N- and C-capping preferences for all 20 amino acids in α-helical peptides

ANDREW J. DOIG[1] AND ROBERT L. BALDWIN[2]

[1] Department of Biochemistry and Applied Molecular Biology, University of Manchester Institute of Science and Technology, P.O. Box 88, Manchester M60 1QD, United Kingdom
[2] Department of Biochemistry, Stanford University Medical Center, Stanford, California 94305

Abstract

We have determined the N- and C-capping preferences of all 20 amino acids by substituting residue X in the peptides NH_2-XAKAAAAKAAAAKAAGY-$CONH_2$ and in Ac-YGAAKAAAAKAAAAKAX-CO_2H. Helix contents were measured by CD spectroscopy to obtain rank orders of capping preferences. The data were further analyzed by our modified Lifson–Roig helix–coil theory, which includes capping parameters (n and c), to find free energies of capping ($-RT \ln n$ and $-RT \ln c$), relative to Ala. Results were obtained for charged and uncharged termini and for different charged states of titratable side chains. N-cap preferences varied from Asn (best) to Gln (worst). We find, as expected, that amino acids that can accept hydrogen bonds from otherwise free backbone NH groups, such as Asn, Asp, Ser, Thr, and Cys generally have the highest N-cap preference. Gly and the acetyl group are favored, as are negative charges in side chains and at the N-terminus. Our N-cap preference scale agrees well with preferences in proteins. In contrast, we find little variation when changing the identity of the C-cap residue. We find no preference for Gly at the C-cap in contrast to the situation in proteins. Both N-cap and C-cap results for Tyr and Trp are inaccurate because their aromatic groups affect the CD spectrum. The data presented here are of value in rationalizing mutations at capping sites in proteins and in predicting the helix contents of peptides.

Keywords: α-helix; C-cap; circular dichroism; helix propensities; N-cap; modified Lifson–Roig theory; protein folding; protein stability

The α-helix is the most commonly occurring secondary structural element in proteins. It has been known for some time that different amino acids show varied preferences for different secondary structures (Chou & Fasman, 1974). Since then, there has been a considerable amount of work on experimentally measuring the preference of the amino acids for the α-helix. Most studies concentrated on the solvent-exposed surface of a central position of an α-helix (reviewed by Scholtz & Baldwin, 1992, and Chakrabartty & Baldwin, 1994). Intrinsic preferences were first measured in a host–guest system of random copolymers (Sueki et al., 1984; Wojcik et al., 1990), but these results gave poor agreement with data from peptides with defined sequences. The reason was investigated by Padmanabhan et al. (1994). Helix preferences have also been measured in shorter isolated helices (Lyu et al., 1990; Merutka & Stellwagen, 1990; Chakrabartty et al., 1994), in proteins (Horovitz et al., 1992; Blaber et al.,

1994), and in trimeric coiled coils (O'Neil & DeGrado, 1990; Lovejoy et al., 1993).

Recently it has been realized that helix preferences differ greatly between interior positions and the helix termini. The N-terminus of a helix has four amide hydrogens that lack the i–(i + 4) backbone hydrogen bonds that characterize the α-helix; similarly, the C-terminus has four free carbonyl groups. Presta and Rose (1988) suggested that hydrogen bonds from polar side chains to these otherwise unsatisfied hydrogen bonding partners are of critical importance in establishing the locations of helix termini. The residue immediately preceding the N-terminal side of the helix has been dubbed the N-cap, whereas the residue immediately following the C-terminal side of the helix is the C-cap. The capping residues have nonhelical (ϕ, ψ) dihedral angles. The observed distributions of amino acids for these sites in protein crystal structures is quite different for preferences for interior positions (Richardson & Richardson, 1988). The importance of capping for helix structure and stability has been studied in peptide (Nicholson et al., 1988, 1991; Fairman et al., 1989; Bruch et al., 1991; Lyu et al., 1992, 1993; Chakrabartty et al., 1993a;

Forood et al., 1993, 1994; Heinz et al., 1993; Regan, 1993; Doig et al., 1994; Zhou et al., 1994a, 1994b) and protein systems (Serrano & Fersht, 1989; Lecomte & Moore, 1991; Bell et al., 1992; Serrano et al., 1992b; Harper & Rose, 1993; Kaarsholm et al., 1993; Zhukovsky et al., 1994).

Although several groups have measured the free energy changes that arise when certain residues are substituted at capping positions (see below), many amino acids have not been investigated at all. In this paper, we determine the intrinsic preferences for both N- and C-cap positions in our α-helical model peptide for all 20 naturally occurring amino acids in various charged states, plus the N-terminal acetyl and C-terminal amide groups. We use the methodology of our previous work, where the N-cap preferences of the 11 nonaromatic, uncharged amino acids were found, together with the acetyl group (Chakrabartty et al., 1993a; Doig et al., 1994).

Results

Theory for the α-helix/coil transition (Doig et al., 1994)

The α-helix is characterized by $i-(i + 4)$ C=O to H–N hydrogen bonds (Fig. 1). Each residue in the polypeptide is classified as being in either a helical (h) or a coil (c) conformation, by its backbone (ϕ, ψ) dihedral angles. Residues are given individual statistical weights based on their conformations and upon the conformations of the residues on either side. Statistical weights are thus given to the central residue in a triplet of conformations. Each statistical weight is related to that of a coil residue between two other coil residues (i.e., a ccc triplet), which is defined as having a weight of 1. To form an $i-(i + 4)$ hydrogen bond, which stabilizes the helix, it necessary to have at least three residues in a helical conformation in a row. An hhh triplet is therefore given a weight of ω that reflects the cost of fixing a residue in an α conformation, offset by the benefit of forming the hydrogen bond. It is unfavorable, for entropic reasons, to fix a residue in a helical conformation with no hydrogen bond being formed. The statistical weight of such a residue is v; it is given to a helical residue adjacent to at least one coil residue (the triplets chh, hhc, and chc). The coil residue immediately preceding a helical stretch (the N-cap) has a weight of n; the coil residue immediately following a helical stretch (the C-cap) has a weight of c. Finally, the rare conformational triplet hch is at present assigned a weight equal to the geometric mean of the N- and C-cap weights $((nc)^{1/2})$. The statistical weights for these conformations are shown in Figure 1. The theory does not include

side-chain–side-chain interactions. It should be adequate, however, for accounting for the stability and determining all the properties of any isolated peptide that can form an α-helix.

Peptide design

A residue at the N-terminal end of an unacetylated peptide cannot adopt an h conformation because it is not bounded by an amide group on both sides (Doig et al., 1994). Such a residue can have a weighting of 1 or n only, depending on whether the following residue is in a c or h conformation, respectively. n-Values can thus be determined by substituting the residue at the N-terminal position of an unacetylated peptide and measuring the change in helix content. The exact structure of the peptide we have used is NH_2-XAKAAAAKAAAAKAAGY-$CONH_2$, where X is the residue being changed. This is identical to the sequence we used previously (Chakrabartty et al., 1993a) for straightforward comparison. If residue X is varied, it affects helix content solely through its N-cap preference. A peptide of this sequence has a helix content of around 50% so that the helix content is maximally sensitive to a change in free energy of helix formation. The Tyr residue is present to give a UV absorption for determination of peptide concentration. The Gly residue ensures that the helix almost always terminates at, or before, this position so that problems arising from Tyr in a helix affecting the CD signal are minimized (Chakrabartty et al., 1993b). All measurements are made at 273 K, as helix content is greater at lower temperature. The peptide we have used to investigate C-capping follows similar principles. The sequence is Ac-YGAA KAAAAKAAAAKAX-CO_2H.

Helix contents of N-cap peptides

The helix content of each peptide was measured at pH 9.55, where the N-terminus is neutral and the Lys residues remain charged, and at pH 5, where the N-terminus is positively charged for residues that have no titratable side chain between pH 2 and 10. The results are given in Table 1. pH titrations were performed on each of these peptides to determine the pK_a values of the N-terminus using the Henderson–Hasselbalch equation (see Fig. 2). The results are given in Table 2.

Titrations of the amino acid side chains in Asp, Cys, Glu, and His gave changes in helix content. The data were fitted to a Henderson–Hasselbalch equation for 2 different pK_a values (Fig. 2). Note that not every combination of charged states could be examined. For example, it is not possible to measure the he-
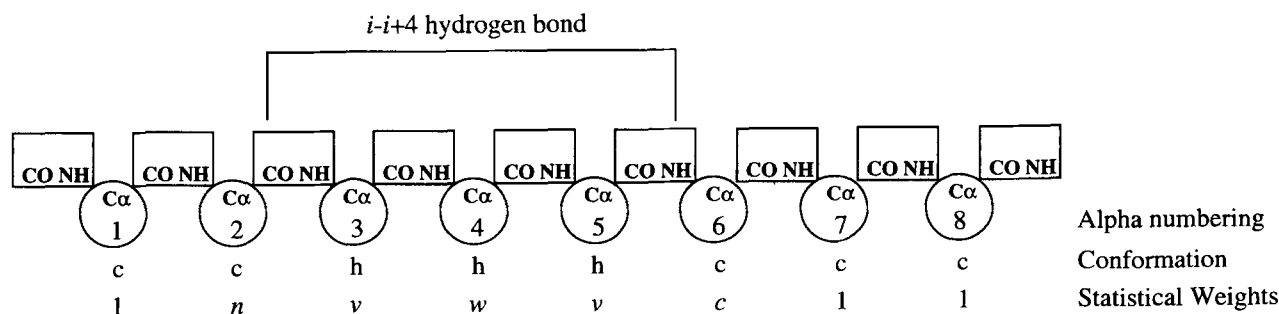

Fig. 1. Statistical weights for residues in an α-helix with modified Lifson-Roig theory (Doig et al., 1994).

Table 1. *Helix contents for peptides with varying N-terminal amino acid*[a]

Residue	$[\theta]_{222}$ (uncharged N-terminus)	Fraction helix[b] (uncharged N-terminus)	$[\theta]_{222}$ (charged N-terminus)	Fraction helix[b] (charged N-terminus)
Asn	$-20,300^c$	0.595	$-17,100^c$	0.501
Asp⁻	$-20,200$	0.591	$-17,900$	0.527
Asp⁰			$-14,200$	0.420
Acetyl	$-19,700^d$	0.578		
Cys⁻	$-19,400$	0.569	$-12,900$	0.379
Cys⁰			$-8,800$	0.257
Trp	$-18,800^e$	$0.552^e/0.516^f$	$-9,900^e$	$0.289^e/0.313^f$
Gly	$-17,900^c$	0.525	$-14,400^c$	0.422
Ser	$-17,900^c$	0.525	$-13,100^c$	0.384
Tyr	$-16,500^g$	$0.485^g/0.556^f$	$-9,300^f$	$0.273^g/0.313^f$
Thr	$-15,600^c$	0.457	$-12,200^c$	0.358
His⁰	$-15,400$	0.450	$-10,900$	0.319
His⁺			$-5,100$	0.150
Leu	$-15,300^c$	0.449	$-10,800^c$	0.317
Phe	$-15,300^g$	0.448^g	$-7,800^g$	0.230^g
Glu⁻	$-15,300$	0.447	$-10,200$	0.298
Glu⁰			$-9,700$	0.285
Ile	$-14,300^c$	0.419	$-10,500^c$	0.308
Pro	$-13,800^c$	0.404	$-8,000^c$	0.234
Met	$-13,700^c$	0.402	$-10,200^c$	0.299
Ala	$-12,900^c$	0.378	$-7,900^c$	0.232
Arg⁺	$-12,900$	0.379	$-5,700$	0.167
Val	$-12,800^c$	0.375	$-9,000^c$	0.265
Lys⁺	$-12,100$	0.355	$-6,100$	0.180
Gln	$-10,000^c$	0.293	$-7,000^c$	0.205

[a] The peptide sequence is NH₂-XAKAAAAKAAAAKAAGY-CONH₂. Data are at pH 9.55 (uncharged N-terminus), pH 5 (charged N-terminus), or from baselines fitted to the Henderson–Hasselbalch equation (Asp, Cys, His, and Glu). All data are at 273 K in 1 M NaCl.
[b] Calculated as $-[\theta]_{222}/34,100$.
[c] From Chakrabartty et al. (1993a).
[d] From Doig et al. (1994).
[e] Probably highly inaccurate because of Trp extinction coefficient and Trp effect on CD spectrum (see text). Calculated as if aromatic effect is not present.
[f] Corrected value (see text).
[g] Inaccurate because of aromatic effect on CD spectrum.

lix content of a peptide with an uncharged N-terminus and a neutral Asp side chain. The results are given in Tables 1 and 2. All buffers used contained 1 M NaCl to screen attractions and repulsions between the charged N-terminus, Lys side chains, and charged capping amino acids.

Helix contents of C-cap peptides

The helix content of each peptide was measured at pH 1, where the N-terminus is neutral and the Lys residues remain charged, and at pH 7, where the N-terminus is positively charged for residues that have no titratable side chain between pH 1 and 10. The results are given in Table 3. Titrations of the amino acid side chains in Asp, Cys, Glu, and His were analyzed as for the N-cap peptides (see Fig. 2). The results are given in Tables 3 and 4.

Determination of n- and c-values from helix contents

In previous work we determined all the w- and some n-values while assuming that the c-value of every amino acid is equal to 1 (Chakrabartty et al., 1994; Doig et al., 1994). In this paper, we consider variation in c-values. Allowing c-values to vary can

lead to changes in w and n. It is therefore necessary to reevaluate w and n, and find c, for Ala, Lys, Gly, Tyr, acetyl, and amide, which occur in the peptides used to give the data in Tables 1 and 3. We therefore used the data in Table 5 to find the results in Table 6. Absolute n- and c-values cannot yet be determined. They are therefore given relative to Ala with an uncharged N- or C-terminus, which is defined to have $n(\text{Ala}) = c(\text{Ala}) = 1$. $w(\text{Gly})$ was fixed at 0.05 (Chakrabartty et al., 1994). The w and n values listed here are in excellent agreement with previously published work, showing that allowing the c-values to vary makes little difference. This is consistent with the conclusion that altering the identity of the C-cap residue makes little difference to helix stability.

The rank order of capping preferences for every amino acid is given by the helix contents listed in Tables 1 and 3. In order to turn this information into a quantitatively useful form, it is necessary to use a helix–coil theory that can give capping preferences and hence free energies. The values of n and c for the amino acids not yet considered were calculated using the parameters listed in Table 6 and the data in Tables 1 and 3. For each sequence, the value of n or c was found that gave agreement with the experimental helix content. The results are listed in Tables 7
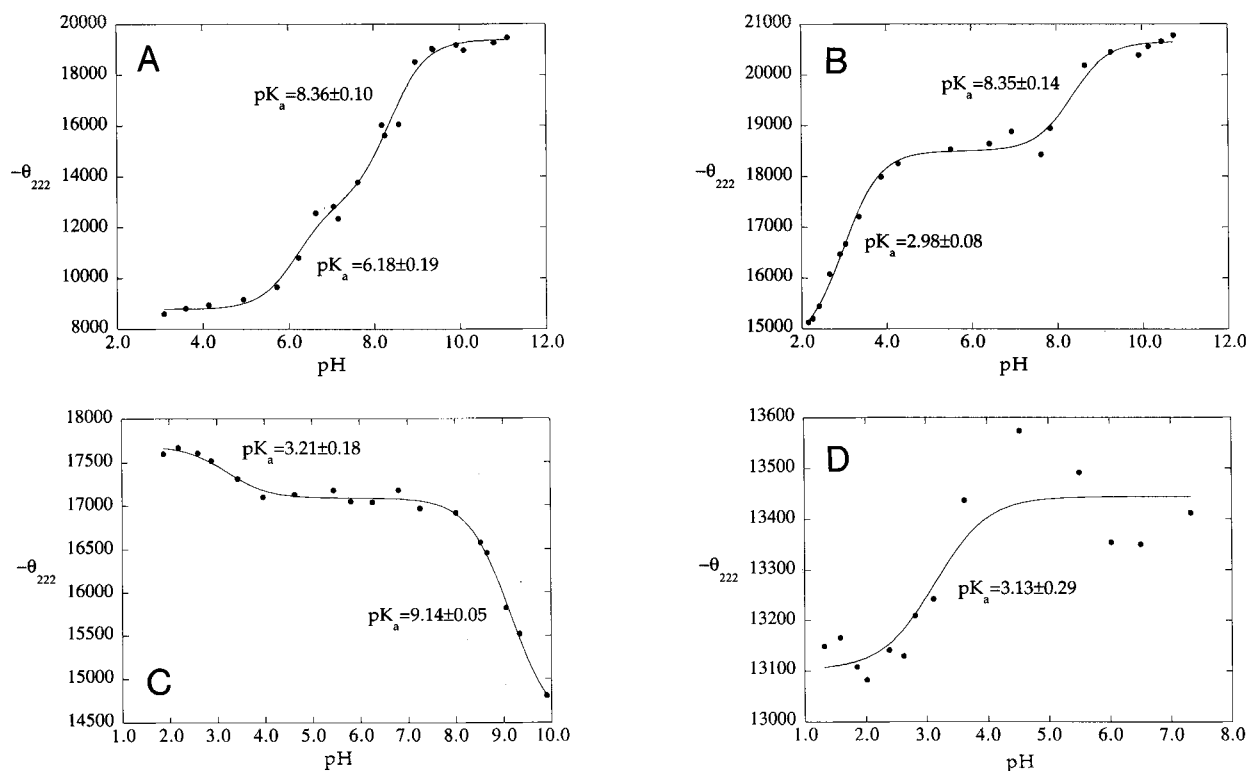
Fig. 2. Representative pH titrations of CD signal for four peptides. Some data for high pH are discarded where neutralization of the Lys side chain leads to aggregation. **A:** Cys N-cap (NH_2-CAKAAAAKAAAAKAAGY-$CONH_2$). **B:** Asp N-cap (NH_2-DAKAAAAKAAAAKAAGY-$CONH_2$). **C:** Cys C-cap (Ac-YGAAKAAAAKAAAAKAC-CO_2H). **D:** Pro C-cap (Ac-YGAAKA AAAKAAAAKAP-CO_2H).

and 8. Free energies of capping (Tables 9, 10) are given by $-RT \ln(n)$ or $-RT \ln(c)$ and give the change in free energy when an amino acid, of intrinsic capping preference n or c, is substituted for Ala. In some cases, the program converged to

Table 2. pK_a values in N-cap peptides

Residue	pK_a, N-terminus	pK_a, side chain
Asn	7.07[a]	
Asp	8.25	3.02
Cys	8.36	6.18
Trp	7.84	
Gly	8.51[a]	
Ser	7.63[a]	
Tyr	7.92	
Thr	7.62[a]	
His	8.12	6.13
Leu	8.31[a]	
Phe	7.75	
Glu	8.51	4.61
Ile	8.24[a]	
Pro	8.85[a]	
Met	7.83[a]	
Ala	8.35[a]	
Arg	7.96	
Val	8.14[a]	
Lys	8.12	
Gln	7.72[a]	

[a] From Chakrabartty et al. (1993a).

negative n- or c-values. This is equivalent to a negative equilibrium constant and is thus meaningless. This arises possibly as a result of experimental error but is more likely caused by fixing all n- and c-values relative to Ala. If the absolute n- or c-value of Ala is greater than 1, then there will be a greater range of helix contents less than that of Ala that will give positive capping preferences.

The most useful sets of results are those that give capping preferences for amino acids at uncharged termini, because these mimic the environment of a protein more closely. Tables 7 and 9 give this result for every amino acid at an N-cap site. Tables 8 and 10 give this result for 15 amino acids at a C-cap site.

Discussion

pK_a values

The pK_a's of the N-termini in the peptides (Table 2) are on average 1.5 pH units below those of the free amino acids (Fersht, 1984). This is presumably caused by the absence of the adjacent CO_2^- group that is present in the free amino acid and the nearby presence of NH groups and the positive end of the helix dipole. In contrast, no universal trends are observed for side chain pK_a's. The pK_a's of Asp and Cys at the N-terminus of the peptides are ≈ 1 and ≈ 2 pH units below those of the free amino acids, respectively, the pK_a of His is unchanged, and the pK_a of Glu is ≈ 0.5 pH units higher than its free amino acid. These results are consistent with the Asp and Cys side chains accepting hydrogen bonds from free amide NH groups at the N-terminus

Table 3. *Helix contents for peptides with varying C-terminal amino acid*[a]

Residue	$[\theta]_{222}$ (uncharged N-terminus)	Fraction helix[b] (uncharged N-terminus)	$[\theta]_{222}$ (charged N-terminus)	Fraction helix[b] (charged N-terminus)
Tyr	$-25,800^c$	$0.757^c/0.869^d$	$-22,700^c$	$0.666^c/0.764^d$
Glu^0	$-18,900$	0.554		
Glu$^-$	$-16,200$	0.475	$-14,200$	0.416
Arg$^+$	$-18,800$	0.551	$-16,700$	0.490
Ala	$-18,500$	0.543	$-15,200^e$	0.446
Gln	$-18,300$	0.537	$-15,500^e$	0.455
Gly	$-18,300$	0.537	$-15,200^e$	0.446
Amide	$-18,200$	0.534		
Met	$-18,200$	0.534	$-15,300^e$	0.449
His$^+$	$-17,900$	0.525	$-14,700$	0.431
His^0			$-13,300$	0.390
Leu	$-17,600$	0.516	$-15,000^e$	0.440
Asn	$-16,900$	0.496	$-14,000^e$	0.411
Cys$^-$			$-14,400$	0.422
Cys^0	$-16,800$	0.493	$-16,300$	0.478
Asp^0	$-16,700$	0.490		
Asp$^-$	$-15,100$	0.443	$-12,800$	0.375
Lys$^+$	$-16,300$	0.478	$-15,000$	0.440
Ser	$-15,800$	0.463	$-13,800^e$	0.405
Val	$-15,800$	0.463	$-13,800^c$	0.405
Thr	$-14,700$	0.431	$-13,100^e$	0.384
Phe	$-14,200^c$	0.416^c	$-14,700^c$	0.431
Trp	$-14,000^c$	$0.411^c/0.385^d$		
Pro	$-13,000$	0.381	$-13,500^e$	0.396
Ile	$-12,600$	0.370	$-13,500^e$	0.396

[a] The peptide sequence is Ac-YGAAKAAAAKAAAAKAX-CO_2H. Data are at pH 1 (uncharged N-terminus), pH 7 (charged N-terminus), or from baselines fitted to the Henderson–Hasselbalch equation (Asp, Cys, His, and Glu). All data are at 273 K in 1 M NaCl.
[b] Calculated as $-[\theta]_{222}/34,100$.
[c] Probably highly inaccurate because of aromatic effect on CD spectrum (see text). Calculated as if aromatic effect is not present.
[d] Corrected value (see text).
[e] From Chakrabartty et al. (1993a).

of the helix while the His and Glu side chains remain in solvent water (see below).

The pK_a's of the C-termini and the C-terminal side chains in the peptides (Table 4) are each around 1 pH unit higher than those of the free amino acids (Fersht, 1984). This is presumably caused by the absence of the adjacent NH_3^+ group in the free amino acids and the nearby presence of CO groups in the helix and the negative end of the helix dipole.

Aromatic residues

It has been shown that the presence of Tyr or Trp residues within an α-helix leads to an error in the determination of helix content from $[\theta]_{222}$ values, as a result of UV absorbance by the aromatic groups at that wavelength (Chakrabartty et al., 1993b). This problem is negligible if the Tyr group, which is necessarily present for concentration determination, is separated from the helix by a helix-breaking Gly spacer. Each of the peptides considered in this paper thus has Gly-Tyr-$CONH_2$ at its C-terminus or Ac-Tyr-Gly at its N-terminus. We consider that results obtained from peptides without this precaution are unreliable and are thus ignored.

When determining the capping preferences of Tyr or Trp, however, an undesirable aromatic CD artifact is unavoidable.

Raw $[\theta]_{222}$ values can be corrected for the presence of Tyr or Trp as described in Chakrabartty et al. (1994). The true fraction helix is given by $[\theta]_{222}^{observed}/([\theta]_{222}^{helix} + [\theta]_{222}^{aromatic})$, where $[\theta]_{222}^{helix}$ is the maximum mean residue ellipticity (given by $-40,000(1 - 2.5/n)$, where n is the number of amino acids in the peptide), and $[\theta]_{222}^{aromatic}$ is $+4,400$ deg·cm^2 dmol^{-1} for Tyr and $-2,300$ deg·cm^2 dmol^{-1} for Trp. Using this equation, the helix contents for Tyr and Trp N- and C-cap peptides were corrected (Tables 1, 3) and new *n*- and *c*-values determined (Tables 7, 8).

Table 4. *pK_a values in C-cap peptides*

Residue	pK_a, N-terminus	pK_a, side chain
Asp	3.12	4.56
Cys	3.21	9.14
His	2.88	7.68
Glu	3.14	4.61
Pro	3.13	
Arg	2.81	
Lys	3.50	

Table 5. *Peptide sequences and helix contents used to determine w, n, and c for Ala, Gly, Lys, Gln, acetyl, and amide*

Name	Sequence	Helix content
YG	Ac-YGKAAAAKAAAAKAAAAK-CONH$_2$	0.737[a]
YGG	Ac-YGGKAAAAKAAAAKAAAAK-CONH$_2$	0.677[a]
YGGG	Ac-YGGGKAAAAKAAAAKAAAAK-CONH$_2$	0.616[a]
YGG-G18	Ac-YGGKAAAAKAAAAKAAGAK-CONH$_2$	0.465[a]
YGG-G8	Ac-YGGKAAGAKAAAAKAAAAK-CONH$_2$	0.373[a]
YGG-G13	Ac-YGGKAAAAKAAGAKAAAAK-CONH$_2$	0.256[a]
NG-GY	NH$_2$-GAKAAAAKAAAAKAAGY-CONH$_2$	0.525[b]
NA-GY	NH$_2$-AAKAAAAKAAAAKAAGY-CONH$_2$	0.378[b]
NQ-GY	NH$_2$-QAKAAAAKAAAAKAAGY-CONH$_2$	0.293[b]
YGG-1Q	Ac-YGGKAAAAKAQAAKAAAAK-CONH$_2$	0.507[c]
YGG-3Q	Ac-YGGKAQAAKAQAAKAQAAK-CONH$_2$	0.312[c]
2Q-GGY	Ac-KAQAAKAQAAKAAAAKGGY-CONH$_2$	0.395[c]
N2Q-GGY	NH$_2$-KAQAAKAQAAKAAAAKGGY-CONH$_2$	0.220[c]
Qref	Ac-YGGQAAAAQAAAAQAAAAQ-CONH$_2$	0.518[c]
Q12	Ac-YGGQAAAAQAQAAQAAAAQ-CONH$_2$	0.378[c]
Q7, 12, 17	Ac-YGGQAQAAQAQAAQAQAAQ-CONH$_2$	0.266[c]
KAK	NH$_2$-KAKAAAAKAAAAKAAGY-CONH$_2$	0.355
XAK	Ac-AKAAAAKAAAAKAAGY-CONH$_2$	0.578
CCA	Ac-YGAAKAAAAKAAAAKAA-CO$_2$H	0.543
CCQ	Ac-YGAAKAAAAKAAAAKAQ-CO$_2$H	0.537
CCG	Ac-YGAAKAAAAKAAAAKAG-CO$_2$H	0.536
CCZ	Ac-YGAAKAAAAKAAAAKA-CONH$_2$	0.535
CCK	Ac-YGAAKAAAAKAAAAKAK-CO$_2$H	0.477

[a] Data taken from Chakrabartty et al. (1993b).
[b] Data taken from Chakrabartty et al. (1993a).
[c] Data taken from Chakrabartty et al. (1994).

Despite this correction, a number of results for Tyr and Trp peptides still appear to be inaccurate. For example, the *c*-value of Tyr with a charged C-terminus was found to be 8.9 (uncorrected) or 94 (corrected), both of which are very high compared to other results in Table 8. This suggests that the aromatic effect on $[\theta]_{222}$ depends strongly on the location of the aromatic ring relative to the α-helix. The corrections of +4,400 deg·cm^2 dmol^{-1} for Tyr and −2,300 deg·cm^2 dmol^{-1} for Trp were found for the side chains in the middle of a helix. They appear

Table 6. *w-, n-, and c-Values for Ala, Gly, Lys, Gln, acetyl, and amide*

w(Ala)	1.58
n(Ala)	1[a]
c(Ala)	1[a]
w(Lys)	0.90
n(Lys)	0.79
c(Lys)	1.09
w(Gln)	0.55
n(Gln)	0.01
c(Gln)	2.38
w(Gly)	0.05[a]
n(Gly)	8.85
c(Gly)	0.88
n(Acetyl)	11.7
c(Amide)	1.30

[a] Fixed (see text).

Table 7. *n-Values for peptides with varying N-terminal amino acid*

Residue	n[a] (uncharged N-terminus)	n[a] (charged N-terminus)	ΔG for protonation of N-terminus[b]
Asn	22	6.3	0.7
Asp$^-$	20	8.3	0.5
Asp^0		2.4	−
Acetyl	12		−
Cys$^-$	14	1.3	1.3
Cys^0		<0[c]	−
Trp	11[d]/7.3[e]	<0[c,d,e]	−
Gly	8.9	2.5	0.7
Ser	8.3	1.4	1.0
Tyr	5.2[d]/12[c]	<0[c,d]/0.07[c]	−/2.8[e,f]
Thr	3.7	0.85	0.8
His^0	3.4	0.16	1.7[f]
His$^+$		<0[c]	−
Leu	3.4	0.13	1.8[f]
Phe	3.4[c]	<0[c]	−
Glu$^-$	3.3	<0[c]	−
Glu^0		<0[c]	−
Ile	2.4	<0c	−
Pro	1.9	<0[c]	−
Met	1.8	<0[c]	−
Ala	1.0	<0[c]	−
Arg$^+$	1.3	<0[c]	−
Val	1.2	<0[c]	−
Lys$^+$	0.78	<0[c]	−
Gln	0.01	<0[c]	−

[a] From fitting program.
[b] $-RT \ln[n(\text{charged})/n(\text{uncharged})]$ in kcal·mol^{-1}.
[c] See text.
[d] Inaccurate because of aromatic effect on CD spectrum.
[e] After estimate of aromatic effect on CD (see text).
[f] Inaccurate as n(charged) is very small.

to be inappropriate for side chains in capping positions. The results given in this paper for Tyr and Trp are therefore likely to be highly inaccurate, and CD is thus probably not a suitable technique for determining the capping preferences of these amino acids.

N-cap preferences

Tables 1, 7, and 9 give, for the first time, complete scales for N-cap preferences. It is important to try to rationalize this information, particularly because there is at present some dispute over the principal mechanism of capping (Zhukovsky et al., 1994). The two extreme models discussed by Zhukovsky et al. are: first, capping by hydrogen bonding from, and to, the exposed NH and CO groups at the N- and C-termini of the helices, and second, charge–dipole interactions, where charged side chains form stabilizing interactions with the helix dipole. These hypotheses are not mutually exclusive, as a charged side chain can also function as a hydrogen bond acceptor or donor. A free energy simulation of Tidor (1994) suggested that helix-stabilizing interactions, as a result of a Tyr → Asp substitution at an N-cap site, arose from hydrogen bonding interactions from its direct

Table 8. *c-Values for peptides with varying C-terminal amino acid*

Residue	c^a (uncharged C-terminus)	c^a (charged C-terminus)	ΔG for protonation of C-terminus[b]
Tyr	$63^c/\infty^d$	$8.9^c/94^d$	$-1.1/-$
Glu^0	2.2		
Glu$^-$	0.41	$<0^c$	$-$
Arg$^+$	2.1	0.66	-0.6
Ala	1.0	$<0^c$	$-$
Gln	2.4	0.11	-1.7
Gly	0.88	$<0^c$	$-$
Amide	1.3		$-$
Met	1.6	0.02	-2.4
His$^+$	1.4	$<0^c$	$-$
His^0		$<0^c$	$-$
Leu	1.2	$<0^c$	$-$
Asn	0.78	$<0^c$	$-$
Cys$^-$		$<0^c$	$-$
Cys^0	0.72	0.46	-0.2
Asp^0	0.66		
Asp$^-$	$<0^c$	$<0^c$	$-$
Lys$^+$	1.1	$<0^c$	$-$
Ser	0.21	$<0^c$	$-$
Val	0.21	$<0^c$	$-$
Thr	$<0^c$	$<0^c$	$-$
Phe	$<0^c$	$<0^c$	$-$
Trp	$<0^{c,e}$		
Pro	$<0^c$	$<0^c$	$-$
Ile	$<0^c$	$<0^c$	$-$

[a] From fitting program.
[b] $-RT \ln[c(\text{uncharged})/c(\text{charged})]$ in kcal·mol^{-1}.
[c] Probably highly inaccurate because of aromatic effect on CD spectrum (see text). Calculated as if aromatic effect is not present.
[d] Corrected value (see text).
[e] See text.

Table 9. *Free energies of N-capping (kcal·mol^{-1})*

Residue	ΔG^a (uncharged N-terminus)	ΔG (charged N-terminus)
Asn	-1.7	-1.0
Asp$^-$	-1.6	-1.2
Asp^0		-0.5
Acetyl	-1.4	
Cys$^-$	-1.4	-0.1
Trp	$-1.3^b/-0.6^c$	
Gly	-1.2	-0.5
Ser	-1.2	-0.2
Tyr	$-0.9^b/-1.5^c$	$-/0.7$
Thr	-0.7	0.1
His^0	-0.7	1.0
Leu	-0.7	1.1
Phe	$-0.7^b/-0.9^c$	
Glu$^-$	-0.7	
Ile	-0.5	
Pro	-0.4	
Met	-0.3	
Ala	0	
Arg$^+$	-0.1	
Val	-0.1	
Lys$^+$	0.1	
Gln	2.5^d	

[a] $-RT \ln n$.
[b] Inaccurate because of aromatic effect on CD spectrum.
[c] After estimate of aromatic effect on CD (see text).
[d] Highly inaccurate as n is very small.

bartty et al., 1993a) and suggests that Gln actively stabilizes a nonhelical state. Glu also has a low N-cap preference relative to other negatively charged side chains, implying that it is subject to the same effect. This is likely to be a hydrogen bond from

hydrogen bonding partners and from more distant electrostatic interactions with groups within the first two turns of the helix.

Table 1 shows that, without exception, the helix content, and hence the N-cap preference, decreases as the N-cap amino acid is titrated so that it becomes more positive. This is shown by titrations of the NH$_2$ group at the N-terminus for all 20 amino acids and also for the side chains of Asp, Glu, His, and Cys. This result can be rationalized as interactions of the N-terminus and side chains with the positive end of the helix dipole. The high n-value of the acetyl group can be explained by the fact that it is the only N-terminal group examined here that has no NH$_2$ or NH$_3^+$ group that can form unfavorable interactions with adjacent NH groups. Negatively charged side chains generally have high n-values, whereas the n-values of positively charged side chains are among the lowest.

On the other hand, there is considerable evidence that some N-cap side chains accept hydrogen bonds from otherwise free NH groups. These include Asn, Asp, Cys, Ser, and Thr, which all have high n-values. Their oxygen or sulfur atoms have been seen to accept hydrogen bonds from a backbone NH group at an $i + 3$ position (Richardson & Richardson, 1988). The remarkable difference in N-cap preference between Asn and Gln has been noted before (Richardson & Richardson, 1988; Chakra-

Table 10. *Free energies of C-capping (kcal·mol^{-1})*

Residue	ΔG^a (uncharged N-terminus)	ΔG (charged N-terminus)
Tyr	$-2.2^b/-$	$-1.2/-2.5$
Glu^0	-0.4	
Glu$^-$	-0.5	
Arg$^+$	-0.4	0.2
Ala	0	
Gln	-0.5	1.2^c
Gly	0.1	
Amide	-0.1	
Met	-0.3	2.1^c
His$^+$	-0.2	
Leu	-0.1	
Asn	0.1	
Cys^0	0.2	0.4
Asp^0	0.2	
Lys$^+$	-0.1	
Ser	0.8	
Val	0.9	

[a] $-RT \ln c$.
[b] Inaccurate because of aromatic effect on CD spectrum.
[c] Highly inaccurate as c is very small.

the Gln or Glu side chain to its own NH group in the nonhelical state, which has been observed for Glu (Ebina & Wüthrich, 1984).

Significant variations in N-cap preference among nonpolar side chains show that not all these results can be explained by hydrogen bonding or electrostatic effects. There appears to be a small preference for larger hydrophobic side chains over smaller hydrophobic side chains (Chakrabartty et al., 1993a). The high N-cap preference of Gly has been attributed to improved solvation of the NH groups when the smallest side chain is at the N-cap (Serrano & Fersht, 1989; Serrano et al., 1992a; Harpaz et al., 1994).

The free energy change for protonation of the N-terminus is given by $-RT \ln[n(\text{charged})/n(\text{uncharged})]$ (Doig et al., 1994). For residues with positive n in both charged and uncharged states, this quantity is listed in Table 7. Similarly, the effect on helix stability for protonating a side chain is given by the change in n-value. This can only be done for Asp as no other side chains have positive n-values in both protonated and deprotonated states. The free energy change for protonating the Asp side chain at the N-cap of a helix was thus found to be 0.7 kcal·mol^{-1}. These free energies for protonation are for 100% helix; the effects of fraying that are present in the experimental data are factored out when finding n-values.

C-cap preferences

There is much less variation in C-cap preference than N-cap (Tables 3, 8, 10). A significant observation is that the β-branched residues Thr, Val, and Ile all have low C-cap preferences. This may be due to poorer solvation of the CO groups at the C-terminus when a bulky β-branched side chain is present.

It is observed that the helix content of the peptide decreases when the C-terminus or side chain becomes more negative, with three exceptions. The first two are Phe and Ile, where this behavior is most likely caused by experimental error as the helix contents are very similar. The third is for Pro. Pro has a very low c-value as it has no NH group that can donate a hydrogen bond within the helix. A pH titration of the Pro COOH group (Fig. 2D) showed that there is a small (≈ 300 deg·cm^2 dmol^{-1}), but real, increase in helix content when the Pro C-terminus becomes negatively charged. All other residues showed a decrease in $[\theta]_{222}$ of $\approx 2,500$ deg·cm^2 dmol^{-1} when the COOH group was titrated. There are several possible explanations for this result. First, the COO$^-$ group may form a stabilizing interaction in the helix, such as a salt bridge with a Lys, that is lost or reduced when the COO$^-$ group is protonated. Second, the Pro amide *cis/trans* equilibrium may be altered when the COO$^-$ group is protonated, leading to a change in helix content if one stereoisomer is more stable in the helix. Third, the COOH group may stabilize a nonhelical state. We cannot distinguish between these possibilities at present. The otherwise universal decrease in helix content when the C-terminus or side chain becomes more negative is readily explained by electrostatic interactions with the negative end of the helix dipole.

The free energy change for protonation of the C-terminus is given as $-RT \ln[n(\text{charged})/n(\text{uncharged})]$ and is shown in Table 8. The free energy change for protonating the Glu side chain at the C-cap of a helix was thus found to be -0.9 kcal·mol^{-1}.

Figure 3 shows a plot of experimental helix content versus helix content calculated using the modified Lifson–Roig theory and
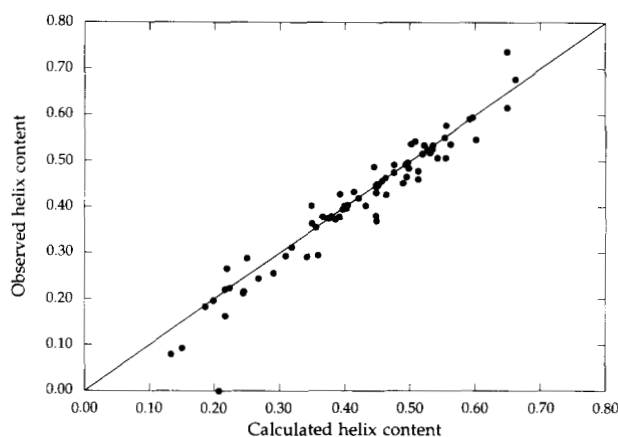

Fig. 3. Calculated helix content of peptides versus experimental helix contents. Helix contents are calculated using the w-, n-, and c-values given here and by Chakrabartty et al. (1994). The straight line is $y = x$, not the least-squares line of best fit.

the parameters listed in Tables 6, 7, and 8 and in Chakrabartty et al. (1994). Peptides with Trp or Tyr at interior or capping positions are excluded. This is the only set of peptides free from aromatic effects on the CD spectrum. The correlation coefficient (R) is 0.97 and the RMS deviation of calculated from experimental values is 3.7%, close to the expected experimental error.

Comparison with capping preferences in crystal structures

Capping preferences for residues in helices in crystal structures have been determined by Richardson and Richardson (1988). At the N-cap position, they found considerable variation, whereas at the C-cap, Gly was the only residue that was strongly preferred. Figure 4 shows a plot of n-values in peptides with uncharged N-termini (Table 7) versus the N-cap preferences of Richardson and Richardson (1988). The correlation coefficient is high ($R = 0.86$), which confirms that the N-cap sites in α-helices in proteins and our peptides are likely to be structur-
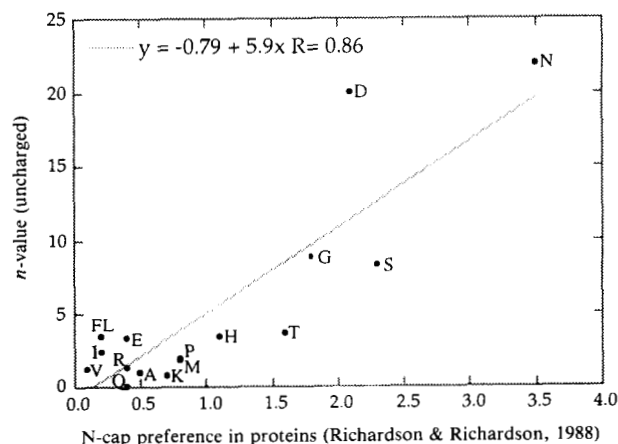

Fig. 4. N-cap preferences in peptide helices (n-values) versus N-cap preferences in crystal structures. Peptides with Tyr or Trp at the N-cap position are not included.

ally identical (Chakrabartty et al., 1993a). The range of N-cap preferences is considerably larger in our peptides than in the proteins, however, as shown by the gradient of 5.9. This probably arises as a result of the larger number of interactions in proteins that ensure that certain residues occupy an N-cap position, even though their intrinsic N-cap preferences are low. The peptide results benefit from being measured in an environment free from tertiary interactions.

Figure 5 shows a plot of our *c*-values with uncharged C-termini (Table 8) versus the C-cap preferences of Richardson and Richardson (1988). In contrast to Figure 4, the correlation is virtually zero ($R = 0.07$). Richardson and Richardson found that there was little variation in C-cap preference, apart from the widespread occurrence of Gly at this position. Our results show that the β-branched residues are disfavored, in agreement with Richardson and Richardson, though their statistics are poorer for rarely occurring residues. We find that Gly has an intrinsic *c*-value no higher than average. The majority of residues show little variation in *c*-value, suggesting that the identity of the C-cap residue is relatively unimportant compared to the N-cap, in peptides at least.

The striking difference in C-cap preferences between Gly in our peptides and in proteins deserves an explanation. Two distinct structures have been observed when Gly is at a C-cap position in a helix, named Schellman and α_L (Schellman, 1980; Baker & Hubbard, 1984; Milner-White, 1988; Bork & Preißner, 1991; Preißner & Bork, 1991; Dasgupta & Bell, 1993; Nagarajaram et al., 1993; Aurora et al., 1994). The Schellman structure is characterized by $6 \rightarrow 1$, $5 \rightarrow 2$ hydrogen bonds between NH and C=O groups in the backbone. The α_L conformation has a $5 \rightarrow 1$ backbone hydrogen bond. Gly is favored over other residues as it can readily adopt unusual (ϕ, ψ) angles. The particular Gly C-cap structure that is adopted depends on the presence or absence of side-chain–side-chain interactions and the nature of side chains at certain critical positions (Aurora et al., 1994). The structure also involves hydrogen bonds from backbone residues beyond the Gly residue. The peptides we have studied have no residue beyond the Gly and have no side-chains to form the side chain interactions that appear essential for these

structures. Indeed, we designed our sequences such that side-chain interactions are avoided wherever possible. We therefore see no preference for Gly in our peptides, in contrast to the situation in proteins.

Comparison with capping free energies in other systems

Free energy changes from mutating N-cap residues have been studied for 7 residues at one site in T4 lysozyme (Bell et al., 1992) and for 11 residues at two sites in barnase (Serrano et al., 1992b). Their results are compared with ours in Table 11 and there is reasonable agreement. The free energy changes arising from mutating two C-cap residues to eight different amino acids were studied by Serrano et al. (1992b) in barnase. These results are not compared to those found here as both these sites are Schellman structures (Aurora et al., 1994) and are thus structurally distinct from the capping sites in our peptides.

Muñoz and Serrano (1994) have recently published complete scales of N- and C-capping energies. They also derived preferences for helix interiors, capping box energies, and all $i-(i + 3)$ and $i-(i + 4)$ interaction energies by fitting data on helix contents of peptides. Because they had 768 parameters and data on only 323 peptides, however, they were forced to group some parameters together. They therefore assumed that His, Phe, Tyr, Trp, Leu, Val, Ile, Met, and Cys have the same N-cap preference, for example. Mainly for this reason, their data are not compared further with the results published here.

Applications and future work

The scales of *n*- and *c*-values presented here and the *w*-values presented in a previous paper (Chakrabartty et al., 1994) should be of use in predicting the helix contents of peptides. We expect that this information will be all that is necessary when there are no side-chain–side-chain interactions that affect the stability of the helix. It should also be noted that the measurement of helix content by CD is subject to possible serious errors when the helix contains Tyr or Trp, especially if these residues populate
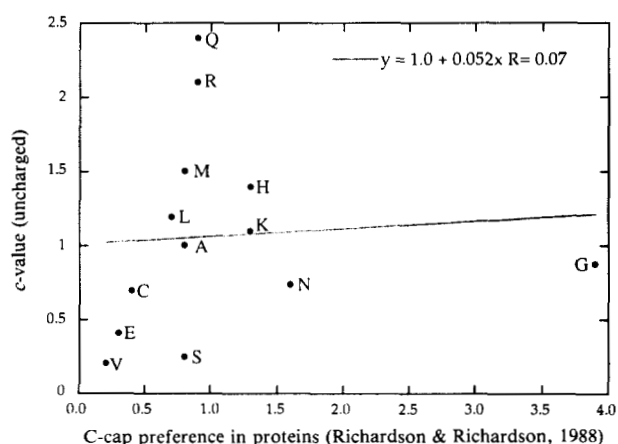
Fig. 5. C-cap preferences in peptide helices (*c*-values) versus C-cap preferences in crystal structures. Peptides with Tyr, Trp, Thr, Phe, Pro, and Ile at the C-cap position are not included (as *c* is negative for Trp, Thr, Phe, Pro, and Ile).

Table 11. *Free energies of N-capping compared in peptides and proteins relative to Ala (kcal·mol^{-1})*

Residue	Peptides[a]	T4 lysozyme, pH 2[b]	T4 lysozyme, pH 6.5[b]	Barnase[c]
Asn	−1.7	−2.2	−0.4	−0.86
Asp$^-$	−1.6	−1.9	−0.3	−2.02
Gly	−1.2	−0.6	0.1	−0.69
Ser	−1.2	−2.1	−1.3	−1.64
Thr	−0.7	−2.8	−1.5	−2.05
His^0	−0.7			−0.16
Glu$^-$	−0.7			−0.25
Pro	−0.4			0.87
Ala	0	0	0	0
Val	−0.1	0	0	0.15
Gln	2.5[d]			−0.42

[a] From Table 7. Uncharged N-terminus.
[b] Bell et al. (1992).
[c] Serrano et al. (1992b). Mean of results from two N-cap sites.
[d] Highly inaccurate as *n* is very small.

capping positions. The condition of excluding helices with side-chain–side-chain interactions naturally applies to almost all sequences. The next logical step is therefore to determine the free energies of interaction between side chains. If this is done for every possible combination of side chains in i-(i + 3) and i-(i + 4) spacings, the problem of predicting the helix content of any peptide is expected to be solved. We have already devised a theory that includes these interactions, together with w-, v-, n-, and c-values (C.A. Rohl & A.J. Doig, unpubl.).

Concluding remarks

By substituting amino acids at the unblocked N- and C-termini of Ala/Lys peptides, we have completed our work on determining the capping preferences of all 20 amino acids. The results were analyzed with our modified Lifson–Roig helix–coil theory which includes capping parameters. We find, as expected, that amino acids that can accept hydrogen bonds from otherwise free backbone NH groups, such as Asn, Asp, Ser, Thr, and Cys, generally have the highest N-cap preference. Gly and the acetyl group are favored, as are more negative charges at the N-terminus and in N-cap side chains. Our N-cap preference scale agrees well with preferences in proteins.

We find little variation when changing the identity of the C-cap residue. We find no preference for Gly at the C-cap, in contrast to the situation in proteins, because the Gly C-cap motifs observed in proteins depend on side-chain interactions that are not present and on additional backbone groups beyond the Gly. Both N-cap and C-cap results for Tyr and Trp are inaccurate because their aromatic groups affect the CD spectrum. The data presented here will be of value in rationalizing mutations at capping sites in proteins and in predicting the helix contents of peptides.

Materials and methods

Peptide synthesis

Peptides were synthesized by the solid-phase method. C-terminal amides were made using Rink resin (Advanced ChemTech) and C-terminal acids using Pepsyn KA resin (Milligen) with the C-terminal amino acid already attached. Pentafluorophenyl esters of 9-fluorenylmethoxycarbonyl amino acids were coupled to the resin using 1-hydroxybenzotriazole. N-termini were acetylated with acetic anhydride/pyridine or left unblocked. Peptides were cleaved from the resin using 95% trifluoroacetic acid/5% anisole. Ethanedithiol was added if Cys was present; indole was added if Trp was present; both ethanedithiol and thioanisole were added if Arg was present. Peptides were purified by C_{18} reverse-phase FPLC and their molecular weights were checked by FAB mass spectrometry. Peptide purity was checked by analytical C_{18} FPLC. Stock solutions of peptides containing Cys were maintained in a reduced state with 10 mM glutathione.

CD measurements

CD measurements were made on an Aviv 60DS spectropolarimeter in a 1-cm quartz cell. The buffer used for pH titrations contained 1 M NaCl, 1 mM sodium phosphate, 1 mM sodium borate, and 1 mM sodium citrate. The pH was adjusted during the titrations with aliquots of 1 M HCl and 1 M NaOH. The pH 5 buffer contained 1 M NaCl, 5 mM sodium phosphate, 5 mM sodium borate, and 5 mM sodium citrate. The pH 9.55 buffer contained 1 M NaCl, 5 mM sodium phosphate, 5 mM sodium borate, and 5 mM sodium citrate. The pH 1 buffer contained 1 M NaCl, 1 mM sodium phosphate, 1 mM sodium borate, and 1 mM sodium citrate. All measurements were made at 273 K.

Peptide concentration was determined by measuring tyrosine UV absorbance of aliquots of stock solution dissolved in 6 M guanidine hydrochloride using $\epsilon_{275} = 1,450$ M^{-1} cm^{-1} (Brandts & Kaplan, 1973). Concentrations of peptides containing Trp were determined using ϵ_{281} (Trp) = 5,690 M^{-1} cm^{-1} + ϵ_{281} (Tyr) = 1,250 M^{-1} cm^{-1} (Edelhoch, 1967).

CD measurements are given as mean residue ellipticity in units of deg·cm^2 dmol^{-1} at 222 nm ($[\theta]_{222}$). Helix content was calculated as $[\theta]_{222}^{observed}/[\theta]_{222}^{max}$. $[\theta]_{222}^{max}$ is given by $-40,000$ ($1 - 2.5/n$), where n is the number of amino acids in the peptide (Chakrabartty et al., 1991).

Determination of helix–coil parameters

The determination of w-, n-, and c-values was achieved through the use of two computer programs. The first implements the modified Lifson–Roig helix–coil theory (Doig et al., 1994) to calculate the helix content of a given peptide sequence. The program draws upon a library of w-, v-, n-, and c-values for each amino acid, plus n for acetyl and c for amide. A file giving the peptide sequence is input and the mean helix content calculated, together with other properties of the helix–coil equilibrium, such as fraction helix for each residue and mean number of helical segments. We assumed throughout that $v = 0.048$ for every residue (the mean value for Ala/Lys peptides measured by Rohl et al. [1992]). Although it may reasonably be expected that different residues will have different v-values (particularly Gly and Pro, which have unusual conformational entropy), we have no measurements at present than can give different v-values for different amino acids. Indeed, there is evidence that v is not influenced greatly by the side chain (Ptitsyn, 1972; Scholtz et al., 1991; Rohl et al., 1992). This program was written by J.A. Schellman and modified by H. Qian, C.A. Rohl, and T. Kortemme.

A second program was used to find the helix–coil parameters that gave the best fit to the experimental data (Chakrabartty et al., 1994). The inputs to the program are: the peptide sequences and corresponding experimental helix contents, library of w-, v-, n-, and c-values for each amino acid, and the acetyl and amide groups, and a list of parameters to be determined. This program used the modified Lifson–Roig theory program to determine the helix content of a peptide sequence for a given set of helix–coil parameters. Parameters were varied until they converged on essentially unchanging values. This program was written by S.L. Mayo and modified by T. Kortemme.

Acknowledgments

References

Aurora R, Srinivasan R, Rose GD. 1994. Rules for α-helix termination by glycine. *Science 264*:1126–1130.

Baker EN, Hubbard RE. 1984. Hydrogen bonding in globular proteins. *Prog Biophys Mol Biol 44*:97–179.

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.

Blaber M, Zhang XJ, Lindstrom JD, Pepiot SD, Baase WA, Matthews BW. 1994. Determination of α-helix propensity within the context of a folded protein. *J Mol Biol 235*:600–624.

Bork P, Preißner R. 1991. On α-helices terminated by glycine. 2. Recognition by sequence patterns. *Biochem Biophys Res Commun 180*:666–672.

Brandts JR, Kaplan KJ. 1973. Derivative spectroscopy applied to tyrosyl chromophores. Studies on ribonuclease, lima bean inhibitor, and pancreatic trypsin inhibitor. *Biochemistry 10*:470–476.

Bruch MD, Dhingra MM, Gierasch LM. 1991. Side chain–backbone hydrogen bonding contributes to helix stability in peptides derived from an alpha-helical region of carboxypeptidase A. *Proteins Struct Funct Genet 10*:130–139.

Chakrabartty A, Baldwin RL. 1994. Comparison of amino acid helix propensities ("s-values") measured in different experimental systems. In: Cleland J, King J, eds. *Protein folding: In vivo and in vitro. ACS Symposium Series 526*. Washington, D.C.: American Chemical Society. pp 166–177.

Chakrabartty A, Doig AJ, Baldwin RL. 1993a. Helix capping propensities in peptides parallel those in proteins. *Proc Natl Acad Sci USA 90*:11332–11336.

Chakrabartty A, Kortemme T, Baldwin RL. 1994. Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing interactions. *Protein Sci 3*:843–852.

Chakrabartty A, Kortemme T, Padmanabhan S, Baldwin RL. 1993b. Aromatic side-chain contributions to far-ultraviolet circular dichroism of helical peptides and its effect on measurement of helical propensities. *Biochemistry 32*:5560–5565.

Chakrabartty A, Schellman JA, Baldwin RL. 1991. Large differences in the helix propensities of alanine and glycine. *Nature 351*:586–588.

Chou PY, Fasman G. 1974. Conformational parameters of amino acids in helical, β-sheet, and random coil regions calculated from proteins. *Biochemistry 13*:211–222.

Dasgupta S, Bell JA. 1993. Design of helix ends. Amino acid preferences, hydrogen bonding and electrostatic interactions. *Int J Pept Protein Res 41*:499–511.

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.

Ebina S, Wüthrich K. 1984. Amide proton titration shifts in bull seminal inhibitor IIA by two-dimensional correlated ¹H nuclear magnetic resonance (COSY). Manifestation of conformational equilibria involving carboxylate groups. *J Mol Biol 179*:283–288.

Edelhoch H. 1967. Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry 6*:1948–1954.

Fairman R, Shoemaker KR, York EJ, Stewart JM, Baldwin RL. 1989. Further studies of the helix dipole model: Effects of a free α-NH₃⁺ or α-COO⁻ group on helix stability. *Proteins Struct Funct Genet 5*:1–7.

Fersht AR. 1984. *Enzyme structure and mechanism*. Reading, Pennsylvania: Freeman. pp 2–3.

Forood B, Feliciano EJ, Nambiar KP. 1993. Stabilization of α-helical structures in short peptides via end capping. *Proc Natl Acad Sci USA 90*:838–842.

Forood B, Reddy HK, Nambiar KP. 1994. Extraordinary helicity in short peptides via end capping design. *J Am Chem Soc 116*:6935–6936.

Harpaz Y, Elmasry N, Fersht AR, Henrick K. 1994. Direct observation of better hydration at the N-terminus of the α-helix with glycine rather than alanine as the N-cap residue. *Proc Natl Acad Sci USA 91*:311–315.

Harper ET, Rose GD. 1993. Helix stop signals in proteins and peptides: The capping box. *Biochemistry 32*:7606–7609.

Heinz DW, Baase WA, Dahlquist FW, Matthews BW. 1993. How amino acid insertions are allowed in an α-helix of T4 lysozyme. *Nature 361*:561–564.

Horovitz A, Matthews JM, Fersht AR. 1992. α-Helix stability in proteins. II. Factors that influence stability at an internal position. *J Mol Biol 227*:560–568.

Kaarsholm NC, Norris K, Jørgenson RJ, Mikkelsen J, Ludvigsen S, Olsen OH, Sørensen AR, Havelund S. 1993. Engineering stability of the insulin monomer fold with application to structure-activity relationships. *Biochemistry 32*:10773–10778.

Lecomte JTJ, 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.

Lovejoy B, Choe S, Cascio D, McRorie DK, DeGrado WF, Eisenberg D. 1993. Crystal structure of a synthetic triple-stranded α-helical bundle. *Science 259*:1288–1293.

Lyu PC, Liff MI, Markey LA, Kallenbach NR. 1990. Side-chain contributions to the stability of alpha-helical structure in peptides. *Science 250*:669–673.

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 helix. *Biochemistry 32*:421–425.

Lyu PC, Zhou HX, Jelveh N, Wemmer DE, Kallenbach NR. 1992. Position-dependent stabilizing effects in α-helices: N-terminal capping in synthetic model peptides. *J Am Chem Soc 114*:6560–6562.

Merutka G, Stellwagen E. 1990. Positional independence and additivity of amino acid replacements on helix stability in monomeric peptides. *Biochemistry 29*:894–898.

Milner-White EJ. 1988. Recurring loop motif in proteins that occurs in right-handed and left-handed forms. Its relationship with alpha-helices and beta-bulge loops. *J Mol Biol 199*:503–511.

Muñoz V, Serrano L. 1994. Elucidating the folding problem of helical peptides using empirical parameters. *Nature Struct Biol 6*:399–409.

Nagarajaram HA, Sowdhamini R, Ramakrishnan C, Balaram P. 1993. Termination of right-handed helices in proteins by residues in left-handed helical conformations. *FEBS Lett 321*:79–83.

Nicholson H, Anderson DE, Dao-Pin S, Matthews BW. 1991. Analysis of the interaction between charged side chains and the α-helix dipole using designed thermostable mutants of phage T4 lysozyme. *Biochemistry 30*:9816–9828.

Nicholson H, Becktel WJ, Matthews BW. 1988. Enhanced protein thermostability from designed mutations that interact with α-helix dipoles. *Nature 336*:651–656.

O'Neil KT, DeGrado WF. 1990. A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science 250*:646–651.

Padmanabhan S, York EJ, Gera L, Stewart JM, Baldwin RL. 1994. Helix-forming tendencies of amino acids in short (hydroxybutyl)-L-glutamine peptides: An evaluation of the contradictory results from host-guest studies and short alanine-based peptides. *Biochemistry 33*:8604–8609.

Preißner R, Bork P. 1991. On α-helices terminated by glycine. 1. Identification of common structural features. *Biochem Biophys Res Commun 180*:660–665.

Presta LG, Rose GD. 1988. Helix signals in proteins. *Science 240*:1632–1641.

Ptitsyn OB. 1972. Thermodynamic parameters of helix-coil transitions in polypeptide chains. *Pure Appl Chem 31*:227–244.

Regan L. 1993. What determines where α-helices begin and end? *Proc Natl Acad Sci USA 90*:10907–10908.

Richardson JS, Richardson DC. 1988. Amino acid preferences for specific locations at the ends of α-helices. *Science 240*:1648–1652.

Rohl CA, Baldwin RL. 1994. Exchange kinetics of individual amide protons in ¹⁵N-labeled helical peptides measured by isotope-edited NMR. *Biochemistry 33*:7760–7767.

Rohl CA, Scholtz JM, York EJ, Stewart JM, Baldwin RL. 1992. Kinetics of amide proton exchange in helical peptides of varying chain lengths. *Biochemistry 31*:1263–1269.

Schellman C. 1980. The α_L conformation at the ends of helices. In: Jaenicke R, ed. *Protein folding*. New York: Elsevier/North-Holland. pp 53–61.

Scholtz JM, Baldwin RL. 1992. The mechanism of α-helix formation by peptides. *Annu Rev Biophys Biomol Struct 21*:95–118.

Scholtz JM, Qian H, York EJ, Stewart JM, Baldwin RL. 1991. Parameters of helix-coil transition theory for alanine-based peptides of varying chain lengths in water. *Biopolymers 31*:1463–1470.

Serrano L, Fersht AR. 1989. Capping and α-helix stability. *Nature 342*:296–299.

Serrano L, Neira JL, Fersht AR. 1992a. Effect of alanine versus glycine in α-helices on protein stability. *Nature 356*:453–455.

Serrano L, Sancho J, Hirshberg M, Fersht AR. 1992b. α-Helix stability in proteins. I. Empirical correlations concerning substitution of side-chains at the N- and C-caps and replacement of alanine by glycine or serine at solvent-exposed surfaces. *J Mol Biol 227*:544–559.

Sueki M, Lee S, Powers SP, Denton JB, Konishi Y, Scheraga HA. 1984. Helix-coil stability constants for the naturally occurring amino acids in water. 22. Histidine parameters from random poly[(hydroxybutyl)glutamine-co-L-histidine]. *Macromolecules 17*:148–155.

Tidor B. 1994. Helix-capping interaction in λ Cro protein: A free energy simulation analysis. *Proteins Struct Funct Genet 19*:310–323.

Wojcik J, Altman KH, Scheraga HA. 1990. Helix-coil stability constants for the naturally occurring amino acids in water. XXIV. Half-cystine

parameters from random poly(hydroxybutylglutamine-co-*S*-methylthio-L-cysteine). *Biopolymers 30*:121-134.

Zhou HX, Lyu PC, Wemmer DE, Kallenbach NR. 1994a. Alpha helix capping in synthetic model peptides by reciprocal side chain-main chain interactions: Evidence for an N-terminal "capping box." *Proteins Struct Funct Genet 18*:1-7.

Zhou HX, Lyu PC, Wemmer DE, Kallenbach NR. 1994b. Structure of a C-terminal α-helix cap in a synthetic peptide. *J Am Chem Soc 116*: 1139-1140.

Zhukovsky EA, Mulkerrin MG, Presta LG. 1994. Contribution to global protein stabilization of the N-capping box in human growth hormone. *Biochemistry 33*:9856-9864.