deviations from the x-ray structure.

¶ Number of violations of upper limit distances derived from 29 NOEs observed in H_2O/D_2O (80:20) at 5 or 17°C, pH 5.6 (Ozaki et al., 1991a).

tively, and these regions were suggested to be rather mobile in aqueous solution (Ozaki et al., 1991a).

Effects of the additional distance restraints

We performed another test of the potential-scaled/hot-solute MD, including additional distance restraints derived from the NOE cross peaks (test E). We considered not all, but only 6 of the 29 NOEs in this test (Table 1). Fig. 9 shows the change in R_g values in test E, compared with that in test D. This figure indicated that additional restraints remarkably accelerated the folding of the molecule. The molecule had already converged in the globular state at 20 ps. The predicted structure from test E and its structural properties compared with the x-ray structure are displayed in Fig. 10 and Table 2, indicating that test E also successfully reproduced the x-ray structure. It should be noted that the additional restraints had little effect on the resulting structures. Both tests D and E generated almost the same structures (Fig. 11), and their structural properties were very similar (Table 2), except the number of NOE violations. The additional restraints reduced the number of violations from two to one. The satisfied NOE in test E was the one observed between the α proton of Ala^{15} and amide proton of Cys^{17}, and it was included in the restraints (Table 1). This satisfied NOE contributed to improvement in the structure of the C-terminal region.

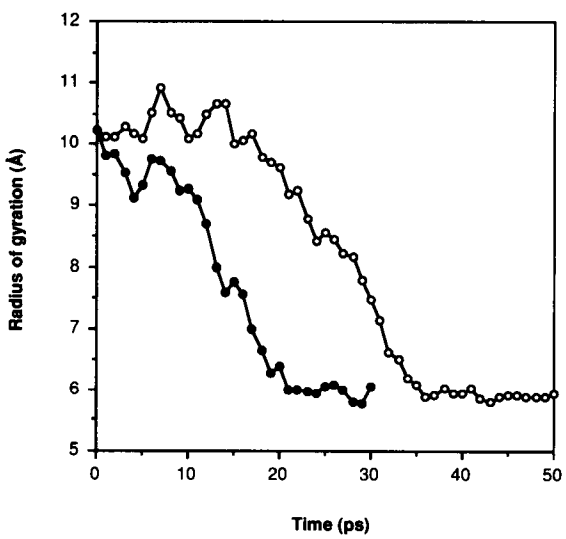

FIGURE 9 Change in the radius of gyration of Mpr^5-STp(5-17). (○) The potential-scaled/hot-solute MD containing distance restraints for disulfide linkages (test D); (●) The calculation containing the restraints in Table 1 in addition to the restraints for disulfide bonds (test E). The radius of gyration was calculated using the heavy atoms in STp.

The role of the explicit solvent

The properties of the final structure obtained from the calculation without the explicit solvent (test F) were also sum-

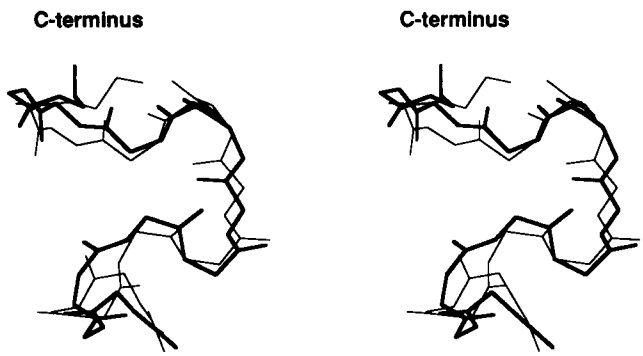

FIGURE 10 Stereo view of the resulting structure from the calculation of the potential-scaled/hot-solute MD including the additional NOE restraints (test E, *thick lines*) and the x-ray structure of Mpr^5-STp(5-17) (*thin lines*). Only the backbone atoms are displayed. Structural superposition was made so that the root mean square of the deviations of the backbone atoms (N, C_α, C, O) could have the smallest value.

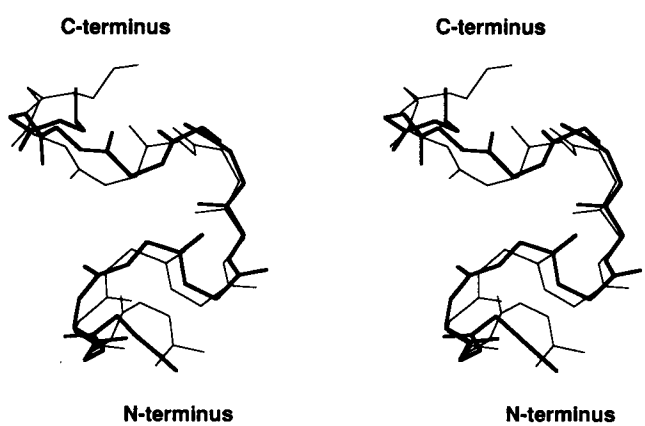

FIGURE 11 Stereo view of the resulting structure of Mpr^5-STp(5-17) from the calculation of the potential-scaled/hot-solute MD including the additional NOE restraints (test E, *thick lines*) and the simulation not including these restraints (test D, *thin lines*). Only the backbone atoms are displayed. Structural superposition was made so that the root mean square of the deviations of the backbone atoms (N, C_α, C, O) could have the smallest value.

marized in Table 2. The smaller R_g value indicates that the resulted structure was more compact than those from the calculations with explicit solvent. This may be due to the lack of the explicit solvent, where intramolecular electrostatic interactions and hydrogen bonding would be estimated too strong. The lack of the solvent also affects the quality of the resulted structure. The structural deviations from the x-ray structure were larger, and more NOE violations occurred, compared with the structures including the explicit solvent. These results strongly suggest that the explicit solvent water would play an important role in reproducing the correct structure and that the implicit (distance-dependent dielectric) solvent model used in test F is still insufficient to compensate the presence of the solvent.

DISCUSSION

The purpose of this study is to develop an efficient and accurate procedure for the modeling of three-dimensional structures of biomolecules in the explicit solvent water. Starting from the extended structure of the test peptide Mpr^5-STp-(5-17), we examined how efficiently and exactly the methods would predict its globular conformation in solution. The most efficient protocol tested in this study was the combination method of the hot-solute and the potential-scaled MD. The final structure obtained from this method exhibits good agreement with both the three-dimensional structure from the x-ray analysis (Ozaki et al., 1991b) and the inter-proton distances derived from NOE cross peaks observed in the NMR study (Ozaki et al., 1991a). These results reflected the accuracy of the methods as tools for the structure modeling. Because our calculations were performed in solution, the three-dimensional structure used in comparison should be the one in solution, not in crystal. We used, however, the crystal structure as the reference because the coordinates of the structure in solution were not available in the literature. The NMR study pointed out that the conformation of STp in solution was almost the same as that in the solid state (Ozaki et al., 1991a). In addition, our structure agreed with the NOE data observed in solution, and the predicted structure reflected the native conformation in aqueous solution.

A main part of our calculation strategy is the potential-scaled or hot-solute process, in which the potential terms in the solute are partially scaled down and the temperature of the solute is kept high. In this process, the entropy of the solute is enlarged and many conformers of the solute would appear in the system. This may correspond to high-temperature state in the simulated annealing and makes it possible to search a wider conformational space. In addition, the distance restraints related to the experimental data work as a guide in this process and help the system to reach the preferable state. The restraint function used here (Fig. 2) begins to work when the system comes close to its target value and then traps the system at the target point. As a result, the solute has completely folded at the end of this process (Figs. 6, 7, and 9). This main process is then followed by the "cooling" process, in which the reduced potential and the elevated temperature are gradually changed to their original levels. The role of this process is to bring the system to a normal equilibrium and to obtain a fine structure.

It is noteworthy that the method successfully predicted the structure of STp, only with the distance restraints representing the disulfide bridges. The additional distance restraints derived from NOEs fairly improved the efficiency of the method, but the resulting structure was almost the same one from the calculation without the additional restraints. This means that the method has the potential for accurate modeling even when less experimental information is available. This would be due to the presence of the explicit solvent water. The solvent water plays an important role in the structural formation of proteins by affecting the electrostatic interactions, hydrogen bonding, and so on (Levitt and Sharon, 1988; van Gunsteren, 1992). The presence of water would make the method more accurate and compensate for the lack of needed information. This is supported by the results summarized in Table 2. The calculation including the explicit solvent (test D) could successfully reproduce the experimentally known structure of the peptide. On the other hand, absence of the explicit water (test F) distorted the resulted structure and caused large deviations from the real structure. These results strongly suggested that the presence of the explicit solvent significantly contributes to the accuracy of the methods. Although the detailed behavior of the solvent molecules in the hot-solute state is unknown, the solvent is still considered to retain its effect on the solute. Moreover, our protocol includes the cooling process to obtain the final structure, where the temperature and the potential energy of the solute are gradually changed to the normal level. In this state, the presence of the solvent might be more critical.

REFERENCES

Aimoto, S., T. Takao, Y. Shimonishi, S. Hara, T. Takeda, Y. Takeda, and T. Miwatani. 1982. Amino acid sequence of a heat-stable enterotoxin produced by human enterotoxigenic *Eschelichia coli*. *Eur. J. Biochem.* 129:257–263.

Aimoto, S., H. Watanabe, H. Ikemura, Y. Shimonishi, T. Takeda, Y. Takeda, and T. Miwatani. 1983. Chemical synthesis of a highly potent and heat-stable analog of an enterotoxin produced by a human strain of enterotoxigenic *Eschelichia coli*. *Biochem. Biophys. Res. Commun.* 112: 320–326.

Berendsen, H. J. C., J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684–3690.

Brünger, A. T., M. Karplus, and G. A. Petsko. 1989. Crystallographic refinement by simulated annealing: application to crambin. *Acta Cryst.* A45:50–61.

Brünger, A. T., J. Kuriyan, and M. Karplus. 1987. Crystallographic R factor refinement by molecular dynamics. *Science.* 235:458–460.

Ferrin, T. E., C. C. Huang, L. E. Jarvis, and R. Langridge. 1988. The MIDAS display system. *J. Mol. Graphics.* 6:13–27.

Gariépy, J., A. K. Judd, and G. K. Schoolnik. 1987. Importance of disulfide bridges in the structure and activity of *Escherichia coli enterotoxin ST1b*. *Proc. Natl. Acad. Sci. USA.* 84:8907–8911.

Gariépy, J., A. Lane, F. Frayman, D. Wilbur, W. Robien, G. K. Schoolnik, and O. Jardetzky. 1986. Structure of toxic domain of the *Escherichia coli* heat-stable enterotoxin ST I. *Biochemistry.* 25:7854–7866.

Holak, T. A., M. Nilges, and H. Oschkinat. 1989. Improved strategies for the determination of protein structures from NMR data: the solution structure of acyl carrier protein. *FEBS Lett.* 242:218–224.

Jorgensen, W. L., J. Chandrasekhar, J. Madura, R. W. Impey, and M. L.

Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

Kirkpatrick, S., C. D. Gelatt, Jr., and M. P. Vecchi. 1983. Optimization by simulated annealing. *Science.* 220:671–680.

Levitt, M., and R. Sharon. 1988. Accurate simulation of protein dynamics in solution. *Proc. Natl. Acad. Sci. USA.* 85:7557–7561.

Mark, A. E., W. F. van Gunstern, and H. J. C. Berendsen. 1991. Calculations of relative free energy via indirect pathways. *J. Chem. Phys.* 94: 3808–3816.

Miller, K. E., C. Mukhopadhyay, P. Cagas, and C. A. Bush. 1992. Solution structure of the Lewis X oligosaccharide determined by NMR spectroscopy and molecular dynamics simulations. *Biochemistry.* 31:6703–6709.

Nilges, M., A. M. Gronenborn, A. T. Brünger, and G. M. Clore. 1988. Determination of three-dimensional structures of proteins by simulated annealing with interproton distance restrains. Application to crambin, potato carboxypeptidase inhibitor and barely serine proteinase inhibitor 2. *Protein Eng.* 2:27–38.

Ohkubo, T., Y. Kobayashi, Y. Shimonishi, Y. Kyogoku, W. Braun, and N. Go. 1986. A conformational study of polypeptide in solution by ^{1}H-nmr and distance geometry. *Biopolymers.* 25:S123-S134.

Ozaki, H., H. Kubota, T. Sato, Y. Hidaka, H. Tamaoki, Y. Kobayashi, Y. Kyogoku, T. Sugimura, A. Tai, and Y. Shimonishi. 1991a. Conformation in solution of the fully toxic domain of heat-stable enterotoxin (STp) produced by enterotoxigenic *Eschelichia coli. Bull. Chem. Soc. Jpn.* 64: 1136–1144.

Ozaki, H., T. Sato, H. Kubota, Y. Hata, Y. Katsube, and Y. Shimonishi. 1991b. Molecular structure of the toxic domain of heat-stable enterotoxin produced by a pathogenic strain of *Escherichia coli.* A putative binding site for a binding protein on rat intestinal epithelial cell membranes. *J. Biol. Chem.* 266:5934–5941.

Pearlman, D. A., D. A. Case, J. C. Caldwell, G. L. Seibel, U. C. Singh, P. Weiner, and P. A. Kollman. 1991. AMBER 4:0. University of California, San Francisco.

Saulitis, J., D. F. Mierke, G. Byk, C. Gilon, and H. Kessler. 1992. Conformation of cyclic analogues of substance-P - NMR and molecular dynamics in dimethyl sulfoxide. *J. Am. Chem. Soc.* 114:4818–4827.

Shimonishi, Y., Y. Hidaka, M. Koizumi, M. Hane, S. Aimoto, T. Takeda, T. Miwatani, and Y. Takeda. 1987. Mode of disulfide bond formation of a heat-stable enterotoxin (STh) produced by a human strain of enterotoxigenic *Escherichia coli. FEBS Lett.* 215:165–170.

Takao, T., T. Hitouji, S. Aimoto, Y. Shimonishi, S. Hara, T. Takeda, Y. Takeda, and T. Miwatani. 1983. Amino acid sequence of a heat-stable enterotoxin isolated from enterotoxigenic *Eschelichia coli* strain. *FEBS Lett.* 152:1–5.

Tsujishita, H., I. Moriguchi, and S. Hirono. 1993. Potential-scaled molecular dynamics and potential annealing: effective conformational search techniques for biomolecules. *J. Phys. Chem.* 97:4416–4420.

van Gunsteren, W. F., and A. E. Mark. 1992. On the interpretation of biochemical data by molecular dynamics computer simulation. *Eur. J. Biochem.* 204:947–961.

van Gunsteren, W. F., and H. J. C. Berendsen. 1977. Algorithms for macromolecular dynamics and constraint dynamics. *Mol. Phys.* 34: 1311–1327.

Weiner, S. J., and P. A. Kollman. 1986. An all atom force field for simulations of proteins and nucleic acids. *J. Comput. Chem* 7:230–252.

Weiner, S. J., P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, and P. Weiner. 1984. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* 106:765–784.

Yoshimura, S., H. Ikemura, H. Watanabe, S. Aimoto, Y. Shimonishi, S. Hara, T. Takeda, T. Miwatani, and Y. Takeda. 1985. Essential structure for full enterotoxigenic activity of heat-stable enterotoxin produced by enterotoxigenic *Escherichia coli. FEBS Lett.* 181:138–142.