Protein Science (1997), 6:1264–1272. Cambridge University Press. Printed in the USA.

The role of context on α-helix stabilization: Host-guest analysis in a mixed background peptide model

JIANXIN YANG, ERIK J. SPEK, YOUXIANG GONG, HONGXING ZHOU, AND NEVILLE R. KALLENBACH

Department of Chemistry, New York University, New York, New York 10003

(RECEIVED October 3, 1996; ACCEPTED March 6, 1997)

Abstract

The helix content of a series of peptides containing single substitutions of the 20 natural amino acids in a new designed host sequence, succinyl-YSEEEEKAKKAXAEEAEKKKK-NH_2, has been determined using CD spectroscopy. This host is related to one previously studied, in which triple amino acid substitutions were introduced into a background of Glu-Lys blocks completely lacking alanine. The resulting free energies show that only Ala and Glu^- prove to be helix stabilizing, while all other side chains are neutral or destabilizing. This agrees with results from studies of alanine-rich peptide models, but not the previous Glu-Lys block oligomers in which Leu and Met also stabilize helix. The helix propensity scale derived from the previous block oligomers correlated well with the frequencies of occurrence of different side chains in helical sequences of proteins, whereas the values from the present series do not. The role of context in determining scales of helix propensity values is discussed, and the ability of algorithms designed to predict helix structure from sequence is compared.

Keywords: α-helix; CD; helix propensities; helix stability; NMR; secondary structure; synthetic peptide

The α-helix is the most common secondary structure in proteins (Levitt, 1978; Fasman, 1989); roughly one third of the amino acid residues in globular proteins occur in α-helical conformation. The energetics of formation of α-helices and the basis for their frequency in the native state of proteins remain imperfectly understood. Several approaches have been taken in an effort to understand α-helix formation and its stability in water. These include mutation of helical sites in native proteins (Blaber et al., 1993a, 1995b), substitution studies in synthetic copolypeptides (Wojcik et al., 1990), in model helical peptides (Lyu et al., 1990; O'Neil & DeGrado, 1990; Chakrabartty et al., 1994), and in prenucleated short chains (Zhou et al., 1994; Groebke et al., 1996). A large number of natural and synthetic peptides form intramolecular α-helical structure in aqueous solution (Muñoz & Serrano, 1994); analysis of these has revealed some important factors which govern the formation and stability of α-helix.

The use of *de novo* designed synthetic short peptides of known amino acid length and sequence in principle opens a direct avenue for understanding the determination of α-helix formation. This approach was pioneered by Baldwin's group (Shoemaker et al., 1987) and extended by several others. The effects of pH, temperature and concentration on the stability of α-helix (Marqusee et al., 1987; Shoemaker et al., 1987; Lyu et al., 1990; Finkelstein et al., 1991), the role of individual side chains and side-chain–side-chain interactions (Kim et al., 1982; Lyu et al., 1990, 1992a; Finkelstein et al., 1991), position effects of amino acid substitutions in short peptides (Padmanabhan et al., 1990; Chakrabartty et al., 1991; Padmanabhan et al., 1991; Lyu et al., 1992a), as well as N and C terminal capping effects (Lyu et al., 1992b, 1993) have all been studied using CD spectroscopy to assess helix content. In addition, the distribution of helical structure has been determined using ^1H NMR (Liff et al., 1991; Lyu et al., 1993) while titration studies and DSC analysis have provided an estimate of the heat of formation for α-helix (Ptitsyn, 1972; Scholtz et al., 1991b), approximately 1 kcal/mol residues.

In many studies of the propensity of amino acids to form α-helix in short peptides, alanine substitution has been emphasized because of the absence of unfavorable entropy upon inserting its $-CH_3$ side chain into helix and its minimal capacity to interact with neighboring side chains. Marqusee et al. (1989) provided a direct indication of the high helix-forming propensity of alanine in their first designed helical peptide. Because oligomers of Ala longer than tetramers tend to associate in water, peptides composed of Ala with interposed Glu or Lys side chains for solubility have been used to obtain substantial α-helix formation in water (Scholtz & Baldwin, 1992). The high apparent helix-forming propensity of

Reprint requests to: Neville R. Kallenbach, Department of Chemistry, New York University, New York, NY 10003; e-mail: kallnbch@is.nyu.edu.

alanine has been confirmed by studies of a completely neutral, water-soluble peptide, containing Ala and Gln residues (Scholtz et al., 1991c). Another approach toward the problem of α-helix formation has been employed by the group of Scheraga. Their results are derived from work on host-guest polypeptides (Wojcik et al., 1990) and tri-block oligopeptide models (Ingwall et al., 1968). The results from both experiments give consistent *s*-values, which disagree with the results derived from short peptides. The reasons for this discrepancy have not been unambiguously resolved. On the one hand, it has been suggested that the host-guest studies include interactions between bulky host side chains and the guest species (Padmanabhan et al., 1994). On the other hand, it has been argued that the short peptides fail to take the hydration of the polar residues that make up part of the sequence into account (Vila et al., 1994). The results obtained by *de novo* designed peptides have also been challenged by Kemp's group, who believe that the intrinsic propensity of Ala is low in the absence of stabilizing interactions due to the presence of long side chains such as Lys, Glu, or Gln, used to solubilize peptides in water. In essence, they concur with Scheraga's report that Ala is effectively neutral with respect to stabilizing helix or coil (Groebke et al., 1996).

In 1990, Lyu et al. (1990) investigated a host system of helical peptides with the sequence:

succinyl-YSEEEEKKKKXXXEEEEKKKK-NH_2

(Y, Tyr; S, Ser; E, Glu; K, Lys; X, any one of 20 amino acid residues) that lacks Ala. One result they obtained was that the alanine substitution (X = Ala) maximized the helix content of these peptides with natural side chains. In this series of peptides, the positions $i \pm 4$ are occupied alternatively by the polar side chains Glu and Lys, on the grounds that such an arrangement stabilizes α-helix by facilitating the formation of salt bridges between acidic and basic side chains (Marqusee et al., 1987, 1989; Merutka & Stellwagen, 1991). The stabilization of α-helices by interactions involving Glu^- was also detected by a statistical analysis of the conformation of amino acid residues in proteins of known structure (Maxfield & Scheraga, 1975). A second result from this series of peptides was that in addition to Ala, Leu and Met are helix stabilizing residues; the order of stabilization and *s* values determined correlate well with the Pα values determined by surveying the composition of helices in proteins (Chou & Fasman, 1978). By contrast, studies of Ala-rich peptides (Chakrabartty et al., 1994) indicate that Leu and Met do not stabilize α-helix, prompting the concern that propensity values might be context dependent in detail and thus not transferable.

The site of substitution in the above peptides was flanked primarily by E and K, both long side chains with the potential to interact with side chains at neighboring sites (most notably, those spaced at sites $i \pm 1$, $i \pm 2$, and $i \pm 4$ from the substitution site). The question we are concerned with here is to what extent do scales of amino acid propensity derived from substitutions in a helical peptide vary with context? To address this, we have investigated the consequences of substituting 20 different amino acids in a host background where the $i \pm 1$ and $i \pm 4$ positions are occupied by A, rather than by E or K. Thus, we have modified the original sequence used by Lyu et al. (1990) to the following sequence:

succinyl-YSEEEEKAKKAXAEEAEKKKK-NH_2

In this sequence, the $i \pm 1$ and $i \pm 4$ positions from the site of substitution (denoted by X) are now occupied by alanine, reducing the potential for side-chain–side-chain interactions. There is still interaction possible at $i \pm 2$ and $i \pm 3$ positions. Although the former are not expected to have significant interactions, the latter can (Huyghues-Despointes et al., 1993). However, the use of charged residues cannot be completely avoided, otherwise solubility and aggregation problems may arise. Possible interactions between charged residues at the X-positions are evaluated by comparing their helix propensities with those of neutral residues at that position. If the correlation with a series that includes the charged residues is significantly different from one without these charged residues, then the interaction at $i \pm 2$ and $i \pm 3$ positions arguably play an important role.

The criterion commonly used to assess helicity is the CD spectral region near 200 nm, in which the α-helix is associated with well-defined minima at 222 nm and 208 nm, a maximum at 195 nm, and helix-coil isodichroic point at 204 nm (Woody, 1985; Yang et al., 1986). We have used CD to compare values of $[\theta]_{222}$ for 20 substituted peptides in this series, denoted as AXA (where X is any of 20 amino acids) at 4 °C, in 10 mM KF and pH 7, conditions that favor the formation of α-helix (Lyu et al., 1989), which is enthalpy driven in water (Scholtz et al., 1991a). Because temperature and peptide concentration also influence the formation and stability of α-helix, we measured the CD for AAA and AGA at different temperatures and for AAA as a function of concentration also. Salt concentration is a third important variable in terms of potential salt bridge formation between E and K. We have measured the CD spectra for AAA at different NaCl concentrations (0–5 M) in order to evaluate this effect. We have also defined the helix distribution within the AAA member of this series using NMR.

A second aspect of this study concerns the practical evaluation of programs designed to predict helical structure a priori from sequence data. Over the years, much effort has been devoted to developing algorithms for prediction of helical secondary structure starting from the primary sequence (Doig et al., 1994; Muñoz & Serrano, 1994; Shalongo & Stellwagen, 1995). We have used the AXA series of peptides to compare predictions of helix content by three such computer algorithms.

Results

CD measurements on synthetic AXA peptides

The spectra of AAA, AGA, and ASA are shown in Figure 1. The spectrum shows that the peptide has adopted an α-helical conformation; minima occur at 222 nm and 208 nm and an isodichroic point at 204 nm. The latter result shows that the spectrum can be deconvoluted in a common two-state helix-coil equilibrium as done by Gans et al. (1991). The results of $[\theta]_{222}$, and the calculated values of $\Delta\Delta G_t$, ΔG_m, $\Delta\Delta G_m$, and S are shown in Table 1. The error in measurement never exceeded 5%. As in the earlier series, the peptide with X = Ala (AAA) has the highest negative values of $[\theta]_{222}$, ΔG_m, and $\Delta\Delta G_m$, and the largest helicity (%), indicating that alanine has the greatest mid-helical α-helix propensity in this system (Lyu et al., 1990). The order of stabilization of α-helix formation by the 10 neutral side chains in the triply substituted peptides studied by Lyu et al. (1990) is A > L > M > Q > I > V > S > T > N > G, which differs from the present series. Except for alanine, the apparent helicities are greater in each case than those reported in the paper of Lyu et al. (1990), which indicates that chains with the $i \pm 4$, $i \pm 1$ alanine background have a greater

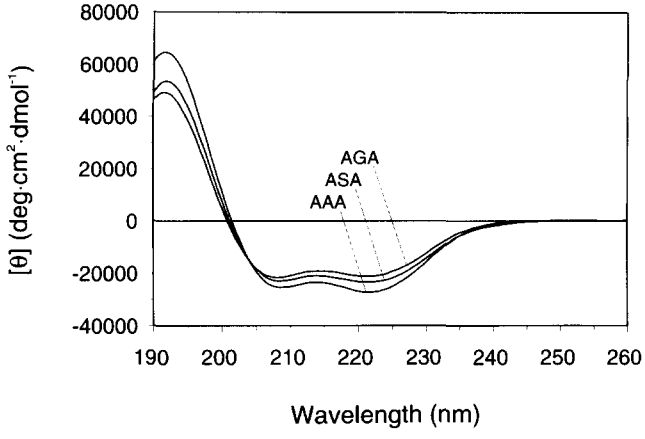

Fig. 1. CD spectra of AAA, AGA, and ASA in 10 mM KF, 1 mM phosphate pH 7 at 4 °C. Peptide concentration is 37 μm.

tendency to form α-helix than E or K residues with or without stabilizing salt bridges.

Temperature dependence of synthetic peptides

We chose the alanine-substituted and glycine-substituted peptides for temperature measurements because the former has the largest tendency (AAA) and the latter a very small tendency (AGA) to favor α-helix formation. The results are shown in Figure 2A. For both samples, the fractional helicity decreased by about two thirds (from 84.9% to 28.9% for AAA and from 67.8% to 17.3% for AGA) as the temperature increased from 4 °C to 75 °C. At any temperature in this range, the ratio of helicity (%) of the AGA peptide to that of AAA was between 0.57 and 0.76.

Concentration dependence

The concentration dependence of AAA was measured by CD using samples with peptide concentrations from 20 μM to 300 μM (Fig. 2B). The helicity (%) varied between 84.0% and 84.9%, indicating that concentration has no significant effect on the helicity. There are amino acid residues with a greater tendency to aggregate than alanine. Lyu et al. (1990) showed that a similar peptide with three leucines in the middle did not aggregate. The peptide employed in this work contains only one such site. Aggregation can therefore be excluded.

Effect of salt (NaCl) concentration on AAA

With increasing salt concentration, the helicity of AAA decreases significantly from 82.2% to 28% (Fig. 2C), showing competition of neutral salt of salt bridges between E and K, as has been reported in several studies (Scholtz et al., 1991c; Lyu et al., 1992a).

The distribution of α-helix in AAA determined by NMR

Proton chemical shifts for the peptide AAA were assigned following standard procedures, using TOCSY and NOESY experiments (Wüthrich, 1986). Figure 3 shows the fingerprint region of the TOCSY spectrum from which the intraresidue spin systems are identified. Tyr 1 and Ser 2 have AMX spin systems, respectively, and are easily determined. Seven lysines and seven glutamic acids are identified by their TOCSY spectrum. Only three A_3X spin systems are resolved for alanine residues; this means that the amide chemical shifts of alanines are degenerate. The pathway of

Table 1. *Helix content, free energy and helicity for amino acid substitution in 20 synthetic peptides*

Peptide	$-[\theta]_{222}$ (deg·cm²/dmol)	ΔG_m[a] (kcal/mol)	$\Delta\Delta G_m$[b] (kcal/mol)	$\Delta\Delta G_t$[c] (kcal/mol)	Helicity[d] F (%)	S[e]	S_x/S_a
Alanine (AAA)	27,200	−0.360	−1.050	−0.664	84.9	1.923	1.000
Glutamate (AEA)	26,400	−0.150	−0.840	−0.560	82.5	1.313	0.683
Leucine (ALA)	25,800	0.000	−0.690	−0.485	80.6	1.000	0.520
Isoleucine (AIA)	25,700	0.020	−0.670	−0.473	80.3	0.964	0.501
Arginine (ARA)	25,500	0.050	−0.640	−0.451	79.7	0.913	0.475
Methionine (AMA)	24,900	0.165	−0.525	−0.384	77.8	0.741	0.385
Histidine (AHA)	24,800	0.175	−0.515	−0.374	77.5	0.728	0.379
Lysine (AKA)	24,500	0.225	−0.465	−0.344	76.6	0.665	0.346
Valine (AVA)	23,500	0.370	−0.320	−0.242	73.4	0.511	0.266
Serine (ASA)	23,200	0.400	−0.290	−0.215	72.5	0.484	0.252
Aspartate (ADA)	23,100	0.420	−0.270	−0.206	72.2	0.466	0.242
Glutamine (AQA)	22,900	0.440	−0.250	−0.189	71.6	0.450	0.234
Asparagine (ANA)	22,500	0.490	−0.200	−0.151	70.3	0.411	0.214
Threonine (ATA)	22,200	0.525	−0.165	−0.126	69.4	0.385	0.200
Glycine (AGA)	20,700	0.690	0	0	64.7	0.286	0.149
Phenylalanine (AFA)	20,100	0.750	0.060	0.049	62.8	0.256	0.133
Tryptophan (AWA)	20,100	0.750	0.060	0.049	62.8	0.256	0.133
Cysteine (ACA)	19,600	0.810	0.120	0.087	61.3	0.230	0.120
Tyrosine (AYA)	11,900	1.585	0.895	0.669	37.2	0.056	0.029
Proline (APA)	8,330	2.200	1.510	0.979	26.0	0.018	0.009

[a] ΔG_m were obtained according to Gans et al. (1991).
[b] $\Delta\Delta G_m = \Delta G_{m,\mathrm{AXA}} - \Delta G_{m,\mathrm{AGA}}$. AXA denoted any peptides.
[c] $\Delta\Delta G_t = \Delta G_{t,\mathrm{AXA}} - \Delta G_{t,\mathrm{AGA}}$, where $\Delta G_{t,\mathrm{AXA}} = -RT\ln[F_x/(1 - F_x)]$. F_x denotes the helicity of any peptides.
[d] Helicity = $F = (-[\theta]_{222,\mathrm{AXA}}/32{,}000) \times 100\%$.
[e] The equilibrium constant, $S = \exp(-\Delta G_m/RT)$, $T = 277.15$ K.

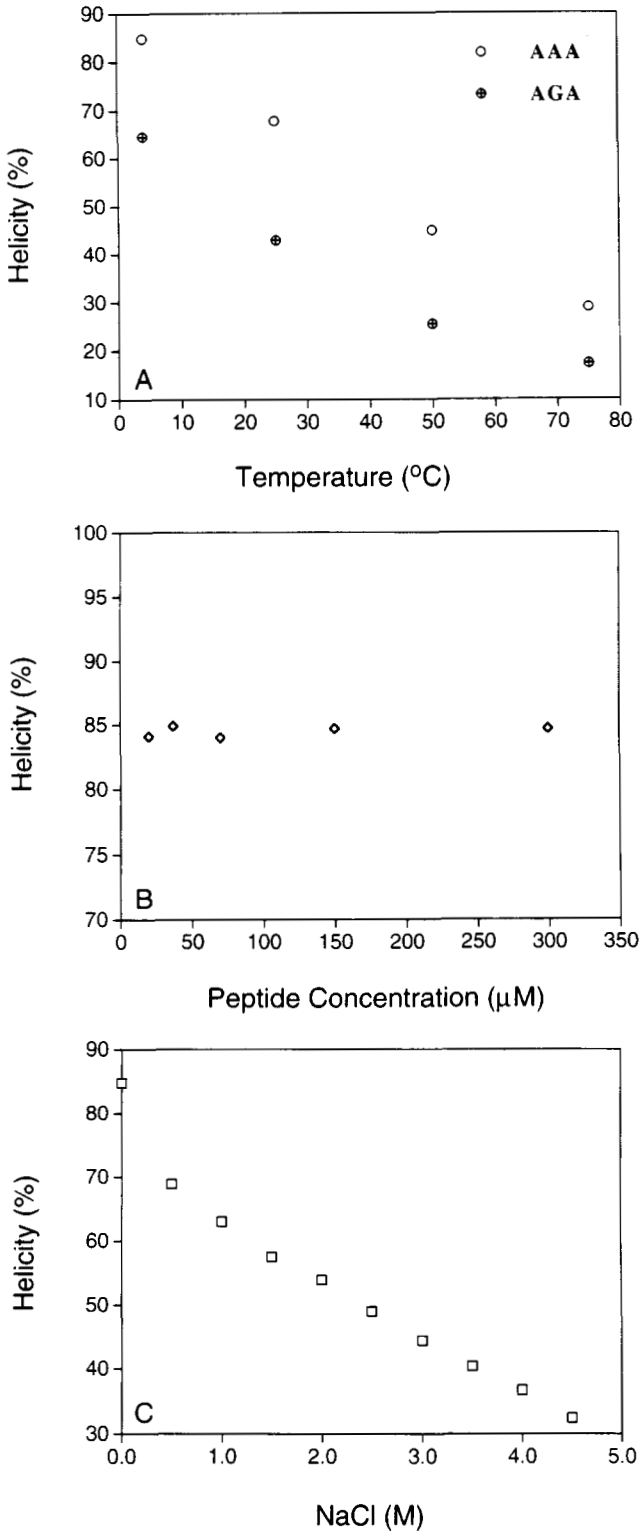

Fig. 2. **A:** Temperature dependence of AAA and AGA peptides. **B:** Concentration dependence of AAA peptide. **C:** Effect of ionic strength on AAA peptide.

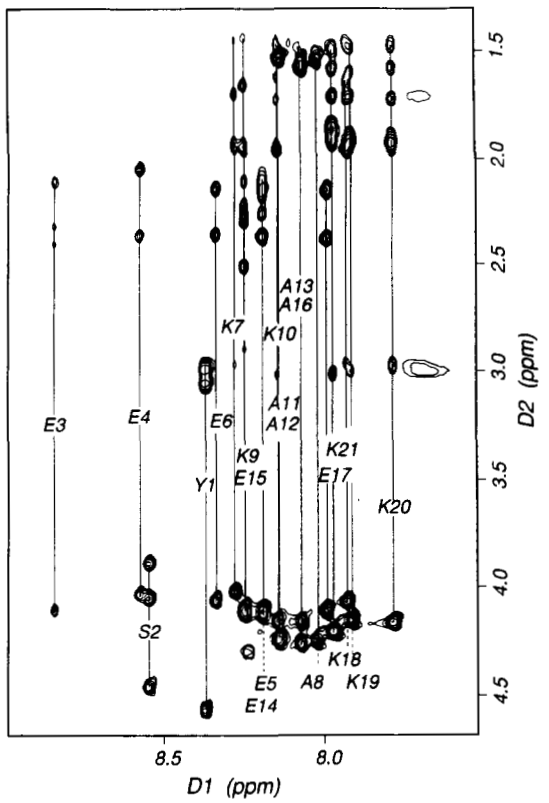

Fig. 3. Fingerprint region of the TOCSY spectrum of X = Ala peptide at ca. 4.0 mM in H_2O/D_2O (90%/10%) pH 5, 10 °C, with a 98-ms mixing time. Each intra residue spin system is connected with solid lines and assigned.

sequential assignment in this peptide is not straightforward due to the frequent occurrence of similar residues in the sequence. Despite this problem, the high helicity of AAA makes it feasible to trace sequential amide–amide connectivities in the NOESY spectrum. A clear $i - i + 1$ NH-NH pattern extending between residues 3 and 10 and 12 and 21 is shown in Figure 4. Amide chemical shift degeneracy obscures the NH-NH connectivities between residues 10 and 12. The NOESY spectra at other temperatures, together with ROESY spectra (data not shown), confirm the above assignments.

The presence of sequential amide–amide NOEs is a necessary but not sufficient criterion for helical structure. The fingerprint region of the NOESY spectrum is shown in Figure 5. A number of sequential αH-NH and βH-NH connectivities, in combination with medium-range $\alpha N(i,i+3)$ connectivities in the same region of the peptide, suggest that residues 3–21 form an α-helix. The sequential NH-NH NOE peaks are strongest near the center of molecule, while their intensity decays toward either terminus, consistent with fraying at the N and C termini. The intensities decay more abruptly for the C-terminus than for the N-terminus, consistent with the previous results on the EK series of peptides (Liff et al., 1991) showing that the distribution of helix is asymmetric in these models, possibly because of the presence in the chain of an SXXE N-terminal capping box (Harper & Rose, 1993; Zhou et al., 1994).

Discussion

The CD spectra of a homologous series of 20 peptides with single residue amino acid substitutions near the central position have been measured. In agreement with earlier results, the alanine sub-

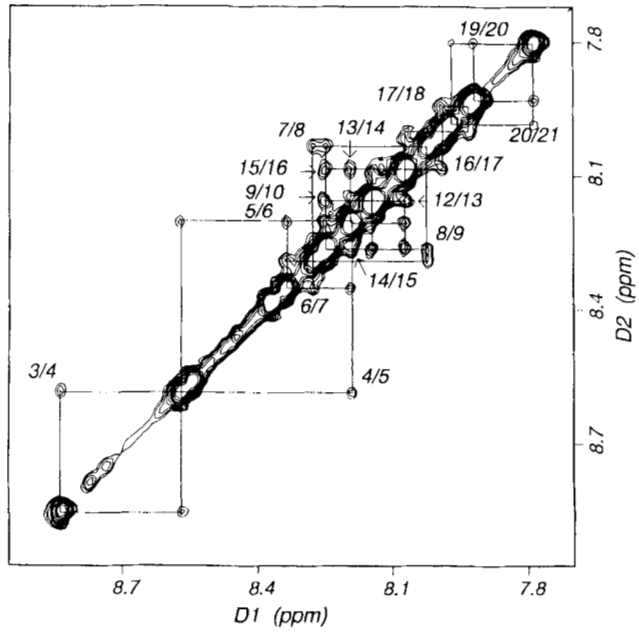

Fig. 4. Section of NOESY spectrum of X = Ala peptide at ca. 4.0 mM in H_2O/D_2O (90%/10%), pH 5, 10 °C, with a 100-ms mixing time indicating cross peaks between backbone amide protons. Sequential crosspeaks are indicated by the sequence number of residues.

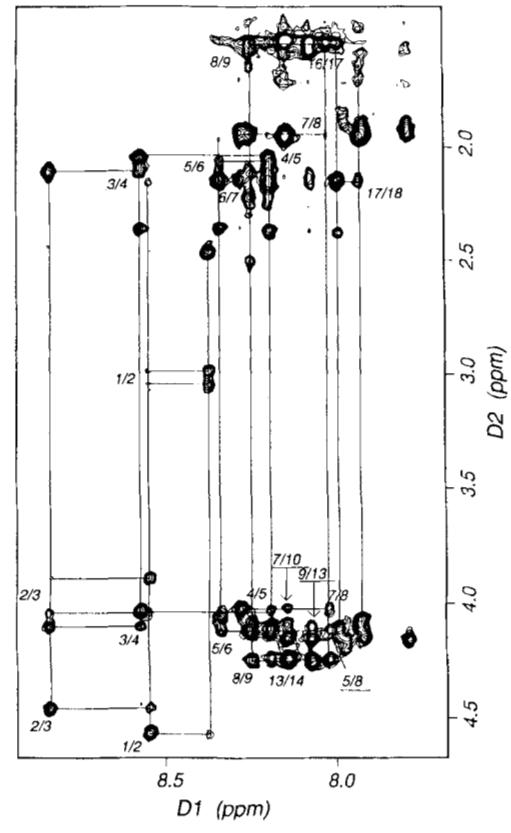

Fig. 5. Section of NOESY spectrum of X = Ala peptide at ca. 4.0 mM in H_2O/D_2O (90%/10%), pH 5, 10 °C, with a 100-ms mixing time indicating cross peaks between backbone amide protons and Cα and side-chain protons. Sequential and medium-range crosspeaks are indicated by the sequence number of residues.

stitution (X = A) shows maximal helicity (84.9%) and the Pro substitution (X = P) has lowest helicity (26.0%), consistent with the fact that the ring structure of proline entails distortion of the helix. In this series, glycine surprisingly does not have the minimum propensity among the true amino acids; Cys does. Peptides containing cysteine can form intermolecular bonds; however, the AXA peptide employed in this work is not dimerized. Adding DTT to the sample does not change the spectrum, and the reason for the low apparent helicity of Cys is not clear. For aromatic side chains, the apparent helicities must be estimated in a different way, because they contribute anomalously to the helical CD signal. This is especially true for tyrosine, and this amino acid will therefore be excluded from the discussion.

In the present series, only Ala and Glu stabilize helix relative to coil; all other side chains are either neutral or destabilizing, as reported by Baldwin's group for Ala-rich model peptides (Chakrabartty et al., 1994). For the aliphatic series of amino acid substitutions (AGA, AAA, AVA, ALA, and AIA), there seems to be no relation between bulkiness of a side chain and its helix-forming propensity. Studies of the helix-forming propensities of unnatural amino acids containing linear aliphatic side chains have shown that linear aliphatic side chains are helix stabilizing (Lyu et al., 1991a; Padmanabhan et al., 1991). On the other hand, replacements of alanine by residues with one, two, and three branches at C_β strongly decreases the helical content (Lyu et al., 1991a), so that restricting the rotamer population is an important factor (Creamer & Rose, 1992, 1994).

Salt bridges between E and K have a stabilizing contribution to the formation of α-helix (Marqusee et al., 1987; Merutka & Stellwagen, 1991; Lyu et al., 1992a; Huyghues-Despointes et al., 1993). In our peptides, there are potentially seven pairs of salt bridges between E and K side chains. Loss of a salt bridge should result in the destabilizing of α-helix, although studies of this effect in coiled-coils have been interpreted in terms of destabilizing coil. Figure 2C shows the result of salt titration for the peptide AAA. With the increase of salt concentration, the $-[\theta]_{222}$ decreased, consistent with competition for salt bridges between E and K. The curvilinear salt dependence of $-[\theta]_{222}$ on NaCl concentration is similar to that reported by Park et al. (1993). They used KCl and $CaCl_2$ to perform salt titrations and found that there was a linear dependence of $-[\theta]_{222}$ on salt concentrations when salt concentrations are above 0.5 M.

With increasing temperature, helicity decreases regardless of sequence. The effect holds for the most favorable helix-forming sequence (AAA) and one of the least favorable (AGA). Therefore, enthalpy drives helix formation in synthetic short peptides as in longer peptide chains (Scholtz et al., 1991c). In our experimental conditions, concentration of peptide has no significant influence on the formation of α-helix (Fig. 2B), which means that the helical structure is intramolecular and not a consequence of intermolecular association of α-helices (Lyu et al., 1990).

An important issue that needs to be addressed is how the free energies derived here correlate with values obtained from other experimental systems. A comparison between the experimental obtained ΔG values can be found in Table 2. Table 3 shows the correlation factors between the data obtained with AXA and other methods, experimental as well as predicted by algorithms. Data published in this paper show a good correlation with the data obtained by Chakrabartty et al. (1994). Without Tyr, the correlation

Table 2. *Free energy of α-helix formation; data are represented as ΔG_m (kcal/mol)*

Amino acid	This work	Lyu[a]	Baldwin[b]	Scheraga[c]
Alanine	−0.360	−0.42	−0.26	−0.04
Glutamate	−0.150		0.43	0.02
Leucine	0.000	−0.15	0.02	0.08
Isoleucine	0.020	0.08	0.45	−0.08
Arginine	0.050		−0.05	−0.02
Methionine	0.165	−0.10	0.25	−0.11
Histidine	0.175		0.53	0.22
Lysine	0.225		0.11	0.04
Valine	0.370	0.13	0.80	0.03
Serine	0.400	0.19	0.53	0.16
Aspartate	0.420		0.64	0.23
Glutamine	0.440	−0.01	0.31	0.01
Asparagine	0.490	0.27	0.64	0.15
Threonine	0.525	0.24	1.07	0.12
Glycine	0.690	0.47	1.62	0.31
Phenylalanine	0.750		0.67	−0.05
Tryptophan	0.750		0.59	−0.06
Cysteine	0.810		0.57	0.01
Tyrosine	1.585		0.43	−0.01
Proline	2.200		~4	1.01

[a]Lyu et al., 1990.
[b]Chakrabartty et al., 1994.
[c]Wojcik et al., 1990.

coefficient for this comparison is 0.90 (Table 3). Tyrosine has been omitted because it is difficult to evaluate the contribution of the aromatic side chain of tyrosine to helicity by means of CD. There is an offset however, which may have to do with different values assigned to specific parameters such as salt bridges. A question that is relevant here is the contribution of $i \pm 3$ interactions. These may arise if X is a charged residue (D, E, K, R). These residues are able to form salt bridges with either lysine or glutamic acid, thereby stabilizing the helix. We have therefore looked at the correlation with and without the charged residues. Omitting the charged residues from the comparison with Chakrabartty et al. (1994) proves to have little influence on the correlation factor ($R = 0.91$). Thus, $i \pm 3$ salt bridges do not seem to be a major contributor to helix stability. Correlation with values obtained using a Glu-Lys block oligomer host peptide is also significant ($R = 0.89$) (Lyu et al., 1990). However, there are differences in the relative contributions of Ile and Gln, and deconvolution indicates that neither Leu nor Met stabilizes helix. Despite evident differences in values from one model to another, the overall correlation remains good, indicating that major features of helix formation have been captured that do not depend on context. There is no correlation between the values measured in the present experiments and the scale obtained from the host-guest polymers of Scheraga's group (Wojcik et al., 1990) or with the Ala and Gly values published by Kemp's group (Groebke et al., 1996). These workers argue that the intrinsic propensity of Ala is low, and that it becomes stabilizing in the presence of charged side chains. The presence of several long side chains in our model precludes a direct test of their hypothesis.

There has been interest in the reason(s) for the discrepancies between the propensity values obtained from different model systems. One suggestion is that effects of context might be more pervasive than originally supposed (Groebke et al., 1996). Scheraga's group pioneered development of a propensity scale using a bulky synthetic host side chain, which demonstrably can influence substitutions at sites several positions away (Padmanabhan et al., 1994). Baldwin's and Stellwagen's groups, on the other hand, have argued that the minimal helix forming side chain of alanine is intrinsically a better reference, although even alanine is capable of promoting interactions from neighboring side chains (Groebke et al., 1996). We show here that deconvolution of CD data on the series of peptides of this study yields propensity values that correlate reasonably with scales obtained by other laboratories using Ala-rich model peptides (Chakrabartty et al., 1994). This does not allow comprehensive analysis of next neighbor and longer range interactions between side chains. So far, despite CD spectral values for many peptides (Muñoz & Serrano, 1994), this complete set of interactions has not been rigorously determined.

Over the years, considerable effort has also been devoted to developing computer algorithms that can predict helical secondary structure for peptides starting from the primary amino acid sequence alone (Wako et al., 1983; Doig et al., 1994; Muñoz & Serrano, 1994; Shalongo & Stellwagen, 1995). We have tested some of these by predicting the helix content of the 20 amino acids substitutions of this study and comparing them to the experimental values. The results are shown in Table 3. Figure 6 shows the

Table 3. *Correlation factors of comparison of AXA with other methods*

	Amino acids excluded from comparison								
Method	None	Tyr	Tyr, Trp	Tyr, Trp, Phe	Tyr, Glu, Lys, Asp, Arg	Tyr, His, Gly	Tyr, Phe, Trp, His, Gly	Tyr, Phe, Asp, Gly	Tyr, Phe, Asp, Gly, Trp
Baldwin[a] (experimental; ΔG)	0.78	0.90	0.91	0.93	0.91				
Baldwin[b] (predicted; F (%))	0.55	0.70	0.72	0.72	0.68	0.86	0.87		
Stellwagen[c] (predicted; F (%))	0.62	0.75	0.75	0.78	0.76				
Serrano[d] (predicted; F (%))	0.57	0.8	0.85	0.91	0.84			0.9	0.96

[a]Chakrabartty et al., 1994.
[b]Doig et al., 1994.
[c]Shalongo & Stellwagen, 1995.
[d]Muñoz & Serrano, 1994.

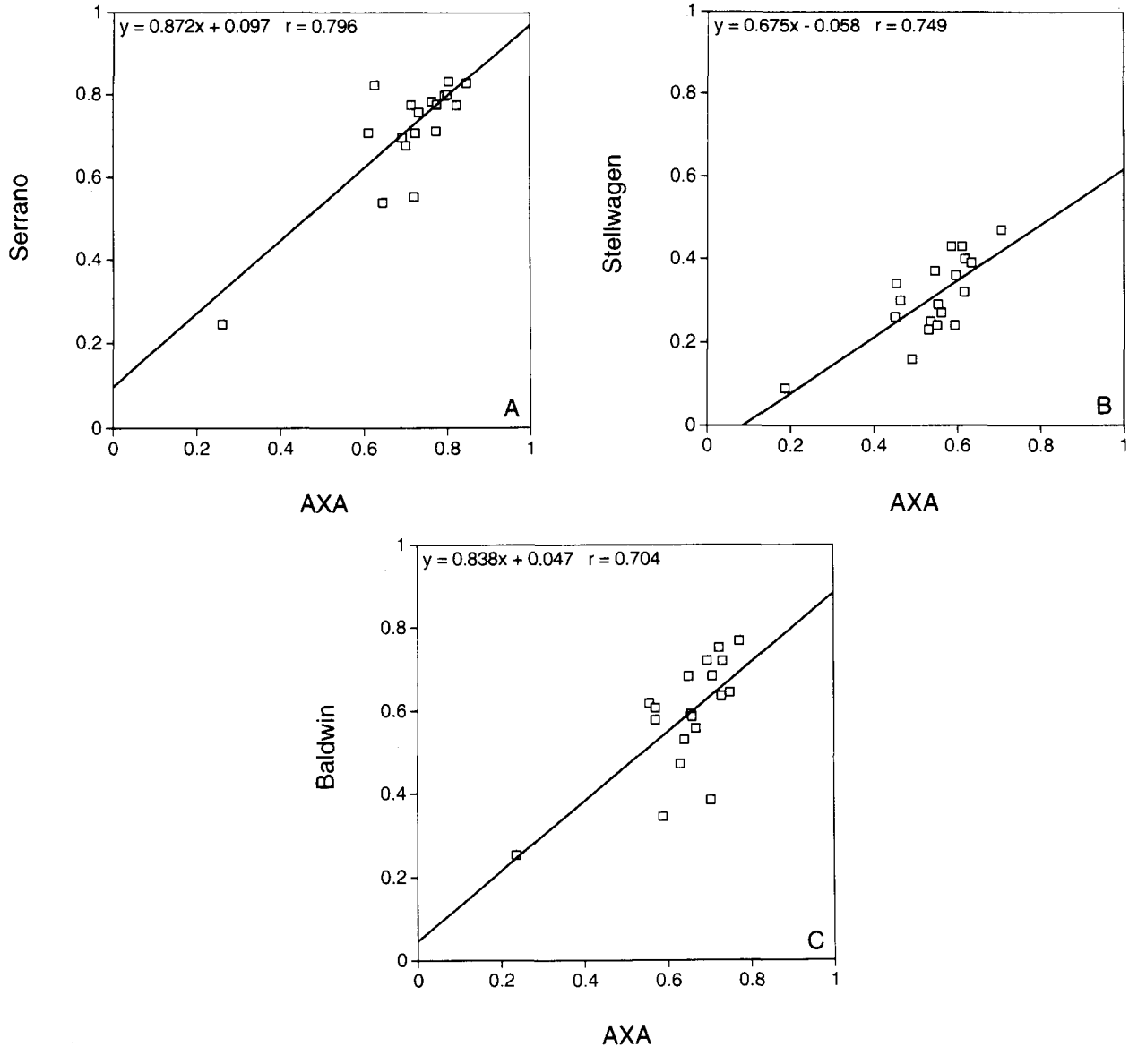

Fig. 6. Correlation of calculated helicity against experimental helicity values. **A:** Prediction of helicity by AGADIR versus experimental values (Muñoz & Serrano, 1994). **B:** Prediction of helicity by Stellwagen versus experimental values (Shalongo & Stellwagen 1995). **C:** Prediction of helicity by Baldwin versus experimental values (Doig et al., 1994).

correlation between AXA and the computer algorithms with tyrosine excluded. It can be seen that the level of correlation of predicted vs. experimental helicity in each case is good ($0.7 < R < 0.8$), but far from perfect. In general, the algorithms deviate most in the cases of peptides containing Pro and Gly, both strong helix destabilizing side chains. In some cases, the correlation can be improved by arbitrarily neglecting certain side chains, in which case the R values can increase to 0.9 or higher. The best correlation between experimental and predicted helicity ($R = 0.8$) was found by AGADIR (Muñoz & Serrano, 1994). The correlation can be improved by neglecting Trp and Phe; in this case, the correlation factor goes up to 0.91. The reason that the aromatic side chains deviate from the experimental values may have to do with unpredictable contributions from the side chains to the helix CD signal. The correlation factor goes down to 0.75 when the experimental values are compared with values generated by the program from Stellwagen (Shalongo & Stellwagen, 1995). The lack of correlation seems to exist for all amino acids. Removal of a few amino acids from the scale will not improve the correlation significantly. The last comparison is made with the program from Baldwin's group ($R = 0.70$; Doig et al., 1994). The correlation improves to 0.86 if Gly and His are excluded. The helix content in peptides with internal Gly or Pro side chains is likely to be difficult to predict satisfactorily with statistical mechanical programs, because both have conformational properties that alter the normal two-state approximation for a given residue. The helix propensity of glycine in the peptides studied by Baldwin's group is lower than in other peptides studied, for reasons we do not yet understand but probably have a physical explanation. NMR studies of Gly-substituted peptides show that a single internal Gly side chain can diminish the

helicity of neighboring side chains two or more residues away in both N and C directions (Lyu et al., 1991b). The His propensity determined by Baldwin's group is also lower than predictions from other models and the experimental values. This might reflect some side-chain–side-chain interaction present in these peptides.

Thus, neighboring side chains indeed exert an effect on the apparent helix content of a side chain at a site i in α-helix. The influence extends to sites at $i \pm 4$ in either direction, and possibly even further in special cases. Given this situation, the matrix of side-chain–side-chain interaction free energies or other effects needed to allow quantitative prediction of helical structure in short sequences is large, and not complete in any database today. What is encouraging is that the incorrect or missing effects do not appear to involve major sources of helix stabilization, because otherwise R values around 0.8 would not be found. Instead, the process of parsing helical free energies into intrinsic propensities and side-chain–side-chain interactions seems to be convergent, and ultimately should allow quantitative prediction for any sequence (Kallenbach et al., 1996). This differs from the situation with respect to β-sheet structure, in which the side-chain–side-chain interaction terms appear to be larger with respect to intrinsic propensities (Minor & Kim, 1994; Smith & Regan, 1995). On the other hand, the ability of theoretical calculations based on approximate force fields to reproduce the small free energy differences seen in single helical models seems limited at present.

Materials and methods

Peptide synthesis and purification

Peptides were synthesized on a Milligen/Biosearch 9600 Peptide Synthesizer by solid-phase peptide synthesis. Fmoc protection was used for α-amino groups and BOP/HOBT method was used for coupling. Following synthesis, peptides were cleaved and deprotected by treatment with TFA reagent (90% TFA, 5% thioanisole, 3% ethanedithiol, and 2% anisole) at room temperature for 4 h and precipitated with cold ether. Crude peptides were purified by RP-HPLC on an Econosil C18-10 μ column (250 × 22.5 mm) monitored at 275 nm with a gradient of 15–25% acetonitrile in 0.1% trifluoroacetic acid. The purified peptides were desalted on a Sephadex G-10 column and checked for purity on an analytic Delta Pak C18-15 μ HPLC column (300 × 3.9 mm) and EI-MS (Lyu et al., 1991b).

CD measurements

CD spectra were recorded on an Aviv DS60 spectropolarimeter equipped with an HP Model 89100A temperature controller. The wavelength was calibrated with (+)-10-camphorsulfonic acid according to the method of Chen and Yang (1977). All measurements were performed in 10 mM KF, 1 mM phosphate buffer pH 7 at 4 °C with a peptide concentration of ca. 37 μM (except for the experiments on peptide concentration dependence) with 1-mm path length cells. Three scans were averaged with a 0.5-nm step in each case. Six to 12 measurements were performed for a peptide sample. CD spectral data were deconvoluted according to the procedure described by Gans et al. (1991).

NMR spectroscopy

Sample preparation

The NMR sample was prepared by dissolving lyophilized peptide AAA in distilled H_2O to a peptide concentration of ca. 4.0 mM. 10% D_2O was used for locking and 0.25 mM of the sodium salt of 3-(trimethylsilyl)-[3,3,2,2,-^{2}H]propionic acid was served as an internal chemical shift reference. We used concentrated sodium hydroxide and hydrogen chloride solution to adjust the pH value until the pH meter reading was 5.0. Another sample in pure D_2O with pH 5.0 (uncorrected for D_2O effect) was prepared by an identical procedure to give a peptide concentration of approximately 4.0 mM.

NMR experiments

All ^{1}H NMR spectra were determined on a Varian UNITY 500 spectrometer. The States method (States et al., 1982) was used to obtain phase-sensitive TOCSY (Bax & Davis, 1985b) with mixing times of 98 ms, NOESY (Jeener et al., 1979; Kumar et al., 1980; Wider et al., 1984) with mixing times of 200 ms, and ROESY (Bax & Davis, 1985a) spectra with a mixing time of 100 ms at both 10 °C and 25 °C. The spectral width in each dimension was 5,000 Hz. The transmitter frequency was placed on the H_2O resonance. Solvent signal suppression was achieved by applying continuous-wave presaturation during the relaxation delay (typically 1.2 s), $t1$ evolution time, and also during the mixing period in the NOESY and ROESY experiments. Elimination of baseline artifacts was achieved by optimizing the "receiver-on" time and the "in-phase" time (Hoult et al., 1983; Davis, 1989). Each 2D data set contained 512 FIDs with 2K complex data points each, obtained by collecting 32 added free induction decays after four dummy scans. Spectra were Fourier transformed in both $t1$ and $t2$ dimensions after apodization with a skewed sine bell function, typically with a 60° phase shift and skew of 0.6. Zero filling was done in the $t1$ dimension to obtain a final matrix of 1,024 × 1,024 real points. In all spectra, the first row of the ($t1$, $\omega2$) matrix was divided by 2 before the second Fourier transform to suppress ridges in the $t1$ dimension (Otting et al., 1986). All NMR data were processed on a Silicon Graphics Crimson workstation using the FELIX 2.3 software (Biosym Technologies Inc.).

Computer algorithms

Several programs were used to evaluate ΔG_m values, the free energy for helix stabilization by a given side chain, or alternatively, to predict the fraction of helix for the series. Dr. Carol Rohl used a version of a Lifson–Roig helix-coil transition model, parameterized as described, to predict the results (Chakrabartty et al., 1994). Stellwagen applied a similar model, parameterized according to Shalongo and Stellwagen (1995). A version of the helix prediction program of Muñoz and Serrano was used to test this model (Muñoz & Serrano, 1994).

Acknowledgments

We thank Drs. E. Stellwagen and R. Baldwin for their cooperation in performing calculations of helix content prior to seeing the data. This work was supported by grant GM 40746 from the NIH and a grant from the Human Frontiers of Science Program.

References

Bax A, Davis G. 1985a. Practical aspects of two-dimensional transverse NOE spectroscopy. *J Magn Reson 63*:207–213.

Bax A, Davis G. 1985b. MLEV-17-based two dimensional homonuclear magnetization transfer spectroscopy. *J Magn Reson 65*:355–360.

Blaber M, Zhang XJ, Matthews BW. 1993a. Structure basis of amino acid α-helix propensity. *Science 260*:1637–1640.

Blaber M, Zhang XJ, Matthews BW. 1993b. Alpha helix propensity of amino acids. *Science 262*:917–918.
Chakrabartty A, Kortemme T, Baldwin RL. 1994. Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interaction. *Protein Sci 3*:843–852.
Chakrabartty A, Schellman JA, Baldwin RL. 1991. Large differences in the helix propensities of alanine and glycine. *Nature 351*:586–588.
Chen GC, Yang JT. 1977. Two-point calibration of circular dichrometer with d-10-camphorsulfonic acid. *Anal Lett 10*:1195–1207.
Chou PY, Fasman GD. 1978. Prediction of the secondary structure of proteins from their amino acid sequence. *Adv Enzymol 47*:45–148.
Creamer TP, Rose GD. 1992. Side chain entropy opposes a helix formation but rationalizes experimentally determined helix forming propensities. *Proc Natl Acad Sci USA 89*:5937–5941.
Creamer TP, Rose GD. 1994. α-Helix forming propensities in peptides and proteins. *Proteins Struct Funct Genet 19*:85–97.
Davis G. 1989. Elimination of baseline distortion and minimization of artifacts from phased 2D NMR spectra. *J Magn Reson 81*:603–607.
Doig AJ, Chakrabartty A, Klinger TM, Baldwin RL. 1994. Determination of free energies of N capping in alpha helices by modification of the Lifson–Roig helix-coil theory to include N and C capping. *Biochemistry 33*:3396–3403.
Fasman GD. 1989. *Prediction of protein structure and the principles of protein conformation*. New York: Plenum.
Finkelstein AV, Badretdinov AYA, Ptitsyn OB. 1991. Physical reasons for secondary structure stability: α-Helices in short peptides. *Proteins Struct Funct Genet 10*:287–299.
Gans PJ, Lyu PC, Manning MC, Woody RW, Kallenbach NR. 1991. The helix-coil transition in heterogeneous peptides with specific side chain interactions: Theory and comparison with circular dichroism spectral data. *Biopolymers 31*:1605–1614.
Groebke K, Renold P, Tsang KY, Allen TJ, McClure KF, Kemp DS. 1996. Template-nucleated alanine-lysine helices are stabilized by position-dependent interactions between the lysine side chain and the helix barrel. *Proc Natl Acad Sci USA 93*:4025–4029.
Harper E, Rose GD. 1993. Helix stop signal in proteins and peptides: The capping box. *Biochemistry 32*:7605–7609.
Hoult DI, Chen CN, Eden H. 1983. Elimination of baseline artifacts in spectra and their integrals. *J Magn Reson 51*:110–117.
Huyghues-Despointes BMP, Scholtz JM, Baldwin RL. 1993. Helical peptides with three pairs of Asp-Arg and Glu-Arg residues in different orientations and spacings. *Protein Sci 2*:80–85.
Ingwall RT, Scheraga HA, Lotan N, Berger A, Katchalski E. 1968. Conformational studies of Poly-L-Alanine in water. *Biopolymers 6*:331–368.
Jeener J, Meier BH, Bachmann P, Ernst RR. 1979. Investigation of exchange processes by two-dimensional NMR spectra. *J Chem Phys 71*:4546–4553.
Kallenbach NR, Lyu PC, Zhou HX. 1996. CD spectroscopy and the helix-coil transition in peptides and polypeptides. In Fasman GD, ed. *Circular dichroism and the conformational analysis of biomolecules*. New York: Plenum Press. pp 201–259.
Kim PS, Bierzinski A, Baldwin RL. 1982. A competing salt-bridge suppresses helix formation by the isolated C-peptide carboxylate of ribonuclease A. *J Mol Biol 162*:187–189.
Kumar A, Ernst RR, Wüthrich K. 1980. A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton–proton cross-relaxation networks in biological macromolecules. *Biochem Biophys Res Commun 95*:1–6.
Levitt M. 1978. Conformational preferences of amino acids in globular proteins. *Biochemistry 17*:4277–4285.
Liff MI, Lyu PC, Kallenbach NR. 1991. Analysis of asymmetry in the distribution of helical residues in peptides by ^{1}H nuclear magnetic resonance. *J Am Chem Soc 113*:1014–1019.
Lyu PC, Gans PJ, Kallenbach NR. 1992a. Energetic contribution of solvent exposed ion pairs to alpha-helix structure. *J Mol Biol 223*:343–350.
Lyu PC, Liff MI, Marky LA, Kallenbach NR. 1990. Side chain contributions to the stability of alpha-helix structure in peptides. *Science 250*:669–673.
Lyu PC, Marky LA, Kallenbach NR. 1989. The role of ion pairs in α-helix stability: Two new designed helical peptides. *J Am Chem Soc 111*:2733–2734.
Lyu PC, Sherman JC, Chen A, Kallenbach NR. 1991a. Alpha helix stabilization by natural and unnatural amino acids with alkyl side chains. *Proc Natl Acad Sci USA 88*:5317–5320.
Lyu PC, Wang PC, Liff MI, Kallenbach NR. 1991b. Local effect of glycine substitution in a model helical peptide. *J Am Chem Soc 113*:3568–3572.
Lyu PC, Wemmer DE. 1993. Capping interactions in isolated α-helices: Position-dependent substitution effects and structure of a serine-capped peptide helix. *Biochemistry 32*:421–425.
Lyu PC, Zhou H, Jelveh N, Wemmer DE, Kallenbach NR. 1992b. Position-dependent stabilizing effects in α-helices: N-terminal capping in synthetic model peptides. *J Am Chem Soc 114*:6560–6562.
Marqusee S, Baldwin RL. 1987. Helix stabilization by $Glu^- \cdots Lys^+$ salt bridge in short peptides of de novo design. *Proc Natl Acad Sci USA 84*:8898–8902.
Marqusee S, Robbins V, Baldwin RL. 1989. Unusually stable helix formation in short alanine-based peptides. *Proc Natl Acad Sci USA 86*:5286–5290.
Maxfield FR, Scheraga HA. 1975. The effect of neighboring charges on the helix forming ability of charged amino acids in proteins. *Macromolecules 8*:491–493.
Merutka G, Stellwagen E. 1991. Effect of amino acid ion pairs on peptide helicity. *Biochemistry 30*:1591–1594.
Minor DL, Kim PS. 1994. Measurement of the β-sheet-forming propensities of amino acids. *Nature 367*:660–663.
Muñoz V, Serrano L. 1994. Elucidating the folding problem of helical peptides using empirical parameters. *Nature Struct Biol 1*:399–400.
O'Neil KT, DeGrado W. 1990. A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science 250*:646–651.
Otting G, Widmer H, Wagner G, Wüthrich K. 1986. Origin of t1 and t2 ridges in 2D NMR spectra and procedures for suppression. *J Magn Reson 66*:187–193.
Padmanabhan S, Baldwin RL. 1991. Straight non-polar amino acids are good helix-formers in water. *J Mol Biol 219*:135–137.
Padmanabhan S, Marqusee S, Ridgeway T, Lau TM, Baldwin RL. 1990. Relative helix-forming tendencies of non-polar amino acids. *Nature 344*:268–270.
Padmanabhan S, York EJ, Gera L, Stewart JM, Baldwin RL. 1994. Helix-forming tendencies of amino acids in short (hydroxybutyl)-L-glutamine peptides: An evaluation of the contradictory results from host-guest studies and short alanine based peptides. *Biochemistry 33*:8604–8609.
Park SH, Shalongo W, Stellwagen E. 1993. Residue helix parameters obtained from dichroic analysis of peptides of defined sequence. *Biochemistry 32*:7048–7053.
Ptitsyn OB. 1972. Thermodynamic parameters of helix-coil transition in polypeptide chains. *Pure Appl Chem 31*:227–244.
Scholtz JM, Baldwin RL. 1992. The mechanism of α-helix formation by peptides. *Annu Rev Biophys Struct 21*:96–118.
Scholtz JM, Marqusee S, Baldwin RL, York EJ, Stewart JM, Santoro M, Bolen DW. 1991a. Calorimetric determination of the enthalpy charge for the α-helix to coil transition of an alanine peptide in water. *Proc Natl Acad Sci USA 88*:2854–2858.
Scholtz JM, Qian H, York EJ, Stewart JM, Baldwin RL. 1991b. Parameters of helix-coil transition theory for alanine-based peptides of varying chain length in water. *Biopolymers 31*:1463–1470.
Scholtz JM, York EJ, Stewart JM, Baldwin RL. 1991c. A natural, water-soluble α-helical peptide: The effect of ionic strength on the helix-coil equilibrium. *J Am Chem Soc 113*:5102–5104.
Shalongo W, Stellwagen E. 1995. Incorporation of pairwise interactions into the Lifson-Roig model for helix prediction. *Protein Sci 4*:1161–1165.
Shoemaker KR, Kim PS, York EJ, Stewart JM, Baldwin RL. 1987. Test of the helix dipole model for stabilization of α-helices. *Nature 326*:563–567.
Smith CK, Regan L. 1995. Guidelines for protein design: The energetics of β-sheet side chain interactions. *Science 270*:960–962.
States DJ, Haberkorn RA, Ruben DJ. 1982. A two-dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants. *J Magn Reson 48*:286–292.
Vila J, Ewilliams RL, Grant JA, Wojcik J, Scheraga HA. 1992. The intrinsic helix-forming tendency of L-alanine. *Proc Natl Acad Sci USA 89*:7821–7825.
Wako H, Saito N, Scheraga HA. 1983. Statistical mechanical treatment of α-helices and extended structures in proteins with inclusion of short- and medium-range interactions. *J Protein Chem 2*:221–249.
Wider G, Macura S, Kumar A, Ernst RR, Wüthrich K. 1984. Homonuclear two-dimensional ^{1}H NMR of proteins. Experimental approaches. *J Magn Reson 56*:207–234.
Wojcik J, Altmann KH, Scheraga HA. 1990. Helix-coil stability constants for the naturally occurring amino acids in water. XXIV half-cysteine parameters from random poly(hydroxypropylglutamine-co-S-methylthio-L-cysteine). *Biopolymers 30*:121–134.
Woody RW. 1985. Circular dichroism of peptides. In: Udenfriend S, Meienhofer J, Hruby JR, eds. *The peptides, vol 7*. New York: Academic Press. pp 15–114.
Wüthrich K. 1986. *NMR of protein and nucleic acids*. New York: Wiley.
Yang JT, Wu CSC, Martinez HM. 1986. Calculation of protein conformation from circular dichroism. *Methods Enzymol 130*:208–269.
Zhou HX, Lyu PC, Wemmer DE, Kallenbach NR. 1994. Alpha helix initiation in synthetic model peptides by a capping box. *Proteins Struct Funct Genet 18*:1–7.