Protein Science (1999), *8*:2144–2150. Cambridge University Press. Printed in the USA.
Copyright © 1999 The Protein Society

Position dependence of amino acid intrinsic helical propensities II: Non-charged polar residues: Ser, Thr, Asn, and Gln

MICHAEL PETUKHOV,[1] KOICHI UEGAKI,[2] NOBORU YUMOTO,[2]
SUSUMU YOSHIKAWA,[2] AND LUIS SERRANO[1]
[1]European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, Heidelberg D-69012, Germany
[2]Osaka National Research Institute, AIST, Ikeda, Osaka 563, Japan

(RECEIVED March 29, 1999; ACCEPTED July 5, 1999)

Abstract

The assumption that the intrinsic α-helical propensities of the amino acids are position independent was critical in several helix/coil transition theories. In the first paper of these series, we reported that this is not the case for Gly and nonpolar aliphatic amino acids (Val, Leu, Met, and Ile). Here we have analyzed the helical intrinsic propensities of noncharged polar residues (Ser, Thr, Asn, and Gln) at different positions of a model polyalanine-based peptide. We found that Thr is more favorable (by ~0.3 kcal/mol) at positions N1 and N2 than in the helix center, although for Ser, Asn, and Gln the differences are smaller (±0.2 kcal/mol), and in many cases within the experimental error. There is a reasonable agreement (±0.2 kcal/mol) between the calculated free energies, using the ECEPP/2 force field equipped with a hydration potential, and the experimental data, except at position N1.

Keywords: α-helix; entropy; folding; hydration; secondary structure; stability

The α-helix is one of the most frequent types of secondary structure elements in proteins and probably the best characterized experimentally (Chakrabartty & Baldwin, 1995; Muñoz & Serrano, 1995b). One of the main contributors to the free energy in α-helices are the so-called helix secondary structure propensities. The intrinsic secondary-structure propensities are defined as the free energy cost required to fix an amino acid in helical angles, excluding the main-chain–main-chain hydrogen contribution, side-chain–side-chain interactions, and electrostatic interactions with the helix macrodipole (Muñoz & Serrano, 1995a). Differences in helix-forming propensities of natural amino acids have been interpreted in terms of configurational entropy (Creamer & Rose, 1994; Lee et al., 1994), hydrophobic effects (Blaber et al., 1993), and electrostatic interactions (Avbelj & Moult, 1995).

The assumption of position independence of α-helix propensities was critical for the development of statistical mechanical helix-coil transition models published since the 1950s. However, the positions in the first α-helix turn are not geometrically equivalent to the rest of the helix. Amino acid side chains at the first helix turn are more solvent exposed; they have a fewer number of intramolecular Van der Waals' contacts and, due to absence steric clashes with the previous helix turn, they posses higher configurational entropy than that in a central position of an α-helix. In the first paper of these series, we reported that the α-helix propensity is indeed different for Gly, and some other nonpolar amino acids at each position of the first helix turn and in the middle of a α-helix (Petukhov et al., 1998). The effect was found to have a complex nature and varies in magnitude and sign for different amino acids.

To study the significance of the positional effect for short polar amino acids, we have synthesized four series of 16 residue Ala-based peptides. These peptides have been substituted with Ser, Thr, Asn, and Gln at N-terminal positions N-cap to N4 and at the central position N_c (which corresponds to position N7 of the established nomenclature) (Richardson & Richardson, 1988). The host peptide was taken to be the same as in our previous paper (Petukhov et al., 1998). This peptide showed no aggregation complications and allows measuring (without the influence of side-chain–side-chain interactions) the helical propensities of the amino acids in the first four and in the central positions of the α-helix. The peptides were analyzed by far-ultraviolet (UV) circular dichroism (CD), and the results interpreted in free energy terms by using the helix/coil transition theory (AGADIR1s; Muñoz & Serrano, 1997), with some modifications (AGADIR1s-2; Lacroix et al., 1998). To explain the observed differences in free energy, we estimated the average free energy of unfolded and folded states using molecular mechanics calculations.

Reprint requests to: Michael Petukhov, European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, Heidelberg D-69117, Germany; e-mail: Petukhov@EMBL-Heidelberg.de.

Throughout the paper we will follow Richardson and Richardson's (1988) nomenclature for indicating the amino acid position in an α-helix:

Ncap	N+1	N+2	N+3	N+4	....	Nc	....	C+4	C+3	C+2	C+1	Ccap
Rc	H	H	H	H	....	H	....	H	H	H	H	Rc

where Rc is a nonhelical residue and H is a helical residue.

Results and discussion

The CD measurements

In this work we used the same 16 residues host peptide:

1 2 3 4 5 6 7
Fr/Ac-Ala-Ala-Ala-Ala-Ala-Ala-Ala-Ala-Arg-Ala-Ala-Ala-Arg-Gly-Gly-Tyr-Am

as in the first paper of these series (Petukhov et al., 1998). The N-termini were free (Fr) or acetylated (Ac), and the C-termini were amidated (Am). C-terminal Tyr, separated by a Gly, has been used for accurate measurements of peptide concentration and to diminish the aromatic contribution (Chakrabartty et al., 1993). There are also two Arg at the C-terminal half of the peptide to favor peptide solubility.

This time we are focused on the intrinsic propensities of short polar guest residues (Ser, Thr, Asn, and Gln) in the conformational context of a right-handed α-helix. The amino acid exchanges were made at N-terminal positions 1 to 4 and at position 7 of the template sequence. The structure of the host sequence and the positions selected for mutations are designed to safely measure the amino acids helical propensities without complications from charge-end effects and side-chain–side-chain interactions. CD measurement of the peptide series with free N-termini at pH 10 having different residues at position 1 allows determination of the N-cap propensities of the corresponding amino acid. The acetylated peptides can be used for determination of the amino acids intrinsic helical propensities at different helix positions (see Materials and methods).

Figure 1 shows the CD spectra obtained for the series of synthetic peptides substituted with Ser, Thr, Asn, and Gln at the positions indicated above. Table 1 summarizes the results of the experiments, the peptide sequences, and the estimated experimental helical content (Chen et al., 1974) for all the peptides analyzed

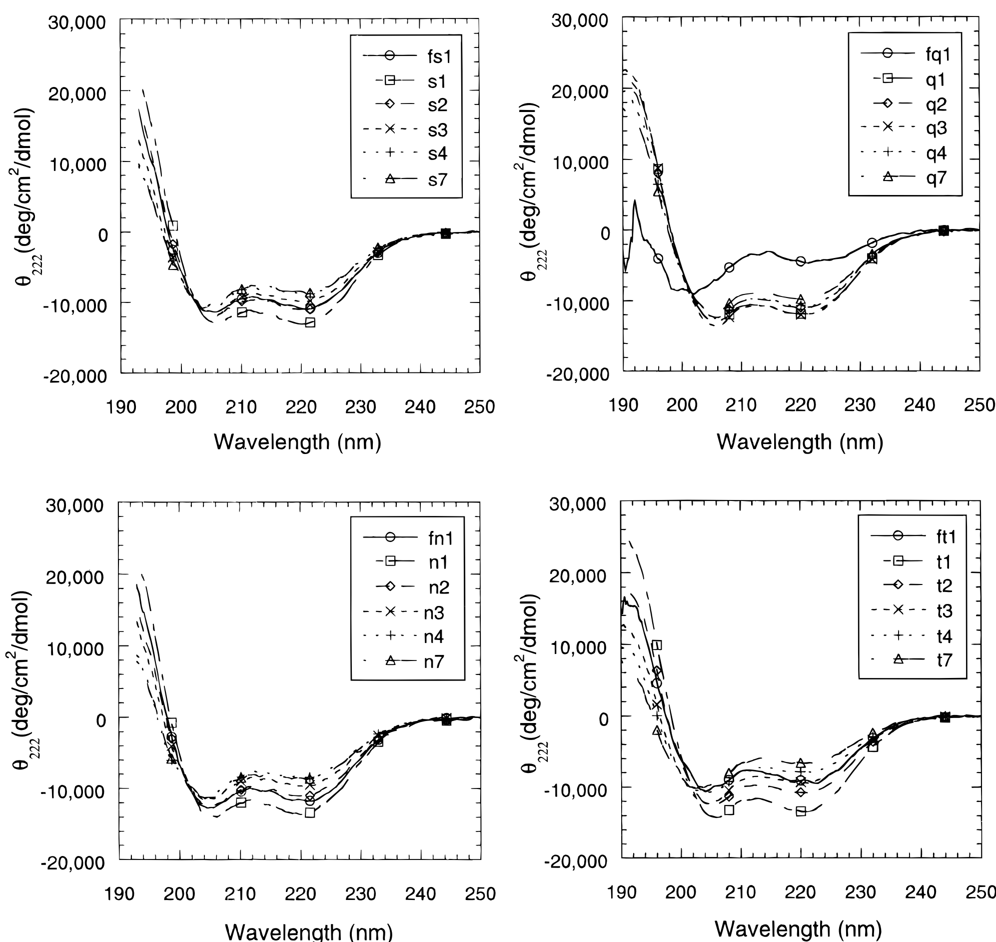

Fig. 1. CD spectra of the series of peptides used in this study. The indexes of the peptides are indicated in the figure panels. The sequences, mean residue ellipticities, and estimations of helical content are given in Table 1. Other experimental conditions are given in Materials and methods.

Table 1. *Sequences and results of CD measurements of the helical contents for the series of synthetic peptides used in this study*[a]

Peptide	Sequence	$-[\theta]_{222}$ (deg·cm^2/dmol)	Helical content (%)
fs	NH2-SAAAAAAARAAARGGY-NH2	10,720 ± 350	32.8 ± 1.0
s1	Ac-SAAAAAAARAAARGGY-NH2	12,700 ± 400	38.3 ± 1.2
s2	Ac-ASAAAAAARAAARGGY-NH2	10,880 ± 350	32.8 ± 1.0
s3	Ac-AASAAAAARAAARGGY-NH2	9,610 ± 300	29.0 ± 1.0
s4	Ac-AAASAAAARAAARGGY-NH2	8,760 ± 300	26.4 ± 1.0
s7	Ac-AAAAAASARAAARGGY-NH2	8,490 ± 300	25.6 ± 1.0
ft	NH2-TAAAAAAARAAARGGY-NH2	9,400 ± 300	28.7 ± 1.0
t1	Ac-TAAAAAAARAAARGGY-NH2	13,230 ± 400	39.9 ± 1.2
t2	Ac-ATAAAAAARAAARGGY-NH2	10,750 ± 350	32.4 ± 1.0
t3	Ac-AATAAAAARAAARGGY-NH2	9,290 ± 300	28.0 ± 1.0
t4	Ac-AAATAAAARAAARGGY-NH2	7,870 ± 250	23.7 ± 1.0
t7	Ac-AAAAAATARAAARGGY-NH2	6,610 ± 200	19.9 ± 1.0
fn	NH2-NAAAAAAARAAARGGY-NH2	11,700 ± 350	35.7 ± 1.0
n1	Ac-NAAAAAAARAAARGGY-NH2	13,210 ± 400	39.8 ± 1.2
n2	Ac-ANAAAAAARAAARGGY-NH2	10,930 ± 350	33.0 ± 1.0
n3	Ac-AANAAAAARAAARGGY-NH2	9,550 ± 300	28.8 ± 1.0
n4	Ac-AAANAAAARAAARGGY-NH2	8,340 ± 300	25.2 ± 1.0
n7	Ac-AAAAAANARAAARGGY-NH2	8,560 ± 300	25.8 ± 1.0
fq	NH2-QAAAAAAARAAARGGY-NH2	4,500 ± 200	13.7 ± 1.0
q1	Ac-QAAAAAAARAAARGGY-NH2	11,730 ± 350	35.4 ± 1.0
q2	Ac-AQAAAAAARAAARGGY-NH2	10,810 ± 350	32.6 ± 1.0
q3	Ac-AAQAAAAARAAARGGY-NH2	11,800 ± 350	35.6 ± 1.1
q4	Ac-AAAQAAAARAAARGGY-NH2	10,700 ± 350	32.3 ± 1.0
q7	Ac-AAAAAAQARAAARGGY-NH2	9,690 ± 300	29.2 ± 1.0

[a]Far-UV CD spectra of the peptides were obtained at pH 10 (FX series), or pH 7 for the rest of the peptides, in 5 mM sodium phosphate buffer at 5 °C. Peptide concentrations were 10 μM. The percentage of α-helix was calculated with the empirical equation $-100*(\theta_{222}/39{,}500(1\text{–}2.57/n)))$ (Chen et al., 1974); where n is a number of peptide bonds, and θ_{222} is an experimentally observed ellipticity of peptide at 222 nm. The error estimates in θ_{222} and in corresponding helical content are based on approximately 3% errors in peptide concentration measurements.

in this work. There is an excellent correlation between the concentration independent parameter R1 (this value is obtained by dividing the ellipticity at 193 nm by the ellipticity minimum in the range 200 to 210 nm; Bruch et al., 1991) and the helical content determined from the ellipticity at 222 nm (data not shown). This indicates that errors in the measurements of peptide concentration are minor.

Determination of the amino acid helical propensities using AGADIR1s-2

In peptide helices there is not a single α-helix in equilibrium with the coil state, but a broad ensemble of helical conformations with different lengths and involving different residues of the sequence. Therefore, the experimental helical contents can only be interpreted in energy terms of intrinsic helical propensities by fitting the far-UV CD data for the different peptides to a statistical mechanical model of the helix/coil transition (see Materials and methods).

First of all, N-capping propensities of Ser, Thr, Asn, and Gln were varied so as to reproduce the experimental numbers of the corresponding N-terminus unprotected peptides (final values: −0.6, −0.5, −0.65, and +0.5 kcal/mol, respectively). These values are in the same order as found by Doig and Baldwin (1995): Asn > Ser > Thr > Gln. Comparison of the N-cap values published by Baldwin's group; those found in proteins and our values shows that all are in the same range of free energies. Moreover, they are almost identical to the values obtained in AGADIR1s-2 after fitting several hundreds of peptides. The differences found in some cases with other systems could be due to context effects of the particular system used. For example, fitting the peptides described by (Doig & Baldwin, 1995) with AGADIR1s-2 results in the N-cap values reported in this paper except for Asn, which is more favorable in Doigs' series. Matthews and co-workers (Bell et al., 1992) indicated that Asn is more favorable than Ser or Thr at the N-cap only when certain dihedral angles for the first helical turn are found. In the case of Doig and Baldwin, there is a Lys at position N2, and this can change the conformation of the first helical turn and, therefore, the preference for Asn at N-cap.

Table 2 shows the results of fitting AGADIR1s-2 set of energy parameters to the CD experimental helical contents as described in Materials and methods. In the majority of the cases, the differences in helical propensities between the central position and the rest is within the experimental error. The main exception is Thr, where the intrinsic helical propensities are found to have a significant position dependence (up to −0.3 kcal/mol increase at position N1 and N2), compared with the center of the α-helix.

Table 2. *Free-energy contributions to the intrinsic helical propensities of short polar amino acids at central and four N-terminal positions of a nine-residue α-helix model*[a]

AA/position	ΔE^{ECEPP} (kcal/mol)	ΔE^{Hydr} (kcal/mol)	$-T*\Delta S$ (kcal/mol)	$\Delta\Delta G_{theor}$ relative to Nc(N+5) (kcal/mol)	$\Delta\Delta G_{exp}$ relative to Nc(N7) (kcal/mol)	$\Delta\Delta G_{exp}$ relative to Ala (kcal/mol)
Ser/N1	−18.12	21.19	0.588	−0.706	−0.10	0.45
Ser/N2	−17.36	21.27	0.403	−0.051	0.00	0.55
Ser/N3	−17.63	21.48	0.549	0.036	0.10	0.65
Ser/N4	−17.60	21.38	0.592	0.009	0.15	0.70
Ser/N_c	−17.65	21.37	0.644	0.000	0.00	0.55
Thr/N1	−17.96	21.53	0.279	−0.942	−0.33	0.50
Thr/N2	−17.55	21.56	0.227	−0.553	−0.33	0.50
Thr/N3	−18.37	22.48	0.408	−0.272	−0.18	0.65
Thr/N4	−18.43	22.86	0.335	−0.025	−0.10	0.73
Thr/N_c	−18.51	22.94	0.360	0.000	0.00	0.83
Asn/N1	−18.11	21.38	0.371	−0.483	0.00	0.60
Asn/N2	−17.31	21.32	0.151	0.037	0.10	0.70
Asn/N3	−17.74	21.45	0.443	0.029	0.10	0.70
Asn/N4	−17.71	21.50	0.517	0.183	0.20	0.80
Asn/N_c	−17.82	21.43	0.514	0.000	0.00	0.60
Gln/N1	−18.54	21.05	0.903	−0.492	−0.10	0.32
Gln/N2	−18.09	21.11	0.696	−0.189	−0.12	0.30
Gln/N3	−18.55	21.48	0.860	−0.115	−0.17	0.25
Gln/N4	−18.43	21.41	0.918	−0.007	−0.09	0.33
Gln/N_c	−18.40	21.36	0.945	0.000	0.00	0.42

[a]The average values of E^{ECEPP} and E^{Hydr} were calculated separately for folded and unfolded states using Boltzmann factor for all possible conformers of a particular guest amino acid, as described in Materials and methods. Values of ΔE^{ECEPP} and ΔE^{Hydr} were calculated as the differences between folded and unfolded states. Configurational entropy S was calculated using classical Boltzmann formula $S = -R\Sigma P_i \ln(P_i)$, where R is a gas constant and P_i are the probabilities of the conformational states, which in turn were calculated from the canonical Boltzmann–Gibbs distribution.

[b]The changes of intrinsic helical propensities between the central and several N-terminal positions were obtained by fitting AGADIR1s set of energy parameters (Muñoz & Serrano, 1995a, 1997) to CD measurements of helical content of the Ala-based peptides listed in the Table 1. The errors in ΔG_{exp} for positions N2 to N4 and N7 is around 0.1 kcal/mol, while at position N1 it is around 0.2 kcal/mol. The error in the estimations was obtained by determining the intrinsic contribution that will give a helical content 3% higher than the experimental one. The difference between the intrinsic contribution obtained by reproducing the experimental value and that described above is the error in the energy estimation.

The energy contributions to the position dependence

We performed molecular mechanics calculations to elucidate the physical reasons behind the observed position dependence of helical propensities of these four short polar amino acids. Preliminary calculations showed very high (up to 2.0 kcal/mol) differences in the intrinsic propensities for Asn and Gln at different positions. A closer inspection of these results indicated that these artifacts are solely associated with a slightly greater exposition of the Asn and Gln side-chain carbonyl carbons at the first α-helix turn. These changes in ASA are as small as 4–5 Å^2, which is even less than the surface area required for one water molecule to make an H-bond in a water–protein interface. The anomalous high value (0.427 kcal/mol/Å^2) of carbonyl carbon solvation (Ooi et al., 1987) can result in large changes in the amino acid helical propensities, when the change in ASA is small. Therefore, we replaced the solvation potential with another one also derived from the experimental data of vapor-to-water transitions of small organic compounds (Wesson & Eisenberg, 1992). It is worth noting that atomic solvation parameters of these two models are close to each other except for the carbonyl carbon. Because we have changed the solvation potential, we have performed the same calculations for Gly, Val, Ile, Leu, and Met, and compared them with the ones previously reported (Petukhov et al., 1998). The results of the calculations are very similar to the old ones (Petukhov et al., 1998) as expected from the absence of a carbonyl carbon in the side chains of these residues (data not shown).

To perform our calculations, we used the same model polyalanine helix (Petukhov et al., 1998) as described in Materials and methods with acetylated N- and amidated C-termini. Table 2 shows the free-energy contributions, from nonbonded interactions, solvation, and configurational entropy, to the intrinsic helical propensities of the amino acids at different positions in helix.

The value of ΔE^{ECEPP} presents the weighted average change of nonbonded interactions between folded and unfolded states. Due to a fewer number of Van der Waals' contacts, the nonbonded interactions between amino acids with a side chain and α-helices are generally less favorable in the fist helix turn than in the center of a helix. This effect is partly compensated by the unfolded state of the peptide molecule that has the same chemical structure and where the terminal residues have less number of Van der Waals' contacts as well. The opposite is true for the solvation and entropic energy terms. The side chains of short polar residues are more solvent exposed at the N-terminal positions than in the helix center

(depending on the amino acids, between 0.1 to 1.0 kcal/mol of additional helix stabilization). The configurational entropy of side chains is higher at the N-terminal positions due to the absence of steric restrictions associated with the previous helix turn. This also contributes favorably to free energy of polar residues at the N-terminal positions. Thus, there is an accurate balance of different nature forces that determinates helical propensities of amino acid residues at a given position inside the α-helix.

Values of $\Delta\Delta G_{theor}$ shows the changes of intrinsic helical propensities relative to the central helix position as calculated with molecular mechanics (Table 2). These data can be directly compared with experimental values determined from CD spectra using AGADIR1s-2. The differences between the theoretical and the experimental propensities are within ± 0.2 kcal/mol, and the sign and magnitude of the energy changes are predicted well, except for position N+1 of an α-helix. Considering all the simplifications and approximations that had to be taken in molecular mechanics calculation, an accuracy of ± 0.2 kcal/mol for all positions, except N+1, should be considered as reasonably good. The correlation between the experimental and the theoretical values of $\Delta\Delta G$, without position N+1, for polar and apolar amino acids is quite high ($R = 0.904$, data not shown).

In general, the theoretical calculations tend to predict that short polar residues are more favorable at position N+1 of α-helices than is found experimentally (Table 2). The most significant differences with the helical propensities of a central residue (Nc) are found for Thr at the N-terminal positions. The qualitative trend is correctly predicted, although the magnitude of the effect is overestimated.

Although there are small energy differences (~0.1 kcal/mol), positions Nc, N+4, and N+3 generally show similar properties of nonbonded interactions, solvation, and entropy losses. Residues at position N+2 have significantly (0.2–0.4 kcal/mol) lower loss of configurational entropy, because they are less conformationally restricted. This in agreement with relatively less favorable nonbonded interactions and significantly more favorable solvent exposition of the amino acid side chains at position N+2 than at positions Nc, N+4, and N+3 (Table 2). Similar trends are observed for the non-polar amino acids, expect that the sign of the solvation term was opposite (Petukhov et al., 1998).

The unique environment of position N+1 (it has an acetyl group preceding it, while all other residues are preceded by another amino acid) can explain the relatively poor correlation between theoretical predictions and experimental data for substitutions at this position of the helix. The average conformational behavior in the denatured state of the polypeptide chain at position N+1 can be different from that at other positions. In addition, the possible formation of non-α-helical conformations, such as 3_{10}-helices that are quite probable at the end of helices, could explain the behavior of the N+1 position. In fact, NMR measurements of hydrogen/deuterium-exchange data for the first three amide protons of similar acetylated 16 residues Ala-based peptides indicate partial protection for the NH group of Ala3. This is in agreement with the presence of an $i \rightarrow i + 3$ hydrogen bond between the carbonyl group of the acetyl blocking group and the NH group of Ala3 (Millhauser et al., 1997). The formation of 3_{10}-helices in the middle of Ala-based peptide helices has been reported to be less probable than at the N-terminus of these peptides. This supports the idea that the errors in the prediction could be due to a different conformational behavior of a guest residue at this position in the folded and unfolded states, more than to the parameter set of model.

Statistical survey of the protein database

The amino acid propensities derived from the statistical analysis of representative set protein crystal structures were found to correlate, at least for some amino acids, with the helical propensities measured by CD (Muñoz & Serrano, 1994; Swindells et al., 1995; Petukhov et al., 1998). Therefore, we expected to find some correlation between the experimental positional helical propensities and the relative preferences of short polar amino acids to occupy N-terminal and central positions of protein helices. Given the rather weak position dependence of helical propensities found for the majority of the amino acids used in this study and the presence of side-chain–side-chain and tertiary interactions, it would be difficult to expect a good correlation in all cases. However, for those positions where significant changes of the intrinsic helical propensities were observed, we can expect some correlation with amino acid frequencies derived form protein database.

The correlation coefficients calculated for the amino acid propensities of Table 2 and those obtained from the protein database are as expected not high ($R \sim 0.5$, data not shown). However, there is a reasonable qualitative correlation. In the majority of cases, the sign of the effect is predicted correctly. Position N+4 is similar, or less favorable, when compared to central helix positions. Position N+3 is more favorable for all amino acids, except Asn. The significant preference of Gln at position N+3 in protein helices is known to be due to Capping Box reciprocal hydrogen bonds between the side chain of residue N+3 and the main-chain amide group of the N-cap residue, and between the side chain of the N-cap residue and the main-chain amide group of residue N+3. However, in our peptides, Ala cannot provide the side-chain–main-chain hydrogen bond as a good N-cap residue (Asn, Asp, Ser, and Thr) does. Moreover, in the case of Gln at position 3 of our peptide, there is no backbone amide at the N-cap position (the acetyl group). Therefore, there is only a minor (0.1 kcal/mol) increase of Gln intrinsic propensity at position N+3 in our theoretical calculations, in agreement with the experimental data.

Recent statistical analysis of first N-terminal turn of the protein α-helices also revealed a strong positional dependence of both the amino acid and the side-chain rotamer propensities at the first free N-terminal positions (N+1, N+2, and N+3) of an α-helix (Penel et al., 1999). Particularly, Ser was found to have a strong preference for positions N+1, N+2, and N+3. Thr and Gln are most favorable only at position N+3 of protein α-helices. Asn showed no significant preference for the first α-helix turn. This is in reasonable quality agreement with our results.

Conclusions

We have experimentally analyzed the intrinsic helical propensity of several noncharged polar amino acids at the first helix turn and at central positions of an α-helix. Our results show minor differences (± 0.2 kcal/mol) for Ser, Asn, and Gln amino acid α-helical propensities at N-terminal positions and central positions. Thr, on the other hand, is significantly more favorable in the first α-helical turn than in the center of the helix. The calculations based on an ECEPP/2 force field equipped with hydration potential indicate that the position effect can be rationalized in terms of three factors that make the first turn different from the center: (a) greater solvent exposure of the side chains, (b) less number of intramolecular nonbonded contacts, and (c) higher configurational entropy.

Therefore, it is expected that the position dependence of the intrinsic helical propensities for these amino acids will vary, de-

pending on solvent and chemical conditions. This is in agreement with experimental measurements of intrinsic helical propensities of natural amino acids measured at high concentration of trifluoroethanol (Rohl et al., 1996) and in apolar membrane environment (Deber & Li, 1995). In both systems it has been found that the amino acid helical propensities significantly differ from that of water solution. Taking into account that three main contributors to the helix intrinsic propensity of the amino acids are different at each position, it is surprising that in water the differences cancel out in the majority of the cases.

Materials and methods

Experimental procedures

Peptide synthesis

The peptides were synthesized in two series. One series of the peptides were synthesized on an automated solid-phase peptide synthesizer (Shimadzu PSSM-8), using Tenta Gel TG-RAM resin and Fmoc chemistry, with benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate and N-hydroxybenzotriazole as coupling reagents. Peptides were cleaved from the resin by trifluoroacetic acid and purified by reverse-phase high-performance liquid chromatography (RP-HPLC) on a C_{18} column. Fmoc-L-amino acids, reagents for peptide synthesis and Tenta Gel TG-RAM resin, were purchased from Shimadzu. The purity of each peptide was assessed by analytical RP-HPLC on a C_{18} column. Molecular masses were confirmed by mass spectrometry on a time-of-flight mass spectrometer (Shimadzu/Kratos Kompact MALDI II) with matrix-assisted laser desorption ionization. Another series of the peptides were synthesized as described in Petukhov et al. (1998).

CD measurements

The CD measurements were done using both peptide series at two concentrations (10 and 50 μM). Each measurement was repeated at least twice, using fresh peptides solutions prepared in the same day. The concentration of peptide was determined by UV absorbance of C-terminal Tyr using the method of (Gill & von Hippel, 1989). No concentration dependence of the CD spectra was found in tested concentration range. All the peptides were found to be soluble, at least, at 500 μM concentration. The error in the concentration determination was around 2%. CD spectra were recorded on a Jasco-710 instrument at a pH 7 and a temperature of 278 K, as described in Petukhov et al. (1998). The helical content of the peptides has been estimated using the mean residue ellipticity at 222 nm (Chen et al., 1974).

Calculations based on statistical mechanics

The intrinsic helical propensities were determined from CD measurements using the AGADIR1s computer program (Muñoz & Serrano, 1997), modified to include the possibility of the residue immediately following an acetyl group, or preceding and amide group, to be helical (Lacroix et al., 1998), as described in Petukhov et al. (1998).

Calculations based on molecular mechanics

Energy profiles for the amino acids were calculated using ECEPP/2 force field for terminal positions N1, N2, N3, N4, and central position N_c [corresponds to position N5 of the standard nomenclature (Richardson & Richardson, 1988)] of nine-residue Ala-based model helices as described in Petukhov et al. (1998).

The solvation energy term was modeled by a continuum approximation model for protein solvent interactions. In the previous work of this series (Petukhov et al., 1998), we used ASA-based solvation potential by Ooi et al. (1987). However, the anomalous high value of the original energy parameter for the carbonyl carbons (427 cal/mol/Å^2) was criticized in the literature (Juffer et al., 1995) and was found to produce too high changes in solvation energy for the Asn and Gln peptide series. Therefore, in this work we used another accessible surface area (ASA) based solvation potential that also was verified on the experimental data of vapor to water transitions of small organic compounds (Wesson & Eisenberg, 1992). ASA was calculated with the NSC program (Eisenhaber et al., 1995). We recalculated the intrinsic helical propensities for Gly, Val, Ile, Leu, and Met previously published using new solvation potential and found them in a good agreement with experimental data (data not shown).

Energy profiles were calculated on the grid of ϕ, ψ, ω, $\chi 1$, $\chi 2$ and, where applicable, $\chi 3$ with grid steps of 20°, and 50 subsequent steps of energy minimization by the conjugate gradient method. The dihedral angles ϕ, ψ, and ω of the backbone of peptides in the folded state were fixed in standard values of $-60°$, $-40°$, and 180°, respectively. In the unfolded state ϕ, ψ, and ω, angles of all residues were initially set to 180° and allowed to vary (except for the guest residue) by the conjugate gradient energy minimization algorithm. Thus, unlike our previous work (Petukhov et al., 1998), the changes of main-chain entropy is included here. The configurational entropy of peptides in folded and unfolded states was calculated from the energy profiles at $T = 278$ K using the classic Boltzmann–Gibbs approach as described in Petukhov et al. (1998).

The conformational states for peptide backbone were separated as described by Zimmerman et al. (1977): there are six conformers for ϕ (0–40°, 40–110°, 110–180°, 180–250°, 250–320°, and 320–360°), and five conformers for ψ (0–20°, 20–110°, 110–220°, 220–270°, and 270–360°). Thus, 30 possible conformers of main chain are considered. The conformational states of side chains were separated as described by Lee et al. (1994).

Survey of the Protein Data Bank

The amino acid frequencies at different helix positions were derived with the WHATIF program from 315 protein crystal structures at better than 2.1 Å resolution, with less than 25% homology, and with R-factor below 0.21 (Vriend, 1990). The crystal structures of the proteins were taken from the Brookhaven Protein Data Bank (Bernstein et al., 1977). We searched for the sequence motif STC/H/H/H/H/H/H/H/H/H, where S is strand, T is turn, C is coil, and H is helix. The amino acid frequencies at central position were calculated as average of three central positions N+5, N+6, and N+7.

Acknowledgments

M.P. acknowledges financial support from EC Grant BIO4-CT97-2086. Dr. F. Eisenhaber, the author of the NSC program, generously provided the source code of this algorithm.

References

Avbelj F, Moult J. 1995. Role of electrostatic screening in determining protein main chain conformational preferences. *Biochemistry 34*:755–764.

Bell JA, Becktel WJ, Sauer U, Baase WA, Matthews BW. 1992. Dissection of helix capping in T4 lysozyme by structural and thermodynamic analysis of six amino acid substitutions at Thr 59. *Biochemistry 31*:3590–3596.

Bernstein FC, Koetzle TF, Williams GJ, Meyer EE Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. 1977. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol 112*:535–542.

Blaber M, Zhang XJ, Matthews BW. 1993. Structural basis of amino acid alpha helix propensity. *Science 260*:1637–1640.

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

Chakrabartty A, Baldwin RL. 1995. Stability of alpha-helices. *Adv Protein Chem 46*:141–176.

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

Chen YH, Yang JT, Chau KH. 1974. Determination of the helix and beta form of proteins in aqueous solution by circular dichroism. *Biochemistry 13*: 3350–3359.

Creamer TP, Rose GD. 1994. Alpha-helix-forming propensities in peptides and proteins. *Proteins 19*:85–97.

Deber CM, Li SC. 1995. Peptides in membranes: Helicity and hydrophobicity. *Biopolymers 37*:295–318.

Doig AJ, Baldwin RL. 1995. N- and C-capping preferences for all 20 amino acids in alpha-helical peptides. *Protein Sci 4*:1325–1336.

Eisenhaber F, Lijnzaad P, Argos P, Sander C, Scharf M. 1995. The double cubic lattice method: Efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *J Comput Chem 16*:273–284.

Gill SC, von Hippel PH. 1989. Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem 182*:319–326.

Juffer AH, Eisenhaber F, Hubbard SJ, Walther D, Argos P. 1995. Comparison of atomic solvation parametric sets: Applicability and limitations in protein folding and binding. *Protein Sci 4*:2499–2509.

Lacroix E, Viguera AR, Serrano L. 1998. Elucidating the folding problem of alpha-helices: Local motifs, long-range electrostatics, ionic-strength dependence and prediction of NMR parameters. *J Mol Biol 284*:173–191.

Lee KH, Xie D, Freire E, Amzel LM. 1994. Estimation of changes in side chain configurational entropy in binding and folding: General methods and application to helix formation. *Proteins 20*:68–84.

Millhauser GL, Stenland CJ, Hanson P, Bolin KA, van de Ven FJ. 1997. Estimating the relative populations of 3(10)-helix and alpha-helix in Ala-rich peptides: A hydrogen exchange and high field NMR study. *J Mol Biol 267*:963–974.

Muñoz V, Serrano L. 1994. Intrinsic secondary structure propensities of the amino acids, using statistical phi-psi matrices: Comparison with experimental scales. *Proteins 20*:301–311.

Muñoz V, Serrano L. 1995a. Elucidating the folding problem of helical peptides using empirical parameters. II. Helix macrodipole effects and rational modification of the helical content of natural peptides. *J Mol Biol 245*:275–296.

Muñoz V, Serrano L. 1995b. Helix design, prediction and stability. *Curr Opin Biotechnol 6*:382–386.

Muñoz V, Serrano L. 1997. Development of the multiple sequence approximation within the AGADIR model of alpha-helix formation: Comparison with Zimm-Bragg and Lifson-Roig formalisms. *Biopolymers 41*:495–509.

Ooi T, Oobatake M, Nemethy G, Scheraga HA. 1987. Accessible surface areas as a measure of the thermodynamic parameters of hydration of peptides. *Proc Natl Acad Sci USA 84*:3086–3090.

Penel S, Hughes E, Doig AJ. 1999. Side-chain structures in the first turn of the alpha-helix. *J Mol Biol 287*:127–143.

Petukhov M, Muñoz V, Yumoto N, Yoshikawa S, Serrano L. 1998. Position dependence of non-polar amino acid intrinsic helical propensities. *J Mol Biol 278*:279–289.

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

Rohl CA, Chakrabartty A, Baldwin RL. 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.

Swindells MB, MacArthur MW, Thornton JM. 1995. Intrinsic phi, psi propensities of amino acids, derived from the coil regions of known structures. *Nature Struct Biol 2*:596–603.

Vriend G. 1990. WHAT IF: A molecular modeling and drug design program. *J Mol Graph 8*:52–56.

Wesson L, Eisenberg D. 1992. Atomic solvation parameters applied to molecular dynamics of proteins in solution. *Protein Sci 1*:227–235.

Zimmerman SS, Pottle MS, Nemethy G, Scheraga HA. 1977. Conformational analysis of the 20 naturally occurring amino acid residues using ECEPP. *Macromolecules 10*:1–9.