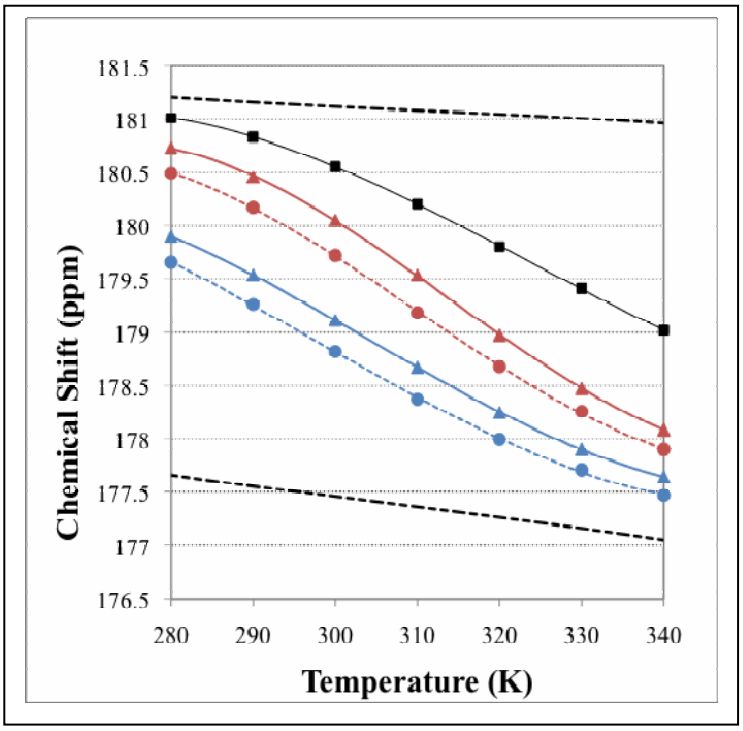
Lysine and arginine residues do not increase the helicity of alanine-rich helices

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

For the Electronic Supplementary Information --



Supplemental Figure 1. The data, in red and black in Figure 1, is repeated here with the 0% and 100% folded lines that were used to obtain the chemical shift deviations reported in the Communication text. The data point in black are for the underlined alanine site in GASEDE(A<u>A</u>AAK)$_3$GY-NH_2, the isolated hairpin system that comes closest to the 100% folded line: CSD(100% helicity) = 3.50 + 0.006·(T – 273).

The chemical shift deviation (CSD) for each site and temperature in $\delta_{obs} - \delta_{(0\%\ line)}$. The fractional helicity is calculated as CSD/(3.50 + 0.006·(T – 273)). These 0 and 100% lines apply only to the central site (shown in red). The C-terminal site (blue) shows smaller downfield shifts for several reasons: 1) C-terminal fraying, 2) the CSD for full local helicity is likely smaller due to a Coulombic effect, the negative end of the helix macrodipole will have a shielding effect, and 3) even if **A** in the -RAA**A**AXAA**A**ARAA-NH_2 has 100% local helicity, C-terminal fraying implies that the ^{13}C=O will not always be H-bonded. H-bonding is responsible for at least a third of the downfield ^{13}C' shift associated central helicity.

The -RAA**A**AXAA**A**ARAA-NH_2 melt also confirms C-terminal fraying: note that the data for the **A** site is leveling out at the low temperature but **A** is not.

The procedure for deriving propagation values is given here.

In the extended Lifson-Roig formalism of Helix1.5 [1] of helix/coil equilibria, the probability (W_{ij}) of a helix spanning form residue i to j is given by eqn 1,

eqn 1 $$W_{ij} = N_{i-1}\, C_{j+1}\, v_N(i)\, v_C(j) \prod_{i+1}^{j-1} w_k$$

where N_{i-1} and C_{j+1} are capping constants, v_N and v_C are nucleation constants, and the w_k are the propagation values of the intervening residues. For very stable helices with clearly defined termini, an extended single sequence assumption leads to the conclusion that $\Delta\Delta G_U$ for residue additions and mutations within the central helical segment should be proportional to ln (Π *w*) where Π *w* is the product of the propagation values for the added residues[2] This is the case for the N-terminal helix of the Trp-cage. Thus, for single site mutations (X→Z) within that helical span, eqn 2 would apply –

eqn 2 $$\Delta\Delta G_U = RT \ln (w_X/w_Z).$$

For the Trp-cage systems,

	w(Lys)
GAAXAAYAQWLKDGGPSSGRPPPS (X = A *vs* K), ΔTm = 6°C ⇒ *w*(Ala) / *w*(Lys) = 1.95	0.79
DAYAQWLXDGGPSSGRPPPS (X = A *vs* K), ΔTm = 5°C ⇒ *w*(Ala) / *w*(Lys) = 1.75	0.88

we can convert ΔTm values to $\Delta\Delta G_U$ values, we have a direct measure of the propagation constant ratio. With an assumed *w*(Ala) w_{Ala} value (1.54), we obtain values of 0.79 for a position slightly toward the N-terminus of the helix and a larger value, 0.88 for the site nearly at the C-terminus.

In the case of less stable helices, the matrix multiplication (summation over all possible helices of varying length) inherent in a Lifson-Roig formalism may not reduce to a single sequence approximation (such as eqns 1 and 2). We used Helix1.5 determine the sensitivity of central helicity measures (for example the simulated local K_H defined as $f_H/(1 - f_H)$, at the labeled central alanine of $A_mAAAKA\underline{A}AA(XAAAA)_nKGNH_2$ or N_{cap}-$A_mAAAKA\underline{A}AA(XAAAA)_nKGNH_2$ sequences to changes in n, m, and w_X. For short sequences (n = 1, N_{i-1} = 0.3 – 6) that display central f_H values between 0.2 and 0.5 for m = 0, a plot of ln(K_H) versus m provides the propagation value: ln(K_H) = mx + k: with $e^x = (1.02 \pm <0.04) \cdot w_{Ala}$. Concordant experimental determinations for such series (data not shown) served as an additional confirmation of the w_{Ala} value we use.

Of more specific pertinence to the present study, we also determined the sensitivity of K_H at the sites shown below to the w_X/w_Z ratio for X→Z mutations.

Ac-WAAAHAAARAA<u>A</u>AXAA<u>A</u>ARAA-NH_2

GG-KAA<u>A</u>AXA<u>A</u>AAKA<u>A</u>AAK-GG

GG-KAA<u>A</u>AKA<u>A</u>AAXA<u>A</u>AAK-GG

In all cases, $K_H(X)/K_H(Z)$ was proportional to w_X/w_Z , but the proportionality constant was not unity. The proportionality constants that applied in the Helix1.5 simulations were used to derive the Lys and Arg propagation values. As an example, for mutations at X in Ac-WAAAHAAARAA<u>A</u>AXAA<u>A</u>ARAA-NH_2 the equations for the Helix1.5 simulation results were $(K_H(X)/K_H(Z)) = 0.835(w_X/w_Z) + 0.166$ and $(K_H(X)/K_H(Z)) = 0.584(w_X/w_Z) + 0.318$, respectively at A12 and A17.

The progression from CSDs, to f_H values, to propagation values for the systems examined is shown below. For the first system, CSD averages were used within the first two repeats.

From the CSDs

	-GKAAAAXAAAAX'AAAAKG-				-AARAAAAXAAAARAA-	
X, X' = K, K	1.92	2.02	1.28	X = A	3.04	2.21
X, X' = A, K	2.39	2.48	1.62	X = R	2.90	2.05
X, X' = K, A	2.11	2.35	1.66	X = K	2.826	1.96

to fractional helicity

	-GKAAAAXAAAAX'AAAAKG-				-AARAAAAXAAAARAA-	
X, X' = K, K	0.542	0.571	0.362	X = A	0.859	0.624
X, X' = A, K	0.675	0.701	0.458	X = R	0.819	0.579
X, X' = K, A	0.596	0.664	0.469	X = K	0.798	0.554

converted to propagation values

	-GKAAAAXAAAAX'AAAAKG-				-AARAAAAXAAAARAA-	
X = A → K	0.781	0.721	0.778	X = A	<1.54>	<1.54>
X' = A → K	0.782	0.846	0.890	X = R	0.947	1.016
				X = K	0.815	0.888

Over a wider range of peptide helix simulations using Helix1.5, the best central site propagation values for Arg and Lys are 1.02 and 0.84, respectively. These are reduced to 0.94 and 0.68, respectively, when the residue is four or less positions from the N-cap, and increased to 1.22 and 1.00 when these positively charged residues occur within 3 residues of the helix C-cap. A more complete re-parameterization of Helix1.5 is in progress.

---

(1) Andersen, N. H.; Tong, H. *Protein Sci* **1997**, *6*, 1920-1936.
(2) Lin, J. C.; Barua, B.; Andersen, N. H. *J Am Chem Soc* **2004**, *126*, 13679-13684.