α-Helix Propensities of Amino Acids Studied by Multicanonical Algorithm

Yuko Okamoto,* Ulrich H.E. Hansmann,† and Takashi Nakazawa††
Department of Physics, Nara Women's University, Nara 630
†*IPS, ETH Zürich, 8092 Zürich, Switzerland*
††*Department of Chemistry, Nara Women's University, Nara 630*

(Received February 9, 1995)

We report the results of multicanonical simulations on α-helix propensities of homo-oligomers of nonpolar amino acids. Three characteristic amino acids are considered: alanine (helix former), glycine (helix breaker), and valine (helix indifferent). Starting simulations from completely random initial conformations, we have calculated the average % helix, free energy differences of helix-coil transitions, and the Zimm-Bragg s parameters over a wide range of temperatures. The results are shown to be in good agreement with experiments.

Prediction of three-dimensional structure of proteins solely from their amino acid sequence remains one of the unsolved problems in bioscience. The difficulty with the conventional methods lies in the fact that simulations tend to get trapped in one of many energy local minima. Hence, novel algorithms are in urgent demand. The *multicanonical algorithm*[1] is a new Monte Carlo method that is known to alleviate the above difficulty in other thermodynamic systems with frustration such as spin glasses.[2-4] The application of this algorithm to the protein folding problem was recently proposed.[5]

In the multicanonical approach Monte Carlo simulations are performed in a *multicanonical ensemble*.[1] This artificial ensemble is defined by the condition that the probability distribution of the energy shall be constant, which in turn implies that a Monte Carlo simulation in this ensemble realizes a 1d random walk in the energy space. This ensures that the simulation avoids getting trapped in any energy local minimum. The probability weight factor in the multicanonical ensemble is not *a priori* known. It has to be constructed in an iterative way described elsewhere.[2,5] Once the multicanonical weight factor is determined, we make a long production run. From the results of this production run, one can calculate the regular canonical expectation values over a wide range of temperatures through the use of the reweighting techniques.[2,5]

We considered three homo-oligomers of 10 amino acids, $(\mathrm{Ala})_{10}$, $(\mathrm{Val})_{10}$, and $(\mathrm{Gly})_{10}$. The computer code KONF90 [6] was modified to accomodate the multicanonical method. The semi-empirical potential energy function that we used is taken from ECEPP/2.[7] The peptide-bond dihedral angles ω were fixed at the value 180° and the dielectric constant was set to 2. Because one can avoid the complications of electrostatic and hydrogen-bond interactions of side chains with the solvent for nonpolar amino acids, explicit solvent molecules were neglected for simplicity. One Monte Carlo step consists of updating all the torsion angles in the backbone and side chains. For each homo-oligomer, several preliminary runs with 10^4 Monte Carlo steps were made to determine the multicanonical weight factors, and then one production run with 10^5 Monte Carlo steps was made from a completely random initial conformation for calculation of thermodynamic quantites.[5]

The criterion we adopt for α-helix formation is as follows:[8] We consider that a residue is in the α-helix state when the dihedral angles (ϕ, ψ) fall in the range $(-70\pm 20°, -37\pm 20°)$. The length ℓ of a helical segment is then defined by the number of successive residues which are in the α-helix state. The number n of helical residues in a conformation is defined by the sum of ℓ over all helical segments in the conformation.

In Figure 1 we show the average % helix per residue $\frac{<n>}{N}$ ($N = 10$) as a function of temperature. $(\mathrm{Ala})_{10}$ is a strong helix former with % helix varying from ~ 80% at $T = 200$ K to ~ 50% at $T = 400$ K, and $(\mathrm{Gly})_{10}$ is a strong helix breaker with % helix varying from ~ 10% at $T = 200$ K to ~ 7% at $T = 400$ K, while $(\mathrm{Val})_{10}$ comes in between the two with % helix varying from ~ 35% at $T = 200$ K to ~ 17% at $T = 400$ K. This is in accord with the results from the recent experiments with oligopeptide systems (for a review, see Ref. 9).

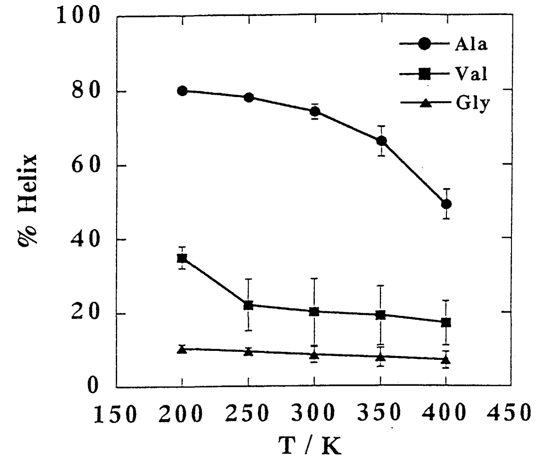

Figure 1. Average % helix per residue $\frac{<n>}{N}$ ($N = 10$) as a function of temperature for the three homo-oligomers.

Table 1. Average number of helical residues $< n >$, average length of a helical segment $< \ell >$, and the Zimm-Bragg s parameter as functions of temperature T (K)

Peptide	T	$< n >$	$< \ell >$	s
$(\mathrm{Ala})_{10}$	250	7.8 (0.1)	7.5 (0.5)	1.6 (0.1)
	300	7.4 (0.2)	6.6 (0.6)	1.5 (0.1)
	350	6.6 (0.4)	5.2 (0.6)	1.3 (0.1)
$(\mathrm{Val})_{10}$	250	2.2 (0.7)	1.6 (0.1)	0.45 (0.03)
	300	2.0 (0.9)	1.5 (0.1)	0.37 (0.09)
	350	1.9 (0.8)	1.4 (0.1)	0.31 (0.08)
$(\mathrm{Gly})_{10}$	250	0.94 (0.10)	1.2 (0.1)	0.16 (0.05)
	300	0.84 (0.21)	1.2 (0.1)	0.13 (0.05)
	350	0.77 (0.26)	1.3 (0.1)	0.11 (0.05)

In Table 1 we list the average number of helical residues $< n >$ and the average length of a helical segment $< \ell >$ as functions of temperature. The numbers in parentheses represent errors. If we neglect the terminal residues, $\ell = 8$ corresponds to the maximum helix length (close to 3 turns of the α-helix) and the conformation can be considered completely helical. As is clear from the table, $(\mathrm{Ala})_{10}$ is almost entirely helical ($< \ell >= 7.5$ and $< n >= 7.8$) at $T = 250$ K and still substantially helical even at $T = 350$ K. On the contrary, $(\mathrm{Val})_{10}$ and $(\mathrm{Gly})_{10}$ are hardly helical at these temperatures.

Table 2. Free energy differences ΔG, enthalpy differences ΔH, and entropy differences $T\Delta S$ (kcal/mol) between helix and non-helix states as functions of temperature T (K)

Peptide	T	ΔG	ΔH	$T\Delta S$
(Ala)$_{10}$	250	−4.4 (1.0)	−10.9 (1.1)	−6.5 (1.5)
	300	−3.1 (0.8)	−10.7 (2.5)	−7.6 (2.6)
	350	−1.9 (0.5)	−10.4 (2.8)	−8.5 (2.9)
(Val)$_{10}$	250	0.4 (0.3)	−2.1 (1.8)	−2.5 (1.8)
	300	0.8 (0.5)	−0.9 (0.5)	−1.7 (0.8)
	350	1.1 (0.5)	−1.8 (1.3)	−2.9 (1.1)
(Gly)$_{10}$	250	2.7 (1.2)	1.3 (2.1)	−1.4 (2.4)
	300	3.1 (1.0)	0.3 (1.7)	−2.8 (2.0)
	350	3.6 (0.9)	−0.3 (2.0)	−3.9 (2.2)

The helix propagation parameter s of the Zimm-Bragg model[10] can be estimated from $< n >$ and $< \ell >$. The results as a function of temperature are listed in Table 1. We have s(Ala) = 1.5 ~ 1.6, s(Val) = 0.37 ~ 0.45, and s(Gly) = 0.13 ~ 0.16 around the experimentally relevant temperature (~ 0° C). These values are in good agreement with the recent experiments,[9] where they give s(Ala) = 1.99 ~ 2.19, s(Val) = 0.20 ~ 0.93, and s(Gly) = 0.02 ~ 0.57.

In Table 2 we present the free energy differences $\Delta G \equiv G_H - G_C$, enthalpy differences ΔH, and entropy differences $T\Delta S$ between helix (H) and non-helix (C) states. Here, a conformation is considered to be in the helix state if it has a segment with helix length $\ell \geq 3$. Note that $\ell = 3$ corresponds to roughly one turn of the α-helix . The free energy differences were calculated from $\Delta G = -RT \ln \frac{N_H}{N_C}$, where N_H and N_C are average numbers of conformations in helix and non-helix states, respectively. The enthalpy differences were obtained from $\Delta H = E_H - E_C$, where E_H and E_C are average potential energies in helix and non-helix states, respectively. Finally, the entropy differences were derived from ΔG and ΔH by the relation $T\Delta S = \Delta H - \Delta G$. It is clear from the table that around temperatures near 0° C (Ala)$_{10}$ favors helix state with $\Delta G = -3 \sim -4$ kcal/mol and (Gly)$_{10}$ favors non-helix state with $\Delta G = 2.7 \sim 3$ kcal/mol, while (Val)$_{10}$ slightly favors non-helix state with $\Delta G = 0.4 \sim 0.8$ kcal/mol. These results again support the experimental fact that Ala is a helix former and Gly is a helix breaker, while Val comes in between the two. Note that ΔH is large negative for (Ala)$_{10}$, whereas it is small for (Val)$_{10}$ and (Gly)$_{10}$, suggesting that ΔH is a key factor for helix formation.[8] Note also that the entropy contribution $-T\Delta S$ tends to monotonically increase with temperature as it should because of the increased thermal fluctuations.

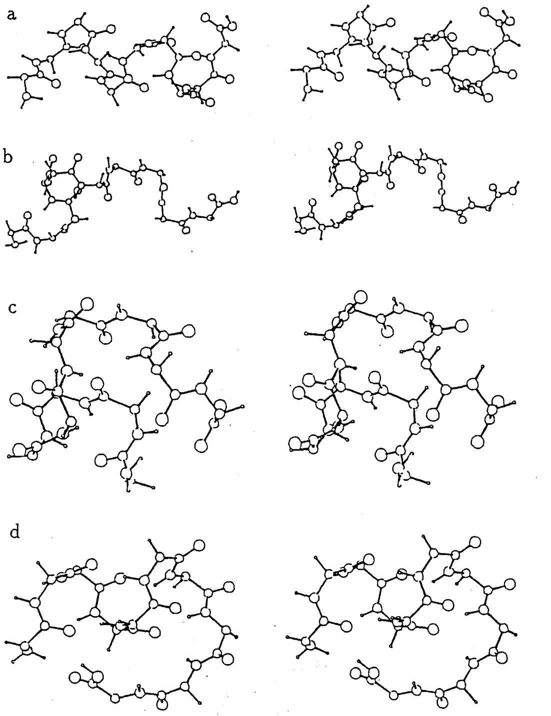

Figure 2. Stereo view of the lowest-energy conformations for a: (Ala)$_{10}$, b: (Val)$_{10}$, c: (Gly)$_{10}$, and d: a conformation with a left-handed α-helix obtained for (Gly)$_{10}$.

Besides thermodynamic quantities, multicanonical simulations can also find the lowest-energy conformations. In Figure 2 we show such conformations that were obtained in the present simulations. These conformations correspond to the structure that would be obtained at a very low temperature. The conformation for (Ala)$_{10}$ (Figure 2a) is a complete right-handed α-helix. That for (Val)$_{10}$ (Figure 2b) is also helical with more than one turn of the α-helix (length $\ell = 4$). On the other hand, the one for (Gly)$_{10}$ (Figure 2c) is not a helix. However, as will be discussed in detail elsewhere, we did obtain a *left-handed* α-helix of length $\ell = 6$ with the energy near the lowest energy (Figure 2d) by small modification of the algorithm.[11] In fact, we found many right-handed and left-handed α-helices of similar size for (Gly)$_{10}$ with roughly the same energy. These findings confirm the conjecture that α-helix is energetically favored in general;[8] even the strongest helix breaker (Gly)$_{10}$ gives a long helix (~ two turns) (Figure 2d), although at an experimentally relevant temperature, helical conformations are hardly observed for this homo-oligomer (Figure 1 and Tables 1 and 2).

The outstanding characteristic of the present work lies in the fact that we observe direct *folding* of helices from completely *random* initial conformations. We have demonstrated the superiority of multicanonical algorithm to conventional methods by calculating various thermodynamic quantities as functions of temperature from only *one* simulation run.

This work was supported, in part, by MK Industries. Our simulations were performed on the cluster of fast RISC workstations at SCRI and NEC SX-3/34R at IMS.

References and Notes

1 B.A. Berg and T. Neuhaus, *Phys. Lett.*, **B267**, 249 (1991).
2 B.A. Berg and T. Celik, *Phys. Rev. Lett.*, **69**, 2292 (1992).
3 B.A. Berg, T. Celik, and U. Hansmann, *Europhysics Lett.*, **22**, 63 (1993).
4 T. Celik, U. Hansmann, and M. Katoot, *J. Stat. Phys.*, **73**, 775 (1993).
5 U. Hansmann and Y. Okamoto, *J. Comp. Chem.*, **14**, 1333 (1993).
6 Y. Okamoto, M. Fukugita, T. Nakazawa, and H. Kawai, *Protein Engng*, **4**, 639 (1991).
7 M.J. Sipple, G. Némethy, and H.A. Scheraga, *J. Phys. Chem.*, **88**, 6231 (1984), and references therein.
8 Y. Okamoto, *Proteins: Struct. Funct. Genet.*, **19**, 14 (1994).
9 A. Chakrabartty and R.L. Baldwin, "*Protein Folding: In Vivo and In Vitro*," ed by J. Cleland and J. King, ACS Press, Washington, D.C. (1993) pp. 166–177 .
10 B.H. Zimm and J.K. Bragg, *J. Chem. Phys.*, **31**, 526 (1959).
11 U. Hansmann and Y. Okamoto, *J. Phys. Soc. Jpn.*, **63**, 3945 (1994).