Biochemistry

 Volume 32, Number 30 August 3, 1993

Accelerated Publications

Helix Stop Signals in Proteins and Peptides: The Capping Box†

Edwin T. Harper‡ and George D. Rose*

Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, Box 8231, 660 South Euclid Avenue, St. Louis, Missouri 63110

Received March 23, 1993; Revised Manuscript Received May 17, 1993

ABSTRACT: The α-helix [Pauling, L., Corey, R. B., & Branson, H. R. (1951) *Proc. Natl. Acad. Sci. U.S.A. 37*, 205–211] is a common motif in both proteins and peptides. Despite intense investigation, predictive understanding of helices is still lacking. A recent hypothesis [Presta, L. G., & Rose, G. D. (1988) *Science 240*, 1632–1641] proposed that the structural specificity of helices resides, in part, in those residues that flank helix termini. If so, then signals that arrest helix propagation—*i.e.*, helix stop signals—should be found among these flanking residues. Evidence is presented for the existence of one such signal, a reciprocal backbone–side-chain hydrogen-bonding interaction, dubbed the *capping box*. In proteins, the capping box is found *uniquely* at helix N-termini. In peptides, the capping box can function as a helix stop signal, as shown in the work of Kallenbach and co-workers.

The α-helix—first proposed as a model structure by Pauling *et al.* (1951) and immediately confirmed in ongoing X-ray studies (Perutz, 1951)—is a common motif in globular proteins. The principal defining characteristic of the α-helix consists of intrasegment main-chain hydrogen bonds between successive amide hydrogen donors and carbonyl oxygen acceptors situated four residues previously in sequence (Pauling *et al.*, 1951), *i.e.*, >NH(*i*)···O=C< (*i* − 4). For the helix of average length (*i.e.*, ∼12 residues), this pattern results in eight intrasegment H-bonds. Unavoidably, end effects deprive the initial four amide hydrogens and final four carbonyl oxygens of intrasegment main-chain hydrogen bonds because, upon termination, no next turn of helix exists to provide such partners (Presta & Rose, 1988). Summing, Pauling–Corey–Branson hydrogen bonds account for only 50% of the total in the helix of average length, with the first four >NH groups and last four >C=O groups accounting for the remaining 50%.

It has been hypothesized that hydrogen bond partners for these otherwise unsatisfied first four amide hydrogens and last four carbonyl oxygens are typically provided by polar groups that flank the helix termini, particularly side-chain polar groups (Presta & Rose, 1988). This hypothesis, which was based upon exhaustive stereochemical modeling, was corroborated by a study showing sharply differentiated positional preferences surrounding helix termini (Richardson & Richardson, 1988), and additional evidence—both positive (Serrano & Fersht, 1989; Bruch *et al.*, 1991; Lecomte & Moore, 1991; Lyu *et al.*, 1992, 1993; Serrano *et al.*, 1992; Forood *et al.*, 1993) and negative (Fairman *et al.*, 1989; Heinz *et al.*, 1993)—has been accumulating during the last 5 years.

Upon further investigation, we find that helix termination in proteins appears to utilize just a few, distinct hydrogen-bonding patterns. In this paper we report on one of these, a reciprocal backbone–side-chain hydrogen-bonding interaction dubbed *the capping box*. As shown below, capping boxes are found *uniquely* at helix N-termini.

IDENTIFICATION OF HELICES

In proteins, helical segments have backbone dihedral angles with φ,ψ angles near the observed mean values of −64° ± 7°, −41° ± 7°, respectively (Presta & Rose, 1988). Helix boundaries are defined by the residues at either end—termed Ncap and Ccap by Richardson and Richardson (1988). Each Ncap and Ccap residue makes one additional intrahelical

† Supported by NIH Grant GM 29458.

* To whom correspondence should be addressed.

‡ Permanent address: Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, 635 Barnhill Drive, Indianapolis, IN 46202-5122.

0006-2960/93/0432-7605$04.00/0

hydrogen bond but departs from helical values of ϕ,ψ angles (Presta & Rose, 1988). Helices and their flanking residues are labeled as follows:

...-N″-N′-Ncap-N1-N2-N3-...-C3-C2-C1-Ccap-C′-C″-...

where N1 through C1 are the residues with helical values of ϕ,ψ angles and the primed residues belong to structures that bracket the helix at either end. Ncap and Ccap serve as bridge residues that belong to both the helix and the adjacent turn. The eight terminal residues lacking intrahelical main-chain H-bond partners are Ncap through N3 and C3 through Ccap, inclusively.

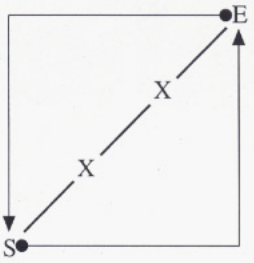
In the current study, 161 α-helices were identified within a set of 42 high-resolution protein structures (resolution ≤ 2.0 Å; R-factor ≤ 20) selected from the Protein Data Bank (Bernstein *et al.*, 1977). Helices were identified as sequences of at least seven consecutive residues including Ncap and Ccap, with backbone dihedral angles of residues N1 through C1 near the observed mean values for α-helices, and containing at least three $i \rightarrow i-4$ backbone hydrogen bonds. The Ncap residue was defined as the first residue of this sequence with an $i \rightarrow i-4$ backbone hydrogen bond. Hydrogen bonds were identified using criteria derived from small-molecule crystal studies, as described in Stickle *et al.* (1992). Briefly, a donor and acceptor with suitable orientation angles were classified as forming an H-bond when their interatomic distance did not exceed the sum of their respective van der Waals radii.

Proteins used in this and an earlier study (Stickle *et al.*, 1992) and their parenthesized Brookhaven identifiers (Bernstein *et al.*, 1977) are as follows: actinidin (2ACT); α-chymotrypsin A (5CHA); amylase inhibitor (1HOE); avian pancreatic polypeptide (1PPT); azurin (2AZA); carbonic anhydrase (2CA2); carboxypeptidase A α (5CPA); citrate synthase (2CTS); crambin (1CRN); cytochrome c (reduced) (5CYT); cytochrome c_3 (2CDV); cytochrome c-551 (oxidized) (351C); D-glyceraldehyde-3-phosphate dehydrogenase (1GD1); dihydrofolate reductase (4DFR); erythrocruorin (deoxy) (1ECD); flavodoxin (semiquinone form) (4FXN); γ-II crystallin (1GCR); glutathione peroxidase (1GP1); glutathione reductase (3GRS); hemerythrin (met) (1HMQ); hemoglobin (cyano, met) (2LHB); IG Bence-Jones protein (2RHE); insulin (3INS); lysozyme (human) (1LZ1); lysozyme (phage) (3LZM); myoglobin (oxy) (1MBO); ovomucoid third domain (2OVO); papain (9PAP); penicillopepsin (3APP); pepsin (4PEP); phospholipase A_2 (1BP2); plastocyanin (apo) (2PCY); ribonuclease A (7RSA); ribonuclease T_1 (3RNT); rubredoxin (1RDG); scorpion neurotoxin (1SN3); staphylococcal nuclease (1SNC); thermolysin (3TLN); trp repressor (2WRP); trypsin (β) (1TPP); trypsin inhibitor (5PTI); and ubiquitin (1UBQ).

THE CAPPING BOX IN PROTEINS

Inspection of hydrogen-bonding patterns at helix termini reveals a recurring motif wherein the side chain of the Ncap residue forms an H-bond with the backbone >NH of N3 and, reciprocally, the side chain of N3 forms an H-bond with the backbone >NH of Ncap (see Figure 1). This boxlike, hydrogen-bonded cycle caps two of the initial four backbone >NH donors of the helix, prompting the name *capping box*.

As shown below, *all* such reciprocal $i \rightarrow i+3$ and $i \rightarrow i-3$ side-chain–backbone hydrogen-bonding interactions found in proteins are localized between Ncap and N3 in helices. In greater detail, among the 42 proteins studied here, 161 α-helices were identified with length seven residues or more. This set includes 15 capping boxes, with 3 additional boxes found among shorter helices. Described initially by Baker and Hubbard (1984), the uniqueness of the capping box has led us to further characterize its frequency, geometry, likelihood, and potential significance, as described below.



SXXE Box

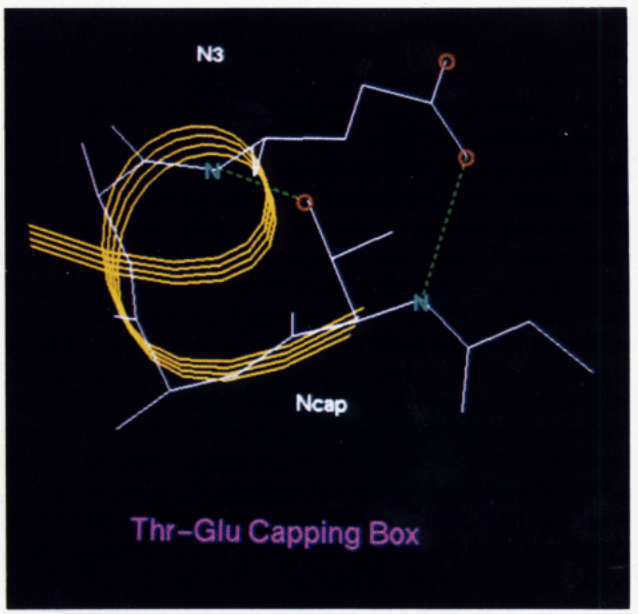


FIGURE 1: The capping box, a hydrogen-bonded cycle localized at the N-terminus of helices. In a capping box, the side chain of Ncap forms an H-bond with the backbone >NH of N3, and reciprocally, the side chain of N3 forms an H-bond with the backbone >NH of Ncap. (a, top) Diagrammatic representation of a Ser-Xaa-Xaa-Glu capping box. Covalent connectivity runs along the diagonal, with each side-chain–backbone H-bond depicted by two sides of the "box". (b, bottom) Computer-generated representation of a Thr-Xaa-Xaa-Glu capping box. The helix is depicted by the yellow ribbon, with residues Ncap and N3 shown in atomic detail. Side-chain acceptor atoms are red, backbone donors are blue, and H-bonds between the two are green broken lines.

A hydrogen bond can be classified by the sequential distance from the donor to the acceptor residue; *e.g.*, the canonical backbone helical H-bond is $i \rightarrow i-4$. In a capping box, where the H-bond donors are backbone >NH groups, the H-bond from the backbone >NH of the N3 residue to the side chain of Ncap is $i \rightarrow i-3$, while the complementary H-bond from the >NH of Ncap to the side chain of N3 is $i \rightarrow i+3$. It should be noted that the convention adopted by Baker and Hubbard (1984) has the opposite sign.

Table I enumerates all $i \rightarrow i+3$ and $i \rightarrow i-3$ backbone >NH to side-chain H-bonds in the set of 42 proteins. This global inventory is subdivided to distinguish H-bonds between Ncap and N3 residues from other positions in helices. A key feature of the table is the demonstration that, in helices, backbone >NH to side-chain acceptor H-bonds are located predominantly between Ncap and N3 residues. There, 94% of the 64 $i \rightarrow i-3$ H-bonds and 100% of the 20 $i \rightarrow i+3$ H-bonds are found between Ncap and N3. Within these 84 H-bonds, 15 capping boxes are found, *i.e.*, reciprocal $i \rightarrow i \pm 3$ H-bonds between the same pairs of residues (Table II).

Table I: Number of $i \rightarrow i-3$ and $i \rightarrow i+3$ Backbone >NH to Side-Chain H-Bonds[a]

	$i \rightarrow i-3$	$i \rightarrow i+3$	total
helix	64	20	84
Ncap-N3[b]	60	20	80
other positions	4	0	4
nonhelix	45	28	73
whole protein	109	48	157

[a] Listed as hydrogen bond donor → acceptor; $i \rightarrow i-3$ and $i \rightarrow i+3$ identify acceptors three positions before and after the donors, respectively. [b] Ncap and N3 are the first and fourth positions of the helix, respectively (see helix nomenclature in the text).

The Ncap and N3 side chains in the identified set of capping boxes have a unique conformation. Figure 2 illustrates the uniformity of side-chain conformations among 10 capping boxes with either Glu or Gln at N3.

The characteristic geometry of capping boxes depends upon nonhelical values of the dihedral angles at Ncap, ψ in particular. In capping boxes, values of these angles are clustered around $\phi = -94° \pm 15°$ and $\psi = 167° \pm 5°$ (Table II). The distribution of Ncap ψ angles suggests that this geometry is associated with the possibility of N-capping interactions between Ncap and N3. In the set of 161 helices, 52 Ncap residues have ψ angles within two standard deviations of the mean for capping boxes; of these, 32 (62%) are in Ncap-N3 pairs where both side chains are capable of capping an >NH group, while only 2 (4%) are found in pairs having no side chains suitable for capping interactions. Viewed from the converse perspective, (coincidentally) 52 Ncap-N3 pairs have both side chains suitable for capping, and 26 (50%) of these have Ncap ψ values within two standard deviations of the mean for capping boxes, while 33 (91%) of the 36 Ncap-N3 pairs without capping side chains have an Ncap ψ more than three standard deviations

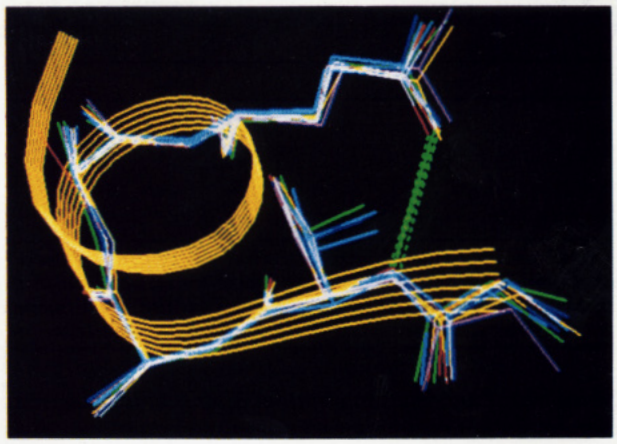


FIGURE 2: Superposition of 10 capping boxes that have either Glu or Gln at N3. Similar to Figure 1, the helix is represented by the yellow ribbon, with residues Ncap and N3 shown in atomic detail. Backbone atoms for nine boxes have been root-mean-square best-fit to an arbitrarily chosen initial box. It is apparent that side-chain residues have a unique conformation in capping boxes (see also Table II).

Table II: Comparison of Observed and Potential Capping Boxes

protein[a]	Ncap to N3 residues[b]	sequence				Ncap dihedrals[c]	
		Ncap	N1	N2	N3	ϕ	ψ
			Observed Capping Boxes				
351C	67–70	Ser	Asp	Asp	Glu	–90	169
1BP2	89–92	Asn	Ala	Cys	Glu	–75	170
2CA2	219–222	Ser	Ser	Glu	Gln	–77	162
5CPA	14–17	Thr	Leu	Asp	Glu	–85	174
2CTS	37–40	Thr	Val	Asp	Met	–101	169
2CTS	70–73	Ser	Ile	Pro	Glu	–74	158
5CYT	60–63	Asn	Asn	Asp	Thr	–156	182
3GRS	383–386	Thr	Glu	Asp	Glu	–78	169
3GRS	456–459	Thr	Lys	Ala	Asp	–108	174
2LHB	12–15	Ser	Ala	Ala	Glu	–95	157
3LZM	59–62	Thr	Lys	Asp	Glu	–97	168
1MBO	3–6	Ser	Glu	Gly	Glu	–77	161
1MBO	51–54	Thr	Glu	Ala	Glu	–108	158
3RNT	12–15	Ser	Ser	Ser	Asp	–95	170
2WRP	44–47	Thr	Pro	Asp	Glu	–88	163
no. = 15						mean –94 ± 15	mean 167 ± 5
			Potential Capping Boxes				
3APP	139–142	Thr	Phe	Phe	Asp	–68	165
1BP2	39–42	Asp	Asp	Leu	Asp	–160	185
5CHA	164–167	Ser	Asn	Thr	Asn	–74	158
2CTS	208–211	Asp	Trp	Ser	His	–82	156
1ECD	93–96	Thr	His	Asp	Gln	–100	166
1GD1	36–39	Asp	Ala	Asn	Thr	–94	165
1GD1	148–151	Ser	Cys	Thr	Thr	–98	166
1GD1	251–254	Thr	Val	Glu	Glu	–114	164
1LZ1	4–7	Glu	Arg	Cys	Glu	–82	156
1LZ1	24–27	Ser	Leu	Ala	Asn	–70	155
3LZM	2–5	Asn	Ile	Phe	Glu	–148	173
2OVO	33–36	Asn	Lys	Cys	Asn	–168	169
5PTI	47–50	Ser	Ala	Glu	Asp	–141	151
3TLN	280–283	Asn	Phe	Ser	Gln	–108	180
1TPP	164–167	Ser	Asp	Ser	Ser	–73	155
1UBQ	22–25	Thr	Ile	Glu	Asn	–84	160
2WRP	67–70	Ser	Gln	Arg	Glu	–80	165
no. = 17						mean –103 ± 32	mean 164 ± 9

[a] Identified by the four-character Protein Data Bank code (Bernstein *et al.*, 1977). [b] Ncap to N3 refers to the first four positions of the helix (see helix nomenclature in the text). [c] Backbone dihedral angles at the Ncap position are listed in degrees.

Table III: Pairwise Occurrence of Potential Capping Residues at Ncap-N3 Positions in α-Helices[a]

	N3						
Ncap	Glu (*3.4*)	Asp (*1.0*)	Gln (*1.4*)	His (*1.1*)	Thr (*0.7*)	Asn (*0.6*)	Ser (*0.3*)
Ser (*2.5*)	8	3	2		1	1	1
Asp (*2.2*)	1	1		1	2		
Thr (*1.9*)	7	2	2		1	1	
Asn (*1.8*)	2		1	1	1	2	
His (*1.0*)	1			1			1
Gln (*0.5*)							1
Glu (*0.4*)	1						

[a] Individual table cells indicate the number of pairs formed by each type of residue at the first and fourth positions of the 161 α-helices in the set of 42 proteins. Blank table entries correspond to pairs with zero occurrence. Italicized numbers indicate the individual preference of each residue for Ncap or N3 positions of the helices.

beyond the mean for capping boxes.

Supplementing the 15 capping boxes found in our survey, 17 additional "near boxes" at Ncap-N3 positions were identified, having side chains capable of reciprocal capping interactions and suitable backbone dihedral angles but lacking one of the complementary H-bonds (Table II). The Ncap dihedral angles of the observed capping boxes and the near boxes are indistinguishable (Table II), and modeling suggests that the missing hydrogen bond can be easily achieved.

RESIDUE PREFERENCES IN CAPPING BOXES

It is of interest to compare the observed and expected frequencies of capping boxes by calculating normalized pairwise residue preferences at Ncap and N3 positions. Residues shown to be stereochemically capable of helix capping (Presta & Rose, 1988) are known to occur preferentially at the Ncap position of helices (Kendrew *et al.*, 1961; Baker & Hubbard, 1984; Argos & Palau, 1982; Richardson & Richardson, 1988; Rose & Wolfenden, 1993). Also, a marked preference for Glu is apparent at the N3 position (Baker & Hubbard, 1984; Richardson & Richardson, 1988; Rose & Wolfenden, 1993).

Table III shows the pairwise frequencies of capping residues at Ncap and N3 positions, along with the individual residue preferences for each position. Except for Asp at Ncap, individual frequencies parallel pairwise frequencies. The individual residue preferences shown in Table III agree with earlier studies, with the notable exception of Asn (Richardson & Richardson, 1988; Rose & Wolfenden, 1993). This exception can be attributed in part to differences in the definition of helix boundaries, but the results also reflect differences in the composition of the chosen proteins, illustrating the difficulty of selecting a truly random set of structures from available high-resolution crystallographic data (Boberg *et al.*, 1992).

The normalized pairwise frequency of occurrence, *f*, of two residues—*e.g.*, Ser and Glu—situated at two arbitrary positions—*e.g.*, Ncap and N3—is given by

$$f = \frac{(\text{no. of Ser-Glu pairs at Ncap-N3})/(\text{no. of all residue pairs at Ncap-N3})}{(\text{no. of Ser-Glu pairs in data set})/(\text{no. of all residue pairs in data set})}$$

A normalized frequency of unity means that the occurrence of the pair Ser-Glu at position Ncap-N3 is the same as its frequency of occurrence at large. An *f*-value of 13, for example, would then mean that the pair occurs at this position 13 times as often as would be expected from the empirically determined percentage composition of such pairs in the data base. Though described here in the context of a specific pair at a particular location, the procedure is general.

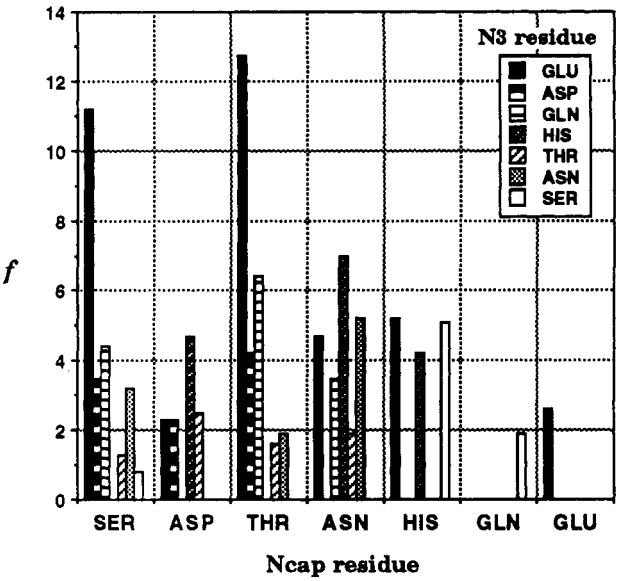


FIGURE 3: Histogram of normalized pairwise residue preferences for positions Ncap and N3 in capping boxes. The normalized frequency of occurrence, *f*, is calculated using the equation given in the text. The seven possible box-forming residues at Ncap are listed on the abscissa. For each, seven histogram bars are shown, corresponding to possible residues at N3. The identity of the N3 residue is indicated by the legend. The respective *f*-value of each Ncap-N3 pair is read on the ordinate. It can be seen that Thr-Glu and Ser-Glu pairs occur at Ncap-N3 positions with more than an order of magnitude higher frequency than expected by chance.

Figure 3 displays the normalized values of the pairwise residue preferences at Ncap and N3 positions listed in Table III. The data show that, in general, pairs capable of reciprocal capping interactions are found preferentially at Ncap-N3 positions. This observation is consistent with the clustering of hydrogen-bonding residues at helix termini noted in earlier work (Argos & Palau, 1982; Presta & Rose, 1988; Richardson & Richardson, 1988) and supports the inference that such clusters serve to define helix boundaries. The fact that the most abundant pair (Ser-Glu) also has the highest normalized frequency suggests that the individual positional preferences observed for Ser and Glu are reflections of their conjoint role in forming capping boxes.

THE CAPPING BOX IN PEPTIDES

Recently, the helix-forming tendencies of amino acid residues in short peptides in water have been studied by a number of groups, both in solution [Sueki *et al.*, 1984; Lyu *et al.*, 1990; Merutka *et al.*, 1990; O'Neil & DeGrado, 1990; Padmanabhan *et al.*, 1990; see review by Scholtz and Baldwin (1992)] and in calculations (Yun *et al.*, 1991a,b; Creamer & Rose, 1992). Several experimental studies have focused on capping residues (Bruch *et al.*, 1991; Lecomte & Moore, 1991; Lyu *et al.*, 1992, 1993; Forood *et al.*, 1993).

A 20-residue peptide studied by Kallenbach and co-workers (Lyu *et al.*, 1993) is of particular interest. The peptide sequence is Tyr-Met-Ser-Glu-Asp-Glu-..., and using 2D ^{1}H NMR, Lyu *et al.* demonstrated a helix boundary at residue 3, consistent with design intentions. Further, the existence of an H-bond between the side-chain O_γ of Ser(3) and the backbone >NH of Glu(6) was shown using NMR. The favorable sequence -Ser-Xaa-Xaa-Glu- together with the presence of the Ser(3) O_γ···HN< Glu(6) H-bond led us to predict the occurrence of a capping box. This prediction has since been confirmed by the presence of NOEs that indicate a complementary Glu(6) O_ϵ···HN< Ser(3) H-bond, as shown by Zhou *et al.* (1993).

Several in a series of 12-residue peptides studied by Nambiar and co-workers (1993) are also suggestive of capping boxes. The sequence of the parent peptide is Asp-Pro-Ala-Glu-..., and five single-residue substitutions at the N-terminus (including Ser and Asn) resulted in substantial differences in helicity, as assessed by circular dichroism. While the precise location of helix boundaries is unknown in these peptides, the marked differences in helicity upon substitution at residue 1 implicate the N-terminus. The presence of a free α-amino group may lead to additional helix-favoring interactions (and multiple states) for some peptides.

These studies suggest that a capping box can function as an effective N-terminal stop signal in peptide helices. Corresponding C-terminal stop signals must also exist (Kim & Baldwin, 1984). Once understood, such signals can be exploited deliberately in peptide design.

DISCUSSION

The capping box appears to function as a helix stop signal, in both proteins (Table II) and peptides (Zhou *et al.*, 1993). In proteins, this structural motif occurs exclusively in helices, where it is unique to Ncap-N3 residue pairs. In peptides, the capping box's reciprocal side-chain–backbone H-bonds lash Ncap to N3 and inhibit N-terminal fraying. Kim and Baldwin (1984) deduced that an embedded C-terminal stop signal must be present in the 20-residue S-peptide of ribonuclease A, since the helix terminates at residue 12. We suspect that such stop signals are a general phenomenon that may utilize a small number of specialized motifs, some of which are already documented in the literature (Schellman, 1980; Sibanda & Thornton, 1985; Milner-White, 1988; Bork & Preissner, 1991; Preissner & Bork, 1991).

In their study of helix capping, Serrano *et al.* (1992) measured the effect of mutating the Ncap residue in the two helices of barnase (a bacterial ribonuclease). Each helix has a threonine residue at the Ncap position that forms a side chain to backbone hydrogen bond with the >NH of the N3 residue. Both threonine and serine at the Ncap position appear to stabilize the helix by about 2 kcal/mol relative to alanine, with the Ncap-N3 side-chain–backbone hydrogen bond a major contributor to the stabilization. The second helix is further stabilized, by about 1 kcal/mol relative to alanine, by an additional hydrogen bond that completes the capping box motif (*i.e.*, the side chain of a glutamate residue at N3 to the >NH of the threonine at Ncap). Capping interactions are clearly not the only source of helix stability, however, since helix 1 appears to be at least as stable as helix 2.

A caution raised by the preceding findings, and underscored by experimental work from many laboratories, is that statistical studies of residue preferences at capping positions are averaged over the ensemble of all stop signals, whereas protein mutational studies probe the energetics of one particular type of stop signal, *viz.*, the one utilized at the site under consideration. For this reason, the two types of results are not precisely comparable.

ACKNOWLEDGMENT

We thank Neville Kallenbach, Rajeev Aurora, Trevor Creamer, and Jeffrey Seale for much useful discussion, Nuria Assa-Munt for her critical reading of the manuscript, and Alan Fersht and K. Henrick for barnase coordinates. We are especially grateful to Leonard Presta, who first called our attention to the frequent occurrence of complementary hydrogen bonds between Ser-Glu pairs at helix capping positions. Computer-generated drawings were made using Insight II (Biosym Technologies).

REFERENCES

Argos, P., & Palau, J. (1982) *Int. J. Pept. Protein Res. 19*, 380–393.
Artymiuk, P. J., & Blake, C. C. F. (1981) *J. Mol. Biol. 152*, 737–762.
Baker, E. N., & Hubbard, R. E. (1984) *Prog. Biophys. Mol. Biol. 44*, 97–179.
Bernstein, F. C., Koetzle, T. G., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rogers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535–542.
Boberg, J., Salakoski, T., & Vihinen, M. (1992) *Proteins 14*, 265–276.
Bork, P., & Preissner, R. (1991) *Biochem. Biophys. Res. Commun. 180*, 666–672.
Bruch, M. D., Dhingra, M. M., & Gierasch, L. M. (1991) *Proteins 10*, 130–139.
Creamer, T. P., & Rose, G. D. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 5937–5941.
Fairman, R., Shoemaker, K. R., York, E. J., Stewart, J. M., & Baldwin, R. L. (1989) *Proteins 5*, 1–7.
Forood, B., Feliciano, E. J., & Nambiar, K. P. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 838–842.
Heinz, D., Baase, W. A., Dahlquist, F. W., & Matthews, B. W. (1993) *Nature 361*, 561–564.
Kendrew, J. C., Watson, H. C., Strandberg, B. E., Dickerson, R. E., Phillips, D. C., & Shore, V. C. (1961) *Nature 190*, 666–670.
Kim, P. S., & Baldwin, R. L. (1984) *Nature 307*, 329–334.
Lecomte, J. T. J., & Moore, C. D. (1991) *J. Am. Chem. Soc. 113*, 9663–9665.
Lyu, P. C., Liff, M. I., Marky, L. A., & Kallenbach, N. R. (1990) *Science 250*, 669–673.
Lyu, P. C., Zhou, H. X., Jelveh, N., Wemmer, D. E., & Kallenbach, N. R. (1992) *J. Am. Chem. Soc. 114*, 6560–6562.
Lyu, P., Wemmer, D. E., Zhou, H. X., Pinker, R. J., & Kallenbach, N. R. (1993) *Biochemistry 32*, 421–425.
Matthews, B. W. (1987) *Biochemistry 26*, 6885–6888.
Merutka, G., Lipton, W., Shalongo, W., Park, S.-H., & Stellwagen, E. (1990) *Biochemistry 29*, 7511–7515.
Milner-White, E. J. (1988) *J. Mol. Biol. 199*, 503–511.
O'Neil, K. T., & DeGrado, W. F. (1990) *Science 250*, 646–651.
Padmanabhan, S., Marqusee, S., Ridgeway, T., Laue, T. M., & Baldwin, R. L. (1990) *Nature 344*, 268–270.
Pauling, L., Corey, R. B., & Branson, H. R. (1951) *Proc. Natl. Acad. Sci. U.S.A. 37*, 205–211.
Perutz, M. F. (1951) *Nature 167*, 1053–1054.
Preissner, R., & Bork, P. (1991) *Biochem. Biophys. Res. Commun. 180*, 660–665.
Presta, L. G., & Rose, G. D. (1988) *Science 240*, 1632–1641.
Richardson, J. S., & Richardson, D. C. (1988) *Science 240*, 1648–1652.
Rose, G. D., & Wolfenden, R. (1993) *Annu. Rev. Biophys. Biomol. Struct. 22*, 381–415.
Schellman, C. (1980) in *Protein Folding* (Jaenicke, R., Ed.) pp 53–61, Elsevier/North-Holland, New York.
Scholtz, J. M., & Baldwin, R. L. (1992) *Annu. Rev. Biophys. Biomol. Struct. 21*, 95–118.
Serrano, L., & Fersht, A. R. (1989) *Nature 342*, 296–299.
Serrano, L., Sancho, J., Hirshberg, M., & Fersht, A. R. (1992) *J. Mol. Biol. 227*, 544–559.
Sibanda, B. L., & Thornton, J. M. (1985) *Nature 316*, 170–176.
Stickle, D. F., Presta, L. G., Dill, K. A., & Rose, G. D. (1992) *J. Mol. Biol. 226*, 1143–1159.
Sueki, M., Lee, S., Powers, S. P., Denton, J. B., Konishi, Y., & Scheraga, H. A. (1984) *Macromolecules 17*, 148–155.
Yun, R. H., & Hermans, J. (1991) *Protein Eng. 4*, 761–766.
Yun, R. H., Anderson, A., & Hermans, J. (1991) *Proteins 10*, 219–228.
Zhou, H. X., Lyu, P., Wemmer, D. E., & Kallenbach, N. R. (1993) *Biochemistry* (submitted for publication).