J|A|C|S
ARTICLES
Published on Web 10/02/2004

The Helical Alanine Controversy: An $(Ala)_6$ Insertion Dramatically Increases Helicity

Jasper C. Lin, Bipasha Barua, and Niels H. Andersen*

Contribution from the Department of Chemistry, University of Washington, Seattle, Washington 98195

Received May 10, 2004; E-mail: andersen@chem.washington.edu

Abstract: Employing chemical shift melts and hydrogen/deuterium exchange NMR techniques, we have determined the stabilization of the Trp-cage miniprotein due to multiple alanine insertions within the N-terminal α-helix. Alanine is shown to be uniquely helix-stabilizing and this stabilization is reflected in the global fold stability of the Trp-cage. The associated free energy change per alanine can be utilized to calculate the alanine propagation value. From the Lifson−Roig formulation, the calculated value (w_{Ala} = 1.6) is comparable to those obtained for short, solubilized, alanine-rich helices and is much larger than the values obtained by prior host−guest techniques or in N-terminally templated helices and peptides bearing long contiguous strings of alanines with no capping or solubilizing units present.

Introduction

The classic studies of helix−coil equilibria[1,2] indicated that polyalanine sequences were only slightly helix-favoring with alanine propagation values (*s* in Zimm−Bragg theory or *w* in the Lifson−Roig formulation[3,4a]) between 1.03 and 1.15. In contrast, studies of shorter designed α helices in a number of laboratories during the 1990s suggested propagation values between 1.5 and 1.8 for alanine,[4] making it the residue with by far the largest helix propensity. The designed heteropolymers were alanine-rich but bore an N-terminal acetyl group (or an N-terminal sequence identified as an N-cap) and had hydrophilic residues (usually lysine or arginine) added to ensure solubility. Scheraga et al.[2] attribute the unexpected helicity of such short peptides to the hydrophilic side chains, which result in backbone desolvation and thus favor the intramolecularly H-bonded helical state. Is, or is not, alanine uniquely favorable for helix formation?[5]

The Kemp laboratory provided a number of systems[7] designed to test the helix propensity of alanine and derived propagation values similar to those of Scheraga. The first systems (Ac-He1-Ala_n-OH) were short, N-templated Ala-rich peptides.[7a] In this series, the helix populations were measured by an NMR probe within the helix cap (Ac-He1). These measures suggested alanine propagation values consistent with Scheraga's values, rather than those from solubilized peptide helices. Later, Kemp designed peptides in which the helix is restricted to a string of contiguous alanines insulated from solubilizing polylysine segments and potential capping interactions by unnatural residues: $WK_4(Inp)_2Z(A)_nZ(Inp)_2K_4$ (Inp = isonipecotic acid, Z = *tert*-Leu).[7c] This strategy also facilitated the examination of considerably longer strings of alanines (up to $n = 28$). CD measures of helicity indicated Ala propagation values on the order of 1.1 for $n = 12$ but values approaching 1.35 for longer strings. Alternative explanations for some of Kemp's results have been proposed.[8] We note another here: within the context of Lifson−Roig theory (see Materials and Methods), the low helicities observed for short, insulated A_n runs could be attributed to very small (≪1) N- and/or C-capping constants for *tert*-Leu. The increasing apparent propagation for longer, uninterrupted A_n runs would imply that Ala still has small capping and nucleation constants but the former are larger than those of *tert*-Leu. However, neither this explanation nor the other rationales presented in the literature resolve the "alanine controversy". In the meantime, position-specific helix propensities continue to be determined by use of solubilized alanine-

(1) (a) Platzer, K. E. B.; Ananthanarayanan, V. S.; Andreatta, R. H.; Scheraga, H. A. *Macromolecules* **1972**, *5*, 177−187. (b) Scheraga, H. A. *Pure Applied Chem.* **1973**, *36*, 1−8. (c) Wojcik, J.; Altmann, K. H.; Scheraga, H. A. *Biopolymers* **1990**, *30*, 121−134.
(2) Scheraga, H. A.; Vila, J. A.; Ripoll, D. R. *Biophys. Chem.* **2002**, *101*−*102*, 255−265.
(3) (a) Lifson, S.; Roig, A. *J. Chem. Phys.* **1961**, *34*, 1963−1974. (b) Doig, A. J.; Chakrabartty, A.; Klingler, T. M.; Baldwin, R. L. *Biochemistry* **1994**, *33*, 3396−3403.
(4) (a) Andersen, N. H.; Tong, H. *Protein Sci.* **1997**, *6*, 1920−1936. (b) Doig, A. J.; Baldwin, R. L. *Protein Sci.* **1995**, *4*, 1325−1336. (c) Shalongo, W.; Stellwagen, E. *Protein Sci.* **1995**, *4*, 1161−1166. (d) Yang, J.; Spek, E. J.; Gong, Y.; Zhou, H.; Kallenbach, N. R. *Protein Sci.* **1997**, *6*, 1264−1272.
(5) Helix propensities of different residues have been rationalized in terms of side-chain entropy differences between the random coil and helical state.[6a−c] Helicity increases for Xaa → Ala mutations can be attributed to a reduction in ΔS_U. In addition, Ala's lack of a bulky side chain allows for stabilization of H-bonding interactions with water in the helical state.[6d]
(6) (a) Creamer, T. P.; Rose, G. D. *Proteins: Struct., Funct., Genet.* **1994**, *19*, 85−97. (b) Blaber, M.; Zhang, X. J.; Matthews, B. W. *Science* **1993**, *260*, 1637−1640. (c) Doig, A. J.; Sternberg, M. J. E. *Protein Sci.* **1995**, *4*, 2247−2251. (d) Luo, P.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 4930−4935.
(7) (a) Kemp, D. S.; Oslick, S. L.; Allen, T. J. *J. Am. Chem. Soc.* **1996**, *118*, 4249−4255. (b) Groebke, K.; Renold, P.; Tsang, K.-Y.; Allen, T. J.; McClure, K. F.; Kemp, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 4025−4029. (c) Miller, J. S.; Kennedy, R. J.; Kemp, D. S. *J. Am. Chem. Soc.* **2002**, *124*, 945−962.
(8) (a) Spek, E. J.; Olson, C. A.; Shi, Z.; Kallenbach, N. R. *J. Am. Chem. Soc.* **1999**, *121*, 5571−5572. (b) Rohl, C. A.; Fiori, W.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3682−3687.

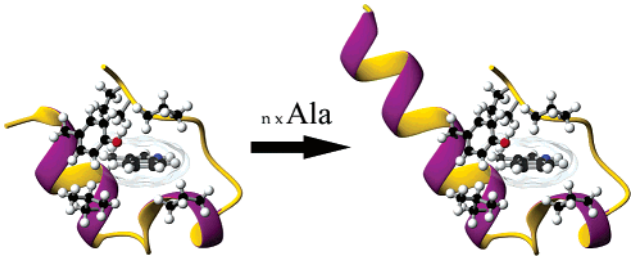

Figure 1. Trp-cage structure and the extension of the N-terminal helix by the insertion of multiple alanines.

rich peptides.[9] These studies, and those of Kemp and Scheraga, rely on CD spectroscopy to define mutation-induced changes in helicity even though there is considerable controversy regarding the $[\theta]_{222}$ value for 100% helicity. For example, Kemp's analysis of the CD data for their insulated $(Ala)_x$ helices requires the $[\theta]_{222}$ value to be 61 000 deg/res·dmol for 100% helicity,[7c] rather than the previously reported 44 000 deg/res·dmol value.[10]

Studies of mutational effects on protein stability rely on the use of backbone amide hydrogen/deuterium exchange studies[11] or changes in melting temperature (T_m, defined as the temperature at which the protein is 50% unfolded). For the former, the equilibrium mole fraction of the unfolded state fraction is given by the reciprocal of the exchange protection factor (PF) for those H_Ns that exchange only upon global unfolding. In this paper, we employ NMR experiments on a miniprotein scaffold, the Trp-cage (which has an N-terminal α helix),[12] to quantify the effect of alanine on helix formation. The Trp-cage and the experimental strategy, the insertion of alanines at the N-terminus, are illustrated in Figure 1. Since the Trp-cage is in fast exchange with its unfolded state, chemical shift melts can be performed and both the helix and global unfolding melting temperatures can be determined. For Trp-cage constructs, PFs can also be readily obtained by NMR experiments.[12] In this system, we find that each alanine insertion increases fold stability to a degree consistent with a propagation value of 1.6. In addition, the present study suggests that an $(Ala)_7$ unit in a helix displays CD values concordant with a $[\theta]_{222}$ value of −44 000 deg/res·dmol for 100% helicity.

Materials and Methods

Statistical Mechanical Treatments of Helix/Coil Transitions. Both the Zimm−Bragg and Lifson−Roig formulations are in common usage. The statistical weight of an α helix spanning from residue i to j (W_{ij}) is given by the following equations:

$$W_{ij} = \sigma(s_{i+2})\prod_{i+3}^{j} s_k = \upsilon^2 \prod_{i+2}^{j} w_k = \upsilon_N(i)\upsilon_C(j)\prod_{i+1}^{j-1} w_k \qquad (1)$$

Zimm−Bragg Lifson−Roig

The second Lifson−Roig formulation includes specific N- and C-terminal nucleation values rather than a generalized value, $\upsilon = \sigma^{1/2}$. As a result, this formalism can be further expanded to include capping effects by the N − 1 and C + 1 residues—for example[4a]

$$W_{ij} = \upsilon_N(i)\upsilon_C(j)\prod_{i+1}^{j-1} w_k \qquad (2)$$

The variable $\upsilon_N(i)$ is the product of the intrinsic probability of residue i being an N-terminal residue of an α helix, $\upsilon_N^0(i)$, and the N-capping constant, N_{i-1} , of the preceding residue (i − 1). Likewise, $\upsilon_C(j)$ is given by $\upsilon_C(j) = C_{j+1}[\upsilon_C^0(j)]$.

For short helices, certainly for those having a span of less than 16 residues, the single sequence approximation can be applied. In such cases, W_{ij} from eq 2 can be viewed as a helix formation equilibrium constant and the stability of the helical state (the free energy cost of unfolding, ΔG_U) includes RT ln W_{ij} terms for the full helical sequence and its frayed versions. When, as is the case for the N-terminal helix of the Trp-cage, there are specific N-capping and C-capping sites, ΔG_U is well approximated by RT ln W_{ij}. Trp-cage formation, which corresponds to the ordering of C-terminal residues (that are not part of the helix) about the indole ring, serves as a highly effective and constant C-cap. In the Ala-insertion series, the N-terminal Gly is the constant N-cap. As a result, the $\Delta\Delta G_U$ associated with residue insertions between the caps is RT ln (Πw), with the product including only the propagation values for the added residues.

Materials and Spectroscopic Methods: (A) Peptide Synthesis and Characterization. Peptides were synthesized on an Applied Biosystems 433A synthesizer employing standard Fmoc solid-state peptide synthesis methods. Wang resins, preloaded with the C-terminal amino acid, were used for the syntheses. Peptides were cleaved by a 90:5:5 trifluoroacetic acid (TFA)/triisopropylsilane/water mixture. The cleaved peptides were then purified by reverse-phase HPLC on a Varian C18 preparatory-scale column with gradients of water and acetonitrile spiked with 0.1% and 0.085% TFA, respectively. Collected fractions were lyophilized and their identity and molecular weight were confirmed by use of a Bruker Esquire ion trap mass spectrometer. The following peptide sequences are examined in the present study:

peptide designation	sequence	observed m/z of $(M + 2H)^{2+}$
TC5b	NLYIQ WLKDG GPSSG RPPPS	1085.6
TC8a	NLYAQ WLKDG GPSSG RPPPS	1064.5
TC9b	NAYAQ WLKDG GPSSG RPPPS	1044.5
TC10b	DAYAQ WLKDG GPSSG RPPPS	1044.9
TC11a	AcAYAQ WLKDG GPSSG RPPPS	1007.7
TC12a	AYAQ WLKDG GPSSG RPPPS	986.0
TC12b	NYAQ WLKDG GPSSG RPPPS	1008.8
TC11b1	GAYAQ WLKDG GPSSG RPPPS	1016.0
TC11b3	GAAAYAQ WLKDG GPSSG RPPPS	1086.8
TC11b4	GAAAAYAQ WLKDG GPSSG RPPPS	1121.6
TC11b5	GAAAAAYAQ WLKDG GPSSG RPPPS	1157.0
TC11b5(A-1K)	GAAKAAYAQ WLKDG GPSSG RPPPS	1185.8
TC11b7	GAAAAAAAYAQ WLKDG GPSSG RPPPS	1228.2

The NMR shift data for TC5b are taken from previous studies;[12] the other peptides displayed NMR spectra in full accord with a Trp-cage structure. The details and an NMR structure ensemble calculation for TC10b will appear in a full account of the optimization of the Trp-cage fold.

(B) CD Spectroscopy. Circular dichroism stock solutions were prepared by dissolving approximate amounts of peptide in 10 mM aqueous pH 7 phosphate buffer to make a solution of about 200 μM. The concentration of the stock solution was measured by the UV absorption of tyrosine and tryptophan (ϵ = 1280 M^{-1} cm^{-1} and 5580 M^{-1} cm^{-1}, respectively, at 280 nm). CD samples were diluted appropriately to obtain 30 μM solutions of the peptide in buffer. Spectra were recorded on a Jasco J715 spectropolarimeter as previously

(9) (a) Sun, J. E.; Penel, S.; Doig, A. J. *Protein Sci.* **2000**, *9*, 750−754. (b) Cochran, D. A. E.; Penel, S.; Doig, A. J. *Protein Sci.* **2001**, *10*, 463−470. (c) Cochran, D. A. E.; Doig, A. J. *Protein Sci.* **2001**, *10*, 1305−1311. (d) Iqbalsyah, T. M.; Doig, A. J. *Protein Sci.* **2004**, *13*, 32−39.
(10) (a) Luo, P.; Baldwin, R. L. *Biochemistry* **1997**, *36*, 8413−8421. (b) Shalongo, W.; Stellwagen, E. *Proteins: Struct., Funct., Genet.* **1997**, *28*, 467−480.
(11) Bai, Y.; Milne, J. S.; Mayne, L.; Englander, S. W. *Proteins: Struct., Funct., Genet.* **1993**, *17*, 75−86.
(12) Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. *Nat. Struct. Biol.* **2002**, *9*, 425−430.

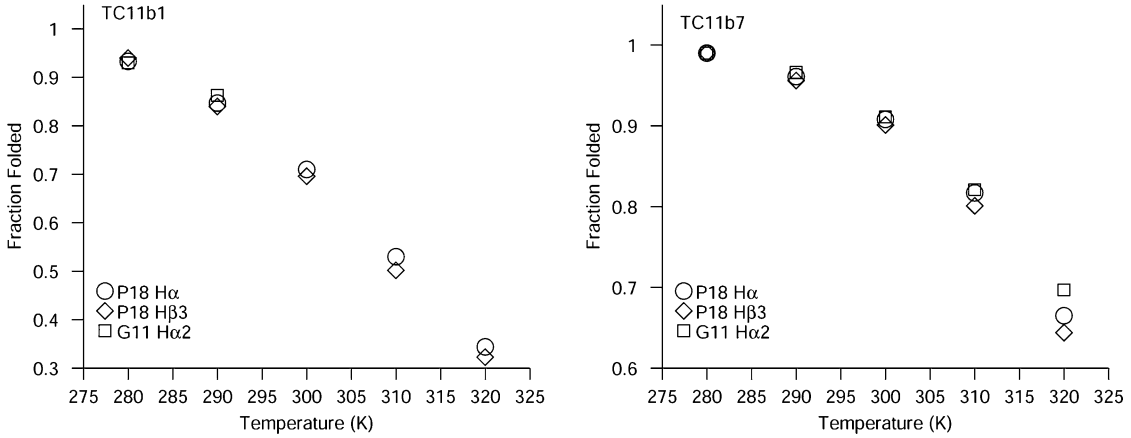

Figure 2. Agreement between different cage measures is illustrated for TC11b1 and TC11b7. The peak for G11 Hα2 in TC11b1 was too broad to detect in the NMR spectra taken at 300 K and above. This is the result of exchange broadening due to microsecond time constants in the folding dynamics.

described.[13,14] For melting experiments, spectra were acquired at 10 °C increments from 5 to 95 °C. The spectrophotometer was equilibrated at each temperature for 5 min before data acquisition.

(C) NMR Spectroscopy. NMR experiments were performed on a Bruker DRX tuned at 500 MHz. A combination of two 2D NMR techniques, total correlation spectroscopy (TOCSY) and nuclear Overhauser effect spectroscopy (NOESY), were used to assign all H_N and Hα resonances. A 60 ms MLEV-17 spin-lock was employed for TOCSY experiments and a 150 ms mixing time for the NOESY experiments. The samples consisted of 1−2 mM peptide in pH 6.5−7.0 phosphate buffer containing 10% D_2O. Sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) was used as the internal chemical shift reference and set to 0 ppm for all conditions. The temperature was changed from 280 to 320 or 325 K in 10 K or smaller increments.

(D) Chemical Shift Deviations and Shift Melting Curves. All chemical shift data are expressed as chemical shift deviations (CSDs) from the random coil value of the specific residues. With the exception of Gly[11], the Hα statistical coil values are the sequence-corrected ones from our current automated program.[15] This program is also available via the Internet (http://andersen.chem.washington.edu/CSDb). The unfolded reference values for Gly[11]-Hα2, Pro[18]-Hβ3, Pro[19]-H$_{\delta 2}$, and Pro[19]-H$_{\delta 3}$ are 4.02, 2.29, 3.59, and 3.74 ppm, respectively.

As previously noted,[15] CSDs can be converted to folded-fraction (f_F) estimates when reliable reference values exist for the fully folded state. These measures of the fold fraction are quite accurate (estimated at ±0.04) through the range $f_F = 0.88-0.30$. In this range, fractional CSDs can thus provide ΔG_U estimates with errors of ±0.4 kJ/mol or less. Outside this range, errors in the applicability of 100% folded reference shifts from more stable analogues or the statistical coil shifts, respectively, compromise this method of fold population estimation. The accuracy of the ΔG_U estimates decreases in the range $f_F = 0.95-0.89$.

NMR melting curves are calculated from the f_F values derived from CSDs. The most stable analogue in each series (TC10b for the 20-mers, TC11b7 for the TC11b*x* series with $x > 1$) is used as the reference. In each case, the reference was considered to be 99% folded at 280 K;[16] the f_F value for each peptide is calculated from the fractional CSD: fraction folded $= 0.99(CSD_{obs}/CSD_{ref})$. The melting temperature, T_m, is defined as the temperature at which the peptide is 50% folded and is, ideally, calculated by a curve fit or linear interpolation including points above T_m. Since our high-temperature limit for the NMR experiments was 320 or 325 K, T_m values greater than 52 °C represent extrapolations. An alternative analysis based on $T_{70\%}$ values appears in the Supporting Information. The agreement between the different individual measures of cage population for TC11b1 and TC11b7 is shown in Figure 2. As has been previously noted for TC5b,[12] chemical shift measures of structuring from different residues, including Hα of Leu[7] and Arg[16] as well as Hβ3 of Pro[12], are in substantial agreement in these new series of Trp-cage analogues.

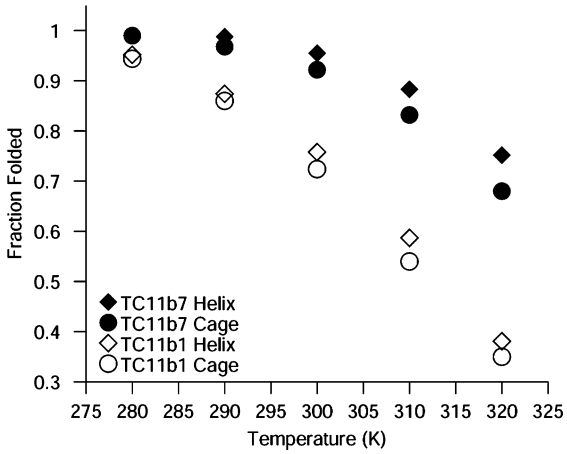

Figure 3. Chemical shift fraction-folded measures for TC11b1 and TC11b7.

Since we could not track Gly[11]-Hα2 for many of the species due to extensive line broadening as the unfolded fraction increases, the NMR cage melts reported herein are, for conformity, based on the sum of the CSDs of P18-Hα, P18-Hβ3, P19-Hδ2, and P19-Hδ3. When possible, we confirmed these using a more inclusive measure of fold population that included the Gly[11]-Hα2 and Leu[7]-Hα CSDs (Leu[7]-Hα is within the helical span but also has a large ring current effect due to cage formation). The measure of helix melting was based on the sum of the CSDs for the α methines of Y3, Q5, W6, and K8, which are common to all of the constructs. It is of particular note that the ΣCSD for P18-Hα, P18-Hβ3, P19-Hδ2, and P19-δ3 are essentially identical for the two reference compounds: −5.66 ppm for the most stable 20-mer (TC10b) and −5.68 ppm for TC11b7. The agreement is nearly as good for the helix Hα sum: −1.50 (for TC10b) and −1.44 ppm (for TC11b7). The correlation between helix and cage melting measures for the two peptides shown in Figure 2 appears as Figure 3.

(E) NH Exchange Studies. Deuterium exchange experiments were used to determine protection factors at specific residues. The pH of a 30 mM phosphate buffer was adjusted by use of phosphoric acid or

(13) Barua, B.; Andersen, N. H. *Lett. Pept. Sci.* **2002**, *8*, 221−226.
(14) Andersen, N. H.; Brodsky, Y.; Neidigh, J. W.; Prickett, K. S. *Bioorg. Med. Chem.* **2002**, *10*, 79−85.
(15) Fesinmeyer, R. M.; Hudson, F. M.; Andersen, N. H. *J. Am. Chem. Soc.* **2004**, *126*, 7238−7243.
(16) From Figure 2 it is apparent that the chemical shifts of TC11b7 are leveling out toward a low temperature value. This sigmoidal behavior is not expected until the folded fraction is greater than 0.96. This is also observed for TC10b and the limiting values are essentially identical. The latter indicates that the same fully folded state is forming. The specific choice of 99% folded rather than some other value between 96% and 100% is based on the extent of NH exchange protection (vide infra) observed within the helix.

sodium hydroxide. The peptide was weighed and added to the solvent to make a 1−2 mM solution. The resulting mixture was frozen in a dry ice/2-propanol bath for one or more hours and lyophilized for at least 24 h. Immediately before acquisition, enough D_2O (containing DSS as a chemical shift standard) is added to obtain the desired peptide concentration. All exchange experiments were run at 280 K. Eight-, 16-, or 64-scan 1D proton NMR spectra were taken at different time intervals. Afterward, the pD values were measured with a microelectrode connected to an Orion 410A pH meter; these values are TC11b3 (5.0, 6.4), TC11b4 (6.5), TC11b5 (5.5, 6.7), TC11b5-A-1K (5.88), and TC11b7 (6.5). Both CD and NMR melting studies indicate that there is no change in fold stability over this pH range (data not shown).

Protection factors were determined for the NHs that were the slowest to exchange. Depending on the construct and the degree of NH signal overlap in 1D spectra, this included the following four common signals: H_N of W6, L7, and G11 and the side-chain NH of W6. For some constructs, we were also able to measure exchange rates for H_N of Q5 and D9. Throughout, the S20 backbone amide resonances displayed slow exchange rates attributed to the negative charge at the C-terminus. The spectra were processed with the Mestre-C program (http://www.mestrec.com), and the peak heights were measured and normalized with respect to a nonexchanging peak in the same spectrum. The heights were then plotted versus time in order to ascertain the exponential decay equation. The experimental rate constant is taken from the fitted data. The exchange rate constants for the unfolded state were calculated by use of the appropriate Molday factors.[11] The ratio of the two rate constants (k_{rc}/k_{exp}) equals the protection factor. This procedure has been described in detail for both TC5b[12] and exendin-4.[17] The PFs for Ser[20] varied from 0.6 to 2.6 (≈1) and are consistent with the unprotected status of the C-terminus. This serves as additional validation of the intrinsic exchange rate values used at each pD value. For the other slow-exchanging sites, ΔG_U(280 K) values were determined from the exchange protection factors: $\Delta G_U = RT \ln (PF)$. This relationship is commonly used for proteins; one example is by Fezoui et al.[18] The mean values of ΔG_U are given in Table 2; when two determinations (at slightly different pD values) are available for any site of a peptide, both are used in calculating the mean.

Results

Trp-cage miniproteins exhibit large chemical shift deviations due to the tight packing of residues around the central tryptophan residue. In the present study, we employ the sum of the ring current shifts for C-terminal residues (P18Hα, P18-Hβ3, P19-Hδ2, and P19-Hδ3) to measure cage formation and the backbone H_Ns of Q5, W6, L7, D9, and G11 and the side-chain H_N of W6 to calculate the overall protection factor (PF). The temperature dependence of the sum of the upfield Hα CSDs of residues Y3, Q5, W6, and K8 is used to measure helix melting. All T_m values reported herein are based on chemical shift melts (see Materials and Methods). The folded reference values are defined separately for the 20-residue constructs and the alanine insertion series (TC11b*x*, $x > 1$) by using the CSD values at 280 K of TC10b (99% folded) and TC11b7 (99% folded), respectively, as the reference (see Materials and Methods[16]). This calibration method is fully supported by exchange protection data;[19] further, the T_m values vary less than 1.7 °C if the reference compounds are taken as 96−100% folded. T_m values from the shift melting curves are reported to the nearest 0.5 °C in the tables; however, all propagation value estimates allow for errors as large as 1.5 °C.

Two things are required to validate the experimental strategy inherent in Figure 1: the Trp-cage unfolding transition needs to be cooperative and ΔG_U for global unfolding must reflect the intrinsic stability of the helical segment. Prior studies indicate that residues 2−9 of the Trp-cage are not helical in the absence of the C-capping associated with the Trp-cage. Both truncated analogues and the W6F analogue display very modest net helicities (<30% in all cases).[13] Mutations near the N-terminus were used to ascertain whether global fold stability (as measured by the temperature dependence of chemical shifts in the Pro[18,19] unit) was a reflection of helix stability (Table 1).

Table 1. Thermodynamic Parameters of Trp Cage Mutants

TC	partial sequence	ΔG_U[a]	cage-T_m[b]
5b	NLYIQ WLKDG GP-	9.2[c]	42
8a	NLYAQ WLKDG GP-		45.5
9b	NAYAQ WLKDG GP-		51.5
10b	DAYAQ WLKDG GP-	13.2[c]	57
11a	AcAYAQ WLKDG GP-		46
11b	GAYAQ WLKDG GP-		39
12a	AYAQ WLKDG GP-	−2.4[d]	≪0
12b	NYAQ WLKDG GP-	4.0[d]	33

[a] Kilojoules per mole at 280 K. [b] Degrees Celsius. [c] Determined from NH protection factors. [d] The fraction folded at 280 K was estimated from the fractional CSDs.

This series of analogues clearly established that global fold stability reflects the intrinsic helix stability, and the alanine mutation effects[5] suggest helix propensities such as those observed in short, designed helices. Furthermore, it is apparent that an N-capping residue is essential for both helix and global fold stability: deletion of the N-cap abrogates Trp-cage folding but it is partially rescued by replacing Ala[2] with Asn. The stabilities of the Asp-, Asn-, acetyl-, and glycine-capped Trp-cages reflect the previously determined N-capping constants for these residues.[4]

Turning to the alanine insertion experiments, residues added between the N-cap and Ala[2] are too far away from the remainder of the Trp-cage structure to have any tertiary interactions (confirmed by NOESY experiments); thus, any increases in stability should reflect only intrinsic helix propensities. We selected the glycine-capped 20-mer as the starting point since we expected it to be sensitive to the helix propensities of added residues. The series of peptides prepared (GA_xYAQ WLKDG GPSSG RPPPS) are designated TC11b*x*, where $x = 1$, 3, 4, 5, and 7. CD spectroscopy revealed a steady increase in both the magnitude (at the low temperature) of the ca. 222 nm minimum and the apparent melting temperature upon serial addition of alanine residues (see Figure 4).

Deriving thermodynamic parameters for the folding transition from this CD data is, however, problematic. It is essential to have, or fit, the temperature dependence of both the folded and unfolded states in order to extract the temperature dependence of the fold population. We have determined the CD spectrum of the unfolded state and its temperature dependence for several Trp-cage constructs in 7 M Gdm^+Cl^-: $[\theta]_{222}$ (deg·cm^2/res·dmol) = −900 − 29*T* (temperature in degrees Celsius). But the 100% folded value and its temperature dependence is

(17) Neidigh, J. W.; Fesinmeyer, R. M.; Prickett, K. S.; Andersen, N. H. *Biochemistry* **2001**, *40*, 13188−13200.
(18) Fezoui, Y.; Braswell, E. H.; Xian, W.; Osterhout, J. J. *Biochemistry* **1999**, *38*, 2796−2804.
(19) The protection factors for the NH of Leu[7], reflecting the residue 7 → 3 H-bond in the helix, are 50.5 (TC10b) and 170 (TC11b7), supporting a 98−99.5% folded assumption. The greater degree of protection in TC11b7, which has protection factors of ca. 500 at Gln[5] and Trp[6], may reflect a small but significant extent of helix formation in the ensemble of structures that lack the Trp-cage (the unfolded state).

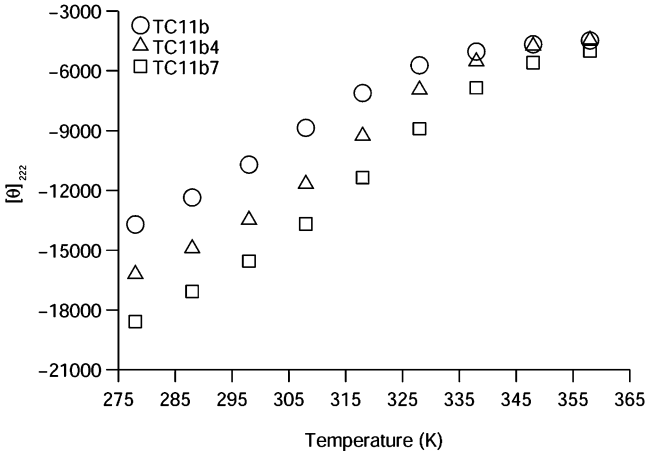

Figure 4. CD melts of TC11b1, -b4, and -b7 monitored at 222 nm.

difficult to predict or extract, particularly since it will be different for each species when the helix is extended. We can derive "T_m estimates" by assuming a sequence-independent unfolded-state spectrum and that NMR measures of fraction folded at 290 K apply. In all cases, CD melting temperatures within 3 °C of the NMR estimates for the helix melting point can be obtained.

While the CD data reveal a steady increase in helix stability, we turned to NMR melting curves and protection factors for quantifying the effect. For most of the TC11bx species, ΔG_U(280 K) values (Table 2) were determined from exchange protection factors: $\Delta G_U = RT \ln (PF)$. In the case of TC11b1, for which the H_N peaks decay too fast at pH > 6 for our rate measurement protocol, the linear relationship between $\Delta\Delta G_U$ and both ΔT_m and $\Delta T_{70\%}$ for Trp cage mutants (see Supporting Information) provided an indirect determination of ΔG_U.

Table 2. Thermodynamic Parameters of TC11b*x*[a]

TC	partial sequence	ΔG_U[b]	cage[c]	helix[c]
	−5 −1 1 5			
11b1	GAYAG	7.3 ± 0.5[d]	39	41
11b3	GA AAYAQ	9.7 ± 0.7[e]	43	46
11b4	GAA AAYAQ	11.0 ± 1.1[e]	50.5	55
11b5	GAAA AAYAQ	11.6 ± 0.8[e]	53.5	57
11b5 A-1K	GAAK AAYAQ	10.7 ± 0.8[e]	47.5	51.5
11b7	G AAAAA AAYAQ	14.1 ± 0.6[e]	59.5	66

[a] *x* denotes the number of Ala residues between the N-terminal Gly and Tyr[3]. The common sequence after Q (Gln[5]) is WLKDG GPSSG RPPPS. [b] Kilojoules per mole at 280 K. [c] T_m in degrees Celsius based on chemical shift melts. [d] Derived from the linear relationship between cage melting temperatures and ΔG_U for Trp-cage species. [e] Based on single or multiple PF determinations for four or more NH sites.

Figure 5 plots the change in ΔG_U observed for serial alanine insertions. A compelling linear correlation is observed. We turned to the Lifson−Roig formalism[4a] (see Materials and Methods) for a derivation of the propagation parameter of alanine. Since the capping and nucleation interactions are constant in the present alanine insertion series (TC11b*x*), $\Delta\Delta G_U = \Delta x(RT \ln w)$ applies, where Δx is the number of alanines added.

From Figure 5, the slope, $m = RT \ln w$, corresponds to an alanine propagation value of 1.62 ± 0.11, which is in agreement with the values reported for short, designed Ala-rich peptides.[4]

Discussion

The CD spectral changes upon $(Ala)_x$ addition can provide the CD signature of the added residues. Figure 6A shows the

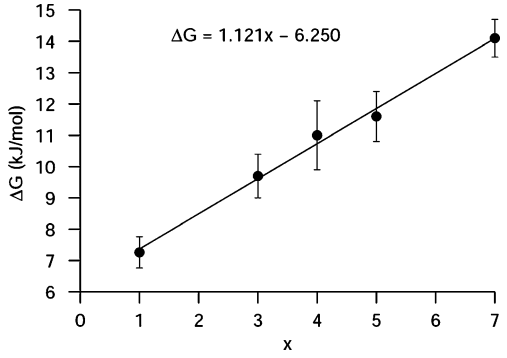

Figure 5. Global fold stability (ΔG_U) versus *x* for the TC11b*x* alanine addition series.

CD spectra of TC11b1 and TC11b7 in molar rather than residue-molar units. Normalizing the difference, dividing by $x = 6$, affords the per-residue spectrum of the $(Ala)_6$ insert. This spectrum (Figure 6B) is a classic helix CD signature. The $[\theta]_{222}$ value (at 5 °C) is −34 200°. This is in excellent agreement with the prediction (−33 600°) for the per-residue contribution expected for a 14-residue span of helix based on the equation of Luo et al.: $[\theta]_{222} = (1 - 3/N_r)[-44\,000 + 250T$ (degrees Celsius)].[10a]

Since the CD melting data (Figure 4) were unlikely to provide a completely satisfying measure of the alanine propagation value and the most dependable measures of global folding are in the C-terminal portion of the sequence (reflecting the ring current shifts in the Trp-cage), the degree of cooperativity of the folding transition is a concern. The agreement between chemical shift melts for the helical sites versus the ring-current-shifted sites in the cage is quite good (Figure 3). The helix melting points are generally higher and the difference increases slightly with the length of the helix. In the case of the longest helix (TC11b7) there is some evidence for partial helicity in the unfolded state.[19] Figure 5, based on NH protection data from both the helix and cage portions of the sequence, yielded $w_{Ala} = 1.62 \pm 0.11$. An alternative analysis (detailed in the Supporting Information) based on the changes in the melting temperature of the cage residue sites, measured as $\Delta T_{70\%}$, affords exactly the same w_{Ala} value. $T_{70\%}$ values are available for all TC11b*x* species without the need for extrapolation.

As is evident in the error bars in Figure 5, there is quite a large variation in the PFs found at different sites. Comparisons at specific sites can also provide the $\Delta\Delta G_U$ increment per alanine addition. If we restrict the analysis to Trp[6] and Leu[7] amide sites, which report on helix stability, the w_{Ala} values are 1.64 and 1.52, respectively. To obtain a more precise measure of the $\Delta\Delta G_U$ increment per alanine addition, we also calculated the protection factor ratios at all individual NH sites for all additions of two or more alanines. These protection factor ratios ($n = 13$) afforded w_{Ala} in the range 1.32−1.75 with a consensus value of 1.55. Thus, all $\Delta\Delta G_U$ determinations, whether based on PFs or changes in T_m and whether based on probes in or remote from the helix, provide w_{Ala} estimates that are much larger than the value (1.1) reported by Kemp[7b] for $(Ala)_x$ units with $x \leq 12$. Rather, the values are in full accord with previous studies[4] on solvated, alanine-rich, helical peptides.

The extended Trp-cage should also be suitable for determining the propagation values of other residues; one particularly pertinent example is included in the present study. Scheraga et

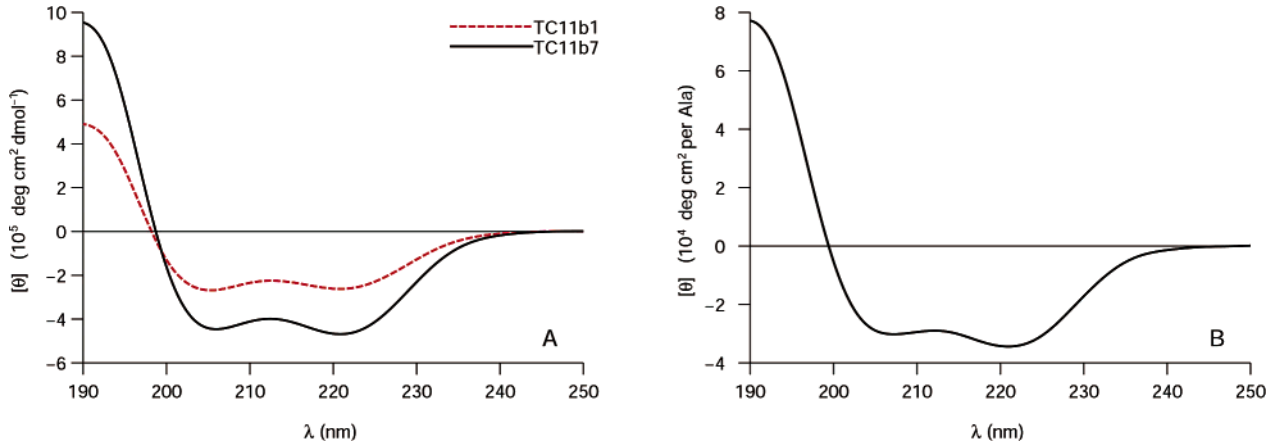

Figure 6. (A) CD spectra of TC11b1 and TC11b7 in molar ellipticity. As expected, TC11b7 exhibits a more helical signature due to added residues of the longer helix. (B) Difference spectra between TC11b1 and TC11b7, divided by the number of alanines added (6), results in a per-alanine spectrum with a classic helix signature.

al.[2] proposed that the insertion of a lysine into a polyalanine sequence increases helix formation. When we replaced the central alanine residue of TC11b5 with a lysine, there was a decrease ($\Delta T_m = -6$ °C, Table 2) in Trp-cage stability. This effect, and the PF change associated with adding Lys to TC11b4, affords w_{Lys} estimates ranging from 0.66 to 1.04. The mean value (0.82 ± 0.13) is comparable to those (e.g., 0.89[4a]) from studies of short designed helices.

The Trp-cage scaffold has provided a new method for determining helix propagation values. This miniprotein provides a constant C-capping interaction and a fold with remote solubilizing side chains that allowed the construction of a helix bearing seven consecutive alanines without the need for inserted lysine residues. On the basis of $w_{Ala} = 1.1$, an $(Ala)_6$ insert should provide only a modest helix stabilization (1.33 kJ/mol); in the present case, a nearly 6 kJ/mol stabilization was observed. The continued use of propagation constants of ca. 1.6 and 0.85 for Ala and Lys, respectively, in peptide helicity studies is validated.

This result clearly does not change the fact that long isolated $(Ala)_x$ strings are less helical than would be expected on the basis of these propagation values and the typically used[20] nucleation constant ($\sigma = v^2 = 0.0013-0.0023$).[7c,21] Why do Kemp's systems and long isolated strings of alanines display so little net helicity? In our view, the simplest explanation is that Kemp's insulating residues, as well as alanine, have very low N- and C-capping constants. Further it is necessary to posit that the N- and C-terminal nucleation constants of alanine, for use in extended Lifson−Roig helix/coil formulations (e.g., eq 2), are lower than those of most other residues. This resolution of the "helical alanine controversy" is not particularly remarkable given that it is well established that both designed[4a,b,22a,b] and natural helices[22c−e] require at least one capping interaction.

Acknowledgment. This work was supported by grants from the NIH (GM59658) and NSF (CHE0315361).

Supporting Information Available: Three figures detailing the correlation between T_m and $T_{70\%}$ and providing an alternative derivation of w_{Ala} based exclusively on $T_{70\%}$ measures (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA047265O

(20) Nucleation constants as low as $\sigma = 0.0001$ are used by Scheraga and coworkers; see ref 1c and Sueki, M.; Lee, S.; Powers, S. P.; Denton, J. B.; Konishi, Y.; Scheraga, H. A. *Macromolecules* **1984**, *17*, 148−155.
(21) Rohl, C. A.; Chakrabarty, A.; Baldwin, R. L. *Protein Sci.* **1996**, *5*, 2623−2637.
(22) (a) Lyu, P. C.; Wemmer, D. E.; Zhou, H. X.; Pinker, R. J.; Kallenbach, N. R. *Biochemistry* **1993**, *32*, 421−425. (b) Schneider, J. P.; DeGrado, W. F. *J. Am. Chem. Soc.* **1998**, *120*, 2764−2767. (c) Presta, L. G.; Rose, G. D. *Science* **1988**, *240*, 1632−1641. (d) Aurora, R.; Srinivasan, R.; Rose, G. D. *Science* **1994**, *264*, 1126−1130. (e) Aurora, R.; Rose, G. D. *Protein Sci.* **1998**, *7*, 21−38.