Ulrich H.E. Hansmann [⋆] Yuko Okamoto [†]

Multicanonical Approach in Statistical Mechanics of Peptides[*]

[⋆] Visiting Scientific Researcher at ZIB

Present address:
IPS
ETH–Zentrum
8092 Zürich
Switzerland

[†] Department of Physics
Nara Women's University
Nara 630
Japan

[*] Talk, presented at the international symposium on Lattice Field Theory, LATTICE '94, Bielefeld, Germany, 27th September – 1st October 1994.

Preprint SC 94–32 (November 1994)

Ulrich H.E. Hansmann Yuko Okamoto

Multicanonical Approach in Statistical Mechanics of Peptides

Abstract

We test the performance of the multicanonical approach for biological molecules. The simulated molecules are frustrated systems with a complicated energy landscape. The resulting slowing down in simulations is alleviated by our ansatz. We perform a multicanonical simulation of nonpolar amino acids and study their α-helix propensities. The results are shown to be in agreement with recent experimental results.

1 INTRODUCTION

Prediction of three-dimensional structure of proteins, which determine their biological functions, solely from their amino acid sequence remains one of the unsolved problems in bioscience (for a recent review, see, for example, Ref. [1]). The difficulty with the conventional methods such as molecular dynamics lies in the fact that simulations strongly depend on the initial conditions and cannot reach thermal equilibrium within the presently available computer resources. To alleviate the above difficulty we recently proposed the application of the *multicanonical algorithm* [2] to the protein folding problem [3]. To further test our approach we performed multicanonical simulations on α-helix propensities of homo-oligomers of nonpolar amino acids [4]. Recent experimental measurements [5, 6] suggest large differences in helix propensities among the amino acids while the older host-guest method [7] indicated small differences (for a review see Ref. [5]). Our aim was to reproduce this experimental results in a numerical simulation for three characteristic amino acids: alanine (helix former), glycine (helix breaker), and valine (helix indifferent).

The multicanonical ansatz consists of three steps. First multicanonical weight factors are constructed in a recursive way [3]. This allows to simulate a "multicanonical" ensemble [2] in which all energies enter with equal probability. With respect to this ensemble equilibrium configurations are generated by the standard Monte Carlo procedure. Since the energy is forced onto a $1d$ random walk by performing a simulation in this new ensemble, one avoids getting trapped in a local minimum and the probability of finding the global minimum is increased. In the last step canonical expectation values are calculated by re-weighting [8] over a wide range of temperatures [3].

2 METHOD

We considered three homo-oligomers of 10 amino acids, $(\mathrm{Ala})_{10}$, $(\mathrm{Val})_{10}$, and $(\mathrm{Gly})_{10}$. The computer code KONF90 [9] was modified to accommodate the multicanonical method. The semi-empirical potential energy function that we used is given by the sum of the electrostatic term, 12-6 Lennard-Jones term, and hydrogen-bond term for all pairs of atoms in the peptide together with the torsion term for all torsion angles, with their parameters adopted from ECCEP/2 as documented in Ref. [10]. The peptide-bond dihedral angles ω were fixed at the value 180° and the dielectric constant was set equal 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. 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.

The criterion we adopt for α-helix formation is as follows: 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 Fig. 1 we show the average % helix per residue $\frac{<n>}{N}$ $(N = 10)$ as a function of temperature for each homo-oligomer. $(\mathrm{Ala})_{10}$ 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 experimental results [5, 6].

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

From the average of n and ℓ one can calculate the helix propagation parameter s of the Zimm-Bragg model [11] by

$$s = \frac{(<\ell> - 1)(1 - \frac{<n>}{N})}{<\ell>(1 - \frac{<n>}{N}) - \frac{<n>}{N}} \ . \tag{1}$$

This parameter was also obtained by recent experiments [5, 6]. We found $s(\mathrm{Ala}) = 1.5 \sim 1.6$, $s(\mathrm{Val}) = 0.37 \sim 0.45$, and $s(\mathrm{Gly}) = 0.13 \sim 0.16$ around the experimentally

Peptide	T	ΔG	ΔH	$T\Delta S$
$(\mathrm{Ala})_{10}$	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)
$(\mathrm{Val})_{10}$	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)
$(\mathrm{Gly})_{10}$	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)

Table 1: Free energy differences ΔG, enthalpy differences ΔH, and entropy differences $T\Delta S$ (all in kcal/mol) between helix and non-helix states as functions of temperature T.

relevant temperature ($\sim 0^\circ$ C). These values are in remarkable agreement with the experiments [5, 6], where they give $s(\mathrm{Ala}) = 1.54 \sim 2.19$, $s(\mathrm{Val}) = 0.20 \sim 0.93$, and $s(\mathrm{Gly}) = 0.02 \sim 0.57$.

In Table 1 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 as 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 $(\mathrm{Ala})_{10}$ favors helix state with $\Delta G = -3 \sim -4$ kcal/mol and $(\mathrm{Gly})_{10}$ favors non-helix state with $\Delta G = 2.7 \sim 3$ kcal/mol, while $(\mathrm{Val})_{10}$ slightly favors non-helix state with $\Delta G = 0.4 \sim 0.8$ kcal/mol. These results again support the experimental fact that alanine is a helix former and

glycine is a helix breaker, while valine comes in between the two. Note that for each homo-oligomer the entropy contribution $-T\Delta S$ monotonically increases with temperature as it should because of the increased thermal fluctuations. Note also that ΔH is large negative for $(\mathrm{Ala})_{10}$, whereas it is small for $(\mathrm{Val})_{10}$ and $(\mathrm{Gly})_{10}$, suggesting that ΔH is a key factor for helix formation.

3 CONCLUSION

We studied helix-propensities of non-polar amino acids. Our numerical results could qualitatively and quantitatively reproduce recent experimental results. 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, which was previously not possible for proteins and peptides.

Acknowledgements

This work was supported by a Grant-in-Aid for Scientific Research from the Japanese Ministry of Education, Science and Culture, by the U.S. Department of Energy, contract DE-FC05-85ER250000, by the Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB) and by MK Industries. Our simulations were performed on the cluster of RISC workstations at SCRI, Tallahassee, USA and the HITAC S820/80 at IMS, Okazaki, Japan.

References

[1] Levitt, M. *Curr. Opin. Struct. Biol.*, **1**, 224 (1991).

[2] Berg, B.A. & Neuhaus, T. *Phys. Lett.* **B267**, 249–253 (1991); Berg, B.A. & Neuhaus, T. *Phys. Rev. Lett.* **68**, 9–12 (1992).

[3] Hansmann, U. & Okamoto, Y. *J. Comp. Chem.* **14**, 1333–1338 (1993).

[4] Okamoto, Y. & Hansmann, U.,*Predicting α-helix propensities of nonpolar amino acids by multicanonical algorithm*, preprint NWU 1/94, submitted for publication.

[5] Chakrabartty, A. & Baldwin, R.L. in *Protein Folding: In Vivo and In Vitro* (eds Cleland, J. & King, J.) 166–177 (ACS Press, Washington, D.C., 1993).

[6] Padamanabhan, S., Marqusee, S., Ridgeway, T., Laue, T.M. & Baldwin, R.L. *Nature* **344**, 268–270 (1990); Lyu, P.C., Liff, M.I., Marky, L.A. & Kallenbach, N.R. *Science* **250**, 669–673 (1990); Chakrabartty, A., Schellman, J.A. & Baldwin, R.L. *Nature* **351**, 586–588 (1991); Gans, P.J., Lyu, P.C., Manning, M.C., Woody, R.W. & Kallenback, N.R. *Biopolymers* **31**, 1605–1614 (1991); Stellwagen, E., Park, S.-H., Shalango, W. & Jain, A. *Biopolymers* **32**, 1193–1200 (1992).

[7] Sueki, M.,Lee, S.,Power, S.P.,Denton, J.B., Konishi, Y. and Scheraga, H.A. *Macromolecules* **17**, 148-155 (1984).

[8] A.M. Ferrenberg and R.H. Swendsen, *Phys. Rev. Lett.*, **61**, 2635 (1988); **63** , 1658(E) (1989), and references given in the erratum.

[9] Kawai,H., Okamoto,Y., Fukugita,M., Nakazawa,T.& Kikuchi,T. *Chem. Lett.* **1991**, 213–216 (1991); Okamoto, Y., Fukugita, M., Nakazawa, T. & Kawai, H. *Protein Engng* **4**, 639–647 (1991).

[10] Momany, F.A., McGuire, R.F., Burgess, A.W. & Scheraga, H.A. *J. Phys. Chem.* **79**, 2361–2381 (1975); Némethy, G., Pottle, M.S. & Scheraga, H.A. *J. Phys. Chem.* **87**, 1883–1887 (1983); Sipple, M.J., Némethy, G. & Scheraga, H.A. *J. Phys. Chem.* **88**, 6231–6233 (1984).

[11] Zimm, B.H. & Bragg, J.K. *J. Chem. Phys.* **31**, 526–535 (1959).