ELSEVIER

Temperature dependence of distributions of conformations of a small peptide

Ayori Mitsutake,* Ulrich H. E. Hansmann,† and Yuko Okamoto*

**Department of Functional Molecular Science, Graduate University for Advanced Studies, and Department of Theoretical Studies, Institute for Molecular Science, Okazaki, Aichi, Japan*
†Department of Physics, Michigan Technological University, Houghton, Michigan USA

Multicanonical Monte Carlo simulations of the pentapeptide Met-enkephalin were used to study its low-energy conformations in detail. The resulting conformations are classified into six categories of similar structures based on the pattern of intrachain hydrogen bonds. Several thermodynamic quantities such as the distributions of hydrogen bonds and those of backbone dihedral angles were obtained as a function of temperature. From these results, it was concluded that at least four of the six categories are well-defined local minimum energy states. These four categories are in agreement with our prior results based on root-mean-square interatomic distances. © 1999 by Elsevier Science Inc.

INTRODUCTION

The pentapeptide Met-enkephalin in gas phase is one of the biomolecules most frequently studied by computer simulations.[1–12] The global minimum energy conformation in gas phase is known[2] and the low-energy conformations have been studied in detail.[4,5,9,12] However, detailed, systematic analyses of the conformations as a function of temperature have yet to be carried out for completeness. It is the purpose of the present article to do so. The method we adopt is the multicanonical algorithm,[13] which allows one to obtain various thermodynamic quantities as a function of temperature. In this article we first classify the low-energy conformations of Met-enkephalin into several groups of similar structures. We then present the distributions of conformations, hydrogen bonds, and dihedral angles as a function of temperature, which gives much information about the free energy landscape of Met-enkephalin.

The Color Plate for this article is on pages 262–263.

Address reprint requests to: Dr. Yuko Okamoto, Department of Functional Molecular Science, Graduate University for Advanced Studies, Okazaki, Aichi 444-8585, Japan.

E-mail addresses: ayori@ims.ac.jp (A. Mitsutake); hansmann@mtu.edu (U.H.E. Hansmann); okamotoy@ims.ac.jp (Y. Okamoto).

METHODS

Multicanonical algorithms

Although the algorithm is explained in detail elsewhere (see, e.g., Refs. 6, 7, and 13), we briefly summarize the idea and implementations of the method for completeness. In the multicanonical (mu) ensemble, the probability distribution $P_{\mathrm{mu}}(E)$ is defined in such a way that a configuration with any energy enters with equal probability [Eq. (1)]:

$$P_{\mathrm{mu}}(E) \propto n(E) w_{\mathrm{mu}}(E) = \text{constant} \tag{1}$$

where $n(E)$ is the density of states. All energies have equal weight, and a one-dimensional random walk in energy space is realized, which ensures that the system can overcome any energy barrier. It then follows that the multicanonical weight factor $w_{\mathrm{mu}}(E)$ should have the form shown in Eq. (2):

$$w_{\mathrm{mu}}(E) \propto n^{-1}(E) \tag{2}$$

Unlike in the case for the canonical ensemble, the multicanonical weight factor $w_{\mathrm{mu}}(E)$ is not a priori known, and an estimation is needed for a numerical simulation. The multicanonical weight factor $w_{\mathrm{mu}}(E)$ is calculated from iterations of short multicanonical simulations. This procedure is summarized as follows:

1. Perform a short canonical Monte Carlo simulation at sufficiently high temperature T_0 (1000 K in the present work). The weight factor for this simulation is given by $w_{\mathrm{mu}}^{(0)}(E) = e^{-\beta_0 E}$ with $\beta_0 = 1/RT_0$.
2. Store the energy distribution obtained with the weight $w_{\mathrm{mu}}^{(E)}$ ($i \geq 0$) as a histogram $H^{i+1}(E)$ with bin width δE (1 kcal/mol in the present work). Defining $n^{(i+1)}(E)$ as in Eq. (3),

$$n^{(i+1)}(E) = H^{(i+1)}(E)/w_{\mathrm{mu}}^{(i)}(E) \tag{3}$$

Journal of Molecular Graphics and Modelling 16, 226–238, 1998

655 Avenue of the Americas, New York, NY 10010

1093-3263/99/$–see front matter
PII S1093-3263(99)00013-3

we set the multicanonical weight factor for the next iteration by Eq. (4):

$$w_{\mathrm{mu}}^{(i+1)}(E) = 1/n^{(i+1)}(E) \tag{4}$$

3. Perform a new short simulation with the obtained multicanonical weight factor $w_{\mathrm{mu}}^{(i+1)}(E)$.
4. Iterate the last two steps from $i = 0$ until the resulting distribution $H(E)$ becomes reasonably flat in the chosen energy range, and the final weight $w_{\mathrm{mu}}^{(i)}$ that gave the flat distribution is the optimal weight factor.

Once the weight factor $w_{\mathrm{mu}}(E)$ is determined, one performs a long multicanonical simulation with this weight factor. Monitoring the energy in this simulation, one would see that a random walk between high-energy states and the ground-state configurations is realized. In this way information is collected over the whole energy range. Finally, from this simulation one not only can locate the energy global minimum but also can calculate the ensemble average of any physical quantity A at any temperature T ($= 1/R\beta$) for a wide range of temperatures by the reweighting techniques[14] shown in Eq. (5):

$$\langle A\rangle_T = \frac{\int dE\; A(E)\, P_{\mathrm{mu}}(E) w_{\mathrm{mu}}^{-1}(E) e^{-\beta E}}{\int dE\; P_{\mathrm{mu}}(E) w_{\mathrm{mu}}^{-1}(E) e^{-\beta E}} \tag{5}$$

where $P_{\mathrm{mu}}(E)$ is the distribution of energy obtained by the final simulation.

Potential energy function and computational details

In this work we study the pentapeptide Met-enkephalin, whose amino acid sequence is Tyr–Gly–Gly–Phe–Met. For our simulations, the backbone was terminated by a neutral NH_2– group at the N terminus and a neutral –COOH group at the C terminus, as in previous works on Met-enkephalin. The potential energy function that we use in the present work is given by the conformational energy E_{P} of the protein molecule itself. We have neglected the effects of solvent in order to compare the results with those of previous works in the gas phase. The conformational energy function E_{P} (in kcal/mol) is given by the sum of the electrostatic term E_{C}, 12-6 Lennard–Jones term E_{LJ}, and hydrogen-bond term E_{HB} for all pairs of atoms in the molecule together with the torsion term E_{TOR} for all torsional angles [Eq. (6)]:

$$\begin{aligned} E_{\mathrm{P}} &= E_{\mathrm{C}} + E_{\mathrm{LJ}} + E_{\mathrm{HB}} + E_{\mathrm{TOR}} \\ E_{\mathrm{C}} &= \sum_{(i,j)} \frac{322 q_i q_j}{\epsilon r_{ij}} \\ E_{\mathrm{LJ}} &= \sum_{(i,j)} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) \\ E_{\mathrm{HB}} &= \sum_{(i,j)} \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) \\ E_{\mathrm{TOR}} &= \sum_{i} U_i [1 \pm \cos(n_i \chi^i)] \end{aligned} \tag{6}$$

Here, r_{ij} is the distance (in Å) between atoms i and j, ϵ is the dielectric constant, and χ^i is the torsional angle for the chemical bond i. Each atom is expressed by a point at its center of mass, and the partial charge q_i (in units of electronic charges) is assumed to be concentrated at that point. The factor 332 in E_{C} is a constant to express energy in units of kilocalories per mole. These parameters in the energy function as well as the molecular geometry are based on ECEPP/2.[15] The computer code KONF90[16,17] was used. We remark that both the parameters and their implementation of KONF90 differ slightly from the original version of ECEPP/2 (for example, ϕ_1 of ECEPP/2 is equal to $\phi_1 - 180°$ of KONF90, and energies are also different by small irrelevant constant terms). The dielectric constant ϵ was set equal to 2 as in the previous works. The peptide bond dihedral angles ω were fixed at the value 180° for simplicity. With this choice of parameters, our results can be directly compared with Refs. 5–7 and 12.

The remaining dihedral angles constitute the variables to be updated in the Monte Carlo simulations: ϕ_i and ψ_i in the main chain ($i = 1, 2, \ldots,$ and 5) and χ_i^j in the side chains (there are three χ angles in Tyr, two χ angles in Phe, and four χ angles in Met). For Met-enkephalin the number of degrees of freedom is 19. One Monte Carlo (MC) sweep consists of updating all these 19 angles once with Metropolis evaluation[18] for each update.

Unlike in regular multicanonical simulations, we optimized our weight factor $w_{\mathrm{mu}}(E)$ so that the resulting energy distribution is not completely flat, but biased toward low energies. This is because we are mainly interested in low-energy conformations, where we want to have more statistics, although occasional excursions into the high-energy region are essential for crossing energy barriers. However, the determination of this optimal weight factor turned out to be tedious. Unlike in the earlier work, where we needed only about 40 000 MC sweeps to calculate the weight factor (using the iterative procedure described in Ref. 7), we needed here 13 iterations of multicanonical simulations with 10^5 MC sweeps. We remark that such a choice of parameters does not affect the thermodynamic averages (see the preceding reweighting equation).

After the optimal multicanonical weight factor $w_{\mathrm{mu}}(E)$ was determined, we then made one production run with 10^6 MC sweeps. Analysis of the time series of energy showed that the present choice of multicanonical weight factor indeed realizes a random walk in potential energy space, which keeps the simulation from becoming trapped in a local minimum. The random walk visited the lowest-energy region ($E \approx -12$ kcal/mol) several times in the 10^6 MC sweeps. The visits are separated by excursions into high-energy regions, which ensures decorrelation of the configuration.

The large changes in energy imply large conformational changes that occur in the course of the simulation. Because large parts of the configuration space are sampled, the use of the reweighting techniques is justified to calculate thermodynamic quantities over a wide range of temperatures [from Eq. (5)]. As an example, we calculated the average energy and specific heat as a function of temperature, and the results were in good agreement with those in the previous multicanonical simulations.[6,7,9]

RESULTS AND DISCUSSION

Previous results

In this section, we summarize the results of our previous papers.[5,9] In Ref. 5 Monte Carlo-simulated annealing[19] of Met-enkephalin in the gas phase was performed and the 40 lowest energy conformations obtained were analyzed by the root-mean-square distances. These conformations were classified into four characteristic groups of similar structures (they were referred to as groups A, B, C, and D). These conformations correspond to the local minimum energy conformations of Met-enkephalin.

On the other hand, in Ref. 9 the pattern of hydrogen bonding was used to characterize groups of low-energy structures of Met-enkephalin. However, the investigations focused on conformations with energies only a little above ground state and therefore found only structures corresponding to groups A (the lowest energy state) and B (the second lowest energy state) in Ref. 5. The conformations in group A have two hydrogen bonds between Gly-2 and Met-5 and form a type II′ β turn involving Gly–Gly–Phe–Met. They correspond to the most stable structure of Met-enkephalin in gas phase. The conformations in group B have hydrogen bonds between Tyr-1 and Phe-4 and form a type II β turn involving Tyr–Gly–Gly–Phe.

In this article we pursue a more detailed classification of low-energy conformations of Met-enkephalin according to their pattern of hydrogen bonding.

Distributions of hydrogen bonds

To study the characteristics of the obtained conformations in detail, we first examine the distributions of intrachain hydrogen bonds in the backbone. In this article, we use the following abbreviations for the atoms in the backbone. The amide nitrogen and carbonyl oxygen of the *i*th residue are referred to as N*i* and O*i*, respectively. For example, O2 and N5 stand for the carbonyl oxygen of Gly-2 and the amide nitrogen of Met-5, respectively. Moreover, the intrachain hydrogen bond between N*i*H and O*j* is denoted by N*i*–O*j* for short. Let d_{HO} be the distance between the amide hydrogen and the carbonyl oxygen, and let θ_{NHO} be the angle spanned by the amide nitrogen, amide hydrogen, and carbonyl oxygen (i.e., the angle between two vectors $\overrightarrow{NH}$ and $\overrightarrow{HO}$). We adopt the following criterion for hydrogen bond formations. We consider that a hydrogen bond between the amide nitrogen (donor) and carbonyl oxygen (acceptor) is formed if [Eq. (7)]

$$\begin{cases} d_L \le d_{HO} \le d_H \\ \theta_{NHO} \le \theta_C \end{cases} \tag{7}$$

where we take [Eq. (8)]

$$d_L = 1.5\ \text{Å}, \quad d_H = 2.5\ \text{Å}, \quad \text{and} \quad \theta_C = 60° \tag{8}$$

Since adjacent residues do not form hydrogen bonds, there are 12 possible hydrogen bond patterns (namely, N1–O3, O1–N3, N1–O4, O1–N4, N1–O5, O1–N5, N2–O4, O2–N4, N2–O5, O2–N5, N3–O5, and O3–N5). Using the reweighting techniques of Eq. (5), we have calculated the probability distributions of these hydrogen bond formations as a function of temperature. The results are shown in Figure 1.

It is found in Figure 1 that the distributions of the hydrogen bonds O2–N5 and N2–O5 are high at low temperatures and monotonically decreasing as the temperature increases. More than 90% of the conformations have these two hydrogen bonds below 150 K. As the temperature is raised, four more hydrogen bonds (N1–O4, O1–N4, O1–N3, and O2–N4) appear. The temperatures of the peak of the distributions are about 300, 300, 350, and 350 K, for N1–O4, O1–N4, O1–N3, and O2–N4, respectively. As the temperature is further raised, other hydrogen bonds will also appear, but their contributions are rather small.

Classification of conformations

As was done in earlier work,[1,4,9] the patterns of intrachain hydrogen bonds can naturally be used for classification of

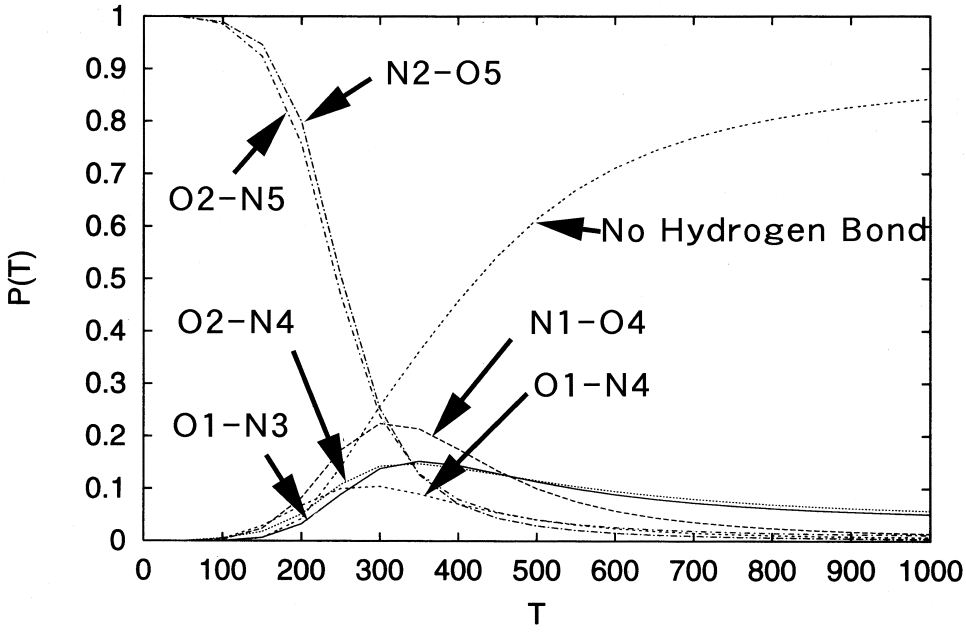

Figure 1. The distributions of hydrogen bonds as a function of temperature.

conformations in a small peptide such as Met-enkephalin. In particular, we try to classify the structure by the patterns of hydrogen bonds that connect a pair of residues. For this purpose we introduce the following notations: we say that the hydrogen bond R_i–R_j between a pair of residues R_i and R_j ($j > i + 1$) is formed when at least one of the hydrogen bonds Ni–Oj and Oi–Nj is formed. For instance, R_1–R_3 can mean that N1–O3 or O1–N3 or both N1–O3 and O1–N3 are formed. There are six such hydrogen bonds (namely, R_1–R_3, R_1–R_4, R_1–R_5, R_2–R_4, R_2–R_5, and R_3–R_5). We can therefore classify the obtained conformations into six groups of similar structures, which we respectively refer to as C13, C14, C15, C24, C25, and C35. Schematic drawings of these six groups are shown in Figure 2.

A conformation can, in principle, have more than one hydrogen bond pattern, and then we have an ambiguity as to which group the conformation belongs to. For example, a conformation can have two hydrogen bonds, R_2–R_4 and R_2–R_5, and the conformation can belong to either group C24 or group C25. In these cases we use the most external hydrogen bond for the classification because external hydrogen bonds will constrain the structure more than internal ones. For instance, if a conformation has two types of hydrogen bonds, R_1–R_3 and R_1–R_5, we consider that the conformation belongs to group C15, because the hydrogen bond R_1–R_3 does not restrict the structure of the C-terminal residue Met-5. The exception is when a conformation has the hydrogen bonds R_1–R_4 and R_1–R_5. By examining in detail the conformations obtained by the present simulation with these hydrogen bonds, we found that their backbone structures have more similarity to group C14 than to group C15. We thus consider that the conformation belongs to group C14 instead of group C15.

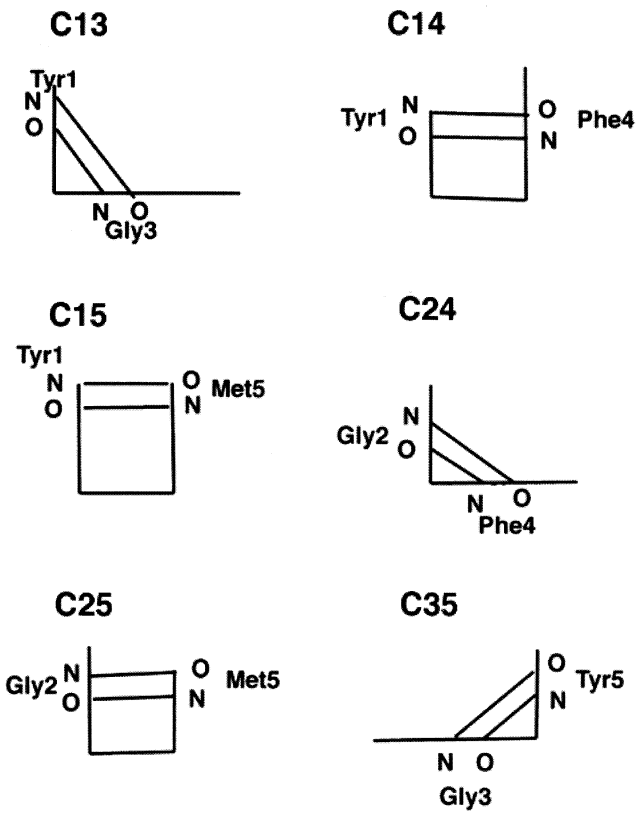

Figure 2. Schematic drawings of six groups of similar structures of Met-enkephalin.

As discussed in the preceding section, the conformations were classified into four groups (A, B, C, and D) by using the root-mean-square distances in the previous work.[5] In the present criterion of classification, the groups A, B, C, and D actually correspond to the groups C25, C14, C15, and C24, respectively. Similarly, groups C25 and C14 correspond to structures 2b and 1a, respectively, in Ref. 9.

After the definition of groups of similar structures is given, we now study the distributions of conformations in these groups as a function of temperature T. The results are shown in Figure 3, which shows group C25 is dominant at low temperatures. Conformations of group C14 start to appear from $T \approx 100$ K. At $T \approx 300$ K, the distributions of these two groups, C25 and C14, balance (≈25% each) and constitute the main groups. Above $T \approx 300$ K, the contributions of other groups become nonnegligible (those of group C24 and group C13 are about 10 and 8%, respectively, at $T = 400$ K). Note that the distribution of conformations that do not belong to any of the six groups monotonically increases as the temperature is raised. This is because random-coil conformations without any intrachain hydrogen bonds are favored at high temperatures. These results are in agreement with those of Refs. 7 and 9.

Details of the six classified structures

After the distributions of the six groups are established, we now study the characteristics of each group in detail. In Color Plate 1 we show the lowest energy conformations in each group. (The figures were created with RasMol.[20]) Conformations were obtained by minimizing (by the Newton–Raphson method with respect to the 19 dihedral angles) the lowest energy conformations found during the multicanonical simulation. We remark that the minimization did not alter the structure much at all, but energy values were lowered from −12.2, −10.1, −8.3, −6.0, −7.0, and −1.7 kcal/mol to −12.2, −11.1, −9.8, −9.1, −8.8, and −5.0 kcal/mol for groups C25, C14, C24, C13, C15, and C35, respectively. The four structures (C25, C14, C24, and C15) in Color Plate 1 are essentially identical to the lowest energy conformations of the four groups found previously in Ref. 5, including the side chain structures, although the previous classification used root-mean-square distances instead of hydrogen bond patterns.

As is shown in Color Plate 1a, the conformation of group C25 has two hydrogen bonds, N2–N5 and O2–N5, and forms a type II′ β turn. The energy of the conformation is −12.2 kcal/mol, and this conformation corresponds to the global minimum energy state of Met-enkephalin in gas phase. The conformation of group C14 (Color Plate 1b) has two hydrogen bonds, N1–O4 and O1–N4, and forms a type II β turn. The energy is −11.1 kcal/mol, and this conformation corresponds to the second lowest energy state. The groups C25 and C14 are the dominant groups in the energy surface of Met-enkephalin as shown above (see Figure 3).

The conformation of group C24 (Color Plate 1c) has only one hydrogen bond, O2–N4 (because the distance between N2 and O4 is larger than 2.5 Å), and forms a γ turn. The energy is −9.9 kcal/mol, and this conformation corresponds to the third lowest energy state. The structures of group C24 can be ob-

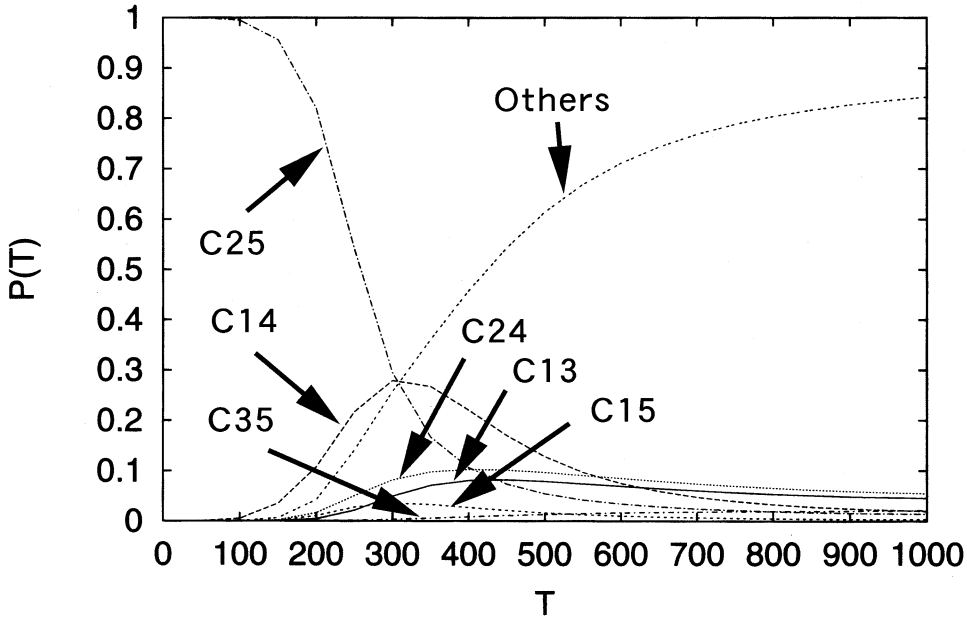

Figure 3. The distributions of each group of similar structures as a function of temperature.

tained from that of group C25 by small rearrangements of the intrachain hydrogen bonds. In group C25 of Color Plate 1a, O2 and N4 are already close to each other. By cutting the two hydrogen bonds, N2–O5 and O2–N5, we can deform the conformation into that of group C24. Note that the OH of the Tyr-1 side chain is hydrogen bonded to the carbonyl oxygen of Gly-3 in both groups C25 and C24.

The conformation of group C13 (Color Plate 1d) has only one hydrogen bond, O1–N3, and also forms a γ turn. The energy is -9.1 kcal/mol, and this conformation corresponds to the fourth lowest energy state. The structure of group C13 can be obtained from that of group C14 by small rearrangements of the intrachain hydrogen bonds. In group C14 of Color Plate 1b, O1 and N3 are already close to each other (see Color Plate 1b also). By cutting the two hydrogen bonds, N1–O4 and O1–N4, we can deform the conformation into that of group C13.

The conformation of group C15 (Color Plate 1e) has only one hydrogen bond, O1–N5. The energy is -8.8 kcal/mol and this conformation corresponds to the fifth lowest energy state. As can be seen in Figure 3, the probability distribution for this group is rather low compared with the four previous groups (C25, C14, C24, and C13).

The conformation of group C35 (Color Plate 1f) has only one hydrogen bond, O3–N5. The backbone of this conformation is extended and the energy is rather high (-5.0 kcal/mol). As can been seen in Figure 3, the probability distribution for this group is also low compared with the first four groups (C25, C14, C24, and C13).

Finally, in Table 1 we list the root-mean-square distances of the atomic coordinates between pairs of the six conformations in Color Plate 1. Only the atoms in the backbone are taken into account. The entries are all more than or equal to 3.0 Å, and this implies that these structures are indeed quite different.

Distributions of hydrogen bonds for each group of similar structures

In Figure 1 we saw the distributions of all possible intrachain hydrogen bonds as a function of temperature. We now examine the distributions of possible hydrogen bonds as a function of temperature for each group separately in order to study how the conformations in each group are disordered as the temperature is raised. The results are shown in Figure 4. We show results only for groups C25 and C14, because these groups are dominant. In Eq. (9) we use weaker conditions for hydrogen bond formations than in Eq. (7) in order to judge which group each conformation belongs to:

$$d_L = 1.5 \text{ Å}, \quad d_H = 3.0 \text{ Å}, \quad \text{and} \quad \theta_C = 90^\circ \qquad (9)$$

because we want to include conformations that do not have hydrogen bonds in a strict sense but yet have similar structures. Namely, Figure 4 shows the probability distributions of hydrogen bond formations [with respect to Eq. (8)] in conformations that belong to groups C25 and C14 [classified with respect to Eq. (9)].

For group C25 (Figure 4a), the distributions of the two characteristic hydrogen bonds, N2–O5 and O2–N5, are equally high at low temperatures. This implies that the lowest energy conformation in Color Plate 1a is quite stable. We remark that at very high temperatures hydrogen bond O2–N5 is stronger

Table 1. Root-mean square distances[a] of the coordinates of the backbone atoms among the lowest energy conformations[b] in groups C13, C14, C15, C24, C25, and C35

	C13	C14	C15	C24	C25	C35
C13	0.0	3.3	3.2	3.1	5.2	3.1
C14	3.3	0.0	4.3	3.0	4.1	4.0
C15	3.2	4.3	0.0	3.4	4.7	3.0
C24	3.1	3.0	3.4	0.0	3.9	3.1
C25	5.2	4.1	4.7	3.9	0.0	5.1
C35	3.1	4.0	3.0	3.1	5.1	0.0

[a]In angstroms.
[b]These six conformations are shown in Color Plate 1.

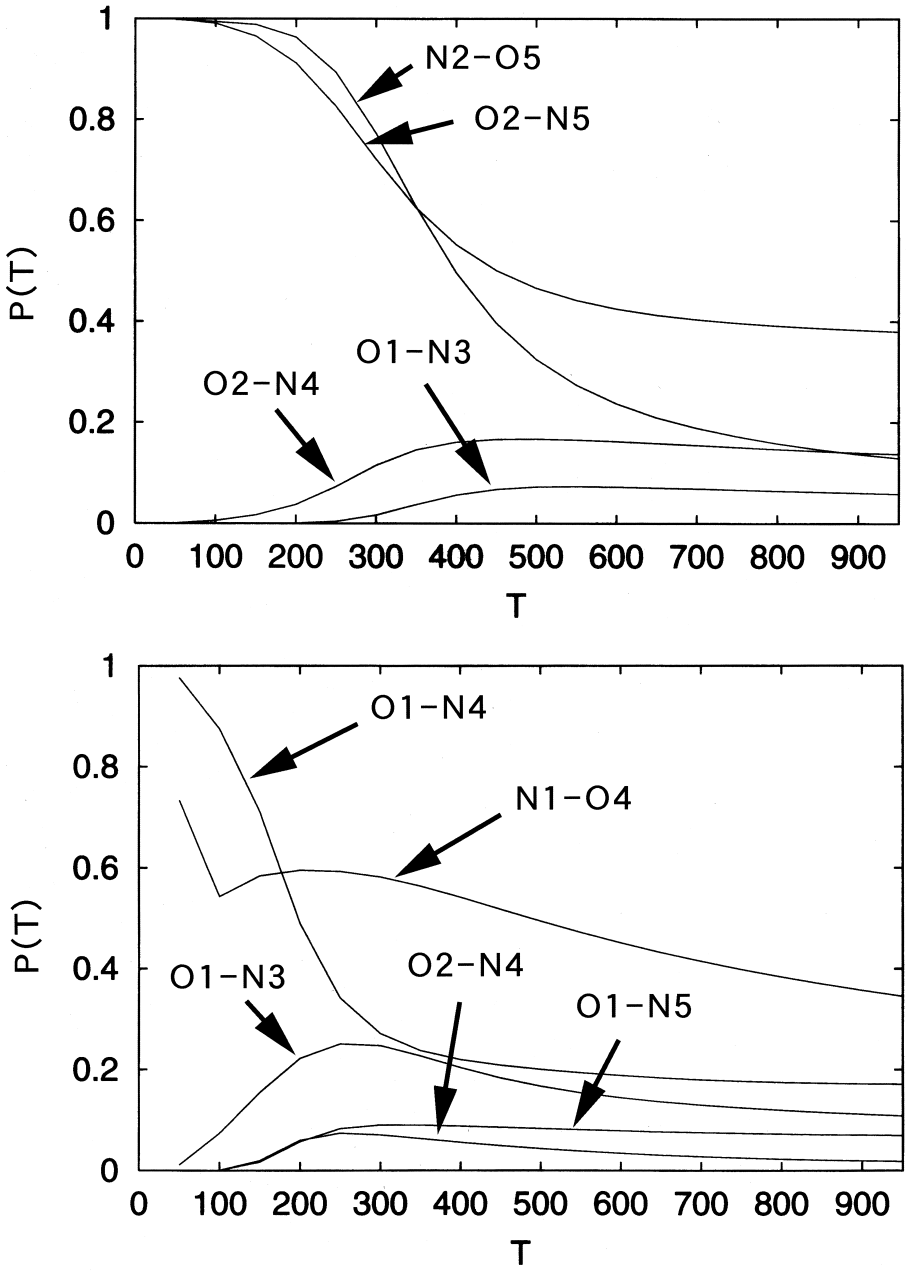

Figure 4. The distributions of 12 possible hydrogen bond patterns in groups C25 and C14 as a function of temperature.

than N2–O5, suggesting that in group C25 the external hydrogen bond (N2–O5) is easier to break than the internal one (O2–N5). The third hydrogen bond, O2–N4, also appears at higher temperatures when the conformation is slightly deformed by thermal fluctuations, opening a possibility of transition of structures from group C25 into group C24 as discussed above.

For group C14 (Fig. 4b), the distributions of the two characteristic hydrogen bonds, N1–O4 and O1–N4, are high but not equal at low temperatures. At low temperatures about 80% of the conformations have the two hydrogen bonds and the remaining 20% have only one hydrogen bond, O1–N4. We remark that at higher temperatures hydrogen bond N1–O4 is stronger than O1–N4, implying that in group C14 the internal hydrogen bond (O1–N4) is easier to break than the external one (N1–O4). The third hydrogen bond, O1–N3, also appears at high temperatures, suggesting that a transition of structures from group C14 to group C13 is possible when the conformation is slightly deformed by thermal fluctuations as discussed above.

For groups C24 and C13 (data not shown), conformations have only one characteristic hydrogen bond (respectively, N2 – O4 and O1–N3) at low temperatures and the other hydrogen bond (respectively, O2–N4 and N1–O3) does not exist, because they both form a γ turn.

Distributions of backbone dihedral angles

Our analyses so far imply that there are two stable structures (groups C25 and C14) at low temperatures. They form, respectively, type II′ β turns and type II β turns. There are two more stable structures (groups C24 and C13), which form γ turns.

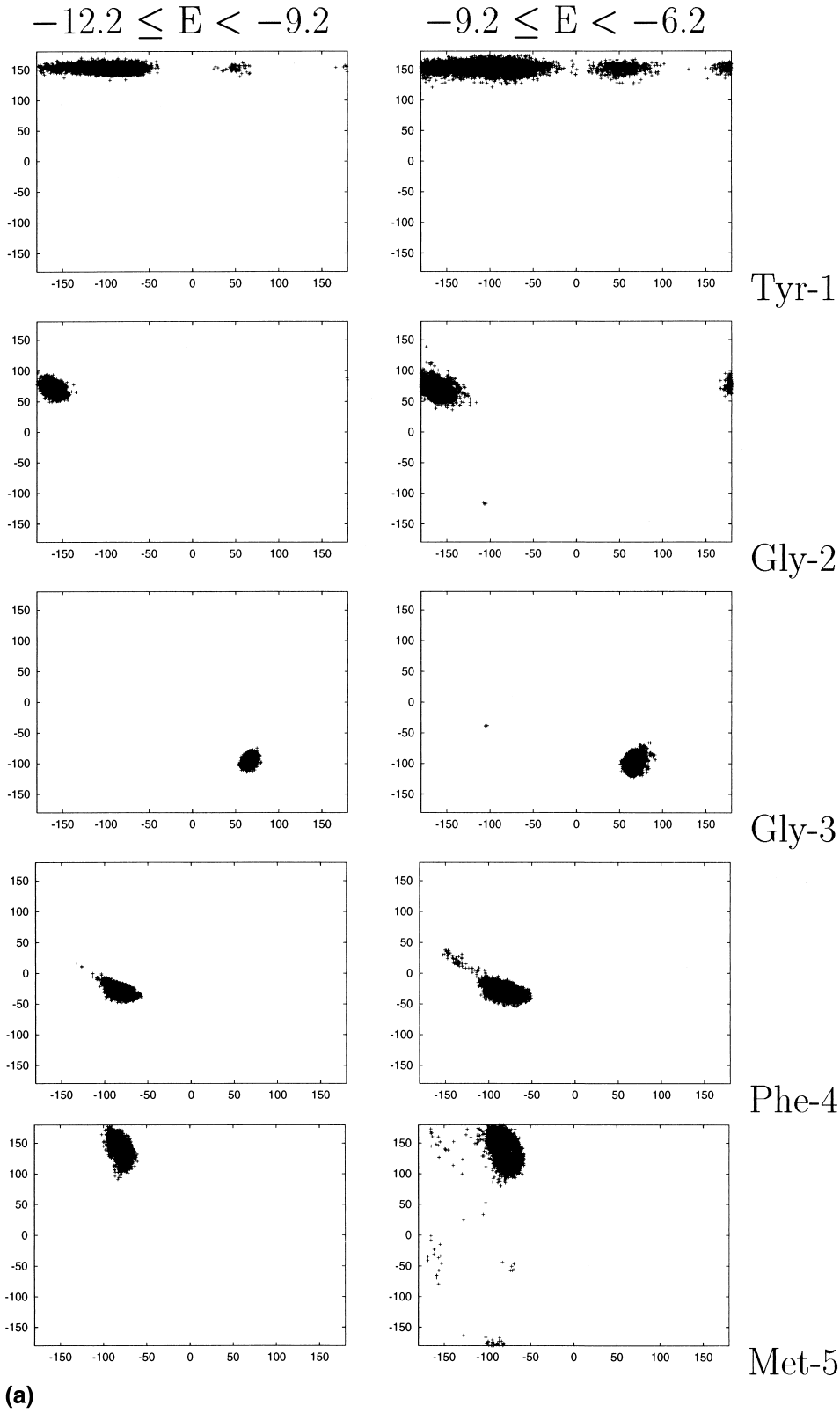

Figure 5. Ramachandran plots for each residue of conformations obtained in the course of the multicanonical simulation. We show the results for C25 (a), C14 (b), C24 (c), and C13 (d). Two sets of Ramachandran plots are made for each residue. The first set is taken from conformations with energy E in the range $E_{min} \leq E < E_{min} + 3$ (kcal/mol) (left-hand side) and the other set is from the range $E_{min} + 3 \leq E < E_{min} + 6$ (kcal/mol) (right-hand side), where E_{min} stands for the lowest energy in each group obtained during the simulation (namely, $E_{\min} = -12.2, -10.1, -8.3$, and -6.0 kcal/mol for groups C25, C14, C24, and C13, respectively). The abscissa stands for the dihedral angle ϕ and the ordinate represents the dihedral angle ψ.

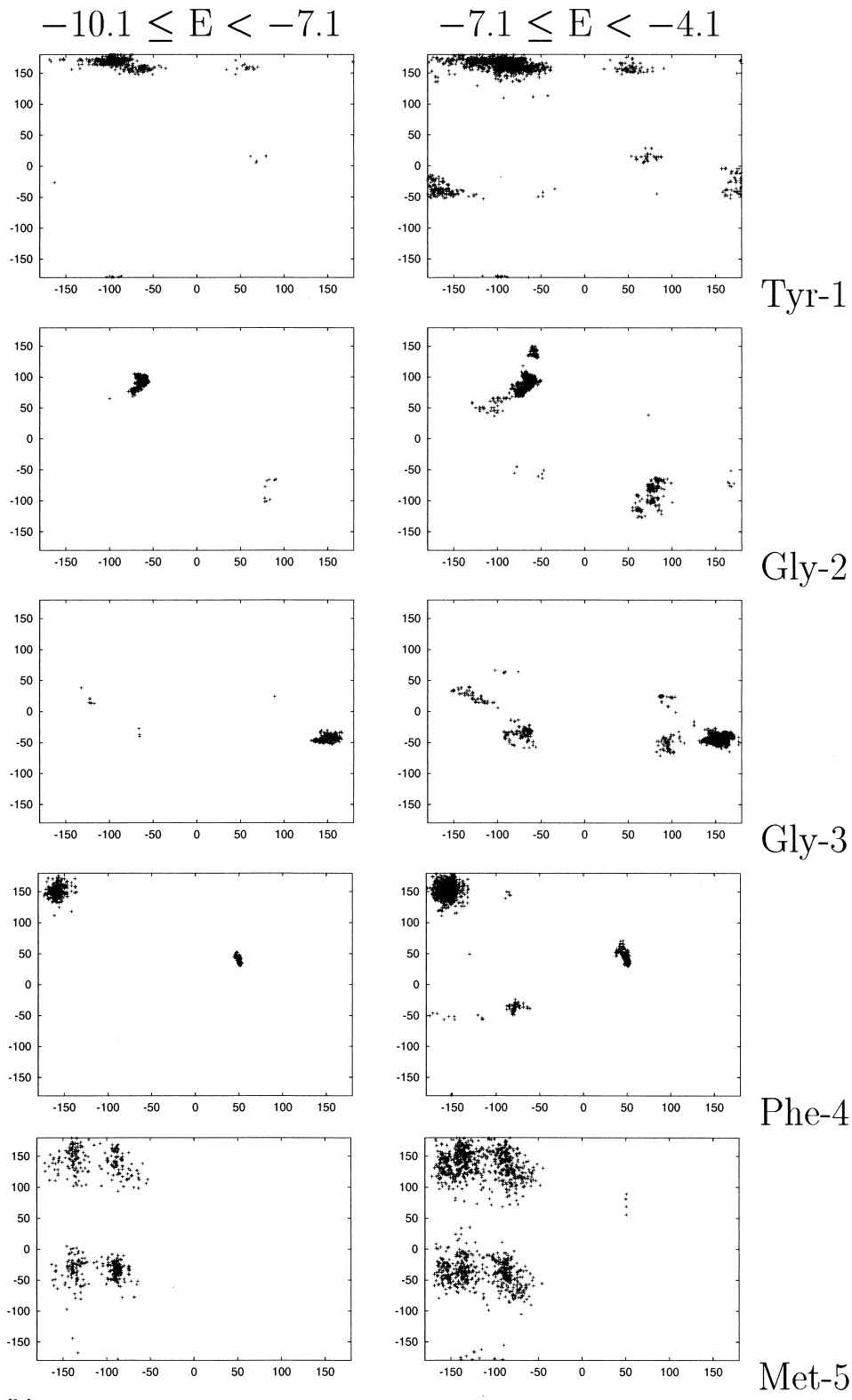

Figure 5 (continued).

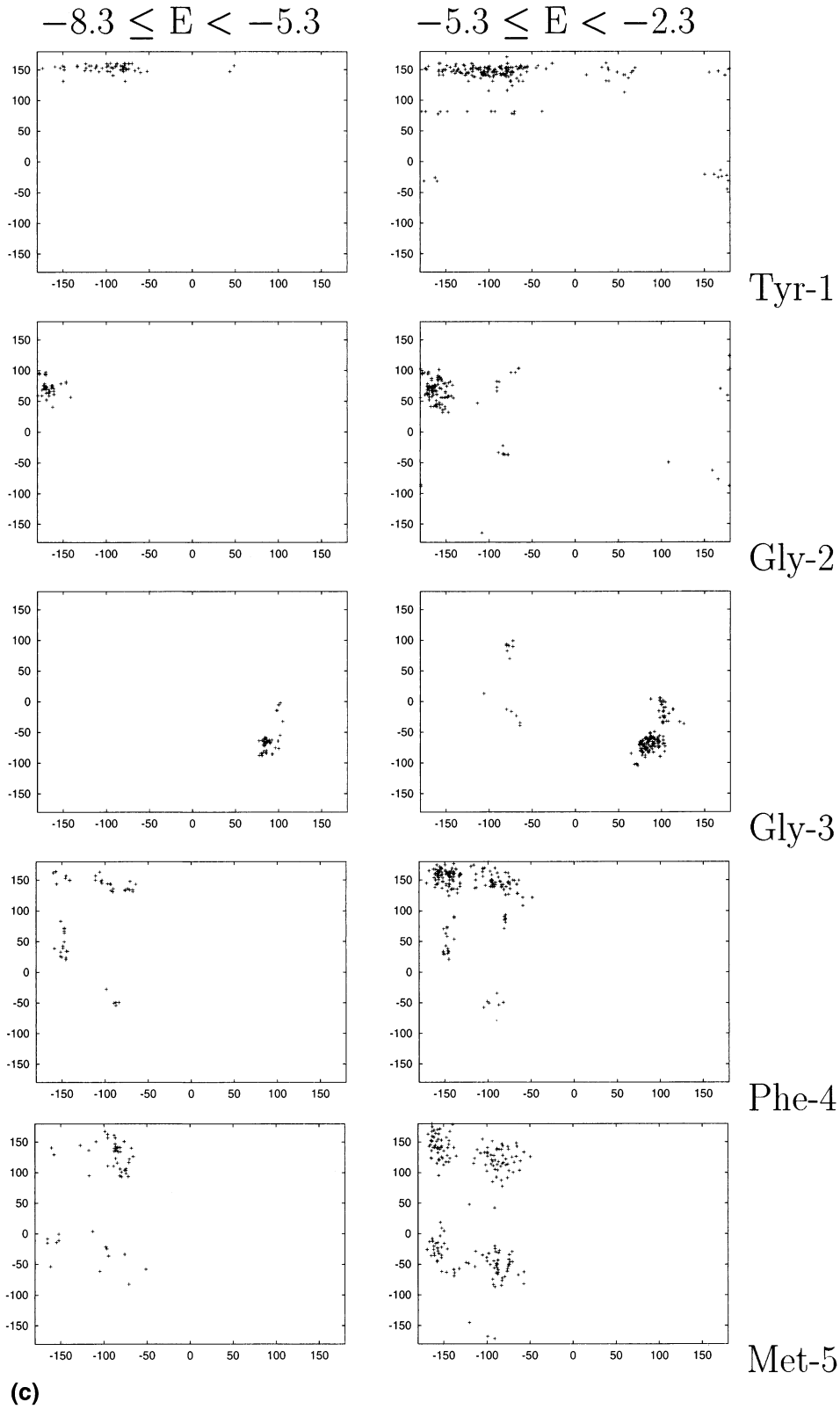

Figure 5 (continued).

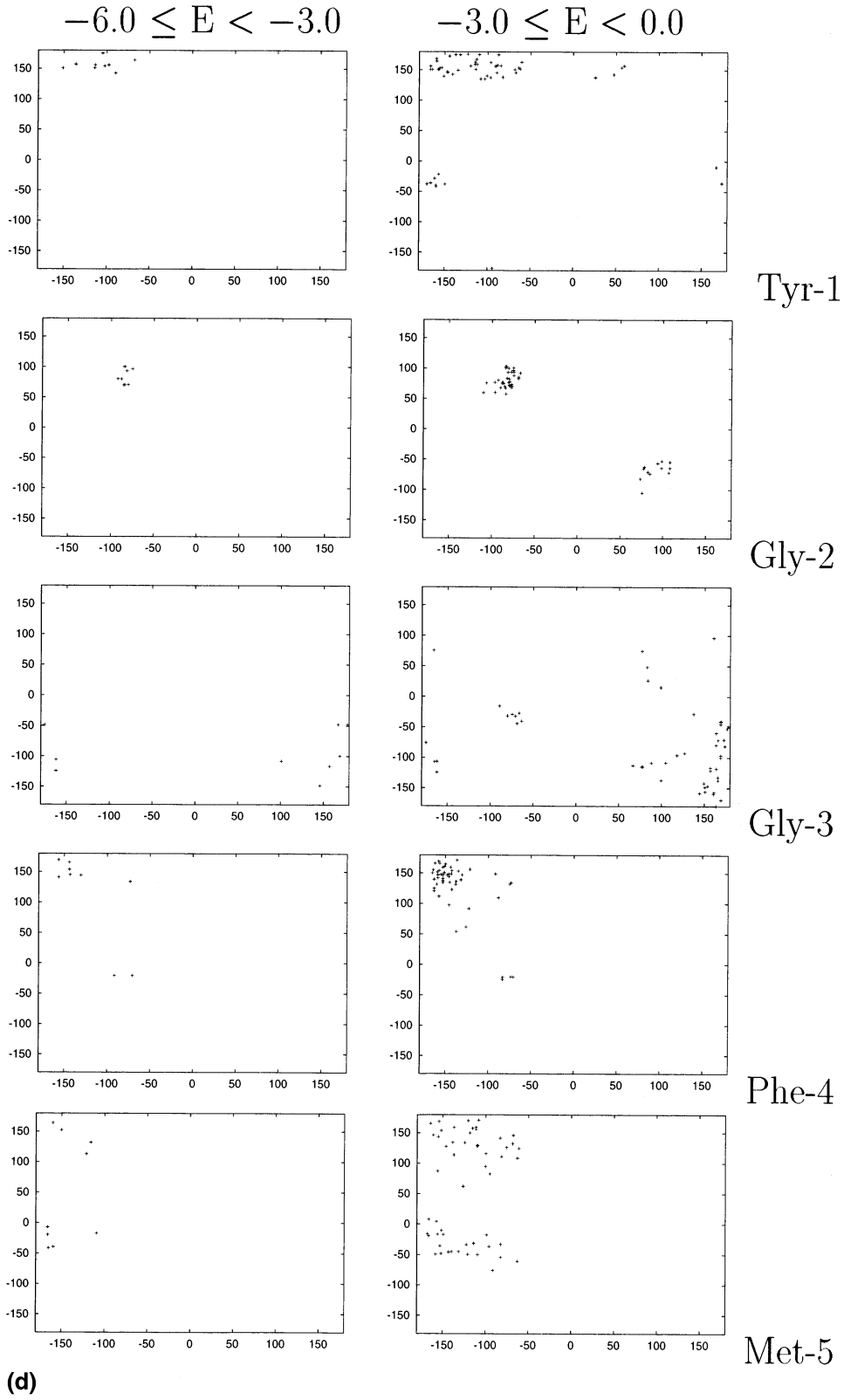

Figure 5 (continued).

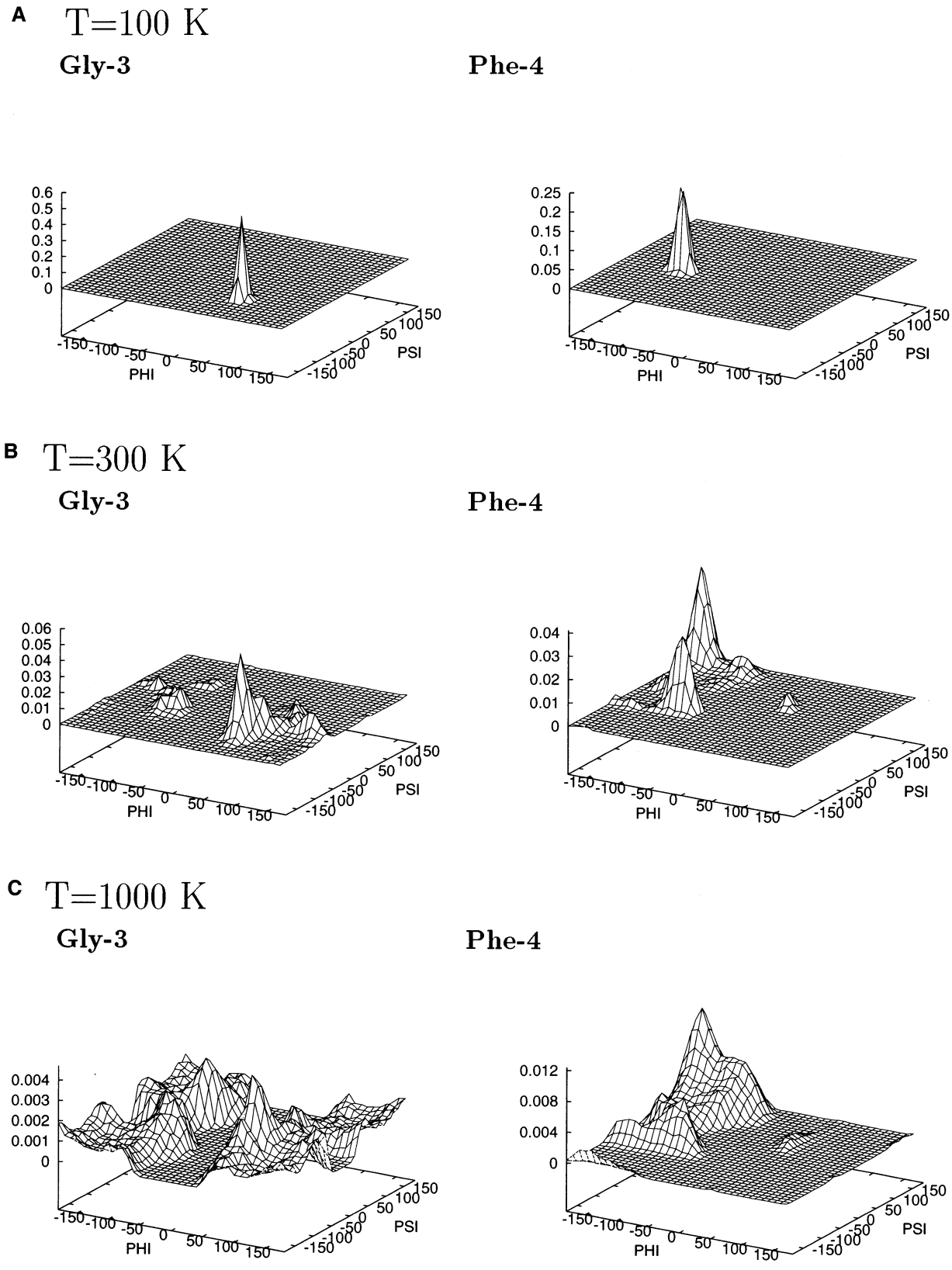

Figure 6. Distributions of the backbone dihedral angles of Met-enkephalin as a function of temperature. The results for Gly-3 and Phe-4 at 100 K (a), 300 K (b), and 1 000 K (c) are shown.

The remaining two groups (C15 and C35) are less stable. These results imply that there are at least four well-defined valleys in the energy landscape of Met-enkephalin in the gas phase. We now study the distributions of backbone dihedral angles. For each group we make a separate Ramachandran plot of conformations obtained in the course of the multicanonical simulation. Here, we again use the weaker definition [Eq. (9)] for hydrogen bonds in Eq. (7) to judge which group each conformation belongs to. Two sets of Ramachandran plots are made for each residue. The first set is taken from conformations with energy E in the range $E_{\min} \leq E < E_{\min} + 3$ (kcal/mol) and the other set is from the range $E_{\min} + 3 \leq E < E_{\min} + 6$ (kcal/mol), where $E_{\min}$ stands for the lowest energy in each group obtained during the multicanonical simulation (namely, $E_{\min} = -12.2$, -10.1, -8.3, -6.0, -7.0, and -1.7 kcal/mol for groups C25, C14, C24, C13, C15, and C35, respectively). We consider that the group in question is a well-defined valley in the energy surface if the second set of Ramachandran plots is distributed more widely than the first set and contains the first set as a subset.

The results are shown in Figure 5. We show only data for groups C25, C14, C24, and C13. For groups C25 (Figure 5a) and C14 (Figure 5b), the numbers of samples are large, and the preceding criterion is clearly met. These groups are thus well-defined valleys in the energy surface. This can also be checked directly by minimizing the configurations. Most of them converge into the same local minimum state. Note that for C14 (Figure 5b), there exist two local minima in the distributions of Gly-2 and Gly-3 dihedral angles (see the second column). This actually implies that besides the type II β turn, there exists a type II′ β turn at a slightly higher energy. Note also that for group C25 the distributions are well localized except for ϕ of Tyr-1 (ϕ_1). This implies that the structure is stable. For ϕ_1 the angle is free to vary because it is located at the end of the peptide and no hydrogen bond stabilizes the structure there. On the other hand, for group C14 the C-terminal dihedral angles (ϕ_5, ψ_5) are free to move because the hydrogen bond R_1–R_4 does not constrain the Met-5 structure.

For group C24 (Fig. 5c) and group C13 (Fig. 5d), the numbers of sample dots are not as large as those for group C25 and group C14, but the criterion (i.e., a valley in the energy surface) is again satisfied. These two groups can thus be said to correspond to local minimum states. We remark that the distributions for group C24 are rather similar to those for group C25, except for those of Phe-4 and Met-5. It is therefore easy for these two structures to interchange with each other as discussed above. Note also that the structure of Met-5 is flexible for C24 because no hydrogen bond stabilizes it. The distributions for C13 are likewise similar to those for C14, except for those of Gly-3.

For C15 and C35 (data not shown), the numbers of samples are rather small and we cannot conclusively claim that these groups correspond to valleys in the energy surface. If they do, they are shallower than those of the previous four groups.

Using the reweighting techniques [Eq. (5)], we can calculate the distributions of the backbone dihedral angles of Met-enkephalin as a function of temperature. The results at 100 K, 300 K, and 1 000 K are shown in Fig. 6 (data are for Gly-3 and Phe-4). At $T = 100$ K (Fig. 6a) the distributions are well-localized, and at this temperature one conformation (C25) is dominant. At $T = 300$ K (Fig. 6a) the distributions are still localized, but we see a trace of the second conformation (group C14) appearing. Note that these results are in complete agreement with our analyses of the energy landscape of Met-enkephalin in Refs. 21 and 22, where we used the overlap functions to characterize the energy landscape. There, the backbone dihedral angles themselves were taken as the order parameter and again we observed the transition between a unique ground state and an ensemble of other well-defined structures. Finally, at $T = 1\,000$ K (Fig. 6c) the distributions are widely spread, implying large thermal fluctuations. These results correspond to the observed transition between an ensemble of well-defined compact conformations and extended coil structures observed in Refs. 21 and 22).

CONCLUSIONS

In this article, we have presented the results of a multicanonical simulation of Met-enkephalin in gas phase. We obtained the distributions of hydrogen bonds in the backbone. We used the patterns of hydrogen bond formations to classify the conformations into groups of similar structures. It was found that there are at least four well-defined conformational categories, and these correspond to local minimum energy states. The global minimum energy state forms a type II′ β turn, the second lowest energy state forms a type II β turn, and the third and fourth states form γ turns. The multicanonical algorithm allowed us to calculate various thermodynamic quantities as a function of temperature from a single simulation run. The results, such as distributions of conformations as a function of temperature, give essential information about the free energy landscape of the peptide.

ACKNOWLEDGMENTS

The simulations were performed on the computers at the Computer Center of the Institute for Molecular Science. This work was supported, in part, by grants from Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists, from the Research for the Future Program of the Japan Society for the Promotion of Science (JSPS-RFTF98P01101), from the Japanese Ministry of Education, Science, Sports, and Culture, and from a Research Excellence Fund (E27448) of the State of Michigan.

REFERENCES

1 Paine, G.H., and Scheraga, H.A. Prediction of the native conformation of a polypeptide by a statistical-mechanical procedure. III. Probable and average conformations of enkephalin. *Biopolymers* 1987, **26,** 1125–1162

2 Li, Z., and Scheraga, H.A. Monte Carlo-minimization approach to the multiple-minima problem in protein folding. *Proc. Natl. Acad. Sci. U.S.A.* 1987, **84,** 6611–6615

3 Kawai, H., Kikuchi, T., and Okamoto, Y. A prediction of tertiary structures of peptide by the Monte Carlo simulated annealing method. *Protein Eng.* 1989, **3,** 85–94

4 Von Freyberg, B., and Braun, W. Efficient search for all low energy conformations of polypeptides by Monte Carlo methods. *J. Comput. Chem.* 1991, **12,** 1065–1076

5 Okamoto, Y., Kikuchi, T., and Kawai, H. Prediction of low-energy structures of Met-enkephalin by Monte

Carlo simulated annealing. *Chem. Lett.* 1992, **1992,** 1275–1278
6 Hansmann, U.H.E., and Okamoto, Y. Prediction of peptide conformation by multicanonical algorithm: New approach to the multiple-minima problem. *J. Comput. Chem.* 1993, **14,** 1333–1338
7 Hansmann, U.H.E., and Okamoto, Y. Comparative study of multicanonical and simulated annealing algorithms in the protein folding problem. *Physica A* 1994, **212,** 415–437
8 Meirovitch, H., Meirovitch, E., Michel, A.G., and Vásquez, M. A simple and effective procedure for conformational search of macromolecules: Application to Met- and Leu-enkephalin. *J. Phys. Chem.* 1994, **98,** 6241–6243
9 Eisenmenger, F., and Hansmann, U.H.E. Variation of the energy landscape of a small peptide under a change from the ECEPP/2 force field to ECEPP/3. *J. Phys. Chem. B* 1997, **101,** 3304–3310
10 Androulakis, I.P., Maranos, C.D., and Floudas, C.A. Prediction of oligopeptide conformations via deterministic global optimization. *J. Global. Opt.* 1997, **11,** 1–34
11 Carlacci, L. Conformational analysis of [Met^5]-enkephalin: Solvation and ionization considerations. *J. Comput. Aided. Mol. Des.* 1998, **12,** 195–213
12 Mitsutake, A., Irisa, M., Okamoto, Y., and Hirata, F. Classification of low-energy conformations of Met-enkephalin in the gas phase and in the a model solvent based on the extended scaled particle theroy. *Bull. Chem. Soc. Jpn.* 1999, **72,** 1717–1729
13 Berg, B.A., and Neuhaus, T. Multicanonical algorithms for first order phase transitions. *Phys. Lett.* 1991, **B267,** 249–253
14 Ferrenberg, A.M., and Swendsen, R.H. Optimized Monte Carlo data analysis. *Phys. Rev. Lett.* 1988, **61,** 2635–2638
15 Momany, F.A., McGuire, R.F., Burgess, A.W., and Scheraga, H.A. Energy parameters in polypeptides. VII. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions, and intrinsic torsional potentials for the naturally occurring amino acids. *J. Phys. Chem.* 1975, **79,** 2361–2381; Némethy, G., Pottle, M.S., and Scheraga, H.A. Energy parameters in polypeptides. 9. Updating of geometrical parameters, nonbonded interactions, and hydrogen bond interactions for the naturally occurring amino acids. *J. Phys. Chem.* 1983, **87,** 1883–1887; Sippl, M.J., Némethy, G., and Scheraga, H.A. Intermolecular potentials from crystal data. 6. Determination of empirical potentials for O—H · · · O═C hydrogen bonds from packing configurations. *J. Phys. Chem.* 1984, **88,** 6231–6233
16 Kawai, H., Okamoto, Y., Fukugita, M., Nakazawa, T., and Kikuchi, T. Prediction of α-helix folding of isolated C-peptide of ribonuclease A by Monte Carlo simulated annealing. *Chem. Lett.* 1991, **1991,** 213–216
17 Okamoto, Y., Fukugita, M., Nakazawa, T., and Kawai, H. α-helix folding by Monte Carlo simulated annealing in isolated C-peptide of ribonuclease A. *Protein Eng.* 1991, **4,** 639–647
18 Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., and Teller, E.J. Equation of state calculations by fast computing machines. *J. Chem. Phys.* 1953, **21,** 1087–1092
19 Kirkpatrick, S., Gelatt, C.D., Jr., and Vecchi, M.P. Optimization by simulated annealing. *Science* 1983, **220,** 671–680
20 Sayle, R.A., and Milner-White, E.J. RasMol: Biomolecular graphics for all. *Trends Biochem. Sci.* 1995, **20,** 374–376
21 Hansmann, U.H.E., Masuya, M., and Okamoto, Y. Characteristic temperatures of folding of a small peptide. *Proc. Natl. Acad. Sci. U.S.A.* 1997, **94,** 10652–10656
22 Hansmann, U.H.E., Okamoto, Y., and Onuchic, J. The folding funnel landscape for the peptide Met-enkephalin. *Proteins* 1999, **34,** 472–483