

Lysine and arginine residues do not increase the helicity of alanine-rich peptide helices†

James M. Stewart, Jasper C. Lin and Niels H. Andersen*

Received (in Cambridge, UK) 28th April 2008, Accepted 8th July 2008

First published as an Advance Article on the web 8th August 2008

DOI: 10.1039/b807101b

The helix-disfavoring, versus alanine, propagation values of lysine (0.8) and arginine (1.0) residues placed centrally in an (Ala)₉ unit have been measured by ¹³C NMR.

The controversy¹ concerning whether alanine is helix indifferent² or has a uniquely high helix propensity^{1,3} has prompted studies of the roles of the solubilizing residues (Arg, Lys, and Gln) that are typically placed in Ala-rich, (AAAXA)-repeat peptide helix models⁴ used to derive residue helix propensities. It has been suggested^{2b,5} that the substantial helicity of designed Ala-rich peptide helix models is due to the solubilizing polar residues rather than a uniquely large Ala helix propensity and experimental data of Kemp and co-workers⁶ have been cited as supporting this view. However, in a more recent account from the Kemp laboratory,⁷ the helix propensity for alanine in an (Ala)₁₅ segment, $w = 1.5$ (w is the propagation constant in the Lifson–Roig helix-coil model⁸), is in agreement with that obtained for (AAAXA)-repeat peptides.^{4a,b} The higher helix propensity of alanine is supported by experimental studies confirming a decrease in helicity associated with increasing Lys content in Ala/Lys copolymers⁹ and with replacing a central Ala residue within a long continuous stretch of Ala residues by Lys.^{1b} Experimental evidence indicating that both Lys and Arg enhance helicity when placed near the C-terminus of designed helices (where there is a favorable Coulombic interaction with the helix macrodipole¹⁰) is clear-cut,¹¹ but there is still some controversy concerning the role of these residues (X) at central positions in designed helices consisting of AAAXA-repeats. The argument centers on whether polar sidechain-induced backbone desolvation increases helicity.^{5b,12} An opposing point of view, that helix backbone solvation is enthalpically favorable and does not decrease helicity, has also been published.^{3c} Conformer ensemble energy calculations by Scheraga^{5b} indicate that Ac-AAAAAKAAAA-NH₂ is 76% helical while Ac-(A)₁₀-NH₂ is less than 10% helical when simulated at the dielectric constant of water. A folding trajectory simulation (continuum solvent) for Ac-YGAAKA-(AAAKA)₂-NH₂ indicated an equilibrium helicity of 64%.¹³ MD simulations of Ala-rich sequences with and without added Arg residues have, depending on the sequence and whether a

continuum or explicit water solvent was employed, indicated both enhanced^{12b} and diminished¹⁴ helix stability due to the inserted Arg residues.

As a result, we saw a need for additional experimental data on the specific effects of Lys and Arg residues at the X positions in AAAXA-repeat peptides. To date, most experimental measures of helix propensities have relied on CD measurements. Since different investigators use CD values ($[\theta]_{222}$) for 100% helicity that can differ by nearly 50%,¹⁵ we turned to other probes to answer questions concerning helix propensities. Changes in NH exchange protection and melting data for mutations within the N-terminal helix of Trp-cage mini-proteins have provided two measures of the relative helix propensities of Lys and Ala.

GAAXAAYAQLKDGPPSSGRPPPS (X = A vs. K),

$$\Delta T_m = 6 \text{ }^\circ\text{C} \Rightarrow w(\text{Ala})/w(\text{Lys}) = 1.95 \text{ [ref. 1a]}$$

DAYAQLXLDGPPSSGRPPPS (X = A vs. K),

$$\Delta T_m = 5 \text{ }^\circ\text{C}^{16} \Rightarrow w(\text{Ala})/w(\text{Lys}) = 1.75$$

Although this was viewed as strong evidence for Lys being significantly helix destabilizing,^{1a} it still leaves the question about the effects of polar sidechains in Ala-rich peptides unanswered.

Recently ¹³C=O (¹³C') chemical shifts, which move downfield upon helix formation, have emerged as residue-specific measures of fractional helicity.^{11b,17} We have used this method to probe the effects of Lys → Ala and Arg → Lys → Ala mutations in two designed helices: Ac-WAAHA-(AARAA)₃-NH₂ and KKGG-K(AAAK)₃-GGKK-NH₂.† The first system has been extensively examined both experimentally and in MD simulations.¹⁸ The second system included multiple Lys residues at the termini to ensure solubility. (¹³C')-Alanine isotopomers were prepared in each case to provide chemical shifts for sites in each of the repeats; NMR spectra were recorded in 20 mM phosphate buffer, pH = 7.0 (10% D₂O) with ¹³C labeled urea as an internal reference. Melting curves (δ_{obs} versus T) over the range 280–340 K were obtained for each site in each peptide. The melts (*vide infra*) reveal helical structuring at low temperatures which melts out (*ca.* 80% loss of structuring shifts) by 340 K. In the case of Ac-WAAHA(AARAA)₃-NH₂, the apparent melting point (the point at which the δ_{obs} versus T slope is maximal), 305 K, is in excellent agreement with the prior study.¹⁸ Chemical shifts were converted to chemical shift deviations (CSDs)§ using the coil values and methods previously

Department of Chemistry, University of Washington, Seattle, WA 98195, USA. E-mail: andersen@chem.washington.edu

† Electronic supplementary information (ESI) available: Fig. S1 and procedure for deriving propagation values. See DOI: 10.1039/b807101b

described.^{11b,17b} The results are shown below; the ¹³C'-labeled sites are in bold. For the second system, CSD averages were used within the first two repeats.

The RAAAAXAAAAR system indicates that propagation values must be in the order $w(\text{Ala}) > w(\text{Arg}) > w(\text{Lys})$. The melting curves for X = A and K appear in Fig. 1.

	-AARAAAXAAAARA-	
X = A	3.04	2.21
X = R	2.90	2.05
X = K	2.83	1.97
	(CSDs in ppm)	

The CSD data for the KAAAAXAAAAX'AAAAK system (below) indicate that the mutational effects can be observed throughout the helical span. The diminished structuring shifts in the final repeat reflect C-terminal fraying of the helix.

	-GKAAAAXAAAAX'AAAAK-		
X, X' = K, K	1.92	2.02	1.28
X, X' = A, K	2.39	2.48	1.62
X, X' = K, A	2.11	2.35	1.66

This result is in accord with expectations based on the cooperativity inherent in helix-coil transition models. Assuming that the value of $w(\text{Ala})$ is not position dependent over the helical span examined, the observation of larger structuring shifts for -AAAAAKA- versus -AKAAAA- suggests, as

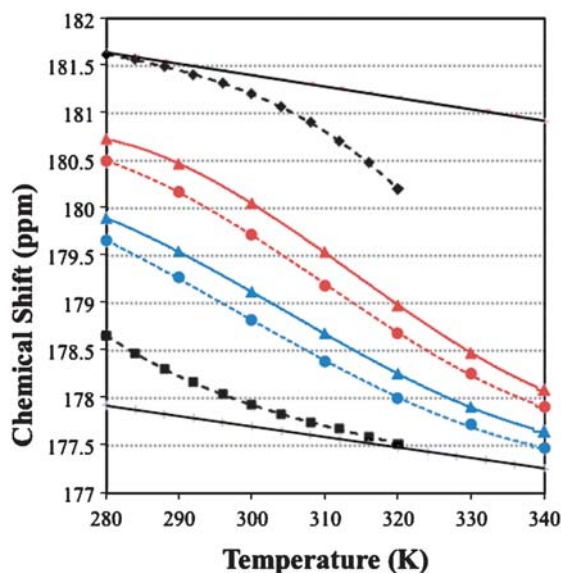


Fig. 1 Melting curves for ¹³C=O chemical shifts. The data for DAYAQWLKDGPPSSGRPPPS and NAYAQLKDGKK appear in black with the derived 100% and 0% helix lines. The data for Ac-WAAAHAAARAAAXAAARAA-NH₂ is color coded by alanine position, with the X = A data shown as ▲ connected by a solid line and X = K as ● connected by a dashed line. The 100% and 0% helix lines for -RAAAXAAAR- appear in the Electronic Supplementary Information together with the method employed to convert CSDs to helix propagation values.†

expected for a sidechain-charge-helix-macro-dipole interaction, that $w(\text{Lys})$ increases as the Lys residue moves closer to the C-terminus of the helix.

To obtain propagation values, the changes in CSDs must be converted to fractional helicity (f_H) changes, and thus, local $\Delta\Delta G_H$ values. This requires a value for the CSD expected for 100% local helicity. The Trp-cage provides a means for obtaining such an estimate. A large body of evidence exists^{1a,16} to support the assertion that a Trp-cage species of the DAYAQWLKDGPPSSGRPPPS sequence is >98% folded at 280 K. The temperature dependence of the ¹³C' CSDs for the sites in bold was measured. The corresponding coil values were obtained from the chemical shift melting behavior of a truncated species (NAYAQLKDGKK),‡ which lacks the features required for Trp-cage formation and displays $f_H = 0.18$ at the central sites at 280 K with melting complete by 320 K. For the intact Trp-cage system, a small degree of melting is observed from 300–320 K with the plot of $\delta^{13}C'$ vs. T for each site over the 280–300 K range suggesting an approach to a linear, pre-melting plot with essentially the same slope as the δ_{coil} line. Together, the two melting studies afforded the following 100% CSD estimates at 280 K: A2 (3.35), A4 (3.71), and L7 (3.25 ppm). The melting curves for A4 of the Trp-cage and the truncated species appear in Fig. 1.

Prior studies, with the $\Delta\delta^{13}C'$ for helix formation calibrated against CD estimates of f_H , had placed the 100% CSD for central residue in a helix at 3.04 ppm.^{17b} Therefore, we designed and prepared an N-capped Ala-rich peptide (GASEDE(AAAAK)₃GY-NH₂)‡ that was expected to have a very high fractional helicity with ¹³C=O units at specific alanine sites. The largest ¹³C' CSD (3.41 ppm) was observed in the first repeat, immediately after the SEDE capping box, with a 3.14 ppm CSD observed in the central repeat. Kemp has also reported data indicating that $\Delta\delta^{13}C' \geq 3.35$ ppm for a central alanine site in a helix.⁷ The prior estimate of the CSD for 100% helix is clearly too small.

As a result, we have adopted a ¹³C' CSD of 3.50 ppm as the 100% folded reference value for both protein and peptide helices. With this value, and the current $w(\text{Ala})$ value (1.54) in Helix1.^{54b} (the same value has been reported recently for helices templated by a La³⁺ binding capping/nucleation unit¹⁹), the propagation values from the data reported herein are: $w(\text{Arg}) = 0.98$ (for a central helical position), $w(\text{Lys}) = 0.80 \pm 0.05$ (for a central helical position), and $w(\text{Lys}) = 0.87 \pm 0.02$ (for locations near the helix C-terminus). While the precise w values may be disputed, there should no longer be any question about the effects of either Lys or Arg inserted in long (Ala)_n sequences. Neither of these polar residues increases helicity; quite the contrary, helicity decreases and the decrease is in proportion to the propagation constants derived in earlier reports^{4a,b} based on CD measurements for Ala-rich helical peptides.

IR studies of unlabeled and ¹³C=O-labeled (AAAAK)-repeat helices^{17b,20} support extensive solvation of the backbone amides of such helices: the amide I' bands appear at 1627–1636 and 1587–1592 cm⁻¹ (or 1598 cm⁻¹)^{20d} (¹²C=O and ¹³C=O, respectively). These are the same values as observed for exposed Ala sites in coiled coil species (which display amide I' bands at 1648–1652 and 1605–1608 cm⁻¹, ¹²C=O and ¹³C=O, respectively, for the buried sites)²¹ and

the IR evidence of differential solvation associated with the relative locations of Ala and Lys residues remains ambiguous.²² A desolvated ¹³C=O amide (1617 cm⁻¹) has, however, been observed upstream from a valine in an Ala-rich helix.^{17b} We will leave the incorporation of the experimental observations in this account into detailed theorizing regarding the effects of helix backbone desolvation to others, but do note that the present results are consistent with the view^{3b} that alanine is uniquely helix-favoring since the small methyl side-chain has no entropic penalty for structuring and does not hinder backbone solvation at its site or upstream in the helix.

This work was supported by the Organic and Macromolecular Chemistry Program of the US National Science Foundation through grant # CHE-0650318. National Institutes of Health support (GM-059658) for Trp-cage studies is also acknowledged.

Notes and references

‡ These peptides were synthesized employing standard Fmoc (9-fluorenylmethoxycarbonyl) solid-phase peptide synthesis methods, cleaved with TFA in the presence of protecting group scavengers and purified using RP-HPLC (using C₁₈ and/or C₈ stationary phases). The sequences of all peptides were confirmed by the molecular ions observed using ion-trap mass spectrometry and a complete TOCSY/NOESY NMR assignment.

§ CSDs are defined as ($\delta_{\text{obs}} - \delta_{\text{coil}}$), where δ_{coil} is the statistical coil expectation δ value. In the present study, the statistical coil values were obtained from very short peptides with the same local sequence. All CSDs reported are positive values indicating a downfield structuring shift. The quite small effect of temperature on δ_{coil} (see for example ref. 17b and Fig. 1) was included in the CSD calculations; potential errors in this term do not change the propagation values significantly ($< \pm 0.04$).

- (a) J. C. Lin, B. Barua and N. H. Andersen, *J. Am. Chem. Soc.*, 2004, **126**, 13679–13684; (b) E. J. Spek, C. A. Olson, Z. Shi and N. R. Kallenbach, *J. Am. Chem. Soc.*, 1999, **121**, 5571–5572.
- (a) D. S. Kemp, J. G. Boyd and C. C. Muendel, *Nature*, 1991, **352**, 451–454; (b) J. A. Vila, R. L. Williams, J. A. Grant, J. Wojcik and H. A. Scheraga, *Proc. Natl. Acad. Sci. U. S. A.*, 1992, **89**, 7821–7825.
- (a) C. A. Rohl, W. Fiori and R. L. Baldwin, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 3682–3687; (b) P. Luo and R. L. Baldwin, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 4930–4935; (c) R. L. Baldwin, *J. Biol. Chem.*, 2003, **278**, 17581–17588.
- (a) A. Chakrabartty, T. Kortemme and R. L. Baldwin, *Protein Sci.*, 1994, **3**, 843–852; (b) N. H. Andersen and H. Tong, *Protein Sci.*, 1997, **6**, 1920–1936; (c) J. E. Sun, S. Penel and A. J. Doig, *Protein Sci.*, 2000, **9**, 750–764; (d) D. A. E. Cochran and A. J. Doig, *Protein Sci.*, 2001, **10**, 1305–1311.
- (a) K. Groebke, P. Renold, K. Y. Tsang, T. J. Allen, K. F. McClure and D. S. Kemp, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 4025–4029; (b) J. A. Vila, D. R. Ripoll and H. A. Scheraga, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 13075–13079.
- (a) L. Williams, K. Kather and D. S. Kemp, *J. Am. Chem. Soc.*, 1998, **120**, 11033–11043; (b) J. S. Miller, R. J. Kennedy and D. S. Kemp, *Biochemistry*, 2001, **40**, 305–309; (c) J. S. Miller, R. J. Kennedy and D. S. Kemp, *J. Am. Chem. Soc.*, 2002, **124**, 945–962.
- G. E. Job, R. J. Kennedy, B. Heitmann, J. S. Miller, S. M. Walker and D. S. Kemp, *J. Am. Chem. Soc.*, 2006, **128**, 8227–8233.
- (a) S. Lifson and A. Roig, *J. Chem. Phys.*, 1961, **34**, 1963–1974; (b) A. J. Doig, *Biophys. Chem.*, 2002, **101–102**, 281–293.
- J. Yang, K. Zhao, Y. Gong, A. Vologodskii and N. R. Kallenbach, *J. Am. Chem. Soc.*, 1998, **120**, 10646–10652.
- V. A. Couch, N. Cheng, K. Nambiar and W. Fink, *J. Phys. Chem. B*, 2006, **110**, 3410–3419; and references cited therein.
- (a) J. P. Schneider and W. F. DeGrado, *J. Am. Chem. Soc.*, 1998, **120**, 2764–2767; (b) K. Song, J. M. Stewart, R. M. Fesinmeyer, N. H. Andersen and C. Simmerling, *Biopolymers*, 2008, **89**, 747–760.
- (a) H. A. Scheraga, J. A. Vila and D. R. Ripoll, *Biophys. Chem.*, 2002, **101–102**, 255–265; (b) A. E. Garcia and K. Y. Sanbonmatsu, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 2782–2787.
- S. Chowdhury, W. Zhang, C. Wu, G. Xiong and Y. Duan, *Biopolymers*, 2003, **68**, 63–75.
- (a) H. Nymeyer and A. E. Garcia, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 13934–13939; (b) W. Zhang, H. X. Lei, S. Chowdhury and Y. Duan, *J. Phys. Chem. B*, 2004, **108**, 7479–7489.
- The classical values for $[\theta]_{222}$ for an infinite 100% helical segment were in the $-38\,000$ to $-40\,000$ deg-cm²/residue-dmol range^{4a,b} with corrections for the non-H-bonded end residues bringing this value down to the $-27\,000$ to $-34\,000^\circ$ range observed for 15-residue helices (A. Chakrabartty, T. Kortemme, S. Padmanabhan and R. L. Baldwin, *Biochemistry*, 1993, **32**, 5560–5565; N. H. Andersen, J. R. Cort, Z. Liu, S. J. Sjöberg and H. Tong, *J. Am. Chem. Soc.*, 1996, **118**, 10309–10310). These values have been used for the majority of the CD studies of residue helix propensities (relative w values). CD studies of fluoroalcohol titrations (P. Luo and R. L. Baldwin, *Biochemistry*, 1997, **36**, 8413–8421) and of short helices of fixed geometry (D.-H. Chin, R. W. Woody, C. A. Rohl and R. L. Baldwin, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 15416–15421) have increased the 100% helical value to $-44\,000^\circ$ and decreased the fraying correction for shorter sequences. Studies of block copolymers have yielded $[\theta]_{222}(100\%)$ values as large as $-61\,000^\circ$.
- B. Barua, J. C. Lin, D. V. Williams, J. W. Neidigh, P. Kummeler and N. H. Andersen, *Protein Eng., Design Selection*, 2008, **21**, 171–185.
- (a) S. H. Park, W. Shalongo and E. Stellwagen, *Proteins: Struct., Funct., Genet.*, 1998, **33**, 167–176; (b) R. M. Fesinmeyer, E. S. Peterson, R. B. Dyer and N. H. Andersen, *Protein Sci.*, 2005, **14**, 2324–2332.
- G. S. Jas and K. Kuczera, *Biophys. J.*, 2004, **87**, 3786–3798.
- G. Goch, M. Maciejczyk, M. Oleszczuk, D. Stachowiak, J. Malicka and A. Bierzyński, *Biochemistry*, 2003, **42**, 6840–6847.
- (a) N. H. Andersen, R. B. Dyer, R. M. Fesinmeyer, F. Gai, Z. Liu, J. W. Neidigh and H. Tong, *J. Am. Chem. Soc.*, 1999, **121**, 9879–9880; (b) J. H. Werner, R. B. Dyer, R. M. Fesinmeyer and N. H. Andersen, *J. Phys. Chem. B*, 2002, **106**, 487–494; (c) C.-Y. Huang, J. W. Klemke, Z. Getahun, W. F. DeGrado and F. Gai, *J. Am. Chem. Soc.*, 2001, **123**, 9235–9238; (d) A. Starzyk, W. Barber-Armstrong, M. Sridharan and S. M. Decatur, *Biochemistry*, 2005, **44**, 369–376.
- (a) E. S. Manas, Z. Getahun, W. W. Wright and W. F. DeGrado, *J. Am. Chem. Soc.*, 2000, **122**, 9883–9890; (b) S. T. R. Walsh, R. P. Cheng, W. W. Wright, D. O. V. Alonso, V. Daggett, J. M. Vanderkooi and W. F. DeGrado, *Protein Sci.*, 2003, **12**, 520–531.
- Calculated IR spectra based on MD ensembles for an (AAAKA)₃ peptide suggest that a shift to higher frequency is expected at the alanine C=O four residues upstream from the lysine (S. Gnanakaran, R. M. Hochstrasser and A. E. Garcia, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 9229–9234). Some experimental evidence for this has been reported by Decatur,^{20d} but that study also indicated that these positions are those that are most notably desolvated by the addition of trifluoroethanol. We have not observed^{17b,20a,b} a resolved ¹³C=O peak for desolvated sites in (AAAKA)-repeat ¹³C'-isotopomers.