Supporting Information

Table S1. Mass Spectroscopy Data.

Peptide	*m/z* Found (Expected)
$WK_4Inp_2{}^tLA_4{}^tLInp_2K_4$	728.64 (728.49), 546.78 (546.62), 437.64 (437.50), 364.98 (364.75), 313.04 (312.79)
$WK_4Inp_2{}^tLA_{12}{}^tLInp_2K_4$	1375.79 (1376.38), 917.54 (917.92), 688.40 (688.69)
$WK_4Inp_2{}^tLA_{13}{}^tLInp_2K_4$	941.79 (941.60), 706.74 (706.45), 565.53 (565.36), 471.49 (471.31), 404.28 (404.12), 353.92 (353.73)
$WK_4Inp_2{}^tLA_{14}{}^tLInp_2K_4$	965.63 (965.28), 724.53 (724.21), 579.70 (579.57), 483.28 (483.14), 414.39 (414.27), 362.78 (362.61)
$WK_4Inp_2{}^tLA_{15}{}^tLInp_2K_4$	989.15 (988.96), 742.25 (741.97), 593.92 (593.78), 495.14 (494.98), 424.56 (424.42), 371.64 (371.49)
$WK_4Inp_2{}^tLA_{16}{}^tLInp_2K_4$	1012.92 (1012.64), 760.03 (759.73), 608.15 (607.99), 507.00 (506.82), 434.73 (434.56), 380.57 (380.37), 338.44 (338.22)
$WK_4Inp_2{}^tLA_{17}{}^tLInp_2K_4$	1035.83 (1036.32), 777.03 (777.49), 662.04 (662.19), 518.45 (518.66)
$WK_4Inp_2{}^tLA_{18}{}^tLInp_2K_4$	1060.15 (1060.00), 795.47 (795.25), 636.61 (636.40), 530.65 (530.50), 455.02 (454.86), 398.29 (398.13), 354.17 (354.00)
$WK_4Inp_2{}^tLA_{19}{}^tLInp_2K_4$	1624.98 (1625.01), 1083.70 (1083.68), 813.11 (813.01)
$WK_4Inp_2{}^tLA_{20}{}^tLInp_2K_4$	1107.26 (1107.36), 830.75 (830.77), 664.79 (664.82), 554.27 (554.18)
$WK_4Inp_2{}^tLA_{21}{}^tLInp_2K_4$	1130.94 (1131.03), 848.25 (848.53), 678.71 (679.02), 566.06 (566.02)
$WK_4Inp_2{}^tLA_{22}{}^tLInp_2K_4$	1154.82 (1154.71), 866.52 (866.29), 693.45 (693.23), 578.00 (577.86), 495.61 (495.45), 433.77 (433.65), 385.77 (385.58)
$WK_4Inp_2{}^tLA_{23}{}^tLInp_2K_4$	1178.49 (1178.39), 884.32 (884.05), 707.87 (707.44), 589.87 (589.70), 505.76 (505.60), 442.70 (442.53), 393.62 (393.47)
$WK_4Inp_2{}^tLA_{24}{}^tLInp_2K_4$	1202.61 (1202.07), 902.00 (901.81), 722.04 (721.65), 601.68 (601.54), 515.91 (515.75), 451.51 (451.41), 401.54 (401.36)
$WK_5Inp_2{}^tLA_{28}{}^tLInp_2K_5$	1382.95 (1382.18), 1037.23 (1036.89), 830.01 (829.71), 692.03 (691.60), 593.19 (592.94), 519.17 (518.95), 461.59 (461.40)
$WK_5Inp_2{}^tLA_{29}{}^tLInp_2K_5$	1406.44 (1405.86), 1054.64 (1054.65), 844.22 (843.92), 703.80 (703.44), 603.28 (603.09), 528.07 (527.83), 469.46 (469.29)
$WK_5Inp_2{}^tLA_{35}{}^tLInp_2K_5$	1547.50 (1547.94), 1160.95 (1161.20), 928.99 (929.17), 774.30 (774.47), 663.92 (663.98), 581.01 (581.11), 516.56 (516.65)
$WK_5Inp_2{}^tLA_{36}{}^tLInp_2K_5$	1571.19 (1571.62), 1178.63 (1178.96), 943.08 (943.37), 786.09 (786.31), 674.00 (674.13), 589.86 (589.99), 524.43 (524.54)
$WK_6Inp_2{}^tLA_{44}{}^tLInp_2K_6$	1108.74 (1108.27), 924.18 (923.73), 792.33 (791.91), 693.50 (693.05), 616.44 (616.15), 554.91 (554.64), 504.58 (504.31)
$WK_6Inp_2{}^tLA_{45}{}^tLInp_2K_6$	936.18 (935.57), 702.69 (701.93), 624.54 (624.05), 562.25 (561.74), 511.28 (510.77)
$WK_2Inp_2{}^tLA_4{}^tLInp_2K_2$	836.35 (836.04), 557.77 (557.70), 418.59 (418.53), 335.11 (335.02)
$WK_3Inp_2{}^tLA_4{}^tLInp_2K_3$	964.58 (964.14), 643.26 (643.10), 482.68 (482.57), 386.45 (386.26), 322.28 (322.05)
$WK_5Inp_2{}^tLA_4{}^tLInp_2K_5$	814.26 (813.89), 611.00 (610.67), 489.08 (488.74), 407.60 (407.45), 349.71 (349.39), 306.09 (305.84)
$WK_6Inp_2{}^tLA_4{}^tLInp_2K_6$	899.88 (899.28), 675.03 (674.72), 540.14 (539.97), 450.42 (450.15), 386.20 (385.98), 338.17 (337.86)
$WK_7Inp_2{}^tLA_4{}^tLInp_2K_7$	739.25 (738.76), 591.58 (591.21), 493.17 (492.84), 422.81 (422.58), 370.19 (369.89), 329.17 (328.90)
$WK_2Inp_2{}^tLA_{12}{}^tLInp_2K_2$	1120.43 (1120.19), 747.43 (747.13), 560.80 (560.60), 448.91 (448.68)

Peptide	*m/z* Found (Expected)
$WK_3Inp_2{}^tLA_{12}{}^tLInp_2K_3$	832.31 (832.53), 624.27 (624.65), 499.76 (499.92), 416.60 (416.77), 357.31 (357.37)
$WK_7Inp_2{}^tLA_{12}{}^tLInp_2K_7$	1173.84 (1174.11), 880.57 (880.84), 704.75 (704.87), 587.27 (587.56), 503.66 (503.77), 440.80 (440.92), 391.97 (392.04)
$WK_3Inp_2{}^tLA_{19}{}^tLInp_2K_3$	998.24 (998.28), 748.98 (748.96), 599.20 (599.37), 499.64 (499.64), 428.35 (428.41)
$WK_7Inp_2{}^tLA_{19}{}^tLInp_2K_7$	804.31 (804.32), 670.48 (670.44), 574.88 (574.80), 503.15 (503.08), 477.31 (447.29), 402.76 (402.67)
$K_4Inp_2{}^tLA_{19}{}^tLInp_2K_4W$	1083.62 (1083.68), 812.99 (813.01), 650.57 (650.61), 542.27 (542.34), 464.92 (465.01), 406.97 (407.01)
$WK_4Inp_2{}^tLA_{19}{}^tLInp_2K_4W$	859.32 (859.53), 687.78 (687.82), 573.23 (573.36), 491.60 (491.59), 430.14 (430.27), 382.53 (382.57)
$WK_4Inp_3A_{19}Inp_3K_4$	1082.40 (1082.33), 812.09 (812.00), 649.87 (649.80), 541.76 (541.67), 464.48 (464.43), 406.59 (406.50)
$WK_4P_3A_{19}P_3K_4$	1054.77 (1054.30), 791.16 (790.98), 633.15 (632.98), 527.79 (527.65), 452.54 (452.42), 396.12 (395.99)
$WK_4Inp_2{}^tLA_{28}{}^tLInp_2K_4$	973.23 (972.84), 778.85 (778.48), 649.18 (648.90), 556.45 (556.34), 487.08 (486.93), 433.08 (432.93)
$WK_7Inp_2{}^tLA_{28}{}^tLInp_2K_7$	1165.24 (1164.98), 932.52 (932.19), 777.26 (776.99), 666.41 (666.14), 583.19 (583.00), 518.54 (518.33), 466.78 (466.60)
$WK_5P_3A_{28}P_3K_5$	1014.96 (1014.86), 812.49 (812.09), 677.26 (676.91), 580.51 (580.35), 508.07 (507.93), 451.78 (451.61), 406.79 (406.55)
$WK_4Inp_2{}^tL$	526.43 (526.36), 351.26 (351.24), 263.64 (263.68)
$W\beta_3G_3A_{12}G_3\beta_3$	1915.40 (1914.97), 957.93 (957.99), 639.02 (639.00), 479.62 (479.50), 383.61 (383.80)
$W^hK_3G_3A_{12}G_3{}^hK_3$	1126.32 (1126.18), 751.25 (751.12), 563.73 (563.59), 451.21 (451.08), 376.14 (376.06), 322.54 (322.48)
$WK_3G_2A_{12}G_2K_3$	1027.32 (1027.11), 685.14 (685.08), 514.17 (514.06), 411.36 (411.45), 343.13 (343.04), 294.24 (294.18)
$WK_3G_2A_{12}G_4K_3$	1084.04 (1084.13), 723.18 (723.09), 542.64 (542.57), 434.21 (434.26), 362.15 (362.05), 310.52 (310.47)
$WK_3G_3A_{12}G_3K_3$	1084.23 (1084.13), 723.24 (723.09), 542.64 (542.57), 434.33 (434.26), 362.09 (362.05), 310.52 (310.47)
$WK_3G_4A_{12}G_2K_3$	1084.04 (1084.13), 723.12 (723.09), 542.64 (542.57), 434.27 (434.26), 362.15 (362.05), 310.52 (310.47)
$WK_3G_4A_{12}G_4K_3$	1141.36 (1141.15), 761.23 (761.10), 571.18 (571.08), 457.20 (457.07), 381.19 (381.06), 326.87 (326.76)
$WK_3G_5A_{12}G_5K_3$	1198.21 (1198.17), 799.31 (799.12), 599.71 (599.59), 479.97 (479.87), 400.10 (400.06), 343.14 (343.06)
$WK_3G_7A_{12}G_7K_3$	1312.62 (1312.22), 875.33 (875.15), 656.70 (656.61), 525.62 (525.49), 438.05 (438.08), 375.76 (375.64)
$WK_3InpA_{12}InpK_3$	1024.38 (1024.13), 683.17 (683.09), 512.64 (512.57), 410.34 (410.26), 342.11 (342.05), 293.35 (293.33)
$WK_3Inp_2A_{12}Inp_2K_3$	1135.42 (1135.20), 757.46 (757.14), 568.17 (568.11), 454.77 (454.69), 379.20 (379.07), 325.21 (325.06)
$WK_3Inp_3A_{12}Inp_3K_3$	831.35 (831.18), 623.75 (623.64), 499.28 (499.11), 416.14 (416.10), 356.94 (356.80)

Table S2. Ellipticities of peptides with Gly IS.

A	$WK_3G_jA_{12}G_kK_3$		
	$-[\theta]_{222,molar} \times 10^{-5}$ deg cm^2 dmol^{-1}		
j,k	2 °C	25 °C	60 °C
2,2	1.41	0.81	0.49
2,4	1.22	0.66	0.34
3,3	1.66	0.91	0.37
4,2	1.95	1.15	0.58
4,4	1.88	0.99	0.44
5,5	1.98	1.11	0.47
7,7[a]	2.07	1.30	0.64
B	$WX_3G_3A_{12}G_3X_3$		
	$-[\theta]_{222,molar} \times 10^{-5}$ deg cm^2 dmol^{-1}		
X	2 °C	25 °C	60 °C
K	1.66	0.91	0.37
β[b]	2.07	1.24	0.77
^{h}K[c]	1.76	1.01	0.49

[a] Precipitated at 60 °C in 100 mM NaCl solution. [b] β = β-aminoalanine. [c] ^{h}K = *S*-$NH_2(CH_2)_5CH(NH_2)CO_2H$ (homolysine).

Table S3. Experimental AUC conditions for peptides of the form $WK_mInp_2{}^tLA_n{}^tLInp_2K_m$.

Plot	n^a	Solvent	Concentration Peptide (μM)	Concentration NaCl (M)	λ (nm)	krpm	T °C	Slope	MW (D)	MW_{exp} (D)
A	12	H_2O	50	0.01	280	40	4.0	0.482	2752	2911
B[b]	19	H_2O	71	0.01	280	40	4.0	0.523	3250	3050
C[c]	28	H_2O	72	0.10	280	47	2.0	0.895	3890	3671
D	28	H_2O	14	—	225	45	5.0	0.235	3890	1054
E	28	D_2O	14	—	223	45	5.0	0.138	3890	968
F	28	D_2O	7.1	0.10	225	45	5.0	0.639	3890	4480
G[d]	45[a]	H_2O	10	—	225	45	2.0	0.518	5098	2204
H	45[a]	D_2O	8.0	—	223	45	5.0	0.272	5098	1790

[a] All of these peptides have solubilizers K_m, $m = 4$ except for those with $n = 45$, in which case $m = 6$. [b] See plot marked "A_{19}" in Fig. 8. [c] See plot marked "A_{28}" in Fig. 8. [d] See plot marked "A_{45}" in Fig. 8.

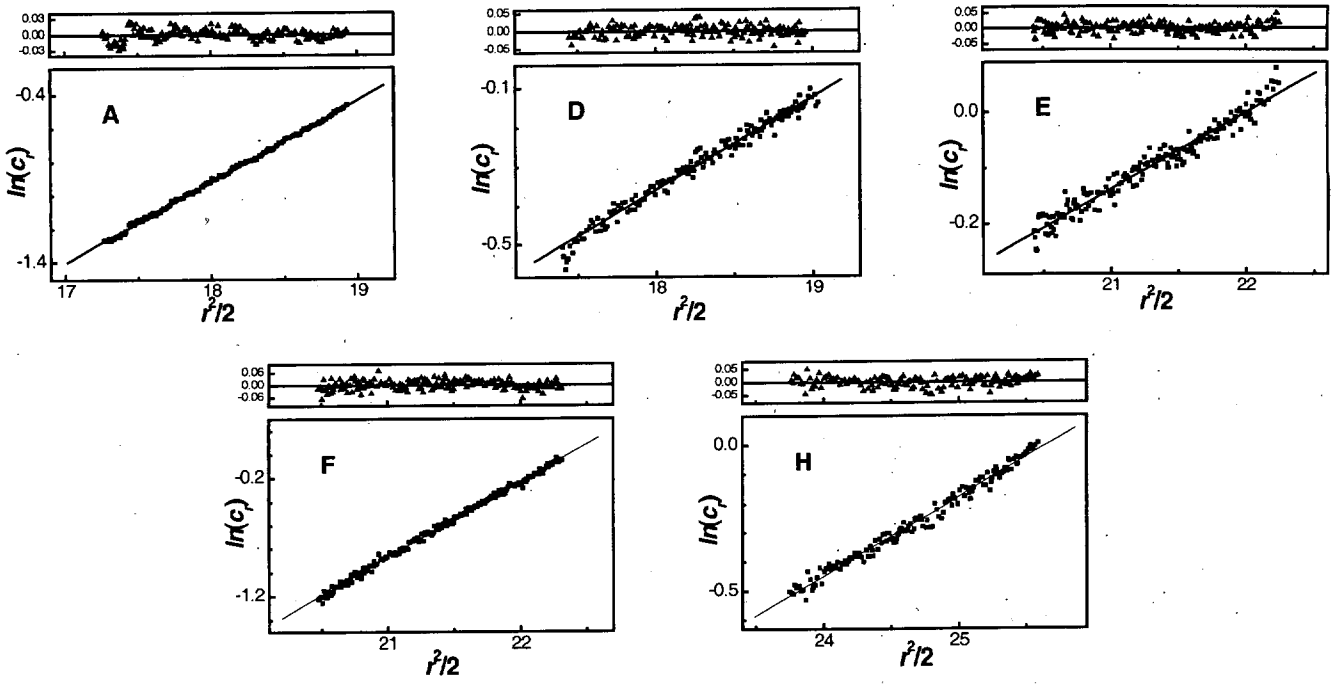

Figure S1. Analytical ultracentrifugation results (squares) for alanine conjugates with $n = 12$ to 45 in pure or deuterated water with and without added sodium chloride. The experimental details are found in Table S3. Each of these plots is fitted to a straight line with negligible residuals (triangles), as expected for unaggregated species.

Table S4. Tryptophan Extinction Coefficient Data.

	$\mathfrak{M}_{280} \times M^{-1}\,cm^{-1}$										
	Number of Solubilizing Lysines									Average	
Solvent	0[a]	0[b]	2	3	4	4[c]	5	6	7	Std. Dev.	
Pure water	5340	5526	5494	5636	5551	5459	5405	5541	5566		
	5533	5535	5549	5629	5745	5495	5673	5691	5476		
	5397	5623	5433	5545	5835	5501	5497	5556	5539	5558	
	5615	5605	5460	5375	5823	5465	5528	5676	5441	118	
	5730	5720									5553
0.1 M NaCl	5610	5542			5583						112
	5353	5570			5661						
	5487	5503			5641					5539	
	5415	5545			5622					94	
	5410	5610									
6.0 M GuHCl	5768	5792			5758						
	5875	5855			6127					5888	
	5880	5827			5816					116	
	6070	5970			5920						

[a] Tryptophan acid. [b] Tryptophan amide hydrochloride. [c] $WK_4Inp_2^tL$.

Mathematical Modeling

All of the mathematical modeling was carried out using Mathematica 3.0 from Wolfram Research. Non-cooperative peptide state sums were generated using Eq S1, setting all $n_i = 1.0$ for $i > 0$, and all $c_i = 1.0$ for $i < n$.

$$SS_n = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & n_0 & 1 \end{pmatrix} \begin{pmatrix} y & tyc_{i+1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & n_i & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & ty & \frac{v}{n_{i-1}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & n_i & 1 \\ y & \frac{v^2}{ty_{i-1}n_{i-2}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & n_i & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & ty & \frac{v}{n_{i-1}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & n_i & 1 \end{pmatrix}^n \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{(S1)}$$

SS_n refers to the state sum for any peptide of length n; ty, the product of a helical initiation parameter t and a helical propensity y, is the weight applied to a helical residue sited at a helix terminus;[1,2] v is the statistical weight applied to a randomly helical residue; y is the helical

[1] The weight normally assigned a helix terminus by L-R modeling is v; the ty weighting allows N- and C-terminal helical residues to be considered as part of the helical length. Although t and v could theoretically be assigned differing values, for the purposes of this study both were assigned values of $t = v = 0.048$.

[2] The variable y is used here for indexing purposes, and is replaced by the conventional helical propensity symbol w in the cooperativity model.

propensity applied to a residue found in the interior of a helix;[2] n_0 is the N-capping parameter for a ^{t}L residue at the N-terminus of a solubilized, spaced polyalanine, and c_{n+1} is similarly the C-capping parameter for a ^{t}L. These capping parameters were initially set equal to unity and therefore did not appear in state sums until capping effects were considered.

Rather than conventional 3×3 or 4×4 matrices, these calculations use the 8×8 matrices that we have introduced previously.[3] All these give equivalent polynomial state sums, with two major differences. First, the state weight for a helical conformation of length k is t^2y^k, and second, in the initial state sum calculation, a conventional v weighting is used for conformations containing isolated helical residues, but a t^2 weight is used for helices. At the evaluation stage of the calculation, we set $t = v = 0.048$, according to common usage.[4] The double t, v weighting system for helix initiation allows easy partitioning of the state sum; for example χ_k, the mole fractions of conformations containing uninterrupted helices of length k, is (1/SS)(sum of all polynomial terms containing t^2y^k as a coefficient) Eq S2.[5] The partitioning of state sums into such terms is rapid and efficient with modern PC-based computational algorithms.

$$\chi_k = \frac{y^k}{\mathrm{SS}_n} \sum_{k=3}^{n} \text{Coefficients in } t^2 \qquad \text{(S2)}$$

As defined by Eq S1, SS_n is a polynomial in v, t, and y (and n_0 and c_{n+1}). We define C_k as the coefficient of the term y^k within this polynomial. Fractional helicity (f_H) is defined as the average number of helical residues divided by the total number of residues in a peptide, and its explicit expression is displayed in Eq S3.

$$f_H = \frac{1}{n \cdot \mathrm{SS}_n} \sum_{k=3}^{n} k\mathrm{C}_k y^k \qquad \text{(S3)}$$

This equation was used to generate plots of f_H *vs.* peptide length.

As noted in the text, the assumption that fractional helicity can be expressed as a simple ratio of $[\theta]_{222}$ and $[\theta_n]_{222}$ is false. The precise expression[5] for modeling $[\theta]_{222}$ is shown in Eq S4

$$[\theta]_{222} = \frac{1}{n \cdot \mathrm{SS}_n} \sum_{k=3}^{n} [\theta_k]_{222} k\mathrm{C}_k y^k \qquad \text{(S4)}$$

in which the $[\theta_n]_{222}$ term has been moved inside the summation and altered to reflect the limiting ellipticity $[\theta_k]_{222}$ for each conformation with a helical length k. The justification for this equation springs from the related Eq 3 of the text. When divided by SS_n each term in $\mathrm{C}_k y^k$ in Eq S4 equals χ_k, the mole fraction of conformations containing a helical region of length k. Multiplication by k/n gives the contribution of each such conformation to the overall fractional helicity as seen in

[3] Wallimann, P.; Kennedy, R. J.; Kemp, D. S. *Angew. Chem., Int. Ed.* **1999**, 38, 1290-1292.
[4] Rohl, C. A.; Scholtz, J. M.; York, E. J.; Stewart, J. M.; Baldwin, R. L. *Biochemistry* **1992**, *31*, 1263-1269.
[5] This expression uses implicitly the single-helix approximation and therefore very slightly overestimates f_H.

Eq 4 of the text. Multiplying in the term $[\theta_k]_{222}$ gives the contribution of the length k helices to the overall $[\theta]_{222}$. Summing over all lengths k then gives $[\theta]_{222}$.

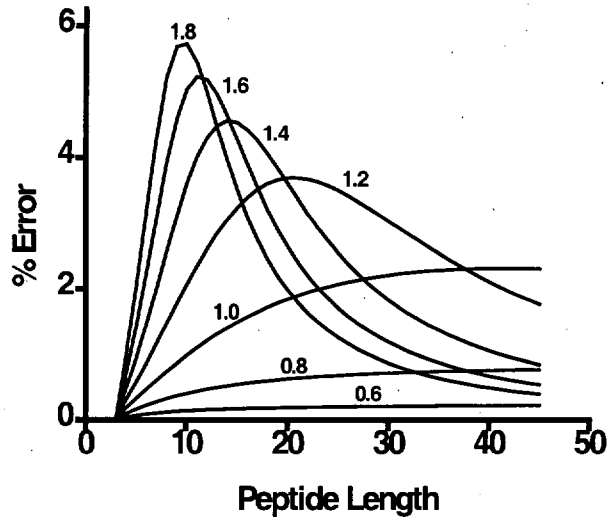

Figure S2. Percent error $(f_H - [\theta]_{222}/[\theta_n]_{222}) \times 100$ associated with standard fractional helicity calculations for homopeptides of varying length and helical propensity. The standard formula for determining f_H takes the ratio of $[\theta]_{222}$ to $[\theta_n]_{222}$; a more precise method is based on Eq S4, and requires mathematical modeling. At helical propensities similar to those reported for alanine near 1.6, f_H for peptides between 10 and 20 residues long is grossly underestimated by the standard calculation.

A plot of the error arising from this modification is shown in Fig. S2 over a length range and for a variety of helical propensities. As noted in the text, a 10 % reduction in least-squares fitting error was obtained when this modification was applied. A further 10 % reduction in error accompanied the inclusion of N- and C-capping values ($n_0 = c_{n+1} = 0.77$ in Eq S1) for the ^{t}L residues in the IS regions.

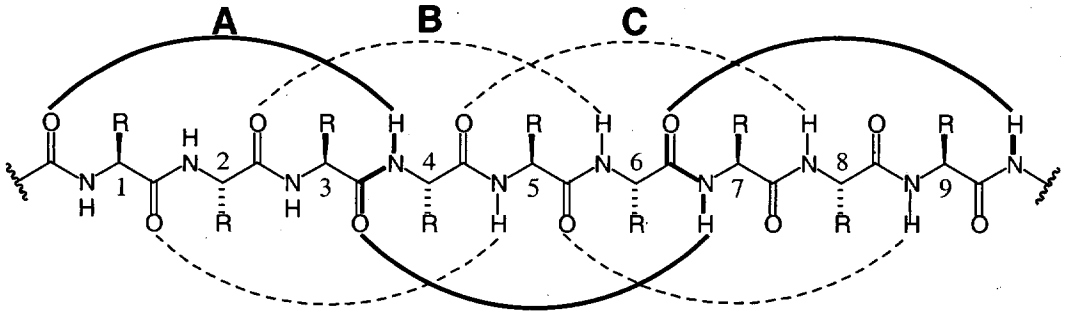

Figure S3. A schematic of a 9-residue α-helix indicating the three independent sets of linked H-bonds. With the ninth residue, chain A grows to three consecutive H-bonds while chains B and C remain at two H-bonds each.

Cooperativity Model. As noted in the text, Guo and Karplus[6] have predicted a cooperative effect resulting from unbroken chains of multiple hydrogen bonds. An α-helix is composed of

[6] Guo, H.; Karplus, M. *J. Phys. Chem.* **1994**, *98*, 7104-7105.

three independent sets of linked hydrogen bonds as illustrated in Fig. S3. Our cooperativity model assigns an increased statistical weight to the terms in L-R state sums that describe helices long enough to exhibit such H-bonding cooperativity. Modified L-R state sums can thus be expressed according to Eq S5

$$SS'_n = \sum_{k=3}^{n} \mathrm{Coop}_k C_k y^k + C_0, \qquad \text{(S5)}$$

with each state containing a helix of length k now assigned an additional length-dependent weight Coop_k; modified state sums are notated with primes as SS'_n.

We have observed and modeled the onset of this cooperativity phenomenon for states containing relatively short helical segments using Ac-Hel nucleated peptides.[7] Helical segments up to four residues long (*i.e.*, $k \leq 4$) were each assigned a weight of $\mathrm{Coop}_k = 1.0$; $\mathrm{Coop}_5 = w^{5(1/3)}$ and $\mathrm{Coop}_6 = w^{6(2/3)}$ to facilitate a smooth mathematical transition; and states with longer helical segments were assigned a weight $\mathrm{Coop}_k = w^k$ as expected in standard L-R modeling.

The weights we assigned previously for short helices reflect cooperativity due to a second contiguous H-bond; modeling in the current study modifies the statistical weights for conformations containing at least four consecutive H-bonds.[8] Thus for $k = 12$ a new cooperativity parameter z is introduced, with $\mathrm{Coop}_{12} = zw^{12}$; as each additional independent H-bonding chain extends to a length of four, the order in z is increased, with $\mathrm{Coop}_{13} = z^2w^{13}$ and $\mathrm{Coop}_{14} = z^3w^{14}$. For every state with a longer helical segment of k residues, $\mathrm{Coop}_k = z^3w^k$. Using these modified state sums we generated Eq S6

$$[\theta]_{222} = \frac{1}{n \cdot SS'_n} \sum_{k=3}^{n} [\theta_k]_{222} k \mathrm{Coop}_k C_k y^k \qquad \text{(S6)}$$

analogous to Eq S4 for modeling $[\theta]_{222}$.[9] It should be noted that the variable y, used here for indexing purposes only, is assigned a value of 1.0 by the cooperativity model. In effect, the cooperativity state sum expressed in Eq S6 is thus defined recursively in terms of the non-cooperative state sum in Eq S4.

---

[7] Kennedy, R. J.; Tsang, K. Y.; Kemp, D. S. *J. Am. Chem. Soc.*, in review

[8] Additional parameters may be added for the third and fifth H-bonds as well, but the data is satisfactorily modeled using only one long-chain cooperativity term.

[9] The use of cooperativity terms renders inappropriate the single-helix approximation, which assumes that a peptide conformation is likely to contain only one helical segment. We considered explicitly states with up to three independent helical segments (*i.e.*, those with terms t^{2j}, $j \leq 3$), and each was assigned a cooperativity term according to the longest possible single helical segment they could contain. For example, terms in t^4y^{17} must describe conformations containing two helical segments of total length 17; one such conformation (with the longest and shortest possible helical segments) contains segments of lengths 3 and 14 residues. Every term in t^4y^{17} was assigned a cooperativity term $\mathrm{Coop}_{14} = z^3w^{14}$ and was multiplied by $[\theta_{14}]_{222}$. In general, terms in $t^{2j}y^k$ were treated explicitly in Eq S6 by assigning cooperativity terms $\mathrm{Coop}_{k-3(j-1)}$ and multiplying by $[\theta_{k-3(j-1)}]_{222}$.