CHEMISTRY LETTERS, pp. 1279−1282, 1991.

Monte Carlo Simulated Annealing Prediction for α-Helix Propensity of Amino Acid Homopolymers

Masataka FUKUGITA,* Takashi NAKAZAWA,† Hikaru KAWAI,†† and Yuko OKAMOTO†††
Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606
† Department of Chemistry, Nara Women's University, Nara 630
†† Department of Physics, University of Tokyo, Hongo, Tokyo 113
††† Department of Physics, Nara Women's University, Nara 630

Monte Carlo simulated annealing calculation is carried out for various homopolymers of amino acid with 5-15 residues. Conspicuous α-helix propensity is found for poly-L-Met and poly-L-Lys. We also found that α-helices are hardly formed in poly-Gly and poly-L-Tyr. This α-helix propensity observed in the simulation is in excellent agreement with the empirical rule derived from synthetic homopolymers and from statistical analyses of peptide data bases.

In the previous publications[1] we have reported an attempt to predict the tertiary structure of the isolated C-peptide of ribonuclease A with Monte Carlo simulated annealing using generic potential energy parameters without referring to any empirical data of C-peptide structure. We have shown that α-helix appears in simulated conformations with a high probability, and that the indication from experiments for the effect of charged residues on helix stability is verified by simulations.

In this Letter we present another example, the prediction of the α-helix propensity for homopolymers. It is known from experimental studies with synthetic homopolymers and statistical analyses of the peptide data base that some amino acids are likely to form α-helices and some others tend to break them.[2−4] While the results of various analyses do not quite agree in details, the most typical of the helix former is Met. Gly, on the other hand, represents an example for the helix breaker. An attempt was made to predict structure of poly-L-Ala with $N = 19$.[5] Systematic studies for homopolymers, however, seem not available to date except for the work[6] based on statistical-mechanical models such as the Zimm-Bragg model.[7] We have obtained the α-helix propensity directly by studying how frequently α-helices are actually formed in simulations without referring to any assumptions of statistical-mechanical theories. Simulations were made for poly-L-Met, poly-L-Lys, poly-L-Ser, poly-L-Tyr, and poly-Gly with the number of amino acid residues $N = 5$, 10, and 15.

The method of simulation is Monte Carlo simulated annealing[8] as described in Refs. 9 and 1. We employed the Hamiltonian given by the sum of the electrostatic energy, the van der Waals energy, the hydrogen-bond energy for all pairs of atoms, and the torsion energy associated with dihedral angles. For energy parameters we adopted those determined for the atoms in each amino acid by Scheraga and

collaborators (ECEPP/2).[10] The effect of surrounding atoms was neglected and the dielectric constant was set equal to 2 according to the prescription given in Ref. 10. Starting simulations with a completely random initial conformation, we updated N sets of dihedral angles (ϕ, ψ) and all torsion angles in side chains (4 for Met, 5 for Lys, 2 for Ser, and 3 for Tyr; no side-chain torsion angles for Gly) with the Metropolis algorithm in a bias-free manner. The dihedral angles ω of the peptide bonds were set equal to 180°. The simulation was started with $T = 1000$ K and the temperature of the system was gradually lowered at each step of Monte Carlo updates till $T = 250$ K. Ten Monte Carlo runs were made for each homopolymer, each run consisting of 10^4 Metropolis updates of each dihedral or torsion angle. We used the computer code KONF90.[1]

The criterion for the α-helix is the same as that described in Ref. 1; we consider that the residue is in the α-helix state when the angles relevant to this residue fall in the range $(\phi, \psi) = (-60 \pm 45°, -50 \pm 45°)$ in the conformational space. The *helixness* $\bar{n}$ of a conformation is defined by the number of successive residues which are in the α-helix state. We note that $\bar{n} = 3$ corresponds to roughly one turn of the helix.

The result of our simulation for $N = 15$ is summarised in Table 1, where the number of conformations with *helixness* $\bar{n}$ is shown for each polypeptide out of 10 Monte Carlo runs. The lowest-energy state is indicated by the symbol # and its dihedral angles are given in Table 2. For poly-L-Met we observed extensive α-helix structure ($\gtrsim$ 2 turns) for 9 conformations out of 10. One of them shows a helix as long as $\bar{n} = 14$ ($\approx$4 turns), essentially all residues being in the α-helix state. Its dihedral angles show values most typical of the α-helix (see Table 2). Poly-L-Lys also tends to form an α-helix, albeit a little less frequently compared with poly-L-Met. The lowest-energy states are given by the longest helix conformations for these homopolymers. On the other hand, poly-Gly and poly-L-Tyr are quite unlikely to yield a helix. Poly-Ser comes in the middle. This trend is in excellent agreement with empirical propensity determined by Chou

Table 1. Helix formation in 10 Monte Carlo runs for homopolymers with $N = 15$. The symbol # indicates the lowest-energy state of the 10 runs

$\bar{n}$	Met	Lys	Ser	Tyr	Gly
3	0	1	2	1	1#
4	0	0	2	0	1
5	0	0	0	0	0
6	1	0	0	0	0
7	0	1	1#	0	0
8	1	0	0	0	0
9	1	0	0	0	0
10	1	0	0	0	0
11	0	1	0	0	0
12	1	1	0	0	0
13	3	3#	0	0	0
14	1#	0	0	0	0
15	0	0	0	0	0
total($\bar{n} \geq 3$)	9	7	5	1	2

Table 2. Dihedral angles for the lowest-energy conformation of the five polypeptides with $N = 15$ as marked by # in Table 1. The symbol * indicates the dihedral angles of the α-helix state

	Met		Lys		Ser		Tyr		Gly	
Residue no.	ϕ	ψ	ϕ	ψ	ϕ	ψ	ϕ	ψ	ϕ	ψ
1	108	163	35	−62	32	157	22	−40	102	59
2	−54	−34*	−63	−40*	−66	−26*	−68	122	−165	43
3	−64	−43*	−62	−37*	−75	−38*	−79	154	73	−150
4	−65	−41*	−72	−37*	−75	−33*	−95	0	−89	115
5	−63	−46*	−71	−36*	−102	38	−151	157	82	−72
6	−64	−40*	−67	−47*	−78	−34*	−147	44	59	−114
7	−61	−44*	−65	−31*	−151	178	−102	−10*	−91	58
8	−67	−40*	−70	−43*	−70	−39*	−114	161	77	−97
9	−62	−45*	−69	−32*	−66	−40*	−84	−17*	−92	−14*
10	−61	−45*	−68	−50*	−68	−40*	−113	157	−52	−59*
11	−60	−41*	−65	−40*	−72	−33*	−137	34	−78	−43*
12	−68	−46*	−63	−41*	−73	−30*	−57	−39*	−85	75
13	−60	−48*	−68	−40*	−79	−31*	−84	97	145	−46
14	−58	−39*	−70	−31*	−74	−46*	−147	157	−81	−22*
15	−76	−46*	−100	75	−68	142	−147	154	158	−71

Table 3. Average fraction of the residue being in the α-helix state $\langle n \rangle / N$

N	Met	Lys	Ser	Tyr	Gly
15	0.77	0.58	0.37	0.12	0.13
10	0.73	0.57	0.40	0.15	0.11
5	0.34	0.28	0.32	0.06	0.08

and Fasman,[3] who gave the average α helix stability constant $s = 1.20$, 1.07, 0.79, 0.61, and 0.53 for Met, Lys, Ser, Tyr, and Gly (the values given by Finkelstein[3] are 1.15, 1.15, 0.75, 1.10, and 0.55 for these 5 amino acids; Scheraga[3] gave 1.17, 0.77, 0.99, and 0.60 for Met, Ser, Tyr, and Gly).

Table 3 presents the average fraction of residues in the α-helix state $\langle n \rangle / N$ for $N = 5$, 10, and 15, where n is the number of residues in the α-helix state and $\langle n \rangle$ is its average over the conformations obtained in the simulation. α-Helix propensity seen in Table 1 is also clearly visible in this Table. The values for $N = 10$ generally agree with those for $N = 15$. The value $\langle n \rangle / N$, however, significantly decreases towards $N = 5$. This is in agreement with the empirical fact that a small peptide tends to have no stable structure. From our $\langle n \rangle / N$ for a larger N we may extract provisionally the s value with the aid of the Zimm-Bragg model,[7] which predicts

$$\frac{\langle n \rangle}{N} = \frac{1}{2} \left[1 + (s-1)/\sqrt{(s-1)^2 + 4s\sigma} \right]$$

with σ the helix nucleation constant. If we assume a value $\sigma = 10^{-2}$, we obtain from the simulation for $N = 10 - 15$ $s = 1.13$ (Met), 1.03 (Lys), 0.95 (Ser), 0.80 (Tyr), and 0.79 (Gly). If we take a more conventional value of $\sigma = 10^{-3} - 10^{-4}$, however, s turns out to be too close to unity.

It was often argued[11] that short α-helices of the size range found in globular protein (6-20 residues) are highly unstable in water and that some specific stabilisation mechanism is necessary to understand, for instance, the α-helix in the isolated C-peptide of ribonuclease A. It is interesting to note that α-helices appear with a high frequency in our simulation even for a peptide as small as $N = 10 - 15$, and sometimes for $N = 5$. This contrasts with what the Zimm-Bragg model predicts for a small peptide.

To conclude, we have obtained the α-helix propensities for a number of amino acids in agreement with the empirical systematics with the direct simulation for homopolymers. Furthermore, our result suggests that the present method would provide a useful tool for testing the validity and the limitation of phenomenological models conventionally used to understand the empirical systematics.

References

1) H. Kawai, Y. Okamoto, M. Fukugita, T. Nakazawa, and T. Kikuchi, *Chem. Lett.*, **1991**, 213; Y. Okamoto, M. Fukugita, T. Nakazawa, and H. Kawai, *Protein Eng.* (1991), in press.

2) E.R. Blout, C. de Lozé, S.M. Bloom, and G.D. Fasman, *J. Am. Chem. Soc.*, **82**, 3787 (1960); D.R. Davies, *J. Mol. Biol.*, **9**, 605 (1964).

3) P.Y. Chou and G.D. Fasman, *Biochem.*, **13**, 222 (1974); A.V. Finkelstein and O.B. Ptitsyn, *J. Mol. Biol.*, **103**, 15 (1976); H.A. Scheraga, *Pure Apple. Chem.*, **50**, 315 (1978); H.A. Scheraga, in T.E. Creighton, "Proteins" (Ref.4).

4) G.E. Schulz and R.H. Schirmer, "Principles of Protein Structure", Springer, Berlin (1979); T.E. Creighton, "Proteins", W.H. Freeman, New York (1984).

5) D.R. Ripoll and H.A. Scheraga, *Biopolymers*, **27**, 1283 (1988).

6) M. Gō, N. Gō, and H.A. Scheraga, *J. Chem. Phys.*, **54**, 4489 (1971) and references therein.

7) B.H. Zimm and J.K. Bragg, *J. Chem. Phys.*, **31**, 526 (1959).

8) S. Kirkpatrick, C.D. Gelatt, Jr., and M.P. Vecchi, *Science*, **220**, 671 (1983).

9) H. Kawai, T. Kikuchi, and Y. Okamoto, *Protein Eng.*, **3**, 85 (1989).

10) F.A. Momany, R.F. McGuire, A.W. Burgess, and H.A. Scheraga, *J. Phys. Chem.*, **79**, 2361 (1975); G. Némethy, M.S. Pottle, and H.A. Scheraga, *ibid.*, **87**, 1883 (1983); M.A. Sippl, G. Némethy, and H.A. Scheraga, *ibid.*, **88**, 6231 (1984).

11) A. Bierzynski, P.S. Kim, and R.L. Baldwin, *Proc. Natl. Acad. Sci. U.S.A.*, **79**, 2470 (1982).

(Received March 15, 1991)