Three-Dimensional Conformations of Polypeptide Chains by a Monte Carlo Calculation. II. Dimensional Change in the Helix–Coil Transition Region

Seiji Tanaka and Akio Nakajima*

Department of Polymer Chemistry, Kyoto University, Kyoto 606, Japan.
Received April 17, 1972

ABSTRACT: The effects of long-range interactions on conformations of polypeptide chains in the helix–coil transition region were investigated. Calculations were carried out for poly(L-alanine) with chain length 151. The generation of helical and coil sequences was performed by means of the conditional probabilities in a Markoffian process translated from the Ising model for helix–coil transitions. The helical content was controlled by the parameters S and σ of the Zimm–Bragg theory. The conformations allowed for the residues involved in randomly coiling sequences are based on the model proposed in the previous paper of this series. Also, the chain conformations generated were classed as "allowed" or "disallowed" in accordance with the nonbonded interaction energies calculated by using the Lennard-Jones potential function for pairs of atoms and atomic groups. The unperturbed chain dimensions were calculated by two different methods, *i.e.*, a Monte Carlo method in which checking of atomic overlaps was omitted, and an exact evaluation by conformational statistics of the polymer chain. The results on unperturbed chain dimensions were compared with those on chain dimensions in the presence of the excluded volume effect.

The effect of long-range interactions should play an important role not only in randomly coiling conformations but also in partially helical conformations of proteins and polypeptides in the region of the helix–coil transition.

In recent years, extensive theoretical[1-8] and experimental[9-11] work has been carried out in order to understand the conformations and physical properties of synthetic polypeptides. Theoretical calculations have been performed to predict the stable helical structures,[12,13] and other conformational properties[14,15] of various polypeptides. In the studies mentioned above[1-8,12-15] the conformational energy calculations were carried out based on semiempirically available information *i. e.*, the potential functions with respect to the torsional energy about single bonds, nonbonded interactions, electrostatic interactions, and hydrogen bonding have been used instead of the hard-sphere potential used previously.[16,17] This energy calculation method has made possible the determination of the most stable helical structure and the evaluation of statistical weights for various conformations in the random-coil state. Statistical-mechanical treatments of helix–coil transitions of polypeptides have been developed by a number of investigators.[18-22] In the statistical mechanical treatments of chain conformation and the helix–coil transition, polypeptide chains have been treated as one-dimensional cooperative systems due to the short-range intramolecular interactions. These treatments do not take into consideration the long-range interactions, except hydrogen bonding, between atoms and atomic groups separated by a considerable number of residues. The conformational properties obtained from such treatments correspond to those observed experimentally in an ideal state called the unperturbed state. As is well known, however, conformations of polymer chains in the nonideal state are essentially affected by long-range intramolecular interactions.

More recently, the excluded-volume problem for polypeptide chains has been investigated, using the Monte Carlo method, by Knaell and Scott[23] and also by ourselves.[24] These papers have proposed simplified models for poly(L-alanine) in the randomly coiling conformation. In our model semiempirical potential functions were used to calculate the nonbonded and electrostatic energies of long-range interactions. For non-self-intersecting chains generated by the Monte Carlo method, we have obtained[24] an excluded-volume effect which is essentially consistent with those obtained for self-avoiding random walks on regular lattices[25,26] or nonlattice.[27,28]

In the present paper, we investigate the effects of long-range interactions on the conformations of polypeptide chains, in particular of poly(L-alanine), in the region of the helix–coil transition. We use a Monte Carlo method in which the model of a randomly coiling polypeptide proposed in the previous paper of this series[24] is connected with the

(1) D. A. Brant and P. J. Flory, *J. Amer. Chem. Soc.*, **87**, 663, 2791 (1965).
(2) W. G. Miller, D. A. Brant, and P. J. Flory, *J. Mol. Biol.*, **23**, 47, 67 (1967).
(3) P. R. Schimmel and P. J. Flory, *Proc. Nat. Acad. Sci. U. S.*, **58**, 52 (1967).
(4) P. J. Flory and P. R. Schimmel, *J. Amer. Chem. Soc.*, **89**, 6807 (1967).
(5) P. J. Flory, "Statistical Mechanics of Chain Molecules," Interscience, New York, N. Y., 1969.
(6) S. Tanaka and A. Nakajima, *Polym. J.*, **1**, 71 (1970).
(7) S. Tanaka and A. Nakajima, *ibid.*, **2**, 717 (1971).
(8) S. Tanaka and A. Nakajima, *ibid.*, **2**, 725 (1971).
(9) D. A. Brant and P. J. Flory, *J. Amer. Chem. Soc.*, **87**, 2788 (1965).
(10) S. Tanaka and A. Nakajima, *Polym. J.*, in press.
(11) A. Nakajima, S. Tanaka, and K. Itoh, *ibid.*, **3**, 398 (1972).
(12) R. A. Scott and H. A. Scheraga, *J. Chem. Phys.*, **45**, 2091 (1966).
(13) T. Ooi, R. A. Scott, G. Vanderkooi, and H. A. Scheraga, *ibid.*, **46**, 4410 (1967).
(14) D. A. Brant, *Macromolecules*, **1**, 291 (1968).
(15) S. Tanaka and A. Nakajima, *Polym. J.*, **1**, 505 (1970).
(16) G. N. Ramachandran, C. Ramakrishnan, and V. Sasisekharan, *J. Mol. Biol.*, **7**, 95 (1963).
(17) G. Némethy and H. A. Scheraga, *Biopolymers*, **3**, 155 (1965).
(18) B. H. Zimm and J. K. Bragg, *J. Chem. Phys.*, **31**, 526 (1959).
(19) S. Lifson and A. Loig, *ibid.*, **34**, 1963 (1961).
(20) K. Nagai, *ibid.*, **34**, 887 (1961).
(21) W. G. Miller and P. J. Flory, *J. Mol. Biol.*, **15**, 298 (1966).
(22) D. Poland and H. A. Scheraga, "Theory of Helix-Coil Transition in Biopolymers," Academic Press, New York, N. Y., 1970.
(23) K. K. Knaell and R. A. Scott, *J. Chem. Phys.*, **54**, 566 (1971).
(24) S. Tanaka and A. Nakajima, *Macromolecules*, **5**, 708 (1972).
(25) F. T. Wall, S. Windwer, and P. J. Gans, "Methods in Computational Physics," Vol. 1, Academic Press, New York, N. Y., 1963, p 217.
(26) F. T. Wall and J. J. Erpenbeck, *J. Chem. Phys.*, **30**, 634, 637 (1959).
(27) S. Windwer, *ibid.*, **43**, 115 (1965).
(28) E. Loftus and P. J. Gans, *ibid.*, **49**, 3828 (1968).

theoretical treatment on the helix–coil transition of polypeptides proposed by Zimm and Bragg.[18] The treatment in this study may be applied to the conformational study of protein molecules.

The present paper, however, is not concerned with the effect of long-range interactions on the helix–coil transition itself. Moreover, solvent effects and hydrophobic interactions are not taken into consideration.

(I) Model for Randomly Coiled and Helical Conformations

Poly(L-alanine) was selected for the calculation presented in this paper. The amide bond is fixed in the planar-trans conformation owing to its partial double bond nature. The chain conformations are described by the rotations about the backbone N–C^{α} and C^{α}–C′ single bonds, whose angles of rotation are denoted by ϕ and ψ, respectively. In the present paper we use the standard conventions[29] for polypeptide chains. The bond lengths and bond angles used in these calculations are the same as those used in a preceding paper[24] of this series. Figure 1 illustrates the chemical structure of a polypeptide of chain length $N = 6$ (R = $-CH_3$ for poly(L-alanine)), division of the molecule into residual units, and the position of hydrogen bonds.

Polypeptide chains in the intermediate region of the helix–coil transition consist of randomly coiling and helical portions. For amino acid residues in the random-coil state, we have employed the three conformational states proposed in our previous study[24] on non-self-interesecting chains of randomly coiling polypeptides. These three states correspond to the minima in the conformational energy map for the random coil of poly(L-alanine), *i.e.*, the right-handed α-helical conformation, the left-handed α-helical conformation, and the antiparallel pleated sheet. The rotational angles for these three conformational states are given in Table V of ref 24. For the structure of the helical portion we adopted the right-handed α helix. As shown in Figure 1, the NH group of the ith residue is hydrogen bonded to the CO group of the $(i - 4)$ residue. The distribution of helical and coiling sequences is determined by the condition of helix–coil equilibrium. The treatments will be described in detail in the following section.

In this paper we use the term "non-self-intersecting chain" with the meaning used in our previous paper.[24] The checking of atomic overlaps for generated chains was carried out by the procedure described in the previous paper.[24] The conformations are classed into "allowed" and "disallowed" according to the nonbonded interaction energies calculated by eq 1. Statistical-mechanical averaging was carried out by two methods. One method takes into consideration the contribution of conformational energies from nonbonded interactions and electrostatic interactions; the other method is that used for the hard-sphere model. The term "hard-sphere model" has been defined previously.[24] Interaction energies used for checking atomic overlaps and conformational energies of generated non-self-intersecting chains are calculated from the semiempirical potential functions which were employed in previous work.[1–8,12–15] With respect to the interaction energies between nonbonded atoms j and k, we used the Lennard-Jones 6–12 potential function

$$V_{(W)jk} = d_{jk}/r_{jk}^{12} - e_{jk}/r_{jk}^{6} \qquad (1)$$

The energy of interaction between dipoles of amide groups

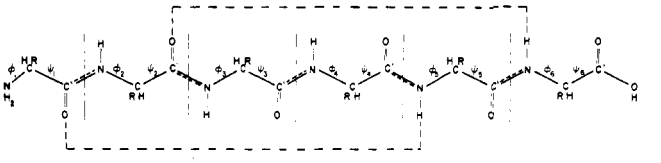

Figure 1. An α-L-peptide chain ($N = 6$) in its fully extended conformation. Hydrogen bonds in the helix conformation are denoted by broken lines. For poly(L-alanine) the R groups are CH_3.

was calculated by means of the point-monopole approximation by using

$$V_{(D)jk} = 332.0\delta_j\delta_k/\epsilon r_{jk} \qquad (2)$$

The method and the parameters used in the calculation of d_{jk} and e_{jk} in eq 1, the partial charges δ and the effective local dielectric constant ϵ are the same as those given in the preceding paper.[24]

(II) On Helix–Coil Transition

Distribution of helical and coil sequences in a polypeptide chain is determined by the equilibrium condition of the helix–coil transition. Theoretical treatments of helix–coil transitions have been carried out[18–22] by applying the one-dimensional Ising model. The process of chain generation employed in this calculation cannot be treated as the problem in the Ising model, because we cannot simultaneously know the conformational states for all the residues in a chain before the chain generation is completed. In other words, when determining the conformational state for the ith residue, we know conformational states of the residues up to the $(i - 1)$ residue, but do not know those of "future" residues, *i.e.*, those of the residues succeeding the ith residue. Therefore, the problem of helix–coil transitions used in this study should be regarded as a problem in a Markoffian process. For the purpose of the present investigation, it is necessary to transform the problem of the helix–coil equilibrium in the one-dimensional Ising model into that of the Markoffian process.

According to the matrix treatment of the nearest-neighbor Ising model,[30,31] the statistical-weight matrix **M**, in which a statistical weight for the ith residue is assigned with respect to the state of the preceding residue, is given by

$$\mathbf{M} = \begin{bmatrix} m_{cc} & m_{ch} \\ m_{hc} & m_{hh} \end{bmatrix} \qquad (3)$$

where the subscripts c and h denote coil and helical conformational states, respectively. According to Zimm and Bragg,[18] the helix–coil transition in the polypeptide chain was interpreted by two parameters, *i.e.*, the internal partition factor S and the helix nucleation parameter σ, and eq 3 is alternatively given by

$$\mathbf{M}_i = \begin{bmatrix} 1 & \sigma S \\ 1 & S \end{bmatrix}_i \qquad (4)$$

for the ith amino acid residue. The Markoffian process is characterized by conditional transition probabilities given by forms corresponding to those of helix–coil transition phenomena, such as

$$\mathbf{P} = \begin{bmatrix} p_{cc} & p_{ch} \\ p_{hc} & p_{hh} \end{bmatrix} \qquad (5)$$

(29) J. T. Edsall, P. J. Flory, J. C. Kendrew, A. M. Liquori, G. Némethy, G. N. Ramachandran, and H. A. Scheraga, *Biopolymers*, **4**, 121 (1966).

(30) H. A, Kramers and G. H. Montroll, *Rev. Mod. Phys.*, **60**, 252 (1953).

(31) G. F. Newell and E. W. Montroll, *ibid.*, **25**, 353 (1953).

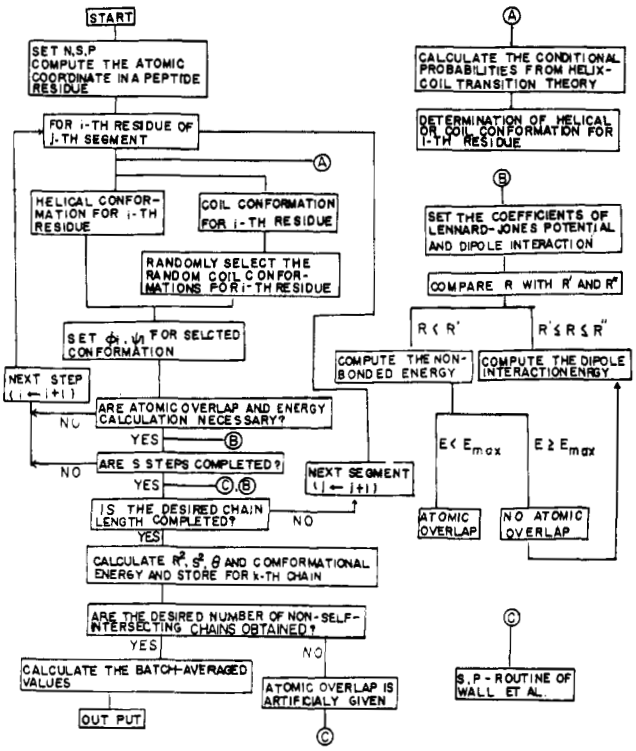

Figure 2. Flow chart for the Monte Carlo calculation.

Correlations of the Ising model with the Markoffian process were treated in general by Birshtein and Ptitysn[32] and by Flory.[5] In the following we shall show the relations between eq 3 or 4 and 5, for both cases of infinite chain length and finite chain length.

(A) For Finite Chain Length. Here we are concerned with an exact treatment of a Markoffian chain of finite chain length.

For the ith residue we define the conditional probability matrix $\mathbf{P}_i$ by

$$\mathbf{P}_i = \begin{bmatrix} p_{cc;\,i} & p_{ch;\,i} \\ p_{hc;\,i} & p_{hh;\,i} \end{bmatrix} \tag{6}$$

and the second-order *a priori* probability matrix $\mathbf{F}_i$ by

$$\mathbf{F}_i = \begin{bmatrix} F_{cc;\,i} & F_{ch;\,i} \\ F_{hc;\,i} & F_{hh;\,i} \end{bmatrix} \tag{7}$$

By using the first-order *a priori* probabilities $F_{c;\,i-1}$ and $F_{h;\,i-1}$ $(=1 - F_{c;\,i-1})$ for the $(i - 1)$ residue, we obtain the following relations

$$\begin{aligned} p_{hh;\,i} &= F_{hh;\,i}/F_{h;\,i-1} \\ p_{ch;\,i} &= F_{ch;\,i}/F_{c;\,i-1} \end{aligned} \tag{8}$$

Stochastic character leads to

$$\begin{aligned} F_{c;\,i-1} + F_{h;\,i-1} &= 1 \\ p_{cc;\,i} + p_{ch;\,i} &= 1 \\ p_{hc;\,i} + p_{hh;\,i} &= 1 \end{aligned} \tag{9}$$

and

$$F_{cc;\,i} + F_{ch;\,i} + F_{hc;\,i} + F_{hh;\,i} = 1$$

(32) T. M. Birshtein and O. B. Ptitysn, "Conformations of Macromolecules," translated from the 1964 Russian edition by S. N. Timasheff and M. J. Timasheff, Interscience, New York, N. Y., 1966.

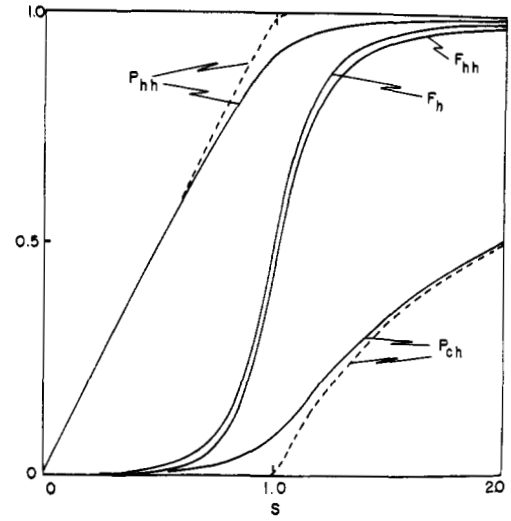

Figure 3. The dependence of the first-order *a priori* probability F_h, the second-order *a priori* probability F_{hh}, and the conditional transition probabilities p_{hh} and p_{ch} on S (Zimm–Bragg parameter) for infinite chain length; solid lines are for $\sigma = 10^{-2}$ and broken lines for $\sigma = 10^{-4}$.

On the other hand, *a priori* probabilities can be readily obtained from the statistical-weight matrix of eq 3 or 4 in the following manner. From eq 4, the partition function Z for a polypeptide molecule is given by

$$Z = \mathbf{e}^* \left[\prod_{i=1}^{N} \mathbf{M}_i \right] \mathbf{e} \tag{10}$$

where N is the degree of polymerization and $\mathbf{e}^*$ and $\mathbf{e}$ are denoted by

$$\begin{aligned} \mathbf{e}^* &= [1 \quad 0] \\ \mathbf{e} &= \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{aligned} \tag{11}$$

A first-order *a priori* probability $F_{h;\,i-1}$ is obtained according to the mathematical method developed by Flory,[5] as

$$F_{h;\,i-1} = Z^{-1}\mathbf{e}^* \left[\prod_{j=1}^{i-2} \mathbf{M}_j \right] \mathbf{M}'_{h;\,i-1} \left[\prod_{k=i}^{N} \mathbf{M}_k \right] \mathbf{e} \tag{12}$$

where

$$\mathbf{M}'_{h;\,i-1} = \partial \mathbf{M}_{i-1}/\partial \ln \mathbf{m}^*_{h;\,i-1} \tag{13}$$

The vector matrix $\mathbf{m}^*_{h;\,i-1}$ appearing in eq 13 is given by the second column in eq 3 as

$$\mathbf{m}^*_{h;\,i-1} = \begin{bmatrix} m_{ch} \\ m_{hh} \end{bmatrix} \tag{14}$$

A second-order *a priori* probability $F_{hh;\,i}$ is obtained by

$$F_{hh;\,i} = Z^{-1}\mathbf{e}^* \left[\prod_{j=1}^{i-1} \mathbf{M}_j \right] \mathbf{M}'_{hh;\,i} \left[\prod_{k=i+1}^{N} \mathbf{M}_k \right] \mathbf{e} \tag{15}$$

where

$$\mathbf{M}'_{hh;\,i} = \partial \mathbf{M}_i/\partial \ln \mathbf{m}_{hh;\,i} \tag{16}$$

and $m_{hh;\,i}$ is given in eq 3.

When $F_{h;\,i}$ is obtained in a manner similar to that described above, $p_{ch;\,i}$, etc. are obtained by the following relations

$$\begin{aligned} F_{c;\,i-1}p_{ch;\,i} + F_{h;\,i-1}p_{hh;\,i} &= F_{h;\,i} \\ F_{h;\,i-1}p_{hc;\,i} + F_{c;\,i-1}p_{cc;\,i} &= F_{c;\,i} \end{aligned} \tag{17}$$

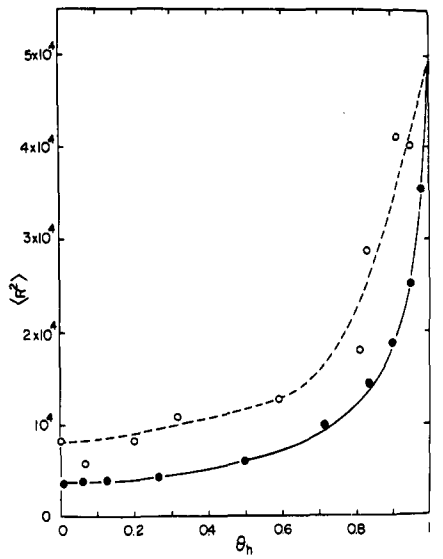

Figure 4. Plots of $\langle R^2 \rangle_H$ against the mean helical content $\langle \theta_h \rangle$ (O), $\langle R^2{}_0 \rangle^*$ against $\langle \theta_h \rangle$ (●) and $\langle R^2{}_0 \rangle$ against $\bar{\theta}_h$ (—). Calculations were performed for $N = 151$ and $\sigma = 10^{-2}$

Thus, we can obtain all the elements in eq 6 for finite chain length from eq 4 by using eq 8, 9, 12, 15, and 17.

(B) For Infinite Chain Length. For sufficiently long chains, we may treat the quantities appearing in eq 6–9 as being independent of i and $i - 1$. Hence, we define the conditional probability matrix as given by eq 5 instead of eq 6. Other quantities are also defined in a similar manner.

With use of the largest eigenvalue λ_1 of eq 4, the partition function Z can be written in the well-known form for the one-dimensional Ising model as

$$Z = \lambda_1{}^N \tag{18}$$

for sufficiently long chain length. Eigenvalues of eq 4 are given by

$$\lambda_1,\lambda_2 = (1/2)\left[(1 + S) \pm \sqrt{(1 + S)^2 + 4\sigma S}\right] \tag{19}$$

The second-order *a priori* probability F_{hh}, for example, is given by

$$F_{hh} = N^{-1}(\partial \ln Z/\partial \ln m_{hh}) = \partial \ln \lambda_1/\partial \ln m_{hh} \tag{20}$$

As proved by Birshtein and Ptitysn[32] and by Flory,[5] the derivatives $\partial\lambda_k/\partial m_{\alpha\beta}$, in which α and β denote appropriate conformational states, helix and coil, respectively, and $k = 1$ or 2 in the present case, can be expressed by the right-handed and left-handed eigenvectors of matrix **M**. It follows that

$$\partial\lambda_k/\partial m_{\alpha\beta} = V_{\beta k}U_{k\alpha}$$

where $V_{\beta k}$ and $U_{k\alpha}$, respectively, are the corresponding elements of the right-handed and left-handed eigenvector matrices **V** and **U** for matrix **M**. By substituting α, β, and k for h, h, and 1, eq 20 becomes

$$F_{hh} = m_{hh}V_{h1}U_{1h}/\lambda_1 \qquad (n \to \infty) \tag{21}$$

The first-order *a priori* probability F_h is also given by

$$F_h = V_{h1}U_{1h} \qquad (n \to \infty) \tag{22}$$

Hence, the conditional probability p_{hh} in eq 5 is given, for sufficiently long chain length, as

$$p_{hh} = F_{hh}/F_h = m_{hh}/\lambda_1 \tag{23}$$

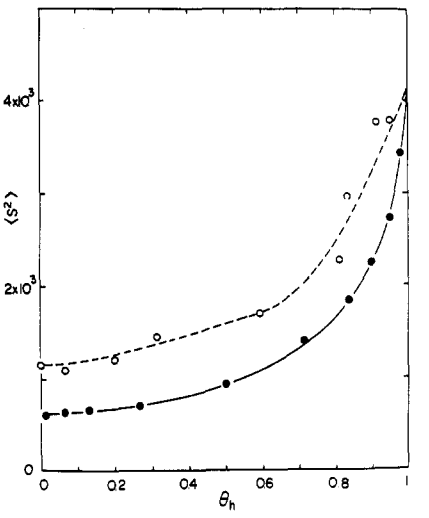

Figure 5. Plots of $\langle S^2 \rangle_H$ against the mean helical content $\langle \theta_h \rangle$ (O), $\langle S_0{}^2 \rangle^*$ against $\langle \theta_h \rangle$ (●), and $\langle S_0{}^2 \rangle$ against $\bar{\theta}_h$ (—). Calculations were performed for $N = 151$ and $\sigma = 10^{-2}$.

By applying eq 4 to the above treatments, we obtain the following relations for finite chain length

$$F_{hh} = S(\lambda_1 - 1)/\lambda_1(\lambda_1 - \lambda_2) \tag{24}$$

$$F_{ch} = (\lambda_1 - 1)(1 - \lambda_2)/\lambda_1(\lambda_1 - \lambda_2) \tag{25}$$

$$F_h = (\lambda_1 - 1)/(\lambda_1 - \lambda_2) \tag{26}$$

$$p_{hh} = S/\lambda_1 \tag{27}$$

and

$$p_{ch} = 1 - (1/\lambda_1) \tag{28}$$

The *a priori* probability F_h is also termed the helical content $\bar{\theta}_h$ in this paper. The helical content in the chains which are generated by Monte Carlo method can be related to the parameters S and σ of Zimm and Bragg's theory through the relations described above.

(III). Calculation Method

The treatment of the helix–coil transition by the Ising model was translated into that of a Markoffian process in the preceding section. Generation of helical and coil sequences along a chain is carried out by using the conditional probabilities which are previously assigned so as to reproduce the desired mean helical content. At each step in the process of generation of a non-self-intersecting chain, the conformational state is determined depending on that for $(i - 1)$ residue. If the ith residue is in a random-coil sequence, the rotational state is then determined by random selection of one from the three conformational states proposed in our previous work.[24] If the ith residue is in a helical sequence, the rotational state is unequivocally that of the right-handed α helix. After the ith residue is linked to its preceding one, residual and atomic overlaps are checked according to the method described in our previous paper.[24] This process is continued until either a non-self-intersecting chain of the desired chain length is completed or an atomic overlap is found. When an atomic overlap is found, the chain was discarded and the above process is repeated from the beginning of the chain. When the above process was completed for a non-self-intersecting chain, the square of the end-to-end distance r^2, the square of the radius of gyration s^2, the helical content, and other data were calculated and stored, and finally

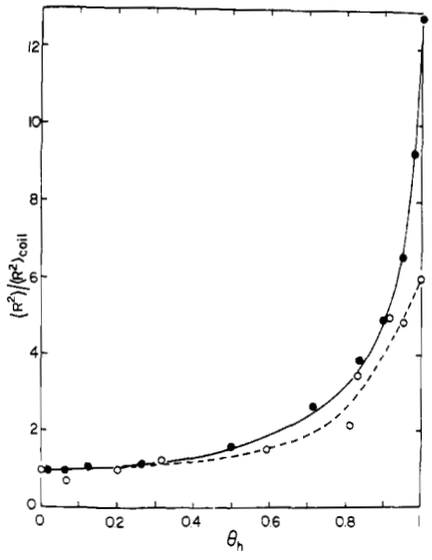

Figure 6. Plots of $\langle R^2\rangle_H/\langle R^2\rangle_{H;coil}$ against the mean helical content $\langle\theta_h\rangle$ (○), $\langle R_0^2\rangle^*/\langle R_0^2\rangle^*_{coil}$ against $\langle\tilde{\theta}_h\rangle$ (●), and $\langle R_0^2\rangle/\langle R_0^2\rangle_{coil}$ against $\tilde{\theta}_h$ (—).

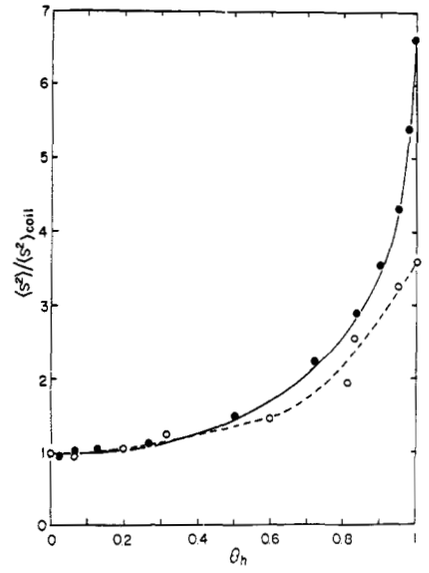

Figure 7. Plots of $\langle S^2\rangle_H/\langle S^2\rangle_{H;coil}$ against the mean helical content $\langle\theta_h\rangle$ (○), $\langle S_0^2\rangle^*/\langle S_0^2\rangle^*_{coil}$ against $\langle\theta_h\rangle$ (●), and $\langle S_0^2\rangle/\langle S_0^2\rangle_{coil}$ against $\tilde{\theta}_h$ (—).

atomic overlap was artificially given at the end of the chain to start a new attempt. The right-handed α-helical residues found in random-coil sequences, which are determined by random selection of one from three rotational states proposed for randomly coiling conformations, were counted out of the helical content, since such residues should be looked upon as those involved in random-coil sequences. The simplified flow chart for the present Monte Carlo calculation is shown in Figure 2.

In the present calculation as well as in the previous one,[24] we employed the chain enrichment technique of Wall, *et al.*[25,26] We carried out a Monte Carlo calculation for non-self-intersecting chains with $N = 151$ by assigning $s = 10$ and $p = 6$. These conditions of the s,p routine are the same as those used previously for the randomly coiling model.[24]

The square of the end-to-end distance r^2, the square of the radius of gyration s^2, and the helical content θ_h (which is not the value of F_h given theoretically in eq 26) were calculated for non-self-intersecting chains completed up to chain length 151. With respect to a fixed conditional probability, the Monte Carlo calculations were performed with five batches, each of which consists of 100 non-self-intersecting chains. For the hard-sphere model, the mean-square end-to-end distance $\langle R^2\rangle_H$, the mean-square radius of gyration $\langle S^2\rangle_H$, and the mean helical content $\langle\theta_h\rangle$ were calculated by averaging over all the batch-averaged values $\langle r^2\rangle$, $\langle s^2\rangle$, and $|\theta_h|$. By taking into consideration the contributions of conformational energies, the mean helical content batch-averaged $|\theta_h|$, the mean-square end-to-end distance $\langle R^2\rangle_{V,D}$, and the mean-square radius of gyration $\langle S^2\rangle_{V,D}$ were calculated for the individual batch consisting of 100 non-self-intersecting chains. The average value for the case of $\langle R^2\rangle_{V,D}$ is given by

$$\langle R^2\rangle_{V,D} = \frac{\sum_{j=1}^{k} r_j^2 \exp(-E_{V,D(j)}/RT)}{\sum_{j=1}^{k} \exp(-E_{V,D(j)}/RT)} \qquad (29)$$

The notations in eq 29 were given in our previous paper.[24]

In the previous Monte Carlo studies,[25-28] the non-self-intersecting chains generated have an equal probability. The conformational properties were obtained after statistical averaging with equal statistical weight for each non-self-intersecting chain. Therefore, it is important that the chain generation be carried out either in a biased or an unbiased fashion. When interaction energies depending on the chain conformation are taken into consideration, it is an important problem that the result of Monte Carlo calculation converges to the thermodynamic ensemble average. For this purpose, the following processes are considered: (1) an unbiased random-chain generation → statistical-mechanical averaging by taking into account both short-range and long-range interaction energies, (2) a biased random-chain generation → conversion into unbiased chains → statistical-mechanical averaging by taking into account both short-range and long-range interaction energies, and (3) chain generation by including the contribution from the short-range interactions → statistical-mechanical averaging by using energies excluding short-range interaction energies. The results from the above three methods converge to the same statistical-mechanical average value. The third process was employed in this paper. Thermodynamic equilibrium in the helix–coil transition is realized in the Monte Carlo method by regarding it as a Markoffian process as treated in the previous section. It is a sufficient condition for convergence of Markoffian chains to a thermodynamical ensemble that the conditional probabilities used in the chain generation satisfy both the conditions of normalization (the second and third equations of eq 9) and of reversibility $F_\beta = \Sigma_\alpha F_\alpha p_{\alpha\beta}$ or $F_\alpha p_{\alpha\beta} = F_\beta p_{\beta\alpha}$ (see eq 8), where α and β denote the conformational states. In fact, as described later, the exact values from the conformational statistical method *i.e.*, $\langle R_0^2\rangle$ and $\langle S_0^2\rangle$, agree quite well with the values $\langle R_0^2\rangle^*$ and $\langle S_0^2\rangle^*$, respectively, from the Monte Carlo calculation.

The unperturbed chain dimensions were calculated by two different methods. One method was determination of the unperturbed dimensions $\langle R_0^2\rangle^*$ and $\langle S_0^2\rangle^*$ by the Monte Carlo calculation in which checking for atomic overlaps was not performed. The actual calculations were carried out for a finite number of chains (6 batches each consisting of 100 unperturbed chains). The mean helical content $\langle\theta_h\rangle$ for the unperturbed chain was also calculated. The other method is the exact calculation by statistical-mechanical averaging over all conformations of polymer chains. Nagai[20] and Miller

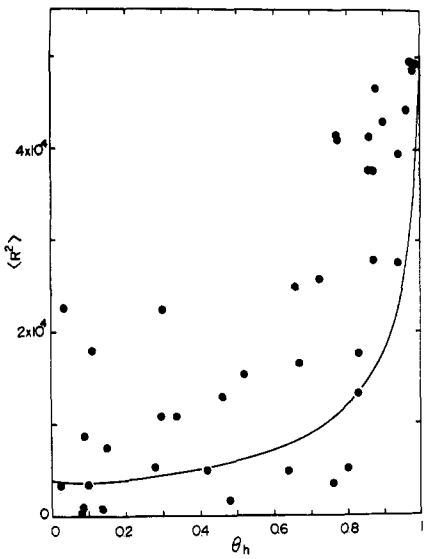

Figure 8. Plots of $\langle R^2\rangle_{V,D}$ against the mean helical content $|\theta_h|$ (●). Each circle was obtained by averaging 100 non-self-intersecting chains, using eq 29. The exact unperturbed chain dimensions $\langle R_0^2\rangle$ are given by the solid curve.

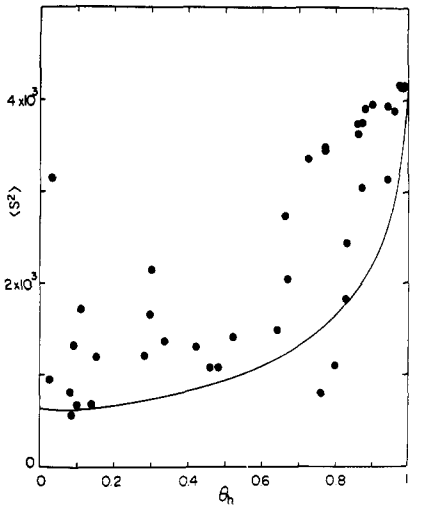

Figure 9. Plots of $\langle S^2\rangle_{V,D}$ against the mean helical content $|\theta_h|$ (●). Each circle was obtained by averaging 100 non-self-intersecting chains, using eq 29, with s_j^2 substituted for r_j^2. The unperturbed chain dimensions $\langle S_0^2\rangle$ are shown by the solid curve.

and Flory[5,21] have proposed calculation methods on the unperturbed dimensions of polypeptide chains in the region of the helix–coil transition. According to the latter authors, the unperturbed dimensions $\langle R_0^2\rangle$ and $\langle S_0^2\rangle$ are given by

$$\langle R_0^2\rangle = 2Z^{-1}[1 \quad \mathbf{0}]\mathbf{G}^{N-1}\begin{bmatrix}\mathbf{0}\\ \mathbf{1}\end{bmatrix} \tag{30}$$

and

$$\langle S_0^2\rangle = 2Z^{-1}[1 \quad \mathbf{0}]\mathbf{K}^{N-1}\begin{bmatrix}\mathbf{0}\\ \mathbf{1}\end{bmatrix} \tag{31}$$

where

$$\mathbf{G} = \begin{bmatrix}\mathbf{M} & (\mathbf{l}^T\mathbf{x}\mathbf{M})\,\mathbf{T} & (l^2/2)\,\mathbf{M}\\ \mathbf{0} & (\mathbf{E}_3\mathbf{x}\mathbf{M})\,\mathbf{T} & \mathbf{l}\mathbf{x}\mathbf{M}\\ \mathbf{0} & \mathbf{0} & \mathbf{M}\end{bmatrix} \tag{32}$$

and

$$\mathbf{K} = \begin{bmatrix}\mathbf{M}\,\mathbf{M} & (\mathbf{l}^T\mathbf{x}\mathbf{M})\,\mathbf{T} & (l^2/2)\,\mathbf{M} & (l^2/2)\,\mathbf{M}\\ \mathbf{0}\;\mathbf{M} & (\mathbf{l}^T\mathbf{x}\mathbf{M})\,\mathbf{T} & (l^2/2)\,\mathbf{M} & (l^2/2)\,\mathbf{M}\\ \mathbf{0}\;\mathbf{0} & (\mathbf{E}_3\mathbf{x}\mathbf{M})\,\mathbf{T} & \mathbf{l}\mathbf{x}\mathbf{M} & \mathbf{l}\mathbf{x}\mathbf{M}\\ \mathbf{0}\;\mathbf{0} & \mathbf{0} & \mathbf{M} & \mathbf{M}\\ \mathbf{0}\;\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{M}\end{bmatrix} \tag{33}$$

In eq 32 and 33, M is given by eq 4, l is the virtual bond vector between adjacent α-carbones,[1-8] and T is given by

$$\mathbf{T} = \begin{bmatrix}\langle\mathbf{T}_c\rangle & \mathbf{0}\\ \mathbf{0} & \mathbf{T}_h\end{bmatrix} \tag{34}$$

$\langle\mathbf{T}_c\rangle$ is the statistical-mechanical average of the transformation matrix of a virtual bond for the randomly coiling polypeptide chain[1-8] and is given, for the three-state model used in this paper, by

$$\langle\mathbf{T}_c\rangle = \begin{bmatrix}0.232 & -0.239 & 0.202\\ -0.080 & 0.226 & 0.074\\ 0.215 & 0.014 & -0.357\end{bmatrix} \tag{35}$$

$\mathbf{T}_h$ is the transformation matrix for the right-handed α-helical transformation, We calculated exactly not only $\langle R_0^2\rangle$ and $\langle S_0^2\rangle$ but also $\tilde{\theta}_h$ for $N = 151$ by the matrix multiplication method. $\tilde{\theta}_h$ was obtained by averaging eq 12 over $1 \leq i \leq N$.

(IV). Results and Discussion

In the preceding section Markoff chain models of the helix–coil transition were theoretically treated for both cases of finite and infinite chain length. For the chain length $N = 151$ the conditional probabilities p_{hh} and p_{ch} calculated by the method for finite chain lengths were approximately equal to those calculated by the method for infinite chain lengths, and they were independent of i except for several residues at the chain end.[33] In fact, as described later, the mean helical contents $\langle\theta_h\rangle$ obtained by use of the conditional probabilities for infinite chain length were in good agreement with F_h or $\tilde{\theta}_h$ predicted by the exact treatments for finite chain length ($N = 151$). Therefore, for the generation of helical and coil sequences in the present Monte Carlo calculation, we employed the conditional probabilities calculated by the method for infinite chain length. Figure 3 illustrates the dependence of the first- and second-order *a priori* probabilities, F_h and F_{hh}, and the conditional probabilities, p_{hh} and p_{ch}, upon the internal partition factor S for the cases of $\sigma = 10^{-2}$ and $\sigma = 10^{-4}$.

The σ values experimentally obtained were 1.4×10^{-4} for poly(L-alanine) in water[34] and 2.5×10^{-3} for poly(L-glutamic acid).[35] Thus σ values experimentally obtained for various polypeptides were of the order of 10^{-3}–10^{-5}. The average length of a helical or coil sequence is approximately given by $\sigma^{-1/2}$ at the midpoint of the helix–coil transition. Therefore, if σ values of the order of 10^{-4} are used in our model of finite chain length $N = 151$, then no (or at most one or two) alternation in helical and coil sequence should result. Hence, in the present calculation, $\sigma = 10^{-2}$ was rather formally employed so that various chain conformations might be obtained with our model.

The results on the mean-square end-to-end distances $\langle R^2\rangle_H$, $\langle R_0^2\rangle^*$, and $\langle R_0^2\rangle$, are shown in Figure 4, in which the first two quantities are plotted against the mean helical

(33) S. Tanaka and A. Nakajima, unpublished data.
(34) R. T. Ingwall, H. A. Scheraga, N. Lotan, A. Berger, and E. Katachalski, *Biopolymers*, 6, 331 (1968).
(35) By the evaluation of Ingwall, *et al.*, in ref 22 and 34 from the data of W. H. Miller, *et al.*, *J. Amer. Chem. Soc.*, 87, 3542 (1965).

content $\langle\theta_h\rangle$, and the last one is plotted against $\bar{\theta}_h$. The results for the mean-square radii of gyration $\langle S^2\rangle_H$, $\langle S_0^2\rangle^*$, and $\langle S_0^2\rangle$ are also shown in a similar manner in Figure 5. It should be noted that $\langle R_0^2\rangle^*$ and $\langle S_0^2\rangle^*$ calculated for a finite number of samples by the Monte Carlo method were, in good agreement with $\langle R_0^2\rangle$ and $\langle S_0^2\rangle$, respectively, from the exact calculation for chain length $N = 151$. It was also found that the mean helical content $\langle\theta_h\rangle$ obtained by the Monte Carlo generation for $N = 151$ under the conditional probabilities for infinite chain length was in good agreement with F_h or $\bar{\theta}_h$ calculated exactly by the matrix multiplication method for $N = 151$.

Figure 6 illustrates the dependence of the mean-square end-to-end distances in the helix–coil transition region relative to that in the randomly coiling conformation, $\langle R^2\rangle_H/\langle R^2\rangle_{H;coil}$, $\langle R_0^2\rangle^*/\langle R_0^2\rangle^*_{coil}$, and $\langle R_0^2\rangle/\langle R_0^2\rangle_{coil}$, on the helical content $\langle\theta_h\rangle$ or $\bar{\theta}_h$. In Figure 7 relations for the mean-square radius of gyration are shown in a similar fashion. The dimensions of $\langle R^2\rangle_{H;coil}$ and $\langle S^2\rangle_{H;coil}$ are those obtained in our previous study[24] on non-self-intersecting chains for a randomly coiling polypeptide, and hence include the excluded-volume effect. The unperturbed dimensions in the random-coil form, $\langle R_0^2\rangle^*_{coil}$, $\langle S_0^2\rangle^*_{coil}$, $\langle R_0^2\rangle_{coil}$, and $\langle S_0^2\rangle_{coil}$, are also cited from the same work.[24] As shown by Nagai,[20] the unperturbed chain dimensions during the transition from coil to helix pass through a minimum. In the model used here we cannot detect the minimum of the unperturbed dimensions in the figure, although they are present. Therefore, we could not compare the dimensions in the unperturbed state with those for non-self-intersecting chains in this respect.

The results on the mean-square end-to-end distance $\langle R^2\rangle_{V,D}$, which was evaluated for $k = 100$, in eq 29 by taking into account the contribution of conformational energies calculated from eq 1 and 2, are plotted against the mean helical content $|\theta_h|$ in Figure 8. The unperturbed dimensions $\langle R_0^2\rangle$ obtained from an exact calculation with eq 30 are also shown as the solid curve in the same figure. The results for the mean-square radius of gyration are shown in a similar manner in Figure 9. The excluded-volume effect during the helix–coil transition could be qualitatively observed in Figure 8, and more distinctly for the case of $\langle S^2\rangle_{V,D}$ in Figure 9. Some points for $\langle S^2\rangle_{V,D}$ are located below the solid curve, *i.e.*, the unperturbed dimension, $\langle S_0^2\rangle$, in the region of helical content near unity; this results from the fact that the rodlike conformation abruptly bends at the middle point of the chain. The position of the junction and the orientation of bending significantly affect $\langle R^2\rangle$ and $\langle S^2\rangle$ in the region of high helical content. The results on $\langle R^2\rangle_{V,D}$ and $\langle S^2\rangle_{V,D}$, as seen in Figure 8 and 9, were therefore scattered. More refined results may be obtained by increasing the number of samples.

Computer Simulation of Polymer Conformation. II. Distribution Function for Polymers with Excluded Volume[1,2]

Steven D. Stellman*[3] and Paul J. Gans

New York University, Department of Chemistry, New York, New York 10003.
Received July 1, 1972

ABSTRACT: Numerical distributions of end-to-end distances were generated by a Monte Carlo method for hard-sphere off-lattice polymers of length $N = 20$, 40, 60, 80, 98, and 298 atoms. Comparison by χ^2 tests against five recently proposed theoretical distribution functions showed that for $N = 80$ and $N = 98$, the data could be described, with 95% confidence, by the equation $f(r) = \exp[-(ar^2 + br + c)]$, where a and b are fitted parameters and c is a normalization constant. For $N = 298$, limitations of sample size lead to lower confidence limits (about 80%), but good fit. The above equation, and not its gaussian counterpart $\exp(-cr^2)$, is probably the limiting distribution function. The function accurately predicts the 1st through 12th observed moments at all chain lengths.

The distribution function for the end-to-end distance of a polymer with excluded volume has been the subject of many recent investigations.[4-15] Nevertheless, its theory remains very much an open question. The importance of knowing the correct function arises from the relationship of its moments, $\langle r^2\rangle$, $\langle r^4\rangle$, etc., to observable physical quantities. Some geometric properties which could be predicted with the aid of the correct distribution function are the mean-squared radius of gyration,[16] light scattering curves,[17] average dipole moment,[18] and viscosity.[19]

(1) Presented in part at the IUPAC Symposium on Macromolecules, July 2–7, 1972, Helsinki.
(2) Paper I: S. D. Stellman and P. J. Gans, *Macromolecules*, **5**, 516 (1972).
(3) To whom correspondence should be addressed at the Department of Biochemical Sciences, Frick Chemical Laboratory, Princeton University, Princeton, N. J. 08540.
(4) H. Yamakawa, *J. Chem. Phys.*, **48**, 3845 (1968).
(5) P. J. Flory, "Principles of Polymer Chemistry," Cornell University Press, Ithaca, N. Y., 1953, Chapter 10.
(6) J. Mazur, *J. Chem. Phys.*, **43**, 4354 (1965).
(7) J. Mazur, *J. Res. Nat. Bur. Stand., Sect. A*, **69**, 355 (1965).
(8) C. Domb, J. Gillis, and G. Wilmers, *Proc. Phys. Soc.*, **85**, 625 (1965).
(9) S. F. Edwards, *ibid.*, **85**, 613 (1965).
(10) Z. Alexandrowicz, *J. Chem. Phys.*, **46**, 3789 (1967).
(11) (a) H. Reiss, *ibid.*, **47**, 186 (1967); (b) J. Curro and P. J. Blatz, *ibid.*, **48**, 2832 (1968).
(12) R. Yeh and A. Isihara, *ibid.*, **51**, 1215 (1969).
(13) T. Ree Chay, *ibid.*, **52**, 1025 (1970).
(14) G. Allegra and G. Avitabile, *ibid.*, **56**, 2385 (1972).
(15) F. T. Hioe and F. T. Wall, *J. Phys. Chem.*, **74**, 4410 (1970).
(16) M. Fixman, *J. Chem. Phys.*, **36**, 306 (1962).
(17) D. McIntyre and F. Gornick, Ed., "Light Scattering from Dilute Polymer Solutions," International Science Review Series, Vol. 3, Gordon and Breach, New York, N. Y., 1963.
(18) M. V, Volkenshtein, "Configurational Statistics of Polymeric Chains," Interscience, New York, N. Y., 1963, pp 331–352.
(19) P. J. Flory and T. G. Fox, *J. Amer. Chem. Soc.*, **73**, 1904 (1951); M. V. Volkenshtein, ref 18, p 376.