Biochemistry

 Volume 34, Number 12 March 28, 1995

New Concepts in Biochemistry

Views of Helical Peptides: A Proposal for the Position of 3_{10}-Helix along the Thermodynamic Folding Pathway†

Glenn L. Millhauser‡

Department of Chemistry and Biochemistry, University of California, Santa Cruz, California 95064

Received December 5, 1994; Revised Manuscript Received February 10, 1995

Proteins contain predominantly two types of helical structures: α-helix (with $i \rightarrow i+4$ hydrogen bonding) and 3_{10}-helix (with $i \rightarrow i+3$ bonding) (Schulz & Schirmer, 1979; Barlow & Thornton, 1988; Toniolo & Benedetti, 1991). 3_{10}-Helix is less common than α-helix, but structural work nevertheless suggests that 3_{10}-helix plays an important biochemical role (Barford & Johnson, 1989; Gerstein & Chothia, 1991; McPhalen et al., 1992; Kavanaugh et al., 1993; Kostrikis et al., 1994; Liu & Day, 1994). Although the existence of 3_{10}-helix in proteins is well established, detecting this conformation in water-soluble helical peptides can be difficult. Over the last two years we have been using double label electron spin resonance (ESR) in an attempt to distinguish helix conformations (Miick et al., 1992; Fiori et al., 1993, 1994). In several designed peptides, evidence has emerged in favor of partial 3_{10}-helix content.

Here we examine the 3_{10}-helix/α-helix equilibrium from the perspective of crystallographic studies, and the results are placed within the context of helix−coil transition theory. Next, recent experiments from NMR, circular dichroism (CD), and double label ESR are discussed. Finally, concepts arising from these diverse views are combined into a single model that attempts to explain the thermodynamic folding pathway of helical peptides.

X-ray Crystallography

α-Aminoisobutyric acid (Aib) is a $C_{\alpha,\alpha}$ disubstituted amino acid. From the geometric restrictions imposed by the two methyls, Aib-rich peptides typically crystallize in a helical conformation. The helix type, α or 3_{10}, depends upon peptide length and the number of Aib's. 3_{10}-Helix is favored by shorter peptides containing one or more Aib residues. Karle and Balaram (1990) have summarized these conformational trends with a graph where they depict helix geometry as a function of both the total number of residues and the Aib content. Interestingly, the 3_{10}-helix and α-helix regions are distinct and separated by a linear boundary. Extrapolation of this boundary to zero Aib content intersects the helix length axis at six residues. This implies that a six-residue helical peptide, containing only L amino acids, is equally likely to form 3_{10}-helix or α-helix.

Protein helices exhibit similar behavior. Barlow and Thornton (1988) surveyed the structures and structural distortions commonly found in protein helices. Of the 362 helices identified, 71 were 3_{10}-helices. As with peptides, the 3_{10}-helices tended to be substantially shorter (mean length = 3.3 residues) than the α-helices (mean length = 12.2 residues). Most of the 3_{10}-helices were found as connections between β-strands in all-β proteins. Although α-helices substantially outnumber 3_{10}-helices, this is not the case if the count is restricted to only short helical segments. Most of the 71 3_{10}-helices are of length 5 residues or less. The number of α-helices containing five residues or less is approximately 67 (or 87 if helices up to six residues are included). Thus, six residues appears to be the limit beyond which α-helix is clearly favored.

Helix−Coil Theory and Importance of the Ensemble

Many designed helical peptides are 16 residues or longer and often exhibit more than 70% fractional helicity (Marqusee & Baldwin, 1987; Marqusee et al., 1989). The average length of a helical domain within such a peptide is ap-

† This work was supported by NIH Grant GM 46870.

‡ Telephone, (408) 459-2176; fax, (408) 459-2935; email, glennm@hydrogen.ucsc.edu.

0006-2960/95/0434-3873$09.00/0

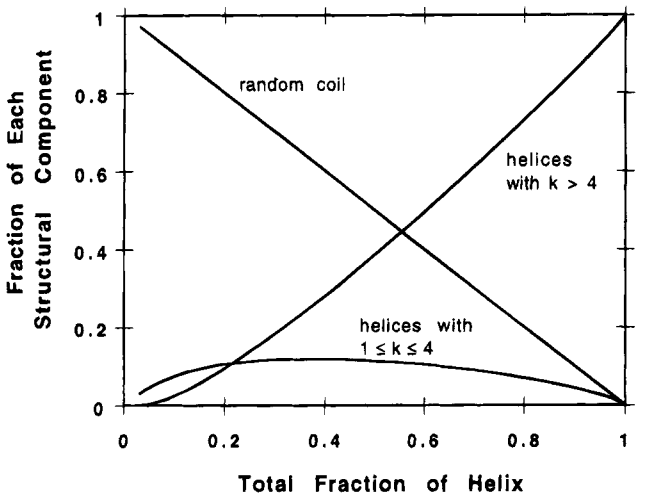

FIGURE 1: α-Helix−coil calculation based on eq 1 for a 16-residue peptide with $\sigma = 0.0029$ (Rohl et al., 1992). The populations are shown of peptides with short helical domains of six residues or less ($1 \leq k \leq 4$; k is the number of hydrogen bonds in a domain), long helical domains ($k > 4$), and random coil ($k = 0$). The short helical domains may switch conformation to 3_{10}-helix. Fractional contributions are computed with the relation

$$f(n \leq k \leq m) = \frac{\sigma}{Z}\sum_{k=n}^{m}(N_h - k + 1)s^k$$

proximately 12 residues. Since this domain length is longer that the apparent 3_{10}-/α-threshold of six, can one ignore the contribution from 3_{10}-helix? Most likely, the answer is no.

Helix−coil transition theory determines helicity from an ensemble average of helix lengths (Schellman, 1958; Zimm & Bragg, 1959; Poland & Scheraga, 1970; Qian & Schellman, 1992). To illustrate the contribution from short helices, we use the partition function for a short, α-helical homopolypeptide:

$$Z = 1 + \sigma\sum_{k=1}^{N_h}(N_h - k + 1)s^k \qquad (1)$$

where N_h is the number of α-helical hydrogen bonds in the fully helical conformation, σ is the initiation constant, and s is the helix propagation equilibrium constant (Qian & Schellman, 1992; note that eq 1 is mathematically identical to their eq 12). In an α-helix, N_h is related to the total number of residues N_r by $N_r = N_h + 2$. Each term in the sum gives the statistical weight for a helical peptide with k hydrogen bonds.

Equation 1 is used to compute peptide populations with helical domains both shorter and longer than the apparent 3_{10}-/α-threshold of six. α-Helical domains of six residues or less contain four or less hydrogen bonds. Figure 1 shows a calculation of the transition for a 16-residue peptide, by variation of s. The fractions of short domains ($1 \leq k \leq 4$), long domains ($k > 4$), and random coil ($k = 0$) are each plotted against the total fraction of helix. At low helicity, the population of short domains accounts for the most of the helices present. However, beyond a fractional helicity of 0.2, longer helix domains begin to dominate. From fractional helicity 0.2−0.8, the contribution of short helices is nearly constant. At 0.5 fractional helicity, the population of short helices accounts for approximately 20% of the total helix content.

If α-helices with $1 \leq k \leq 4$ are able to largely interconvert to 3_{10}-helices, as suggested by crystallographic studies, Figure 1 shows that the 3_{10}-helix population may be significant throughout the helix−coil transition. However, this elementary calculation probably undercounts the total quantity of 3_{10}-helices. Studies have shown that the termini of α-helices often follow a 3_{10}-conformation (Némethy et al., 1967; Barlow & Thornton, 1988), and inclusion of these mixed helices would increase the total 3_{10}-helix count. Furthermore, an exact calculation with a partition function that accounts for the differential hydrogen-bonding pattern of 3_{10}- and α-helix would require σ and s values specific for the two helix types. The helix initiation constants should be different for the two conformations with, perhaps, $\sigma(3_{10}) > \sigma(\alpha)$ since there are only two intervening pairs of ϕ and ψ dihedral angles in the first turn of 3_{10}-helix and three pairs in an α-helix. A ratio of $\sigma(3_{10})/\sigma(\alpha) > 1$ could lead to a preponderance of short 3_{10}-helices (C. L. Brooks, personal communication, manuscript submitted). Nevertheless, the calculation presented here demonstrates that the population of short helices is substantial for peptides of intermediate helix domain length. Therefore, the conformational preferences for short helices must play a role in determining average helix geometry.

CD, NMR, and ESR

Circular dichroism remains the standard technique for identifying global helical structure. The unmistakable CD signal of a right-handed helix is characterized by transitions with negative maxima at 208 and 222 nm. Total helix content is often measured by the signal strength at 222 nm. However, recent calculations suggest that 3_{10}-helix may be characterized by a weak 222 transition along with a strong 208 transition (Manning & Woody, 1991). It is noteworthy, therefore, that helical peptides often show exactly this signature [e.g., Bradley et al. (1990), Merutka et al. (1990), Scholtz et al. (1991), Storrs et al. (1992), and Zhou et al. (1994)]. In addition, partial thermal unfolding of a helical peptide often results in a preferential weakening of the CD signal at 222 nm. These results may indicate that partially unfolded helical peptides contain 3_{10}-helix. Unfortunately, such an assignment may be ambiguous because random coil also contributes to the CD signal, giving a maximum near 222 nm (Yang et al., 1986).

NMR is widely used to localize helical regions in designed peptides and peptides derived from protein sequences (Wüthrich, 1986; Wright et al., 1988). 2D nuclear Overhauser effect spectroscopy (2D NOESY) reveals geometry through NOE cross peaks. A continuous series of cross peaks revealing short NN($i,i+1$) and αN($i,i+3$) interproton distances identifies helical structure (Wüthrich, 1986). Likewise, the αβ($i,i+3$) NOE is evidence for helix structure although the distance is longer in 3_{10}-helix than in α-helix. While these short-range (<3.5 Å) NOE's demonstrate helical structure, they do not distinguish 3_{10}-helix from α-helix. However, the longer range αN($i,i+2$) and αN($i,i+4$) NOE's can distinguish the helix types. The former is present in 3_{10}-helix and the latter in α-helix.

NMR experiments on helical peptides often reveal conflicting NOE's. While the expected NN($i,i+1$) and αN($i,i+3$) NOE's are present, many peptides also reveal simultaneous αN($i,i+2$) and αN($i,i+4$) (Wright et al., 1988;

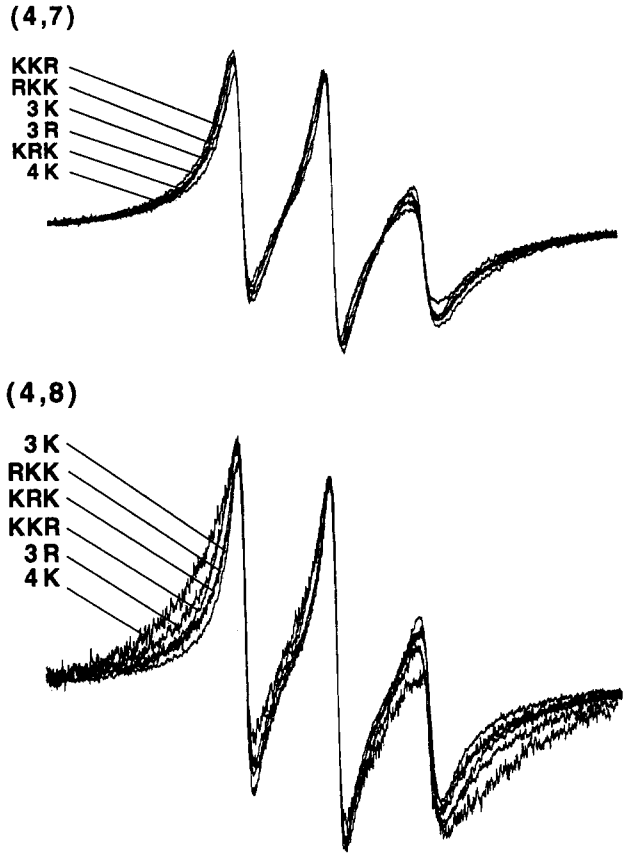

FIGURE 2: ESR spectra from (4,7) and (4,8) doubly labeled analogs of Ala-rich peptides with line widths indicating conserved $d(i,i+3)$ and variable $d(i,i+4)$, respectively, which is consistent with a coexistence of 3_{10}-helix and α-helix. Broader spectra indicate a closer distance between the label positions. The 21-residue 4K peptide, which follows the sequence Ac-$(AAAAK)_4A$-NH_2, is the most helical of the peptides studied here. The 3K and 3R peptides follow Ac-$(AAAAX)_3A$-NH_2, where X is K or R, respectively. The remaining peptides are analogous 16-mers that contain K or R at positions 5, 10, and 15, as indicated. Spectra are labeled from top to bottom in order of increasing line width.

Table 1: Distances between Side Chains[a]

distance (Å)	strand	3_{10}-helix	α-helix
$i \rightarrow i+3$	⟨15.9⟩	7.8 ± 0.7	8.6 ± 1.0
$i \rightarrow i+4$	⟨15.4⟩	12.4 ± 0.9	8.4 ± 0.7

[a] Molecular dynamics (Insight and Discover from BioSym) was used to determine the average distances and standard deviations for the two helical conformations. During the MD trajectories, the side-chain atoms were unconstrained, but the helix backbone atoms were constrained using a restoring potential to maintain the helical structure. The distance rankings hold as shown even if rigid side chains are assumed. Distances for the strand conformation (in broken brackets) were estimated on the basis of a fixed side chain geometry since the labels are too far apart to exhibit observable J coupling.

Osterhout et al., 1989; Bradley et al., 1990; Merutka et al., 1993). The latter NOE is used as confirmation of α-helix structure. Interpretation of the αN($i,i+2$) cross peak, however, is more challenging because isolated turns also give this particular NOE (Wüthrich, 1986). [In fact, 3_{10}-helix is structurally identical to a continuous series of type III β-turns (Schulz & Schirmer, 1979).] Isolated turns are an expected component of random coil structure so the conflicting αN($i,i+2$) and αN($i,i+4$) NOE's may be from a mixture of random-coil and α-helix. Alternatively, these NOE's may arise from a mixture of 3_{10}-helix and α-helix.

It appears that CD and NMR are challenged by mixtures of 3_{10}-helix and α-helix. In order to add to the information available, we initiated studies using doubly spin labeled peptides. In double label ESR, spin labels are placed $i+3$ or $i+4$ positions apart in a pair of companion peptides (Miick et al., 1992; Fiori et al., 1993). The spin labels interact largely by the biradical mechanism of through-space J coupling. Similar to the through-bond J coupling responsible for NMR multiplets, this static interaction does not depend strongly on the rotational correlation time. Thus, double label ESR allows for the ranking of distances within a peptide and comparison of analogous distances among peptides of different molecular weights. 3_{10}- and α-helix are distinguished by the relative $i \rightarrow i+3$ and $i \rightarrow i+4$ distances with $d(i,i+3) \geq d(i,i+4)$ for α-helix and $d(i,i+3) < d(i,i+4)$ for 3_{10}-helix. These rankings are detected by line broadening originating from J coupling and, perhaps, other biradical mechanisms.

We have studied peptides based on the designs from R. L. Baldwin's laboratory (Marqusee et al., 1989) in order to examine how length and sequence influence the 3_{10}-helix → α-helix equilibrium. One of the common findings is that the N-termini of these peptides are in a well-structured helical conformation. Figure 2 shows the ESR spectra from the (4,7) and (4,8) doubly labeled analogs from several sequences. Biradical line broadening is present in all the spectra, and by overlapping the spectra, an interesting pattern emerges. The line shapes of all the (4,7) analogs are nearly identical, indicating that $d(4,7)$ is approximately constant among these peptides. However, the line shapes of the (4,8) peptides show variation. For example, the 3K-(4,8) gives a spectrum narrower than the companion 3K-(4,7) spectrum but the 4K-(4,8) spectrum, is much broader than the 4K-(4,7) spectrum.

Do these ESR results arise from a mixture of random coil and α-helix or, alternatively, a mixture of 3_{10}-helix and α-helix? Table 1 provides the necessary distances for interpreting these data. Comparison of 3_{10}-helix and α-helix shows that the $i \rightarrow i+3$ distance between side chains is conserved in the two conformations to within less than 1.0 Å. However, the $i \rightarrow i+4$ distance shows considerable variation. Thus, ESR distance rankings from a collection of doubly labeled peptides with varying amounts of 3_{10}- and α-helix, and little β-strand, should reveal line-shape variation from only the $i+4$ and not the $i+3$ doubly labeled peptides. This is consistent with our observations. It is noteworthy that peptides exhibiting broader (4,8) spectra yield stonger helical CD signals (Fiori et al., 1993, 1994), which suggests that 3_{10}-helix is significantly populated only in partially helical Ala-based peptides (as expected from Figure 1).

Double label ESR measurements on the Ala-rich sequences do not contradict CD or NMR studies on analogous peptides. In fact, these ESR experiments provide two new distance constraints that aid in the interpretation of NMR spectra. Although signal overlap inhibits NMR studies of the exact Ala-rich sequences presented here, NMR experiments on analogous peptides have revealed αN($i,i+2$) NOE's [e.g., Osterhout et al. (1989) and Merutka et al. (1993)]. Conservation of the $i \rightarrow i+3$ distance for the Ala-rich peptides suggests that 3_{10}-helix is a significantly populated conformation and may be responsible for the often observed αN($i,i+2$) NOE's.

Nascent helix has emerged as a key concept for explaining the NMR of partially folded peptides (Dyson et al., 1988).

	random coil	↔ nascent helix	↔ 3_{10}-helix	↔ α-helix
Hydrogen Bonding	none	transient $i \rightarrow i+3$	$i \rightarrow i+3$	$i \rightarrow i+4$
CD ($-[\theta]_{222}$ in deg cm^2 dmol^{-1})	< 0	induced by TFE to ≤15,000	15,000 to 30,000 $-[\theta]_{222} \leq -[\theta]_{208}$	≥30,000 $-[\theta]_{222} \geq -[\theta]_{208}$
NMR (NOE)	αN(i,i+1)	NN(i,i+1) αN(i,i+2)	NN(i,i+1) αN(i,i+2) αN(i,i+3)	NN(i,i+1) αN(i,i+3) αN(i,i+4) αβ(i,i+3)
ESR (line broadening)	weak (i,i+3) weak (i,i+4)	weak (i,i+3) weak (i,i+4)	**strong** (i,i+3) weak (i,i+4)	**strong** (i,i+3) **strong** (i,i+4)

FIGURE 3: A proposal for the thermodynamic folding pathway of helical peptides. Reconciling various experimental techniques suggests that 3_{10}-helix is a thermodynamic intermedate between nascent helix and α-helix. This scheme suggests that highly helical peptides exist as mixtures of 3_{10}- and α-helix with proportions that depend on the total helix content.

Sequences derived from helical regions of proteins often have weak CD signals but can give a series of NN($i,i+1$) and αN($i,i+2$) NOE's [but no αN($i,i+3$) NOE's]. This behavior is explained by postulating an interconverting ensemble of unfolded strand-like structures each containing perhaps a single turn. Mixtures of α-helix and nascent helix have been used to explain the observation of αN($i,i+2$) NOE's in some helical peptides. Would such an explanation apply to the Ala-rich peptides? Because nascent helix involves extended strands, the average $i \rightarrow i+3$ side-chain distance is not equivalent to that found in helices. Therefore, mixing various proportions of nascent helix and α-helix will result in $i \rightarrow i+3$ distance variations, which is not observed in the peptides here.

A Proposal for the Thermodynamic Folding Pathway of Helical Peptides

Combining the double label ESR data with observations from CD and NMR suggests that 3_{10}-helix is an important thermodynamic intermediate between nascent helix and α-helix. This view is consistent with crystallographic findings and the expected populations of short helices from helix−coil theory. Furthermore, molecular dynamics studies suggest that 3_{10}-helix exists as a kinetic intermediate along the α-helix unfolding pathway (Soman et al., 1991; Tirado-Rives & Jorgensen, 1991), and free energy calculations have revealed a distinct 3_{10}-minimum (Tobias & Brooks, 1991). Thus, single $i \rightarrow i+3$ hydrogen bonds are readily formed and may propagate, giving short stretches of 3_{10}-helix. However, calculations have also suggested that a purely 3_{10}-helical structure may be unfavored in longer Ala peptides presumably due to a buildup of unfavorable intramolecular steric and electrostatic interactions (Tirado-Rives et al., 1993).

A proposal for incorporation of 3_{10}-helix into the general α-helix folding scheme is shown in Figure 3. Below each helix conformation are listed various spectroscopic signatures. The CD assignments are estimates based on comparison between double label ESR and CD. Clearly, a peptide exhibiting the observables of a particular conformation will not exist solely in that conformation. Nearby species are also populated. Thus, a peptide giving $-[\theta]_{222}$ between 15 000 and 30 000 deg cm^2 dmol^{-1} probably exists as a mixture of nascent helix, 3_{10}-helix, and α-helix.

An appealing feature of the proposal is the nature in which hydrogen bonding develops along the pathway. Nascent helix contains a mixture of largely unfolded conformers with transient localized turns containing characteristic $i \rightarrow i+3$ hydrogen bonds. As helical structure becomes more favored, this hydrogen-bonding pattern propagates, giving sequential $i \rightarrow i+3$ bonds, i.e., a 3_{10}-helix. As helical domains lengthen further, the strain of the 3_{10}-conformation builds and the α-helix becomes the most stable species with, perhaps, some remaining 3_{10}-type hydrogen bonds at the α-helix termini. This thermodynamic scheme provides a basis for understanding the mixed conformations often observed in helical peptides and suggests that local 3_{10}-helix → α-helix switching often observed in proteins is consistent with peptide models.

ACKNOWLEDGMENT

I thank C. L. Brooks, G. Montelione, S. Mronga, and A. Kuki for providing many helpful comments on the manuscript. Special acknowledgment is due to C. Stenland for improving our methodologies of peptide labeling and purification and for repeating many of the measurements presented here. Important insights were also provided by M. P. Hanson, K. Lundberg, G. Martinez, L. Howe, J. Trulson, T. Block, and D. Thompson. M. Rabenstein and Y. K. Shin are gratefully acknowledged for suggestions regarding peptide purification and line-shape analysis.

REFERENCES

Barford, D., & Johnson, L. N. (1989) *Nature 340*, 609−616.
Barlow, D. J., & Thornton, J. M. (1988) *J. Mol. Biol. 201*, 601−619.

Bradley, E. K., Thomason, J. F., Cohen, F. E., Kosen, P. A., & Kuntz, I. D. (1990) *J. Mol. Biol. 215*, 607−622.
Dyson, H. J., Rance, M., Houghten, R. A., Wright, P. E., & Lerner, R. A. (1988) *J. Mol. Biol. 201*, 201−217.
Fiori, W. R., Miick, S. M., & Millhauser, G. L. (1993) *Biochemistry 32*, 11957−11962.
Fiori, W. R., Lundberg, K. M., & Millhauser, G. L. (1994) *Nature Struct. Biol. 1*, 374−377.
Gerstein, M., & Chothia, C. (1991) *J. Mol. Biol. 220*, 133−149.
Karle, I. L., & Balaram, P. (1990) *Biochemistry 29*, 6747−6756.
Kavanaugh, J. S., Moopenn, W. F., & Arnone, A. (1993) *Biochemistry 32*, 2509−2513.
Kostrikis, L. G., Liu, D. J., & Day, L. A. (1994) *Biochemistry 33*, 1694−1703.
Liu, D. J., & Day, L. A. (1994) *Science 265*, 671−674.
Manning, M. C., & Woody, R. W. (1991) *Biopolymers 31*, 569−586.
Marqusee, S., & Baldwin, R. L. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8898−8902.
Marqusee, S., Robbins, V. H., & Baldwin, R. L. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5286−5290.
McPhalen, C. A., Vincent, M. G., Picot, D., Jansonius, J. N., Lesk, A. M., & Chothia, C. (1992) *J. Mol. Biol. 227*, 197−213.
Merutka, G., Lipton, W., Shalongo, W., Park, S. H., & Stellwagen, E. (1990) *Biochemistry 29*, 7511−7515.
Merutka, G., Morikis, D., Brüschweiler, R., & Wright, P. E. (1993) *Biochemistry 32*, 13089−13097.
Miick, S. M., Martinez, G. V., Fiori, W. R., Todd, A. P., & Millhauser, G. L. (1992) *Nature 359*, 653−655.
Némethy, G., Phillips, D. C., Leach, S. J., & Scheraga, H. A. (1967) *Nature 214*, 363−365.
Osterhout, J., Jr., Baldwin, R. L., York, E. J., Stewart, J. M., Dyson, H. J., & Wright, P. E. (1989) *Biochemistry 28*, 7059−7064.
Poland, D., & Scheraga, H. A. (1970) *Theory of helix-coil transitions in biopolymers*, Academic Press, New York.
Qian, H., & Schellman, J. A. (1992) *J. Phys. Chem. 96*, 3987−3994.
Rohl, C. A., Scholtz, J. M., York, E. J., Stewart, J. M., & Baldwin, R. L. (1992) *Biochemistry 31*, 1263−1269.
Schellman, J. A. (1958) *J. Phys. Chem. 62*, 1485−1494.
Scholtz, J. M., Hong, Q., York, E. J., Stewart, J. M., & Baldwin, R. L. (1991) *Biopolymers 31*, 1463−1470.
Schulz, G. E., & Schirmer, R. H. (1979) *Principles of Protein Structure*, Springer, New York.
Soman, K. V., Karimi, A., & Case, D. A. (1991) *Biopolymers 31*, 1351−1361.
Storrs, R. W., Truckses, D., & Wemmer, D. E. (1992) *Biopolymers 32*, 1695−1702.
Tirado-Rives, J., & Jorgensen, W. L. (1991) *Biochemistry 30*, 3864−3871.
Tirado-Rives, J., Maxwell, D. S., & Jorgensen, W. L. (1993) *J. Am. Chem. Soc. 115*, 11590−11593.
Tobias, D. J., & Brooks, C. L. (1991) *Biochemistry 30*, 6059−6070.
Toniolo, C., & Benedetti, E. (1991) *Trends Biochem. Sci. 16*, 350−353.
Wright, P. E., Dyson, H. J., & Lerner, R. A. (1988) *Biochemistry 27*, 7167−7175.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley and Sons, Inc., New York.
Yang, J. T., Wu, C.-S. C., & Martinez, H. M. (1986) *Methods Enzymol. 130*, 208−269.
Zhou, H. X. X., Lyu, P. C., Wemmer, D. E., & Kallenbach, N. R. (1994) *Proteins 18*, 1−7.
Zimm, B. H., & Bragg, J. K. (1959) *J. Chem. Phys. 31*, 526−535.

BI942791L