Effect of the N1 residue on the stability of the α -helix for all 20 amino acids

DUNCAN A.E. COCHRAN, SIMON PENEL, AND ANDREW J. DOIG

Department of Biomolecular Sciences, UMIST, Manchester M60 1QD, UK

(RECEIVED July 27, 2000; FINAL REVISION November 20, 2000; ACCEPTED November 28, 2000)

Abstract

N1 is the first residue in an α -helix. We have measured the contribution of all 20 amino acids to the stability of a small helical peptide $\text{CH}_3\text{CO-XAAAAQAAAAQAAGY-NH}_2$ at the N1 position. By substituting every residue into the N1 position, we were able to investigate the stabilizing role of each amino acid in an isolated context. The helix content of each of the 20 peptides was measured by circular dichroism (CD) spectroscopy. The data were analyzed by our modified Lifson-Roig helix-coil theory, which includes the nI parameter, to find free energies for placing a residue into the N1 position. The rank order for free energies is Asp^- , $\text{Ala} > \text{Glu}^- > \text{Glu}^0 > \text{Trp}$, Leu , $\text{Ser} > \text{Asp}^0$, Thr , Gln , Met , $\text{Ile} > \text{Val}$, $\text{Pro} > \text{Lys}^+$, Arg , $\text{His}^0 > \text{Cys}$, $\text{Gly} > \text{Phe} > \text{Asn}$, Tyr , His^+ . N1 preferences are clearly distinct from preferences for the preceding N-cap and α -helix interior. pK_a values were measured for Asp, Glu, and His, and protonation-free energies were calculated for Asp and Glu. The dissociation of the Asp proton is less favorable than that of Glu, and this reflects its involvement in a stronger stabilizing interaction at the N terminus. Proline is not energetically favored at the α -helix N terminus despite having a high propensity for this position in crystal structures. The data presented are of value both in rationalizing mutations at N1 α -helix sites in proteins and in predicting the helix contents of peptides.

Keywords: α -Helix; N1 position; circular dichroism; protein folding; protein stability; helix propensities; helix-coil theory; macrodipole

The α -helix is the most commonly occurring secondary structural element in proteins. Positional preferences for amino acids in the interior positions of the α -helix have been measured experimentally in peptides and proteins (for review, see Pace and Scholtz 1998). Amino acid preferences for the helix differ greatly between terminal and interior positions (Argos and Palau 1982; Richardson and Richardson 1988; Kumar and Bansal 1998; Penel et al. 1999). The N terminus of an α -helix has three amide hydrogens that lack the $i, i+4$ backbone hydrogen bonds that characterize the α -helix; similarly, the C terminus has three 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 α -helix termini. The residue immediately preceding the N-terminal side of the α -helix has been dubbed the N-cap, while the residue immediately following the C-terminal side of the α -helix is the C-cap. The capping residues have nonhelical ϕ , ψ dihedral angles, although they form helical $i, i+4$ hydrogen bonds. Between the capping residues lie the residues with α -helical ϕ , ψ dihedral angles, designated as N1, N2, N3 . . . C3, C2, and C1. While the N terminus of the α -helix is structurally well characterized (Penel et al. 1999), it is less well understood in thermodynamic terms, and this represents one barrier to successfully predicting secondary structure from primary sequence. The N-cap, N1, N2, and N3 residues can potentially interact with the α -helix macrodipole, free backbone amides, and local side chains in different ways and, thus, contribute differently to α -helix stability.

While the Serrano group and our group have measured the free energy changes that arise when certain residues are

Reprint requests to: Dr. Andrew J. Doig, Department of Biomolecular Sciences, UMIST, Manchester M60 1QD, UK; e-mail: Andrew.Doig@umist.ac.uk; fax: 44-161-236-0409.

Article and publication are at www.proteinscience.org/cgi/doi/10.1110/ps.31001.

substituted at N1 positions (Petukhov et al. 1998, 1999; Sun et al. 2000), many amino acids have not been investigated, notably the charged residues. In this article, we determine the intrinsic preferences for N1 positions in our α -helical model peptide for all 20 naturally occurring amino acids. We use the methodology of previous work, where the N1 preferences of 11 amino acids were determined using new helix-coil theory (Sun et al. 2000) and previously published peptide data (Chakrabartty et al. 1993a). Here, we study 20 peptides of improved design that differ only in their N1 amino acid.

Results and discussion

Helix-coil theory

α -Helices in aqueous solution adopt a large number of structures, with fully helical, fully coil, and partly helical conformations all populated. For a complete understanding of helix formation and stability, all the factors contributing to this equilibrium need to be assessed thoroughly. These include the helix-forming tendencies of constituent amino acids, capping preferences at the C and N termini and preferences for the N1, N2, and N3 positions. Our helix-coil model N1N2N3 (Sun et al. 2000) takes all these factors into account. Briefly, every residue is considered to be in either a helix or coil conformation, giving a total of 2^N conformations for N residues. In each conformation, the N-cap residue has a weight n , followed by residue weights of $n1$, $n2*w$ and $n3*w$. The C1 residue has a weight of $c1$, the C-cap c , and all helix interior residues from N4 to C2 have weights of w . Coil residues before and after the caps have weights of 1. Side-chain interactions between residues spaced $i, i+3$ and $i, i+4$ are also of considerable importance to helix stability, though irrelevant here. The statistical weight of each conformation is the product of the weights of each residue, taking into account the residue conformations. The sum of the statistical weights for every conformation is the partition function, and the population of each conformation is the statistical weight of the conformation divided by the partition function. All properties of the helix/coil equilibrium can be calculated from the populations and the partition function. Increasing the $n1$ value of a residue X will thus increase the population of all conformations where X is at the N1 position of a helix.

Design of the N1 peptide sequence

A residue at the N-terminal position in an acetylated peptide adopts a helical conformation when it has helical dihedral angles, and the acetyl CO group is hydrogen bonded to the peptide NH of the $i+4$ residue. This residue is bounded by an amide group on both sides, and the N-terminal acetyl group can act as an N-cap (Doig et al. 1994). N1 preferences can therefore be determined by substituting the first residue

in an acetylated peptide and measuring the change in helix content.

Our peptide ($\text{CH}_3\text{CO-XA}_4\text{QA}_4\text{QA}_2\text{GY-NH}_2$) is intrinsically helical because of its high Ala content. X is the variable residue, and it affects helix content through both its N-cap (if X acts as the N-cap) and N1 (if acetyl is the N-cap) preferences. N-cap preferences (Serrano and Fersht 1989; Lecomte and Moore 1991; Bell et al. 1992; Serrano et al. 1992; Chakrabartty et al. 1993a; Forood et al. 1993; Lyu et al. 1993; Yumoto et al. 1993; Doig et al. 1994; Doig and Baldwin 1995) and preferences for the peptide interior (O'Neil and DeGrado 1990; Horovitz et al. 1992; Park et al. 1993; Blaber et al. 1994; Chakrabartty et al. 1994; Rohl et al. 1996; Myers et al. 1997; Yang et al. 1997; Pace and Scholtz 1998) have already been measured, leaving the N1 preference ($n1$ in helix-coil theory) values as the only unknown parameter. The N1 preference can thus be found by finding the value of $n1$ that gives predicted helix content in agreement with experiment. No side-chain interactions are present in this sequence. The energy scales of Rohl (Rohl et al. 1996) are used in the calculations, as they are complete and derived in similar peptide systems for the interior and N-cap positions. The Tyr residue is present to give the peptide a UV absorption for determination of concentration. The penultimate Gly ensures that the α -helix almost always terminates at, or before, this position so that problems arising from a helical Tyr affecting the CD signal are minimized (Chakrabartty et al. 1993b). A peptide of this sequence has a helix content of $\sim 40\%$ so that the CD signal is sensitive to a change in free energy of α -helix formation. The design of the AQ peptide ensures that the contribution of the N1 residue to stability is not influenced by a charge on a free N terminus or by neighboring Gln residues included for solubility, as the Q residues are spaced $i, i+5$.

Helix contents of N1 peptides

The helix content of each peptide was measured by CD at 222 nm at a concentration of 10–30 μM in 10 mM NaCl, 5 mM sodium phosphate (pH 7.01) at 273 K. Measurements are made at 273 K because the helix-coil equilibrium favors the helical conformation at lower temperature and CD signals are consequently larger. CD measurements for these types of simple peptides are known to be proportional to concentration in this range (Stapley and Doig 1997) indicating that aggregation does not occur. The CD signal at 222 nm was converted to $[\theta]_{222}$ (see Materials and Methods), and the percentage helicity calculated with $\pm 3\%$ experimental error (Table 1).

We observe that Asp^- in the N1 position of the α -helix promotes an increase in helix content from 42% to 45% relative to Ala, and that Asp^- stabilizes the α -helix more than Asp^0 (34%). All the other peptides are less helical than N1(Ala). Interestingly, the helix content of N1(Glu $^-$) is the

Table 1. Rank order of *nI* values and free energies for the helix-coil equilibrium for peptides with varying N1 amino acid

N1	$[\theta]_{222}^a$	% helix ^b	<i>nI</i> value ^c ($\pm 3\%$ helicity)	ΔG for helix-coil transition ^d	$\Delta\Delta G^e$	$\Delta\Delta G^e$ Petukhov et al. 1998, 1999	$\Delta\Delta G$ Sun et al. 2000
D ⁻	-15100	45	0.28 (0.16–0.49)	0.7 (1.0–0.4)	0	—	—
A	-13900	42	0.27 (0.34–0.20)	0.7 (0.6–0.9)	0	0	0
E ⁻	-13500	41	0.22 (0.15–0.31)	0.8 (1.0–0.6)	0.1	—	—
E ⁰	-12000	36	0.15 (0.11–0.21)	1.0 (1.2–0.9)	0.3	—	—
W	-12800	38	0.13 (0.08–0.21)	1.1 (1.4–0.9)	0.4	—	—
L	-12100	36	0.13 (0.09–0.19)	1.1 (1.3–0.9)	0.4	0.2	0.5
S	-12800	38	0.13 (0.08–0.20)	1.1 (1.4–0.9)	0.4	0.5	0.6
D ⁰	-11200	34	0.12 (0.09–0.16)	1.2 (1.3–1.0)	0.5	—	—
T	-11700	35	0.11 (0.07–0.16)	1.2 (1.4–1.0)	0.5	0.5	0.5
Q	-10200	31	0.11 (0.08–0.15)	1.2 (1.4–1.0)	0.5	0.3	0.4
M	-11000	33	0.11 (0.07–0.15)	1.2 (1.4–1.0)	0.5	0.06	0.5
I	-11200	34	0.11 (0.07–0.15)	1.2 (1.4–1.0)	0.5	0.4	0.5
V	-10600	32	0.10 (0.07–0.13)	1.3 (1.5–1.1)	0.6	—	0.4
P	-10471	31	0.09 (0.06–0.12)	1.3 (1.6–1.1)	0.6	—	0.5
K ⁺	-9900	30	0.08 (0.06–0.11)	1.4 (1.6–1.2)	0.7	—	—
R	-10000	30	0.08 (0.05–0.11)	1.4 (1.6–1.2)	0.7	—	—
H ⁰	-10800	32	0.08 (0.05–0.12)	1.4 (1.7–1.2)	0.7	—	—
C	-8400	25	0.04 (0.02–0.07)	1.7 (2.0–1.5)	1.0	—	—
G	-10400	31	0.04 (0.01–0.08)	1.7 (2.4–1.4)	1.0	0.7	0.6
F	-8500	26	0.02 (0.0005–0.04)	2.1 (4.2–1.7)	1.4	—	—
N	-10800	32	0	— ^g	— ^g	0.6	1.4
Y ^f	-8400	25	0	— ^g	— ^g	—	—
H ⁺ f	-5500	17	0	—	—	—	—

^a The peptide sequence is CH₃CO-XA₄QA₄QA₂GY-NH₂. $[\theta]_{222}$ units are deg · cm² · dmol⁻¹.

^b Calculated as $(-[\theta]_{222} \times 100)/(-40000 \times (1 - 2.5/N))$.

^c From fitting program.

^d $-RT \ln(nI)$, in kcal mol⁻¹.

^e $\Delta G_{Xaa} - \Delta G_{Ala}$, in kcal mol⁻¹, for helix-coil transition relative to Ala.

^f Not possible to fit a positive value for *nI*.

^g Too high to measure.

same as N1(Ala), where the negative charge might be expected to stabilize the α -helix N terminus. Positioning of the basic residues, Lys, Arg, and His⁺ (pH 2–4) at the N terminus results in a substantial reduction in helix content to 30%, 30%, and 17%, respectively. Two other residues are more destabilizing to the α -helix than Arg, namely Phe (26%) and Cys (25%). The $[\theta]_{222}$ value measured for Trp and Tyr at the N1 position is unreliable, as their UV absorption is known to perturb the CD signal (Chakrabartty et al. 1993b).

It is tempting to draw conclusions about the observed stabilizing effects of residues at N1 on the basis of helix content relative to Ala. However, it is the underlying equilibrium constant for the helix-coil transition at N1, the *nI* parameter, that defines the relative free energies for stabilization according to modern helix-coil theory. The stabilizing effects of the N-terminal residues are, therefore, described by both the *n-cap* and *nI* energies, and helix-coil theory is able to separate the two parameters.

Determination of *nI* values and ΔG from helix contents

In all previous work, *w*, the equilibrium constant for adding a given residue to an α -helix, has been determined assuming

that the values of *nI*, *n2*, and *n3* are no different from any other internal position. However, we have now examined the contributions of the N1 residues to stability in isolation from other interactions by considering the magnitude of *nI* values specifically. *nI* is formally defined as the equilibrium constant for the helix-coil transition at the N1 position (Sun et al. 2000). The percentage helicities measured for the N1 peptides have been converted to *nI* for the helix-coil transition using a modified Lifson-Roig theory (see Materials and Methods). Initially, *nI* for the four sequences with an X residue matching residues of the generic peptide (Ala, Gly, Gln, and Tyr) were solved simultaneously by fitting the equilibrium equation to the observed helicities using the N1N2N3 program (see Materials and Methods). Fitting a value for Tyr was problematic, as convergence to a positive value for the *nI* parameter was not possible. It was, therefore, set to zero. This did not affect other *nI* values significantly, as Tyr is found only at their C termini and, therefore, is unlikely to be in position N1 of an α -helix in these peptides. Subsequently, these four *nI* values were fixed for the calculation of *nI* for the remaining peptides.

The residue *nI* values are presented in rank order in Table 1. The order is different when compared with the order of

helicity values. This indicates that the observed helicity depends on the relative magnitudes of the nI parameter compared with the $n\text{-cap}$ parameter and confirms that analysis using helix-coil theory is essential for understanding the equilibrium.

The $n\text{-cap}$ and nI parameters represent the probabilities that a particular residue will adopt either an N-cap conformation or an N1 conformation, respectively. When the $n\text{-cap}$ value for an amino acid is large and of similar magnitude to the acetyl group ($n\text{-cap} = 5.9$; Doig et al. 1994), it will spend much of its time in a nonhelical N-cap conformation. Indeed, the large $n\text{-cap}$ values available for Cys (5.4), Asn (6.8), and Tyr (4.9) tend to shift the equilibrium in favor of the N-cap state for the residue X . This competition between parameters results in low estimates for the nI values.

All helix-coil parameters are interdependent, so the nI values could be perturbed by errors in n values, in particular. This problem is most important when X in the sequence we used, $\text{CH}_3\text{CO-XAAAAQAAAAQAAGY-NH}_2$ has a high N-cap preference, such as Asp. In this case, the helices will often nucleate with X as the N-cap. When X has a low N-cap preference, the acetyl will usually be the N-cap, X will be at N1, and an error in $n(X)$ will be small, as the conformation when X is the N-cap will be rarely populated. To investigate this problem, we recalculated nI for Asp after recalculating the error range in n (found by varying the helix contents of the N-cap peptide [Doig and Baldwin 1995] by $\pm 3\%$ and refitting).

$n(\text{Asp})$ is in the range 4–7, and these values put $nI(\text{Asp})$ in the range 0.330 to 0.249 or $\pm 0.07\text{kcal/mol}$. Other amino acids are less sensitive to errors in n .

$nI(\text{Pro})$, for example, varies from only 0.086 to 0.111 ($\pm 0.07\text{kcal/mol}$) if $n(\text{Pro})$ is varied from 1.35 to 0.1 (much larger than the real error range). Our results are therefore insensitive to errors in n -values and were indeed designed so that this is the case.

The N-cap energy of Cys⁰ has not been measured previously. To estimate $nI(\text{Cys})$, we set its $n\text{-cap}$ parameter to that used for Ala; that is, 1. This yielded $nI(\text{Cys}) = 0.04$, which is comparable with that measured for a similar residue, Ser, $nI = 0.13$, given the experimental uncertainty. However, calculations with the apparently more appropriate $n\text{-cap}(\text{Ser})$ value did not allow determination of $nI(\text{Cys})$ with a positive value. It is therefore likely that the real $n\text{-cap}$ parameter for Cys in an uncharged environment is somewhere between that of Ala and Ser.

From the nI values, the free energy change for the transfer of a residue from coil conformation to the N1 helical conformation, ΔG , is given by $-RT \ln(nI)$. The free energy change relative to Ala, $\Delta\Delta G_{\Delta G\text{Xaa}-\Delta G\text{Ala}}$, was also calculated (Table 1). The data show that the N1 energies for the majority of the residues lie within 1 kcal mol^{-1} of that measured for Ala. The hypothesis that stabilization of the α -he-

lix N terminus results from the presence of negatively charged amino acids is not universally supported by the $\Delta\Delta G$ values measured in the AQ peptides. However, the neutral species of Asp and Glu are less stabilizing relative to their anionic forms.

Comparison with other reported nI values

Recently, free energies for N1 have been reported for a series of AR peptides (Petukhov et al. 1998, 1999), and we have recently published a reanalysis of previously reported AK peptide data (Chakrabartty et al. 1993a) using the N1N2N3 program (Sun et al. 2000). The previously reported free energies for several N1 residues are very similar (Table 1), with Leu, Ser, Thr, Gln, Met, Ile, Val, Pro, and Gly having $\Delta\Delta G \sim 0.5 \pm 0.1\text{ kcal mol}^{-1}$. The one residue that does not follow this trend is Asn, which is most destabilizing at N1 for the short AK peptides.

The peptides studied here are 15 mers, with approximately double the helical content of those studied previously. These data, therefore, give more accurate N1 energies. Our new data indicate that the reported 1.4 kcal mol^{-1} (Sun et al. 2000) is too low an estimate of nI for Asn because we were unable to determine a positive value. This may be because the previously reported N-cap value (Doig and Baldwin 1995) used in the calculation is too high.

The $\Delta\Delta G$ values presented in this study for Leu, Ser, Thr, Gln, Val, Pro, and Ile are similar (within error) compared to those described previously (Table 1). Proline destabilizes the α -helix by 0.6 kcal mol^{-1} relative to Ala, which is moderately high relative to the other 18 amino acid residues. Structurally, Pro is apparently well suited to the N1 position, as its ϕ dihedral angle is constrained to a helical conformation and it has no free peptide NH group. Interestingly, statistical preferences for Pro at N1 are the highest for any amino acid in both the global (compared to all protein positions) and local (compared to other helical positions) context, 2.6 and 5.8, respectively (Penel et al. 1999). This demonstrates that a high propensity does not necessarily indicate a large stabilization.

While propensities are measured relative to all other positions in a protein, N1 energies are measured relative to the coil state. The high propensity for Pro at N1 is thus likely to be because of the fact that Pro is highly destabilizing at most other locations within a protein, such as within an α -helix or β -sheet, rather than because of any intrinsic energetic preference for N1. The average global propensity for Pro in a type I turn is similarly high at 2.45 (Hutchinson and Thornton 1994). We suggest that this high propensity is, again, because Pro is tolerated at this location rather than because it stabilizes turns more than do other amino acids. The propensity measurements from statistics in protein structures cannot therefore substitute for free energy measurements. There are no other reported measurements of nI preferences

for Glu, Trp, His, Lys, Arg, Cys, Phe, or Tyr for comparison.

pH titrations

pH titrations were performed on those peptides with titratable groups to determine the pK_a values of the residue side chains in the N1 position (Fig. 1). The asymptotic values of θ_{222} for the different protonation states were used to fit helix-coil theory and, hence, find nI values. The ellipticity data from the titrations were fitted to a Henderson-Hasselbach equation for two different pK_a s, as all peptides contain Tyr at the C terminus. In general, we measured pK_a values within the broad range observed in proteins, that is, Asp/Glu, $pK_a = 2$ –5.5; and His, $pK_a = 5$ –8 (Fersht 1999). However, these ranges are too broad for comparisons with those pK_a s measured here.

A range of pK_a s have been reported for amino acid side chains measured in a host of different model compounds (Kyte 1995). Measurements of dissociation constants have been made when Asp or Glu are located centrally in a poly-Gly heptamer. The pK_a s measured in these compounds are lower than in other models, reflecting the unstructured conformation of the peptide chain and the interaction of polar groups with the side chains of the acid residues (Keim et al. 1974). Intrinsic pK_a s (measured in the absence of electrostatic effects) have been reported for Asp (4.1) and Glu (4.5) in N-acetyl acid amides (Nozaki and Tanford 1967) and for His (6.4) in Pro-His-glycinamide (Tanford 1962). Our pK_a s were measured at 273 K in 0.01 M NaCl, while those for the model compounds are reported at 298 K in water.

The pK_a for N1(Asp) (3.93 ± 0.03) is marginally lower than that seen in appropriate model compounds, and the pK_a of N1(Glu) is unperturbed (4.56 ± 0.09). The pK_a for N1(His) of 6.79 ± 0.03 is also normal given that the imidazole group has two microscopic dissociation constants, 6.5 and 7.1 (Kyte 1995).

In the AQ peptides, there can be no side chain–side chain interactions, so stabilization must come from other effects. The acetyl group is very probably hydrogen bonded to the NH group of the $i+4$ residue in a capping interaction and is not free to interact with the side chains at N1. In crystal structures of proteins, it is generally observed that the side chains of Asp and His are sterically constrained from forming favorable hydrogen bonds with the main chain (Penel et al. 1999) and, therefore, potentially interact solely with the macrodipole. However, an M2S1 hydrogen bond between the Glu side-chain carbonyl and the main chain of the N2 residue is possible, though energetically costly and unfavorable, as it is not in the preferred rotamer conformation (Penel et al. 1999). These facts support the hypothesis that the effect of N1 side-chain titration on α -helix stability involves an interaction with the α -helix macrodipole.

The free energy for protonation of Asp at N1 in the AQ peptide is $0.5 \text{ kcal mol}^{-1}$ (range of 0.3–0.6), and the anionic form is clearly more stable than the neutral species (Fig. 1a). Doig and Baldwin (1995) measured the free energy for protonation of Asp at the N-cap and reported a value $0.7 \text{ kcal mol}^{-1}$. The difference between these two values reflects the stronger interaction of the N-cap(Asp) with the free NH groups at the N terminus, as an N-cap Asp usually forms strong hydrogen bonds to the N2 and N3 NH groups (Doig et al. 1997). The protonated state of Glu was also found to be marginally disfavored at N1 by $0.2 \text{ kcal mol}^{-1}$ (range of 0.17–0.21 kcal mol^{-1}) and destabilization of the α -helix at low pH is clearly observed in Figure 1b. Hence, negatively charged side chains stabilize the α -helix compared with the neutral species by virtue of their larger nI parameter.

Asp[−] at N1 may be more stabilizing to the α -helix compared with Glu[−] because of the differences in distance between the deprotonated side chain and the macrodipole field and not hydrogen bonding per se. Similarly His⁺ may be more destabilizing than Lys⁺ and Arg⁺, as its charge is closer to the helix terminus.

The free energy for protonation of the His side chain is not determined, as an nI value for His⁺ could not be estimated. However, titration of the N1(His) peptide to pH 12 demonstrates that the α -helix is progressively stabilized from pH 5–8 and destabilized with ionization of the Tyr–OH between pH 8 and 12 (Fig. 1c). From this, we were able to measure the pK_a of the C-terminal Tyr to be 10.60 ± 0.13 . The pK_a of Tyr is consistently measured to be 9.8 in a range of model compounds (Kyte 1995). The increased pK_a of Tyr may result from an unfavorable electrostatic interaction with the C-terminal dipole or partial negative charges on the terminal CO groups.

Other studies on alanine-based helical peptides have not reported the pK_a of the C-terminal Tyr. We believe that α -helix destabilization between pH 9 and 12 was present in a previous study (Doig and Baldwin 1995), but the Tyr titration was obscured by the effect of three Lys residues included for solubility. The pK_a s of peptides containing Arg, Lys, Cys, and Tyr side chains at N1 were not evaluated, as their measurements are complicated by overlapping titration of Tyr at the C terminus.

Direct observation of the α -helix macrodipole is not possible, and its existence is inferred by the observation that charged side chains and other ionic species are commonly found near α -helix termini in crystal structures. The pK_a s of the side chains at the N1 position of the AQ α -helix are not significantly perturbed by a macrodipole interaction. Stabilization of the α -helix on ionization of the acidic side chains is indisputable, and protonation of His is so unfavorable that we cannot measure this nI . The helicity of these peptides is very sensitive to interactions at the N1 position such that a small interaction of an anionic side chain with the dipole

Fig. 1. pH titrations of CD signal at 222 nm for (a) N1(Asp), (b) N1(Glu), and (c) N1(His) peptides. N1-Xaa: $\text{CH}_3\text{CO-XAAAAQAAAAQAAGY-NH}_2$.

leads to measurable changes in $[\theta]_{222}$ (but no detectable changes in pK_a). It is also possible that the coil form of the polypeptide is destabilized by the anionic side chain or stabilized by the protonated form in any number of possible interactions with the rest of the molecule.

Correlation of $n1$ energies with N1 propensities and w values

We recently analyzed the local and global propensities for every amino acid in first turn of the α -helix (Penel et al. 1999). The global propensity value is an indicator of the abundance of a particular amino acid in the N1 position relative to the whole molecule. The relationship between the observed frequency of a particular residue at N1 and the experimentally determined $n1$ values presented in this report has been investigated (Fig. 2). A correlation between the two parameters ($R = 0.49$) is clearly observed, with Pro as an outlier (see above). Thus, the energy cost for placing a residue at N1 in our model peptide is a reasonable reflection of the process in natural proteins. The fact that the correlation is not perfect shows the necessity of measuring free energies rather than solely determining propensities from crystal structures. Ala has a high N1 preference, while that for Gly is low, as in α -helix interior positions. This is presumably because the flexible backbone of Gly loses more backbone conformational entropy when it folds than for other amino acids.

Fig. 2. The relationship between $n1$ and N1 residue propensity. The N1 propensities (the observed relative abundance in crystal structures) are taken from Penel et al. (1999). Proline was not included in the regression analysis because it has an abnormally high propensity for N1 (see text).

The relationship between w and nI is shown in Figure 3. In helix-coil theory, the w value is the equilibrium constant for extending a helical conformation by one amino acid; that is, α -helix propagation. The observed correlation between w and nI implies that the energetics of producing helical dihedral angles is an intrinsic property of the residue to some extent. If a residue is easily incorporated into the extending α -helix, it is also easily accommodated in the N1 position. This is logical, as both involve fixing to the same dihedral angles. The acidic side chains are found together in the same region of the plot, as are the basic side chains, and this again demonstrates the preference for negatively charged groups at the N terminus of the α -helix.

Conclusion

The substitution of all 20 amino acids at the N1 position in an AQ peptide has yielded the range of intrinsic preferences for residues at this position in an α -helix. We suggest that this information can be used to rationally modify protein stability, particularly as helix N termini very rarely form tertiary interactions (Doig et al. 1997). The low nI value for Pro compared to its propensity for N1 is contrary to the popularly held view that Pro stabilizes α -helix N termini. The difference in stability between the charged and uncharged basic and acidic side chains clearly shows the effect of placing charged groups proximal to the peptide N terminus and C terminus. However, we cannot attribute the observed effect at the N terminus to hydrogen bonding, and instead, the effect must come from an interaction with the macrodipole. This interaction is small, however, as it does

not grossly perturb the pK_a s of the side chains, and within proteins it may easily be masked by additional dipolar interactions.

Materials and methods

Peptide synthesis

Peptides were synthesized by the solid-phase method. C-terminal amides were made using Rink Amide resin (CN Biosciences). N termini were acetylated with acetic anhydride/pyridine. Peptides were cleaved from the resin using 95% trifluoroacetic acid, 2.5% triisopropylsilane, 2.5% water. Peptides were purified by C_{18} reverse phase HPLC, and their molecular weights were checked by ES or MALDI mass spectrometry. Peptide purity was checked by analytical C_{18} HPLC. Peptide concentration was determined by measuring tyrosine UV absorbance of aliquots of stock solution dissolved in water or 6 M guanidine hydrochloride using $\epsilon_{275} = 1390 \text{ M}^{-1} \text{ cm}^{-1}$ or $\epsilon_{275} = 1450 \text{ M}^{-1} \text{ cm}^{-1}$, respectively (Brandts and Kaplan 1973). Concentrations of peptides containing Trp were determined in 6 M GuHCl (pH 7.0), 25°C using $\epsilon_{281}(\text{Trp}) = 5690 \text{ M}^{-1} \text{ cm}^{-1} + \epsilon_{281}(\text{Tyr}) = 1250 \text{ M}^{-1} \text{ cm}^{-1}$ (Edelhoch 1967).

Circular dichroism measurements

Equilibrium CD measurements were made at 222 nm with a Jasco J810 spectropolarimeter at 273 K. Ellipticity was measured in 10 mM NaCl, 5 mM sodium phosphate (pH 7.0) in a quartz cell with a 0.1 cm path length. The Cys peptide was measured in the presence of 1 mM dithiothreitol. CD data in millidegrees was converted to mean residue ellipticity using $[\theta]_{222} = \theta/(\text{molar concentration} \times 15 \text{ residues})$, in $\text{deg cm}^2 \text{ dmol}^{-1}$. Helix content was calculated as $[\theta]_{222(\text{observed})}/[\theta]_{222(\text{max})}$. $[\theta]_{222(\text{max})}$ is given by $-40,000(1-2.5/n)$, where n is the number of amino acids in the peptide (Chakrabarty et al. 1991).

pH titrations were performed in a 1.0-cm quartz cell at 273 K. The buffer used for pH titrations contained 10 mM NaCl, 1 mM sodium phosphate, 1 mM sodium borate, and 1 mM sodium citrate. The pH meter was calibrated for measurements at 273 K. The pH was adjusted during the titrations with aliquots of HCl or NaOH. Data was only used in the pH range where the equilibrium is reversible (thus ruling out deamidation or other irreversible effects as the cause of the pH changes). The ellipticity data from the titrations were fitted to a Henderson-Hasselbach equation for two different pK_a s

$$[\theta]_{222nm} = [\theta]_{222,\text{mid_pH}} * \left(1 - \frac{1}{1 + 10^{\text{pH}-\text{pKa1}}}\right) + [\theta]_{222,\text{low_pH}} * \left(\frac{1}{1 + 10^{\text{pH}-\text{pKa1}}}\right) + [\theta]_{222,(\text{mid_pH}-\text{high_pH})} * \left(1 - \frac{1}{1 + 10^{\text{pH}-\text{pKa2}}}\right).$$

Fig. 3. The relationship between w and nI in the α -helix. The nI values are calculated from helix-coil theory as described herein. The w values are taken from Rohl and Baldwin (1996).

where, pK_a1 and pK_a2 are the pK_a s measured for the acid–base equilibrium at low and high pH, respectively. $[\theta]_{222\text{ low_pH}}$, $[\theta]_{222\text{ mid_pH}}$, and $[\theta]_{222\text{ high_pH}}$ are the molar ellipticities measured at 222 nm at the titration end points at low, mid, and high pH. $[\theta]_{222\text{ mid_pH-high_pH}}$ is the change in molar ellipticity associated with pK_a2 . The free energies for protonation were calculated using $\Delta G = -RT \ln[nI(\text{charged})/nI(\text{uncharged})]$.

Determination of helix-coil parameters

nI values were determined using a statistical mechanics algorithm implemented in the program N1N2N3 (<http://www.bi.umist.ac.uk/users/mjfajdg/n1n2n3.htm>). N1N2N3 implements the modified Lifson-Roig helix-coil theory (Sun et al. 2000) to calculate the helix content of a given peptide sequence. The inputs to the program are the peptide sequences and corresponding experimental helix contents; a library of w , $n\text{-cap}$, $n1$, $n2$, $n3$, $c1$, and $c\text{-cap}$ values for each amino acid; an $n\text{-cap}$ value for acetyl and a $c\text{-cap}$ value for amide; and a list of parameters to be determined. All $n2$, $n3$, and c values were set to 1. Parameters were varied until they converged on essentially unchanging values.

Acknowledgments

The Michael Barber Mass Spectrometry Facility at UMIST is thanked for verifying peptide identity. This work was supported by the Biotechnology and Biological Sciences Research Council (Grant 36/B09794).

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Argos, P. and Palau, J. 1982. Amino acid distribution in protein secondary structures. *Int. J. Peptide Protein Res.* **19**: 380–393.
- Bell, J.A., Becktel, W.J., Sauer, U., Baase, W.A., and Matthews B.W. 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, X.J., Lindstrom, J.D., Pepiot, S.D., Baase, W.A., and Matthews, B.W. 1994. Determination of α -helix propensity within the context of a folded protein: Sites 44 and 131 in bacteriophage-T4 lysozyme. *J. Mol. Biol.* **235**: 600–624.
- Brandts, J.R. and Kaplan, K.J. 1973. Derivative spectroscopy applied to tyrosyl chromophores: Studies on ribonuclease, lima bean inhibitor, and pancreatic trypsin inhibitor. *Biochemistry* **10**: 470–476.
- Chakrabarty, A., Schellman, J.A., and Baldwin, R.L. 1991. Large differences in the helix propensities of alanine and glycine. *Nature* **351**: 586–588.
- Chakrabarty, A., Doig, A.J., and Baldwin R.L. 1993a. Helix capping propensities in peptides parallel those in proteins. *Proc. Natl. Acad. Sci.* **90**: 11332–11336.
- Chakrabarty, A., Kortemme, T., Padmanabhan, S., and Baldwin, R.L. 1993b. Aromatic side-chain contribution to far-ultraviolet circular dichroism of helical peptides and its effect on measurement of helix propensities. *Biochemistry* **32**: 5560–5565.
- Chakrabarty, A., Kortemme, T., and Baldwin, R.L. 1994. Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. *Protein Sci.* **3**: 843–852.
- Doig, A.J. and Baldwin, R.L. 1995. N- and C-capping preferences for all 20 amino acids in α -helical peptides. *Protein Sci.* **4**: 1325–1336.
- Doig, A.J., Chakrabarty, A., Klingler, T.M., and Baldwin, R.L. 1994. Determination of free energies of N-capping in α -helices by modification of the Lifson-Roig helix-coil theory to include N- and C-capping. *Biochemistry* **33**: 3396–3403.
- Doig, A.J., MacArthur, M.W., Stapley, B.J., and Thornton, J.M. 1997. Structures of N-termini of helices in proteins. *Protein Sci.* **6**: 147–155.
- Edelhoch, H. 1967. Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry* **6**: 1948–1954.
- Fersht, A.R. 1999. Structure and mechanisms in protein science: A guide to enzyme catalysis and protein folding. W.H. Freeman, New York.
- Forood, B., Feliciano, E.J., and Nambiar, K.P. 1993. Stabilization of α -helical structures in short peptides via end capping. *Proc. Natl. Acad. Sci.* **90**: 838–842.
- Horovitz, A., Matthews, J.M., and Fersht, A.R. 1992. α -helix stability in proteins. II. Factors that influence stability at an internal position. *J. Mol. Biol.* **227**: 560–568.
- Hutchinson, E.G. and Thornton, J.M. 1994. A revised set of potentials for β -turn formation in proteins. *Protein Sci.* **3**: 2207–2216.
- Keim, P., Vigna, R.A., Nigen, A.M., Morrow, J.S., and Gurd, F.R. 1974. Carbon 13 nuclear magnetic resonance of pentapeptides of glycine containing central residues of methionine, proline, arginine, and lysine. *J. Biol. Chem.* **249**: 4149–4156.
- Kumar, S. and Bansal, M. 1998. Dissecting α -helices: Position-specific analysis of α -helices in globular proteins. *Proteins* **31**: 460–476.
- Kyte, J. 1995. Structure in protein chemistry, Garland, New York.
- Lecomte, J.T.J. and Moore, C.D. 1991. Helix formation in Apocytochrome-B5—The role of a neutral histidine at the N-cap position. *J. Am. Chem. Soc.* **113**: 9663–9665.
- Lyu, P.C., Wemmer, D.E., Zhou, H.X., Pinker, R.J., and Kallenbach, N.R. 1993. Capping interactions in isolated α -helices: Position-dependent substitution effects and structure of a serine-capped peptide helix. *Biochemistry* **32**: 421–425.
- Myers, J.K., Pace, C.N., and Scholtz, J.M. 1997. A direct comparison of helix propensity in proteins and peptides. *Proc. Natl. Acad. Sci.* **94**: 2833–2837.
- Nozaki, Y. and Tanford, C. 1967. Intrinsic dissociation constants of aspartyl and glutamyl carboxy groups. *J. Biol. Chem.* **242**: 4734–4735.
- O’Neil, K.T. and DeGrado, W.F. 1990. A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science* **250**: 646–651.
- Pace, C.N. and Scholtz, J.M. 1998. A helix propensity scale based on experimental studies of peptides and proteins. *Biophys. J.* **75**: 422–427.
- Park, S.H., Shalongo, W., and Stellwagen, E. 1993. Residue helix parameters obtained from dichroic analysis of peptides of defined sequence. *Biochemistry* **32**: 7048–7053.
- Penel, S., Hughes, E., and Doig, A.J. 1999. Side-chain structures in the first turn of the α -helix. *J. Mol. Biol.* **287**: 127–143.
- Petukhov, M., Muñoz, V., Yumoto, N., Yoshikawa, S., and Serrano, L. 1998. Position dependence of non-polar amino acid intrinsic helical propensities. *J. Mol. Biol.* **278**: 279–89.
- Petukhov, M., Uegaki, K., Yumoto, N., Yoshikawa, S., and Serrano, L. 1999. Position dependence of amino acid intrinsic helical propensities. II. Non-preferred polar residues: Ser, Thr, Asn, and Gln. *Protein Sci.* **8**: 2144–2150.
- Presta, L.G. and Rose, G.D. 1988. Helix signal in proteins. *Science* **240**: 1632n–1641.
- Richardson, J.S. and Richardson, D.C. 1988. Amino acid preferences for specific locations at the ends of a helices. *Science* **240**: 1648–1652.
- Rohl, C.A., Chakrabarty, A., and Baldwin, R.L. 1996. Helix propagation and N-cap propensities of the amino acids measured in alanine-based peptides in 40 volume percent trifluoroethanol. *Protein Sci.* **5**: 2623–2637.
- Serrano, L. and Fersht, A.R. 1989. Capping and α -helix stability. *Nature* **342**: 296–299.
- Serrano, L., Sancho, J., Hirshberg, M., and Fersht, A.R. 1992. α -helix stability in proteins. I. Empirical correlations concerning substitution of side-chains at the N- and C-caps and the replacement of alanine by glycine or serine at solvent-exposed surfaces. *J. Mol. Biol.* **227**: 544–559.
- Stapley, B.J. and Doig, A.J. 1997. Hydrogen bonding interactions between glutamine and asparagine in α -helical peptides. *J. Mol. Biol.* **272**: 465–473.
- Sun, J.K., Penel, S., and Doig, A.J. 2000. Determination of α -helix N1 energies by addition of N1, N2 and N3 preferences to helix-coil theory. *Protein Sci.* **9**: 750–754.
- Tanford, C. 1962. The interpretation of hydrogen ion titration curves of proteins. *Adv. Prot. Chem.* **17**: 69–161.
- Yang, J., Spek, E.J., Gong, Y., Zhou, H., and Kallenbach, N.R. 1997. The role of context on α -helix stabilization: Host-guest analysis in a mixed background peptide model. *Protein Sci.* **6**: 1264–1272.
- Yumoto, N., Murase, S., Hattori, T., Yamamoto, H., Tatsu, Y., and Yoshikawa, S. 1993. Stabilization of α -helix in C-terminal fragments of neuropeptide Y. *Biochem. Biophys. Res. Comm.* **196**: 1490–1495.