Determination of α -helix N1 energies after addition of N1, N2, and N3 preferences to helix/coil theory

JIA KE SUN, SIMON PENEL, AND ANDREW J. DOIG

Department of Biomolecular Sciences, UMIST, P.O. Box 88, Manchester M60 1QD, United Kingdom

(RECEIVED November 29, 1999; FINAL REVISION January 21, 2000; ACCEPTED January 28, 2000)

Abstract

Surveys of protein crystal structures have revealed that amino acids show unique structural preferences for the N1, N2, and N3 positions in the first turn of the α -helix. We have therefore extended helix-coil theory to include statistical weights for these locations. The helix content of a peptide in this model is a function of N-cap, C-cap, N1, N2, N3, C1, and helix interior (N4 to C2) preferences. The partition function for the system is calculated using a matrix incorporating the weights of the fourth residue in a hexamer of amino acids and is implemented using a FORTRAN program. We have applied the model to calculate the N1 preferences of Gln, Val, Ile, Ala, Met, Pro, Leu, Thr, Gly, Ser, and Asn, using our previous data on helix contents of peptides Ac-XAKAAAAKAAGY-CONH₂. We find that Ala has the highest preference for the N1 position. Asn is the most unfavorable, destabilizing a helix at N1 by at least 1.4 kcal mol⁻¹ compared to Ala. The remaining amino acids all have similar preferences, 0.5 kcal mol⁻¹ less than Ala. Gln, Asn, and Ser, therefore, do not stabilize the helix when at N1.

Keywords: α -helix; helix-coil theory; hydrogen bonding; N-terminus; protein structure

The residues at the N-terminus of the α -helix are called N'-N-cap-N1-N2-N3-N4, etc., where the N-cap is the residue with nonhelical ϕ, ψ angles immediately preceding the N-terminus of an α -helix and N1 is the first residue with helical ϕ, ψ angles (Richardson & Richardson, 1988). The N1, N2, and N3 residues are unique because their amide NH groups do not participate in $i, i + 4$ backbone-backbone hydrogen bonds. The N-cap residue is now well understood in terms of its effect on the stability and structure of the helix (Serrano & Fersht, 1989; Lecomte & Moore, 1991; Bell et al., 1992; Serrano et al., 1992; Chakrabartty et al., 1993; Forood et al., 1993; Lyu et al., 1993; Yumoto et al., 1993; Doig et al., 1994, 1997; Doig & Baldwin, 1995).

Davies (1964) first showed that some amino acids occur more often than others in the helix. Further examination of amino acid preferences in the α -helix revealed a greater complexity, because the frequency of amino acid depends upon its position in the helix. In particular, the helix N-terminus shows significantly different residue frequencies for the N-cap, N1, N2, N3, and helix interior positions (Argos & Palau, 1982; Richardson & Richardson, 1988; Kumar & Bansal, 1998). Penel et al. (1999) made a detailed survey of the structures adopted by the amino acids at N1, N2, and N3 and identified many new structural features. They found that the most significant structural preferences can be explained by short range hydrogen bonding to the free N-terminal NH groups, as foreseen by Presta and Rose (1988), with the strongest trends being the N2

amino acids Gln, Glu, Asp, Asn, Ser, Thr, and His preferentially forming i, i or $i, i + 1$ hydrogen bonds to the backbone.

Experimental studies on helical peptides support the uniqueness of the N1, N2, and N3 locations. Chakrabartty et al. (1993) synthesized helical peptides beginning Ac-XAKAA. In the most populated helical conformations, Ac is the N-cap, so the following amino acids X, A, and K are therefore usually the N1, N2, and N3 residues. They varied X (N1) to observe its effect on helix content and found that Ala was favored over 10 small polar and nonpolar amino acids. While these results demonstrate qualitatively that Ala is favored, it was not possible to quantify this preference in the form of a free energy. Petukhov et al. (1998, 1999) synthesized a 16-residue peptide helical peptide and analyzed the behavior of small residues (Val, Ile, Leu, Met, Gly, Ser, Thr, Asn, and Gln) at the N-cap, N1, N2, and N3 positions and at a central helical position, also finding significant differences.

There is therefore a need to measure the free energies of substituting each amino acid at the N1, N2, and N3 position in the α -helix. A straightforward experimental method to achieve this is to synthesize helical peptides where the N1, N2, and N3 positions are varied. A helical peptide in solution forms a complex mixture of fully helical, fully unfolded (coil), and, most often, partially helical conformations, typically with frayed ends. To extract free energies for making substitutions at N1, N2, and N3, it is essential to include parameters for N1, N2, and N3 preferences; current Lifson-Roig based helix/coil models do not do this and implicitly assume that the N2 and N3 positions are identical to the helix interior and that the N1 position is identical to C1, by assigning them the same parameter. In this paper, we extend helix/coil theory to include parameters for N1, N2, and N3. We apply the model to previously published experimental data to give free energies for substitution at N1.

Reprint requests to: Dr. Andrew J. Doig, Biomolecular Sciences, UMIST, P.O. Box 88, Manchester M60 1QD, United Kingdom; e-mail: Andrew.Doig@umist.ac.uk.

Theory

This work builds upon previous Lifson–Roig based models for the α -helix/coil equilibrium (Lifson & Roig, 1961; Doig et al., 1994; Rohl et al., 1996). The partition function for a partially helical peptide is a function of the chain length and the statistical weights of each residue in the chain describing helix initiation, propagation, and capping. The chain length N is the number of residues, where a residue is defined as a C_α with peptide bonds on both sides. Since the N- and C-terminal C_α atoms are not residues according to this definition, a peptide with N residues is $N + 2$ units long. Each residue in a given polypeptide conformation is classified as being in either the α -helical (h) or coil (c) conformation as defined by its backbone (ϕ, ψ) dihedral angles. The N- and C-terminal units, which are not residues, cannot be constrained to the helical (ϕ, ψ) space and are in the c conformation by definition.

The coil conformation flanked by two coil neighbors is defined as the reference state and assigned a statistical weight of unity. All other statistical weights are relative to this reference. In the original Lifson–Roig model, only two nonunity statistical weights are assigned. Nucleating residues, assigned weight v , are residues in the h conformation with one or more of the nearest neighbors in the c conformation. Propagating residues, assigned weight w , are those in the h conformation whose nearest N- and C-terminal neighbors are also in the h conformation. These weights and assignments lead to a physical meaning for the statistical weights. v is the equilibrium constant, relative to a nonhelical conformation, for constraining a residue to the helical region of the Ramachandran map. w is the equilibrium constant for constraining a residue to the helical region of the Ramachandran map when it is part of an α -helix. Each helix propagating residue assigned weight w is consequently associated with an $i, i + 4$ main-chain–main-chain hydrogen bond. The original Lifson–Roig model has been previously modified to include N- and C-capping by defining new statistical weights for the coil residues adjacent to helical neighbors (Doig et al., 1994). Residues in the c state with a h state C-terminal neighbor are assigned weight n , and those with an N-terminal helical neighbor are assigned weight c . Rohl et al. (1996) modified the use of the N-cap parameter by assigning it only when the N-cap precedes three or more consecutive helical residues. This is because the N-cap energy usually arises by an interaction between the N-cap side chain and the N3 backbone NH group (Presta & Rose, 1988; Doig et al., 1997). Side-chain interactions (Stapley

et al., 1995) and 3_{10} -helices (Rohl & Doig, 1996; Sun & Doig, 1998) have also been added to Lifson–Roig helix/coil models. These are not considered here as they are not present in the peptides that we analyze.

We have added weights for the N1, N2, and N3 ($n1$, $n2$, and $n3$) positions in the α -helix as follows: The $n1$ -value is assigned to a helical residue immediately following a coil residue. In previous helix/coil models, residues in this location, and in the final helical position (C1), were assigned the helix initiation parameter v . The new model therefore distinguishes between N1 and C1. v -Values are always low (≈ 0.04) as they take into account the energetic cost of localizing a residue to helical (ϕ, ψ) space without the benefit of hydrogen bond formation. $n1$ -Values are therefore also expected to be small. The penalty for helix initiation is now $n1 \cdot v$, instead of v^2 . An N2 helical residue is assigned a weight of $n2 \cdot w$, instead of w . The weight w is maintained to keep the useful definition of the number of residues with a w weighting being equal to the number of residues with an $i, i + 4$ main-chain–main-chain hydrogen bond. The $n2$ -value is an adjustment to the weight of an N2 residue that takes into account the structures that can be adopted by side chains uniquely at this position. Similarly, an N3 residue is now assigned the weight $n3 \cdot w$, instead of w . Residues in the centre of the rare hch conformation have a weight of 1.

The assignment of weights to very short helical segments requires careful consideration of the structures that can be adopted. An isolated helical residue, in the center of a chc conformation, is assigned a weight $n1$, rather than v . This is because the most important structures adopted by residues at N1 are hydrogen bonds between the N1 side chain and the N1 backbone NH group (Penel et al., 1999). These interactions can form if only a single residue is in an h conformation. All N1 residues commencing any helical segment have the weight $n1$. The N2 residue in a chhc sequence is assigned the weight $n2 \cdot v$ for the same reason: N2 structures are typically hydrogen bonds between the N2 side chain and the N2 backbone NH group (Penel et al., 1999). These can therefore form even if only two consecutive helical residues are present. The N3 residue in a chhch sequence has the weight $n3 \cdot v \cdot n$ since both a capping box interaction between the N3 side chain and the N-cap backbone (assigned $n3$) and the N-cap side chain to N3 backbone interaction (assigned n) require a minimum of three consecutive helical residues (Harper & Rose, 1993). A weight of v is assigned rather than w as the final helical residue in a sequence does not have an associated hydrogen bond. Figure 1 shows these assign-

Fig. 1. Assignment of statistical weights for residues in the α -helix.

ments for a short helix. To define the weight of a residue, it is necessary to consider the conformation of six consecutive residues. Table 1 lists all possible conformations and the weight of the fourth residue in the hexamer. Note that the weight is for the amino acid found at the fourth position with the exception of n , which is the weight of the first residue in the hexamer.

All the properties of an equilibrium are present within its partition function. Partition functions in helix/coil theory are efficiently calculated using matrices. These weights can be written in the form of a 32×32 matrix (not shown), which is simplified by combining identical columns to a 16×16 matrix (M):

					$i+2$	(hUc)	(hUc)	(hUc)	(hUc)	(hUc)	(hUc)	(hUc)	(hUc)	(hUc)	(hUc)	(hUc)	(hUc)	(hUc)		
					$i+1$	c	h	c	h	c	h	c	h	c	h	c	h	c	h	
					i	c	c	h	h	c	c	h	h	c	c	h	h	c	h	
					$i-1$	c	c	c	c	h	h	h	c	c	c	c	h	h	h	
$i-3$	$i-2$	$i-1$	i	$i+1 \setminus i-2$	c	c	c	c	c	c	c	c	h	h	h	h	h	h	h	
c	c	c	c	(hUc)	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
c	c	c	h	(hUc)	0	0	$n1$	$n1$	0	0	0	0	0	0	0	0	0	0	0	
c	c	h	c	(hUc)	0	0	0	0	c	c	0	0	0	0	0	0	0	0	0	
c	c	h	h	(hUc)	0	0	0	0	0	0	$n2 \ v$	$n2 \ w$	0	0	0	0	0	0	0	
c	h	c	c	(hUc)	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	
c	h	c	h	(hUc)	0	0	0	0	0	0	0	0	0	0	$n1$	$n1$	0	0	0	
c	h	h	c	(hUc)	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
c	h	h	h	(hUc)	0	0	0	0	0	0	0	0	0	0	0	0	$n_{i-3}n3 \ v$	$n_{i-3}n3 \ v$	$n_{i-3}n3 \ w$	$n_{i-3}n3 \ w$
h	c	c	c	(hUc)	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
h	c	c	h	(hUc)	0	0	$n1$	$n1$	0	0	0	0	0	0	0	0	0	0	0	0
h	c	h	c	(hUc)	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
h	c	h	h	(hUc)	0	0	0	0	0	0	$n2 \ v$	$n2 \ w$	0	0	0	0	0	0	0	0
h	h	c	c	(hUc)	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
h	h	c	h	(hUc)	0	0	0	0	0	0	0	0	0	0	$n1$	$n1$	0	0	0	0
h	h	h	c	(hUc)	0	0	0	0	0	0	0	0	0	0	0	0	c	1	0	0
h	h	h	h	(hUc)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	v	w

The partition function of the α -helix with N amino acids (Z) is given by $Z = U \cdot M \cdot V$, where

$$U = (1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0)$$

$$V = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

All the properties of a system at equilibrium are contained within the partition function. In particular, the mean number of residues with a weight x is given by $(\partial \ln Z)/(\partial \ln x)$. Circular dichroism is commonly used to give the mean helix content of a helical peptide, namely the fraction of residues that have a weight of w . Our model can thus be related to experimental data by equating the measured mean helix content to $[(\partial \ln Z)/(\partial \ln w)]/N$, where N is the number of residues in the peptide.

Results

Here we demonstrate the utility of the model by applying it to experimental data to determine $n1$ -values and hence the free energy changes for making substitutions at the N1 position of the α -helix. Our previous work (Chakrabartty et al., 1993) measured the helix content of peptides with sequence Ac-XAKAAAKAAGY-CONH₂, where X = Gln, Val, Ile, Ala, Met, Pro, Leu, Thr, Gly, Ser, or Asn, though we did not interpret these results quantitatively, in the absence of suitable theory. Table 2 shows the ellipticities at 222 nm and mean helix contents of these peptides at pH 9.55.

Previous Lifson–Roig based helix-coil models have assigned the weight v to both the N1 and C1 positions in the helix and have found a value of 0.036 for v (Rohl et al., 1996). Now we retain this value for the C1 position, but allow the N1 statistical weight ($n1$) to vary. In fitting the experimental data, we use the w -, n - and c -values of Rohl et al. and set all $n2$ - and $n3$ -values to 1, leaving the only unknown to be $n1$. We assume that $n1(K) = v = 0.036$. This is an excellent assumption as the population of conformations in this sequence that have Lys as the N1 residue is very low. Varying $n1(K)$ therefore makes almost no difference to the calculated helix content. First, we determined $n1(\text{Ala})$ by finding the value of $n1(\text{Ala})$ that gives a calculated helix content of the peptide Ac-XAKAAAKAAGY-CONH₂ in agreement with a experimentally measured helix content of 27.1%. This give $n1(\text{Ala}) = 0.0231$, a similar value to v , as expected. We then retain this value and determine the $n1$ -values of the other amino acids in the same way, giving the results shown in Table 2. Experimental error in helix contents is $\pm 3\%$ (Stapley et al., 1995). The errors in the $n1$ -values were therefore determined by adding or subtracting 3% to the helix contents and refitting. Petukhov et al. (1998, 1999) also studied a series of helical peptides with nonpolar groups at N1, N2, and N3. Free energies for substituting these amino acids were found by analyzing the experimental results with AGADIR (Lacroix et al., 1998) and by molecular modeling. We are unable to refit their data

Table 1. Assignment of statistical weights

Conformation	Weight of the fourth residue	Conformation	Weight of the fourth residue
cccccc	1	hcccc	1
cccch	1	hcccch	1
ccch	1	hccch	1
ccchh	1	hccchh	1
ccchc	$n1$	hcchc	$n1$
ccchh	$n1$	hcchh	$n1$
ccchhc	$n1$	hcchhc	$n1$
ccchhh	$n1$	hcchhh	$n1$
cchccc	1	hcchcc	1
cchcch	1	hcchch	1
cchchc	1	hcchhc	1
cchchh	1	hcchhh	1
cchhcc	$n2 \cdot v$	hchcc	$n2 \cdot v$
cchhch	$n2 \cdot v$	hchch	$n2 \cdot v$
cchhcc	$n2 \cdot w$	hchchc	$n2 \cdot w$
cchhhh	$n2 \cdot w$	hchhhh	$n2 \cdot w$
chcccc	1	hhcccc	1
chcccch	1	hhcccch	1
chccch	1	hhccch	1
chccchh	1	hhccchh	1
chccchc	$n1$	hhcchc	$n1$
chccchh	$n1$	hhcchh	$n1$
chchcc	$n1$	hhchcc	$n1$
chchch	$n1$	hhchch	$n1$
chchhc	$n1$	hhchhc	$n1$
chchhh	$n1$	hhchhh	$n1$
chhccc	1	hhhccc	c
chhcch	1	hhhcch	c
chhchc	1	hhhchc	1
chhchh	1	hhhchh	1
chhhcc	$n_{i-3} \cdot n3 \cdot v$	hhhhcc	v
chhhch	$n_{i-3} \cdot n3 \cdot v$	hhhhch	v
chhhcc	$n_{i-3} \cdot n3 \cdot w$	hhhhhc	w
chhhhh	$n_{i-3} \cdot n3 \cdot w$	hhhhhh	w

with our model, however, as their sequences include an $i, i + 4$ Arg-Arg interaction that has not been measured accurately.

We find that within this set of 11 amino acids, Ala has the highest preference for the N1 position. Asn is clearly the most unfavorable, though the errors in determining its $n1$ -value are very high. We can conclude that Asn destabilizes a helix at N1 by at least 1.4 kcal mol⁻¹ compared to Ala. This is in agreement with surveys of protein crystal structures that find Asn to have among the lowest propensity for N1 of all 20 amino acids (Argos & Palau, 1982; Richardson & Richardson, 1988; Kumar & Bansal, 1998; Penel et al., 1999). The remaining nine amino acids have the same $n1$ -value within experimental error and are significantly less stable than Ala by ≈ 0.5 kcal mol⁻¹. These conclusions are not apparent from examining the helix contents of the peptides, illustrating the necessity of using a helix-coil model to determine helix parameters. For example, while the peptide with X=Ala indeed has the highest helix content, the peptide with X=Asn has a similar helix content to all the others. The reason these data lead us to conclude that Asn has a low N1 preference is that Asn has the highest N-cap preference (Doig & Baldwin, 1995). This means that conformations with Asn at the N-cap will be strongly favored. To only give a moderate mean helix content, conformations with Acetyl as the N-cap and Asn at N1 must be strongly disfavored. The wide variation in N1 preferences that we find from Ala to Asn demonstrates that the derivation of a new model to include N1 is necessary and the previous assumption that the N1 preference of all amino acids is identical is incorrect. Our results agree well with Petukhov et al. (Table 2), except our N1(Asn) preference is rather lower and our N1(Met) preferences is slightly higher.

Ala has the highest preference for interior positions of the α -helix. This has been attributed to the fact that it has no side chain that loses conformational entropy when restricted in a helix (Creamer & Rose, 1992). The results here suggest that a similar effect is present at N1. Our previous work has shown that Gln, Asn, and Ser

Table 2. Determination of $n1$ -values by fitting experimental results of Ac-XAKAAAAKAAGY-CONH₂

Residue	θ_{222}	% Helix ^b	$n1$	Minimum value of $n1$ ^c		Maximum value of $n1$ ^e		$\Delta G(N1)^f$	Range of
$\Delta G(N1)^a$	$\Delta \Delta G(N1)$	Relative to Ala		$\Delta \Delta G(N1)$ (relative to Ala) ^d					
Gln	-5,950	18.2	0.0120	0.0086	0.0156	2.4	2.3 to 2.6	0.4	0.32
Val	-6,360	19.5	0.0119	0.0084	0.0157	2.4	2.3 to 2.6	0.4	
Ile	-6,350	19.4	0.0106	0.0071	0.0144	2.5	2.3 to 2.7	0.5	0.41
Ala	-8,850	27.1	0.0231	0.0177	0.0283	2.0	1.9 to 2.2	0	0
Met	-6,050	18.5	0.0100	0.0067	0.0137	2.5	2.3 to 2.7	0.5	0.06
Pro	-6,290	19.2	0.0108	0.0073	0.0146	2.5	2.3 to 2.7	0.5	
Leu	-6,490	19.9	0.0103	0.0067	0.0142	2.5	2.3 to 2.7	0.5	0.21
Thr	-6,640	20.3	0.0104	0.0069	0.0144	2.5	2.3 to 2.7	0.5	0.50
Gly	-6,450	19.7	0.0080	0.0045	0.0120	2.6	2.4 to 2.9	0.6	0.66
Ser	-7,740	23.4	0.0114	0.0074	0.0159	2.4	2.2 to 2.7	0.4	0.45
Asn	-6,680	20.4	0.0018	<0 ^g	0.0059	3.4	2.8 to ∞	1.4	0.60

^aRange is ΔG (minimum value of $n1$) to ΔG (maximum value of $n1$).

^bUsing θ_{222} of 100% helix = -42,500 (1 - x/N), where $x = 3$ and $N = 12$ (Rohl et al., 1996).

^cFound by refitting $n1$ after subtracting 3% from helix content.

^dPetukhov et al. (1998, 1999). Errors are ± 0.2 kcal mol⁻¹.

^eFound by refitting $n1$ after adding 3% to helix content.

^fCalculated as $\Delta G(N1) = -RT \ln(n1)$, where $T = 273.1$ K.

^gIt was impossible to fit a positive value of $n1$ here. A negative statistical weight is meaningless as this is equivalent to an imaginary free energy.

can form weak hydrogen bonds to free NH groups at the N-terminus of the helix (Penel et al., 1999). Despite this, Gln, Asn, and Ser do not stabilize the helix when at N1 compared to Ala. Indeed, Asn at N1 is worse than a nonpolar group that forms no hydrogen bonds at all. This may be a result of Asn stabilizing a nonhelical conformation. This is in contrast to the adjacent N-cap position, where the side chains with highest preference hydrogen bond to the backbone NH groups of N2 and N3. This is likely to be because N1 hydrogen bonds with these side chains are far from being linear (Penel et al., 1999), unlike N-cap hydrogen bonds. This suggests that hydrogen bonding can make a substantial contribution to protein stability, but only if the bond geometry is close to optimal.

Conclusion

Inspired by empirical data, we have now extended the Lifson–Roig helix-coil model to include capping, side-chain interactions, 3_{10} -helices, and N1, N2, and N3 preferences. These models are essential if helix formation in peptides is to be put on a quantitative basis, such as extracting free energies from experimental data and in predicting the helix contents of unknown peptides. We have shown that N1 preferences vary significantly between different amino acids and that Ala is the most favorable and Asn the least, out of 11 small nonpolar, nonaromatic amino acids. The design of helices and the rational modification of protein helices at N1 could take these results into account.

Materials and methods

A FORTRAN program implementing the N1N2N3 helix-coil model is freely available at <http://www.bi.umist.ac.uk/users/mjfajdg/HC.htm>.

Acknowledgments

This work was funded by the Wellcome Trust (Grant number 051374). J.K.S. thanks Hulme Hall for a Student Scholarship.

References

- Argos P, Palau J. 1982. Amino acid distribution in protein secondary structures. *Int J Pept Prot Res* 19:380–393.
- Bell JA, Becktel WJ, Sauer U, Baase WA, Matthews BW. 1992. Dissection of helix capping in T4 lysozyme by structural and thermodynamic analysis of six amino acid substitutions at Thr 59. *Biochemistry* 31:3590–3596.
- Chakrabartty A, Doig AJ, Baldwin RL. 1993. Helix capping propensities in peptides parallel those in proteins. *Proc Nat Acad Sci USA* 90:11332–11336.
- Cremer TP, Rose GD. 1992. Side-chain entropy opposes α -helix formation but rationalizes experimentally determined helix-forming propensities. *Proc Nat Acad Sci USA* 89:5937–5941.
- Davies DR. 1964. A correlation between amino acid composition and protein structure. *J Mol Biol* 9:605–609.
- Doig AJ, Baldwin RL. 1995. N- and C-capping preferences for all 20 amino acids in α -helical peptides. *Protein Sci* 4:1325–1336.
- Doig AJ, Chakrabartty A, Klingler TM, Baldwin RL. 1994. Determination of free energies of N-capping in α -helices by modification of the Lifson–Roig helix-coil theory to include N- and C-capping. *Biochemistry* 33:3396–3403.
- Doig AJ, MacArthur MW, Stapley BJ, Thornton JM. 1997. Structures of N-termini of helices in proteins. *Protein Sci* 6:147–155.
- Forood B, Feliciano EJ, Nambiar KP. 1993. Stabilization of α -helical structures in short peptides via end capping. *Proc Nat Acad Sci USA* 90:838–842.
- Harper ET, Rose GD. 1993. Helix stop signals in proteins and peptides: The capping box. *Biochemistry* 32:7605–7609.
- Kumar S, Bansal M. 1998. Dissecting α -helices: Position-specific analysis of α -helices in globular proteins. *Proteins Struct Funct Genet* 31:460–476.
- Lacroix E, Viguera AR, Serrano L. 1998. Elucidating the folding problem of α -helices: Local motifs, long-range electrostatics, ionic-strength dependence and prediction of NMR parameters. *J Mol Biol* 284:173–191.
- Lecomte JTT, Moore CD. 1991. Helix formation in apocytochrome-B5—The role of a neutral histidine at the N-cap position. *J Am Chem Soc* 113:9663–9665.
- Lifson S, Roig A. 1961. On the theory of the helix-coil transition in polypeptides. *J Chem Phys* 34:1963–1974.
- Lyu PC, Wemmer DE, Zhou HX, Pinker RJ, Kallenbach NR. 1993. Capping interactions in isolated α -helices: Position-dependent substitution effects and structure of a serine-capped peptide helix. *Biochemistry* 32:421–425.
- Penel S, Hughes E, Doig AJ. 1999. Side-chain structures in the first turn of the α -helix. *J Mol Biol* 287:127–143.
- Petukhov M, Muñoz V, Yumoto N, Yoshikawa S, Serrano L. 1998. Position dependence of non-polar amino acid intrinsic helical propensities. *J Mol Biol* 278:279–289.
- Petukhov M, Uegaki K, Yumoto N, Yoshikawa S, Serrano L. 1999. Position dependence of amino acid intrinsic helical propensities II: Non-charged polar residues: Ser, Thr, Asn, and Gln. *Protein Sci* 8:2144–2150.
- Presta LG, Rose GD. 1988. Helix signals in proteins. *Science* 240:1632–1641.
- Richardson JS, Richardson DC. 1988. Amino acid preferences for specific locations at the ends of α -helices. *Science* 240:1648–1652.
- Rohl CA, Chakrabartty A, Baldwin RL. 1996. Helix propagation and N-cap propensities of the amino acids measured in alanine-based peptides in 40 volume percent trifluoroethanol. *Protein Sci* 5:2623–2637.
- Rohl CA, Doig AJ. 1996. Models for the 3_{10} -helix/coil, π -helix/coil, and α -helix/ 3_{10} -helix/coil transitions in isolated peptides. *Protein Sci* 5:1687–1696.
- Serrano L, Fersht AR. 1989. Capping and α -helix stability. *Nature* 342:296–299.
- Serrano L, Sancho J, Hirshberg M, Fersht AR. 1992. α -Helix stability in proteins. I. Empirical correlations concerning substitution of side chains at the N and C-caps and the replacement of alanine by glycine or serine at solvent-exposed surfaces. *J Mol Biol* 227:544–559.
- Stapley BJ, Rohl CA, Doig AJ. 1995. Addition of side chain interactions to modified Lifson–Roig helix-coil theory: Application to energetics of phenylalanine-methionine interactions. *Protein Sci* 4:2383–2391.
- Sun JK, Doig AJ. 1998. Addition of side-chain interactions to 3_{10} -helix/coil and α -helix/ 3_{10} -helix/coil theory. *Protein Sci* 7:2374–2383.
- Yumoto N, Murase S, Hattori T, Yamamoto H, Tatsu Y, Yoshikawa S. 1993. Stabilization of α -helix in C-terminal fragments of neuropeptide Y. *Biochem Biophys Res Comm* 196:1490–1495.