Structure of a C-Terminal α-Helix Cap in a Synthetic Peptide

Hongxing X. Zhou,† Pingchiang C. Lyu,‡
David E. Wemmer,‡ and Neville R. Kallenbach*,†

Departments of Chemistry
New York University, 4 Washington Place
New York, New York 10003
University of California at Berkeley
Berkeley, California 94720

Received September 15, 1993

The ends of α-helices play a crucial role in stabilizing the helical structure in proteins.[1] While a central amino acid in a long α-helix has both its NH and CO groups H-bonded to peptide CO and NH groups four residues away,[2] at the ends of a helix the first four NH donors and the last four CO acceptors remain unsatisfied. The mean length of α-helices in globular proteins is only about 12 amino acids,[1] so the groups at the ends actually constitute a majority of helical residues present. Polar side chains or peptide groups next to flexible glycine residues can interact with the main chain at the ends so as to fulfill the H-bonding potential of these NH and CO groups.[1,3-7] This process is referred to as helix capping. The sequences at the ends of helices in proteins reflect strong preferences for particular side chains in capping positions.[4] N-Terminal helix capping has been demonstrated by sequence analysis and site-directed mutations of native proteins[6] and in synthetic peptide models.[7] The presence of an unusual type of H-bond at the C-terminus of the H helix in myoglobin was first noted by Watson.[8] Subsequently, Schellman reported that about one-third of helices in proteins end with a residue in a left-handed (α_L) conformation, with positive ϕ,ψ angles favored by the presence of a Gly residue.[3a] This analysis has since been confirmed and extended.[3b-e] Experimentally, a C-cap mutation from Gly to Ala destabilizes barnase by 3.1 kcal/mol.[6b] How amino acids other than glycine participate in C-terminal capping is not so clearly understood, and experimental evidence in isolated helices is lacking.

We report here a novel C-terminal capping structure in a peptide helix, in which the NH of the side chain of asparagine forms an H-bond with the helix main chain CO four residues away. The backbone forms a local 3_{10} helix at the C-terminus, with the side chain contributing an additional H-bonded loop. This structure reveals formation of H-bonds by the side chain and main chain of a single residue that serve as a fundamental signal at the C-terminus of helices. The structure formed in this way blocks continuation of the α-helix, hence providing a stronger C-termination signal than Pro 19, as seen in the relative CD values in Table 1.

Table 1. Peptide Sequences and Relative Helical Content[a]

ac-YMSEDELKAAEAAFKRH X PT-NH_2				
peptide	X	$-[\theta]_{222}$[b]	f,[c] %	P[d]
N18	Asn	20 030	63	1.6
Q18	Gln	17 930	56	0.9
A18	Ala	17 360	54	0.8
S18	Ser	17 370	54	0.8
T18	Thr	15 950	50	0.3
G18	Gly	16 000	50	3.9

[a] The peptide sequence includes one of the most frequently occurring residues at the corresponding position according to the Richardson and Richardson survey[4] except that the middle four residues of the sequence contain a glutamic acid to increase the peptide solubility and permit a salt bridge with the lysine 15.[11] Apart from alanine, neutral residues with H-bond donor side chains were substituted at position 18. Single-letter abbreviations for these amino acids are A, Ala; D, Asp; E, Glu; F, Phe; G, Gly; H, His; K, Lys; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; Y, tyr. CD measurements and peptide synthesis were performed as described.[7b] [b] $[\theta]_{222}$, mean residue ellipticity (deg-cm^2-dmol^{-1}) at 222 nm, 4 °C, pH = 6. The independence of $[\theta]_{222}$ on peptide concentration of N18 was checked between 30 and 300 μM. [c] f, $([\theta]_{obsd}/-32\,000) \times 100\%$ = percent helix content.[9] [d] Preferences of the amino acids at the C-cap position according to Richardson and Richardson.[4]

The sequences of several synthetic peptides of consensus sequence containing C-terminal substitutions are shown in Table 1,[4,7a] together with their apparent helix content determined by CD spectroscopy in the far UV region.[9] The asparagine side chain enhances the apparent helix content significantly over either Ala or Gly; as noted, Gly occurs with very high frequency at the C-termini of helices in proteins because of its flexibility.[4]

How asparagine enhances the apparent helix structure in N18 was investigated by ^{1}H NMR spectroscopy. Identification of the spin systems and sequential assignment of the peptide were carried out using 2D TOCSY and NOESY experiments.[10,11] A clear pattern of cross peaks links each amide from residue 4 in the sequence to the NH of Asn 18, with longer range connectivities between the C_αH and amide NH from residue 5 to residue 14 (data not shown). Together with the α-helical CD spectrum of N18 (not shown), these are sufficient to identify the structure of the molecule as an α-helix, as in the related peptide S3.[7b]

Figure 1 shows three regions of the NOESY spectrum that provide details of the structure of the C-terminus of the molecule. The following NOE cross peaks define a C-cap structure involving the side chain of Asn at position 18. (i) The two β protons of Asn 18 have distinct chemical shifts; one proton (β′) is closer to its peptide NH than the other (β) (panel a). (ii) One side chain amide proton of Asn 18 is close to the C_αH of Phe 14 (panel b) and to its own backbone amide (panel c). A separate ROESY experiment showed that these weak cross peaks are not due to spin diffusion (data not shown). With the interamide connectivities, these data allow us to construct a model for the C-cap of the peptide. As seen in Figure 2, an H-bond between one of the side-chain amide protons of Asn 18 and the CO of Phe 14 accounts for the NOEs described (indicated by curves with arrows). The peptide backbone NH of Asn 18 bonds with the CO of Lys 15, forming a 3_{10} structure; this bond might bifurcate to include an interaction with the CO of Phe 14 as well. However, the side-chain amide of Asn 18 bonds only to the CO of Phe 14.

* To whom correspondence should be addressed.
† New York University.
‡ University of California at Berkeley.

(1) Presta, L. G.; Rose, G. D. Science **1988**, *240*, 1632.
(2) Pauling, L.; Corey, R. B.; Branson, H. R. *Proc. Natl. Acad. Sci. U.S.A.* **1951**, *37*, 205.
(3) (a) Schellman, C. in *Protein Folding*; Jaenicke, R., Ed.; Elsevier/North Holland Biomedical Press: Amsterdam, 1980; p 53. (b) Milner-White, E. J. *J. Mol. Biol.* **1988**, *199*, 503. (c) Preissner, R.; Bork, P. *Biochem. Biophys. Res. Commun.* **1991**, *180*, 660; (d) **1991**, *180*, 666. (e) Nagarajaram, H. A.; Sowdhamini, R.; Ramakrishnan, C.; Balaram, P. *FEBS* **1993**, *321*, 79.
(4) Richardson, J. S.; Richardson, D. C. *Science* **1988**, *240*, 1648.
(5) (a) Harper, E.; Rose, G. D. *Biochemistry* **1993**, *32*, 7605. (b) Zhou, H. X.; Lyu, P. C.; Wemmer, D. E. Kallenbach, N. R. *Protein Struct. Funct. Genet.*, in press.
(6) (a) Lecomte, J. T. J.; Moore, C. D. *J. Am. Chem. Soc.* **1991**, *113*, 9663. (b) Serrano, L.; Fersht, A. R. *Nature* **1989**, *342*, 296. (c) Serrano, L.; Neira, J.-L.; Sancho, J.; Fersht, A. R. *Nature* **1992**, *356*, 453.
(7) (a) Lyu, P. C.; Zhou, H. X.; Jelveh, N.; Wemmer, D. E.; Kallenbach, N. R. *J. Am. Chem. Soc.* **1992**, *114*, 6560. (b) Lyu, P. C.; Wemmer, D. E.; Zhou, H. X.; Pinker, R.; Kallenbach, N. R. *Biochemistry* **1993**, *32*, 421. (c) Forood, B., Feliciano, E. J.; Nambiar, K. P. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 838. (d) Bruch, M. D.; Dhingra, M. M.; Gierasch, L. M. *Protein Struct. Funct. Genet.* **1991**, *10*, 130.
(8) Watson, H. C. *Prog. Stereochem.* **1969**, *4*, 299.
(9) Gans, P. J.; Lyu, P. C.; Manning, M. C.; Woody, R. W.; Kallenbach, N. R. *Biopolymers* **1991**, *31*, 2283.
(10) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, 1986.
(11) (a) Perutz, M. F.; Fermi, G. *Protein Struct. Funct. Genet.* **1988**, *4*, 294. (b) Marqusee, S.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8898. (c) Lyu, P. C.; Gans, P. J.; Kallenbach, N. R. *J. Mol. Biol.* **1992**, *223*, 343.

0002-7863/94/1516-1139$04.50/0

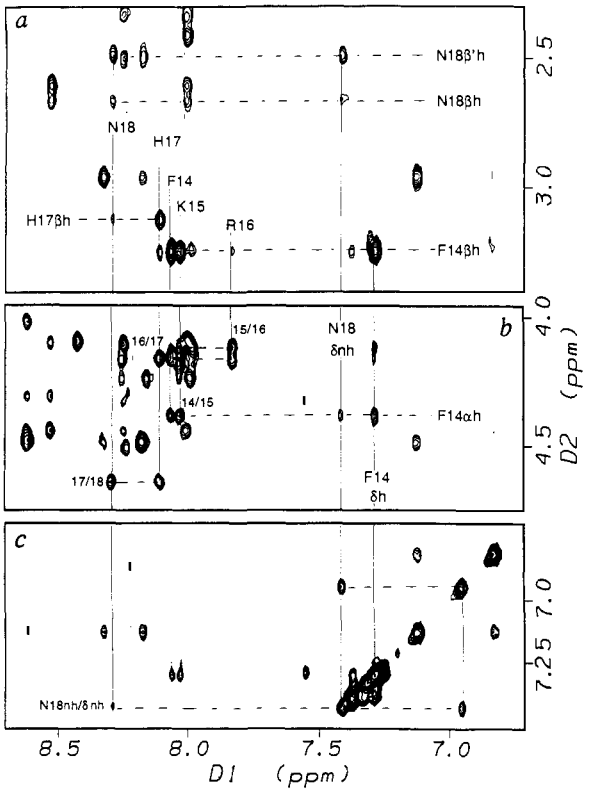

Figure 1. (a) Amide β Hs, (b) amide α Hs, and (c) aromatic and asparagine side-chain amide regions of N18 NOESY. Each solid line connects the cross peaks of amide protons to α and β protons. The dashed lines connect NOE cross peaks in the Ccap region. The numbers labeled in (b) show the sequential $d_{\alpha N}(i,i+1)$ cross peaks in the C-cap region. This spectrum was taken on a Bruker AMX-600 in 85% H_2O/15% D_2O at 10 °C, pH = 5.5, concentration about 4 mM. The mixing time τ_m = 250 ms.

The weak NOE connecting the N18 backbone amide to the β proton of His 17 (panel a) is seen in internal α-helical structure[10] and indicates that the C-terminal cap inhibits the progressive fraying of this end of the helix.[5b,12] This cross peak is usually not observed in peptides with normal fraying at the end.

According to the substitution series (Table 1), the order of helix stabilization at the C-terminus is Asn > Gln > (Ala, Ser) > (Thr, Gly). We interpret this from our model as follows. (i) Gln has one methylene unit more than Asn, so H-bond formation is entropically less favorable. (ii) Ala still has a stabilizing effect on helical structure at the C-terminus and so does not destabilize maximally. (iii) Ser and Thr share the potential to donate γH to form a H-bond with CO of Phe 14, although we have no direct evidence for this. (iv) Gly should stabilize helix better than Ala at the C-terminus,[3,4] but this is not observed in these peptides. We believe the reason is that Gly is followed by Pro, which has

(12) Liff, M. I.; Lyu, P. C.; Kallenbach, N. R. *J. Am. Chem. Soc.* **1991**, *113*, 1014.

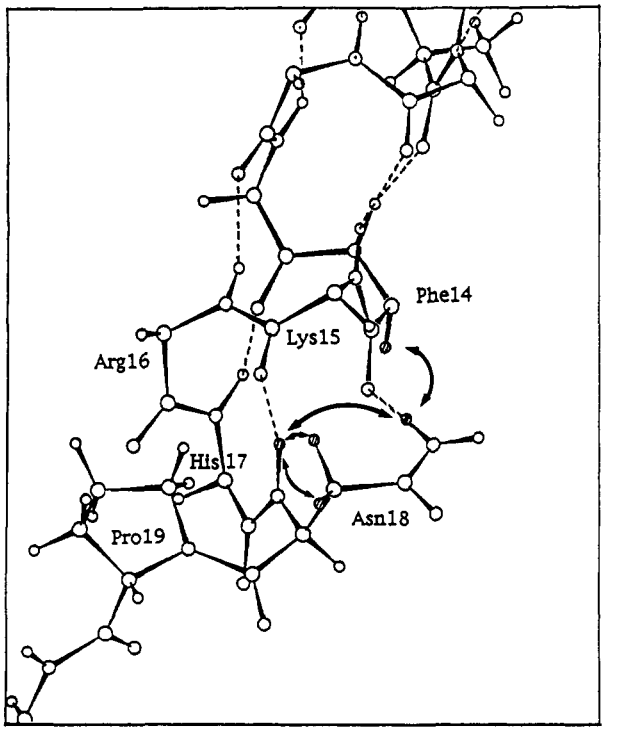

Figure 2. Molecular model illustrating the C-terminus cap structure in N18 peptide. Curves with arrows indicate protons showing NOE connectivities.

no amide H-bonding potential. A significant fraction of protein helices terminate in Gly,[3] which allows the formation of two H-bonds using main-chain NH donors.[3a] The model in Figure 2 shows that the amide side chain of Asn can replace the amide NH from the residue following Gly, bypassing the need for the left-handed (α_L) backbone conformation. The common feature in N-[5b] and C-caps demonstrated so far in model peptides is that H-bonds between a side chain and the backbone restrict the conformational freedom present at the ends in normal helices because of the unfulfilled H-bond potential of the terminal NH and CO of the main chain.[1] The structure of both these terminal caps prevents normal α-helix propagation beyond them, so that they terminate helical structure as well. These represent strong signals for locating helical structure in an amino acid sequence.

Acknowledgment. We thank George Rose and Oleg Ptitsyn for helpful discussions and insights. This work was supported by research grants GM40746 from the NIH, DMB 88-15998 from the NSF, and HFSP RE-331/93 from the International Human Frontier of Science Program. Instrumentation grants DE FG05-86ER75281 (U.S. DOE) and DMB 86-09305 and BBS 87-20134 (NSF) are also gratefully acknowledged. P.C.L. holds a NRSA postdoctoral fellowship from the N.I.H. (GM 14803); H.X.Z. is a Sokol Fellow.