α-Helix Nucleation Constant in Copolypeptides of Alanine and Ornithine or Lysine

Jianxin Yang, Kang Zhao, Youxiang Gong, Alexander Vologodskii, and Neville R. Kallenbach*

Contribution from the Department of Chemistry, New York University, New York, New York 10003

Received July 2, 1998

Abstract: The α-helix is one of a small number of fundamental structural motifs of the peptide backbone that are abundant in native proteins. The forces responsible for its stability have attracted a great deal of theoretical and experimental attention. Helix formation can naturally be considered in terms of two processes, initial nucleation of a helix from a sequence of disordered residues and propagation or growth of helical structure from such a nucleus. Data from a variety of short peptide and polypeptide models have revealed more details about the propagation of helix structure than about helix nucleation. To investigate helix nucleation, we have synthesized two series of high molecular weight polypeptides containing differing ratios of alanine and one of the basic side chains, ornithine or lysine; poly L-alanine is insoluble in water. The CD signals of these copolymers have been analyzed by a program that evaluates the helix content in terms of the experimental chain-length distribution and composition, with the helix nucleation constant (σ value) and the propagation constants (s values) for the amino acids involved. Fitting the CD data allows the determination of the propagation constants for Orn ($s = 0.45$) and Lys ($s = 0.8$) in addition to that of the helix nucleation constant once the s value for Ala is specified. The value of σ is sensitive to the dependence of the CD signal on helix length; using the Yang equation, $[\theta]_{222} = -41000(n - x)/n$, with $x = 2.5$, the nucleation constant value is $\sigma = 0.004 \pm 0.002$ at 4 °C in the presence of 1 M salt. This value is consistent with earlier estimates based on analysis of the helix−coil transition in poly(Lys), poly(Glu), and shorter Ala-rich peptides. However, if x is taken to be zero, the resulting σ value is 0.02, considerably larger than the above estimates.

Introduction

Studies of a large number of short helical model peptides in aqueous solution have defined many of the dominant factors that govern α-helix stabilization in aqueous solution.[1] These include the H-bonding of the backbone NH and CO groups originally postulated by Pauling,[2] electrostatic[3] and van der Waals interactions[4] between appropriately spaced side chains, capping of residues at the ends of short chains,[5] and interactions between charged or polar side chains and the helix dipole.[6] Data on model peptides have been interpreted by means of helix−coil transition models[7] that include contributions from each of the above interactions.[8] Much of the information available has resulted from analysis of short helical peptide models or fragments from proteins, primarily relying on CD spectroscopy to assess the extent of helical structure present. While the helix propagation constants of most natural side chains apart from alanine do not stabilize α-helical structure, hydrophobic interactions, salt bridges, and packing interactions can all exert significant helix-stabilizing effects.[1,8] One objective of these efforts is then to determine the equilibrium effect of these interactions on helix propagation.[8,9] In the case of helix-capping interactions, side chains such as Ser and Glu near the N termini of helices and the Gly residue at the C terminus can interact with NH or CO groups of the main chain, stabilizing helical structure.[5] However, since most chains that have been studied experimentally contain between 11 and 22 residues, detailed information concerning the nucleation of helical structure is sparse, relative to that for the propagation reaction. The process of nucleating α-helix formation is of fundamental importance in developing a detailed understanding of the kinetics and mechanism of helix formation. In this study, we investigate high molecular weight Ala-rich polypeptide chains containing Orn or Lys side chains for solubility. The chains are pseudo-random copolymers containing different ratios of Ala and Orn or Lys. Measurements of the CD spectra of these polypeptides allow determination of the helix nucleation constant, assuming a value for the propagation constant of alanine.

(1) Kallenbach, N. R.; Lyu, P. C.; Zhou, H. X. *Circular Dichroism and the Conformational Analysis of Biomolecules*; Plenum Press: New York, 1996; pp 201−259.

(2) Pauling, L.; Corey, R. B.; Branson, H. R. *Proc. Natl. Acad. Sci. U.S.A.* **1951**, *37*, 205−211.

(3) (a) Lyu, P. C.; Gans, P. J.; Kallenbach, N. R. *J. Mol. Biol.* **1992**, *223*, 343−350. (b) Scholtz, J. M.; Qian, H.; Robbins, V. H.; Baldwin, R. L. *Biochemistry* **1993**, *32*, 9668−9676.

(4) Richards, F. M. *Annu. Rev. Biophys. Bioeng.* **1977**, *6*, 151−176.

(5) (a) Presta, L. G.; Rose, G. D. *Science* **1988**, *240*, 1632−1641. (b) Lyu, P. C.; Zhou, H. X.; Jelveh, N.; Wemmer, D. E.; Kallenbach, N. R. *J. Am. Chem. Soc.* **1992**, *114*, 6560−6562. (d) Lyu, P. C.; Wemmer, D. E.; Zhou, H. X.; Pinker, R. J.; Kallenbach, N. R. *Biochemistry* **1993**, *32*, 421−425. (e) Aurora, R.; Srinivasan, R.; Rose, G. D. *Science* **1994**, *264*, 1126−1130.(f) Doig, A. J.; Chakrabartty, A.; Klinger, T. M.; Baldwin, R. L. *Biochemistry* **1994**, *33*, 3396−3403.

(6) (a) Hol, W. G. J. *Prog. Biophys. Mol. Biol.* **1985**, *45*, 149−195. (b) Lockhart, D. J.; Kim, P. S. *Science* **1993**, *260*, 198−202.

(7) (a) Zimm, B. H.; Bragg, J. K. *J. Chem. Phys.* **1959**, *31*, 526−535. (b) Lifson, S.; Roig, A. *J. Chem. Phys.* **1961**, *34*, 1963−1974.

(8) (a) Rohl, C. A.; Chakrabartty, A.; Baldwin, R. L. *Protein Sci.* **1996**, *5*, 2623−2637; (b) Muñoz, V.; Serrano, L. *Nat. Struct. Biol.* **1994**, *1*, 399−409. (c) Andersen, N. H.; Tong, H. *Protein Sci.* **1997**, *6*, 1920−1936. (d) Yang, J.; Spek, E. J.; Gong, Y.; Zhou, H. X.; Kallenbach, N. R. *Protein Sci.* **1997**, *6*, 1264−1272.

(9) Finkelstein, A. V.; Ptitsyn, O. B. *J. Mol. Biol.* **1976**, *103*, 15−24.

10.1021/ja982319d CCC: $15.00
Published on Web 10/01/1998

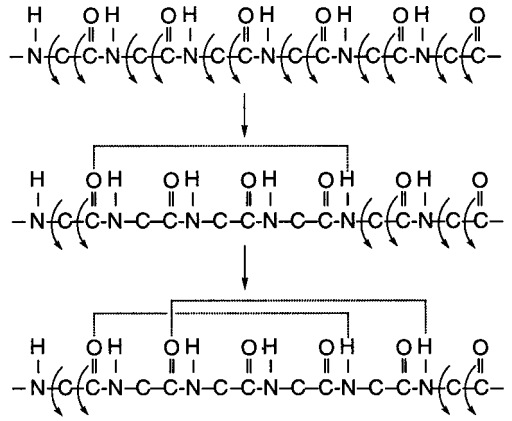

Figure 1. Nucleation process in an α-helix. The figure shows that six dihedral angles are fixed to form the first hydrogen bond, while two are needed to form a bond adjacent to an existing one. Curved arrows represent the dihedral angles, and dotted lines represent the hydrogen bond (adapted from Cantor, C. R. and Schimmel, P. R. *Biophysical Chemistry III*; W. H. Freeman: New York, 1980; p 1049.

Figure 1 illustrates the process of nucleating an α-helix, with the successive addition of a helical residue to the nucleated helix. As shown, the process of nucleation mandates a set of three adjacent residues in helical conformation, that is, having values of the ϕ and ψ dihedral angles within the right-handed α-helix domain, $\phi = -64° \pm 6°$ and $\psi = -40° \pm 8°$. Following an argument due to Zimm and Bragg,[7a] if each of two dihedral angles per segment (in the nucleation of a homopolypeptide) is restricted by a fraction r of the phase space available to random coil, then the propagation constant would be proportional to r^2h, where h is the equilibrium constant for H-bonding in the backbone at the expense of solvent H-bonds. Forming the first H-bond would then correspond to r^6h, since six restricted angles are involved, so that the nucleation constant, σ in Zimm and Bragg's[7a] notation or v^2 according to Lifson and Roig,[7b] should be roughly proportional to r.[4] Thus, if the restriction on each dihedral angle compatible with nucleation is 0.1, σ will be 10^{-4}, leading to highly cooperative formation of helical structure. If the restriction is less severe, for example, 0.3, helix formation becomes less cooperative, and values of σ around 10^{-2} result. Reported values are within these limits, as discussed below. An equivalent view of helix cooperativity is that it reflects the loss of favorable bonding interactions in four residues at each end of a helix segment relative to that in residues in the middle.[10] The extent of fraying at the ends has been used directly to evaluate s in short, artificially nucleated chains,[11b] but this does not allow precise determination of the nucleation constant.

Since helix initiation may well include transient 3_{10} helix formation in addition to that of the α-helix and conceivably other conformations as well, helix initiation may not be an all-or-none process in the sense just described and may, in fact, vary with the number and kind of residues present. In the context of a high molecular weight polypeptide, the nucleation process can be approximated as all-or-none so that a chain of 50% helix consists of alternating and fluctuating sequences of helix and coil residues with equal numbers of residues on average. The mean length of each sequence depends both on the nucleation constant and the composition of the side chains in the molecule.

We apply a statistical−mechanical analysis to describe the helix−coil transition in long chains composed of two types of residues.[12] It is assumed that the corresponding sequences are random. Therefore, to simulate the properties of polypeptides with specific average Ala/Orn or Ala/Lys composition, we generated a large population of random sequences with corresponding composition and calculated the average properties of the ensemble, including the helix content. The value(s) of the model parameters, σ, s_{Ala}, and s_{Orn} or s_{Lys} which give the best fit to the measured helix content values were determined. The value of s_{Ala} alone is held fixed since this has been accurately determined using both hydrogen exchange[11] and CD experiments on short peptides.[1] Values for s_{Orn} and s_{Lys} have been determined in an alanine-rich environment by Baldwin's group;[13] these values can be specified as well. In practice, the given values are slightly different from the ones that give minimal residuals in our calculations.

Results

Synthesis of Poly(Ala-Orn) and Poly(Ala-Lys) Polypeptides. The preparation of monomers and their polymerization are illustrated in Figure 2. Experimental properties of the two sets of polypeptides we synthesized, one containing Ala and Orn side chains and the second containing Ala and Lys side chains, are summarized in Table 1. The composition of the polypeptides does not match the input NCA ratios. Thus, for an input reaction ratio of Ala/Orn = 2/1, the polymer contains 1.84/1, determined using NMR analysis. For a reaction ratio of Ala/Orn = 5/1, the resulting polymer contains 4.59/1. The ratio (output/input) varies between 0.89 and 0.92 for the Ala/Orn copolymers. For polypeptides consisting of alanine and lysine, the ratio varies between 0.86 and 1.01. We did not obtain a product with reaction ratio of Ala/Lys = 5/1, possibly due to its limited solubility in water.

CD Measurements. The CD spectra of all polypeptides showed α-helix signals (minima at 222 and 208 nm, maximum at 195 nm) characteristic of a mixture of α-helix and coil (Figure 3). The isodichroic point at around 203 nm is consistent with a two-state description for each residue in the polypeptides.[14] Taking the mean residue molecular weights (see Experimental Section) as the theoretical molecular weight and the limiting CD of an infinitely long helix to be $[\theta]_{222} = -41\,000$ which was measured using a synthetic poly(Ala) linked to PEG for water solubility (U. Skarpidis, unpublished results), we can estimate apparent helicity from the molar residue ellipticities for all polypeptides at 222 nm (Table 1). The helix content of the chains ranges from 50−80%, depending on the composition. As expected if Ala is helix-stabilizing, the helix content from the CD signal of the Ala/Orn polypeptides becomes progressively greater as the alanine content increases. By contrast, in the Ala/Lys series, the effect is much weaker, despite the difference between the s values for Lys and Ala that have been determined in short-chain peptide models (see Table 1 and Figure 3).[13,15]

The effects of temperature and peptide concentration on the CD signal of the polypeptides are illustrated in Figure 4. There

(10) Schellman, J. A. *C. R. Lab. Carlsberg Ser. Chim.* **1955**, *29*, 230−259.

(11) (a) Rohl, C. A.; Baldwin, R. L. *Biochemistry* **1994**, *33*, 7760−7764. (b) Zhou, H. X.; Hull, L. A.; Kallenbach, N. R.; Mayne, L.; Bai, Y.-W.; Englander, S. W. *J. Am. Chem. Soc.* **1994**, *116*, 6482−6483.

(12) Vedenov, A. A.; Dykhne, A. M.; Frank-Kamenetskii, M. D. *Sov. Phys. Usp.* **1972**, *14*, 715−736.

(13) Chakrabartty, A.; Kortemme, T.; Baldwin, R. L. *Protein Sci.* **1994**, *3*, 843−852.

(14) (a) Yang, J. T.; Wu, C.-S. C.; Martinez, H. M. *Methods in Enzymol.* **1986**, *130*, 208−269. (b) Manning, M. C.; Woody, R. W. *Biopolymers* **1991**, *31*, 569−586. (c) Gans, P. L.; Lyu, P. C.; Manning, M. C.; Woody, R. W.; Kallenbach, N. R. *Biopolymers* **1991**, *31*, 1605−1614.

(15) Padmanabhan, S.; York, E. J.; Stewart, J. M.; Baldwin, R. L. *J. Mol. Biol.* **1996**, *257*, 726−734.

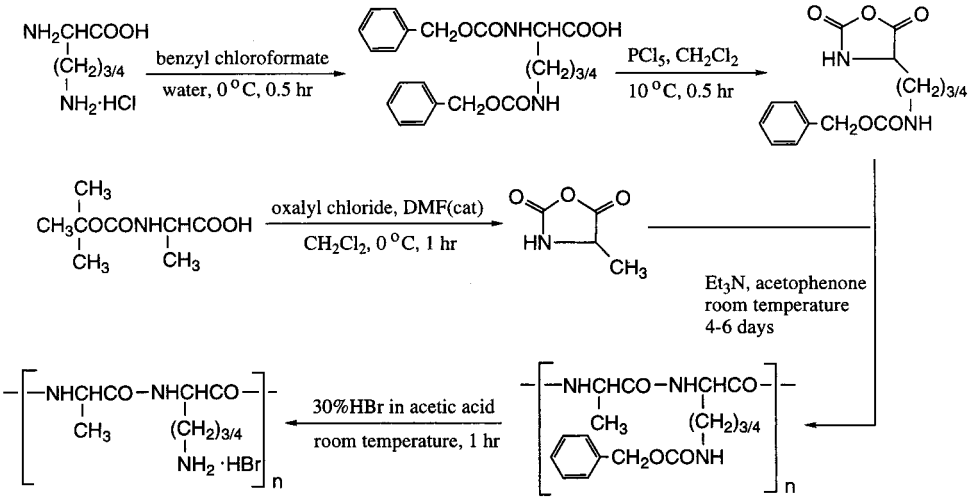

Figure 2. Synthetic scheme for the polypeptides of this study.

Table 1. Experimental Properties of Polypeptides

	molecular weights[a]			Ala/X Ratio				
polypeptide	average	mean residue	DP	input	observed[b]	$[\theta]_{222}$[c]	x_{Ala}[d]	helicity(%)[e]
X = Orn [Poly(Ala-Orn)]								
55−04	23200	114.6	202	2/1	1.84/1	−21400	0.650	0.528
55−13	18200	103.1	176	3/1	2.84/1	−27800	0.740	0.688
55−23	25500	98.2	260	4/1	3.56/1	−30100	0.780	0.741
55−33	17600	93.2	189	5/1	4.59/1	−31900	0.820	0.788
X = Lys [Poly(Ala-Lys)]								
62−04	31000	116.8	265	2/1	2.01/1	−30800	0.670	0.758
62−13	18700	108.0	173	3/1	2.73/1	−31700	0.730	0.785
62−23	17400	102.0	171	4/1	3.45/1	−32400	0.775	0.802

[a] Average molecular weights were estimated by size exclusion chromatography and represented as number average molecular weight M_n.[30] Mean residue molecular weights were calculated from the ratio of Ala/Orn(Lys) determined by NMR. [b] Calculated from the integrals of NMR spectra. [c] CD spectra were measured in 1 M NaCl/20 mM Na/PO_4, pH 7, 4 °C in a 1-mm path-length cell. The polypeptide concentrations were 50 μg/mL. The ellipticities ($[\theta]_{222}$ (deg·cm^2/dmol) in this table) were calculated by taking the mean residue molecular weights as the molecular weights used in calculation. [d] The molar fraction of alanine in polypeptides. [e] The helicities were calculated by dividing $-41000(n - x)/n$, where $x = 2.5$, from $[\theta]_{222}$ of all polypeptides.

is significant residual ellipticity at 85 °C (Figure 4, panel A) for polypeptides with higher alanine ratio (such as 55-33). Moreover, the peptide concentration has no effect over the concentration range of 12.5 μg/mL to 200 μg/mL (Figure 4, panel B).

Determination of the Nucleation Constant of Alanine. Determining the nucleation constant σ requires specification of the propagation constants s_{Ala}, s_{Lys}, and s_{Orn} together with the compositions and length distributions of the polypeptides. Values for the former parameters have been determined in short Ala-rich peptides by Baldwin's group, who reported s values of 1.4−1.6 for Ala, 0.90 for Lys, and 0.56 for Orn in an aqueous buffer with 1 M NaCl and at 0 °C.[15] Other estimates of s_{Ala} have been published; hydrogen exchange rate measurements on a prenucleated helical peptide, for example, give a value $s_{Ala} = 1.7$ at 4 °C in good agreement with the preceding estimates. However, the differences prove to have a negligible effect on the σ value determined in this study: using $s_{Ala} = 1.4-1.7$, the σ values change only by 10%. Similarly, it can be seen that the residuals are lower for $s_{Orn} = 0.45$ and $s_{Lys} = 0.80$ than for the values determined by Baldwin or in our previous work.[15] The effect on the value of σ is not large, that is, less than 10%. We have also investigated the effect of the chain-length distribution on the σ value. This again proves to be weak, so that the value of σ is not affected by assuming forms for the chain-length distribution different from the experimentally measured one.

By contrast, the deconvolution of CD values using eq 2 reveals that the σ value is sensitive to the value of the cutoff constant, x, in the Yang equation.[29] This effect is shown in Table 2, where it can be seen that σ varies from 0.02 if $x = 0$

(16) (a) Barskaya, T. V.; Ptitsyn, O. B. *Biopolymers* **1971**, *10*, 2181−2197. (b) Bychkova, V. E.; Ptitsyn, O. B.; Barskaya, T. V. *Biopolymers* **1971**, *11*, 2161−2179.

(17) (a) Wojcik, J.; Altmann, K.-H.; Scheraga, H. A. *Biopolymers* **1990**, *30*, 121−134. (b) Scholtz, J. M.; Qian, H.; York, E. J.; Stewart, J. M.; Baldwin, R. L. *Biopolymers* **1991**, *31*, 1463−1470.

(18) Padmanabhan, S.; York, E. J.; Gera, L.; Stewart, J. M.; Baldwin, R. L. *Biochemistry* **1994**, *33*, 8604−9809.

(19) von Dreele, P. H.; Lotan, N.; Ananthanarayanan, V. S.; Scheraga, H. A. *Macromolecules* **1971**, *4*, 408−417.

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

(21) (a) Blaber, M.; Zhang, X.-J.; Matthews, B. W. *Science* **1993**, *260*, 1637−1640. (b) Blaber, M.; Zhang, X.-J.; Matthews, B. W. *Science* **1993**, *262*, 917−918. (c) Myers, J. K.; Pace, C. N.; Scholtz, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2833−2837.

(22) (a) Vasquez, M.; Scheraga, H. A. *Biopolymers* **1988**, *32*, 41−58. (b) Ingwall, R. T.; Scheraga, H. A.; Lotan, N.; Berger, A.; Katchalski, E. *Biopolymers* **1968**, *6*, 331−368; (c) Groebke, K.; Renold, P.; Tsang, K. Y.; Allen, T. J.; MaClure, K. F.; Kemp, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 4025−4029. (d) Kemp, D. S.; Oslick, S. L.; Allen, T. J. *J. Am. Chem. Soc.* **1996**, *118*, 4249−4255.

(23) Bierzynski, A.; Kim, P. S.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 2470−2474.

(24) Waley, S. G.; Watson, J. *Proc. R. Soc.* **1949**, *A199*, 499−517.

(25) (a)Katchalski, E.; Spitnik, P. *J. Am. Chem. Soc*, **1951**, *73*, 3992−3995. (b) Blout, E. R.; Karlson, R. H. *J. Am. Chem. Soc*. **1956**, *78*, 941−946.

(26) (a) Ben-Ishai, D.; Berger, A. *J. Org. Chem*. **1952**, *17*, 1564−1570. (b) Ben-Ishai, D. *J. Org. Chem*. **1954**, *19*, 62−66.

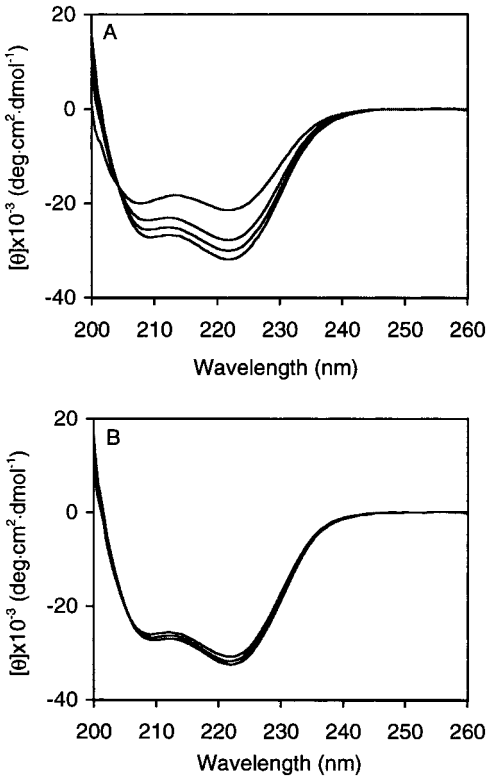

Figure 3. CD spectra of polypeptides. (A) Poly(Ala-Orn). The curves correspond to the samples 55-04, 55-13, 55-23, 55-33, respectively, in descending order near 222 nm. (B) Poly(Ala-Lys). The curves are 62-04, 62-13, 62-23, respectively, in descending order near 222 nm. CD spectra were measured in 1 M NaCl/20 mM Na/PO_4, pH 7, 4 °C in a 1-mm path length cell. The polypeptide concentrations were 50 μg/mL.

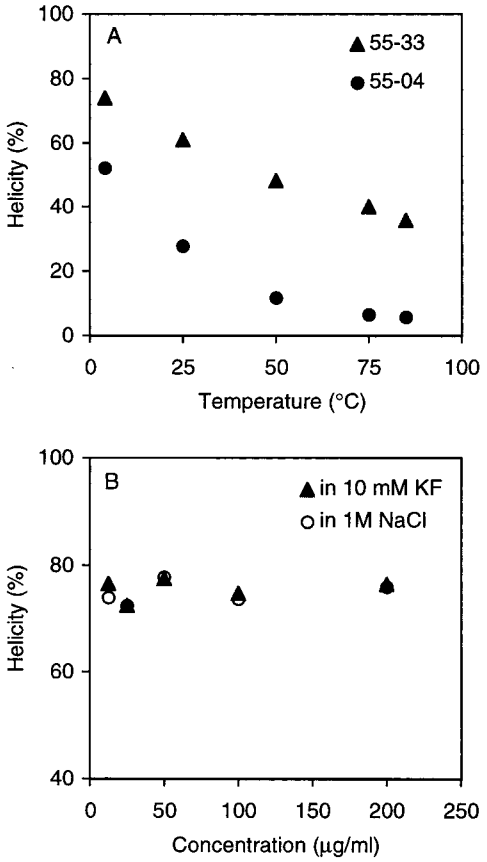

Figure 4. The temperature and concentration dependence of polypeptides 55-04 and 55-33. (A) Temperature measurement points were 4 °C, 25 °C, 50 °C, 75 °C, and 85 °C. (B) The concentration range of polypeptide 55-33 was from 12.5 μg/mL to 200 μg/mL in 10 mM KF/1 mM phosphate and 1 M NaCl/20 mM Na/PO_4.

to 0.004 for $x = 4$, holding the s values constant. If we assume that the cutoff is close to 3, the value of σ minimizes the residual, and an uncertainty of a factor of 2 is introduced. The reason for this behavior can be understood in terms of the variation of the mean helix length in each polypeptide as a function of x; for $x = 0$, these range from 8 to 13 residues, while for $x = 4$, the mean helix length is significantly greater, 17−36 residues. The mean helix lengths in these polypeptides are short enough for the correction imposed by the length dependence of the CD signal to be significant.

Table 2. Relation Between σ Value and Cut-Off Constant[a]

cut-off (x)	sigma	residual
0	0.016	0.0040
2	0.008	0.0026
3	0.004	0.0015
4	0.004	0.0032

[a] Value of σ is sensitive to the value of the cutoff constant, x. It can be seen that σ varies from 0.02 if $x = 0$ to 0.004 for $x = 4$.

Discussion

Polypeptide Design. Various polypeptide models have previously been used to estimate σ; values have been proposed that vary from below 10^{-4} to about 10^{-2}, depending on the polypeptide used and the solution conditions.[16,17a] In short, synthetic peptides of variable chain length, up to a 50mer, fitting the temperature-dependent CD spectra or HX rates to s values, including an enthalpy term for the propagation constant, led to an estimate of $\sigma = 0.002$ for the nucleation constant,[17b,11a] which agrees closely with our determination here.

The large variation in σ values obtained from Scheraga's[17] host−guest polypeptide models seems puzzling in the light of our results. A possible reason may be the particular helix-forming properties of the derivatized (hydroxybutyl- or hydroxypropyl)-L-glutamine side chains used as the host residue in the host−guest copolymer series.[18] The values of s (and potentially σ as well) determined for guest residues in host−guest experiments might be strongly influenced by helix-stabilizing (side chain−side chain) interactions between the host residues. To determine whether the values of σ for this system of host polypeptides are sensitive to the value of x, we recalculated the CD data of von Dreele et al.[19] on copolymers of HBG and HPG, the hydroxy-butyl and hydroxy-propyl glutamine residues. With $x = 2.5$ and s values close to those specified by von Dreele et al.,[19] we find a good fit to the data for $\sigma = 0.0002$, suggesting higher cooperativity than we find in the Ala-based polypeptides of the present study. If the s values are allowed to vary, σ and the s values for HBG and HPG cannot all be specified uniquely from the data we used.

The Helix Nucleation and Helix−Coil Transition Models. A variety of model helical peptides, polypeptides, and proteins have been investigated in the effort to define the determinants of α-helix stability. Short oligopeptide models tend to provide evidence that consistent scales of helix propagation constants (s values in Zimm−Bragg[7a] notation, or w values in Lifson−Roig[7b] notation) for individual amino acids can be determined experimentally, provided a sufficient set of side-chain−side-chain interactions can be calibrated and included to allow for the numerous potential modulating effects of sequence.[8] Thus, a matrix of side-chain−side-chain interactions,[9] together with

N- and C-terminal capping effects,[5] must be supplied in order to predict the helix content of oligopeptides regardless of context.[1,8d] Alanine-rich peptide models have played a notable role in these studies, since Ala proves to be strongly helix-stabilizing in many peptide[13] and protein[21] systems. This is attributed to the fact that its methyl side chain minimizes the unfavorable entropic effect required to confine bulkier and longer side chains into the helical geometry.[20] A roughly consistent set of *s* values for each of the standard side chains has been determined, using both peptide and protein models.[8,13,20]

The results of the above studies are at striking variance with substitution data on the HBG or HPG polypeptides from Scheraga's laboratory,[17a] block oligopeptide data from Scheraga's group,[22b] and data on a series of short, prenucleated peptides studied by Kemp's group.[22c,d] In these dissenting reports, Ala is found to be helix-indifferent rather than helix-stabilizing. The helix-stabilizing effect of Ala found in other model systems is attributed instead to different effects exerted by neighboring side chains, in particular, the charged side chains needed to maintain solubility. According to Scheraga,[22a] solvation effects due to the presence of charged side chains such as Lys or Arg confer an apparent helix-stabilizing effect on neighboring alanines. A different mechanism is invoked by Kemp's group; the presence of three or more adjacent Ala side chains allows a long, charged, side chain such as Lys to interact favorably with the helix backbone via a charge−dipole interaction.[22c,d] In these experiments, Lys interacts strongly via its CH_2 groups, while the shorter Orn side chain is much less effective.

The picture derived from nucleated short chains by Kemp's group is inconsistent with the results of the present study. The difficulty is that $s_{Ala} = 1.03$, the unperturbed propagation constant determined by Kemp's group, does not lead to a satisfactory fit to the CD data for the series of Ala/Orn copolymers; the values of the residuals are more than 100 times greater than those obtained with $s_{Ala} = 1.6$, regardless of the σ value or s_{Orn} value chosen. In experiments on short prenucleated helices, the presence of Lys was found to stabilize Ala neighbors more effectively than that of either Orn or Ala itself.[22c] This means that in the Ala/Lys series, the effective *s* value for Ala should be larger than that in the Orn series. In fact, the two are indistinguishable, as shown by fitting the data for each series independently. While we cannot exclude an effect of Lys on the propagation constant of adjacent Ala residues in the Ala/Lys series, the low value of s_{Ala} seen in HPG or HBG polypeptides or short Ala-containing blocks[22] is inconsistent with our CD data. Moreover, if the helix propensity of Ala is low but stimulated by neighboring Lys side chains, the dependence of the CD signal on Lys composition should be greater than in the Orn copolymer series, but it is not (Table 1).

In summary, we have used two series of high molecular weight copolypeptides containing Ala and Orn or Lys to determine the helix nucleation constant. Our result, $\sigma = 0.004 \pm 0.002$ at 4 °C and 1 M salt, agrees well with several values determined by earlier studies, based on poly(Lys),[27] poly(Glu),[28] and a series of peptides with varying chain lengths, containing repeats of Ala and Lys.[30] Our results show that a key parameter is the value of the cutoff constant *x* expressing the dependence of the CD signal on helix length (eq 2), indicating that the mean helix lengths in these molecules is low. This proves to be true also in the case of copolypeptides of HBG and HPG, the host residues used in experiments by the Scheraga group.[17a] Questions that remain to be addressed include (i) whether the value of the nucleation constant is independent of the side chain, as appears to be the case from comparison of the available results, (ii) whether the nucleation constant is temperature-dependent, as it arguably should be,[10] and (iii) if different values for the N and C nucleation events may need to be introduced, rather than the symmetric process implied by use of σ or v^2 in the standard Lifson−Roig formalism.[7,23] It is clear that the presence of short, capping sequences such as SxxE at the N terminus can strongly influence the nucleation probability for α-helical structure.[5] The nucleation process in the absence of specific signals establishes a baseline for assessing such effects relative to the intrinsic nucleation process in peptides, polypeptides, or nascent proteins.

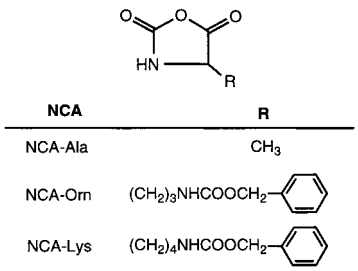

Figure 5. Structures of NCAs of alanine, ornithine, and lysine.

Experimental Section

Materials and Reagent Purification. *t*-Boc-L-alanine was purchased from Bachem, CA and used after drying on vacuum line overnight. All other reagents were from Aldrich, WI. Acetophenone was purified according to the procedure of Waley and Watson.[24] Triethylamine was dried over CaH_2. L-Ornithine and L-lysine monohydrochloride were used without further purification. Dichloromethane (DCM), *N*,*N*-dimethylformamide (DMF), hexane, petroleum ether (PE), and ethyl acetate (EtOAc) were all anhydrous form and used without further purification. All glassware was thoroughly cleaned and dried in an oven at 200 °C before use.

Peptide Synthesis and Purification. NCA polymerization has been widely used for the synthesis of polypeptides with high molecular weights.[19,25,28b] A series of Ala-Orn and Ala-Lys polypeptides were synthesized by the copolymerization of NCAs of alanine, ornithine, and/or lysine (Figure 5). After polymerization, the side-chain protection groups on ornithine and lysine residues were removed by treatment with hydrogen bromide.[22b,26] The final purification was achieved by dialysis against 0.5 M NaBr and distilled water and then lyophilization. The polypeptides were stored at −80 °C until use. Figure 2 illustrates the synthetic scheme used in this study.

Synthesis of NCAs of Alanine, Ornithine, and Lysine. To synthesize Ala-NCA, we modified the procedure developed by Johnson's group.[27] A typical synthesis is as follows: To 6 g of *t*-Boc-Ala (31.8 mmol) in 20 mL anhydrous CH_2Cl_2 was added oxalyl chloride (3.6 mL, 41.5 mmol) at 0 °C under argon atmosphere, followed by 2 drops of DMF. The reaction mixture was allowed to warm to room temperature and additional DMF (2−3 drops) was added dropwise until no further gas evolved (approximately 1 h). After the reaction was complete, the solvent was evaporated in vacuo, and the residue was purified by column chromatography on predried silica gel. Pure Ala-NCA was obtained by several recrystallizations in a DCM/hexane system in 26% yield (940 mg): ^{1}H NMR (200 MHz, acetone-d_6) δ 1.50 (d, 1H, $J = 7$ Hz), 4.59 (dq, 1H, $J = 1, 7$ Hz), 7.84 (b, 1H); ^{13}C NMR δ 18.0, 54.5, 152.9, 172.9.

The synthesis of Orn-NCA and Lys-NCA was achieved by the method of Bergmann et al. with slight modification.[28a] A typical run was as follows: To 5.3 g of dibenzyloxycarbonyl-L-ornithine (12.8 mmol[28a] in 18 mL anhydrous CH_2Cl_2 was added PCl_5 (4.0 g, 19.2 mmol)

(27) Mobashery, S.; Johnston, M. *J. Org. Chem.* **1985**, *50*, 2200−2202.
(28) (a) Bergmann, M.; Zervas, L.; Ross, W. F. *J. Biol. Chem.* **1935**, *111*, 245−260. (b) Bodanszky, M.; Bodanszky, A. *The practice of peptide synthesis*, Springer-Verlag: New York, 1994.
(29) Chen, G. C.; Yang, J. T. *Anal. Lett.* **1977**, *10*, 1195−1207.
(30) Cowie, J. M. G. *Polymers*: *Chemistry and Physics of Modern Materials*. Chapman and Hall: New York, 1991; pp 8−10.

at 10 °C under argon atmosphere, and the reaction mixture was stirred for $^1/_2$ h at 10 °C. After the reaction was complete, the solvent was evaporated in vacuo, and the residue was purified by column chromatography on predried silica gel. Pure Orn-NCA was obtained after several recrystallizations in a DCM/PE system with 29% yield (1.09 g): ^{1}H NMR (200 MHz, DMSO-d_6) δ 1.49 (m, 1H), 1.69 (m, 1H), 3.03 (m, 1H), 4.46 (m, 1H), 5.03 (s, 1H), 7.36 (m, 6H), 9.10 (b, 1H); ^{13}C NMR δ 25.4, 28.9, 40.0, 57.2, 65.5, 127.9, 128.0 (2C), 128.6 (2C), 137.5, 152.1, 156.4, 171.7. Lys-NCA was synthesized in a similar way: ^{1}H NMR (200 MHz, DMSO-d_6) δ 1.39 (m, 2H), 1.70 (m, 1H), 3.01 (m, 1H), 4.45 (m, 1H), 5.04 (s, 1H), 7.35 (m, 6H), 9.10 (b, 1H); ^{13}C NMR δ 22.0, 29.2, 31.0, 40.3, 57.4, 65.5, 128.0 (3C), 128.6 (2C), 137.5, 152.2, 156.3, 171.8.

Polymerization and Purification of Polypeptides. Polymerization reactions were carried out in acetophenone with triethylamine as initiator, with molar ratios of Ala to Orn or Lys in the range of 2/1, 3/1, 4/1, and 5/1. The ratio of amino acids to initiator (A/I ratio) was 15/1.[25] After polymerization, the side-chain protecting groups of ornithine/lysine were removed using 30% HBr in glacial acetic acid.[26] A typical run for polyAla-Orn is described below.

In a round-bottom flask protected by argon and sealed by septa, a mixture of 2 mmol NCAs (with appropriate molar ratio of 2/1, 3/1, 4/1, or 5/1 of Ala-NCA/Orn-NCA) and 18.6 μL triethylamine (0.133 mmol, A/I ratio = 15/1) were added, respectively. The flask was incubated at room temperature for 4−6 days until no increasing of viscosity was observed.

After polymerization, the viscous product was dropped into a large volume of MeOH to produce a gelatinous material, which was washed several times with MeOH, dissolved in 20 mL 30% HBr/AcOH, and stirred at room temperature for about 1 h. The mixture was washed with PE and EtOAc several times, and the resulting precipitates were dried on a vacuum line. The white precipitate was dissolved in 15 mL D_2O then dialyzed overnight in dialysis tubing with a cutoff molecular weight of 12 500 against 0.5 M NaBr and finally pure D_2O, respectively. After lyophilization, the polypeptide was obtained as a loose, white, fibrous solid.

Polypeptides of Ala-Lys were synthesized according to the same procedure as that for polypeptides of Ala-Orn. The successful synthesis of poly(Ala-Orn) and poly(Ala-Lys) polypeptides was confirmed by ^{1}H NMR (Figure 6), CD spectroscopy (Figure 3), and their solubility in water.

Determination of Residue Ratio in Synthetic Polypeptides. The residue ratio in each polypeptide was determined by ^{1}H NMR.[32] The NMR samples were prepared by dissolving about 5 mg of polypeptide into 0.6 mL D_2O with 0.1% TSP. The NMR resonances used for composition determination were at 3.1 ppm (assigned to be δ/ϵ-CH_2 of Orn/Lys residues) and at 4.15 ppm (assigned to the α-CH of Ala and Orn/Lys residues). The integrals of these two peaks were used to determine the residue ratio, giving the results in Table 1.

Determination of Average Molecular Weight and Mean Residue Molecular Weight of Polypeptides. The molecular weight distribution in each of the copolypeptides was determined by size exclusion chromatography under denaturing conditions, using a calibrated column system that detects polypeptide by absorbance at 219 nm. (Michigan Molecular Institute, Midland, MI); average and mean residue molecular weights are listed in Table 1. The weight average data were deconvoluted to yield the chain-length distribution, from which the values of M_n and M_z were calculated. The mean residue molecular weights used to calculate molar residue ellipticities in CD spectra were determined from the chemical composition measured by NMR spectroscopy. The polypeptides are treated as repeats of a unit consisting of x residues Ala:1 residue Orn/Lys, where x is the experimental ratio of Ala/Orn(Lys). For example, the unit in polypeptide 55-04 consists of $1.84 + 1 = 2.84$ residues with molar mass $195 \times 1 + 1.84 \times 71 = 325.6$. The mean residue molecular weight is then 325.6 g/2.84 mol = 114.6 g/mol. Mean residue molecular weights of the other polypeptides were determined similarly and are listed in Table 1.

Circular Dichroism (CD) Measurements. CD spectra were recorded on an Aviv DS60 spectropolarimeter equipped with a thermoelectric temperature control. The wavelength was calibrated with (+)-10-camphorsulfonic acid according to the method of Chen and Yang.[29] About 5 mg of each polypeptide sample was accurately weighed and added in D_2O to make a 5 mg/mL first stock solution. Samples for measurements were prepared by diluting the first stock solutions with 100 mM KF, 10 mM phosphate buffer, pH 7, to a final polypeptide concentration of 50 μg/mL and a buffer concentration of 10 mM KF, 1 mM phosphate buffer (containing 1 M NaCl/20 mM Na/PO_4 when needed) using a 1-mm path length cell. Three scans were averaged with a 0.5-nm step in each case. Three to six measurements were performed on each polypeptide sample.

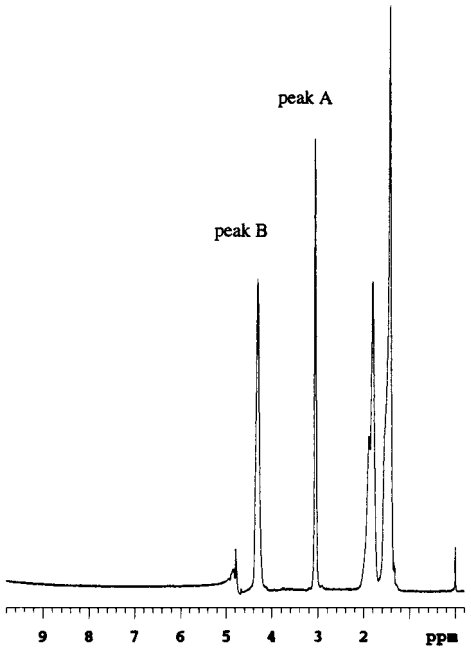

Figure 6. The ^{1}H NMR of polypeptide 55-13. The spectrum shows no peak in the range of 6−8 ppm, indicating completion of HBr cleavage of side-chain protection groups of poly(Ala-Orn) and poly-(Ala-Lys). Peaks A and B are used to calculate the residue ratio in polypeptides. Peak A is assigned to be the α-CH of Ala and the peak B the δ/ϵ-CH_2 of Orn/Lys.

Helix−Coil Transition and the Determination of Nucleation Constant. The standard Zimm−Bragg model of the helix−coil transition was used to calculate circular dichroism values for the copolypeptides. According to the model, each residue can adopt one of two states which correspond to helix or coil conformations. Thus, to define a state of a polypeptide one needs to specify the states of each residue. The statistical weight of a polypeptide state with k helical regions at specific locations would be written as

$$\sigma^k \prod_i s_i \tag{1}$$

where the product is taken over all H-bonded groups in the helical state, σ is the helix nucleation constant, and s_i are the different equilibrium constants for helix propagation.[7a,12] The values of s_i depend on the type of amino acid residues, and thus, only two different values of s_i in our polypeptide system are required for a copolypeptide in a particular solution condition, provided neighbor interactions are absent.

A chain of n residues assumes a large number of possible states according to this model, and special algorithms are required for efficient calculation of different chain characteristics for sufficiently large n values. We used an iterative algorithm developed by Vedenov et al.[12] and the procedure described below to calculate the average fraction of residues in helix state, f, and the average number of helix regions, m, for a heterogeneous polypeptide of particular sequence.

CD values were computed from the statistical mechanical analysis using the set of helix−coil transition parameters, σ, s_{Ala}, s_{Orn}, and s_{Lys}. The CD signal can be calculated from the mean helix length according to the formula described by Yang et al.,[14a] a helix of n residues has a CD signal at 222 nm given by

$$[\theta]_{222}(n) = -41000\,\frac{n-x}{n} \qquad (2)$$

where x is a cutoff constant representing the effect of nonhelical amide and carbonyl groups at the ends of peptide chains. The value $x = 2.5$ was suggested by Yang et al.;[14a] theory suggests a slightly larger number, $x = 4.0$.[14b] The limiting value of $[\theta]_{222} = -41\,000$ is that for a long helix containing only Ala residues, which was measured using a synthetic poly(Ala) linked to PEG for solubility (U. Skarpidis, unpublished results).

To model the copolypeptides used in the experiments, we generated 1000 random sequences of two kinds of residues with appropriate composition. Each sample population was taken to have a Gaussian distribution of lengths with known mean value and variance. This analysis of the helix−coil transition in long chains of two types of residues assumes that the corresponding sequences are random. To simulate the properties of polypeptides with specific average Ala/Orn or Ala/Lys composition, we generated a large set of random sequences with appropriate composition and calculated the average properties of these. The value(s) of the model parameters, σ, s_{Ala}, and s_{Orn} or s_{Lys} which give the best fit to the measured values, that is, which minimize the residual between theoretical and experimental CD values, were then identified. This procedure is quite similar to the one used by Scheraga's group in evaluating helix−coil parameters for the natural amino acids in the HBG or HPG copolypeptides.[19]

Helix−coil transition parameters were fit to experimental data as follows. The CD values of each of the series of Ala-Orn and Ala-Lys copolymers represent the set of observables. For a series, given values of s_{Ala} and s_{Orn} or s_{Ala} and s_{Lys}, the mean and standard deviation of the chain-length distribution, and the composition of each copolymer are specified. A trial value of σ is assumed. A population of chains conforming to the experimental composition and chain length is generated, and the helix content, number of helical regions and mean length of helix regions in each simulated chain are computed recursively.[12] This is essentially the Zimm−Bragg next-neighbor interaction model[7a] adapted to heteropolymers allowing calculation of the helix and coil probabilities of residue i from those of the preceding sequence of residues in the chain. The CD is then calculated using eq 2 above, for several values of x. The residual between the sum of squares of the calculated and measured CD values is computed and stored. Values of this residual are tabulated for different trial values of σ and/or x, and the process is continued until the residuals reach a minimum.

Acknowledgment. Drs. Yuejun Xiang and Hung-Jang Gi contributed helpful advice on the synthesis of NCA's. We are grateful to Mr. Andrew H. Wood of the Michigan Molecular Institute for determining the molecular weight distributions of the polypeptides of this study. This work was supported by grant GM 40746 from the NIH and a grant to N.R.K. from the Human Frontiers of Science Program.

JA982319D