Peptide Structures

Calibrated Calculation of Polyalanine Fractional Helicities from Circular Dichroism Ellipticities**

*Gabriel E. Job, Björn Heitmann, Robert J. Kennedy, Sharon M. Walker, and Daniel S. Kemp**

In characterizing peptides, fractional helicities (FHs),[1] are calculated from per-residue molar ellipticities at $\lambda = 222$ nm [Eq. (1)].[2] Here, for a partially helical peptide of length n,

$$FH = [\theta]_{222,Exp,n}/[\theta]_{222,n} \quad (1)$$

$[\theta]_{222,Exp,n}$ is the observed per-residue molar ellipticity at $\lambda = 222$ nm, and $[\theta]_{222,n}$ is the corresponding length-dependent calibration value for a completely helical peptide. FHs are often used to construct quantitative helicity algorithms,[3] but for the important cases of alanine-rich peptides and polyalanines, the calculation is inaccurate[4] and requires calibration. Spaced, solubilized, helical Ala_n peptides (Figure 1) that span a large range of lengths and are characterized by FHs that approach 1.0 provide this calibration.

NMR experiments show that strong helix-stabilizing caps, like β-aminoalanine *beta* or the *N*-acyl-Pro-Pro analogue Hel, restrict the helical regions of these peptides to the Ala_n sequences.[5] NMR studies on simpler analogues of the sequence in Figure 1b[6] demonstrate literature-precedented,[7] short-range stabilizing contacts between caps and Ala_n termini by ^{1}H TOCSY, ROESY, and NOESY experiments, but tertiary interactions are not observed. ^{1}H, ^{13}C, and ^{15}N NMR chemical shifts, $^3J_{HNH\alpha}$ values, and α-helical structure are assigned from HNCA, E.COSY HNCA, and

[*] G. E. Job, Dr. B. Heitmann, Dr. R. J. Kennedy, Dr. S. M. Walker, Prof. D. S. Kemp
Department of Chemistry, Room 18-296
Massachusetts Institute of Technology
Cambridge, MA 02139 (USA)
Fax: (+1) 617-258-7500
E-mail: kemp@mit.edu

[**] This research was supported by the NSF CHE-0131250 and the NIH GM 13453 (S.M.W.).

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

 DOI: 10.1002/ange.200460536

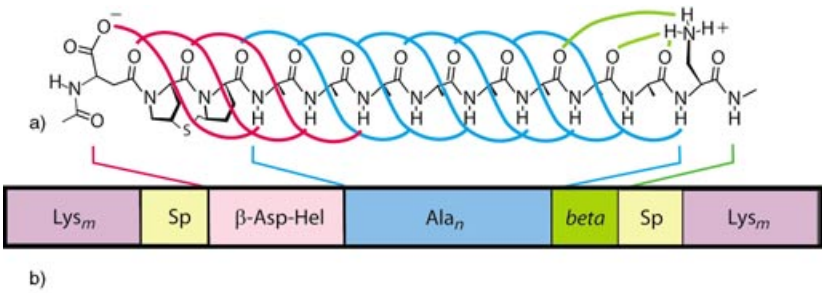

Figure 1. a) Molecular structure of the helical $^{\beta}$AspHel-Ala$_n$-*beta* subunit shared by all spaced, solubilized peptides of this study. Blue: α-helical H-bonds within the Ala$_n$ region; green: H-bonds that link *beta* to the Ala$_n$ region; red: H-bonds that link $^{\beta}$AspHel to the Ala$_n$ region. b) Functional regions within the peptide series used for CD calibration Ac-Trp-Lys$_5$-Inp$_2$-^{t}L-$^{\beta}$AspHel-Ala$_n$-*beta*-^{t}LInp$_2$-Lys$_5$-NH$_2$, n = 4–11; n = 12–24, n even. Only residues of the blue Ala$_n$ core are helical.[5b,c] Pink: helix-stabilizing N-cap; green: helix-stabilizing C-cap; purple: polyLys solubilizer; yellow: spacing elements[5a,d] drawn from the list: ^{t}L = *tert*-leucine, Inp = 4-carboxypiperidine, Acc = *trans*-4-aminocyclohexanecarboxylic acid. See Experimental Section and Supporting Information for details.

H(N)CO experiments on simple and ^{13}C,^{15}N-labeled Ala$_n$ peptides. The Ala$_n$ peptides are unaggregated in water by analytical ultracentrifugation.

The chemical shifts for amide NH protons at sites 5 to (n−4) of Ala$_n$ sequences with $n > 8$ are independent of both site and length n. For the first four alanines of each peptide, the HN resonances are resolved, and for a particular site their chemical shifts are constant throughout the series; this property is also seen for the three resonances assigned to the last four alanine NHs. NMR-assigned rate constants for backbone NH→ND exchange in D_2O at 2 °C, pH 4.5–6.0, yield protection factors PF$_i$ for each peptide,[8] from which the FH$_i$s and an average FH can be calculated. Most FH$_i$s are equal to or greater than 0.985, but at C-terminal sites NH$_{(n-3)}$ through NH$_{(n-1)}$, a monotonic decrease is seen from 0.98 to 0.94. The FH$_n$, defined by PF$_{n+1}$ of *beta*, have respective lower and upper limits of 0.1 and 0.7.[6] For the overall Ala$_n$ series, FHs lie in the range 0.90 to 0.98,[9] but for the central Ala$_n$ region, residues 2 through n−4, FH$_i$ is consistently equal to or greater than 0.993.[10] The calibrating circular dichroism (CD) relations [Eq. (2) and Eq. (3)] apply rigorously to a series with the above properties that meets the linearity test described below.[6]

$$[\theta]_{\lambda,n} = [\theta]_{\lambda,\infty}(1-X/n) \quad (2)$$

$$[\theta_{\mathrm{Molar}}]_{\lambda,n} = n[\theta]_{\lambda,n} = [\theta]_{\lambda,\infty}(n-X) = n[\theta]_{\lambda,\infty} - X[\theta]_{\lambda,\infty} \quad (3)$$

A CD-based calculation of FH for a partially helical peptide requires $[\theta]_{222,n}$, the ellipticity for an analogous peptide with FH = 1.0, [Eq. (1)]. Estimates for $[\theta]_{222,\infty}$ and X are usually applied to Equation (2) to obtain $[\theta]_{222,n}$,[2] but their unambiguous assignment has been problematic.[11] Both $[\theta]_{\lambda,\infty}$ and X can now be assigned from those $[\theta_{\mathrm{Molar}}]_{\lambda,n}$ values of our Ala$_n$ series that correlate linearly with n, as required by Equation (3).

Figure 2 shows CD spectra of five peptides from Figure 1b measured in water at 2 °C, pH > 4.5. A linear regression in n for the ten-member data set of $[\theta_{\mathrm{Molar}}]_{222,n}$ values, $9 \le n \le 24$, yields a slope of −59600 (standard deviation (SD) 1300) and an intercept of 202000 (SD 20700); the calculated error in slope at 222 nm corresponds to a relative precision of ±2%, within measurement error.[12] Restrictions of the data set to $11 \le n \le 24$ and $14 \le n \le 24$ yield respective slopes of −59300 (SD 1700) and −58800 (SD 3100). Within SD limits, these assignments are indistinguishable. For $n > 8$ and for all λ, the slopes, which are per-residue ellipticity increments, converge to limiting values, plotted as the red curve in Figure 2. This CD spectrum of $[\theta]_{\lambda,\infty}$

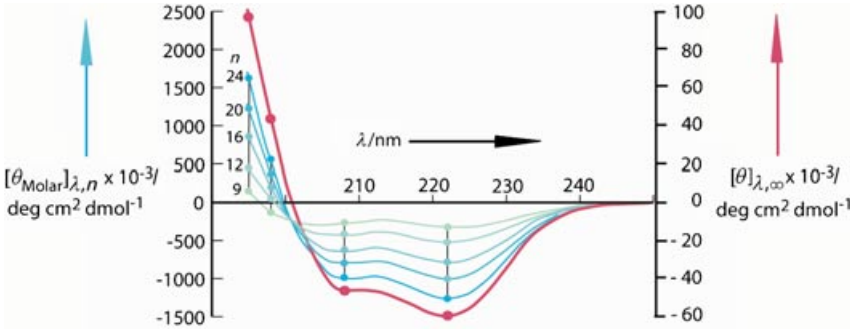

Figure 2. Length-dependent CD spectra, $[\theta_{\mathrm{Molar}}]_{\lambda,n}$, for the Ala$_n$ peptides described in Figure 1, in water, pH > 4.5, T = 2 °C, (blue-green curves, left-hand axis, blue arrow). For clarity, five representative spectra from the data base of ten are shown. Linear regressions for each wavelength λ (see text) yield slopes (red dots) from $[\theta_{\mathrm{Molar}}]_{\lambda,n}$ data (blue-green circles linked by vertical lines). These λ-dependent slopes (red curve, right-hand axis, red arrow) yield a CD spectrum of $[\theta]_{\lambda,\infty}$ values calculated for length-independent cores of completely helical polyalanines.

values reflects the length-independent core properties of completely helical polyalanines. A notable feature is a value of 1.3 for the ratio $[\theta]_{222}/[\theta]_{208}$, typical for alanine-rich helical peptides.[4c]

The value of X must reflect deviations from FH$_i$ = 1.0 within the Ala$_n$ termini. A linear regression on n and $[\theta_{\mathrm{Molar}}]_{222,n}$ data that have been corrected by subtracting $[\theta_{\mathrm{Molar}}]_{222,0}$ values and dividing the resulting differences by the length series of FH values yields $[\theta]_{222,\infty} = -60600$ (SD 1200), $X = 3.0$ (SD 0.3) for FH$_n$ = 0.1, and $[\theta]_{222,\infty} = -60500$ (SD 1200), $X = 3.2_5$ (SD 0.3) for FH$_n$ = 0.7. This X range is consistent with earlier reports.[10,13]

Applied to Equation (2) these parameters yield the calculated length dependences for $[\theta]_{222,n}$ given in Figure 3, which includes estimates of the precision of currently feasible assignments of FH from CD data. For a 24-residue peptide, the maximum error in X translates into a 10% error in the assignment of $[\theta]_{222,n}$ and in a calculation of FH that uses this parameter. For a 12-residue peptide, the error increases to 16%.

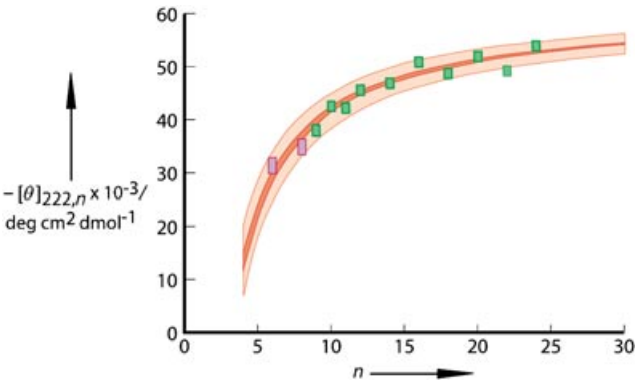

Figure 3. Length dependence of $[\theta]_{222,n}$, [Eq. (1)] calculated [Eq. (2)] from $[\theta]_{222,\infty} = -60600$ (SD 1200), $X = 3.0$ (SD 0.3) for $FH_n = 0.1$, and $[\theta]_{222,\infty} = -60500$ (SD 1200), $X = 3.2_5$ (SD 0.3) for $FH_n = 0.7$.[11] The red region of the graph $[\theta]_{222,n}$ was calculated from these mean parameter values. The pink region defines boundaries of values calculated from variations of X and $[\theta]_{222,\infty}$ by one SD unit. Values of the ten experimental $([\theta_{Molar}]_{222,n} - [\theta_{Molar}]_{222,1})/FH$ used to assign X and $[\theta]_{222,\infty}$ are plotted (green). For comparison, two corresponding values calculated from $[\theta_{Molar}]_{222,n}$, $n = 6$ and 8 are also shown (violet).

Inspection of Figure 3 shows that for $n < 12$, the value of $[\theta]_{222,n}$ is very sensitive to errors in the X assignment; X is expected to reflect changes in solvation at the helix termini, the presence terminal charges, and end-region contributions from 3_{10}-helical structure. For short peptides that belong to a particular structural series, maximal precision for calculation of FH from $[\theta]_{222,Exp,n}$ may require tailoring of X to mirror the CD properties of that series.

We have validated our earlier estimates[4] of $[\theta]_{222,\infty}$ for alanine-rich peptides by a method that can be generalized to other peptide series. The literature $[\theta]_{222,\infty}$ values of -37000[11a] to -44000[11b] underestimate $[\theta]_{222,n}$ for alanine-rich peptides, but they almost certainly remain relevant to most highly helical fragments derived from natural protein sequences.[11a] For what cases should our $[\theta]_{222,\infty}$ value assignment be used? The best criterion is the value of $[\theta]_{222}/[\theta]_{208}$, measured for members of a new peptide structural series under helix-stabilizing conditions. If this ratio exceeds 1.2,[4c] $[\theta]_{222,\infty}$ values of -60000 ± 1000 are appropriate choices.

Experimental Section

Peptides were synthesized, purified by repeated reverse-phase HPLC, characterized by electrospray ionization mass spectrometry (EI-MS), and analyzed by CD spectroscopy (Aviv 62DS circular dichroism spectrometer) following experimental protocols published previously.[4,5] The instrument was calibrated as described in its operating manual using titrated water solutions of sublimed 9-camphorsulfonic acid. Peptide concentrations were determined on a Cary 300 UV/Vis spectrometer utilizing the Trp chromophore of the peptide, as previously reported. A Bruker Avance 600 instrument (Karlsruhe, Germany) equipped with four channels and a pulsed field gradient triple probe with z gradients was used for all NMR studies. For convenience, NMR studies were initially carried out on $Ac^{\beta}AspHel-Ala_8$-*beta*-NH_2 and the series $Ac^{\beta}AspHelAla_n$*beta*$AccLys_2TrpNH_2$, where Acc is derived from the spacer *trans*-4-aminocyclohexanecarboxylic acid. Where appropriate, parallel NMR studies were carried out on the CD-calibration peptides of Figure 1b. PF measurements and chemical shift assignments for the Ala_n cores of these peptide series were in agreement. PF measurements on a similar Ala_n peptide have been reported previously.[4b]

The Supporting Information contains peptide AUC (Analytical Ultracentrifugation Equilibrium) and MS characterization, details of NMR structural assignments and chemical shift correlations, PF data and calculation of FH_i from PF_i data, and discussion of correlations underlying the analysis of Figure 2.

Received: May 1, 2004

Keywords: circular dichroism · helical structures · NMR spectroscopy · peptides

[1] FH, the fraction of potentially helical α-carbon atoms that are actually in a helical structure (range: 0.0 to 1.0); FH equals the average of site FH_i at each α-carbon of Ala_n cores.
[2] a) R. W. Woody, *J. Polym. Sci. Macromol. Rev.* **1977**, *12*, 181–320; b) M. C. Manning, R. W. Woody, *Biopolymers* **1991**, *31*, 569–586; c) Y.-H. Chen, J. T. Yang, H. M. Martinez, *Biochemistry* **1972**, *11*, 4120–4131.
[3] a) A. Chakrabartty, T. Kortemme, R. L. Baldwin, *Protein Sci.* **1994**, *3*, 834–853; b) S. H. Park, W. Shalongo, E. Stellwagen, *Biochemistry* **1993**, *32*, 7038–7053.
[4] a) P. Wallimann, R. J. Kennedy, D. S. Kemp, *Angew. Chem.* **1999**, *111*, 1377–1379; *Angew. Chem. Int. Ed.* **1999**, *38*, 1291–1292; b) R. J. Kennedy, K. Y. Tsang, D. S. Kemp, *J. Am. Chem. Soc.* **2002**, *124*, 934–944; c) P. Wallimann, R. J. Kennedy, J. S. Miller, W. Shalongo, D. S. Kemp, *J. Am. Chem. Soc.* **2003**, *125*, 1203–1220.
[5] a) J. S. Miller, R. J. Kennedy, D. S. Kemp, *Biochemistry* **2001**, *40*, 305–309; b) S. Deechongkit, R. J. Kennedy, K.-Y. Tsang, P. Renold, D. S. Kemp, *Tetrahedron Lett.* **2000**, *41*, 9679–9683; c) W. Maison, E. Arce, P. Renold, R. J. Kennedy, D. S. Kemp, *J. Am. Chem. Soc.* **2001**, *123*, 10245–10254; d) J. S. Miller, R. J. Kennedy, D. S. Kemp, *J. Am. Chem. Soc.* **2002**, *124*,945–962.
[6] For details, see the Experimental Section and the Supporting Information.
[7] a) D. S. Kemp, S. L. Oslick, T. J. Allen, *J. Am. Chem. Soc.* **1995**, *117*, 6641–6657; b) H. A. Nagarajaram, R. Sowdhamini, C. Ramakrishnan, P. Balaram, *FEBS Lett.* **1993**, *321*, 79–83.
[8] S. W. Englander, N. R. Kallenbach, *Q. Rev. Biophys.* **1984**, *16*, 621–655.
[9] If $FH_n = 0.1$, the FHs increase monotonically for Ala_9 to Ala_{24} within the following range: $0.90 \le FH \le 0.96$; if $FH_n = 0.7$, the corresponding range is $0.94 \le FH \le 0.98$.
[10] Since the value of a PF_i that exceeds 100 largely reflects the trace of peptide that assumes a nonhelical conformation at site i, a large measurement error in PF_i corresponds to a much smaller error in FH_i. Thus $90 \le PF_i \le 110$ corresponds to $0.988 \le FH_i \le 0.992$ (see the Supporting Information).
[11] a) N. A. Besley, J. D. Hirst, *J. Am. Chem. Soc.* **1999**, *121*, 9636–9644; b)P. Luo, R. L. Baldwin, *Biochemistry*, **1997**, *36*, 8413–8421.
[12] Units of $deg\,cm^2\,dm^{-1}$; average residuals for these regressions over the λ range are $\pm 3\%$ of $[\theta_{Molar}]_{\lambda,n}$ and lie within the measurement error of ca 3–5% estimated from CD spectra,[5d] $14 < n < 20$, for a series $WK_mInp_2{}^tL\text{-}Ala_n\text{-}{}^tLInp_2K_m\text{-}NH_2$.
[13] a) P. J. Gans, P. C. Lyu, M. C. Manning, R. W. Woody, N. R. Kallenbach, *Biopolymers* **1991**, *31*, 1605–1614; b) J. M. Scholtz, H. Qian, E. J. York, J. M. Stewart, R. L. Baldwin, *Biopolymers* **1991**, *31*, 1463–1470.

