

Published in final edited form as:

Chem Commun (Camb). 2008 October 21; (39): 4765–4767. doi:10.1039/b807101b.

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

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

Department of Chemistry, University of Washington, Seattle, WA 98195, USA

Abstract

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 in-different² 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 coworkers⁶ 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-YGAACA- (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

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

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 miniproteins have provided two measures of the relative helix propensities of Lys and Ala.

GAAXAAYAQLKDGPPSSGRPPPS (X = A vs. K),

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

DAYAQLXDGGPSSGRPPPS (X = A vs. K),

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

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 $^{13}\text{C}=\text{O}$ ($^{13}\text{C}'$) 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 \rightarrow Ala and Arg \rightarrow Lys \rightarrow Ala mutations in two designed helices: Ac-WAAHA-(AARAA)₃-NH₂ and KKGG-K(AAAAK)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. ($^{13}\text{C}'$)-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 ^{13}C 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 described.^{11b,17b} The results are shown below; the $^{13}\text{C}'$ -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.

[‡]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 ($< \pm 0.04$).

-AARAAAAXAAAARAA-		
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 KAAAAAXAAAAX'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.

-GKAAAAAXAAAAX'AAAAKG-			
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 -AKAAAAA- suggests, as expected for a sidechain-charge-helix-macrodipole 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 $^{13}\text{C}'$ 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 $\delta^{13}\text{C}'$ 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 $\Delta\delta(^{13}\text{C}')$ 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 $^{13}\text{C}=\text{O}$ units at specific alanine sites. The largest $^{13}\text{C}'$ 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 $\Delta\delta(^{13}\text{C}') \geq 3.35$ 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 $^{13}\text{C}'$ 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 Helix 1.5^{4b} (the same value has been reported recently for helices templated by a La^{3+} 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 $(\text{Ala})_n$ 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 $^{13}\text{C}=\text{O}$ -labeled (AAAKA)-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} ($^{12}\text{C}=\text{O}$ and $^{13}\text{C}=\text{O}$, 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} , $^{12}\text{C}=\text{O}$ and $^{13}\text{C}=\text{O}$, 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 $^{13}\text{C}=\text{O}$ 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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

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.

References

- (a) Lin JC, Barua B, Andersen NH. *J Am Chem Soc.* 2004; 126:13679–13684. [PubMed: 15493925] (b) Spek EJ, Olson CA, Shi Z, Kallenbach NR. *J Am Chem Soc.* 1999; 121:5571–5572.
- (a) Kemp DS, Boyd JG, Muendel CC. *Nature.* 1991; 352:451–454. [PubMed: 1861726] (b) Vila JA, Williams RL, Grant JA, Wojcik J, Scheraga HA. *Proc Natl Acad Sci U S A.* 1992; 89:7821–7825. [PubMed: 1502201]
- (a) Rohl CA, Fiori W, Baldwin RL. *Proc Natl Acad Sci U S A.* 1999; 96:3682–3687. [PubMed: 10097097] (b) Luo P, Baldwin RL. *Proc Natl Acad Sci U S A.* 1999; 96:4930–4935. [PubMed: 10220396] (c) Baldwin RL. *J Biol Chem.* 2003; 278:17581–17588. [PubMed: 12582164]
- (a) Chakrabartty A, Kortemme T, Baldwin RL. *Protein Sci.* 1994; 3:843–852. [PubMed: 8061613] (b) Andersen NH, Tong H. *Protein Sci.* 1997; 6:1920–1936. [PubMed: 9300492] (c) Sun JE, Penel S, Doig AJ. *Protein Sci.* 2000; 9:750–764. [PubMed: 10794417] (d) Cochran DAE, Doig AJ. *Protein Sci.* 2001; 10:1305–1311. [PubMed: 11420432]
- (a) Groebke K, Renold P, Tsang KY, Allen TJ, McClure KF, Kemp DS. *Proc Natl Acad Sci U S A.* 1996; 93:4025–4029. [PubMed: 8633010] (b) Vila JA, Ripoll DR, Scheraga HA. *Proc Natl Acad Sci U S A.* 2000; 97:13075–13079. [PubMed: 11078529]

6. (a) Williams L, Kather K, Kemp DS. *J Am Chem Soc.* 1998; 120:11033–11043. (b) Miller JS, Kennedy RJ, Kemp DS. *Biochemistry.* 2001; 40:305–309. [PubMed: 11148022] (c) Miller JS, Kennedy RJ, Kemp DS. *J Am Chem Soc.* 2002; 124:945–962. [PubMed: 11829602]
7. Job GE, Kennedy RJ, Heitmann B, Miller JS, Walker SM, Kemp DS. *J Am Chem Soc.* 2006; 128:8227–8233. [PubMed: 16787087]
8. (a) Lifson S, Roig A. *J Chem Phys.* 1961; 34:1963–1974. (b) Doig AJ. *Biophys Chem.* 2002; 101–102:281–293.
9. Yang J, Zhao K, Gong Y, Vologodskii A, Kallenbach NR. *J Am Chem Soc.* 1998; 120:10646–10652.
10. Couch VA, Cheng N, Nambiar K, Fink W. *J Phys Chem B.* 2006; 110:3410–3419. and references cited therein. [PubMed: 16494355]
11. (a) Schneider JP, DeGrado WF. *J Am Chem Soc.* 1998; 120:2764–2767. (b) Song K, Stewart JM, Fesinmeyer RM, Andersen NH, Simmerling C. *Biopolymers.* 2008; 89:747–760. [PubMed: 18428207]
12. (a) Scheraga HA, Vila JA, Ripoll DR. *Biophys Chem.* 2002; 101–102:255–265. (b) Garcia AE, Sanbonmatsu KY. *Proc Natl Acad Sci U S A.* 2002; 99:2782–2787. [PubMed: 11867710]
13. Chowdhury S, Zhang W, Wu C, Xiong G, Duan Y. *Biopolymers.* 2003; 68:63–75. [PubMed: 12579580]
14. (a) Nymeyer H, Garcia AE. *Proc Natl Acad Sci U S A.* 100:13934–13939. [PubMed: 14617775] (b) Zhang W, Lei HX, Chowdhury S, Duan Y. *J Phys Chem B.* 2004; 108:7479–7489.
15. The classical values for $[\theta]_{222}$ for an infinite 100% helical segment were in the -38000 to -40000 $\text{deg}\cdot\text{cm}^2/\text{residue}\cdot\text{dmol}$ range^{4a,b} with corrections for the non-H-bonded end residues bringing this value down to the -27000 to -34000° range observed for 15-residue helices (Chakrabarty A, Kortemme T, Padmanabhan S, Baldwin RL. *Biochemistry.* 1993; 32:5560–5565. [PubMed: 8504077] Andersen NH, Cort JR, Liu Z, Sjoberg SJ, Tong H. *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 (Luo P, Baldwin RL. *Biochemistry.* 1997; 36:8413–8421. [PubMed: 9204889]) and of short helices of fixed geometry (Chin DH, Woody RW, Rohl CA, Baldwin RL. *Proc Natl Acad Sci U S A.* 2002; 99:15416–15421. [PubMed: 12427967]) have increased the 100% helical value to -44000° and decreased the fraying correction for shorter sequences. Studies of block co-polymers have yielded $[\theta]_{222}(100\%)$ values as large as -61000° ⁷.
16. Barua B, Lin JC, Williams DV, Neidigh JW, Kummeler P, Andersen NH. *Protein Eng, Design Selection.* 2008; 21:171–185.
17. (a) Park SH, Shalongo W, Stellwagen E. *Proteins: Struct, Funct, Genet.* 1998; 33:167–176. [PubMed: 9779786] (b) Fesinmeyer RM, Peterson ES, Dyer RB, Andersen NH. *Protein Sci.* 2005; 14:2324–2332. [PubMed: 16131660]
18. Jas GS, Kuczera K. *Biophys J.* 2004; 87:3786–3798. [PubMed: 15339816]
19. Goch G, Maciejczyk M, Oleszczuk M, Stachowiak D, Malicka J, Bierzyński A. *Biochemistry.* 2003; 42:6840–6847. [PubMed: 12779338]
20. (a) Andersen NH, Dyer RB, Fesinmeyer RM, Gai F, Liu Z, Neidigh JW, Tong H. *J Am Chem Soc.* 1999; 121:9879–9880. (b) Werner JH, Dyer RB, Fesinmeyer RM, Andersen NH. *J Phys Chem B.* 2002; 106:487–494. (c) Huang CY, Klemke JW, Getahun Z, DeGrado WF, Gai F. *J Am Chem Soc.* 2001; 123:9235–9238. [PubMed: 11562202] (d) Starzyk A, Barber-Armstrong W, Sridharan M, Decatur SM. *Biochemistry.* 2005; 44:369–376. [PubMed: 15628879]
21. (a) Manas ES, Getahun Z, Wright WW, DeGrado WF. *J Am Chem Soc.* 2000; 122:9883–9890. (b) Walsh STR, Cheng RP, Wright WW, Alonso DOV, Daggett V, Vanderkooi JM, DeGrado WF. *Protein Sci.* 2003; 12:520–531. [PubMed: 12592022]
22. 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 (Gnanakaran S, Hochstrasser RM, Garcia AE. *Proc Natl Acad Sci U S A.* 2004; 101:9229–9234. [PubMed: 15197256]). 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 $^{13}\text{C}=\text{O}$ peak for desolvated sites in (AAAKA)-repeat $^{13}\text{C}'$ -isotopomers.

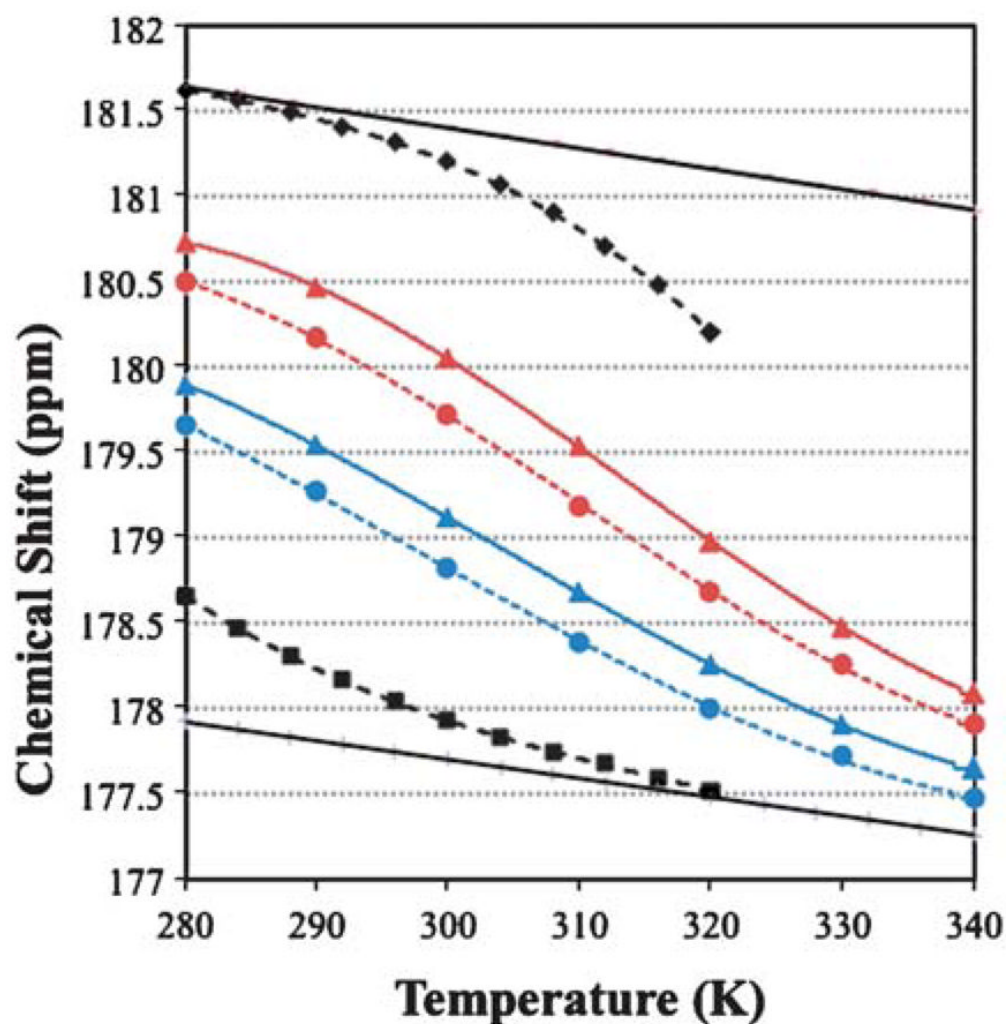


Fig. 1. Melting curves for $^{13}\text{C}=\text{O}$ chemical shifts. The data for DAYAQLKDGK and NAYAQLKDGK appear in black with the derived 100% and 0% helix lines. The data for Ac-WAAAHAAARAA \blacktriangle AXAA \blacktriangle ARAA-NH₂ is color coded by alanine position, with the X = A data shown as \blacktriangle connected by a solid line and X = K as \bullet connected by a dashed line. The 100% and 0% helix lines for -RAA \blacktriangle AXAA \blacktriangle AR- appear in the Electronic Supplementary Information together with the method employed to convert CSDs to helix propagation values.[†]

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