Published in final edited form as:
J Am Chem Soc. 2005 December 7; 127(48): 16961–16968.

Energetic Characterization of Short Helical Polyalanine Peptides in Water: Analysis of $^{13}C{=}O$ Chemical Shift Data

Robert J. Kennedy, **Sharon M. Walker**, and **Daniel S. Kemp***
Contribution from the Department of Chemistry, Room 18-296, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Abstract

Measured at 2 °C in water, NMR chemical shifts of $^{13}C{=}O$ labeled central alanine residues of peptides W-Lys$_5$-^{t}L$_3$-Ala$_n$-^{t}L$_3$-Lys$_5$NH$_2$, n = 9, 11, 13, 15, 19 and W-Lys$_5$-^{t}L$_3$-a-Ala$_n$-A-Inp-^{t}L$_2$-Lys$_5$NH$_2$ (a = D-Ala; ^{t}L = *tert*-leucine; Inp = 4-carboxypiperidine) are used to assign $j_L{}^t$ and $c_L{}^t$, the N- and C-terminal ^{t}L capping parameters and length-dependent values for $w_{Ala}(n)$, the alanine helical propensity for Ala$_n$ peptides. These parameters allow Lifson–Roig characterization of the stabilities of Ala$_n$ helices in water. To facilitate chemical shift characterization, different $^{13}C/^{12}C$ ratios are incorporated into specific Ala sites to code up to six residue sites per peptide. Large left/right chemical shift anisotropies are intrinsic to helical polyalanines, and a correcting L–R-based model is introduced. Capping parameters $j_L{}^t = c_L{}^t$ lie in the range of 0.3 to 0.5; the ^{t}L residues are thus moderately helix-destabilizing. For helical conformations of lengths shorter than eight residues, assigned values for w_{Ala} approach 1.0 but increase monotonically with length to a value of 1.59 for $w_{Ala}(19)$.

Introduction

The stabilities of helical conformations of proteins[1] and peptides[2] depend on the structures of their N- and C-terminal caps, which provide the first and last H-bonding sites within a fully helical conformation. Reported values for parameters j (N-terminal) and c (C-terminal) that reflect helix-capping efficiency vary substantially.[2] The constructs **1** allow characterization of polyalanine helicities. In these, N- and C-^{t}L (*tert*-leucine) caps link a helical Ala$_n$ core to spaced, solubilizing terminal regions.[2b,3] Values for jt$_L$ and ct$_L$ are essentials for rigorous analysis of all construct helicity data. Here we report and analyze experimental $^{13}C{=}O$ NMR chemical shifts for Ala$_n$ peptides of medium length, measured in water at 2 °C, and from them we assign capping parameters and corresponding Ala$_n$ helicities. Jointly, these define the formation energetics for Ala$_n$ helices, which are helicity archetypes.

For short Ala$_n$ peptides, we have previously assigned a low value for the alanine helical propensity w_{Ala},[3] which is consistent with early reports.[4] For longer Ala$_n$ peptides, we have proposed that w_{Ala} increases to values that are consistent with later reports.[5] Both assignments rely on preliminary evidence[6] that ^{t}L caps are moderate helix destabilizers. There is an alternative interpretation: ^{t}L caps strongly destabilize Ala$_n$ helices, and our assignments of low w_{Ala} values and their length dependence are spurious. We now resolve this issue.

E-mail: Kemp@mit.edu.

Supporting Information Available: Typical ^{13}C labeling examples, peptide MS, characterization tables, and an example of the FH$_i$ partitioning used to model data in Figures 8 and 9. This material is available free of charge via the Internet at http://pubs.acs.org.

Capping Parameters and Caps that Terminate Helices; the Series ^{t}L-An-^{t}L and a-A$_n$-A; Deductions from CD Data

$$W-Lys_m-X-Ala_n-Y-Lys_mNH_2 \quad \mathbf{1}$$

$$X = Inp_2\ {}^tLeu \text{ and } Y = {}^tLeu-Inp_2,$$

$$\text{or } X = Inp_2 ala \text{ and } Y = Ala-Inp_2,\ \ ala \equiv D-Ala$$

The standard Lifson–Roig energetic algorithm reflects the helicity-controlling role of *j* and *c*.[2a] A completely helical N-and C-capped ^{t}L-Ala$_n$-^{t}L peptide conformation is assigned a weight $j_L{}^t\upsilon^2w^nc_L{}^t$, and a second conformation that contains no helical residues is assigned a weight 1.0. The corresponding helical state sum ss is $(1 + j_L{}^t\upsilon^2w^nc_L{}^t)$, and the fraction FH of helical α-carbons within the Ala$_n$ region is $j_L{}^t\upsilon^2w^nc_L{}^t$/ss.[7] Here, υ is the classical L–R initiation parameter,[8] usually assigned a value of ca. 0.05,[9] and *w* is the helical propensity, the tendency of a single residue to join a preexisting helical conformation to which it is linked.

In this all-or-none model, FH is clearly zero if any one of the length-independent parameters *j*, *c*, or *w* is zero. Values for $j_L{}^t$ and $c_L{}^t$ are defined relative to values for j_{Ala} and c_{Ala}, set by definition equal to 1.0.[2a] Structural factors influence *j* or *c*. The side chains of capping residues may contain charges that stabilize the helix or H-bonding sites that stabilize the terminal helical amide functions that lack a full complement of intrahelical H-bonds.[2c,d,10] The caps N-acetyl and N-formyl[2c] are also helix stabilizers; thus, replacing an N-terminal alanine by an N-acetyl function increases the FH values of helically disposed, medium-sized heteropeptides by ca. 70%. This corresponds to a j_{Ac} of ca. 10.

In addition to their Ala$_n$ cores and ^{t}L residues, which terminate helices, constructs **1** contain other essential elements. The N-and C-terminal K$_m$ sequences ensure aggregation-free core solubilization, and for pH < 10, charge-repelled residues in these regions adopt extended conformations. Lys-induced perturbations of Ala$_n$ helicity are minimized by rigid Inp$_2$ spacers.

What properties allow a ^{t}L cap to both H-bond to the Ala$_n$ core and terminate its helix? As an N-cap, ^{t}L provides the N-terminal amide oxygen that H-bonds to the amide NH at site four of an α-helix, Figure 1a. As a C-cap, it provides the C-terminal NH that H-bonds to the carbonyl of the residue at site $(n-3)$, Figure 1b. Formation of these H-bonds places little or no restriction on backbone φ and ψ angles of either ^{t}L. These can form second intrahelical H-bonds, join the helical structure, and extend the helix if their helical φ, ψ space is energetically accessible. A helix-terminating residue must be a strong helix breaker, and ^{t}L was selected[3] since its *w* value is small.[11] In work reported below, a conjectural helix-inhibiting synergy between linked ^{t}L Inp$_2$ and ^{t}L is probed and values are assigned to $j_L{}^t$ and $c_L{}^t$.

The N-cap $_D$-alanine ≡ "a" and the C-cap A-Inp, both helix-terminating, are needed as capping standards. For these, j_a and c_{A-Inp} should closely approximate 1.0. Like ^{t}L, "a" is a strong helix breaker,[12] but its φ, ψ rotamers can provide polar, hydrophobic, and steric environments such as those of an $_L$-alanine, and as an H-bond acceptor, its C-terminal carbonyl oxygen is a near-equivalent. A C-cap that is linked to its successive residue through a tertiary amide bond lacks the second NH H-bonding donor that is the prerequisite for helix extension; thus the A-Inp pair must act as a helix terminator and a normal donor for the first H-bond. The accessible Ramachandran space for the Ala of A-Inp should closely match that reported for the Ala of Ala-Pro.[13] Both can access sheets and extended conformations, which have positive values of ψ, but owing to steric crowding, the negative ψ values of the helical region correspond to destabilized, inaccessible conformations. In the discussions that follow, the constructs derived from the two generic sequences of **1** are abbreviated as ^{t}L-Ala$_n$-^{t}L and a-Ala$_n$-A.

The CD ellipticities of peptides in the short UV wavelength region are largely defined by the conformations of their backbone amide sequences, and at 222 nm the CD ellipticity $[\theta]_{222}$ primarily reflects the presence of helical structure. We have previously shown that to a good approximation, CD spectra of ^{t}L-Ala$_n$-^{t}L constructs can be analyzed as sums of independent contributions of caps and cores, and $[\theta]_{222}$ data are particularly useful for detecting changes in core helicity that result from local changes in ^{t}L-Ala$_n$-^{t}L constructs.[2b,3b] For example, changes in the lengths of either the Lys$_m$ regions or of the Ala$_n$ cores contribute independently to CD spectra, which can be cap-corrected to yield core values. A comparison of $[\theta]_{222}$ data of Table 1, entries 2 and 3 show that, as a solubilizer, Arg can replace Lys.

Further table data address the important issue of helicity-inhibiting effects of Inp$_2$^{t}L and ^{t}LInp$_2$. Does the hydrophobic bulk of the piperidine rings of Inp$_2$ alter helicity at sites once-removed from the core? A comparison of entry 6 with 8 shows that, for A$_{15}$ peptides, replacement of Inp by the substantially less bulky Aze 3-carboxyazetidine does not perturb $[\theta]_{222}$ within measurement error.[14] Does the ^{t}L-Inp tertiary amide strongly perturb helicity? Entries 4 and 5, 6 and 7 show that without significant change in $[\theta]_{222}$ either Inp$_2$^{t}L or ^{t}LInp$_2$ can be replaced by the secondary amide-linked tripeptide sequence ^{t}L$_3$, which is their functional equivalent.[2b] We conclude that any strong inhibition of core helicity by Inp$_2$^{t}L and ^{t}LInp$_2$ caps must be attributed solely to the ^{t}L residue itself.

We have noted previously that, inconsistent with large values for w_{Ala}, classical unstructured CD spectra are observed for ^{t}L-Ala$_n$-^{t}L, n = 10–12. Can lack of CD helicity be caused by ^{t}L? Entries 1, 2, and 3 show that $[\theta]_{222}$ lacks helical values if ^{t}L-Ala$_{10}$-^{t}L is replaced by a-Ala$_{10}$-A; clearly Ala$_{10}$ helicity is undetectable, even for $j = c = 1$.

The Table also presents CD data for ^{t}L-Ala$_{15}$-^{t}L, ^{t}L-Ala$_{15}$-A, a-Ala$_{15}$-^{t}L, and a-Ala$_{15}$-A peptides, all of which exhibit weakly helical CD spectra. This series allows comparisons of the results of mutation of one or both of the ^{t}L residues. If the value of one capping parameter is reduced from 1.0 to 0.0 by a mutation, the resulting $[\theta]_{222}$ change corresponds to a single residue deletion: e.g., if $c_L{}^t = 0$, $[\theta]_{222}$ for a-Ala$_{15}$-^{t}L should closely approximate that of a-Ala$_{14}$-A. In either ^{t}L-Ala$_{15}$-^{t}L or a-Ala$_{15}$-A series, a polynomial length fit of $[\theta]_{222}$ values has a local per-residue change of $(-4.3 \pm 0.2) \times 10^3$ deg dm^{-1} cm^{-1} for n = 14 through 18.[3b,15] Comparisons of this slope with results of single and double ^{t}L substitutions in a-Ala$_{14}$-A, entries 6, 9, 10, and 11, imply that values for $j_L{}^t$ and $c_L{}^t$ are bounded by 1.0 and 0, with likely values that are close to 0.5. A similar conclusion follows from our previously reported C-terminal $[\theta]_{222}$ values.6

Lifson–Roig Modeling of Site Helicities FH$_i$ for Conformational Ensembles

Both NMR ^{13}C═O and ^{13}Cα chemical shifts are used to assign secondary structure within protein sequences[16] as well as helicities for peptides[17] and proteins in native and partially structured states.[18] For selectively labeled constructs a-A$_{15}$-A and ^{t}L-A$_n$-^{t}L, n = 7 through 19, we now analyze ^{13}C═O NMR chemical shift data measured in water at 2 °C. In principle, a ^{13}C chemical shift of a residue i of a helical peptide sequence is proportional to the site helicity FH$_i$, the mole fraction of peptide conformations that are helical at residue i. An experimental FH$_i$ is thus an abundance-weighted average of values for members of an ensemble containing fully and partially helical conformations. For an Ala$_n$ peptide, the weight of a single ensemble conformation corresponds to one of the 2^n terms in an L–R state sum.[8]

Individual conformations cannot be characterized, but ensemble averages of these properties can be calculated from L–R parameters such as w, c, and j. If these have been assigned, FH$_i$ can be calculated as $ss_{helix,i}/ss$, in which $ss_{helix,i}$ is the part of the state sum that includes weights of all conformations in which the ith residue is part of a helix. The average of FH$_i$ over all potentially helical sites i is the fractional helicity FH of the full Ala$_n$ peptide, which is relatable

to the CD ellipticity $[\theta]_{222}$, the most commonly used tool for characterizing global peptide helicity.

A data set of FH or FH_i for an Ala_n library allows the reverse of this process: use of L–R algorithms to assign the best-fit parameters *w*, *j*, and *c*. Libraries must be constructed to minimize covariance between parameters, since inclusion of each new amino acid residue within library sequences adds three to the size of the parameter set. The polyalanines are uniquely suitable for falsification tests of the L–R hypothesis that the value of *w* is length-independent, since a library of ^{t}L-A_n-^{t}L peptides allows FH_i to be defined by just three parameters: w_{Ala}, $j_L{}^t$, and $c_L{}^t$; for a-A_{15}-A peptides, w_{Ala} alone is required.

For a peptide such as ^{t}L-Ala_n-^{t}L, four types of conformations contribute to FH_i. Most partially helical conformations of length $k < n$ extend to neither ^{t}L residue, and these have L–R weights $j_{Ala}\upsilon^2 w_{Ala}{}^k c_{Ala} = \upsilon^2 w_{Ala}{}^k$. Two types of helical conformations are capped by one but not both ^{t}L residues; depending upon whether they extend to the N- or the C-terminus, these are weighted $j_L{}^t\upsilon^2 w_{Ala}{}^k$ or $\upsilon^2 w_{Ala}{}^k c_L{}^t$. Only the completely helical conformation is weighted $j_L{}^t\upsilon^2 w_{Ala}{}^k c_L{}^t$. If both $j_L{}^t$ and $c_L{}^t$ are assigned very large values, this conformation dominates the ensemble, and the site-dependent FH_i plot approaches 1.0 for all *i*. The opposite extreme is provided by a helical peptide characterized by a *w* that approaches 1.0 and that lacks strongly helix-stabilizing N- and C-caps. All its helical conformations, short and long, make significant contributions to the FH_i, which are maximal in the central region and decrease substantially toward each terminus. Such a helix is commonly described as “frayed”.

An introduction to the interpretation of ^{13}C-derived FH_i data is provided by Figure 2a–c, which deconstructs each modeled FH_i value at a series of Ala_{15} sites as a sum of contributions from a length series of ensemble members. A key question is how are changes in the magnitudes of *j* and *c* mirrored in the site values of FH_i? Data of Figure 2a–c were L–R modeled for a hypothetical ^{t}L-A_{15}-^{t}L peptide in which the capping parameters are assigned three values: $j_L{}^t = c_L{}^t = 0.0$, 1.0, or 10.0. To facilitate comparisons, the calculated value of the central residue FH_8 is set as 0.66, which corresponds to the experimental δ^{13}C═O at site 8 of ^{t}L-Ala_{15}-^{t}L at 2 °C in water. To achieve this value, the w_{Ala} must vary inversely with $j_L{}^t = c_L{}^t$. Thus for $j_L{}^t = c_L{}^t = 10$, the required w_{Ala} is 1.1854, but for $j_L{}^t = c_L{}^t = 0.0$, w_{Ala} is 1.5215. These w_{Ala} values nearly span the range reported for helix-stabilizing amino acids.[4,5a]

In each plot, the upper boundary of the yellow region defines FH_i, and differences between ordinate values at the upper and lower boundaries of each colored region correspond to contributions to FH_i from conformers within a defined length range. Nonzero weights for the three conformers of lengths 15 and 14 (purple and blue) contribute to all FH_i. If the values of $j_L{}^t$ and $c_L{}^t$ approach zero, their weights become insignificant, but if $j_L{}^t = c_L{}^t$ is increased from 1 to 10, the weights are respectively increased by 100 and 10. In this case, as seen in Figure 2c, these three conformers dominate the helical ensemble.

Changes in capping parameters thus control mole fractions of longer helical conformers, and these in turn largely determine the site dependences of the FH_i. If the capping parameters assume large values, the site dependence of FH_i is remarkably small, and the local curvature in the central region of a FH_i plot is decreased by any effect that stabilizes long over short helical conformers. If $j_L{}^t = c_L{}^t$ approaches zero, only conformations that lack N- and C-terminal capping weights contribute significantly to the helical ensemble, and the longest helical conformer that contributes to the state sum has length $(n - 2)$. In effect, for these capping values, the residue length of the helical region for the Ala_n peptide is truncated by two. Thus if its $j = c = 0$, an Ala_{15} peptide behaves like a corresponding a-Ala_{13}-A peptide, modeled by the same *w*. Finally, comparison of Figure 2a and 2b shows that a change in $j_L{}^t = c_L{}^t$ from 0 to 1 dramatically increases FH_3, FH_2, and FH_1, decreasing the difference between these values.

The corresponding site dependences within the central region are much less sensitive to changes in $j_L{}^t = c_L{}^t$.

Assignment of Site-Dependent Chemical Shifts and Preliminary Studies of $^{13}C{\equiv}O$ Data

Prior to helicity analysis, the ^{13}C NMR chemical shift for each Ala_n residue must be assigned. For heteropeptides, standard ^1H, ^{13}C, and ^{15}N NMR experiments are used for these assignments. Poor dispersion prevents an analogous Ala_n assignment, but the alternative, synthesis and NMR analysis of *n* singly labeled position isomers, is prohibitively time-intensive. We used a new, highly efficient ^{13}C tagging of multiple sites within a single peptide. A series of distinct $^{13}C/^{12}C$ ratios were incorporated at up to six residue sites; the relative ^{13}C intensity of a resonance then identifies its site.

As reported previously for ^1H spectra,[19] ^{13}C Ala resonances for Inp-containing peptides exhibit anomalous multiplicities. These reflect mixtures of slowly equilibrating *c*- and *t*-state diastereomers formed at the tertiary amide linkages. For studies of ^{13}C-labeled ^{t}L-Ala_n-^{t}L peptides, we turned to the functionally equivalent W-Lys_m-^{t}L$_3$-Ala_n-^{t}L$_3$-Lys_mNH$_2$ series, which show normal singlets. A single Inp residue appears in members of the ^{13}C-labeled a-Ala_n-A peptide series, and NMR resonances of ^{13}C-labeled Ala residues within the C-terminal Ala_n region appear as close doublets, but FH_i values can be assigned from area-weighted averages of their integrated intensities.

$$FH_i = ((\delta^{13}C_{Exp-i},\ \text{ppm}) - (\delta^{13}C_{Coil},\ \text{ppm}))\Big/((\delta^{13}C_{\alpha-Helix},\ \text{ppm}) - (\delta^{13}C_{Coil},\ \text{ppm})) \quad (1)$$

$$FH_i = ((\delta^{13}C{=}O_{Exp-i},\ \text{ppm}) - 177.5)\Big/(180.9 - 177.5) \quad (2)$$

Under the assumption that end and position effects are insignificant, $^{13}C{=}O$ and $^{13}C\alpha$ chemical shifts have been previously used to calculate FH values from eq 1.[17,18] The required, recently assigned[19] $^{13}C{=}O$ calibrations for α-helical Ala (FH_i ϒ 1.0) and unstructured Ala (FH_i ϒ 0.0) in an Ala_n context are incorporated in eq 2. If position effects are small, the FH_i calculated at a particular site from $^{13}C{=}O$ and $^{13}C\alpha$ data should be identical. Data of Table 2 show that these FH_i are in reasonable agreement at sites 3 through 8, but the agreement is much less satisfactory in end regions. The data also reveal a striking lack of left/right symmetry for both data sets. This observation is consistent with a recent ab initio modeling study of ^{13}C chemical shifts for short, completely helical Ala_n peptides, which demonstrates a large position dependence within the C-terminus.[20]

Our initial strategy for assigning *j* and *c* was based on site-dependent ^{13}C data of Figure 3. The red circles correspond to experimental $^{13}C{=}O$ chemical shifts for the ^{t}L-Ala_{15}-^{t}L peptide measured in water at 2 °C, with open red circles used for data that are influenced by end effects. The black curves correspond to site-dependent chemical shifts calculated from five fixed *w* L–R modelings for $j = c$ assignments of 0, 0.5, 1.0, 2.0, 10. This assignment strategy clearly fails. The dispersion of chemical shifts required to assign $j_L{}^t$ and $c_L{}^t$ are seen only for the less reliable open circle data. Moreover the data at sites 4–8, which are expected to better correlate with FH_i, show almost no dispersion. These data suggest dominance of longer helical conformers within the helical ensemble, which is inconsistent with conventional L–R modeling. Moreover the data are approximated only by values of $j_L{}^t = c_L{}^t$ that exceed 1.0, contrary to CD evidence that both $j_L{}^t$ and $c_L{}^t$ lie between 1 and 0.

More useful information is obtained from a comparison of $^{13}C{=}O$ chemical shifts for ^{t}L-Ala$_{15}$-^{t}L and a-Ala$_{15}$-A, measured at 2 °C in water, Figure 4. The larger central residue chemical shift for a-Ala$_{15}$-A peptide vs ^{t}L-Ala$_{15}$-^{t}L proves that ^{t}L caps are helix-destabilizing. The data also prove that left/right asymmetries observed for ^{t}L-Ala$_{15}$-^{t}L must be intrinsic to Ala$_n$ and are not attributable to the ^{t}L caps.

What feature of Ala$_n$ is responsible for these asymmetries? Each site-dependent ^{13}C chemical shift is an abundance-weighted average over the ensemble conformers. At the N-terminal and C-terminal sites, where the asymmetries are largest, all helical conformers that contribute to the average are respectively initiated and terminated. Plausibly the asymmetries reflect chemical shift differences of terminal Ala amides caused by helix dipoles[21] and distinct local solvation environments.[22] To explore this hypothesis, an L–R model was constructed in which each FH$_i$ reflects contributions from three subensembles of helices that are initiated, continued, or terminated at the site *i*. Figure 5 shows that broad central regions are dominated by the second population, but the first and third dominate end regions. Qualitatively, these results are consistent with the experimental left/right asymmetries of Table 2 and Figure 4.

$^{13}C{=}O$ Chemical Shifts of Central Residues as a Test of the Length-Independence of *w* Values; A Length-Dependent *w* Series; Assignment of ^{t}L Capping Parameters

In this section we resolve the two central issues of this report: the possible dependence of w_{Ala} on conformer length and the assignment of values for w, $j_L{}^t$, and $c_L{}^t$. As seen in Figure 6, for the plausible limits of $j_L{}^t$, and $c_L{}^t$, the central residue $^{13}C{=}O$ chemical shifts measured in water at 2 °C for the ^{t}L-Ala$_n$-^{t}L series, n = 9, 11, 13, 15, 19 cannot be modeled by a length-independent w_{Ala}.

To assign conjointly both a length dependent series of $w_{Ala}(n)$ and a value for ($j_L{}^t = c_L{}^t$), this set of shifts was expanded to include the central residue $^{13}C{=}O$ shift at 2 °C for the peptide a-Ala$_n$-A. From the set of ^{t}L-Ala$_n$-^{t}L data, paired with each of 11 choices for $0 \leq (j_L{}^t = c_L{}^t) \leq 1.0$, taken at intervals of 0.1, we constructed a set of 11 length-dependent w_{Ala} (n). Each assignment was made recursively, using, in turn, central residue ^{13}C chemical shift data for n = 9, 11, 13, 15, 19. Previously reported t/c data[3a] were used in the first iteration, as described in the Experimental Section. From each of these 11 ($j_L{}^t = c_L{}^t$) + w_{Ala} (n) pairs, by setting $j_a = c_{A-Inp}$) 1, we calculated a central residue chemical shift for the a-Ala$_{15}$-A peptide. These are shown as blue squares in Figure 7. The experimental chemical shift for the a-Ala$_{15}$-A peptide21 appears in red on the ordinate and corresponds to $j_L{}^t = c_L{}^t = 0.4 \pm 0.1$.[23] The values in this range, paired with their $w_{Ala}(n)$, thus mirror the experimental central residue chemical shifts of the ^{t}L-Ala$_n$-^{t}L series, n = 9, 11, 13, 15, 19, as well as that for the calibrating a-Ala$_{15}$-A peptide.

$$\delta^{13}\mathrm{C{=}O}_i,\ \ \mathrm{ppm} = 177.5 + (\mathrm{FH}_i - (\mathrm{FH}_{i-\mathrm{I}} + \mathrm{FH}_{i-\mathrm{T}}))$$
$$(\delta^{13}\mathrm{C}_{\alpha-\mathrm{Helix-E}} - 177.5) + \mathrm{FH}_{i-\mathrm{I}}(\delta^{13}\mathrm{C}_{\alpha-\mathrm{Helix-I}} - 177.5) + \qquad (3)$$
$$\mathrm{FH}_{i-\mathrm{T}}(\delta^{13}\mathrm{C}_{\alpha-\mathrm{Helix-I}} - 177.5)$$

Can one use the hypothesis underlying Figure 5, together with assigned $j_L{}^t = c_L{}^t$ and $w_{Ala}(n)$, to model 2 °C values of noncentral chemical shifts for the ^{t}L-Ala$_n$-^{t}L length series? To explore this possibility, eq 3 was derived from eq 1 by partitioning each FH$_i$ into contributions from conformations that are initiated at site i, FH$_{i-I}$; extended at site i, (FH$_i$ − (FH$_{i-I}$ + FH$_{i-T}$)); or terminated at site i, FH$_{i-T}$. This model requires three values for $\delta^{13}C_{\alpha\,-Helix}$. Our prior value, 180.9 ppm, is used for $\delta^{13}C_{\alpha\,-Helix-E}$. For $\delta^{13}C_{\alpha\,-Helix-I}$, 180.1 ppm is used, extrapolated from

the experimental value at site 3 of the a-Ala_{15}-A peptide. For $\delta^{13}C_{\alpha\ -Helix-T}$ we used 172.2 ppm, extrapolated from the experimental value at site 13 of this peptide. Data are compared with modeling results in Figure 8. Although data for $n = 9$ and for sites to the left of center are underestimated, the asymmetries are modeled successfully and most modeled points agree with data within experimental error.

Discussion

This is the first of a series of reports in which length- and temperature-dependent $w_{Ala}(n)$ values are assigned from independent sets of helicity data. No single type of helicity measurement is optimally precise throughout the span of a full Ala_n series. Within the length range of this study, experimental $^{13}C{=}O$ chemical shifts are ideal, but neither can they be applied to peptides shorter than 10 residues, for which chemical shifts at 2 °C approach $\delta^{13}C_{Coil}$, nor can they be applied to lengths longer than ca. 20 residues, for which shifts approach $\delta^{13}C_{\alpha\ -Helix}$. Subsequent reports present CD and protection factor data that allow precise assignment of $w_{Ala}(n)$ for longer peptides. Previously reported t/c values do permit an independent confirmation of the assignments made in this report. The t/c helicity assay has an optimal Ala_n length range of 4 to 15 residues and is uniquely sensitive to length-dependent changes in helical stability.[3a]

A t/c value is the ratio of integrated intensity of ^1H NMR resonances for the *s-trans* and *s-cis* tertiary acetamido rotamers of the helix-initiating peptide N-cap Ac-Hel. Ac-Hel is a conformationally restricted derivative of Ac-Pro-Pro that is preorganized to initiate helices in a linked peptide. If Ac-Hel is linked to an Ala_n sequence, t/c is effectively proportional to a ratio of state sums for *s-trans* and *s-cis* conformational ensembles. Figure 9 shows t/c values measured in water at 2 °C for Ac-Hel-Ala_n-^{t}L-Inp_2-K_4-W-NH_2, $n = 4$ through 14, together with modeled t/c data using the $c_L{}^t$ and $w_{Ala}(n)$ assignments of the preceding section. For the Ala_n length range of 4 through 19 at 2 °C, our cap and *w* assignments quantitatively model both ^{13}C and t/c data within measurement and assignment error.

In any classical L–R model, the stability of a helical homopeptide increases linearly with its length, but the values we have assigned for $w_{Ala}(n)$ imply that the actual increase must be greater than linear. H-bonding cooperativity is the likely cause. In accord with conclusions of earlier modeling results,[24] ab initio calculations by Wieczorek and Dannenberg[25] show that for Ala_n helices the strengths of amide–amide H-bonds increase significantly with peptide length *n*. Testable predictions with paradigm-breaking potential follow from this nonlinearity. For example if the length increase of $w_{Ala}(n)$ is attributable to H-bonding cooperativity, this effect, which is a property of the helix backbone, is likely to be largely independent of residue and sequence. It then follows that helical propensities for all residues are expected to increase proportionately with length, but their ratios should be length-independent. Residues that are borderline breakers ($w < 1$) in short helices may be transformed into helix formers ($w > 1$) in longer helices.

A final issue concerns the magnitude of changes in helical structure that result from a length dependence for $w_{Ala}(n)$. Inspection of the TOC graphic for this report reveals the changes in the ^{t}L-Ala_{15}-^{t}L helical ensemble population that results if L–R modeling using a constant $w = 1.496$ is replaced by our new length-dependent values. At all sites the relative contribution to FH_i by helical conformations of lengths 12 to 13 is nearly constant, but the contribution from lengths 14–15 increases substantially, at the expense of shorter conformations.

Experimental Section

Synthesis, Purification, and Characterization of Peptides

Peptides were synthesized on a 0.02–0.03 mol scale by automated continuous-flow solid-phase synthesis using a PE Biosystems Pioneer Peptide Synthesizer with standard 9-fluorenylmethyleneoxy-carbonyl (Fmoc/HATU) chemistry, and peptides were cleaved from the resin as reported previously.[3] Each unlabeled coupling used 9 equiv of Fmoc-AA-OH and HATU. The labeled sites used HATU and 3 equiv of a mixture of ^{13}C labeled and nonlabeled Fmoc-Ala-OH with ^{13}C percentages that varied from 100% to 25%. Each labeled site was then double coupled with 9 equiv of nonlabeled Fmoc-Ala-OH. Although the ^{13}C-α chemical shifts appear to be less sensitive to position effects, the cheaper $^{13}C{=}O$ labels were used for this study.

Purification was carried out on a Waters 2690 HPLC with autoinjector and 996 detector using YMC ODS-AQ 200 Å 4.5 mm × 150 mm columns. Closely related unlabeled peptides were previously demonstrated to be unaggregated under measurement conditions by AUC.[3,19] MS data obtained on a Waters ZMD mass spectrometer and labeling patterns used to assign ^{t}L-Ala$_n$-^{t}L $^{13}C{=}O$ chemical shift values of Figure 9 are reported in the Supporting Information.

CD Experiments

Circular dichroism measurements were obtained on an Aviv 62DS spectrometer, and peptide concentrations were determined on a Cary 300 UV–vis spectrometer utilizing the Trp chromophore of the peptide, as previously reported.[3,5,12b] The 62DS was calibrated using titrated water solutions of sublimed 9-camphor-sulfonic acid. [3]

NMR Experiments

Our ^{13}C labeling technique, based on varying $^{13}C/^{12}C$ ratios, significantly improves the speed and efficiency of ^{13}C chemical shift characterization. Since an Ala$_n$ is a homopeptide, residues isolated from end regions are expected to experience very similar environments. Gratifyingly very similar line widths were observed for all ^{13}C resonances of this study. (See the Supporting Information for a representative spectrum.) Prior to chemical shift assignment, peptide samples were subjected to NH → ND exchange by at least three overnight lyophilizations from D_2O. Concentrations in D_2O within the range of 2–10 mM were used for NMR measurements. The ^{13}C spectra were taken on a 500 MHz Oxford Magnet, with a Unity INOVA console using a Varian 500 SW/PFG broadband probe. Standard ^{13}C parameters were used: delay, 0.763 s; pulse width 6.9 ms, acquisition time, 1.736 s; line broadening, 2 Hz; 128 to 512 scans were needed to get adequate signal-to-noise. The ^{13}C chemical shifts were referenced relative to 1 mM DSS.[26] With precautions to ensure equilibration, temperature was measured using neat methanol. Owing to the temperature dependence of FH$_i$, measurement errors can bias assigned $^{13}C{=}O$ chemical shifts. Calibration of limiting ^{13}C shifts for any new peptide series is an essential precondition for calculation of FH$_i$ from chemical shift data. Our value of 180.9 ppm for the limiting helical chemical shift of an Ala residue[19] substantially exceeds the average of reported Ala values from the protein database. It was measured for an unaggregated Ala$_{12}$ peptide in which high helicity (99+ %, assigned from NH → ND exchange kinetics) was induced by optimal N- and C-caps. The method of Cordier and Grzesiek[27] was used to establish dominance of α-helical structure and absence of detectible 3_{10} character.

Lifson–Roig Modeling

Previously reported algorithms[2e] were used to calculate FH$_i$ values and modeling data that appear in Figures 2, 4, and 6 and in the TOC. For the modeled t/c data of Figure 9, the previously reported algorithm[3a] was used with one change. Helical conformations initiated by the Ac-

Hel c state were assigned an N-capping parameter with a value chosen to fit the experimental t/c observed for Ala_{14}; as a helix-inducing N-cap the Ac-Hel c state has ca. two-thirds the efficiency of an *N*-acetyl group.

The Ala_n data of Figures 5 and 8 were calculated for each *i* site of each peptide of length *n* by sorting the explicit h,c list of major L–R helical conformers[2e] into three groups. Each h,c list corresponds to groups of terms within the helical state sum, which is then partitioned analogously. This sorting yields $ss_{helix,i,I}$ and $ss_{helix,i,T}$, but the remaining terms of the site *i* state sum $ss_{helix,i}$ are contributed by helices that are neither initiated or terminated at site *i*: $ss_{helix,i} - (ss_{helix,i,I} + ss_{helix,i,T})$. Division by $ss_{helix,i}$ then gives the three ordinate values of Figure 5; division by the overall state sum ss gives the values of FH_{i-I} and FH_{i-T} required for eq 3 and used in Figure 8.

Iterative L–R regressions were used to assign each of the 11 $w_{Ala}(n)$ series used for the modeling of Figure 7. Length-dependent $w_{Ala}(n)$ values are assumed to increase monotonically by small increments and converge to a limiting value for sufficiently large *n*. For a first iteration we started with previously reported t/c data in the length range of 1 to 8;[3a] these are compatible with $w_{Ala}(n)$ values of 1, 1, 1, 1, 1.28, 1.31, 1.33, 1.34. Using ^{t}L-Ala_n-^{t}L chemical shift data, $n = 9$ and 11, under the assumption that $j = c = 0$, values for $w_{Ala}(n)$, $n = 7–11$ were adjusted or assigned that minimized residuals. By an analogous recursive process, the $w_{Ala}(n)$ data set was extended to $n = 19$. In nearly all cases the condition $w_{Ala}(n+1) - w_{Ala}(n) \leq 0.03$ was met. This recursion process was then repeated 10 times, with a progressive increase of $j = c$ by 0.1; for each of these, perturbations were made in the previously defined $w_{Ala}(n)$ data set. The 19-member $w_{Ala}(n)$ series at 2 °C, $j = c = 0.5$, from t/c, and ^{13}C data are (in numerical order from 1 to 19) as follows: 1, 1, 1, 1, 1.28, 1.31, 1.33, 1.34, 1.39, 1.41, 1.425, 1.455, 1.48, 1.505, 1.525, 1.54, 1.565, 1.58, 1.59. Alternative regression strategies provide close approximations to this series. A conventional L–R state sum is a polynomial in υ and *w*, in which each w^k power term corresponds to the overall weight of helices of length *k*. It is converted to the modified L–R state sum required for the $w_{Ala}(n)$ calculation by treating *w* as a dummy variable ($w = 1.0$) and multiplying each w^k term by $w_{Ala}(k)^k$.[2e,3a]

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Financial support was provided by NIH Grant GM13453 (S.M.W. and R.J.K.).

References

1. Aurora R, Rose GD. Protein Sci 1998;7:21–38. [PubMed: 9514257]
2. Doig AJ, Baldwin RL. Protein Sci 1995;4:1325–1336. [PubMed: 7670375](a) b Miller JS, Kennedy RJ, Kemp DS. Biochemistry 2001;40:305–309. [PubMed: 11148022] c Maison W, Arce E, Renold P, Kennedy RJ, Kemp DS. J Am Chem Soc 2001;123:10245–10254. [PubMed: 11603974] d Deechongkit S, Tsang KL, Renold P, Kennedy RJ, Kemp DS. Tetrahedron Lett 2000;41:9679–9683. e Kemp DS, Helυ. Chim Acta 2002;85:4392–4423.
3. Kennedy RJ, Kwok KY, Kemp DS. J Am Chem Soc 2002;124:934–944. [PubMed: 11829601](a) b Miller JS, Kennedy RJ, Kemp DS. J Am Chem Soc 2002;124:945–962. [PubMed: 11829602]
4. Wojcik J, Altmann KH, Scheraga HA. Biopolymers 1990;30:11–143.(a)(b) Scheraga, H. A. In Perspectives in Structural Biology; Vijayan, M., Yathindra, N., Kolaskar, A. S., Eds.; Indian Academy of Sciences; Bangalore, 1999; pp 275–292.
5. Chakrabartty A, Kortemme T, Baldwin RL. Protein Sci 1994;3:843–952. [PubMed: 8061613](a) b Spek EJ, Olson CA, Shi Z, Kallenbach NR. J Am Chem Soc 1999;121:5571–5572.

6. Maison W, Kennedy RJ, Miller JS, Kemp DS. Tetrahedron Lett 2001;42:4975–4977.
7. In a conventional Lifson–Roig algorithm, the helical weight is $j\upsilon^2 w^{(n-2)}c$. In this report we use the weight $j\upsilon^2 wc$, which has significant advantages.[2e]
8. Lifson S, Roig A. J Chem Phys 1961;34:1963–1874.
9. Rohl CA, Scholtz JM, York EJ, Stewart JM, Baldwin RL. Biochemistry 1992;31:1263–1269. [PubMed: 1310608]
10. Forood B, Feliciano EJ, Nambiar KP. Proc Natl Acad Sci USA 1993;90:838–842. [PubMed: 8430094]
11. Lyu PC, Sherman JC, Chen A, Kallenbach NR. Proc Natl Acad Sci USA 1991;88:5317–5320. [PubMed: 2052608](a) b Hermans J, Anderson AG, Yun RH. Biochemistry 1992;31:5646–5653. [PubMed: 1610812]
12. Fairman R, Anthony-Cahill SJ, DeGrado WF. J Am Chem Soc 1992;114:5458–5459.(a)(b) Allen, T. J. Ph.D. Dissertation, M.I.T., 1993.
13. Schimmel PR, Flory PJ. J Mol Biol 1968;3:105–120. [PubMed: 5760450]
14. From $[\theta]_{222}$ values for the ^{t}L-Ala$_n$-^{t}L length series,3b the error in a $[\theta]_{222}$ difference is estimated as 8–10%; values less than ±2000 have an error margin of ca. 100%.
15. Walker, S. M.; Kemp, D. S. Unpublished observations.
16. Wishart DS, Richards FM, Sikes BD. J Mol Biol 1991;222:311–333. [PubMed: 1960729](a) b Wishart DS, Sikes BD. Methods Enzymol 1994;239:363–392. [PubMed: 7830591]
17. Shalongo W, Dugad L, Stellwagen E. J Am Chem Soc 1994;116:2500–2507.(a) b Werner JH, Dyer RB, Fesinmeyer RM, Andersen NH. J Phys Chem B 2002;106:487–494.
18. Eliezer D, Yau J, Dyson HJ, Wright PE. Nat Struct Biol 2001;5:148–155. [PubMed: 9461081]
19. Heitmann B, Job GE, Kennedy RJ, Walker SM, Kemp DS. J Am Chem Soc 2005;127:1690–1704. [PubMed: 15701003]
20. Vila JA, Baldoni HA, Scheraga HA. Protein Sci 2004;13:2939–2948. [PubMed: 15498939]
21. Eaton G, Symons MCR. J Chem Soc, Faraday Trans 1 1988;84:3459–3473.(a) b Eaton G, Symons MCR, Rastogi PP, O'Duinn C, Waghorne WE. J Chem Soc, Faraday Trans 1 1992;88:1127–1142.
22. Tidor B. Proteins: Struct, Funct, Genet 1994;19:310–323. [PubMed: 7984627]
23. Equivalent fits to these data also result from a range of $j \neq c$, provided their average lies between 0.3 and 0.5.
24. Mehler EL. J Am Chem Soc 1980;102:4051–4056.(a) b Guo J, Karplus M. J Phys Chem 1994;98:7104–7105.
25. Wieczorek, R.; Dannenberg, J. J. J. Am. Chem. Soc. 2003, 125, 8124–8129
26. Markley JL, Bax A, Arata Y, Hilbers CW, Kaptein R, Sykes BD, Wright PE, Wüthrich K. Pure Appl Chem 1998;70:117–142.
27. Cordier F, Grzesiek S. J Am Chem Soc 1999;121:1601–1602.

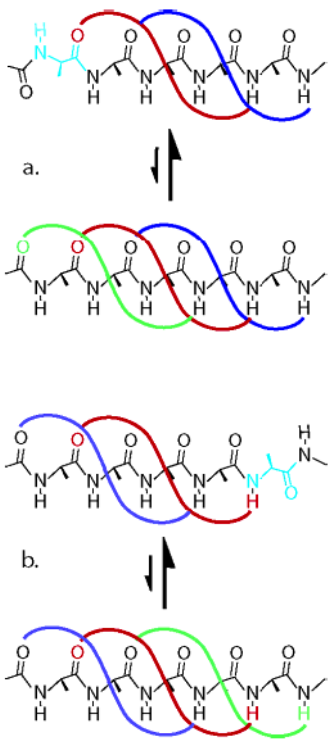

Figure 1.
Conformational features of residues that cap and terminate or cap and extend peptide helices. In the terminating conformation, only the red H-bond forms; in the extending conformation both red and green H-bonds must form, implying presence of comparable amounts of the upper and lower conformations of both a and b. Any helix-terminating residue must strongly favor nonhelical conformations, shown in cyan. A C-terminal cap must be helix-terminating if it is linked to a following peptide sequence through a tertiary amide, which lacks the NH required to form the green H-bond of b.

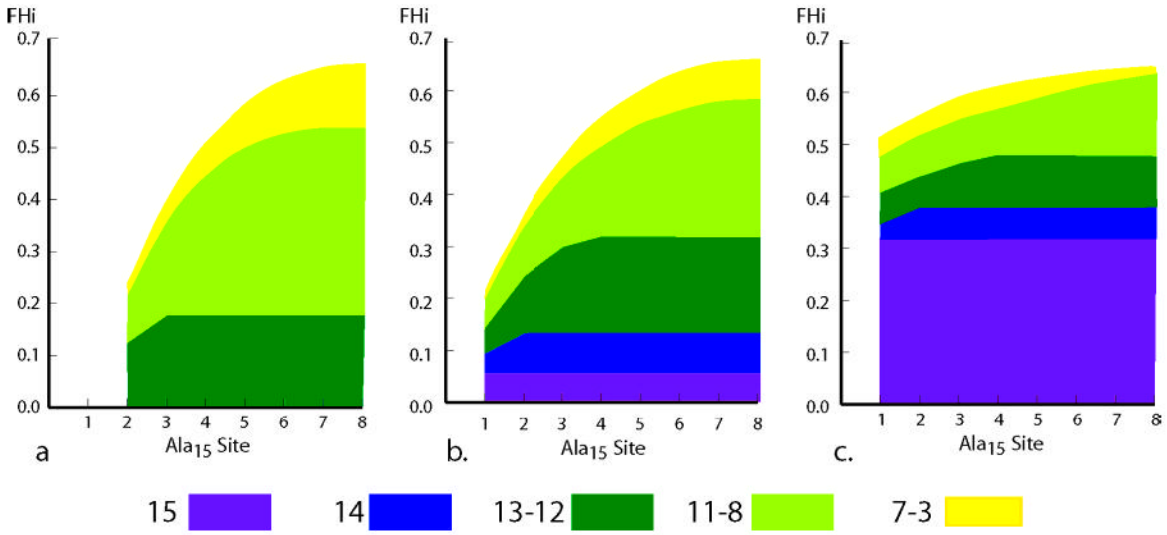

Figure 2.

L–R-modeled magnitudes of capping contributions to site-dependent FH_i for an Ala_{15} peptide. (a) $j = c = 0$; (b) $j = c = 1.0$; (c) $j = c = 10$. Values for FH_i correspond to the upper boundary of yellow regions, and fractional contributions to each FH_i from helical conformations of different lengths are symbolized by colored zones. The values of w were chosen to yield a constant value of $FH_8 = 0.66$. FH_i for sites 9–15 mirror corresponding FH_i at sites 7–1.

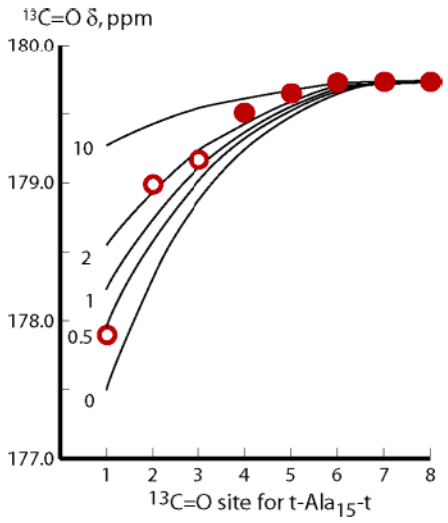

Figure 3.

Experimental $^{13}C{=}O$ chemical shift values for $^{t}L\text{-}A_{15}\text{-}^{t}L$, measured at 2 °C in water, red closed and open circles, compared with values calculated from eq 2. Required FH_i values were calculated from a conventional L–R model with fixed *w*, adjusted to yield the experimental central residue chemical shift. The curves for $j = c = 0$, 1, and 10 thus correspond to the FH_i values of the three Figure 2a–c. For each curve, *w* was normalized to yield 179.25 ppm for chemical shifts at sites 7 and 8.

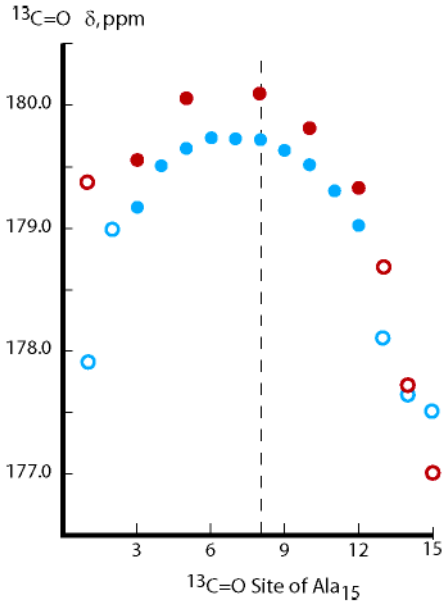

Figure 4.

Site-dependent $^{13}C{=}O$ chemical shifts at 2 °C in water for the ^{t}L-Ala$_{15}$-^{t}L peptide, blue circles, and the a-Ala$_{15}$-A-Inp peptide, red circles.

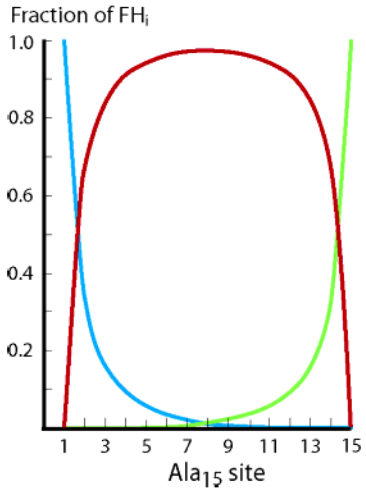

Figure 5.

A L–R deconvolution of site helicities FH_i as contributed by three types of helices for $j = c = 1.0$ and for subsequently assigned values for $w_{Ala}(n)$. The blue curve reflects the weight of helices initiated at site *i*; the green, helices terminated at site i; the red, all other helices. Helices containing 9 or more residues are used to calculate blue and green curves. Similar plots result if blue and green curves include shorter helices. A subsequent application of these modeling results appears in Figure 8.

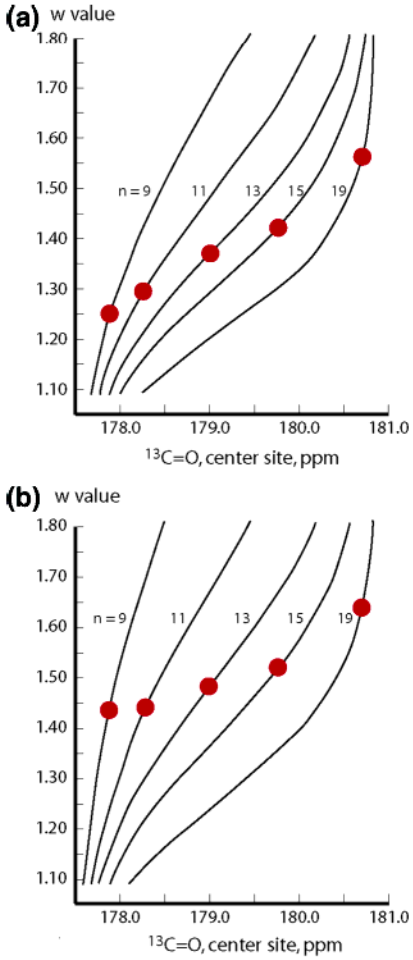

Figure 6.

Tests of the length independence of helical propensities, w_{Ala}. (a) $j = c = 1.0$; (b) $j = c = 0.0$. Red dots correspond to experimental central Ala residue $^{13}C{=}O$ chemical shifts for the length series of ^{t}L-Ala$_{15}$-^{t}L peptides at 2 °C. For each length n, lines correlate constant w values with calculated chemical shifts. The hypothesis of a length-independent w_{Ala} is falsified.

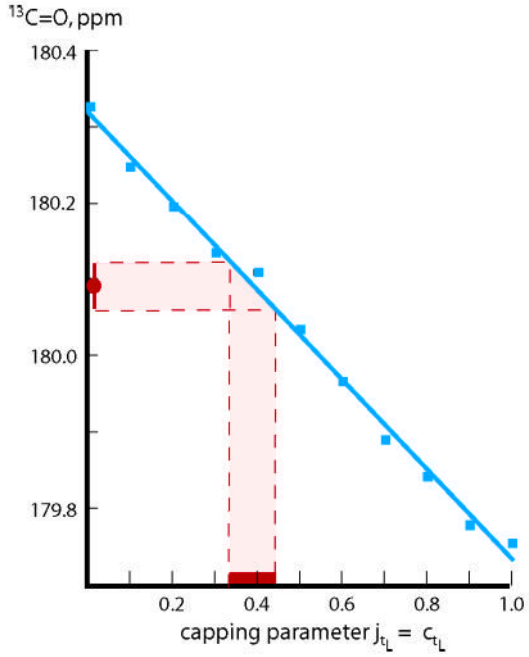

Figure 7.

Assignment of best-fit values for $j_L{}^t = c_L{}^t = 0.4 \pm 0.1$ (red abscissa region). The blue data points are 11 L–R shifts calculated for the central residue of a-Ala$_{15}$-A-Inp ^{13}C═O by setting $n = 15$, ($j_a = c_{A–Inp}$) = 1.0. For each calculation, a particular $w_{Ala}(n)$ set is used, assigned from central residue ^{t}L-Ala$_n$-^{t}L data with $j_L{}^t = c_L{}^t$ equal to the abscissa value. The red ordinate region corresponds to the experimental chemical shift for the a-Ala$_{15}$-A-Inp peptide.

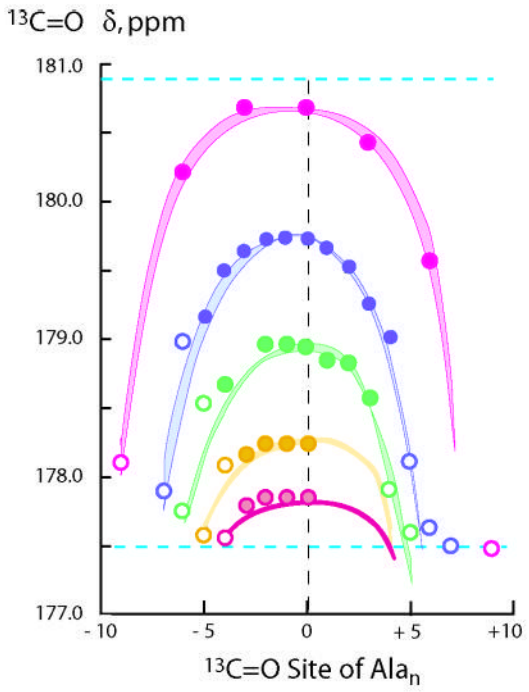

Figure 8.

Quantitative L–R modeling using eq 3 of site-dependent $^{13}C{=}O$ chemical shifts measured at 2 °C in water for the ^{t}L-Ala$_n$-^{t}L peptide series. Magenta, data points for n = 19; blue, n = 15; green, n = 13; orange, n = 11; pink, n = 9. The cyan dotted lines show the limiting chemical shift values for Ala residues that are fully helical, top, or nonhelical, bottom. Boundaries of each curved region correspond to L–R modeling for $j_L{}^t = c_L{}^t = 0.3$ and 0.5, each paired with its $w_{Ala}(n)$ set.

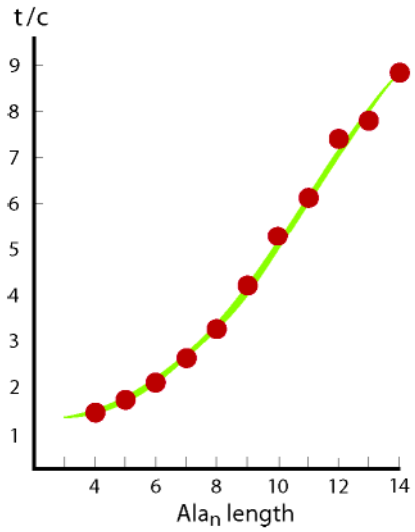

Figure 9.

A fit of t/c data, red circles, for the series Ac-Hel-Ala$_n$-^{t}L-Inp$_2$-K$_4$-W-NH$_2$, n = 4 through 14, measured in water at 2 °C. The upper boundary of the green zone corresponds to t/c calculated for $c_L{}^t$ = 0.5, paired with its $w_{Ala}(n)$ parameter set. The lower boundary, which gives a slightly poorer fit, is calculated for $c_L{}^t$ = 0.3, paired with its $w_{Ala}(n)$ set. For the former, as noted in the Experimental Section, a c state N-capping parameter of 6.4 was fitted to the value of the n = 14 data point; for the latter, the corresponding value was 6.2. The remaining required modeling parameters A = 0.80 and B = 0.17 have been assigned previously from t/c data.[3a]

Table 1

Circular Dichroism Ellipticity Data[a] for Peptides **1** at 2 °C, in Water[b]

Ala_{10} sequences	**$[\theta]_{222}$**
1. W-K_4-Inp_2-^{t}L-A_{10}-^{t}L-Inp_2-K_4-NH_2	+400
2. W-K_4-Inp_2-a-A_{10}-A-Inp_2-K_4-NH_2	−2100
3. W-R_4-Inp_2-a-A_{10}-A-Inp_2-R_4-NH_2	−1500

Ala_{12} sequences	**$[\theta]_{222}$**
4. W-K_4-Inp_2-^{t}L-A_{12}-^{t}L-Inp_2-K_4-NH_2	−3900
5. W-K_4-^{t}L_2-^{t}L-A_{12}-^{t}L-^{t}L_2-K_4-NH_2	−2400

Ala_{15} sequences	**$[\theta]_{222}$**
6. W-K_4-Inp_2-^{t}L-A_{15}-^{t}L-Inp_2-K_4-NH_2	−15 400
7. W-K_4-^{t}L_2-^{t}L-A_{15}-^{t}L-^{t}L_2-K_4-NH_2	−15 500
8. W-K_4-Aze_2-^{t}L-A_{15}-^{t}L-Aze_2-K_4-NH_2	−16 700
9. W-K_4-Inp_2-^{t}L-A_{15}-A-Inp_2-K_4-NH_2	−17 200
10. W-K_4-Inp_2-a-A_{15}-^{t}L-Inp_2-K_4-NH_2	−17 700
11. W-K_4-Inp_2-a-A_{15}-A-Inp_2-K_4-NH_2	−21 100

[a] Per-residue ellipticity data, deg cm^2 $dmol^{-1}$, are cap-corrected as previously reported.[3b]

[b] ^{t}L, tert-leucine; Inp, 4-carboxypiperidine; Aze, azetidine-3-carboxylic acid; a, D-alanine.

Table 2

Values for FH_i Calculated from Eq 1 from $^{13}C{=}O$ and α-C Chemical Shifts for ^{13}C Site-Labeled ^{t}L-Ala_{15}-^{t}L in Water at 25 °C, Using Previously Assigned Limiting Chemical Shift Parameters[19]

A. $^{13}C{=}O$-derived FH_i (first site data appear in first row; second site data, in second row)

sites 1,15	sites 2,14	sites 3,13	sites 4,12	sites 5,11	sites 6,10	sites 7,9	site 8
0.05	0.21	0.26	0.32	0.35	0.37	0.375	0.37
—	0.02	0.07	0.22	0.29	0.325	0.34	

B. $^{13}C\alpha$-derived FH_i

sites 1,15	sites 2,14	sites 3,13	sites 4,12	sites 5,11	sites 6,10	sites 7,9	site 8
0.00	0.26	0.29	0.31	0.36	—	0.39	0.39
0.00	0.02	0.15	0.25	—	0.35	—	