PROTEINS: Structure, Function, and Genetics 9:23–27 (1991)

Rate of β-Structure Formation in Polypeptides

A. V. Finkelstein
Institute of Protein Research, Academy of Sciences of the USSR, Pushchino, Moscow Region, USSR

ABSTRACT **An explanation is suggested for why a marginally stable β-structure folds extremely slowly; it is predicted that even a small increase in stability drastically accelerates β-folding. According to the theory, this folding is a first-order phase transition, and the rate-limiting step is nucleation. The rate-determining "nucleus" (transition state) is the smallest β-sheet that is sufficiently large to provide an overall free energy reduction during subsequent folding. If the stability of the β-structure is low, the nucleus is large and possesses a high free energy due to having a large perimeter. When the net stability of the final β-structure increases (due to either an increase of the β-sheet stability or a decrease in stability of the competing structures, e.g., α-helices), the size and energy of a nucleus decrease and the rate of folding increases exponentially. This must result in a fast folding of polypeptides enriched by β-forming residues (e.g., protein chains). The theory is developed for intramolecular β-structure, but it can also explain the overall features of intermolecular β-folding; it is applicable both to antiparallel and parallel β-sheets. The difference in folding of β-sheets, α-helices, and proteins is discussed.**

Key words: polypeptides, protein folding, β-structure, free energy barrier, nucleation of folding, rate-limiting step

INTRODUCTION

Two puzzles are connected with the rate of β-folding. First, β-structure (both intra- and intermolecular[1] may form extremely slowly in synthetic polypeptides. It may take minutes or hours,[1–6] while an α-helix folds within microseconds.[7,8] Second, changes in conditions may increase the rate of β-folding in polypeptides so that it occurs within milliseconds,[5,6] which is already close to the rate of β-folding in protein chains[9] and to the folding of α-helices. The estimated activation energy of β-folding (≈60 kcal/mol)[2] is many-fold higher than that of α-folding (≈10 kcal/mol).[7]

This difference in the rate-limiting steps of β- and α-folding requires an explanation, since the forces stabilizing these structures are similar. It is also of interest to understand whether or not the rate-limiting step of β-folding in polypeptides is similar to the rate-limiting step in protein self-organization.

I shall consider the folding of intramolecular β-structure in a homopolypeptide, even though aggregation occurs in the majority of experiments. With a long polypeptide, however, intermolecular β-structure is formed only slightly faster than intramolecular β-structure: the former requires 100 sec, the latter 500 sec under the same conditions.[1] The reason for the 100 to 500 sec difference in folding is outside the scope of this paper. This paper addresses the question of why β-folding may take minutes while α-folding takes microseconds.

"COIL"-β PHASE TRANSITION

It is useful to recollect some thermodynamic results before studying the kinetics.

A β-coil transition is rather different from a helix-coil transition. The latter[10] is *never* a first-order phase transition because it is a transition between two one-dimensional states.[11] This means that α-helices arise and melt independently if they are remote from each other in the chain. A β-sheet is a two-dimensional body, illustrated schematically in Figure 1. It has been shown[12,13] that the "coil"-β transition is a first-order phase transition (i.e., an "all-or-none" conversion of a *whole* molecule), *provided*

$$\exp(-F_t/RT) \cdot (2RT/\Delta F_e) < 1 \qquad (1)$$

Here F_t is the free energy of a "turn," the totality of all possible bends and loops connecting the adjacent strands of a sheet (Fig. 1), ΔF_e is the free energy loss for a residue transferred from the inside to the edge of a sheet, R is the gas constant, and T the temperature. The first term of the product in Eq. (1) can be interpreted (according to the model used[12–15]) as a partition function of all kinds of bends and loops connecting the adjacent strands. The second term of this product can be interpreted (according to the same model) as $\int \exp(-|n|\,\Delta F_e/RT)dn$, the partition function of the intrahairpin edge effects, n being the difference in the lengths of the adjacent strands. Al-

Received September 4, 1989; revision accepted June 5, 1990.
Address reprint requests to Dr. A.V. Finkelstein, Institute of Protein Research, Academy of Sciences of the USSR, Pushchino, Moscow Region, USSR.

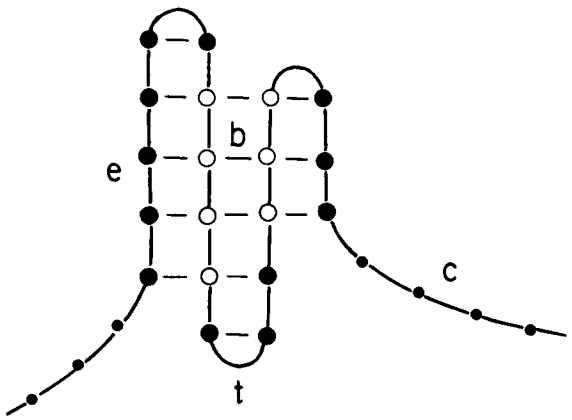

Fig. 1. β-Sheet in a polypeptide. b (open circle)—internal residue of a sheet with the free energy F_b; e (closed circle)— edge residue, its free energy is $F_b + \Delta F_e$; t—"turn," the totality of all possible bends and loops connecting the strands in a sheet; c (point)—the "coil" residue (see the text), its free energy is F_c. The net stability of β-structure per residue is determined by $\Delta F_\beta = F_b - F_c$.

together this means that each hairpin, despite fluctuations, gives a positive contribution

$$\tilde{F}_t \equiv F_t - RT \cdot \ln(2RT/\Delta F_e) > 0 \qquad (2)$$

to the free energy of a sheet boundary. In other words, a first-order phase transition will occur if the total free energy of sheet boundary is positive despite all the boundary fluctuations. Slow kinetics is typical of first-order phase transitions if the approaching phase is at the border of stability.[16] The reason for this slowness is simply an unfavorable free energy at the boundaries of the nucleus.

If $\tilde{F}_t < 0$, a first-order transition is impossible.[13] In this case the equilibrium fraction of the β-state will change gradually, in a manner similar to the helix–coil transition. Such a case was investigated.[17] There is no reason why this transition must be slow compared to the elementary acts of β-turn formation and strand elongation (cf. 18).

A first-order phase transition is the key to the peculiarities of β-folding. This is a transition between two phases.[13] One phase is a huge regular β-sheet, the other a "coil" that contains not only a mixture of irregular and extended sections, but also some [small—under condition (1)] admixture of fluctuating hairpins and sheets. It is convenient to take the free energy of this mixture to be zero. The "coil" (more precisely, the "non-β-sheet") phase may also include α-helices (cf. 13). In this case, according to the Zimm–Bragg theory,[10] the free energy of a residue in this phase (F_c) is less than zero due to an admixture of helices with a "pure" coil. The "coil" →"β-sheet" transition occurs when $\Delta F_\beta \equiv F_b - F_c \leq 0$, i.e., the free energy of a sheet residue F_b is less than F_c, the free energy of a residue in the "non-β-sheet" phase (see legend to Fig. 1).

THE RATE-LIMITING STEP OF "COIL"-β PHASE TRANSITION

The peculiarities of β-folding cannot be attributed to chain diffusion, which takes microseconds (see Table 5.3 in 19); instead, they imply an unusually low stability of the transition state.

Folding pathways may go through different intermediates. The most unstable intermediate (transition state) will limit the rate of a given path. Our first aim is to understand why the stability of at least one intermediate of *any* β-sheet pathway is many-fold less than that of a turn or of a residue in a β-strand.

The number N of residues in the β-sheet grows gradually in the course of folding. It is easy to show that *any* sheet of a definite size N^* has a relatively high free energy. Consider a sheet that has N residues in K β-strands, m residues being at the edges (in Fig. 1, $N = 19$, $K = 4$, $m = 12$), and the "turns" are fluctuating. The free energy of formation of this sheet from a "coil" is

$$N \cdot \Delta F_\beta + (K-1) \cdot F_t + m \cdot \Delta F_e$$

A β-sheet with minimal free energy contribution from edges will have $m \cong 2N/K$. The fluctuations of shape of a sheet are not essential under condition (1). Optimizing the number of strands under the condition $m = 2N/K$, we estimate (with a negligible error of $\sim F_t + \sim \Delta F_e$) the minimal possible free energy of a sheet of N residues:

$$F(N) \approx N \cdot \Delta F_\beta - F_t + 2(N \cdot F_t \cdot 2\Delta F_e)^{1/2} \qquad (3)$$

The "coil"–β phase transition occurs spontaneously only if $\Delta F_\beta < 0$ (though some amount of the unstable fluctuating β-structure is always present in "coil," cf. 13). However, even if $\Delta F_\beta < 0$, the minimal energy $F(N)$ first grows with N, reaches the maximum

$$F_{max} \approx -F_t + 2(F_t \cdot \Delta F_e)/(-\Delta F_\beta) \approx 2(F_t \cdot \Delta F_e)/(-\Delta F_\beta) \qquad (4)$$

at $N^* = 2F_t \Delta F_e (\Delta F_\beta)^{-2}$ and only then falls (Fig. 2). *No* pathway of β-folding may avoid such a barrier. The activation free energy (a minimal estimate of which is given by F_{max}) may be much higher than the free energy F_t of β-hairpin initiation if

$$\Delta F_\beta + \Delta F_e > 0 \qquad (5)$$

i.e., if the β-hairpin stability decreases with increasing elongation.

A possible pathway of β-folding has been examined.[20] Folding can be initiated by a nucleus that looks like a β-hairpin with a subsequent turn (Fig. 3). As this hairpin must be the most unstable intermediate, its minimal length n^* is determined by the condition that the free energy gain from ad-

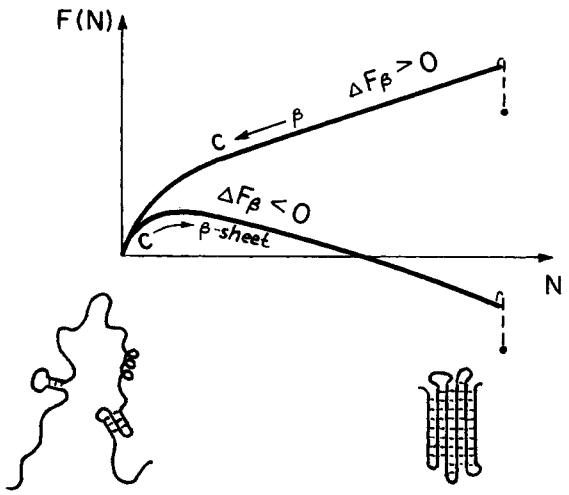

Fig. 2. β-Sheet free energy F as a function of N, the number of residues in the sheet. The maximum (at $\Delta F_\beta<0$) corresponds to a transition state and determines a rate-limiting step of β-sheet folding. This maximum divides two stable phases: the "coil" (C), which may include (along with irregular sections and α-helices) some admixture of small fluctuating β-sheets, and the continuous "β-sheet" phase. A rise and a subsequent slump at the right of the curves (dashed line) reflect possible gluing of the β-sheet edges.

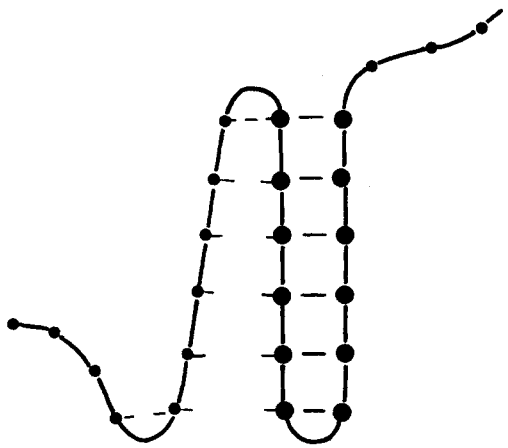

Fig. 3. Transition state for β-folding. The rate-limiting step is formation of a "nucleus" (β-hairpin ending with a turn) that is long enough to provide a net free energy decrease during the subsequent folding, an elementary step of which is the addition of a new β-strand with a new turn.

dition of the next strand must compensate for the free energy loss for the addition of the next turn:

$$n^* \approx F_t/(-\Delta F_\beta) \tag{6}$$

The minimal free energy of this hairpin nucleus is

$$F^* \approx F_t + 2n^*(\Delta F_\beta + \Delta F_e) + F_t = 2(F_t \cdot \Delta F_e)/(-\Delta F_\beta) \tag{7}$$

and coincides in the main term with the above calculated F_{max}. A hairpin shorter than n^* cannot correspond to a transition state: the addition of the next strand and turn (see Fig. 3) increases its free energy. A hairpin longer than n^* could serve as a transition state, but it would need a longer time to arise.

The value of F^* tends to infinity as ΔF_β tends to zero.

The possibility of different folding pathways may be taken into account[20] by some change of F_t that resembles the change of F_t to $\tilde{F}_t$ in Eqs. (1) and (2). This numerical change does not alter the overall conclusion: a first-order "coil"→β phase transition must be extremely slow (cf. 16) *if* the β-structure is not very stable, and the rate of the transition will depend drastically on the net stability of the β-state.

It would be expected that the theory is precise when $-\Delta F_\beta$ is small and we can readily neglect the items F_t and ΔF_e as compared to F^*.

The characteristic rate of β-nucleus formation at a given chain point is $(t_o)^{-1}\exp(-F^*/RT)$, where $t_o \sim 10^{-8}$ sec is the time of secondary structure elongation by one residue.[7,8] A time

$$t_{in} \sim t_o \exp(F^*/RT)\cdot M^{-1} \tag{8}$$

is necessary for initiation of a sheet in a chain of length M (a nucleus may form at any point). A time

$$t_{el} \sim t_o \cdot M \tag{9}$$

is then necessary to involve all M residues of the chain in the sheet. It can be readily shown that the time of β-folding, $t_\beta = t_{in} + t_{el}$, is determined by nucleation, given $M < M_1 \equiv \exp(F^*/2RT)$. If $M > M_1$, then about M/M_1 nuclei arise and grow in a time of $\sim 2t_o \cdot M_1$. This is surely not the case for slow folding polypeptides.

Comparison of Eq. (8) with experiment gives $F^* \approx 20$ kcal/mol for uncharged polylysine[1,2] where $M \sim 10^3$ and $t_\beta \sim 50$ or 500 sec at $T = 328$ K. Taking into account that $\Delta F_\beta \approx -0.1$ kcal/mol at the same temperature,[21] we obtain

$$F_t \cdot \Delta F_e \approx 1.0 \text{ (kcal/mol)}^2 \approx 2.5\ (RT)^2 \tag{10}$$

Substituting (10) into (1) we can prove that condition (1) for a first-order phase transition is fulfilled for the investigated polypeptide even without knowing the values of F_t and ΔF_e. If we assume F_t to be ≈ 4 kcal/mol, as for the initiation of an α-helix[22] (a β-turn resembles a helix turn), we can even estimate the edge free energy: $\Delta F_e \approx 0.25$ kcal/mol. This estimate is useful for computation of β-sheet stability.[23]

DISCUSSION

Equations (7) and (8) show that the time of β-folding may be enormously long if net stability of β-structure is low, and it may be very short if the stability is high (Fig. 4).

The net stability is determined by the free energy difference between the β and "coil" phases: $\Delta F_\beta =$

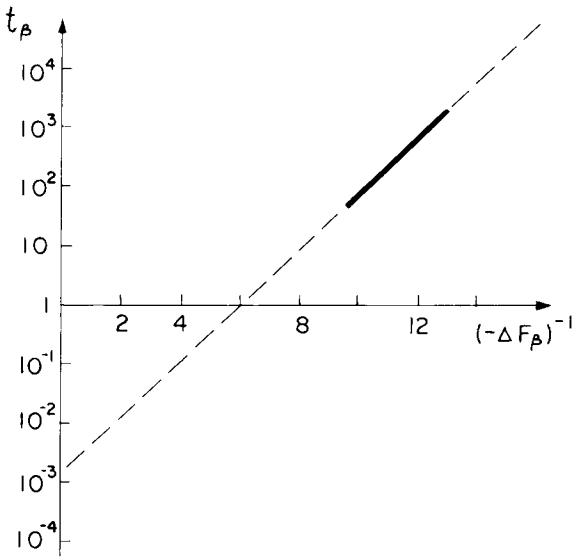

Fig. 4. Time for β-structure formation t_β (sec) vs. the reciprocal of the net stability $(-\Delta F_\beta$, kcal/mol)$^{-1}$ of the final β-structure. The dashed line is a linear extrapolation of the thick line, which corresponds to the experimental data.[2,21,22] The change of $\Delta F_\beta = F_b - F_c$ in a given case is due mainly to the change of F_c, which, in turn (see the text), reflects the melting of α-helices with temperature[22] in the "coil" (non-β-sheet) phase.

$F_b - F_c$. The "coil" may include α-helices and any other structures, except for large regular β-sheets (cf. 13). It is worthwhile to consider this case because α→β transitions are rather common. If the helices are unstable, then F_c is close to zero. If the helices are stable (free energy of helix elongation $F\alpha \leq 0$), then $F_c \approx F_\alpha < 0$ (cf. 10). This decreases $-\Delta F_\beta = F_c - F_b$ even if the free energy difference F_b between β and irregular states remains the same. Thus β-folding must be hindered by the competing α-helices that arise in a "coil," as is observed.[21] In particular, when the concentration of Leu is increased in copolymers, the rate of β-folding does not change initially, but then decreases[2,24] in a reciprocal correlation with the increasing stability of α-helices.[24] A similar effect seems to be caused by the emerging globules in poly-Tyr.[4]

The rate of β-folding must be comparatively low when $\Delta F_\beta + \Delta F_e > 0$. But when $\Delta F_\beta + \Delta F_e < 0$ and elongation stabilizes the β-hairpin, the rate is limited only by an initiating β-turn; then β-folding must be fast. This may take place, for example, in polypeptides with Ile and Val residues, where the stability parameter $-\Delta F_\beta$ seems to be as large as 0.6 or 0.8 kcal/mol.[26,27] It should be noted that the β-structure of globular proteins is enriched by just these residues[28]; thus, *β-folding in proteins can be relatively fast.*

The above results concern folding both of antiparallel and parallel β-sheets. The formulas are essentially the same for both kinds of structure; the parameters differ mainly in the value of the term F_t ($F_t^{par} > F_t^{antipar}$, because only long loops are appropriate for a parallel β), while ΔF_β and ΔF_e may be more or less close in value.

In accord with the theory, β-folding is indeed observed[2,3,6] to be accelerated in both cases when the β-structure stability is increased, i.e., either the temperature is raised or the polymer is discharged.

It has also been observed[6] that the rate of folding is independent of the amount of "preexisting" β-structure. This was treated by Fukada et al.[6] as a discrepancy with my statement[20] that the rate-limiting step is β-nucleation. However, the independence observed is not a contradiction. This independence follows directly from the theory[20] which (1) assumes β-folding to be an intramolecular process, and (2) shows that each chain awaits β-nucleation and then folds immediately and completely [see Eqs. (7)–(9): they contain no terms that depend on the amount of preexisting β-structure]. In other words, because the β-nuclei are *unstable*, all the "preexisting" β-structure must belong only to the completely folded chains and cannot help the folding of others. Even though aggregation of polypeptides occurs in experiments, the folded sheets are not able to nucleate further β-folding.[6] This experimental result supports the assumption (see Introduction) that the rate-limiting step of β-folding is due mainly to effects that are not changed grossly by aggregation.

The β-structure seems to arise through an "all-or-none" phase transition, at least in the case of slow β-folding when $F^* >> F_t$. Single-exponential kinetics were observed[6] for a very slow folding ($t_\beta \sim 10^3$–10^5sec at pH>4.7). However, a freshly folded sheet must be far from equilibrium. The loops are not tight and fixed; the strand length is $\sim n^* \sim F_t/(-\Delta F_\beta)$, while the equilibrium length must be $\approx (M \cdot F_t / 2\Delta F_e)^{1/2}$. Some rearrangement of the folded sheets must follow. Possibly, this is an ultimate, very slow phase (or phases) of folding, as observed at pH<4.6.[6] This process may also include an inter- or intramolecular gluing of sheet edges (see right part of Fig. 2): note that β-unfolding is also slow,[2,6] which implies that the "free" sheet edges (where unfolding may start immediately, see line $\Delta F_\beta > 0$ in Fig. 2) are absent. A discussion of all these intriguing events is outside the scope of this paper. I will also not elaborate upon the fast folding of some amount of the short (with a length $< n^*$) fluctuating hairpins within the "coil" phase (see left part of Fig. 2); their folding must resemble the folding of unstable α-helices. A very fast "pretransition" phase was observed at pH<4.7.[6]

CONCLUSION

The rate of β-structure folding depends on the net stability of the β-structure to be formed. It may be very fast or extremely slow. This is a consequence of an "all-or-none" transition between the "coil" and

β-sheet phases. This transition is very different from fast folding of α-helices, but the "all-or-none" type of β formation resembles protein folding.[29,30] However, the rate-limiting step of β-folding corresponds to a sheet *nucleation*. In proteins, on the contrary, the rate is limited by the *late step* of folding.[30] Both folding and unfolding of proteins go through the same (contrary to β-sheet) transition state, which corresponds to a somewhat distorted native globule. The rate-limiting event in proteins (contrary to β-sheets) seems to be side chain rather than backbone ordering.[31,32] Therefore, despite some similarity in the order of phase transition, β-folding and protein folding are principally different.

ACKNOWLEDGMENTS

I am grateful to Dr. T.E. Creighton for stimulating discussions and to Dr. D. Webster for reading the manuscript. This work was partly supported by the EMBO short-term fellowship ASTF 5702.

REFERENCES

1. Wooley, S.-Y.C., Holzwarth, G. Intramolecular β-pleated sheet formation by poly-L-lysine in solution. Biochemistry 9:3604–3608, 1970.
2. Snell, C.R., Fasman, G.D. Kinetics and thermodynamics of the α-helix–β-structure transition of poly(L-lysine) and and L-leucine. Biochemistry 12:1017–1028, 1973.
3. Hartman, R., Schwaner, R.C., Hermans, J., Jr. β-Poly(L-lysine): A model system for biological self-assembly. J. Mol. Biol. 90:415–429, 1974.
4. Auer, H.E., Patton, E. Kinetics of the disordered chain—to—β transformation of poly(L-tyrosine) in aqueous solutions. Biophys. Chem. 4:15–21, 1976.
5. Auer, H.E., Miller-Auer, H. Dynamics of the disordered—β transition in poly(L-tyrosine) determined by stopped-flow spectrometry. Biopolymers 25:1607–1613, 1986.
6. Fukada, K., Maeda, H., Ikeda, S. Kinetics of pH-induced random coil—β structure conversion of poly[S-(carboxymethyl)-L-cysteine]. Macromolecules 22:640–645, 1989.
7. Cummings, A.L., Eyring, E.M. On the rate determining step for helix propagation helix—coil transition kinetics in aqueous poly(α-L-glutamic acid). Biopolymers 14:2107–2114, 1975.
8. Zana, R. On the rate determining step for helix propagation in the helix—coil transition of polypeptides in solution. Biopolymers 14:2425–2428, 1975.
9. Kuwajima, K., Yamaya, H., Sugai, S., Nagamura, T. Rapid formation of secondary structure framework in protein folding studied by stopped-flow circular dichroism. FEBS Lett. 221:115–118, 1987.
10. Zimm B.H., Bragg J.R. Theory of the phase transition between helix and random coil in polypeptide chains. J. Chem. Phys. 31:526–535, 1959.
11. Landau, L.D., Lifshitz, E.M. "Statistical Physics." London: Pergamon, 1959.
12. Birshtein, T.M., Elyashevich, A.M., Skvortsov, A.M. The theory of intramolecular cross β-coil transitions. Mol. Biol. (USSR) 5:78–90, 1971.
13. Adonts, V.G., Birshtein, T.M., Elyashevich, A.M., Skvortsov, A.M. Intramolecular conformational transitions "random coil—helix—folded structure" in polypeptides. Biopolymers 15:1037–1059, 1976.
14. Zwanzig, r., Lauritzen, J.I., Jr. Exact calculation of the partition function for a model of two-dimensional polymer crystallization by chain folding. J. Chem. Phys. 48:3351–3360, 1968.
15. Lauritzen, J.I., Jr., Zwanzig, R. Exact calculation of the partition function for a generalized model of two-dimensional polymer crystallization by chain folding. J. Chem. Phys. 52:3740–3751, 1970.
16. Ubbelohde, A.R. "Melting and Crystal Structure." Oxford: Clarendon Press, 1965.
17. Mattice W.L., Scheraga H.A. Role of interstrand loops in the formation of intramolecular cross-β-sheets by homopolyamino acids. Biopolymers 24:565–579, 1985.
18. Schwarz, G. On the kinetics of helix-coil transition of polypeptides in solution. J. Mol. Biol. 11:64–77, 1965.
19. Creighton, T.E. "Proteins." New York: W.H. Freeman, 1984.
20. Finkelstein, A.V. Kinetics of the antiparallel β-structure formation. Bioorgan. Khim. (USSR) 4:340–344, 1978.
21. Cosani, A., Terbojevich, M., Romanin-Jacur, L., Peggion, E. Potentiometric and circular dichroism studies of the coil-β transitions of poly-L-lysine. Peptides, Polypeptides and Proteins. Proc. Symp. Israel, pp. 166–176, 1974.
22. Barskaya, T.V., Ptitsyn, O.B. Thermodynamic parameters of helix—coil transition in polypeptide chains. II. Poly-L-lysine. Biopolymers 10:2181–2197, 1971.
23. Finkelstein, A.V. Secondary structure of the unfolded protein chain. Bioorgan. Khim. (USSR) 4:345–348, 1978.
24. Arfmann, H.A., Labitzke, R., Wagner, K.G. Conformational properties of L-leucine, L-isoleucine and L-norleucine side chains in L-lysine copolymers. Biopolymers 16:1815–1826, 1977.
25. Bychkova, V.E., Gudkov A.T., Miller, W.G., Mitin Yu.V., Ptitsyn, O.B., Shpungin, I.L. Thermodynamic parameters of the helix—coil transitions in polypeptides. III. Random copolymers of L-leucine with L-glutamic acid. Biopolymers 14:1739–1753, 1975.
26. Walter, B., Fasman, G.D. The random coil—β transitions of copolymers of L-lysine and L-isoleucine: Potentiometric titration and circular dichroism studies. Biopolymers 16: 17–32, 1977.
27. Mandel, R., Fasman, G.D. The random coil—β transitions of L-lysine and L-valine: Potentiometric titration and circular dichroism studies. Biopolymers 14:1633–1650, 1975.
28. Finkelstein, A.V., Ptitsyn, O.B. Statistical analysis of the correlation among amino acid residues in helical, β-structural and non-regular regions of globular proteins. J. Mol. Biol. 62:613–624, 1971.
29. Privalov, P.L. Stability of proteins. Small globular proteins. Adv. Protein Chem. 33:167–241, 1979.
30. Creighton, T.E. Toward a better understanding of protein folding pathways. Proc. Natl. Acad. Sci. U.S.A. 85:5082–5086, 1988.
31. Shakhnovich, E.I., Finkelstein, A.V. Theory of cooperative transitions in protein molecules. I. Why denaturation of globular protein is the first order phase transition. Biopolymers 28:1667–1680, 1989.
32. Finkelstein, A.V., Shakhnovich, E.I. Theory of cooperative transitions in protein molecules. II. Phase diagram for protein molecule in solution. Biopolymers 28:1681–1694, 1989.