PROTEINS: Structure, Function, and Genetics 10:287–299 (1991)

Physical Reasons for Secondary Structure Stability: α-Helices in Short Peptides

A.V. Finkelstein, A.Y. Badretdinov, and O.B. Ptitsyn
Institute of Protein Research, Academy of Sciences of the USSR, 142292 Pushchino, Moscow Region, USSR

ABSTRACT **It was recently found that some short peptides (including C- and S-peptide fragments of RNase A) can have considerable helicity in solution,[1–12] which was considered to be surprising. Does the observed helicity require a new explanation, or is it consistent with previous understanding? In this work we show that this helicity is consistent with the physical theory of secondary structure[12–19] based on an extension of the conventional Zimm-Bragg model.[20] Without any special modifications, this theory explains reasonably well almost all the experimentally observed dependencies of helicity on pH, temperature, and amino acid replacements. We conclude that the observed "general level" of helicity of C- and S-peptides (5–30% at room temperature and 10–50% near 0°C) is "normal" for short peptides consisting mainly of helix-forming and helix-indifferent residues. The helicity is modified by a multitude of weak specific side chain interactions, many of which are taken into account by the present theory;[13–19] some discrepancies between the theory and experiment can be explained by weak side-chain-side chain interactions that were neglected. A reasonable coincidence of the theory with experiment suggests that it had been used to investigate the role of local interactions in the formation of α-helical "embryos" in unfolded protein chains.**

Key words: polypeptides, α-helix, secondary structure, protein folding

INTRODUCTION

The nativelike "embryos" of protein structure have attracted great interest. It has been shown[1–12] that N-terminal fragments of ribonuclease A (C- and S-peptides, which are helical in the native protein) and their analogs,[1–8,11] as well as some other short peptides,[9,10,12] have considerable helicity in solution, depending upon temperature, pH, ionic strength, and replacements in the sequence. This indicates that a helical "autonomous folding unit"[8] may be as small as ~10 residues. Does this contradict theoretical expectations as proposed?[1,4] Does the observed helicity demand very specific and powerful contacts of side chains? To elucidate, we have calculated the expected helicity of the peptides investigated[1–12] for all the experimental conditions, using our previous computer program ALB[13] based on the physical theory of secondary structure of nonglobular polypeptides.[14–19] A reasonable concordance between theory and experiment enables us to speculate on the physical basis of stability of α-helices.

We describe here a basic (simplified) version of the theory, which gives results rather close to those obtained by the strict version,[13] to stress the important features without going into detail.

METHODS

Theory

Our treatment of helix-coil equilibrium is based on the Zimm-Bragg model.[20] We apply this model to heteropolypeptides and take into account specific side chain-backbone and side chain-side chain interactions, including interactions of side chains with the N- and C-termini of α-helices. The statistical weight of an α-helix spanning from a residue i to a residue j in a given peptide is

$$W_{i,j} = \sigma^{N_i} \cdot \prod_{k=i}^{j} S_k \cdot \sigma^{C_j} \qquad (1)$$

where S_k is the helix-elongation parameter for a residue k, σ^{N_i} is the initiation parameter of a helix whose N-end is residue i, and σ^{C_j} is the termination parameter for a helix whose C-end is residue j.

According to the simplest version of the Zimm-Bragg model, the helicity of a peptide consisting of M residues is equal to

$$\vartheta = \frac{1}{M}\sum_{k=1}^{M} \vartheta_k = \frac{1}{M}\sum_{k=1}^{M} \frac{\partial \ln Z}{\partial \ln S_k} \qquad (2)$$

where ϑ_k is the probability of a helical state for residue k (i.e., $\Sigma_k \vartheta_k$ is the average number of helical residues in a peptide), and:

Received December 14, 1989; revision accepted December 20, 1990.

Address reprint requests to O.B. Ptitsyn, Institute of Protein Research, Academy of Sciences of the USSR, 142292 Pushchino, Moscow Region, USSR.

$$Z = [1, 0] \times \prod_{k=n'}^{n''} \begin{bmatrix} 1 & \sigma^{N_k} S_k \\ \sigma^{C_{k-1}} & S_k \end{bmatrix} \times \begin{bmatrix} 1 \\ \sigma^{C_{n''}} \end{bmatrix} \quad (3)$$

is the partition function of the peptide. The matrix product is taken over all the residues whose ϕ, ψ angles can be fixed in the α-helix by hydrogen bonds. If a chain has -$CONH_2$-groups at the ends (e.g., if the N-terminus is acetylated and/or the C-terminus is amidated), $n' = 1$ and/or $n'' = M$. Otherwise $n' = 2$ or/and $n'' = M-1$.

Equations (2) and (3) permit calculation of the helicity of any peptide. Actually, our program[13] is based upon the precise version of the Zimm-Bragg model,[20] which takes into account that the α-helix cannot be less than three residue long, using 4×4 matrices rather than the 2×2 ones that are used in the simplest version and are explained above. The results of "precise" and "simplified" calculations coincide, however, within ~1%.

For a homopolypeptide (where $S_k = S$ for all k, and where the helix initiation/termination parameter $\sigma = \sigma^{N_i}\sigma^{C_j}$ is the same for all i and j) the equations (2) and (3) result in

$$\vartheta = \frac{\lambda_1 - 1}{\lambda_1 - \lambda_2} \cdot \frac{1 + \left(\frac{\lambda_2}{\lambda_1}\right)^{n+1} - \frac{2}{n}\frac{\lambda_2}{\lambda_1 - \lambda_2}\left(1 - \left(\frac{\lambda_2}{\lambda_1}\right)^{n}\right)}{1 + \frac{\lambda_1 - 1}{1 - \lambda_2}\left(\frac{\lambda_2}{\lambda_1}\right)^{n}} \cdot \frac{n}{M} \quad (4)$$

where $n = n'' - n' + 1$ and

$$\lambda_{1,2} = \frac{S + 1}{2} \pm \sqrt{\left(\frac{S-1}{2}\right)^2 + \sigma S}, \quad .$$

In heteropolypeptides all $S_{k'}$ and σ^{N_i} and σ^{C_j} are sequence-dependent.

The helix-elongation parameters S_k can be divided in two parts: $S^0(a_k)$, which depends only upon the type of residue a_k (=Ala, Gly, Lys . . .) that occupies a position k in the sequence, and a product of the contributions $\gamma_{k,k'}(a_k,a_{k'})$ of pairwise interactions of a_k with neighboring residues $a_{k'}$; these interactions are typically important only when $|k-k'| \leq 4$:

$$S_k = S^0(a_k) \cdot \left[\prod_{k'=k-4}^{k-1} \gamma_{k',k}(a_{k'},a_k) \cdot \prod_{k'=k+1}^{k+4} \gamma_{k,k'}(a_k,a_{k'})\right]^{1/2} \quad (5)$$

The values of σ^{N_i}, σ^{C_j} are represented as

$$\sigma^{N_i} = \sigma_0 \cdot \tilde{\sigma}^{N_i}_{i-1}(a_{i-1}) \cdot \tilde{\sigma}^{N_i}_{i}(a_i) \cdot \tilde{\sigma}^{N_i}_{i+1}(a_{i+1}) \cdot \tilde{\sigma}^{N_i}_{i+2}(a_{i+2}) \quad (6)$$

$$\sigma^{C_j} = \tilde{\sigma}^{C_j}_{j-2}(a_{j-2}) \cdot \tilde{\sigma}^{C_j}_{j-1}(a_{j-1}) \cdot \tilde{\sigma}^{C_j}_{j}(a_j) \cdot \tilde{\sigma}^{C_j}_{j+1}(a_{j+1})$$

where σ_0 is a sequence-independent parameter, $\tilde{\sigma}^{N_i}_k(a_k)$ is the contribution of residue a_k to initiation of a helix starting at residue i and $\tilde{\sigma}^{C_j}_k(a_k)$ is the contribution of residue a_k to termination of a helix ending at residue j.

Estimation of Stability Parameters

Helix-forming abilities of many amino acid residues have been estimated as a result of extensive experimental investigations of a vast number of different polypeptides. These data, however, concern mostly the helix-elongation and generalized helix initiation/termination parameters of residues in polypeptides consisting mainly of residues with long polar side chains, such as lysine, glutamic acid, and (hydroxypropyl)- or (hydroxybuthyl)-glutamine. Not all S parameters are known at present even for such an environment (and much less was known a decade ago when the theory[14–16] was suggested). Information on particular values of such parameters as $\tilde{\sigma}^N$, $\tilde{\sigma}^C$, and γ is practically absent. Therefore, we have to estimate the necessary parameters theoretically.

It is very difficult to calculate the S and σ values from first principles. It is much easier to take the experimental parameters for some residue "a" and then to evaluate the relative values (e.g., $S^0(a')/S^0(a)$) for all other residues. In this case we have to calculate only the differences between side chain interactions for different residues.

There are only two polypeptides for which parameters S(a) and σ have been carefully studied by direct thermodynamic methods (i.e., without any model assumptions)—poly(Lys) and poly(Glu).[21–27] Their S values at 25°C in uncharged states are 1.13 and 1.34, respectively; the difference can be explained by side chain-side chain hydrogen bonds in poly (Glu^O).[28] Direct thermodynamic methods[28,29] as well as the "guest-host" technique[30] have been applied also to measure the S value of Ala in copolymers with lysine,[29] glutamic acid,[28] and (hydroxypropyl)-glutamine,[30] all the results are in the range of $s = 1.07$–1.09.

Temperature dependence of S for uncharged poly(Lys) and poly(Glu) is described by equation

$$S(T) = S(T_*) \cdot \exp\left(\frac{\Delta H}{RT_*} - \frac{\Delta H}{RT}\right) \quad (7)$$

with $\Delta H = H_{helix} - H_{coil} \approx -1$ kcal/mol being the enthalpy of the coil-helix transition per residue.[22–27] We used this value of ΔH also for all other residues, which makes $S(0°C) = 1.16 \times S(25°C)$.

Experimental data on the charge-induced helix-coil transition in poly(Lys) and poly(Glu) show that the value of the helix initiation/termination parameter σ does not depend on temperature and is close to 0.0025 ± 0.0005 for these polypeptides in a half-charged state,[21,25–27] An estimation of electrostatic interactions[18] shows that the σ value must be approximately half this value when these polypeptides are uncharged. As "typical" water-soluble polypeptides have numerous long polar side chains, it is

convenient to take this value ($1.3 \cdot 10^{-3}$) as a value of σ_0, a sequence-independent part of the helix initiation/termination parameter.

The parameters that are determined by side chain-backbone and side chain-side chain interactions have been estimated from conformational analysis of polypeptides.[13-16,18] These parameters are the relative values of S°(a) as well as the values of $\tilde{\sigma}^N$, $\tilde{\sigma}^C$, and γ. All the parameters $\tilde{\sigma}^N$, $\tilde{\sigma}^C$, γ as well as all the ratios S°(a')/S° (a) were considered as temperature-independent.

An estimation of electrostatic interactions can illustrate a general method of evaluating the helix stability parameters.

The free energy of interaction of the charges q_1 and q_2 at distances r_{12} is evaluated[18] by the usual equation

$$G^{el}(q_1,q_2,r_{12}) = \frac{q_1 \cdot q_2}{\epsilon \cdot r_{12}} \cdot e^{-r_{12}/R_D} \qquad (8)$$

where ϵ (≈80) is the dielectric constant of the solution and R_D the Debye-Huckel radius (≈8 Å for 0.15 M and ≈3 Å for 1 M NaCl).

An increment of the electrostatic interactions between the side chains is equal to

$$\gamma^{el}_{k,k'} = \left\{\exp\left[-G^{el}(q_{a_k},q_{a_{k'}},r^{helix}_{k,k'})/RT\right]\right\}: \qquad (9)$$
$$:\left\{\exp\left[-G^{el}(q_{a_k},q_{a_{k'}},r^{coil}_{k,k'})/RT\right]\right\}.$$

Here q_{a_k}, $q_{a_{k'}}$ are the side group charges. Each charge can vary from zero to ±e (+e = proton charge), depending on the pK_a of the side chain group and on the pH. The distance $r_{k,k'}$ between the charges is averaged over the conformations of the side chains; the backbone in the coil is taken to be an extended one. For example, $r^{helix}_{k,k+4}$ is taken as 7Å and $r^{coil}_{k,k+4}$ as 12Å for all the side chains. This leads to $\gamma^{el}_{k,k+4}(Lys^+,Lys^+) = \gamma^{el}_{k,k+4}(Glu^-,Glu^-) \approx 0.65{:}0.9 = 0.7$ and $\gamma^{el}_{k,k+4}(Lys^+,Glu^-) \approx 1.4$ $(= 0.7^{-1})$ at 300°K in water solution with ionic strength of 0.15 M.

The estimate of Equation (9) can be improved if it is taken into account that side chains can have different conformations. The distance between the charges in some conformations (approximately in a quarter of them) is smaller, and in others greater, than the average one by a value of 2L (L≈2Å being the amplitude of side chain fluctuations):

$$\gamma^{el}_{k,k'} = \left\{\frac{1}{4}\exp\left[-G^{el}(q_{a_k},q_{a_{k'}},r^{helix}_{k,k'}-2L)/RT\right]+ \qquad (9a)\right.$$
$$+\frac{1}{2}\exp\left[-G^{el}(q_{a_k},q_{a_{k'}},r^{helix}_{k,k'})/RT\right] +$$
$$\left.\frac{1}{4}\exp\left[-G^{el}(q_{a_k},q_{a_{k'}},r^{helix}_{k,k'}+2L)/RT\right]\right\}:$$
$$:\left\{\frac{1}{4}\exp\left[-G^{el}(q_{a_k},q_{a_{k'}},r^{coil}_{k,k'}-2L)/RT\right]+\right.$$
$$+\frac{1}{2}\exp\left[-G^{el}(q_{a_k},q_{a_{k'}},r^{coil}_{k,k'})/RT\right] +$$
$$\left.\frac{1}{4}\exp\left[-G^{el}(q_{a_k},q_{a_{k'}},r^{coil}_{k,k'}+2L)/RT\right]\right\}.$$

In this case $\gamma^{el}_{k,k+4}(Lys^+,Lys^+) = \gamma^{el}_{k,k+4}(Glu^-,Glu^-) \approx 0.7$ but $\gamma^{el}_{k,k+4}(Lys^+,Glu^-) \approx 2.0$ $(> 0.7^{-1})$. This difference reflects an attraction between the oppositely charged side chains and thus accounts for the presence of salt bridges.

The interaction of charged side chains with partial charges $q_N \approx +e/2$ and $q_C \approx -e/2$ at the N- and C-ends of the α-helix have been estimated in a similar way.[18] Equation (8) is used with $q_1 = q_{a_k}$, $q_2 = q_N$ (or q_C), where $r_{12} = r_{k,N_i}$ (or r_{k,C_i}) is now the mean distance between the charge of residue a_k and the center of the N- (or C-) terminal turn of the α-helix positioned at residue *i*. For the side chains near to the N- and C-termini (i−1,i,i+1,i+2 and j−2, j−1,j,j+1), *r* is close to 4 Å and $\Delta G^{el} \approx 0.4$ kcal/mol. For the residues further from the helix termini, the distances are greater and the interactions decrease according to eq. (8).

We have considered the effect of the side chain charge-helix dipole interactions as a contribution to σ^N and σ^C values.[15,16,18] Later, Scheraga and coworkers[31,32] have shown that it can also be described through a change of *S* values for the residues that are situated in a sequence before and after each charged side chain. They have also shown that the experimental data on the helicity of C- and S-peptides can be rationalized after some fitting of the parameters of electrostatic interactions, such as the dielectric constant. In this work, however, we did not try any special fitting of the parameters, but used just a priori estimates of their values.

The interactions between nonpolar side chains are estimated from the size of the hydrophobic surface screened in a close side chain-side chain contact, and from the fraction of rotamers of side chains that provides these close contacts. For example, the side chains of the residues *k* and *k*+4 interact in a helix only if the $C_\beta - C_\gamma$ bond in residue *k* has $\chi_1 \approx 180°$ and if the same bond in residue *k*+4 has $\chi_1 \approx -60°$. As a result, their hydrophobic interaction can be estimated as

$$\gamma^{hp}_{k,k+4}(a_k,a_{k+4}) = 1 +$$
$$w^{\gamma}_{k,k+4} \cdot \{\exp\left[-A_{a_k,a_{k+4}} \cdot \xi/RT\right] - 1\}. \qquad (10)$$

Here $w^{\gamma}_{k,k+4}$ is the probability of a contact between the side chains of residues *k* and *k*+4 (equal to $n^{\gamma}_k \cdot n^{\gamma}_{k+4}/4$, where n^{γ} is the number of C_γ atoms in a corresponding residue), A is the accessible surface of

TABLE I. Theoretical Estimates of Helix-Coil Stability Parameters for Natural Amino Acid Residues at 300°K in Water at Ionic Strength of 0.15 M

Residue	Helix elongation* S°	Main contributions to helix initiation and termination†			
		$\tilde{\sigma}^{N_i}_{i-1}$	$\tilde{\sigma}^{N_i}_{i,i+1,i+2}$	$\tilde{\sigma}^{C_j}_{i-2,i-1,i}$	$\tilde{\sigma}^{C_j}_{i+1}$
Leu	1.25				
Met	1.15‡				
Lys°, pH > 10.5	1.15				
Glu°, pH < 4.3	1.15				
Phe	1.10				
Trp	1.10‡				
Ala	1.10			1.2§	
Ile	1.10‡	0.6		0.8§	
Arg°, pH > 12	1.10‡				
Tyr	1.10‡				
Gln	1.00‡				
Lys^+, pH < 10.5	1.00‡	0.7	0.7	1.7	1.7
His°, pH > 6.3	1.00‡			1.2″	
Glu^-, pH > 4.3	1.00	1.7	1.7	0.7	0.7
Arg^+, pH < 12	0.95‡	0.7	0.7	1.7	1.7
Val	0.95	0.6		0.8§	
Cys	0.95				
His^+, pH < 6.3	0.85‡	0.7	0.7	1.7	1.7
Asn	0.85‡	1.7		1.2″	
Asp°, pH < 4	0.85‡	1.7			
Thr	0.75‡	1.7			
Ser	0.75	1.7		1.5″	
Asp^-, pH > 4	0.65‡	3.2	1.7	0.7	0.7
Gly	0.6		1.3§	0.8§	1.5
Pro	0.1		40; 13; 13		0.1
Nonspecific parameters	ΔH = −1kcal/mol			$\sigma_o = 0.0013$	

*Values refer to the residues surrounded by long polar side chains (such as Lys° or Gln). They practically coincide (usually with an error of 5% and not greater than 10%) with experimental values for residues in poly(hydroxypropyl)- and poly(hydroxybuthyl)-glutamine.[43]

†Only the values that differ from 1 are listed.

‡Values that have been first predicted and then confirmed experimentally.

§Only for position i and $i + 1$ or $j - 1$ and j.

″Only for position j.

nonpolar atoms of the side chains that become screened in a contact, and $\xi = -25$ cal/mol$Å^2$ is the increment of the hydrophobic effect.[23]

The estimates of H-bonds between the side chains (e.g., between Glu^O and Gln in k, $k+3$ pairs) were made in a way similar to that used earlier for the estimation of H-bonds between side chains and the backbone.[15,16] In these estimates, the free energy of the protein H-bond in water was taken as −2 kcal/mol and the free energy of fixation of one angle of internal rotation as +1 kcal/mol (cf.[27]).

The set of stability parameters used in the theory is presented in Tables I and II. The parameters indicated in Table I were calculated in 1975–1977, included into the ALB program,[13] and published separately.[14–19] The parameters of side chain-side chain interactions (Table II) were included in the ALB program, but have not been published separately (except for the estimates of electrostatic interactions).[18] The only differences between the parameters used by the ALB program in 1983 and now are:

1. Now we use ΔH = 1 kcal/mol, which better fits the available experimental data on natural amino acids (see[22–27]), whereas in 1983 we used ΔH = 0.6 kcal/mol, the average over poly(Lys), poly(Glu), poly(hydroxypropyl)-, and (hydroxybutyl)-glutamine.

2. The estimate (9) for electrostatic interactions is replaced by a more precise one (9a).

3. If NH_3^+- and COO^--groups exist at the termini of a polypeptide chain, their charges are now taken into account in the same way as the charges of side chains.

None of these modifications affects greatly the calculated helicity of short peptides.

Figure 1 shows how these parameters are used to calculate the statistical weight of a given α-helix. For an α-helix starting at Trp 3 and ending at Arg 10, the statistical weight is 0.0029 at 300°K (27°C). At 3°C the statistical weight of the same helical region is $0.0029 \cdot (1.16)^8 = 0.0095$. However, this is only one term of the partition function of a helical state, as α-helices can start and end at all other points of a polypeptide. As a result the total statistical weight of α-helical states in this peptide (~10 positions for the N-end and ~10 positions for the C-end of a helix) is as large as ~1/3.

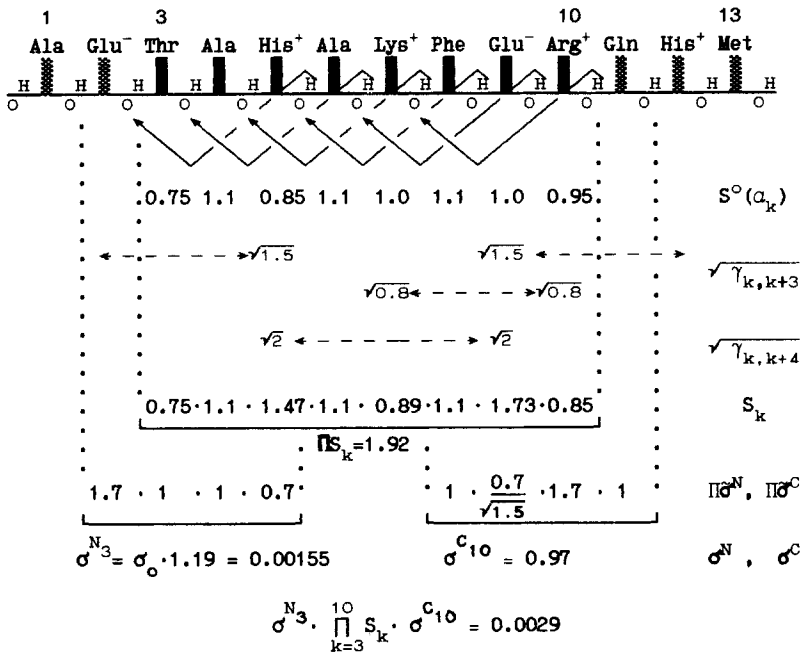

Fig. 1. Parameters for the stability of the α-helical region in peptide 2 (see Table IV) starting at Thr3 (i.e., Thr3 is the first residue whose φ, ψ angles are fixed by a hydrogen bond) and ending at Arg10 (the last residue whose φ, ψ angles are fixed by a hydrogen bond). Values of all parameters at 300°K have been taken from Tables I and II for pH 5.0 (when all side chains are charged) and ionic strength 0.15 M.

TABLE II. Theoretical Estimates of γ, Increments of the Most Important Side Chain-Side Chain Interactions to Stability of α-Helices

$\gamma_{k,k+3}(a_k,a_{k+3})$

a_k \ a_{k+3}	Leu,Phe, Trp,Tyr	Ile Met	Val	Arg^+,His^+ Lys^+	Asp^- Glu^-	Glu°	Gln
Leu,Phe,Trp, Tyr	1.4	1.3	1.2				
Ile,Met	1.3	1.2	1.1				
Val	1.2	1.1	1.0				
Arg^+,His^+,Lys^+				0.8	1.5		
Asp^-,Glu^-				1.5	0.8		
Glu°						1.1	1.05
Gln						1.05	1.0

$\gamma_{k,k+4}(a_k,a_{k+4})$

a_k \ a_{k+4}	Ile,Leu, Phe	Trp, Tyr	Met Val	Arg^+,His^+ Lys^+	Asp^- Glu^-
Ile,Leu,Phe	1.8	1.5	1.5		
Trp,Tyr	1.5	1.0	1.3		
Met,Val	1.5	1.3	1.3		
Arg^+,His^+,Lys^+				0.7	2.0
Asp^-,Glu^-				2.0	0.7

RESULTS AND DISCUSSION

The "General Level" of Helicity for Short Peptides of Helix-Forming Residues

It is worthwhile to show first that the rather high helicity of short peptides found by Baldwin et al.[1–12] is not surprising and that the oligopeptides of "normal" helix-forming residues (e.g., uncharged Glu or Lys) must also have similar helicities at the same temperatures; see also[34].

Table III shows the calculated helical contents for 13 residue peptides from uncharged lysine and glutamic acids. Calculations have been done (see Fig. 2) using equation (4) for S values taken from direct experiments with poly(Lys) and poly(Glu).[23–26] Two values of σ have been used: σ = 0.0025, which was experimentally obtained for half-charged poly(Lys) and poly(Glu),[21,25–27] and σ = 0.0008 (used in[4]), which is smaller than our estimate (σ = 0.0013) for uncharged poly(Lys°) and poly(Glu°). As the majority of peptides studied in[1–12] have acetylated N-ends

TABLE III. Helical Content of 13-Residue Peptides

	$(Lys^{o})_{13}$ $\vartheta(\%)$			$(Glu^{o})_{13}$ $\vartheta(\%)$			13-residue heteropep-
T(°C)	s(T)	$\sigma = 0.0008$	$\sigma = 0.0025$	s(T)	$\sigma = 0.0008$	$\sigma = 0.0025$	tides[3,4,6,7,8,11]
3	1.27	12	24	1.53	38	53	10–50%
27	1.13	6	13	1.34	18	33	5–30%
45	1.05	4	8	1.22	9	19	4–20%

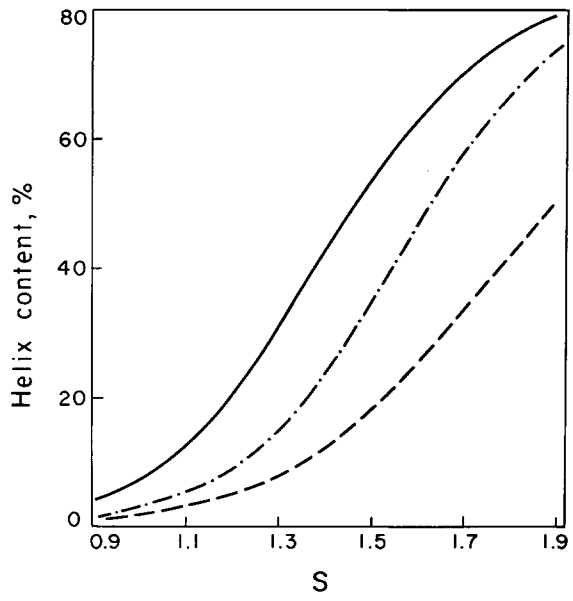

Fig. 2. Helical content (i.e., the average fraction of residues fixed by hydrogen bonds) of a 13-residue peptide as a function of *S*, the helix-coil equilibrium constant. Curve ——— corresponds to equation (4) with $n = M = 13$ (i.e., with $n' = 1$ and $n'' = 13$) and $\sigma = 0.0025$. Curve –·–·– is the same equation, but with $\sigma = 0.0008$. Curve – – – is taken from[4] and corresponds to $M = 12$, $n = 11$ and $\sigma = 0.0008$.

and amidated C-ends, we use n = M = 13 in equation (4).

It follows that the expected helical contents of $(Lys^{O})_{13}$ and $(Glu^{O})_{13}$ are quite consistent with those of 13-residue heteropeptides studied in.[1,3,4,6–8,11] Longer peptides with similar amino acid contents have even greater predicted helicities.

α-Helices in C-, S-, and Alanine-Based Short Peptides

Figure 3 presents a comparison of calculated and experimental helical contents for 40 different C-, S-, and alanine-based peptides and protein fragments studied experimentally.[1,3–12,35–37] (Table IV). The pH-dependencies of helical content are considered below (see Figs. 5–8). Calculations have been done using equations (3) and (5)–(7), with the set of parameters listed in Tables I and II.

Calculations for temperatures other than 27°C have been carried out using equation (7) with $\Delta H = -1$ kcal/mol, which fits the experimental data for poly(Glu^{O}) and poly(Lys^{O}) and have been arbitrarily used for other amino acid residues. The reasonable agreement between our calculations and the experimental data suggests that this ΔH value can be used for all the residues involved, particularly for alanine, which is the main component of the majority of these peptides. This is emphasized in Figure 4, which compares the experimental temperature dependence of helical content of peptide 21 with the calculations based on $\Delta H = -1$ kcal/mol.

Of course, this does not mean that ΔH must be ≈ -1 kcal/mol for all amino acid residues. Very few reliable experimental data are available on ΔH for other than poly(Glu^{O}), poly(Lys^{O}), and alanine-based short peptides. For long polypeptides, the temperature dependence of the helical content is determined by $\Delta H/\sqrt{\sigma}$ rather than by ΔH, and therefore ΔH can be evaluated only if σ is known. For homopolypeptides, σ can be obtained from the dependence of helical content on the chain length. Application of this method to poly(hydroxypropyl)-glutamine[40–42] and poly(hydroxybutyl)-glutamine[42] gave $\Delta H = -0.1 \div -0.2$ kcal/mol. Applied to heteropolypeptides ("guest- host" technique),[42] this method gives very different ΔH's for different residues.[43] However, the application of this method to heteropolypeptides would be reliable only with a set of samples with different chain lengths and the same amino acid content, which is difficult to achieve. Therefore, the S(T) dependencies and σ estimations are uncertain, and, in fact, the "guest-host" method can give reliable results only for *S* averaged over the studied intervals of temperatures. As the middle points of these intervals are usually close to room temperature, the "guest-host" technique gives *S* values at room temperature much more reliably than the dependencies S(T). For example, experimental data on copolymers of alanine with (hydroxypropyl)-glutamine[30] can be described by the values $\Delta H_{Ala} = -1$ kcal/mol and $\sigma_{Ala} = 0.002$ (see Table I) almost as well as by the values $\Delta H = -0.24$ kcal/mol and $\sigma = 0.0008$ used in.[30]

The general conclusion which we can drawn from Figure 3 is that the existing theory[13–19] of helical structure of unfolded polypeptides explains not only the general level of helicity of short peptides, but, in the majority of cases, also the dependence of their helical content on their sequences. The calculation actually describes all the observed range of se-

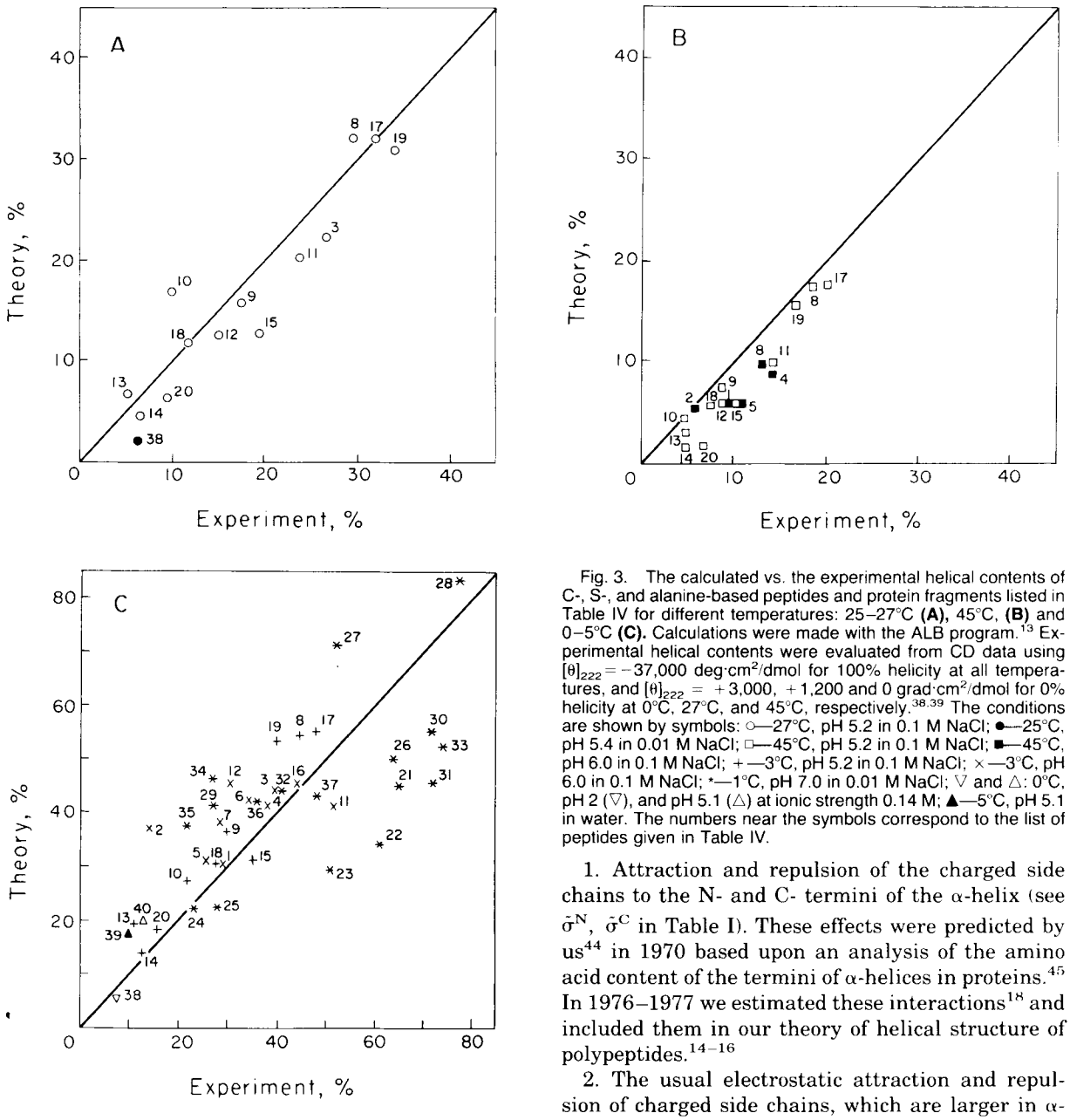

Fig. 3. The calculated vs. the experimental helical contents of C-, S-, and alanine-based peptides and protein fragments listed in Table IV for different temperatures: 25–27°C **(A)**, 45°C, **(B)** and 0–5°C **(C)**. Calculations were made with the ALB program.[13] Experimental helical contents were evaluated from CD data using $[\theta]_{222} = -37{,}000$ deg·cm²/dmol for 100% helicity at all temperatures, and $[\theta]_{222} = +3{,}000$, +1,200 and 0 grad·cm²/dmol for 0% helicity at 0°C, 27°C, and 45°C, respectively.[38,39] The conditions are shown by symbols: ○—27°C, pH 5.2 in 0.1 M NaCl; ●—25°C, pH 5.4 in 0.01 M NaCl; □—45°C, pH 5.2 in 0.1 M NaCl; ■—45°C, pH 6.0 in 0.1 M NaCl; +—3°C, pH 5.2 in 0.1 M NaCl; ×—3°C, pH 6.0 in 0.1 M NaCl; *—1°C, pH 7.0 in 0.01 M NaCl; ▽ and △: 0°C, pH 2 (▽), and pH 5.1 (△) at ionic strength 0.14 M; ▲—5°C, pH 5.1 in water. The numbers near the symbols correspond to the list of peptides given in Table IV.

quence-dependent variations of helical content (from ~5 to ~35% at room temperature and from ~10 to ~75% at low temperatures). The maximal deviations between our calculations and experiment are usually no more than twofold, which corresponds to a free energy deviation of less than 0.5 kcal/mol for a polypeptide.

Electrostatic Interactions

Let us consider, as an example, the effect of electrostatic interactions on the helical content, which determines its pH-dependence. These interactions are:

1. Attraction and repulsion of the charged side chains to the N- and C- termini of the α-helix (see $\tilde{\sigma}^N$, $\tilde{\sigma}^C$ in Table I). These effects were predicted by us[44] in 1970 based upon an analysis of the amino acid content of the termini of α-helices in proteins.[45] In 1976–1977 we estimated these interactions[18] and included them in our theory of helical structure of polypeptides.[14–16]

2. The usual electrostatic attraction and repulsion of charged side chains, which are larger in α-helices than in the coil state (see γ in Table II).

3. A decrease in S° upon ionization of side chains (see Table I) due to greater repulsion of charges from a bulk nonpolar α-helix than from a thin extended chain.[18]

Figures 5–8 compare the experimental pH-dependencies of helical contents of seven peptides with the calculations. The theory explains the helicity decrease at low pH for peptides 3, 6, and 7 by the protonation of Glu^-2 and the disappearance of its electrostatic attraction to the partial positive charge at the N-end of the α-helix dipole. In a similar way, the decrease of helicity at neutral pH for peptides 3, 4, 5, and 6 is explained by the disappearance of the His^+12 attraction to the negative charge at the C-

TABLE IV. List of Peptides

#	Name	Sequence	Reference
	C-peptides		
1	C-I	ac-AETAAAKFERQHM-amide	4
2	C-I (A5→H)	ac-AETAHAKFERQHM-amide	4
3	C-III	ac-AETAAAKFLRAHA-amide	6
4	C-III (R10→A)	ac-AETAAAKFLAAHA-amide	8
5	C-III (E2→A)	ac-AATAAAKFLRAHA-amide	6
6	C-III (F8→A)	ac-AETAAAKALRAHA-amide	8
7	C-III (H12→A)	ac-AETAAAKFLRAAA-amide	6
8	C-III (ac-A1→suc-A)	suc-AETAAAKFLRAHA-amide	8
9	C-III (ac-A1→A)	AETAAAKFLRAHA-amide	8
10	C-III (ac-A1→K)	KETAAAKFLRAHA-amide	8
11	RN44	suc-AATAAAKFLAAHA-amide	7
12	RN54	ac-AATAAAKFLAAHA-amide	7
13	RN56	KATAAAKFLAAHA-amide	7
	S-peptides		
14	S (1–15)	KETAAAKFERQHMDS	6
15	SI	KETAAAKFLREHMDS-amide	6
16	SIII	ac-AETAAAKFLREHMDS-amide	6
17	SIV	suc-AETAAAKFLREHMDS-amide	6
18	SV	KETAAAKFLRQHMDS-amide	6
19	SVI	suc-AETAAAKFLRQHMDS-amide	6
20	S (1–19)	KETAAAKFERQHMDSSTSA	1
	Alanine-based peptides		
21	3K (I)	ac-AAAAKAAAAKAAAAKA-amide	10
22	3K (II)	ac-AKAAAAKAAAAKAAAA-amide	10
23	4K	ac-AKAAKAAAAKAAAAKA-amide	10
24	6K (I)	ac-AKAAKAKAAKAKAAKA-amide	10
25	6K (II)	ac-AKAAAKKAAAKKAAAKA-amide	10
26	3E	ac-AEAAAAEAAAAEAAAA-amide	10
27	(i+3) E,K	ac-AEAAKAEAAKAEAAKA-amide	9
28	(i+4) E,K	ac-AEAAAKEAAAKEAAAKA-amide	9
29	(i+3) K,E	ac-AKAAEAKAAEAKAAEA-amide	9
30	(i+4) K,E	ac-AKAAAEKAAAEKAAAEA-amide	9
31	3ALA	ac-YKAAAAKAAAAKAAAAK-amide	12
32	3ILE	ac-YKAAIAKAAIAKAAIAK-amide	12
33	3LEU	ac-YKAALAKAALAKAALAK-amide	12
34	3PHE	ac-YKAAFAKAAFAKAAFAK-amide	12
35	3VAL	ac-YKAAVAKAAVAKAAVAK-amide	12
36	2VAL	ac-YKAAVAKAAVAKAAAAK-amide	12
37	1VAL	ac-YKAAAAKAAVAKAAAAK-amide	12
	Protein fragments with low helicity		
38	C-helix of ribonuclease	SLADVQAVCSQK	35
39	C-helix of myohemerythrin	EVVPHKKMHKDFLEKGGL	36
40	α-helix of BPTI	SAEDAMRTAGGA	37

end of the α-helix dipole. The first effect is absent in peptide 5, where Glu 2 is replaced by Ala. The second disappears in peptide 7, where His 12 is replaced by Ala. Quite unexpected is the observed behavior of peptide 4 (which differs from peptide 3 only by the replacement Arg10 → Ala): its helicity does not decrease at low pH despite the presence of Glu 2.

In the symmetrical peptide 24, the N- and C-terminal effects compensate each other. The increase of its helicity at high pH (when Lys residues become uncharged) is explained by the disappearance of repulsions between charged lysines in positions k and $k+3$ and by the increase of S°(Lys°) as compared with S°(Lys^+) (see[18]).

In peptide 21 there are no electrostatic side chain-side chain interactions within a turn of a helix. The disappearance of attraction of Lys^+15 to the C-end of the α-helix at high pH in this case is just compensated (according to our parameters) by the increase of S^O upon uncharging of three lysine residues. Therefore our calculations do not predict an increase in helicity of this peptide at high pH. A simple estimate shows that the observed increase may be due to the disappearance of a repulsion between Lys^+5 and Lys^+15 (which is not included in the present theory).

Discrepancies Between Calculations and Experiments

There are only two systematic deviations of our calculations from experimental data: (1) the observed helicity of the peptides consisting mainly of

alanine is usually ~1.5 times greater than that calculated; and (2) the point replacements of alanines for valines, phenylalanines, and isoleucines decrease markedly the observed helicities of these peptides.[12]

These deviations can be explained by weak side chain-side chain interactions. Only the strongest pairwise interactions (mainly those between bulk hydrophobic or between charged residues) have been included in the calculations[13] and in Table II of this study. The weaker ones (e.g., the interactions between the hydrophobic parts of long polar side chains)[46] are neglected at present; however, in some cases they seem to alter considerably the apparent helix-forming abilities of amino acid residues.

In the usual treatment of experimental data for watersoluble homo- and heteropolypeptides (e.g., in the "guest-host" technique), the side chain-side chain interactions are actually included in the apparent values of the S constants. This is true also for the S^O values listed in Table I, which may be treated as S values of residues surrounded by long polar side chains (such as Lys^O). In another environment the S parameters of the same residues can be different (see eq. (5)).

A replacement of one "host" residue (h) by a "guest" one (g)

$$...hhhhhhhhh... \rightarrow ..hhhhghhhh...$$

changes the free energy of coil-helix transition by a value

$$\Delta G_{g\rightarrow h} = (G_{hg} + G_g^O + G_{gh}) - (G_{hh} + G_h^O + G_{hh}). \quad (11)$$

Here G_g^O and G_h^O are the contributions from the backbone-backbone and side chain-backbone interactions of the residues g and h, while G_{hh}, G_{hg} and G_{gh} are the contributions of the interactions between their side chains.

The free energy of the coil-helix transition per residue in poly(h) is equal to

$$(G_h)_{in\ h} = G_h^O + G_{hh} \quad (12)$$

so equation (11) can be represented as the difference between the apparent value of the free energy of the helix-coil transition for the guest residue in a poly(h) environment

$$(G_g)_{in\ h} = G_g^O + G_{hg} + G_{gh} - G_{hh} \quad (12a)$$

and the corresponding free energy (12) of the host residue:

$$\Delta G_{g\rightarrow h} = (G_g)_{in\ h} - (G_h)_{in\ h}.$$

As a result, the apparent S value of a "guest" in a given host polypeptide is given by

$$(S_g)_{in\ h} = (S_g)_{in\ Ala} \cdot \gamma_{hg} \cdot \gamma_{gh} / \gamma_{hh} \cdot \quad (13)$$

Here

$$(S_g)_{in\ Ala} = \exp(-G_g^O/RT)$$

(as the small alanine side chain does not interact with any others, i.e., all $\gamma_{Ala,g} = \gamma_{g,Ala} = 1$),

$$\gamma_{hg} = \exp(-G_{hg}/RT) = \prod_{k'=k+1}^{k+4} \gamma_{k,k'}\ (h,g)$$

$$\gamma_{gh} = \exp(-G_{gh}/RT) = \prod_{k'=k-4}^{k-1} \gamma_{k',k}\ (g,h)$$

and so on. The difference between the apparent S value given by equation (12), and that given by equation (5) is due to the side chain-side chain interactions modifying not only the S value of a "guest," but also those values of the neighboring "host" residues.

Let us compare the apparent S values for two host polypeptides, poly(Ala) and poly(p),p being a residue with a long polar side chain, e.g., p = Lys° or p = (hydroxybutyl)-glutamine. As $\gamma_{Ala,p} = \gamma_{p,Ala} = 1$,

$$(S_{Ala})_{in\ p} = (S_{Ala})_{in\ Ala} / \gamma_{pp} \quad (14)$$

where $\gamma_{pp} = \exp(-G_{pp}/RT)$ and G_{pp} is the free energy of interactions of long polar side chains. This mean that small alanine side chains break down the interactions between long polar side chains in poly(p), which destabilizes the α-helical state. Therefore the apparent S value of Ala in p is smaller than its intrinsic value (i.e., than its value in poly(Ala)).

In a similar way, the effective S value of a residue with a long polar side chain in poly(p) includes the side chain-side chain interactions between these groups and therefore is larger than their intrinsic values:

$$(S_p)_{in\ p} = (S_p)_{in\ Ala} \cdot \gamma_{pp}. \quad (15)$$

The bulky nonpolar side chains, with two C_γ-atoms or aromatic rings (Val, Ile, Phe, Tyr, Trp), interact more strongly with long polar side chains than long polar side chains interact with each other. If we assume that $G_{p,b} \approx G_{b,p} \approx 2G_{pp}$ (where b designates a bulky non-polar side chain), we get

$$(S_b)_{in\ p} = (S_b)_{in\ Ala} \cdot \gamma_{p,b} \cdot \gamma_{b,p} / \gamma_{pp} \approx (S_b)_{in\ Ala} \cdot (\gamma_{pp})^3 \quad (16)$$

(note that the interactions of bulky non-polar groups with each other have been already included in Table II).

It is worthwhile to mention that the main contributions in γ_{pp} are provided by the interactions of side chains in k and $k \pm 4$ positions of an α-helical state, i.e., between the adjacent turns of an α-helix. Therefore the terms "in Ala" or "in p" mean that positions $k \pm 4$ relative to residue k are occupied either by alanine (or other residues with short side chain) or by residues with long polar side chains.

According to our previous theoretical estimates, G_{pp} (which is due mainly to the attraction of C_γ atoms of residues k and $k+4$ within an α-helix) was about $-0.03 \div -0.06$ kcal/mol, and we neglected

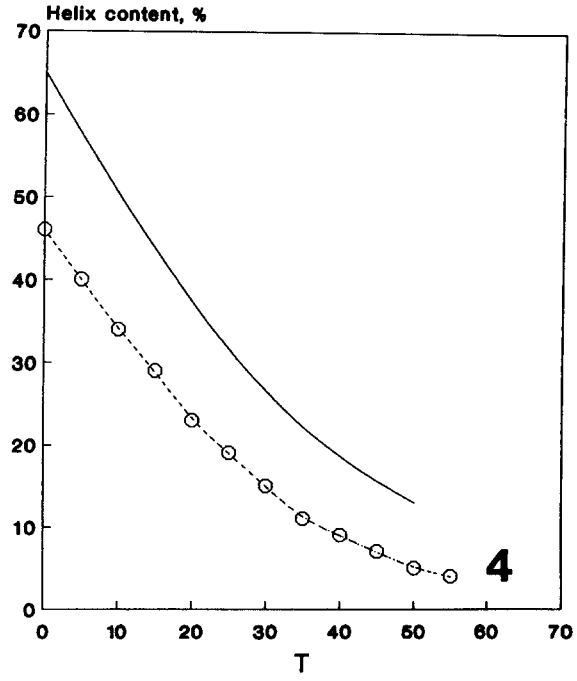

Fig. 4.

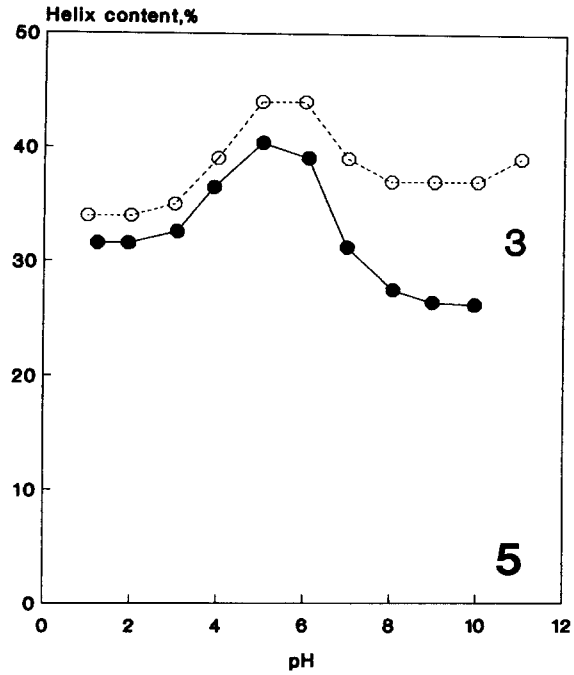

Fig. 5.

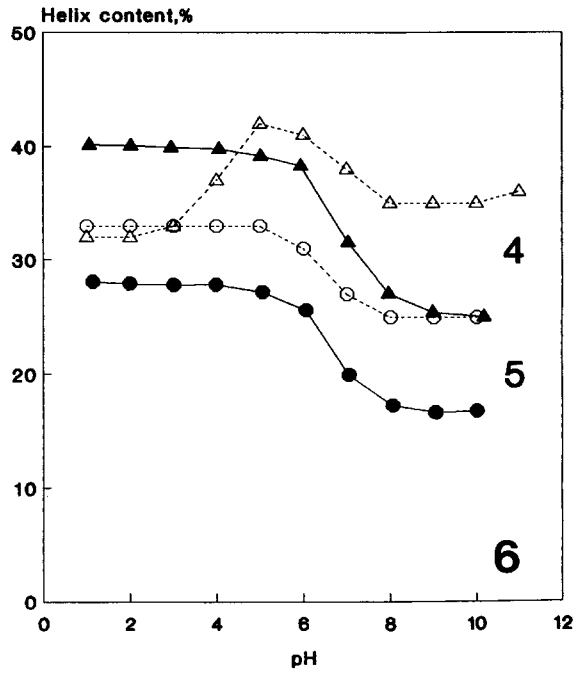

Fig. 6.

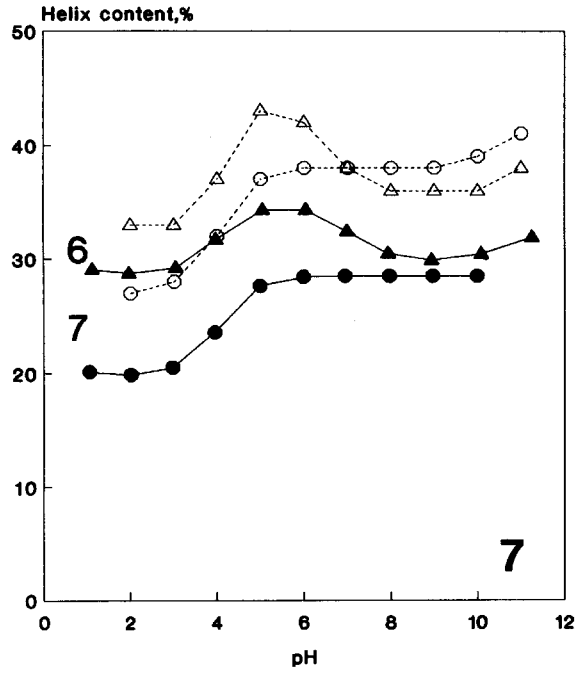

Fig. 7.

Figures 4–7. Legends appear on page 297.

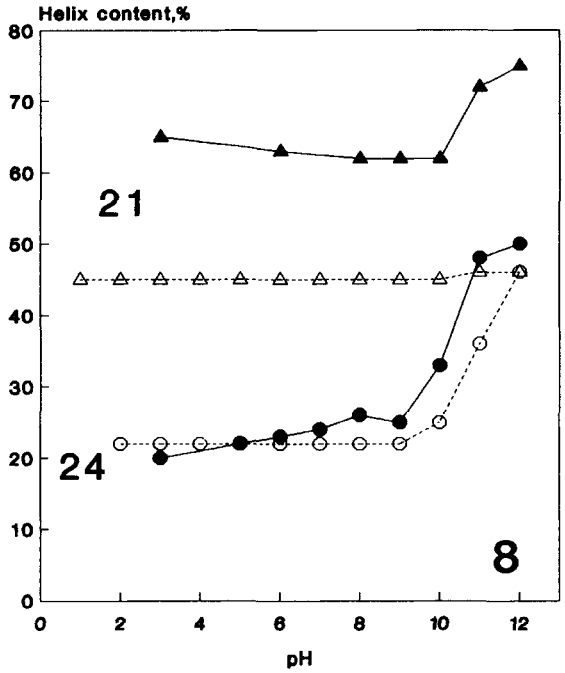

Fig. 8. pH dependence of helical content for peptides 21, i.e., 3K(I), (triangles), and 24, i.e., 6K(I), (circles) at 1°C in 0.01 M NaCl.

such a small interaction. It seems, however, that this attraction is actually greater and must be taken into account. Experimental data on alanine-based peptides[9,10,12] are consistent with $G_{pp} \approx -0.1$ kcal/mol (i.e., $\gamma_{pp} \approx 1.2$).

If this estimate is correct it follows that the "intrinsic" S values of residues are

$$(S_{Ala})_{in\ Ala} = (S_{Ala})_{in\ p} \cdot \gamma_{pp} \approx \begin{cases} 1.1 \times 1.2 \approx 1.3 \text{ at } 25°C \\ 1.3 \times 1.2 \approx 1.5 \text{ at } 0°C. \end{cases} \quad (17)$$

$$(S_p)_{in\ Ala} = (S_p)_{in\ p}/\gamma_{pp} \approx \begin{cases} 1.1_5/1.2 \approx 0.9_5 \text{ at } 25°C \\ 1.3/1.2 \approx 1.1 \quad \text{at } 0°C. \end{cases} \quad (18)$$

$$(S_{val})_{in\ Ala} = (S_{val})_{in\ p}/\gamma_{pp}^3 \approx 0.9_5/(1.2)^3 \approx 0.5_5 \text{ at } 25°C. \quad (19)$$

$$(S_{I,F,Y,W})_{in\ Ala} = (S_{I,F,Y,W})_{in\ p}/\gamma_{pp}^3 \approx 1.1/(1.2)^3 \approx 0.6_5 \text{ at } 25°C. \quad (20)$$

Here the values of S in p at 25°C (practically coinciding[19,47] with the experimental values[43]) have been taken from Table I and the S values at 0°C have been obtained as S(0°C) = 1.16·S(25°C) (see above).

These S values of residues in an Ala environment describe the deviations of the experimental data on Ala-based peptides from the results of our calculation based on the ALB program. For example, an Ala → Val replacement in the environment of long polar side chains will lead only to a small destabilization of the α-helix by a factor of $(S_{Val}/S_{Ala})_{in\ p} = 0.95/1.1 \approx 0.85$, whereas the same replacement in an alanine environment results in a much greater destabilization: $(S_{Val}/S_{Ala})_{in\ Ala} = 0.55/1.3 \approx 0.4$, which is consistent with the experimental data for Ala-based peptides.[12]

Of course, the effect of replacements of Ala by residues with bulky non-polar side chains also can be influenced by differences in S(T) dependencies for Ala and these groups. Little is known of the magnitudes of these differences. To discriminate their effect, it would be worthwhile to measure the effects caused, for example, by an Ala → Val replacement at room temperature.

It follows that "intrinsic" S values of residues (in an Ala environment) can be sometimes quite different from their effective values (e.g., in the environment of long polar side chains). However, these "intrinsic" values operate only in special cases, when the given residue is surrounded mainly by alanines or other residues with small side chains (such as the Ala-based peptides studied in).[9,10,12] This is not the case for a typical protein and peptide, for which the "old" S values obtained by the "guest-host" technique[43] can be used as a reasonable approximation. This is the reason why our calculations based on the "guest-host" S values coincide reasonably with the majority of the experimental data on small peptides.

CONCLUSIONS

We conclude that the "general level" of helicity observed in[1–12] is "normal" for short peptides consisting mainly of helix-forming and helix- indifferent residues. The "general level" is considerable at room temperatures and even higher at low temperatures. This "level" is due to the backbone-backbone interactions ($C_\beta H_3$ group being taken as a part of the backbone) and to fluctuations of helix termini. The helical content can be modified by a multitude of weak specific side chain-backbone and side chain-

Fig. 4. Temperature dependence of helical content for peptide 21, i.e., 3K(I), in 1.0 M NaCl at pH7. _____experiment,[10] – –○– –theory. The coincidence in slopes reflects a reasonable agreement between the accepted (−1 kcal/mol) and experimental enthalpies of helix-coil transitions.

Fig. 5. pH dependence of helical content for peptide 3, i.e., CIII, at 3°C in 0.1 M NaCl; ———●——— -experiment[6], – –○– – -calculations.

Fig. 6. pH dependence of helical content for peptides 4, i.e., CIII(R10→A) (triangles), and 5, i.e., CIII(E2→A) (circles). Each peptide differs from CIII only by one "point mutation" shown in parentheses.

Fig. 7. pH dependence of helical content for peptides 6, i.e., CIII(F8→A) (triangles) and 7, i.e., CIII (H12→A) (circles).

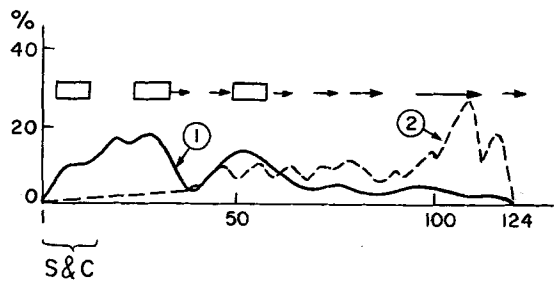

Fig. 9. The predicted probability of α-helical (1) and β-structural (2) states of residues in the unfolded chain of ribonuclease A reproduced from[48]; calculations performed for pH7, 27°C, 0.15 M NaCl. The native location of α (□) and β (→) regions is shown above the plot. The abscissa shows residue numbers in the primary structure. The fragment that corresponds to S- and C-peptides is marked. It would be timely now to prove or disapprove experimentally also the high predicted stability of the β-hairpin at the C-terminus!

side chain interactions. Many of them (particularly electrostatic ones) have been taken into account.[13–19] Some interactions, however, await a quantitative explanation.

Though each separate specific interaction is weak (contributing less than 0.5 kcal/mol, according to both the theoretical estimates and experiments with short peptides), a multitude of these interactions can lead to a relatively defined localization of α-helices (as well as of β-hairpins) in a heteropolypeptide (protein) chain, even in the absence of a globular structure. This is illustrated in Figure 9, which presents the results of our previous calculations[48] of secondary structure for unfolded ribonuclease A. The predicted probability of α-helical structure near the N-end reaches 20% and is very small near the C-end, whereas the probability of the β-structure is very small near the N-end and as large as 30% near the C-end.

It is worthwhile to mention that the NH-groups involved in the β-sheet of ribonuclease A are protected very early during its folding.[49] It would be interesting to compare the theory with the experimental data on secondary structure formation in proteins at very early steps of their folding (which are now becoming available; see[50–53] and the data of Semisotnov and Kuwajima[54] and to learn the role of specific local interactions in the formation of the "embryos" of protein secondary structure.

This study attempts to compare systematically and quantitatively the experimental data with the results of the physical theory of secondary structure of unfolded protein chains. The mathematical apparatus of the theory[17] is strict and permits inclusion of all possible interactions in unfolded chains. Therefore, the reliability of the theory depends only upon the values of the parameters. The calculated values of helix-coil equilibrium constants are in agreement with the experimental ones[19,47] in the corresponding environment (see above). Moreover, these values, as well as the evaluations of effects at N- and C-termini of α-helices, are in qualitative agreement with the amino acid content of α-helices and their N- and C-termini in globular proteins.[16] Therefore, it is not surprising that the physical theory can predict secondary structure of proteins at a level of the best empirical and semiempirical methods. But it can do more: it can predict the absolute values of helicities for fluctuating peptides and unfolded protein chains.

REFERENCES

1. Kim, P.S., Bierzinski, A., Baldwin, R.L. A competing salt-bridge suppresses helix formation by the isolated C-peptide carboxylate of ribonuclease A. J. Mol. Biol. 162:187–199, 1982.
2. Bierzinski, A., Kim, P.S., Baldwin, R.L. A salt bridge stabilizes the helix formed by isolated C-peptide of RNase A. Proc. Natl. Acad. Sci. U.S.A. 79:2470–2474, 1982.
3. Kim, P.S., Baldwin, R.L. A helix stop signal in the isolated S-peptide of ribonuclease A. Nature 307:329–334, 1984.
4. Shoemaker, K.R., Kim, P.S., Brems, D.N., Marqusee, S., York, E.J., Chaiken, I.M., Stewart, J.M., Baldwin, R.L. Nature of the charged-group effect on the stability of the C-peptide helix. Proc. Natl. Acad. Sci. U.S.A. 82:2349–2353, 1985.
5. Mitchinson, C., Baldwin, R.L. The design and production of semisynthetic ribonucleases with increased thermostability by incorporation of S-peptide analogs with enhanced helical stability. Proteins 1:23–33, 1986.
6. Shoemaker, K.R., Kim, P.S., York, E.J., Stewart, J.M., Baldwin, R.L. Tests of the helix dipole model for stabilization of α-helices. Nature 326:563–567, 1987.
7. Shoemaker, K.R., Kim, P.S., York, E.J., Srewart, J.M., Baldwin, R.L. Circular dichroism measurements of peptide helix unfolding. Proc. 10th. Am. Pept. Symp.:15–20, 1987.
8. Shoemaker, K.R., Fairman, R., Kim, P.S., York, E.J., Stewart, J.M., Baldwin, R.L. The C-peptide helix from ribonuclease A considered as an autonomous folding unit. Cold Spring Harbor Symp. Quant. Biol. LII:391–398, 1987.
9. Marqusee, S., Baldwin, R.L. Helix stabilization by $Glu^-\cdots Lys^+$ salt bridges in short peptides of de novo design. Proc. Natl. Acad. Sci. U.S.A. 84:8898–8902, 1987.
10. Marqusee, S., Robbins, V.H., Baldwin, R.L. Unusually stable helix formation in short alanine based peptides. Proc. Natl. Acad. Sci. U.S.A. 86:5286–5290, 1989.
11. Fairman, R., Shoemaker, K.R., York, E.J., Stewart, J.M., Baldwin, R.L. Further studies of the helix dipole model: Effects of a free $\alpha\text{-}NH_3^+$ or $\alpha\text{-}COO^-$ group on helix stability. Proteins 5:1–7, 1989.
12. Padmanabhan, S., Marqusee, S., Rigeway, T., Lane, T.M., Baldwin, R.L. Relative helix-forming tendencies of nonpolar amino acids. Nature 344:268–270, 1990.
13. Finkelstein, A.V. Program "ALB" for protein and polypeptide secondary structure calculation and prediction, 1983. Deposited at the Brookhaven Protein Data Bank, Upton, NY, and (a more user friendly version) at the EMBL, Heidelberg, Germany; distributed by request.
14. Finkelstein, A.V., Ptitsyn, O.B. A theory of protein molecule self-organization. IV. Helical and irregular local structures of unfolded protein chains. J. Mol. Biol. 103: 15–24, 1976.
15. Finkelstein, A. V., Ptitsyn, O.B. Theory of protein molecule self-organization. I. Thermodynamical parameters of local secondary structures in the unfolded protein chain. Biopolymers 16:469–495, 1977.
16. Finkelstein, A.V., Ptitsyn, O.B., Kozitsyn, S.A. Theory of protein molecule self-organization. II. A comparison of calculated thermodynamic parameters of local secondary structures with experiment. Biopolymers 16:497–524, 1977.
17. Finkelstein, A.V. Theory of protein molecule self-organization. III. A calculating method for the probabilities of the secondary structure formation in an unfolded protein chain. Biopolymers 16:525–529, 1977.
18. Finkelstein, A.V. Electrostatic interactions of charged groups in water environment and their influence on the

polypeptide chain secondary structure formation. Molek. Biol. (U.S.S.R.) 10:811–819, 1977.
19. Ptitsyn, O.B., Finkelstein, A.V. Similarities of protein topologies: evolutionary divergence, functional convergence or principles of folding? Quart. Rev. Biophys. 13:339–386, 1980.
20. 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.
21. Snipp, R.L., Miller, W.G., Nyland, R.E. The charge-induced helix—random coil transition in aqueous solution. J. Amer. Chem. Soc. 87:3547–3553, 1965.
22. Rialdi, G., Hermans, J., Jr. Calorimetric heat of the helix-coil transition of poly-L-glutamic acid. J. Amer. Chem. Soc. 88:5719–5720, 1966.
23. Hermans, J., Jr. Experimental free energy and enthalpy of formation of the α helix. J. Phys. Chem. 70:510–515, 1966.
24. Olander, D.S., Holtzer, A. The stability of the polyglutamic acid α helix. J. Amer. Chem. Soc. 90:4549–4560, 1968.
25. Bychkova, V.E., Ptitsyn, O.B., Barskaya, T.V. Thermodynamic parameters of helix-coil transition in polypeptide chains. I. Poly-(L-Glutamic Acid). Biopolymers 10:2161–2179, 1971.
26. Barskaya, T.V., Ptitsyn, O.B. Thermodynamic parameters of helix-coil transition in polypeptide chains. II. Poly-L-lysine. Biopolymers 10:2181–2197, 1971.
27. Ptitsyn, O.B. Thermodynamic parameters of helix-coil transition in polypeptide chains. Pure Appl. Chem. 31: 227–244, 1972.
28. Bychkova, V.E., Ptitsyn, O.B. Thermodynamic parameters of helix-coil transitions in polypeptide chains. IV. Statistical copolymers of L-alanine with L-glutamic acid. Molek. Biol. (U.S.S.R.) 10:916–922, 1976.
29. Sugiyama, H., Noda, H. Stability of the helical conformation of random L-alanine-L-lysine copolymers in aqueous solution. Biopolymers 9:459–469, 1970.
30. Platzer, K.E.D., Ananthanarayanan, V.S., Andreatta, R.H., Scheraga, H.A. Helix-coil stability constants for the naturally occurring amino acids in water. IV. Alanine parameters from random poly[(hydroxypropyl)glutamine-co-L-alanine]. Macromolecules 5:177–187, 1972.
31. Vasquez, M., Pinkus, M.R., Scheraga, H.A. Helix-coil transition theory including long-range electrostatic interactions: Application to globular proteins. Biopolymers 26: 351–371, 1987.
32. Vasquez, M., Scheraga, H.A. Effect of sequence-specific interactions on the stability of helical conformation in polypeptides. Biopolymers 27:41–58, 1988.
33. Chothia C. Hydrophobic bonding and accessible surface in proteins. Nature 248:338–339, 1974.
34. Finkelstein, A. V., Badretdinov, A.Ya., Ptitsyn, O.B. Concerning α-helix stability in short peptides. Nature 345: 300–300, 1990.
35. Jimenez, M.A., Nieto, J.L., Herranz, J., Rico, M., Santoro, J. ^{1}H NMR and CD evidence of the folding of the isolated ribonuclease 50-61 fragment. FEBS Letters 221:320–324, 1987.
36. Dyson, H.J., Rance, M., Houghten, R.A., Wright, P.E., Lerner, R.A. Folding of immunogenic peptide fragments of proteins in water solution. II. The nascent helix. J. Mol. Biol. 201:201–217, 1988.
37. Goodman, E.M., Kim, P.S. Folding of a peptide corresponding to α-helix in bovine pancreatic trypsin inhibitor. Biochemistry 28:4343–4337, 1989.
38. Greenfield, N., Fasman, G.D. Computed circular dichroism spectra for the evaluation of protein conformation. Biochemistry 8:4108–4116, 1969.
39. Tiffany, M.L., Krimm, S. Effect of temperature on the circular dichroism spectra of polypeptides in the extended state. Biopolymers 11:2309–2316, 1970.
40. Lotan, N, Yaron, A., Berger, A., Sela, M. Conformational changes in the nonionizable water-soluble synthetic polypeptide poly-N^5-(3-hydroxypropyl)-L-glutamate. Biopolymers 3:625–655, 1965.
41. Okita, K., Teramoto, A., Fujita, H. Solution properties of synthetic polypeptides. VI. Helix-coil transition of poly-N^5—(3-hydroxypropyl)-L-glutamine. Biopolymers 9: 717–738, 1970.
42. von Dreele, P.H., Lotan, N., Ananthanarayanan, V.S., Andreatta, R.H., Poland, D., Scheraga, H.A. Helix-coil stability parameters for the naturally occurring amino acids in water. II. Characterization of host polymers and application of host-guest technique to random poly[(hydroxypropyl)glutamine-co-(hydroxybuthyl)glutamine]. Macromolecules 4:408–417, 1971.
43. Sueki, M., Lee, S., Powers, S.P., Denton, J.B., Konishi, Y., Scheraga, H.A. Helix-coil stability constants for the naturally occurring amino acids in water. 22. Histidine parameters from random poly[(hydroxybuthyl)glutamine-co-L-histidine]. Macromolecules 17:148–155, 1984.
44. Ptitsyn, O.B., Finkelstein, A.V. Relation of the secondary structure of globular proteins with their primary structure. Biofizika (U.S.S.R.) 15:757–767, 1970.
45. Ptitsyn, O.B. Statistical analysis of the distribution of amino acid residues among helical and non-helical regions in globular proteins. J. Mol. Biol. 42:501–510, 1969.
46. Lotan, N, Yaron, A., Berger, A. The stabilisation of the α-helix in aqueous solution by hydrophobic side-chain interaction. Biopolymers 4:365–368, 1966.
47. Ptitsyn, O. B., Finkelstein, A.V. Theory of protein secondary structure and algorithm of its prediction. Biopolymers 22:15–25, 1983.
48. Finkelstein, A.V., Ptitsyn, O.B. Theory of the protein secondary structure self-organization: the dependence of the native structure on the unfolded chain secondary structure. Doklady Akad. Nauk SSSR 242:1226–1228, 1978.
49. Udaomkar, J.B., Baldwin, R.L. NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A. Nature 335:694–699, 1988.
50. Kuwajima, K., Yamaya, H., Miwa, S., 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.
51. Gilmanshin, R.I., Ptitsyn, O.B. Early intermediate at α-lactalbumin refolding forms within 20 msec. FEBS Letters 223:327–329, 1987.
52. Kuwajima, K., Sakuroaka, A., Fueki, S., Yoneyama, M., Sugai, S. Folding of carp parvalbumin studied by equilibrium and kinetic circular dichroism spectra. Biochemistry 27:7419–7423, 1989.
53. Goldberg, M.E., Semisotnov, G.V., Friguet, B., Kuwajima, K., Ptitsyn, O.B., Sugai, S. An early immunological intermediate of tryptophan synthase β_2 subunit is a "molten globule". FEBS Lett. 263:51–56, 1990.
54. Ptitsyn, O.B., Semisotnov, G.V. Mechanism of protein folding. In: Conformations and Forces in Protein Folding (B.T. Nall and K.A. Dill, eds), Proc. of AAAS Meeting on Protein Folding, New Orleans, 1990, 1991:153–166.