Helix capping propensities in peptides parallel those in proteins

(α-helix/α-helix stability/N-cap and C-cap residues/protein folding)

AVIJ CHAKRABARTTY, ANDREW J. DOG, AND ROBERT L. BALDWIN*

Department of Biochemistry, Beckman Center, Stanford University Medical School, Stanford, CA 94305-5307

Contributed by Robert L. Baldwin, August 27, 1993

ABSTRACT Helix content of peptides with various uncharged nonaromatic amino acids at either the N-terminal or C-terminal position has been determined. The choice of N-terminal amino acid has a major effect on helix stability: asparagine is the best, glycine is very good, and glutamine is the worst helix-stabilizing amino acid at this position. The rank order of helix stabilization parallels the frequencies of these amino acids at the N-terminal boundary (N-cap) position of helices in proteins found by Richardson and Richardson [Richardson, J. S. & Richardson, D. C. (1988) Science 240, 1648–1652], and the N-terminal amino acid in a peptide composed of helix-forming amino acids may be considered as the N-cap residue. The choice of C-terminal amino acid has only a minor effect on helix stability. N-capping interactions may be responsible for the asymmetric distribution of helix content within a given peptide found by various workers. An acetyl group on the N-terminal α-amino function cancels the N-cap effect and the acetyl group is equivalent to N-terminal asparagine in an unacetylated peptide. Our results demonstrate a close relationship between the mechanisms of α-helix formation in peptides and in proteins.

It has been established for about 20 years that certain amino acid residues occur more frequently than others in helices, sheets, and reverse turns of globular proteins (1). Frequency of occurrence of amino acid residues in helices displays greater complexity than in other secondary structures, because the residue frequencies in a helix are different at the N-terminal end (N-cap), at the C-terminal end (C-cap), and at interior positions (1). A careful analysis by Richardson and Richardson (2) of 215 helices in 45 proteins revealed that residue frequency distributions within helices are even more complex than previously thought. The analysis by Richardson and Richardson (2) indicates that helices of proteins can be subdivided into five types of positions, each of which has a unique residue frequency distribution. These types of positions are: the two boundary positions (N- and C-caps), positions following the N-cap (N1, N2, N3, N4, and N5) or preceding the C-cap (C5, C4, C3, C2, and C1), and interior positions. The N-cap position is dominated by polar residues with small side chains (such as Asp, Asn, Ser, and Thr) as well as Gly; positions N1 to N3 display a preference for negatively charged residues; Ala is the most common residue in the helix interior; positively charged residues are found frequently at positions C3 to C1; and Gly is found very often at the C-cap position.

The obvious question that the study by Richardson and Richardson (2) evokes is: why do amino acid residues occur with particular frequencies at these positions in protein helices? One explanation, provided by Presta and Rose (3), is that the residue preferences, which are most pronounced at the helix boundaries, result from clusters of residues that initiate helix formation during protein folding. Their proposed mechanism for helix initiation involves formation of H-bonds between the side chains of polar residues flanking the helix termini and the unpaired main-chain NH and CO groups of the first and last turn of the helix. Richardson and Richardson (2), after observing many of these side-chain–main-chain H-bonds in crystal structures of proteins, suggested that the residue preferences reflect the thermodynamic role of H-bonds in stabilizing the final helical structure, and they coined the term ‘‘capping’’ to describe this type of interaction. In addition to helix capping, interactions between the charged side chains of acidic and basic residues and the α-helix macrodipole provide a different reason for the asymmetric distribution of charged residues in protein helices (4, 5).

As suggested by Presta and Rose (3), a simple method to test whether helix-capping interactions contribute to helix stability is to determine whether substituting polar residues at the ends of a helical peptide favors helix formation. The tendency of isolated helices to fray could complicate the interpretation of such an experiment. Recent experiments by Bruch et al. (6), Lyu et al. (7), and Forood et al. (8) have shown, however, that relatively large effects are observable in peptide helices that are consistent with the side-chain–main-chain H-bond hypothesis. These observations suggest that the extent of helix fraying can be greatly altered by placing appropriate residues at the ends of peptide helices.

Here we examine the propensities of uncharged nonaromatic amino acids in the N-cap and C-cap positions to stabilize the helices formed by Ala-based peptides. We find significant differences in the helix-capping propensities of these residues. Moreover, we find that helix capping must involve interactions in addition to side-chain–main-chain H-bonding.

MATERIALS AND METHODS

Peptide Synthesis. Peptides were synthesized by the solid-phase method on the Milligen model 9050 peptide synthesizer, as either peptide amides or peptide acids using Rink-resin (Advanced ChemTech) or Pepsyn KA resin (MilliGen), respectively. An active-ester-coupling procedure, employing pentafluorophenyl esters of 9-fluorenylmethoxycarbonyl amino acids, was used. The N termini were either acetylated with acetic anhydride or left unblocked. The peptides were cleaved from the resin with trifluoroacetic acid/anisole, 95/5 (vol/vol). The peptides were purified by C18 reverse-phase chromatography, and peptide identity was confirmed by fast atom bombardment/mass spectrometry and amino acid analysis. Peptide purity was assessed by analytical C18 reverse-phase chromatography using the Pharmacia FPLC system.

CD Measurements. CD measurements were made using an Aviv model 60DS spectropolarimeter and a 1.0-cm quartz cell. Measurements were made on peptides dissolved in 1.0 M NaCl/1 mM sodium borate/1 mM sodium citrate/1 mM sodium phosphate, at 0°C. pH was adjusted with 0.5 M HCl.

To whom reprint requests should be addressed.
or 0.5 M NaOH. Peptide concentration was determined by measuring tyrosine absorbance of aliquots of stock solutions diluted in 6.0 M guanidine hydrochloride solutions using $\varepsilon_{275} = 1450 \text{ M}^{–1} \text{cm}^{–1}$ (9). CD measurements are reported as mean residue ellipticity in units of deg$\cdot$cm$^{2}$$\cdot$dmol$^{–1}$, where deg is degree.

RESULTS

N- and C-Cap Positions in Peptide Helices. The N- and C-cap positions in protein helices have been defined by Presta and Rose (3) as the positions occupied by the first and last residues of the helix that form helical H-bonds and yet possess nonhelical dihedral angles. Richardson and Richardson (2) defined the N- and C-cap residues as the first and last residues in a helix whose $\alpha$-carbons are within the cylinder defined by the helix. Either definition will produce similar lists of N- and C-cap positions when applied to the protein structures in the Brookhaven Protein Data Bank. Demarcating helix boundaries in helical peptides, on the other hand, is much more difficult because the peptide does not exist as a well-populated single conformation. Instead, the peptide interconverts between helical, partly helical, and nonhelical conformations, causing the helix boundaries to fluctuate. Helix boundaries in peptides must, therefore, be defined by a distribution. The shape of the distribution depends on the amino acid sequence and the helix propensities of the amino acid residues. An example of the helix boundary distribution for a hypothetical 17-residue homopolymer containing a strong helix former is shown in Fig. 1. The distribution was calculated using Lifson-Roig theory, which assumes that the choice of N- and C-cap amino acids does not influence the stability of the helix. Note that in the distribution (Fig. 1), the N- and C-terminal amino acids have the greatest probabilities of occupying the N-cap and C-cap positions. The residues in the interior also display, however, a finite probability of being either an N-cap or C-cap residue; this effect is caused by helix fraying.

Peptide Design. It is our intention to measure the helix capping propensities of the uncharged nonaromatic amino acids by substitution experiments at the N- and C-cap positions in a helical peptide. Aromatic amino acids contribute side-chain bands to the CD spectrum around 222 nm and complicate the measurement of helix content (10). Charged amino acids contribute to helix stability by interacting with the helix dipole. All of the peptides used in this study are based on the sequence (AAKAA), and are similar to other Ala-Lys peptides previously used by this laboratory (11–14), with the exception that a single Gly residue has been inserted between the terminal Tyr and the rest of the sequence. This modification allows us to retain the Tyr residue for accurate concentration determination while eliminating the induced CD band resulting from interactions between the phenolic chromophore and the helix; the induced CD band introduces error in measurements of helix content (10).

Since Ala is known to be a strong helix former, we assumed that the helix boundary distribution of Ala-Lys peptides in 1.0 M NaCl (where electrostatic repulsion between Lys' residues is screened) resembles that of the hypothetical homopolymer in Fig. 1. Based on this assumption, we surmise that the N- and C-terminal positions should be the most sensitive sites for measuring capping propensity because they display the highest helix boundary probabilities.

Effect of N-Terminal Substitutions and N-Acetylation. We initially examined the N-cap propensities of residues with high (Asn, Ser, and Gly) and low (Val, Leu, Ala, and Pro) N-cap preferences (2) in a series of 20-residue peptides with the generic sequence: acetyl-XAKA$_4$KA$_4$KA$_4$KAY-amide. These peptides were synthesized before we discovered that aromatic contributions from the terminal Tyr residue can cause significant error in estimating helix content from ellipticity measurements at 222 nm (10); consequently, these peptides lack the penultimate Gly present in the other peptides. Differences between the [0]$_{222}$ values within this set of peptides should, however, accurately reflect relative differences in their helix contents.

The results of this experiment (Table 1) indicate that variation of the N-terminal residue in this series of peptides has no discernable effect on helix content. The differences in [0]$_{222}$ values of the peptides in this series are <1300 deg$\cdot$cm$^{2}$$\cdot$dmol$^{–1}$, which is within the ±2.5% error limit of the measurement. These results are surprising in light of the work of Forood et al. (8), who demonstrated significant N-cap effects in a series of 12-residue Ala-based peptides. Two possibilities may account for this apparent inconsistency. (i) End-capping effects may be more pronounced in shorter peptides because the relative proportions of end residues increase with decreasing length. (ii) Our peptides, unlike those of Forood et al. (8), are acetylated at the N terminus, and an H-bond formed by the acetyl-carbonyl group to a main-chain NH group could interfere with capping interactions. To test both possibilities, two sets of 12-residue peptides of the sequence XAKA$_{11}$KA$_4$AGY-amide were synthesized. The N-terminal amino group was acetylated in one set, and it was left unblocked in the other set. In turn, 11 nonaromatic uncharged residues were placed at the N terminus in both sets of peptides, and the effect of each one on helicity was examined.

To assess the effect of acetylation on helicity, the helix contents of the acetylated and unacetylated peptides must be compared at identical extents of ionization. Consequently, the helix content of each unacetylated peptide must be measured at a pH where the $\alpha$-amino group is neutral and the $\varepsilon$-amino groups of Lys residues are charged. To determine this pH, the $pK_a$ of each peptide $\alpha$-amino group was determined by pH titration of its helix content. Titration of the $\alpha$-amino group causes changes in helix content because of the interaction of the charge with the helix dipole, and this measurement can be used to calculate the $pK_a$ of the free amino acid (15), indicating that the
differences in pKₐ between residues are caused by local effects of the side chain of the N-terminal residue. The pKₐ of the ε-amino group of the Lys residues in these peptides varies between 11.3 and 11.5 at 0°C in 1.0 M NaCl. From the pKₐ data, we find that at pH 9.55 the ε-amino group in each peptide is predominantly neutral while the ε-amino group of Lys is predominantly charged. The helix contents were, consequently, measured at this pH (Table 1).

With two notable exceptions, the [θ]₂₂₂ values of the acetylated set of 12-residue peptides are around −6500 deg-cm²-dmol⁻¹. The exceptions are the Ala-peptide and Ser-peptide, which possess [θ]₂₂₂ values of −8850 deg-cm²-dmol⁻¹ and −7740 deg-cm²-dmol⁻¹, respectively. In contrast to the slight variations in helix content of the acetylated peptides, the unacetylated peptides differ vastly in their helix contents. The unacetylated Gln-peptide is practically nonhelical and has a [θ]₂₂₂ value of only −820 deg-cm²-dmol⁻¹. The unacetylated Asn-peptide is the most helical peptide of the set with a [θ]₂₂₂ value of −6900 deg-cm²-dmol⁻¹. The [θ]₂₂₂ values of the other unacetylated peptides are relatively evenly distributed between the bounds of the Asn- and Gln-peptides; there is, however, some clustering of [θ]₂₂₂ values of peptides with nonpolar N-terminal residues around −1500 to −2000 deg-cm²-dmol⁻¹. The other notable finding from this experiment is that N-acetylation increases the helix contents of all except the Asn-peptide.

**Effect of N-Terminal vs. C-Terminal Substitutions.** N-terminal substitution experiments in unacetylated 12-residue peptides show that the identity of the N-terminal amino acid influences helix content greatly (Table 1). The next logical experiment to perform was to examine the effects of varying the C-terminal amino acid. Because the 12-residue helices are quite unstable, the absolute CD signals are very small and the CD signal of the unacetylated Gln- and Val-peptides is only slightly greater than the baseline value. Thus, to obtain more reliable data, we made N- and C-terminal substitutions in 17-residue peptides, which form more stable helices. Moreover, when the C-terminal amino acid was varied, the C-terminal α-carboxyl group was not amidated, as C-terminal amidation may produce effects similar to N-terminal acetylation. The helix contents of the N-terminal substitution peptides were measured at pH 5.00, where the ε-amino group is charged, and at pH 9.55, where it is neutral (Table 2). Similarly, helix contents of C-terminal substitution peptides were measured at pH 7.00 and pH 2.00, where the α-carboxyl group is, respectively, charged and neutral.

The results of making N-terminal substitutions in unacetylated 17-residue peptides (Table 2) produced results that are qualitatively similar to those in the unacetylated 12-residue peptides (Table 1). The rank order of helix contents is similar when the ε-amino group of the peptides is either charged or neutral, with the exception of the Pro-peptide, which has a much lower relative helix content when its ε-amino group is charged than when it is neutral. The helix contents of the peptides increased by an average value of −4500 deg-cm²-dmol⁻¹ when the ε-amino group is neutralized. This effect is caused by elimination of the repulsive interaction between the charged ε-amino group and the positive pole of the ε-helix macrodipole (5).

As opposed to the large variation in helix contents we observe for N-terminal substitutions, C-terminal substitutions produced only minor changes (Table 2). Peptides with

**Table 1. Helix content of peptides, both acetylated and unacetylated, with various N-terminal amino acids**

<table>
<thead>
<tr>
<th>Residue X</th>
<th>Ac-XAKAAAAKAAAAYAKAY-COH₂ (pH 7.00)</th>
<th>Ac-XAKAAAAKAAAGY-COH₂ (pH 7.00)</th>
<th>NH₂-XAKAAAAKAAAGY-COH₂ (pH 9.55)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gln</td>
<td>—</td>
<td>5950</td>
<td>820</td>
</tr>
<tr>
<td>Val</td>
<td>28,800</td>
<td>6360</td>
<td>930</td>
</tr>
<tr>
<td>Ile</td>
<td>—</td>
<td>6350</td>
<td>1400</td>
</tr>
<tr>
<td>Ala</td>
<td>29,000</td>
<td>8850</td>
<td>1510</td>
</tr>
<tr>
<td>Met</td>
<td>—</td>
<td>6050</td>
<td>1570</td>
</tr>
<tr>
<td>Pro</td>
<td>27,800</td>
<td>6290</td>
<td>1935</td>
</tr>
<tr>
<td>Leu</td>
<td>28,900</td>
<td>6490</td>
<td>2070</td>
</tr>
<tr>
<td>Thr</td>
<td>—</td>
<td>6640</td>
<td>3590</td>
</tr>
<tr>
<td>Gly</td>
<td>27,700</td>
<td>6450</td>
<td>4610</td>
</tr>
<tr>
<td>Ser</td>
<td>28,800</td>
<td>7740</td>
<td>5030</td>
</tr>
<tr>
<td>Asn</td>
<td>28,000</td>
<td>6680</td>
<td>6900</td>
</tr>
</tbody>
</table>

Conditions: 0°C and 1.0 M NaCl. Peptides in the first column have 20 residues; those in the second and third columns have 12 residues.

**Table 2. N- and C-terminal substitutions in 17-residue peptides**

<table>
<thead>
<tr>
<th>Residue X</th>
<th>NH₂-XAKAAAAKAAAAYAKAY-COH₂ (pH 5.00)</th>
<th>NH₂-XAKAAAAKAAAAYAKAY-COH₂ (pH 9.55)</th>
<th>Ac-YGAAKAAAAKAAAAKAX-COH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gln</td>
<td>7,000</td>
<td>10,000</td>
<td>15,500</td>
</tr>
<tr>
<td>Ala</td>
<td>7,900</td>
<td>12,900</td>
<td>15,200</td>
</tr>
<tr>
<td>Val</td>
<td>9,040</td>
<td>12,800</td>
<td>13,750</td>
</tr>
<tr>
<td>Met</td>
<td>10,200</td>
<td>13,700</td>
<td>15,300</td>
</tr>
<tr>
<td>Pro</td>
<td>8,000</td>
<td>13,800</td>
<td>16,600</td>
</tr>
<tr>
<td>Ile</td>
<td>10,500</td>
<td>14,300</td>
<td>13,500</td>
</tr>
<tr>
<td>Leu</td>
<td>10,800</td>
<td>15,300</td>
<td>15,030</td>
</tr>
<tr>
<td>Thr</td>
<td>12,200</td>
<td>15,600</td>
<td>13,050</td>
</tr>
<tr>
<td>Gly</td>
<td>14,400</td>
<td>17,900</td>
<td>15,200</td>
</tr>
<tr>
<td>Ser</td>
<td>13,100</td>
<td>17,900</td>
<td>13,800</td>
</tr>
<tr>
<td>Asn</td>
<td>17,100</td>
<td>20,300</td>
<td>14,000</td>
</tr>
</tbody>
</table>

Conditions: 0°C and 1.0 M NaCl.
DISCUSSION

Helix Capping by Amino Acids and the Acetyl Group. The N-terminal substitution experiments in 12-, 17-, and 20-residue peptides clearly show that the N-terminal unit of the peptide, whether it is an amino acid or an acetyl group, can have a profound effect on the extent of helix formation (Tables 1 and 2). N-terminal substitutions in 17-residue peptides provide the most reliable data for evaluating the rank order of N-cap propensities of the nonaromatic uncharged amino acids (Table 2); that order is Asn > (Ser, Gly) > (Thr, Leu, Ile) > (Pro, Met) > Val > Ala > Gln. Inspection of the helix contents of acetylated and unacetylated 12-residue peptides indicated that the helix-stabilizing effect of the acetyl group is roughly equivalent to that of Asn in the N-cap position (Table 1).

In contrast to the varied N-cap effects, the observed C-cap effects of the uncharged nonaromatic residues are very uniform (Table 2). Based on the C-terminal substitution data in Table 2, the rank order of C-cap propensities is (Glu, Ala, Met, Leu, Gly) > Asn > (Ser, β-branched residues). It should be stressed, however, that the differences in helix contents of these peptides are very small: there is only a 20% difference between the highest and lowest values. We therefore conclude that there are only minor differences in the C-cap propensities of the nonaromatic uncharged amino acids in isolated α-helices.

The large changes in helix content caused by varying the N-terminal residue are not predicted by classical helix-coil theories, such as the Lifson–Roig theory (16). In fact, Lifson–Roig theory predicts that substitution of a helix breaker or helix former into a helical peptide produces the greatest effect at the middle, and substitutions at either end are predicted not to have any effect on helix stability. We have modified the Lifson–Roig theory to include end effects (A.J.D., A.C., T. M. Klingler, and R.L.B., unpublished data).

One additional point should be made about the measurement of helix-capping propensities by substitution experiments. Although these experiments allow measurement of the rank orders of N- and C-cap propensities, they cannot be used to compare N-cap propensity with C-cap propensity. For example, the substitution data show that Gly is helix-stabilizing relative to Ala at the N terminus and equivalent to Ala at the C terminus. These data cannot, however, be used to determine whether the N-cap propensity of Gly is greater than its C-cap propensity.

Helix-Capping Propensities Measured in Other Systems. Helix-capping propensities of a select number of nonaromatic uncharged residues have been measured in other peptide (7, 8) and protein (17–19) systems. Because these workers have not examined the same set of amino acids, a proper comparison is not possible. Nevertheless, there is a general consensus that the small polar residues (Asn, Ser, and Thr) and Gly have greater N-cap propensities than Ala in each of the systems. We also find (Fig. 2) that the N-terminal substitution data from 17-residue peptides at pH 9.55 correlate with the N-cap preference of those residues in protein helices (2). Even though the relationship between helix content and N-cap preference is probably not linear, the good correlation indicates that N-cap propensity measured in peptides is a basic determinant of where helices start in proteins.

Is N-Capping Responsible for Helix Asymmetry? There is accumulating evidence that the distribution of helicity in α-helical peptides with rather symmetrical sequences can be quite skewed. Using NMR spectroscopy, Lyu et al. (20) studied a 21-residue peptide with the sequence: succinyl-YSE_{4}K_{4}X_{4}E_{4}K_{4}-amide, where X is Ala, Leu, or Gly. The CαH chemical shifts suggested that residues near the N terminus were more helical than corresponding residues near the C terminus. Miick et al. (21) studied a series of Ala-based peptides [sequence, acetyl-(AAAALK)_{n}amide] where a nitroxide spin label was incorporated at various sites along the sequence. Measurements of rotational correlation time of the spin label by ESR spectroscopy indicated that the C terminus was more mobile than the N terminus, suggesting that the N-terminal half of the peptide was more helical than the C-terminal half. Both peptides contain units that can participate in strong N-cap interactions. The peptides of Lyu et al.

![Graph showing correlations between helix content of 17-residue N-substituted peptides and N-cap preference in proteins. The helix content at pH 9.55 of various 17-residue peptides with a particular residue at the N terminus (data from Table 2) is plotted against the N-cap preference value for that amino acid [data are from Richardson and Richardson (2)].](image-url)
(20) contain a Ser residue that has a high N-cap propensity at position 2. The peptides of Miick et al. (21) are acetylated at the N terminus. The results of our study indicate that the acetyl group possesses strong N-capping characteristics (Table 1). Consequently, the asymmetric distribution of helicity in these peptides may be caused by N-cap interactions that reduce the extent of N-terminal fraying.

Possible Mechanisms of N-Capping Interactions. What determines the rank order of N-cap propensities? Presta and Rose (3) attribute N-capping to formation of side-chain–main-chain H-bonds at the N-terminal end of the helix. Consistent with their idea, our measurements indicate that residues with polar side chains such as Asn, Ser, and Thr have high N-cap propensities. Our data also indicate, however, that Glu has the lowest N-cap propensity and it differs by only a single methylene group from Asn, the residue with the highest N-cap propensity. This interesting result can also be reconciled with the side-chain–main-chain H-bond hypothesis if it is assumed that, with Glu present, side-chain–main-chain H-bonds stabilize nonhelical conformations, whereas with Asn present, side-chain–main-chain H-bonds stabilize helical conformations. The side chain of Glu is known to form a H-bond with its own peptide NH, both in unstructured peptides (22) and in a protein (23). Lyu et al. (7) detected long-range nuclear Overhauser effect connectivities that are consistent with a side-chain–main-chain H-bond conformation at the N terminus of their helical peptide.

Our data indicate that Gly has a very high N-cap propensity; this has also been observed in barnase by Serrano et al. (17, 18). The high N-cap propensity of Gly cannot be explained by the side-chain–main-chain H-bond hypothesis. Serrano et al. (17, 18) have, however, provided an alternative explanation that may also explain our data. They propose that side chains of nonpolar residues at the N-cap position cause steric hindrance to solvation of non-H-bonded NH groups of the first turn of the helix. Because Gly does not have a side chain, it should not hinder solvation of the NH groups, and hence, it will have a higher N-cap propensity than the nonpolar residues. Our data indicate that the nonpolar residues differ significantly in N-cap propensity. Neither the mechanism of Presta and Rose (3) nor that of Serrano et al. (17, 18) can explain the rank order of N-cap propensities of the nonpolar residues, which is Leu, Ile > Met > Val > Ala. Because the rank order correlates with the size of the alkyl side chain, burial of hydrophobic surface may contribute to the N-cap propensity.

Concluding Remarks. A role for helix-capping interactions in specifying the locations of helices in proteins has recently been proposed (24). Our study provides experimental evidence to support this role. We find that the helix-capping propensities of the nonaromatic uncharged residues vary widely, that there are only minor differences in C-cap propensities, that N-cap interactions can make large contributions to helix stability, and that N-capping involves other interactions in addition to side-chain–main-chain H-bonding.

We thank the Mass Spectrometry Facility, University of California, San Francisco (supported by a National Institutes of Health grant) for verification of peptide identity by mass spectrometry. This work was supported by National Institutes of Health Grant GM 31475. A.J.D. thanks the Science and Engineering Research Council (United Kingdom) for a North Atlantic Treaty Organization postdoctoral fellowship.