For Short Alanine-Lysine Peptides the Helical Propensities of Lysine Residues (*s* Values) Are Strongly Temperature Dependent

Peter Renold, Kwok Yin Tsang, Linda S. Shimizu, and D. S. Kemp*

Room 18-582, Department of Chemistry
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Received May 16, 1996

Short peptides that are helically disposed in water often become more helical near 0 °C. First noted for the C- and S-peptides from ribonuclease A,[1] this effect is exhibited strikingly by the simple peptide Ac-$A_4KA_4KA_4KA$-NH_2 and its analogs[2,3] and has been attributed by Baldwin and co-workers to an intrinsic property of the helical polypeptide backbone.[4] For N-templated Ala-Lys conjugates, we find that enhanced helicity near 0 °C results from a large temperature dependence of the helical propensities (*s* values) of non-neighboring lysine residues rather than from an alanine backbone effect. This and other results[5] led us to reassign the causes of helicity for the Baldwin−Marqusee peptides $(A_4K)_n$ and their analogs.

When N-terminally linked to polypeptides, a conformationally restricted derivative of acetyl-prolyl-proline Ac-Hel_1 initiates α-helices and allows measurement of their helicity.[5,6] The ^1H NMR spectra of peptide−Ac-Hel_1 conjugates show two sets of resonances that correspond to slowly equilibrating *s-cis* and *s-trans* acetamido conformations (*c* and *t* states). In water the *c* state conformation is a random coil, but the *t* state is helical. For each member of a series of homologous derivatives that differ only in peptide length, the *t*/*c* state ratio derived from NMR integration is relatable to helical stability by eqs 1−3 where *A* and *B* are constants characteristic of the template and its peptide junction and *s* and s_K are the respective *s* values for alanine and lysine residues.[5−8] A plot of *t*/*c* vs length for such a homologous series is linear for $s = 1.0$ but must show positive curvature if $s > 1.0$.[5,6]

$$t/c = A + B(1 + s + s^2 + s^3 + s^4 + s^5 + ... + s^n) = B(s^{n+1} - 1)/(s - 1); \quad \text{for Ac-Hel}_1\text{-A}_n\text{-NH}_2 \quad (1)$$

$$t/c = A + B((s^{m+1} - 1)/(s - 1) + s^m s_K(s^{n+1} - 1)/(s - 1)); \quad \text{for Ac-Hel}_1\text{-A}_m\text{-K-A}_n\text{-NH}_2,\ m = 4 \text{ or } 5 \quad (2)$$

$$t/c = A + B((s^{j+1} - 1)/(s - 1) + s^j s_K((s^{j+1} - 1)/(s - 1) + s^j s_K(s^{k+1} - 1)/(s - 1))); \quad \text{for Ac-Hel}_1\text{-A}_j\text{-K-A}_j\text{ K-A}_k\text{-NH}_2,\ j = 4 \text{ or } 2,\ k = 2 \text{ or } 4\ s = s_{Ala};\ s_K = s_{Lys} \quad (3)$$

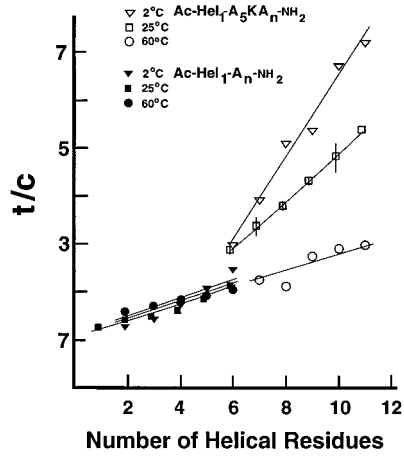

Figure 1. Temperature dependences of *t*/*c* ratios as a function of peptide helical length *n* (or *m*) for two peptide series, Ac-Hel_1-A_n-NH_2, $n = 1-6$ and Ac-Hel_1-A_5KA_m-NH_2, $m = 0-5$ ($n = m + 6$) at 2, 25, and 60 °C. Data at 25 °C are averages of duplicate measurements with SD error bars. Both *t*/*c* values and CD spectra were measured after dilution to the micromolar range for selected members of each series and were unchanged, implying insignificant aggregation. At each temperature the observed series *t*/*c* linearities imply that $s = s_{Ala}$ is close to 1.0 and temperature independent within measurement error. Temperature-dependent changes in the *t*/*c* discontinuity at $n = 7$ between the two series and for the slopes for the Lys series imply that $s_K = s_{Lys}$ is strongly temperature dependent. The lines of the figure were calculated[9] from s_K obtained from model 2 with fixed *A*, *B*, and *s*. Lines drawn using the two models[9] were nearly indistinguishable.

Figure 1 depicts the experimental dependence of *t*/*c* on peptide length in D_2O at 2°, 25°, and 60 °C for two series, Ac-Hel_1-A_n-NH_2 where $n = 1-6$ and Ac-Hel_1-A_5KA_m-NH_2 where $m = 0-5$. If the temperature dependence of *t*/*c* depends primarily on a backbone effect, it must be shared among all amino acid residues and both *s* and s_K must decrease with temperature. A plot of *t*/*c* vs helix length for each series must then show positive, temperature-dependent curvature. Alternatively, if s_K decreases significantly with temperature but *s* does not, the plot must be nearly linear within each alanine series, reflecting the lack of temperature dependence of *s*, but both $\Delta(t/c)$ at the junction between series and the slope of *t*/*c* vs *n* for the lysine-containing series must decrease with increasing temperature. Figure 1 embodies all of the latter features.

For a given data set, the statistically allowable choices for *s* and s_K vary inversely, with s_K showing the larger range. At each temperature, *s* and s_K can be calculated from *t*/*c* values by an iterative least squares analysis,[6c] yielding optimal values at 2, 25, and 60 °C of $s = 0.96$, 0.97, and 0.97 (SD 0.055) and $s_K = 4.3$, 2.9, and 1.2.[9] Alternatively, from 35 *t*/*c* values for Ac-Hel_1-A_iKA_j-NH_2 where $i = 2-6$ and $j = 1-6$ in D_2O at 25 °C we have assigned *s* more precisely as 1.02 (SD 0.035) with template constants $A = 0.832$, $B = 0.156$.[10] Since both *s* and these constants[6a] are temperature independent within the errors of the analysis, they can be used in a noniterative least squares analysis[9] to assign values at 2, 25, and 60 °C of $s_K = 4.96$,

(1) Brown, J. E.; Klee, W. A. *Biochemistry* **1971**, *10*, 470−476. Silverman, D. N.; Kotelchuck, D.; Taylor, G. T.; Scheraga, H. A. *Arch. Biochem. Biophys.* **1972**, *150*, 757−766. Bierzynski, A.; Kim, P. S.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 2470−2474.

(2) Marqusee, S.; Robbins, V. H.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5286−5290.

(3) Merutka, G.; Lipton, W.; Shalongo, W.; Tark, S.-H.; Stellwagen, E. *Biochemistry* **1990**, *29*, 7511−7515. Merutka, G.; Shalongo, W.; Stellwagen, E. *Biochemistry* **1991**, *30*, 4245−4248. Todd, A. P.; Millhauser, G. L. *Biochemistry* **1991**, *30*, 5515−5523. Fiori, W. R.; Miick, S. M.; Millhauser, G. L. *Biochemistry* **1993**, *32*, 11957−11962. Jasanoff. A.; Fersht, A. R. *Biochemistry* **1994**, *33*, 2129−2135. Albert, J. S.; Hamilton, A. D. *Biochemistry* **1995**, *34*, 984−990.

(4) (a) Scholtz, J. M.; Marqusee, S.; Baldwin, R. L.; York, E. J.; Stewart, J. M.; Santoro, M.; Bolen, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 2854−2858. (b) Scholtz, J. M.; York, E. J.; Stewart, J. M.; Baldwin, R. L. *J. Am. Chem. Soc.* **1991**, *113*, 5102−5104.

(5) Groebke, K.; Renold, P.; Tsang, K.-Y.; Allen, T. J.; McClure, K. F.; Kemp, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **1996,** 2025−2029.

(6) (a) Kemp, D. S.; Allen, T. J.; Oslick, S. L. *J. Am. Chem. Soc.* **1995**, *117*, 6641−6657. (b) Kemp, D. S.; Allen, T. J.; Oslick, S.; Boyd, J. G. *J. Am. Chem. Soc.* **1996**, *118*, 4240−4248. (c) Kemp, D. S.; Oslick, S. L.; Allen, T. J. *J. Am. Chem. Soc.* **1996**, *118*, 4249−4255.

(7) Cammers-Goodwin, A.; Allen, T. J.; Oslick. S. L.; McClure, K.; Lee, J. H.; Kemp, D. S. *J. Am. Chem. Soc.* **1996**, *118*, 3082−3090.

(8) Kemp, D. S.; Curran, T. P.; Davis, W. M.; Boyd, J. G.; Muendel, C. C. *J. Org. Chem.* **1991**, *56*, 6672−6682. Kemp, D. S.; Boyd, J. G.; Muendel, C. C. *Nature* **1991,** *352*, 451−454.

S0002-7863(96)01646-0 CCC: $12.00

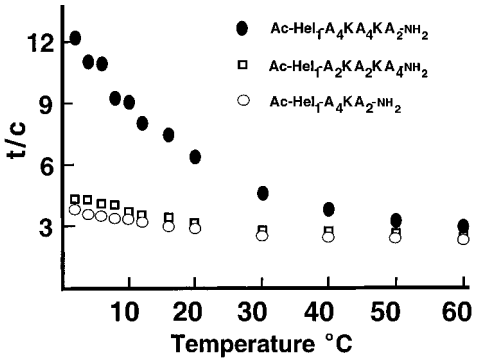

Figure 2. Temperature dependences of *t*/*c* values for monolysine derivative Ac-Hel$_1$-A$_4$KA$_2$-NH$_2$ and dilysine derivatives Ac-Hel$_1$-A$_4$-KA$_4$KA$_2$-NH$_2$ and Ac-Hel$_1$-A$_2$KA$_2$KA$_4$-NH$_2$. For the latter, the small change in *t*/*c* implies a corresponding small temperature dependence for $\langle s_K \rangle$, as analyzed in Figure 3.

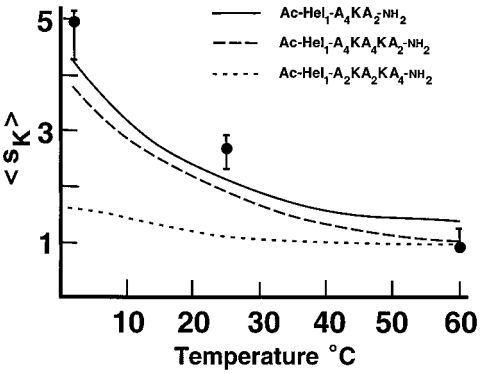

Figure 3. Calculated temperature dependences for s_K values. Solid and dashed curves correspond to average s_K values $\langle s_K \rangle$ calculated from the data of Figure 2, using *A*, *B*, and *s* of model 2,[9] eq 2 with $m = 4$ and $n = 2$ for the mono-Lys derivative, and eq 3 with $j = 4$ or 2 and $k = 2$ or 4 for the di-Lys derivatives. The significantly smaller $\langle s_K \rangle$ values for the A$_2$KA$_2$K conjugate is consistent with the postulated large charge destabilization for lysine residues separated by less than one helical loop. The filled circles correspond to the s_K values obtained for the Ac-Hel$_1$-A$_5$KA$_m$-NH$_2$ series of Figure 1 and derived from model 2; the lower limit of the error bars correspond to model 1-derived values.

2.69, and 0.91, which are linked to the larger, self-consistent data set. A van't Hoff analysis of these s_K yields a ΔH of -5.3 kcal/mol for the transfer of one lysine residue at the terminus of a pre-existing templated pentapeptide helix from a random coil to a helical state.

Figure 2 shows the temperature dependences of *t*/*c* for a mono-A$_4$K conjugate Ac-Hel$_1$-A$_4$KA$_2$-NH$_2$, for a di-A$_4$K conjugate Ac-Hel$_1$-A$_4$KA$_4$KA$_2$-NH$_2$, and for Ac-Hel$_1$-A$_2$KA$_2$KA$_4$-NH$_2$, which bears proximate dilysines in an A$_2$KA$_2$K motif. Figure 3 shows s_K values calculated from these *t*/*c* data using eqs 2 and 3, with $A = 0.832$, $B = 0.156$, and $s = 1.02$. Included for comparison are the three s_K values obtained by analysis of the data of Figure 1, which are expected to be larger than those obtained from the A$_4$K peptides, given that s_K is position dependent at 25 °C, increasing with separation from the template.[5] The agreement of s_K values for the A$_4$KA$_2$ and A$_4$-KA$_4$KA$_2$ conjugates is excellent at all temperatures, but much smaller s_K values are observed at temperatures below 40 °C for the A$_2$KA$_2$K conjugate.

These observations are pertinent to results of Marqusee, Robbins, and Baldwin,[2] who noted that the helicity at 1 °C for Ac-(A$_4$K)$_3$A-NH$_2$ decreases by ca 75% when three additional lysines are incorporated, forming Ac-(AKA$_2$K)$_3$A-NH$_2$. They concluded that lysine is helix-destabilizing relative to alanine and that the high helicity observed for A$_4$K peptides must be attributed to an exceptionally large s_{Ala}. Given our observation that s_K is abnormally small near 0 °C for sequences KA$_2$K, this conclusion and the analyses that follow from it appear to be unwarranted.

Consistent with analyses of Scheraga et al.,[13] we have found that at 25 °C the lysine side chain of Ac-Hel$_1$-A$_n$KA$_m$-NH$_2$ conjugates packs against the helix barrel and s_K values are strongly site dependent.[5] We hypothesize that owing to the operation of well-known entropic, hydrophobic, and charge−dipole effects,[11−13] the magnitude of s_K is governed by a strongly temperature dependent equilibrium between two helical manifolds, one characterized by helix-packed lysine side chains, the other by solvent-exposed lysine side chains. Electrostatic repulsion between proximate, charged lysine residues should selectively destabilize the helix packed manifold for KA$_n$K motifs with $n < 4$. Moreover, the amino acids that bear long, straight side chains terminating in charged or polar groups (Lys, Arg, Gln, Glu) are expected to show maximal $\Delta(s_X)/\Delta T$ values that may be strongly sequence-sensitive, since packing results in contacts at $(i - 3)$ or $(i - 4)$ sites, and replacement of alanine at these sites by bulkier amino acids may perturb packing interactions.

Lys and Arg residues have strong solubilizing and helix-enhancing properties, and they are often incorporated into Ala-rich host peptides designed to study the relationships between amino acid composition and helicity. Owing to the contributions of side chain packing and the resulting temperature and context sensitivities of s_K and s_R, interpretations of helicities for these peptides can be equivocal. Studies directed at clarifying the roles of Arg, Lys, Orn, and His in stabilizing alanine-rich helical Ac-Hel$_1$ conjugates will be reported subsequently.

Acknowledgment. Financial support from the National Institutes of Health Grant 5 R37 GM 13453, from the National Science Foundation Grant No. 9121702-CHE, and from Pfizer Research is gratefully acknowledged. We thank Professor K. Biemann and co-workers for mass spectrometric analyses.

JA961646N

(9) Two models were used to obtain optimal values of s_K at 2, 25, and 60 °C. In model 1, iterative least squares analysis was applied to data at each temperature to obtain the s value that minimizes $\Sigma\delta_i 2$, together with an s_K value.[5,6] This model yields template constants *A* and *B* that are sensitive to errors in series data points for small *m*: for 2, 25, and 60 °C, $A = 0.64$, 0.75, 0.90; $B = 0.26$, 0.21, 0.19; $s = s_{Ala} = 0.96$, 0.97, 0.97 and $s_K = s_{Lys} = 4.3$, 2.9, 1.2, respectively. Model 2 uses *A*, *B*, and *s* obtained by least squares analysis of *t*/*c* data at 25 °C for 35 derivatives Ac-Hel$_1$-A$_i$KA$_j$-NH$_2$, $i = 2-6$, $j = 1-6$, under the assumption that s_K is dependent on *i*. Temperature dependent s_K values assigned in this study are tied to constants assigned from the large 25 °C data base by a linear regression in *t*/*c* with respect to s_K using $A = 0.832$, $B = 0.156$, and $s = 1.02$ at all temperatures; at 2, 25, and 60 °C, this yields $s_K = 4.96$, 2.69, and 0.91, shown as filled circles in Figure 3. The precision in *t*/*c* is ±5%,[6a,c] and a 5% increase over the minimum $\Sigma(t/c_{exp} - t/c_{calcd})^2$ results in a 4% change in *s*; at fixed T, pairs of values for *s* and s_K are inversely correlated, and a small change in *s* results in a large change in s_K.[5] Likely error ranges for s_K are ±10%.

(10) Tsang, K.-Y.; Renold, P.; Kemp, D. S. Unpublished observations.

(11) Creamer, T. P.; Rose, G. D. *Proc. Natl. Acad. Sci U.S.A.* **1992**, *89*, 5937−5941.

(12) Fairman, R.; Schoemaker, K. H.; York, E. J.; Stewart, J. M.; Baldwin, R. L. *Proteins: Struct., Funct., Genet.* **1989**, *5*, 1−7. Perutz, M. F.; Fermi, G. *Proteins: Struct., Funct., Genet.* **1988**, *4*, 294−295.

(13) Vila, J.; Williams, R. L.; Grant, J. A.; Wojcik, J.; Scheraga, H. A. *Proc. Nat. Acad. Sci. U.S.A.* **1992**, *89*, 7821−7825. Nemethy, G.; Scheraga, H. A. *J. Phys. Chem.* **1962**, *66*, 1773−1789. Nemethy, G.; Steinberg, I. Z.; Scheraga, H. A. *Biopolymers* **1963**, *1*, 43−63.