November, 1987] © *1987 The Chemical Society of Japan* *Bull. Chem. Soc. Jpn.*, ***60***, *3895—3901 (1987)*

A Simple Model for the Kinetics of Beta Meander Formation of Homopolypeptide Chains

Hiroshi MAEDA
Department of Chemistry, Faculty of Science, Nagoya University, Nagoya 464
(Received June 5, 1987)

A simple kinetic theory is presented concerning the irreversible formation of intramolecularly folded β-structure of homopolypeptide chains through a bend formation. To reduce the number of parameters, the number of strands is limited to be less than five and the growth rate of the third and fourth strands is assumed to be so fast that the final strand length is attained immediately after the bend formation. Numerical calculations were carried out on relevant ranges of rate costants for relatively short chains. It was shown that experimentally found diverse kinetic patterns were all explicable based on the present model.

Folding of single homopolypeptide chains into the β-structure occurs, in principle, at any site on the chain and hence it is a stochastic process. Therefore, it differs from that of globular proteins, where the folding pathway is determined nearly uniquely by the amino acid sequence.[1—3] It is interesting and also important to know the difference between the folding of homopolypeptide chains and proteins containing the β-structure.

Several experimental studies have been done on the kinetics of intramolecular β-structures of poly(L-lysine) (PLL)[4—8] and its side-chain homolog,[9] poly(L-tyrosine) (PLT),[10,11] and poly(*S*-carboxymethyl-L-cysteine) (poly[Cys(CH_2COOH)]).[12] Thermally induced α-helix-to-β-structure conversion of PLL showed single exponential behavior,[5,6,8] while its pH-induced coil-to-β-structure conversion at high temperatures showed a lag phase when followed with light scattering or fluorescence.[7] In the case of poly(L-tyrosine), pH-induced coil-to-β-structure change was successfully described with two exponential terms and no lag phase was observed.[10] Time domains in these studies for the formation of the β-structure were about $10—10^2$ min for PLL and poly[Cys(CH_2COOH)][12] but $10—10^2$ s for PLT. Different behaviors of PLT from those of PLL were noticable as pointed out.[10] Fast changes in $10^{-2}—0.1$ s region were observed in a recent study on PLT; the results were approximately fitted with two exponential terms but multi-relaxation time character was also observed.[11] Thermal coil-β-structure conversion of poly[Cys(CH_2COOH)] showed multi-relaxation time behavior.[12]

Since intramolecular folding of single polypeptide chain certainly consists of several steps, rate constants of these steps should be related to the time constants observed in experiments. However, such an attempt has been scarcely reported.

In view of the complex nature of the whole process of chain folding and also to reduce the number of parameters, we wish to present a highly approximate model in the present report. The model is of phenomenological nature and quantitative analyses of experimental data based on the present model are not intended. However, various kinetic behaviors such as single exponential, tow exponential terms, or lag phase, are all predicted from this simple model. Therefore, all the experimental data on homopolypeptides to data can be consistently interpreted according to this model.

Model. A symmetric polymer chain consisting of L units is considered. Directional nature of polypeptide chain is tentatively neglected. Each unit can assume three states: disordered state with no hydrogen bonds, β-conformation with one and two hydrogen bonds. Three basic assumptions of the model are:

1) Only the formation of the β-conformation is considered. No reversible reaction is considered.
2) Initiation of β-strand formation is thoroughly due to bend formation. Initiation by collision of two units distant apart on the chain is excluded.
3) Multiple initiation is inhibited.

Other assumptions are introduced below in the course of the description of the present model depicted in Fig. 1.

First bend occurs with a rate constant k_i between the η-th and $(\eta+1)$-th units of a random coil chain. Units are counted from one end of a symmetric chain. No units are incorporated in the bend portion (assumption 4). It is simple to modify the present theory to take into account any (fixed) number of units in the bend portion. We make assumption 4 only for the sake of simplicity. Two β-strands thus initiated grow with a rate constant k_+. At any stage of the growth, second bend can be formed with a rate constant k_b at either of the two junctions between random coil region and the growing β-strands. Growth of the third strand is assumed to be very fast compared with k_+ and k_b and hence final strand length is attained immediately after the second bend is formed (assumption 5). Third bend is restricted to take place, in our model, at the remaining junction between random coil region and the first two β-strands. Formation of the fourth strand is assumed to take place instantly, just as the third strand. No further bend formation is

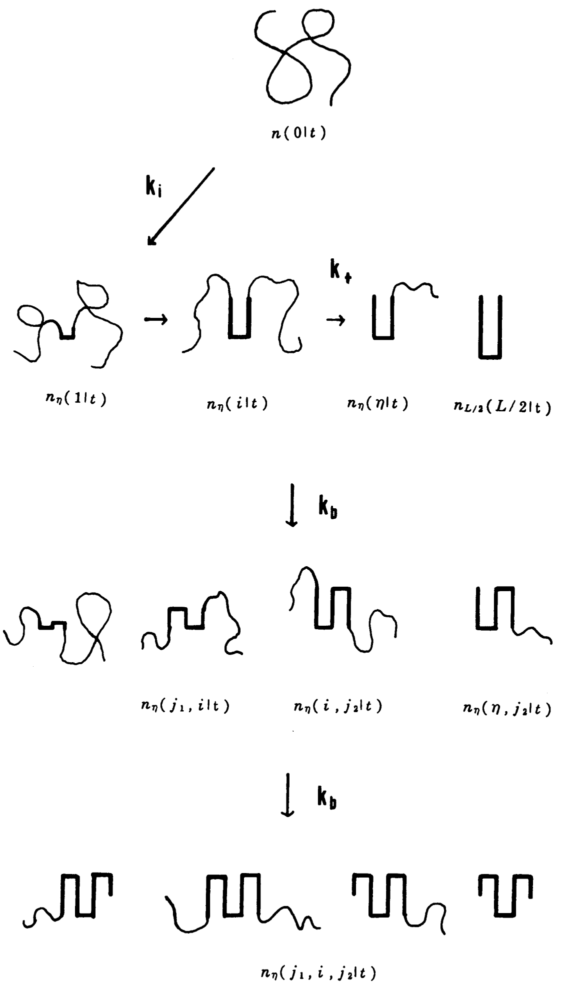

Fig. 1. Schematic representation of the model. Thick lines and thin curves represent the β-strands and random coil portions, respectively.

allowed (assumption 6). Some of these assumptions will be discussed later.

In this way, only those species consisting of two, three and four strands are considered in this model. We hope this simple model contains essentials of a general model where any number of strands is allowed. It is to be noted that three-strand species known as β-meander,[13] which are frequently found in proteins, are major products in this model.

At this stage, it is pertinent to consider the nature of the unit in the model. The unit does not necessarily correspond to a peptide unit nor an amino acid residue. Instead, it may correspond to p peptide groups together. Hence, chain length L of the model does not necessarily represent the degree of polymerization of actual polypeptide chain (DP). By virtue of this extension of the definition of unit, we need not calculate these cases with large L, say more than 100. If p is taken to 10, L=10 is sufficient to describe a chain of DP=100. With this definition of unit, the present model is also applicable to random and sequential copolypeptides, if p is taken to the repeating unit (or greater) in the latter case. Initiation rate constant k_i is not taken as position-dependent. Actually, however, an initiation near chain end is not effective, because two unstable short strands thus formed are dissociated and other initiation at the site much distant from chain ends will occur subsequently. If we take a value of p as such that formation of the β-conformation at p-th site from the chain end is stable enough, then in the present model we can assign the same value of k_i to the units near chain end as well as to other units in the interior of the chain. On the other hand, this extention of the definition of unit invalidates quantitative analyses of experimental data based on this model, unless p is evaluated reasonably. This is the reason why present model is of phenomenological nature.

Rate Equations and Solutions. All number concentrations n's are normalized to give unity at any time when they all are summed up. Time course of the concentration of random coil species, $n(0|t)$, is governed by Eq. 1.

$$\mathrm{d}n(0|t)/\mathrm{d}t = -Lk_i n(0|t) = -k_0 n(0|t) \tag{1}$$

where k_0 is defined as Lk_i. Let $n_\eta(i|t)$ denotes the concentration of a species at time t with two β-strands of length i starting from the first bend between η- and (η+1)-th units. Since the chain is symmetric, only those species of $\eta \leqq L/2$ are considered explicitly. Final results will be doubled in due course.

Rate equations for $n_\eta(i|t)$'s are as follows. For $\eta \leqq L/2-1$,

$$\mathrm{d}n_\eta(1|t)/\mathrm{d}t = k_i n(0|t) - (k_+ + 2k_b)n_\eta(1|t) \tag{2}$$

$$\mathrm{d}n_\eta(i|t)/\mathrm{d}t = k_+ n_\eta(i-1|t) - (k_+ + 2k_b)n_\eta(i|t) \quad (2 \leqq i \leqq \eta-1) \tag{3}$$

$$\mathrm{d}n_\eta(\eta|t)/\mathrm{d}t = k_+ n_\eta(\eta-1|t) - k_b n_\eta(\eta|t) \tag{4}$$

$$\mathrm{d}n_{L/2}(L/2|t)/\mathrm{d}t = k_+ n_{L/2}(L/2-1|t) \quad (\eta = L/2) \tag{5}$$

For three strand species, $n_\eta(j_1,i|t)$ and $n_\eta(i,j_2|t)$ and four strand species, $n_\eta(j_1,i,j_2|t)$, their concentration changes are described by the following equations.

$$\mathrm{d}n_\eta(j_1,i|t)/\mathrm{d}t = k_b n_\eta(i|t) - k_b n_\eta(j_1,i|t) \tag{6}$$

$$\mathrm{d}n_\eta(i,j_2|t)/\mathrm{d}t = k_b n_\eta(i|t) - k_b' n_\eta(i,j_2|t) \tag{7}$$

$$\mathrm{d}n_\eta(j_1,i,j_2|t)/\mathrm{d}t = k_b n_\eta(j_1,i|t) + k_b n_\eta(i,j_2|t) \tag{8}$$

Here, j_1, j_2, and k_b' are defined as follows: j_1=Min(i, η–i), j_2=Min(i, $L-\eta-$i) and $k_b' = k_b (i \leqq \eta-1)$ or 0 ($i=\eta$).

Initial conditions at t=0 are: concentrations of all species except random coil are zero and $n(0|0)=1$. Solutions of rate equations consistent with the initial conditions are given below.

$$n(0|t) = \exp(-k_0 t) \tag{9}$$

$$n_\eta(i|t) = (k_i/k_+)(k_+/k)^i \exp(-k_0 t)\gamma(i,kt)/(i-1)! \tag{10}$$

$$n_\eta(\eta|t) = (k_+/k_b)n_\eta(j_1,i=\eta-1|t) \quad (n \leqq L/2-1) \tag{11}$$

$$n_{L/2}(L/2\,|\,t) = k_i(k_+/k)^{L/2-1}\left[\frac{(1-\exp(-k_0t))}{k_0} - \sum_{r=0}^{L/2-2}\frac{k^r}{(k_++2k_b)^{r+1}}\gamma[r+1,(k_++2k_b)t]/r!\right] \quad (12)$$

For $1\leqq i\leqq \eta$,

$$n_\eta(j_1,i\,|\,t) = k_i(k_+/k)^{i-1}(k_b/k)\left[\frac{(\exp(-k_0t)-\exp(-k_bt))}{k_b-k_0} - \frac{\exp(-k_bt)}{k_++k_b}\sum_{r=0}^{i-1}\left(\frac{k}{k_++k_b}\right)^r\gamma[r+1,(k_++k_b)t]/r!\right] \quad (13)$$

$$n_\eta(i,j_2\,|\,t) = n_\eta(j_1,i\,|\,t) \quad (i\leqq \eta-1) \quad (14a)$$

$$n_\eta(\eta,j_2\,|\,t) = (k_+/2k_b)n_\eta(j_1,i=\eta-1,j_2\,|\,t) \quad (14b)$$

$$n_\eta(j_1,i,j_2\,|\,t) = \frac{2k_ik_b^2}{k(k_0-k_b)}(k_+/k)^{i-1}\times \left[\frac{(1-\exp(-k_bt))}{k_b}\left(\frac{k}{k_++k_b}\right)^i - \frac{(1-\exp(-k_0t))}{k_0}\right] + \frac{2k_ik_b^2}{k(k_++k_b)}(k_+/k)^{i-1}\sum_{r=0}^{i-1}\left(\frac{k}{k_++k_b}\right)^r\times \left[\sum_{l=0}^{r}\frac{(k_++k_b)^l}{(k_++2k_b)^{l+1}}\gamma[l+1,(k_++2k_b)t]/l!\right] \quad (15)$$

In the above equations, k is defined as $k_++2k_b-k_0$ and $\gamma(i,x)$ is the incomplete gamma function defined below.

$$\gamma(i,x) = \int_0^x e^{-y}y^{i-1}dy = (i-1)![1-e^{-x}\sum_{r=0}^{i-1}\frac{x^r}{r!}] \quad (16)$$

From the concentrations given above, fraction f of the β-conformation or hydrogen-bonded units can be readily calculated. Contributions from two-, three-, and four-strand species are denoted as f_1, f_2, and f_3, respectively. Calculations and the results are given in Appendix.

The extent of reaction Δf is defined as follows.

$$\Delta f(t) = [f(\infty)-f(t)]/f(\infty) \quad (17)$$

Here, $f(t)=f_1+f_2+f_3$ and $f(\infty)$ represents the final amount at $t=\infty$. The final amounts are also given in Appendix. After substituting these results in Appendix into the above three terms, the extent of reaction is explicitly represented in the below as a function of time. In the case of the fraction of the β-conformation f_β,

$$\Delta f_\beta(t) = a_1\exp(-k_0t)+a_2\exp(-k_bt) + \left[\sum_{r=0}^{L/2-2}a_{3,r}\frac{(kt)^r}{r!}\right]\exp\{-(k_++2k_b)t\} + \left[\sum_{r=0}^{L/2-2}a_{4,r}\sum_{l=0}^{r}(k_++k_b)^lt^l/l!\right]\exp\{-(k_++2k_b)t\} + \left[\sum_{r=0}^{L/2-2}a_{5,r}\sum_{l=0}^{r}(k_++2k_b)^lt^l/l!\right]\exp\{-(k_++2k_b)t\} \quad (18)$$

Here, $a_{i,r}$'s are complex functions of k_b/k_+, k_i/k_+, and L.

We confine ourselves for the moment to the case $k_+\gg k_b$. Under this condition, strands can grow up to a considerable length before the next bend is formed. Since number of strands is limited to be smaller than five, fraction of the β-conformation will be very small in otherwise case $k_+<k_b$. Under the limiting case considered here, the last three terms of Eq. 18 become functions of k_+t and Eq. 18 is reduced to a simple form. For a small t ($k_+t\ll 1$), first two terms can be approximated to be constant and the last three terms are expanded in power series of k_+t. After truncation of higher order terms in k_+t, Eq. 19 results.

$$\Delta f_\beta(k_+t\ll 1) = 1 - (a_{3,0}-a_{3,1}-a_{4,0}-a_{5,0})k_+t + O(k_+^2t^2) \quad (19)$$

Here, the following obvious relation is used.

$$\Delta f_\beta(0) = a_1 + a_2 + a_{3,0} + \sum_{r=0}^{L/2-3}a_{4,r} + \sum_{r=0}^{L/2-2}a_{5,r} = 1 \quad (20)$$

Since,

$$a_{3,1} - a_{3,0} = -(1-2/L)2k_+k_b/k^2 \approx O(k_i/k_+)$$

and

$$a_{4,0} \approx a_{5,0} \approx O(k_ik_b/k_+^2),$$

the coefficient of k_+t term in Eq. 19 is much smaller than unity under the limiting case considered. This means the presence of a lag phase. However, whether or not a lag phase is significant depends on the ratio t_0/τ_1, where t_0 and τ_1 denote a lag time and the time constant of the reaction, respectively.

Results of Numerical Calculations. Numerical calculations were carried out on the following sets of rate constants for chains of $L=12$, 24, and 48. (For simplicity, L is taken to be a multiple of 12)

$k_b/k_+ = 1\times 10^{-1}$, 1×10^{-2}, and 1×10^{-3}

$k_b/k_i = 1$, 10, and 10^2.

One of the most complex behavior obtained is shown in Fig. 2 which was obtained for $L=24$, $k_b/k_+=$

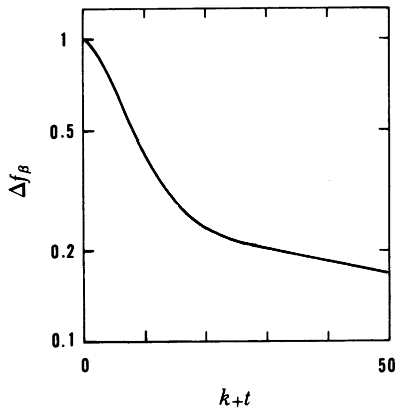

Fig. 2. A general kinetic pattern where log Δf_β is given as a function of reduced time k_+t. $L=24$, $k_b/k_+=k_i/k_+=1\times 10^{-2}$, $k_0/k_+=0.24$.

$k_i/k_+=1\times10^{-2}$. The time course is well approximated with the following formula of four parameters for a time range $t>t_0$.

$$\Delta f_\beta(t) = A_1 \exp[-(t-t_0)/\tau_1] + (1-A_1)\exp[-(t/\tau_2)] \tag{21}$$

In Table 1, values for the parameters are given for 26 cases. As seen from Table 1, many kinetics examined show simpler behavior than that shown in Fig. 2 or represented by Eq. 21. For about 9 examples, the time course can be approximated with single exponential term. About 6 examples show single exponential behavior accompanying a lag time t_0. Time course with two exponential terms without a lag time is found on 6 examples. A lag phase is discriminated when t_0/τ_1 exceeds 0.03.

At $k_b/k_+=1\times10^{-3}$ ($k_+\gg k_0$ or k_b), a general trend is seen which is valid irrespective of chain length. Two exponential character $0.7\exp(-k_0t)+0.3\exp(-k_bt)$ prevails if $k_0\gg k_b$ but single exponential dependence $\exp(-k_0t)$ is dominant if $k_b\gg k_0$. Examples of time course of various species are shown in Fig. 3 for a chain of $L=24$. In Fig. 3(a) where $k_0\gg k_b$ the contribution from two-strand species first reaches about 65% of the final fraction of the β-conformation, which is close to $A_1(0.73)$, with a rate constant close to k_0. The rest amount (35%) comes out as two-strand species are converted into three-strand species with a rate constant k_b. When k_0/k_b is reduced, as shown in Figs. 3(b) and (c), three-strand species are formed before two-strand species fully grow up, which is seen by a decrease of amplitude A_1 in Table 1. The rate of the formation of three-strand species is given by k_0 (which is now smaller than k_b).

This behavior resembles what is predicted from a following simple Scheme 1.

$$\text{Random coil} \xrightarrow{k_0} \text{Two-strand species} \xrightarrow{k_b} \text{Three-strand species}$$

Scheme 1.

At two other values of k_b/k_+, transition from two exponential character to single exponential one similarly takes place, as k_0/k_b is reduced. The following features are further noticed. In initiation-controlled regime ($k_b>k_0$), single exponential character prevails but the presence of a lag phase becomes significant, as k_+ approaches k_b, i.e., $k_b/k_+=1\times10^{-2}$ and 0.1. In bend-formation-controlled regime ($k_0>k_b$), on the other hand, two exponential character remains but the contribution of a lag phase becomes vanishing, as k_b/k_+ increases.

According to Scheme 1, a lag phase is expected if $k_0\sim k_b$. Some lag phases in Table 1 can be understood based on Scheme 1. However, lag phase in two exponential regime as well as in the initiation-controlled regime cannot be interpreted based on simple Scheme 1. In this case, recourse should be made to Eq. 18 to understand the origin and nature of the lag phase.

Table 1. Summary of Kinetic Patterns

$k_b/k_+=1\times10^{-3}$			
k_b/k_i	1	10	10^2
$L=12$	$0.71\ e^{-1.08\,k_0t}+0.29\ e^{-1.10\,k_bt}$	$e^{-0.94\,k_bt}$	$e^{-0.92\,k_0t}$
$L=24$	$0.73\ e^{-1.08\,k_0(t-t_0)}+0.21\ e^{-1.0\,k_bt}$ $(k_+t_0=4;\ 3\%)$	$0.24\ e^{-1.42\,k_0t}+0.76\ e^{-1.3\,k_bt}$	$e^{-0.96\,k_0t}$
$L=48$	$0.73\ e^{-0.96\,k_0(t-t_0)}+0.27\ e^{-1.3\,k_bt}$ $(k_+t_0=5;\ 5\%)$	$0.64\ e^{-1.08\,k_0t}+0.36\ e^{-1.1\,k_bt}$	$e^{-0.86\,k_0t}$
$k_b/k_+=1\times10^{-2}$			
k_b/k_i	1	10	10^2
$L=12$	$0.69\ e^{-1.08\,k_0t}+0.31\ e^{-1.1\,k_bt}$	$e^{-0.94\,k_bt}=e^{-0.78\,k_0t}$	$e^{-1.0\,k_0t}$
$L=24$	$0.73\ e^{-0.73\,k_0(t-t_0)}+0.27\ ^{-0.97\,k_bt}$ $(k_+t_0=1.3;\ 6\%)$	$e^{-0.71\,k_0t}=e^{-1.7\,k_bt}$	$e^{-0.96\,k_0t}$
$L=48$	$0.71\ e^{-0.29\,k_0(t-t_0)}+0.29\ e^{-1.0\,k_bt}$ $(k_+t_0=1;\ 3\%)$	$0.52\ e^{-1.15\,k_0(t-t_0)}+0.48\ e^{-1.2\,k_bt}$ $(k_+t_0=5;\ 6\%)$	$e^{-0.86\,k_0t}$
$k_b/k_+=1\times10^{-1}$			
k_b/k_i	1	10	10^2
$L=12$	$0.73\ e^{-0.43\,k_0t}+0.28\ e^{-0.87\,k_bt}$	$e^{-0.92\,k_b(t-t_0)}=e^{-0.77\,k_0(t-t_0)}$ $(k_+t_0=1.3;\ 5\%)$	$e^{-1.0\,k_0(t-t_0)}$ $(k_+t_0=6;\ 3\%)$
$L=24$	$0.54\ e^{-0.17\,k_0t}+0.46\ e^{-1.20\,k_bt}$	$e^{-1.3\,k_b(t-t_0)}$ $(k_+t_0=1.2;\ 7\%)$	$e^{-0.96\,k_0(t-t_0)})$ $(k_+t_0=4;\ 4\%)$
$L=48$		$e^{-1.3\,k_b(t-t)}$ $(k_+t_0=0.6;\ 3\%)$	$e^{-0.92\,k_0(t-t_0)}$ $(k_+t_0=5.5;\ 10\%)$

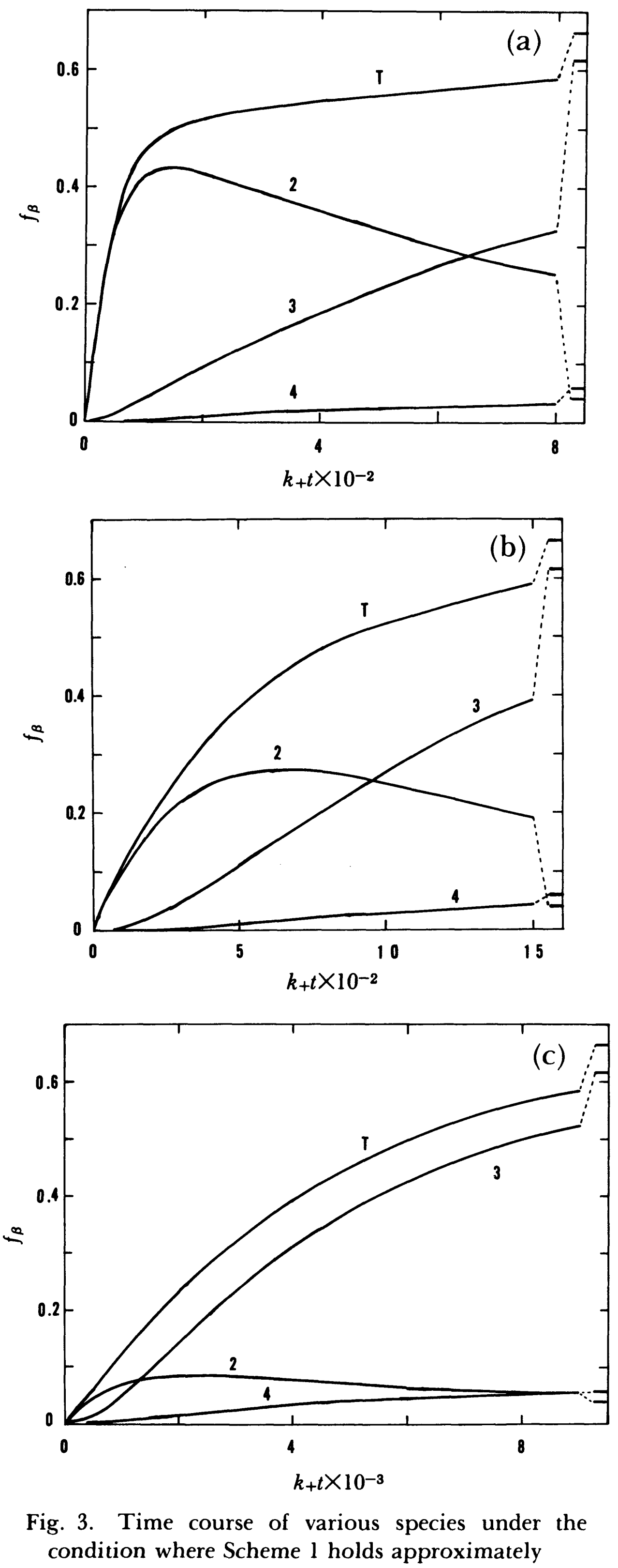

Fig. 3. Time course of various species under the condition where Scheme 1 holds approximately L=24, $k_b/k_+=1\times10^{-3}$. k_i/k_b: (a) 1, (b) 0.10, and (c) 1×10^{-2}.
In each figure (a)—(c), curves labelled as 2, 3, 4, and T represent the contributions from two- three-, four-strand species and their sum. Final values at $t=\infty$ are indicated by horizontal bars.

It is to be noted in Fig. 3 that final products are mostly three- strand species (β-meander). The population of four-strand species is significant under the situation $k_b/k_+=0.1$ in the case of L=24.

Experimental data on homopolypeptides published to date are systematically interpreted in terms of the present model. Their kinetics are not different from each other in essence but they all are explicable with the present model. The time constants obtained are expected to be related not to k_+ but rather to k_0 in the case of PLL[6–8] or to k_b and k_0 in the case of PLT.[10]

Recently, a progress has been made on the equilibrium theory of coil-β-structure conversion.[14–17] However, no attempt was made to evaluate various parameters from the comparison with experimental data. This indicates the difficult situation in the study of the β-structure to obtain proper experimental results consistent with the requirements of theories. The situation is worse in kinetic aspects than equilibrium ones. Hence, in the present study, no attempt is made to evaluate various rate constants from the comparison with published kinetic data.

Discussion

Some assumptions introduced in the present study are discussed briefly here.

(1) **Irreversibility.** It is desirable to take the reverse reaction into account. However, most experimental data were obtained under the condition that β-structure was nearly completely formed. Reverse reaction can be neglected under this condition. Furthermore, even under the condition where the fraction of the β-conformation is appreciably smaller than unity, there are several indications suggesting two processes; the first fast process followed by a slow annealing process.[12] It is likely that the fast growth step occurs irreversibly in a practical sense. Reversible nature of the reaction is partly taken into account even in the present model, for example, in the definition of unit or in the following discussion on (5).

(2) **Initiation through Bend Formation.** Formation of the β-structure was much slower in buffers than in unbuffered solutions under otherwise identical conditions in the case of PLL[7] and poly[Cys-(CH_2COOH)] (unpublished result). The findings indicate that collision of two random coil chains is much less effective than bend formation to initiate the β-structure. Since reaction between two random coil portions within a chain does not differ much from that between two random coil chains, the assumption is pertinent.

(3) **Multiple Initiation is not Considered.** For long chains, the fraction of the β-conformation reaches a limit, which is far below completion, even the limitation on the number of strands is removed. Therefore, initiation occurs at a unit outside of the

first β-sheet region and hence multiple initiation is a generally expected event. When all these locally formed β-sheets encounter each other, the first irreversible growth step terminates. Then, second annealing step follows where reversible nature of the reaction plays an essential role. The net increase in the fraction of the β-conformation is not big, however, accompanying this second step, although stability of the β-structure increases. Hence, the second step can be tentatively ignored. Then, present model can be applied to the case of multiple initiation, only by modifying the original chain length L_0 to L, the latter being roughly given by $L_0 f_\beta(\infty; L_0)$.

(4) **Calculations are Restricted under the Condition $k_+ > k_b$.** In the reverse situation of $k_b > k_+$, major products will be such that chains consisting of many bends connected by short strands. We are not interested in these species in the present study. Further, to examine the kinetics of the formation of such species, limitation on the number of strands (assumption 6) should be removed.

(5) **Propagation Rate.** Propagation rate of the first two strands is taken to be much smaller than that of the following strands. This partly takes reversible nature of strand growth into account in the present model. Since the first two strands are not as stable as the following ones, their propagation step proceeds in a more stochastic way.

(6) **Comparison with Helix-Coil Conversion.** Present analysis suggests that information about k_+ is not obtained from published experimental data. Consequently, it is misleading to compare slow formation rate of the β-structure with results on the fast propagation rate of helix formation.[18]

Since β-bend type III is equivalent to 3_{10} helix,[19] elementary processes in the α-helix initiation and bend formation are expected to be rather similar. However, there is at least one difference. For strand growth to occur significantly in the case of the β-structure, sufficient length will be required for both sides of an initiation site. Therefore, a given β-bend formation is not always valid for subsequent strand growth. For a helix to develop, on the other hand, it is required that either side of initiation site has a sufficient length. This different situation certainly makes considerable difference with respect to effective initiation rate constants between the α-helix and the intramolecular β-structure.

Appendix

Calculations of the Fraction of the β-Conformation. From the expressions on the concentrations given in the text, fraction f of the β-conformation or hydrogen-bonded units can be readily calculated.

The contribution from two-strand species f_1 is explicitly written as follows.

$$f_1 = f_{11} + f_{12} + f_{13} \tag{A1}$$

$$\begin{aligned} f_{11} &= (2m_1/L) \sum_{i=1}^{L/2-1} \sum_{\eta=i+1}^{L/2} i n_\eta(i|t) \\ &= (2m_1/L) \sum_{i=1}^{L/2-1} i(L/2-i) n_\eta(i|t) \end{aligned} \tag{A2}$$

$$f_{12} = (2m_1/L) \sum_{i=1}^{L/2-1} i\, n_i(i|t) \tag{A3}$$

$$f_{13} = (m_1/2) n_{L/2}(L/2|t) \tag{A4}$$

Similarly, contributions from three-strand species f_2 and four-strand species f_3 are given as follows.

$$f_2 = f_{21} + f_{22} + f_{23} \tag{A5}$$

$$\begin{aligned} f_{21} &= (2m_2/L)[\sum_{i=1}^{L/4} \sum_{\eta=2i}^{L/2} i\, n_\eta(j_1=i,i|t) \\ &\quad + \sum_{i=1}^{L/4} \sum_{\eta=i+1}^{L/2} i\, n_\eta(i,j_2=i|t) \\ &\quad + \sum_{i=L/4+1}^{L/3-1} \sum_{\eta=i+1}^{L-2i} i\, n_\eta(i,j_2=i|t)] \\ &= (2m_2/L)[\sum_{i=1}^{L/4} i(L/2-2i+1) n_\eta(j_1=i,i|t) \\ &\quad + \sum_{i=1}^{L/4} i(L/2-i) n_\eta(i,j_2=i|t) \\ &\quad + \sum_{i=L/4+1}^{L/3-1} i(L-3i) n_\eta(i,j_2=i|t)] \end{aligned} \tag{A6}$$

$$\begin{aligned} f_{22} &= (2/L)[\sum_{i=2}^{L/4} \sum_{\eta=i+1}^{2i-1} (\eta+m_3 i) n_\eta(j_1=\eta-i,i|t) \\ &\quad + \sum_{i=L/4+1}^{L/2-1} \sum_{\eta=i+1}^{L/2} (\eta+m_3 i) n_\eta(j_1=\eta-i,i|t) \\ &\quad + \sum_{i=L/4+1}^{L/3} \sum_{\eta=L-2i+1}^{L/2} (L-\eta+m_3 i) n_\eta(i,j_2=L-\eta-i|t) \\ &\quad + \sum_{i=L/3+1}^{L/2-1} \sum_{\eta=i+1}^{L/2} (L-\eta+m_3 i) n_\eta(i,j_2=L-\eta-i|t)] \\ &= (2/L)[\sum_{i=2}^{L/4} i(i-1)(3/2+m_3) n_\eta(j_1=\eta-i,i|t) \\ &\quad + \sum_{i=L/4+1}^{L/2-1} (L/2-i)\left(m_3 i + \frac{L+2i+2}{4}\right) n_\eta(j_1=\eta-i,i|t) \\ &\quad + \sum_{i=L/4+1}^{L/3} (i-L/4)(2i+L/2+2m_3 i-1) \\ &\qquad\qquad n_\eta(i,j_2=L-i-\eta|t) \\ &\quad + \sum_{i=L/3+1}^{L/2-1} \left(\frac{L-2i}{4}\right)(2m_2 i-1-i+3L/2) \\ &\qquad\qquad n_\eta(i,j_2=L-i-\eta|t)] \end{aligned} \tag{A7}$$

$$\begin{aligned} f_{23} &= (2/L)[\sum_{i=1}^{L/3} m_2 i\, n_i(i,j_2=i|t) \\ &\quad + \sum_{i=L/3+1}^{L/2-1} (L-i+m_3 i)\, n_i(i,j_2=i|t)] \end{aligned} \tag{A8}$$

$$f_3 = f_{31} + f_{32} + f_{33} \tag{A9}$$

$$\begin{aligned} f_{31} &= (2m_4/L) \sum_{i=1}^{L/4} \sum_{\eta=2i}^{L/2} i n_\eta(j_1=i,i,j_2=i|t) \\ &= (2m_4/L) \sum_{i=1}^{L/4} (L/2-2i+1) i\, n_\eta(j_1=i,i,j_2=i|t) \end{aligned} \tag{A10}$$

$$f_{32} = (2/L)[\sum_{i=2}^{L/4} \sum_{\eta=i+1}^{2i-1} (\eta+m_1 i)\, n_\eta(j_1=\eta-i,i,j_2=i|t)$$

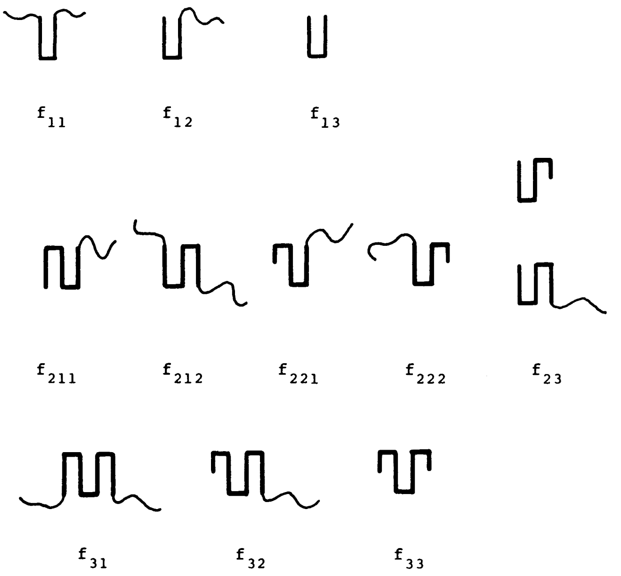

Fig. 4. Schematic representation of various species corresponding to respective contributions given in Eqs. A1—A12. Thick lines and thin curves represent the β-strands and random coil portions, respectively.

$$\begin{aligned} &+\sum_{i=L/4+1}^{L/3}\sum_{\eta=i+1}^{L-2i}(\eta+m_1 i)\,n_\eta(j_1=\eta-i,i,j_2=i\,|\,t) \\ &=(2/L)[\sum_{i=2}^{L/4}i(i-1)(3/2+m_1)\,n_\eta(j_1=\eta-i,i,j_2=i\,|\,t) \\ &\quad+\sum_{i=L/4+1}^{L/3}(L+2m_1 i-i+1)\left(\frac{L-3i}{2}\right) \\ &\qquad n_\eta(j_1=\eta-i,i,j_2=i\,|\,t)] \end{aligned} \tag{A11}$$

$$\begin{aligned} f_{33}&=(2/L)[\sum_{i=L/4+1}^{L/3}\sum_{\eta=L-2i+1}^{L/2}(L-i+m_3 i) \\ &\qquad n_\eta(j_1=\eta-i,i,j_2=L-\eta-i\,|\,t) \\ &\quad+\sum_{i=L/3+1}^{L/2-1}\sum_{\eta=i+1}^{L/2}(L-i+m_3 i) \\ &\qquad n_\eta(j_1=\eta-i,i,j_2=L-\eta-i\,|\,t)] \\ &=(2/L)[\sum_{i=L/4+1}^{L/3}(2i-L/2)(L-i+m_3 i) \\ &\qquad n_\eta(j_1=\eta-i,i,j_2=L-\eta-i\,|\,t) \\ &\quad+\sum_{i=L/3+1}^{L/2-1}(L+m_3 i-i)(L/2-i) \\ &\qquad n_\eta(j_1=\eta-i,i,j_2=L-\eta-i\,|\,t)] \end{aligned} \tag{A12}$$

In the above equations except Eq. A4, a factor of 2 appears in order to take species with η greater than $L/2$ into account. In the above, (m_1, m_2, m_3, m_4) is (2, 3, 1, 4) for f_β (fraction of the β-conformation) or (1, 2, 0, 3) for f_{hb} (fraction of hydrogen bonded units). In Fig. 4, various species corresponding to each f_{ij} term are schematically represented.

Final Amounts of the β-Conformation or Hydrogen-Bonded Units. As the reaction proceeds, f reaches the final value $f(\infty)$ at $t=\infty$.

$$f(\infty)=f_{13}(\infty)+f_{23}(\infty)+f_{31}(\infty)+f_{32}(\infty)+f_{33}(\infty) \tag{A13}$$

These final values are functions of k_i/k_0 and k_b/k_+. Since $k_i/k_0=1/L$, the final values are solely determined by the ratio k_+/k_b for a given chain length L. Introducing $\xi=k_+/(k_++2k_b)$, they are written as follows.

$$\begin{aligned} f_{13}(\infty)&=(m_1k_i/2k_0)\xi^{L/2-1} \\ f_{23}(\infty)&=(2k_i/Lk_0)[m_2\sum_{i=1}^{L/3}i\xi^{i-1}+\sum_{i=L/3+1}^{L/2-1}(L-i+m_3 i)\xi^{i-1}] \\ f_{31}(\infty)&=\frac{4m_4k_ik_b}{Lk_0k_+}\sum_{i=1}^{L/4}i(L/2-2i+1)\xi^i \\ f_{32}(\infty)&=\frac{4k_ik_b}{Lk_0k_+}[(3/2+m_1)\sum_{i=2}^{L/4}(i^2-i)\xi^i \\ &\quad+(1/2)\sum_{i=L/4+1}^{L/3}(L-3i)(L-i+2m_1+1)\xi^i] \\ f_{33}(\infty)&=\frac{4k_ik_b}{Lk_0k_+}[\sum_{i=L/4+1}^{L/3}(2i-L/2)(L-i+m_3 i)\xi^i \\ &\quad+\sum_{i=L/3+1}^{L/2-1}(L/2-i)(L-i+m_3 i)\xi^i] \end{aligned} \tag{A14}$$

References

1) Ch. Ghelis and J. Yon, "Protein Folding," Academic Press, New York (1982).

2) P. S. Kim and R. L. Baldwin, *Ann. Rev. Biochem.*, **51**, 459 (1982).

3) Th. E. Creighton, "Proteins," W. H. Freeman & Co., New York (1983).

4) B. Davidson and G. D. Fasman, *Biochemistry*, **6**, 1616 (1967).

5) S.-Y. C. Wooley and G. Holzwarth, *Biochemistry*, **9**, 3604 (1970).

6) C. R. Snell and G. D. Fasman, *Biochemistry*, **12**, 1017 (1973).

7) R. Hartman, R. C. Schwaner, and J. Hermans, Jr., *J. Mol. Biol.*, **90**, 415 (1974).

8) H.-A. Arfmann, R. Labitzke, and K. G. Wagner, *Biopolymers*, **16**, 1815 (1977).

9) H. Yamamoto and J. T. Yang, *Biopolymers*, **13**, 1109 (1974).

10) H. E. Auer and E. Patton, *Biophys. Chem.*, **4**, 15 (1976).

11) H. E. Auer and H. Miller-Auer, *Biopolymers*, **25**, 1607 (1986).

12) K. Fukada, H. Maeda, and S. Ikeda, *Polymers*, **28**, 1887 (1987).

13) W. L. Mattice, *Macromolecules*, **17**, 407 (1984).

14) W. L. Mattice and H. A. Scheraga, *Biopolymers*, **23**, 1701 (1984).

15) W. L. Mattice and H. A. Scheraga, *Macromolecules*, **17**, 2690 (1984).

16) W. L. Mattice and H. A. Scheraga, *Biopolymers*, **24**, 565 (1985).

17) W. L. Mattice, E. Lee, and H. A. Scheraga, *Can. J. Chem.*, **63**, 140 (1985).

18) J. Janin and S. J. Wodak, *Prog. Biophys. Mol. Biol.*, **42**, 21 (1983).

19) C. M. Venkatachalam, *Biochim. Biophys. Acta*, **303**, 211 (1973).