Acta Biochimica Polonica

Vol. 44 No. 3/1997

423–432

QUARTERLY

*Critical review*

# Helix-coil transition theories. Are they correct?

Andrzej Bierzyński and Krzysztof Pawłowski

*Institute of Biochemistry and Biophysics, Polish Academy of Sciences, A. Pawińskiego 5a, 02-106 Warszawa, Poland*

**Key words**: α-helix, random coil, thermodynamics, peptides folding, hydration

**Principles of contemporary theoretical description of α-helix formation by polypeptide chains in water solution are shortly presented and critically discussed.**

**The theory treats the unfolded state of a peptide as "random coil" — an ideal conformation quite distant from reality. We suggest that for this reason the helix propagation parameters of amino-acid residues, determined using series of model peptides with different sequential patterns, are not the same.**

**Interpretation of the so called "nucleation parameter" is erroneous. In fact, it is not determined by the helix nucleation process but rather by a specific situation of residues at the helix N- and C-termini, and it strongly depends on solvation of their NH and CO groups, respectively. Consequently, helical segments with terminal sequences dominated by residues with strongly hydrophobic, bulky side chains can be very unstable.**

**We postulate that an unexpectedly high stability of very short, pre-nucleated helices studied by us arises from a "helix end separation effect": separated helix termini are better solvated than when they overlap each other. Because of this effect, helix initiation may be much more difficult than predicted by the theoretical "helix nucleation parameters".**

Helix-coil transition theories of Zimm & Bragg [1] and of Lifson & Roig [2] were originally developed to describe α-helix formation in long homopolypeptide chains. Their modified versions are still used today to calculate helix content and distribution of helical segments of various length in short heteropolypeptides — protein fragments or their synthetic models — peptides of *de novo* design. We shortly present here the basic ideas of the two theories and their modifications. We refer mostly to review articles where further references to original papers can be found.

In a number of cases, results of theoretical calculations compare favorably with experiment. Nevertheless, as we show in this paper, the very foundations of these theories are incorrect, mainly for two reasons: they refer to a "random coil" state of a polymer — a model very distant from real conformations assumed by polypeptide chains in water, and they falsely interpret one of their basic parameters, i.e., the so called "nucleation parameter".

We believe that these faults of the theory are responsible for its inability to explain an extreme stability of short, pre-nucleated α-

helix studied by us, as well as for large differences between helix propensities determined for amino-acid residues in various laboratories using various model peptides.

## THERMODYNAMICS OF α-HELIX FORMATION

Distribution of φ and ψ torsional angles of residues in the α-helical conformation are centered around –57° and –47°, respectively [3]. A system of hydrogen bonds is formed between the carbonyl groups of residues $j$ and the amide groups of residues $j + 4$ as shown in Fig. 1. The process of helix formation can be described as follows. Anywhere in an unfolded polypeptide two consecutive residues can accidentally adopt the helical conformation (e.g. residues 3 and 4 or 4 and 5 in Fig. 1) forming the helix nucleus, although no helix bond is formed, yet. Then, one of the adjacent residues (5 or 3, respectively, in Fig. 1) can adopt the helical conformation and the shortest possible helical segment is formed with one hydrogen bond. The residues adjacent to its two ends can undergo, in turn, a conformational transition with concomitant formation of hydrogen bonds, and the helix propagates in both directions. Of course, the helix nucleation and each step of its propagation are fully reversible. Consequently, a wide variety of peptide conformational species are present, containing helical segments of various length situated at various positions within the polypeptide chain*. Each of these species can be identified by the number of the first residue in the helical conformation ($i$) and by the number of residues forming the helical segment ($n$), e.g. in Fig. 1 $i = 3$ and $n = 3$.

Each helical segment is flanked at its N terminus by $i - 1$ residue, called N-cap, that is hydrogen-bonded by its CO group to the amide group of residue $i + 3$, although it is not in the helical conformation because rotational freedom around its $C_\alpha$–CO and $C_\alpha$–NH bonds is preserved. Analogously, the C-cap residue, $i + n$, adjacent to the C-terminal of the helix, is not in the helical conformation, although its NH group is hydrogen-bonded to the CO group of residue $i + n-4$.

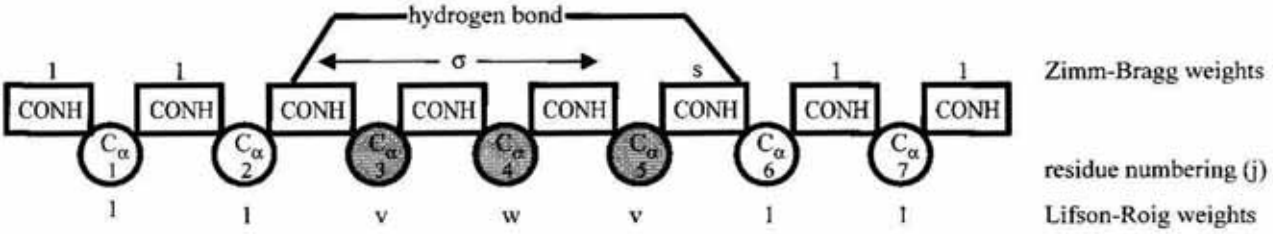


**Figure 1. Schematic representation of the shortest possible α-helical segment containing only one hydrogen bond.**

Basic units of helix formation and their weighting according to Zimm-Bragg [1] and Lifson-Roig [2] theories are shown. The residues in the helical conformation are indicated by shaded circles.

Therefore, in a polypeptide chain N residues long, $n$ can vary from 3 to N – 2 and $i$ from 2 to N – 3. The amide groups of the first three helical residues ($i$, $i$+1, and $i$+2), as well as the carbonyl groups of the last three residues ($i + n - 3$ , $i + n - 2$, and $i + n - 1$), do not participate in the helical hydrogen bonds. The statistical weight of species $i,n$, equal to the equilibrium constant between this species and the unfolded peptide, is given by the equation:

$$k_{i,n} = \exp(-\Delta G_{i,n}/RT) \qquad (1)$$

*In principle not only one but two or more separate helical segments can be formed within a long polypeptide chain. Nevertheless, in models or analogues of helical protein fragments not longer than 30 residues and not containing proline it is an event of a very low probability and can be neglected.

where $\Delta G_{i,n}$ is the Gibbs free energy of species $i,n$, relative to the unfolded peptide.

The population of each of the species equals:

$$x_{i,n} = k_{i,n}/Z \tag{2}$$

where the partition function

$$Z = 1 + \sum_{i=2}^{N-3} \sum_{n=3}^{N-i} k_{i,n} \tag{3}$$

Equations 1–3 are of general validity as derived uniquely from thermodynamic principles.

## THE HELIX FORMATION THEORIES

The aim of these theories is to express the statistical weights $k_{i,n}$ as functions of, as few as possible, parameters that can be determined experimentally.

The most widely used theory of the helix formation was developed by Lifson & Roig [2]. In its original version it describes a conformational equilibrium in homopolypeptide chains. In the Lifson-Roig theory the statistical weight of an $n$-residue long helix segment, independently of its position ($i$), is given by the expression:

$$k_n = v^2 w^{(n-2)} \tag{4}$$

$v^2$ is interpreted as the "nucleation parameter" because it is equal to the statistical weight of the helix nucleus formed by two residues adopting the α-helical conformation before the helix starts to propagate. It must be small, much smaller than unity, because of a large entropy drop related to the helix nucleation not compensated by hydrogen bonding. $w$ is called the "propagation parameter". Its value is determined not only by the conformational entropy drop of a residue but also by free enthalpy of CO–NH hydrogen bond it forms, as well as of short-range interactions, i.e. not exceeding $j-j\pm4$ range, of its backbone and side chain groups with the helix. Formally, the $v$ parameter is attributed to the first (i) and the last ($i+n-1$) residues of the helical segment and $w$ to all other residues within the helix (see Fig. 1).

The Lifson-Roig theory is based on two assumptions.

- 1) The long-range interactions within the helix can be neglected. This is not quite correct. There are long-range interactions between the peptide group dipoles that additionally stabilize long helices relative to short ones [4]. Nevertheless, this effect is of minor importance and will not be discussed in this paper.
- 2) In the unfolded peptide, as well as in its non-helical segments flanking the helix, each amino-acid residue behaves as a separate unit and does not interact in any way — apart from covalent bonds with its neighbors — with the rest of the polypeptide chain. Consequently, enthalpies and entropies of unfolded segments are simple sums of enthalpies and entropies of the residues they consist of. This is the reason why the Lifson-Roig formalism is called a "helix-coil transition" theory.

Apart from that theory an alternate description of the helix-coil transition, formulated by Zimm & Bragg [1], is widely used today. In this theory, not amino-acid residues, but rather peptide groups, are treated as the polypeptide chain units. Besides, the conformational criterion is not used to classify the units. Instead, a peptide group is defined as helical, if its amide hydrogen is engaged in a helical bond. Similarly as in the Lifson-Roig theory, the statistical weight of an $m$ unit long helical segment does not depend on its position within a polypeptide chain and it is given by the equation:

$$k_n = \sigma s^m \tag{5}$$

The nucleation (σ) and propagation ($s$) parameters are very close to, though not identical with [5], $v^2$ and $w$ in Lifson-Roig, respectively, and $m = n - 2$. σ is formally attributed to the N-terminus of the helix.

Both theories, used originally to describe the helix stability in long synthetic homopolypeptides, were soon modified to make

them suitable for studying much more interesting systems, namely heteropolypeptides.

Since the nucleation and propagation parameters are expected to be different for each of the 20 amino-acid residues, in a heteropolymer the statistical weight of a helical segment depends on its sequence and, consequently, on its position. Therefore:

$$k_{i,n} = v_i v_{i+n-1} \prod_{j=i+1}^{i+n-2} w_j \qquad (6)$$

according to Lifson-Roig and

$$k_{i,n} = \sigma_i \prod_{j=i+3}^{i+m+2} s_j \qquad (7)$$

in Zimm-Bragg formalism, where $i$ has the same meaning as in Eqn. 6 and $\sigma_i$ is attributed to residues $i$, $i + 1$, and $i + 2$.

These equations are correct, provided that an additional assumption, apart from the two discussed previously, holds true. Namely, that the side chains of residues within the helix do not interact in any way with each other or with the helix [6].

Investigations on conformational properties of C-peptide of RNase A and short synthetic model peptides with high α-helix propensities, initiated by R.L. Baldwin and then undertaken in a few other laboratories [7, 8], have shown that this assumption is incorrect. Electrostatic and hydrophobic interactions between side chains of residues $j$, $j\pm4$ and $j$, $j\pm3$ can promote the helix or destabilize it. Besides, electrostatic interactions of charged side chain groups of residues situated at the helix termini, including N- and C-cap ones, with dipoles of the terminal peptide groups of the helix (so called charge-helix dipole interactions), stabilize or destabilize the helical conformation. The N-cap and C-cap residues can interact with the helix and contribute to its stability, e.g. by forming hydrogen bonds between their side chains and free NH and CO backbone groups of N and C helix termini, respectively. Moreover, some specific sequences of residues preceding or following the N- and C-cap, respectively, can promote the helical conformation by incorporating the helix ends into relatively stable, specific structures (cap boxes [9, 10]), such as the so called paperclip conformation [11].

Following experimental investigations, and sometimes even promoting them (e.g. [12]), theoretical studies on side chain-side chain interactions within the helix and on the helix stabilizing cap effects have been made. Experimental and theoretical works have provided a large, though far from complete, library of free enthalpies arising from such effects [13, 14]. The Lifson-Roig and Zimm-Bragg theories have been modified to include them into the calculations of statistical weights of α-helical segments in polypeptides with various sequences [7, 8, 14, 15].

## TODAY "STATE OF THE ART" OF THE α-HELIX STABILITY CALCULATIONS

The helix propagation parameters ($s$ or $w$) have been determined for all the 20 amino-acid residues from circular dichroism studies of various model peptides [7, 14]. The "helix nucleation parameter" ($\sigma$ or $v^2$) is believed to be sequence- and temperature-independent and equal to about 0.0025 [7, 8, 15]. Using these parameters and the modified helix-coil transition theories it was possible, for a large variety of model peptides, to calculate not only the global helix content (as measured by CD), but also the probabilities of each residue of a polypeptide chain to be in the α-helical conformation (determined experimentally by NMR methods). Munoz & Serrano [14] have calculated the helix content in 17 protein fragments. They also calculated the temperature and pH dependence of the helix content in a number of model peptides and protein fragments [16]. Their results compare favorably with the experimental data.

These undeniable successes of the theory are rather puzzling, since the basic assumptions it is based upon have not, in fact, been satisfied.

## THE RANDOM COIL AS THE REFERENCE STATE

Although the modified versions of Zimm-Bragg and Lifson-Roig theories include specific interactions between residues within the helix, as well as cap effects, they still are based on the concept of the random coil conformation of unfolded peptide chains.

Actually, in the reference unfolded state, the peptide chain is assumed to have even more conformational freedom than in the random coil as defined by Brant & Flory [17] because the excluded volume effect is not taken into account: conformations of a polymer chain with interatomic distances shorter than van der Waals contact distances are not feasible. This effect increases with the polymer length. Therefore, the entropy of a random coil polymer is smaller than the sum of its unit entropies and the difference between the two values grows with the polymer length.

Consequently, the total entropy of two separate random coil segments $n$ and $k$ residues long is higher than that of one segment of $n + k$ residues. For this reason, short helical segments situated in the middle of a peptide should be favored against N- or C-terminal helices of the same length. Moreover, longer helices should be additionally stabilized relative to shorter ones. The effective $s$ or $w$ values should, therefore, depend on the length and position of a helical segment.

More important is that, in reality, conformation of unfolded peptides in aqueous solution is always very far from random coil. Nemethy & Scheraga [18] shown long ago that even when any side chain–side chain interactions between amino-acid residues were neglected the number of possible conformations of an unfolded peptide and their populations are sequence-dependent.

However, the conformational space of an unfolded peptide is also greatly reduced by various, more or less specific, side chain–side chain interactions. It was shown that some of them can lead to a high population of well-defined, non-helical conformations, even in very short peptides [19, 20]. It is true that only a few sequences were found to assume such structures at a detectable population, e.g. exceeding a few per cent. Nevertheless, by extrapolation, one can suspect that a large variety of sequences assume specific conformations that are marginally stable in comparison with sensitivity of available experimental techniques (population less than 2%) although, if a considerable number of them is formed by a peptide, its conformational entropy is dramatically lower, relative to that of the random coil.

The state of hydrophobic collapse is a particularly important exemplification of such an ensemble of conformations. By CD or NMR they may be indistinguishable from the random coil. Yet, the entropy of the collapsed state is much lower. The hydrophobic collapse has been proven for many unfolded protein chains [20], including some very short ones, e.g. the reduced bovine pancreatic trypsin inhibitor (BPTI) chain of 58 residues [21, 22]. There is no reason to expect that also much shorter peptides, particularly those with relatively high content of hydrophobic residues, would not tend to assume a globular conformation. Theoretical modeling of peptide folding has shown that such a conformation is indeed highly populated at temperatures close to the folding-unfolding transition, even by random-sequence polypeptide chains of 10–20 residues [23]. There is also no reason to assume that helical segments do not interact with the unfolded part of a peptide. They can be incorporated, e.g., into a wide variety of hydrophobically collapsed conformations.

So, why the helix-coil transition theories have been so successful in describing the helix content and helix segment distribution within the model peptides, if their basis are false? Most probably because the series of the models used to study the helix formation were always constructed on the basis of well-defined sequence patterns, conserved in all the peptides of a given series and only the effects of a single, or very few, sequence mutations have been measured [7]. Conformational properties of unfolded peptides are not perturbed significantly by such mutations. Nevertheless, the helix formation parameters measured for different series of model peptides should be different, and indeed they are.

The ranking of the amino-acid residues — beginning with the helix-breakers, characterized by the smallest $s$ or $w$ values, up to the best helix-makers, with alanine at the top of the list — is almost the same, but the absolute values of the helix propagation parameters are different. At temperatures close to 0°C the $s$ value for Ala measured by Baldwin and his coworkers [24] is 1.54 and that determined in Kallenbach's laboratory [25] equals 2.19. The helix contents in the peptide $AcAla_8NH_2$, calculated using the former and the latter $s$ values, including the cap effects of acetyl and $NH_2$ groups [26] equal 44% and 73%, respectively.

We suspect that these discrepancies arise from different conformational states of unfolded model peptides used in the two laboratories. The Baldwin's peptides are polyalanine-based with only a few lysine or glutamine residues incorporated to make them soluble. Such peptides can be expected to assume readily a globular conformation stabilized by hydrophobic interactions. On the other hand, the models used in Kallenbach's laboratory are not prone to the hydrophobic collapse, being built of polylysine and polyglutamate blocks, and their conformation in the unfolded state, although also quite different, probably, from the random coil, may be, nevertheless, much closer to it. Remarkably, the $s$ values determined by Park *et al.* [27] using a system being a hybrid of the two other peptide models lie in between the values obtained for them.

## PHYSICAL MEANING OF THE "NUCLEATION PARAMETERS"

Another objection that should be raised against the existing helix formation theories concerns their interpretation of $\sigma$ or $v^2$ parameters. They suggest that the free energy of a helical segment is related to the history of its formation, namely to its nucleation. This is misleading. Enthalpy and entropy, as the state functions, depend uniquely on the final state of the system and not on the way it has been attained. $\sigma$ or $v^2$ are not related, therefore, to the helix nucleation but they describe the special situation of the helix N- and C-termini with their NH and CO groups, respectively, not involved in the network of the helical hydrogen bonds. These groups must be still solvated, to a large extent, by water molecules. If they were not, the free enthalpy of each of them would increase by about 3–6 kcal/mol [28] and the total price paid for the helix formation would amount to 18–36 kcal/mol. The free enthalpy of the helix end solvation, relative to that of the unfolded state — $\Delta G_N$ and $\Delta G_C$ — contribute to the total value of $\Delta G_v = -RT \ln v$ in Lifson-Roig notation. Another contribution comes, of course, from conformational entropy drop of the helix ends not compensated by the energy of helical hydrogen bonds.

Therefore, the $\sigma$ and $v^2$ parameters should not be attributed to one of the helix termini ($\sigma$) or to the terminal residues ($v^2$) but rather to the whole terminal sequences $i, i+1, i+2$ and $i+n-3, i+n-2, i+n-1$, including the N- and C-cap residues as well.

Since $\sigma$ and $v^2$ depend on helix termini solvation they can be expected to be sequence-dependent. Why such a dependence has not been observed and it is believed that the "nucleation parameters" are, for all amino-acid residues, the same? We think that it is because the model peptides used for determination of the "helix nucleation" parameters consisted predominantly of alanine. In such peptides the ends of all possible helical segments were either built of, or dominated by, alanine residues.

Another possible explanation is that, indeed, the differences in the solvation free enthalpies of unfolded peptide and the helix ends are very similar for at least a great majority of sequences. Nevertheless, for some specific terminal sequences of residues with large hydrophobic side chains the $\sigma$ and $v^2$ values can be by orders of magnitude smaller than for polyalanine.

The thermodynamically correct interpretation of $\sigma$ and $v^2$ parameters makes it possible to suggest a plausible explanation for unexpectedly high stability of the helical conformation in the system studied by us. It consists of a calcium-binding loop (Fig. 2), analogous to the III$^{rd}$ loop of calmodulin, with a short sequence of amino-acid residues (AlaAlaAlaGln) with high helix propensity, at-

tached to its C-terminus. The loop coordinates lanthanide ions with a high binding constant ($10^5$ $M^{-1}$) [29, 30]. When saturated with the metal, it adopts a very rigid structure with residues Ala-10, Ala-11, and Glu-12 in the α-helical conformation. The NMR measurements of NH proton exchange rates, $C_\alpha H$ proton chemical shifts and NOE contact maps show that residues Ala13–Gln16 form also an α-helix of very high stability (Siedlecka *et al.*, in preparation). Comparative CD measurements in water and in 75% (v/v) trifluoroethanol indicate that at 1°C the helix population is close to 100%. Even assuming the largest possible errors of our measurements it must exceed 90%.

The existing theory is unable to explain this phenomenon. To begin with, there are some formal problems. According to Lifson-Roig the statistical weight in the $n$-residue long helical segment formed at the C-terminus of an α-helix nucleus is given by the equation:

$$k_n = w^n v \tag{8}$$

and the calculated helix content would be low because of the small $v$ value.

It seems much more reasonable, then, to use the expression proposed by Zimm-Bragg:

$$k_n = s^n \tag{9}$$

that gives much higher $k_n$ values. But, in fact, it is an accident. If our nucleation sites were situated not at N- but at C-terminus of the helix the equation:

$$k^n = \sigma s^n \tag{10}$$

should be used and the calculated helix content would be much lower. These paradoxes are a consequence of treating $v$ or σ parameters in an artificial way, without attributing to them a well-defined, physical meaning.

Since in our system the helix is nucleated, Eqn. 9 seems the most adequate to describe its formation. But even using the largest reported $s$ values [7] for alanine (2.19) and glutamine (1.2), measured by Gans *et al.* [25], and C-cap values for glutamine and the $NH_2$ group [26] the calculated average helical conformation of residues Ala13-Gln16 does not exceed 77%. To obtain a reasonable agreement with the experimental data (the helix content greater than 90%) the $s$ values of Gans *et al.* [25] have to be multiplied by a factor of at least 3. This means that the helical conformation of Ala13-Gln16 segment is additionally stabilized by $\Delta\Delta G < -4RT\ln 3 = -2.4$ kcal/mol. We believe that this free enthalpy comes from a "separation effect" of the helix ends.

When the peptide segment outside the loop is unfolded, the N- and C-ends of the helix formed by residues Ala-10, Ala-11, and Glu-12 merge into each other: both NH and CO groups of these residues are not hydrogen-bonded. As the helix propagates, the free groups move away. In the helix formed by residues 10–15 the end separation is completed: the residues with free NH groups are hydrogen-bonded by their carbonyls and *vice-versa*. Now, the solvation of the separated helix ends can be expected to be much more effective. The free amide and carbonyl groups of the helix ends present to water a well-defined, rigid structure. Water molecules tend to arrange themselves into clusters matching the helical configuration of hydrogen donors and acceptors, respectively. In the case of the three-residue long helix, the two clusters of water molecules collide with each other.

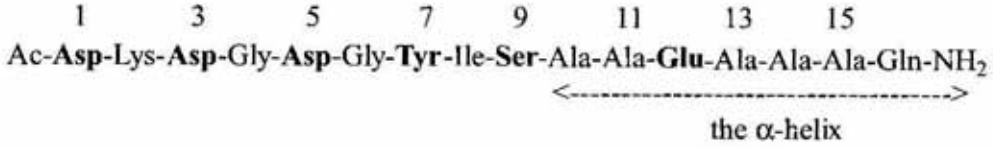


**Figure 2. The sequence of the studied peptide containing a calcium-binding loop (residues Asp1-Glu12) with residues Ala10-Glu12 fixed in the α-helical conformation by a coordinated lanthanide ion.** The residues providing the metal ligands are indicated by bold characters.

Therefore, the solvation energy of the merged helical ends is smaller than the sum of the solvation energies of the separated ends.

This hypothesis is not a product of pure speculation. Preliminary calculations made by us using the SYBYL program with energy parameters of Weiner *et al.* [31] have shown that, indeed, the absolute value of the solvation energy of the ends of the Ac-$Ala_{12}$-$NH_2$ helix is by about 30% greater than that of the helix formed by Ac-$Ala_3$-$NH_2$ peptide. In absolute values this effect can be estimated to contribute a few kcal/mol to the stability of Ala3-Glu16 helix in our peptide.

## CONCLUSIONS

The Zimm-Bragg and, particularly, Lifson-Roig theories of helix-coil transition, modified by introduction of parameters describing the sequence-dependent side chain–side chain interactions, N- and C-cap effects, and helix dipole–charge interactions have been shown to be useful tools for prediction of helix content in a wide variety of short polypeptide chains — protein fragments or model peptides.

Nevertheless, both theories are not thermodynamically sound. They refer to the conformational state of a polypeptide chain that goes even beyond the concept of "random coil" — conformation that is never, in fact, attained by real polypeptide chains in water solutions. This is probably the reason why different helix-coil transition parameters have to be used to calculate the helix content in different model peptides.

Interpretation of $\sigma$ and $v^2$ in Zimm-Bragg and Lifson-Roig theories, respectively, as "nucleation parameters" is misleading. Actually, these parameters are determined by free enthalpies of the helix ends. These, in turn, depend on a large variety of sequence-specific interactions: hydrogen bonds with side chain groups of N- and C-cap residues, electrostatic interactions of side chain charged groups with the helix dipole, and, last but not least, the solvation of their unbonded NH and CO backbone groups. All these effects are interrelated and should not be treated separately, as it is done now. For example, a hydrogen bond between a side chain of an N-cap residue with NH groups of $i + 2$ or $i + 3$ residues must have a large effect on solvation of the N-terminus of the helix. So far the cap effects and the charge–helix dipole interactions have been treated separately, and the helix end solvation has been largely neglected. Only on rare occasions the cap preferences of some residues (e.g. Gly [32]) have been correlated with end solvation effects.

Finally, one can ask whether it is worthwhile to analyze more deeply helix formation in peptides? After all, it is not a problem of crucial importance *per se*. The main point is to understand early steps in the protein folding process that are related with formation of secondary structures, the $\alpha$-helix being, perhaps, the most important of them. If the existing helix formation theory is good enough, why to criticize its basis and show that, in some extreme cases, as those of very short, pre-nucleated helices, it fails? Particularly, as the theory is approaching now the acceptable sophistication limits. Any further refinement would make it inapplicable in practice.

The answer is that if, for practical reasons, we cannot improve a theoretical model we use, it is always important to realize that it is only a model, more or less adequate to describe the reality, and to be aware of its limitations. It is also of crucial importance to realize what is the physical meaning of its parameters.

Interpretation of $\sigma$ parameter proposed in this paper has important consequences. For some sequences of helix ends, particularly those dominated by residues with bulky hydrophobic side chain groups, $\sigma$ can be much smaller than 0.0025 determined for polyalanine-based peptide chains. The helices with strongly hydrophobic ends would be very unstable and, indeed, such helices have not been observed either in model peptides or in proteins. Nevertheless, hydrophobic sequences are quite common in the interior of protein helical segments. The kinetics of $\alpha$-helix formation by such segments and, consequently, of protein folding, can be considerably influenced by them. They can impose quite large free energy barriers on the helix

propagation and unwinding. Therefore, these processes may not resemble free, one-dimensional diffusion, but rather they proceed by a series of conformational jumps.

If, as we argue, the solvation energy of the helix ends increases dramatically when they merge into each other (the end separation effect), the $\Delta G$ determining the helix initiation rate must be much higher than $-RT\ln\sigma$. In other words, the α-helix initiation is probably much more difficult than it has been supposed. Moreover, only the helices formed by at least 6 residues with well solvated helix ends can be expected to be of a stability high enough to play a kinetically significant role as protein folding seeds.

We would like to thank Dr. Jacek Wójcik for critical reading of the manuscript and helpful discussion.

## REFERENCES

1. Zimm, B.H. & Bragg, J.K. (1959) Theory of the phase transition between helix and random coil in polypeptide chains. *J. Chem. Phys.* **31**, 526–535.
2. Lifson, S. & Roig, A. (1961) On the theory of helix-coil transition in polypeptides. *J. Chem. Phys.* **34**, 1963–1974.
3. Pauling, L., Corey, R.B. & Branson, H.R. (1951) The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci. U.S.A.* **37**, 205–211.
4. Aqvist, J., Luecke, H., Quiocho, F.A. & Warshel, A. (1991) Dipoles localized at helix termini of proteins stabilize charges. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 2026–2030.
5. Qian, H. & Schellman, J.A. (1992) Helix-coil theories: A comparative study for finite length polypeptides. *J. Phys. Chem.* **96**, 3987–3994.
6. Bierzyński, A. (1987) The α-helical conformation of short natural polypeptide chains in water solutions. *Comments Mol. Cell. Biophys.* **4**, 189–214.
7. Chakrabartty, A. & Baldwin, R.L. (1995) Stability of α-helices. *Adv. Protein Chem.* **46**, 141–176.
8. Baldwin, R.L. (1995) α-Helix formation by peptides of defined sequence. *Biophys. Chem.* **55**, 127–135.
9. Zhou, H.X., Lyu, P., Wemmer, D.E. & Kallenbach, N.R. (1994) Alpha helix capping in synthetic model peptides by reciprocal side chain-main chain interactions: Evidence for an N terminal "capping box". *Proteins* **18**, 1–7.
10. Seale, J.W., Srinivasan, R. & Rose, G.D. (1994) Sequence determinants of the capping box, a stabilizing motif at the N-termini of α-helices. *Protein Sci.* **3**, 1741–1745.
11. Milner-White, E.J. (1988) Recurring loop motif in proteins that occurs in right-handed and left-handed forms. *J. Mol. Biol.* **199**, 503–511.
12. Presta, L.G. & Rose, G.D. (1988) Helix signals in proteins. *Science* **240**, 1632–1641.
13. Creamer, T.P. & Rose, G.D. (1995) Interactions between hydrophobic side chains within α-helices. *Protein Sci.* **4**, 1305–1314.
14. Munoz, V. & Serrano, L. (1995) Elucidating the folding problem of helical peptides using empirical parameters. II. Helix macrodipole effects and rational modification of the helical content of natural peptides. *J. Mol. Biol.* **245**, 275–296.
15. Stapley, B.J., Rohl, C.A. & Doig, A.J. (1995) Addition of side chain interactions to modified Lifson-Roig helix-coli theory: Application to energetics of phenylalanine-methionine interactions. *Protein Sci.* **4**, 2383–2391.
16. Munoz, V. & Serrano, L. (1995) Elucidating the folding problem of helical peptides using empirical parameters. III. Temperature and pH dependence. *J. Mol. Biol.* **245**, 297–308.
17. Brant, D.A. & Flory, P.J. (1965) The configuration of random polypeptide chains. II. Theory. *J. Am. Chem. Soc.* **87**, 2791–2800.
18. Nemethy, G. & Scheraga, H.A. (1977) Protein folding. *Quat. Rev. Biophys.* **10**, 239–352.
19. Dyson, H.J. & Wright, P.E. (1993) Peptide conformation and protein folding. *Curr. Opin. Struct. Biol.* **3**, 60–65.

**20.** Brooks III, C.L. (1993) Molecular simulations of peptide and protein unfolding: In quest of molten globule. *Curr. Opin. Struct. Biol.* **3**, 92–98.

**21.** Amir, D. & Haas, E. (1988) Reduced bovine pancreatic inhibitor has a compact structure. *Biochemistry* **27**, 8889–8893.

**22.** Gussakovsky, E.E. & Haas, E. (1992) The compact state of reduced bovine pancreatic trypsin inhibitor is not the compact molten globule. *FEBS Lett.* **308**, 146–148.

**23.** Fukugita, M., Lancaster, D. & Mitchard, M.G. (1993) Kinematics and thermodynamics of a folding heteropolymer. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6365–6368.

**24.** Chakrabartty, A., Kortemme, T. & Baldwin, R.L. (1994) Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. *Protein Sci.* **3**, 843–852.

**25.** Gans, P.J., Lyu, P.C., Manning, M.C., Woody, R.W. & Kallenbach, N.R. (1991) The helix-coil transition in heterogeneous peptides with specific side-chain interactions: Theory and comparison with CD spectral data. *Biopolymers* **31**, 1605–1614.

**26.** Doig, A.J. & Baldwin, R.L. (1995) N- and C-capping preferences for all 20 amino acids in alpha-helical peptides. *Protein Sci.* **4**, 1325–1336.

**27.** Park, S.-H., Shalongo, W. & Stellwagen, E. (1993) Residue helix parameters obtained from dichroic analysis of peptides of defined sequence. *Biochemistry* **32**, 7048–7053.

**28.** Rose, G.D. & Wolfenden, R. (1993) Hydrogen bonding, hydrophobicity, packing, and protein folding. *Annu. Rev. Biophys. Biomol. Struct.* **22**, 381–415.

**29.** Dadlez, M., Góral, J. & Bierzyński, A. (1991) Luminescence of peptide-bound terbium ions. Determination of binding constants. *FEBS Lett.* **282**, 143–146.

**30.** Wojcik, J., Góral, J., Pawłowski, K. & Bierzyński, A. (1997) Isolated calcium-binding loops of EF-hand proteins can dimerize to form a native-like structure. *Biochemistry* **36**, 680–687.

**31.** Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S.J. & Weiner, P. (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **106**, 765–784.

**32.** Harpaz, Y., Elmasry, N., Fersht, A.R. & Henrick, K. (1994) Direct observation of better hydration at the N-terminus of an α-helix with glycine rather than alanine as the N-cap residue. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 311–315.