Thermal Unfolding in a GCN4-like Leucine Zipper: $^{13}C^{\alpha}$ NMR Chemical Shifts and Local Unfolding Curves

Marilyn Emerson Holtzer, Eva G. Lovett, D. André d'Avignon, and Alfred Holtzer
Department of Chemistry, Washington University, St. Louis, Missouri 63130 USA

ABSTRACT $^{13}C^{\alpha}$ chemical shifts and site-specific unfolding curves are reported for 12 sites on a 33-residue, GCN4-like leucine zipper peptide (GCN4-lzK), ranging over most of the chain and sampling most heptad positions. Data were derived from NMR spectra of nine synthetic, isosequential peptides bearing 99% $^{13}C^{\alpha}$ at sites selected to avoid spectral overlap in each peptide. At each site, separate resonances appear for unfolded and folded forms, and most sites show resonances for two folded forms near room temperature. The observed chemical shifts suggest that 1) urea-unfolded GCN4-lzK chains are randomly coiled; 2) thermally unfolded chains include significant transient structure, except at the ends; 3) the coiled-coil structure in the folded chains is atypical near the C-terminus; 4) only those interior sites surrounded by canonical interchain salt bridges fail to show two folded forms. Local unfolding curves, obtained from integrated resonance intensities, show that 1) sites differ in structure content and in melting temperature, so the equilibrium population must comprise more than two molecular conformations; 2) there is significant end-fraying, even at the lowest temperatures, but thermal unfolding is not a progressive unwinding from the ends; 3) residues 9–16 are in the lowest melting region; 4) heptad position does not dictate stability; 5) significant unfolding occurs below room temperature, so the shallow, linear decline in backbone CD seen there has conformational significance. It seems that only a relatively complex array of conformational states could underlie these findings.

INTRODUCTION

The GCN4 leucine zipper (GCN4-lz) is a 33-residue peptide whose sequence displays the pseudo-repeating heptad characteristic of coiled coils (O'Shea et al., 1989). In benign aqueous media, it exists as a parallel, two-stranded coiled coil (O'Shea et al., 1989, 1991). GCN4-lz unfolding equilibria are of great interest biochemically, because of its biological role. The simplicity of the structure also recommends it as a model for exploring the physical questions of folding.

In a recent paper, physical studies are reported on a close analog of GCN4-lz, having conservative substitutions at four sites: R1K, H18K, R25K, and R33K (Lovett et al., 1996). This pseudo-wild-type analog (GCN4-lzK) also forms a two-chain coiled coil, the thermal stability of which is slightly less than that of parent GCN4-lz, as determined by CD. Site-specific unfolding equilibria were examined via NMR of a series of synthetic isosequential peptides, each bearing 99% $^{13}C^{\alpha}$ substitution at a few selected sites. That study provides new information on the population of conformational states in the unfolding equilibria, because folded and unfolded states at the same site are in slow exchange on the NMR time scale, yielding separate resonances. Moreover, many sites show two folded forms in slow exchange near room temperature. Not all sites show the same percentage of unfolded form, also indicating some complexity in the conformational population.

This approach thus allows a great deal to be learned from one-dimensional NMR on this system. Accordingly, we report here further thermal unfolding studies on the nine variously labeled isosequential GCN4-lzK peptides shown in Fig. 1. In the course of our studies, a great many $^{13}C^{\alpha}$ chemical shifts were observed. These values in themselves can be informative. Although the theory for chemical shifts remains extremely difficult, empirical methods have provided limits for chemical shift values for the various amino acids in random (Spera and Bax, 1991; Wishart et al., 1991; Wishart and Sykes, 1994) versus α-helical (Spera and Bax, 1991; Wishart et al., 1991) conformation. Below, we present chemical shift values, including the effects of temperature, for folded and unfolded forms at several sites on GCN4-lzK, to learn more about the nature of its local conformational states and to aid in the ongoing effort to provide a sound interpretation of peptide chemical shifts.

A physical understanding of folding/unfolding equilibria requires insight into the population of conformational states that is deeper than chemical shift values alone can supply. For larger coiled coils, such as tropomyosin, there is general agreement that this population does not comprise only native and more or less randomly coiled species. However, opinions differ on the number and type of partially folded states (Privalov, 1982; Holtzer et al., 1990). For smaller coiled coils, such as GCN4-lz, opinions also differ on these questions. Some employ an all-or-none or a simple two-state model, for which there is some evidence (Kenar et al., 1995; Sosnick et al., 1996). However, H/D exchange experiments on GCN4-lz (Goodman and Kim, 1991) indicate wide variations in local dynamics, and data on the stability

Received for publication 3 February 1997 and in final form 9 May 1997.
Address reprint requests to Dr. Alfred Holtzer, Department of Chemistry, Washington University, Campus Box 1134, One Brookings Drive, St. Louis, MO 63130-4899. Tel.: 314-935-6572; Fax: 314-935-4481; E-mail: holtzer@wuchem.wustl.edu.

0006-3495/97/08/1031/11 $2.00

g a b c d e f g a b c d e f g a b c d e f g a b c d e f g a b c d

Ac-K M K Q L E D K [V] E E L L S K N Y K [L] E N E V [A] K L K K L V [G] E K-Am
Ac-K M K Q L E D K [V] E E L L S K N Y K [L] E N E V A K L K K L V [G] E K-Am
Ac-K M K Q [L] E D K V E E L L S K N Y K L E N E [V] A K L K K L V [G] E K-Am
Ac-K M K Q L E D K V E E [L] L S K N Y K L E N E V A K L K K L [V] [G] E K-Am
Ac-K M K Q L E D K V E E L [L] S K N Y K L E N E V A K L K K L V [G] E K-Am
Ac-K M K Q L E D K V E E L L S K N [Y] K L E N E V A K L K K L V [G] E K-Am
Ac-K M K Q L E D K V E E L L S K N Y K L E N E V [A] K L K K L V [G] E K-Am
Ac-K M K Q L E D K V E E L L S K N Y K L E N E V A K [L] K K L V [G] E K-Am
Ac-K M K Q L E D K V E E L L S K N Y K L E N E V A K L K K [L] V [G] E K-Am

5 10 15 20 25 30

FIGURE 1 Catalog of nine synthesized, isosequential, $^{13}C^{\alpha}$-enriched GCN4-lzK peptides. Boxed residues mark sites with 99% $^{13}C^{\alpha}$. All peptides are acetylated at the N-terminal and amidated at the C-terminal. Heptad designations are given above in the standard *abcdefg* notation, in which *a* and *d* designate interior hydrophobes (McLachlan and Stewart, 1975). The residue number is given at the bottom.

of its subsequences (Lumb et al., 1994) suggest that the molecule contains conformationally inequivalent domains. Moreover, evidence from $^{13}C^{\alpha}$-NMR on GCN4-lzK (Lovett et al., 1996) points to a rather rich array of local conformational states.

The question of the conformational population is important in its own right. However, it has further implications. The model one uses can become a self-fulfilling prophecy, because it is often used itself for interpreting data. For example, the backbone CD (at 222 nm) of coiled coils changes almost linearly below room temperature, followed by a more precipitous, cooperative change at higher temperatures. Interpretion of the linear region is model-dependent and, in turn, affects the helix content deduced from the data (Holtzer and Holtzer, 1992). Believers in a two-state molecular population deny conformational significance to the linear region, attributing it to the temperature dependence of the chiro-optical properties of the native coiled coil (Hvidt et al., 1985; Lehrer and Stafford, 1991). For them, it is baseline. Others give it conformational significance, ascribing it to end fraying (Holtzer et al., 1983, 1990; Holtzer and Holtzer, 1992). The resulting "data," as helix content versus temperature, necessarily differ.

To answer these questions requires experimental methods that reveal local conformations. One such method investigates the kinetics of H/D isotopic exchange at amide protons. To our knowledge, there is only one such study extant (Goodman and Kim, 1991). Its important conclusion is that protection against exchange is highly dependent upon position within the peptide chain. Even neighboring residues can display very different protection factors. This suggests that a rather rich array of intermediate molecular states may exist. However, as discussed below, the use of cross-linked chains in the exchange experiments makes caution necessary in applying the findings to the more ordinary, non-cross-linked case. Moreover, this exchange method, although highly useful in protein studies, is susceptible to more than one interpretation (Miller and Dill, 1995). It is not a direct measure of local structure content.

For revealing local information, NMR is the method of choice. Yet, although there have been several NMR investigations of leucine zippers, to our knowledge, none has hitherto examined the temperature dependence so as to reveal local thermal unfolding curves. $^{13}C^{\alpha}$-NMR has many advantages as a local probe. The $^{13}C^{\alpha}$ is a backbone atom, and its resonance is conformation-dependent (Spera and Bax, 1991; Wishart et al., 1991; Wishart and Sykes, 1994). Moreover, substitution of ^{13}C for ^{12}C is virtually conformationally neutral, unlike other spectroscopic probes such as fluorescent chromophores, spin labels, or foreign nuclei (e.g., fluorine). Indeed, $^{13}C^{\alpha}$-NMR has been used to obtain site-specific unfolding curves for small single-stranded α-helical peptides from chemical shifts, folded and unfolded forms being in fast exchange in that system (Shalongo et al., 1994). In GCN4-lzK, however, folded and unfolded forms are in slow exchange, leading to well-separated $^{13}C^{\alpha}$ resonances for locally folded and locally unfolded forms (Lovett et al., 1996). It thus becomes possible to obtain NMR-derived local thermal unfolding curves from the resonance intensities in the one-dimensional spectra.

The use of isosequential peptides, each bearing high $^{13}C^{\alpha}$ substitution at sites chosen to avoid spectral overlap, has several strengths and therefore complements more complex NMR techniques. First, the high substitution and lack of overlap facilitate assignment of the spectra and shorten data acquisition times, avoiding sample damage at the higher temperatures. Reliable data are accessible over the span 8.5–73°C. Second, interpretation is rapid, direct, and unambiguous, placing no reliance on model-dependent interpretive schemes. Third, minimal demands are made on the nature of the aqueous solvent, which can be at neutral pH and of natural isotopic abundance. Fourth, the use of relatively dilute solutions (i.e., peptide-chain formalities of a few hundred micromolarity) is feasible. For H/D exchange, on the other hand, concentrations 20 times higher have been used (Goodman and Kim, 1991).

Below we report the results of such NMR-intensity measurements on the nine isotopically labeled peptides shown in Fig. 1. These isosequential peptides allow study of the unfolding at 12 amino acid sites, covering most of the molecular length and sampling a variety of coiled-coil heptad positions. Because the thermal unfolding of GCN4-lzK is reversible, provided exposure to temperatures above 65°C is not protracted, these unfolding curves describe equilibrium states (Lovett et al., 1996).

As will be seen, our results do not support any of the extant simple models, but display some features of each. It seems that only a relatively complex array of intermediate states can explain these findings, which are characterized by strong local differences in the degree of unfolding and by significant conformational change, even in the region of temperature wherein the backbone CD changes are linear.

MATERIALS AND METHODS

Peptide synthesis, purification, and analysis

Nine labeled isosequential peptides (Fig. 1) and one at natural abundance were synthesized as described earlier (Holtzer et al., 1995; Lovett et al., 1996) via the solid-phase method, using fluoren-9-ylmethoxy carbonyl (Fmoc) chemistry and a RaMPS multiple-peptide synthesis system (DuPont Co., Wilmington, DE). Procedures used to prepare and purify the $^{13}C^{\alpha}$-labeled Fmoc-amino acids have also been described previously (Lovett et al., 1996). These derivatives were characterized by ^{1}H NMR and by reversed-phase, high-performance liquid chromatography (HPLC), using a Beckman Ultrasphere ODS C-18 column (4.6 × 250 mm, 5 μ, 80-Å pore), and were coupled to the growing peptide chain by using diisopropylcarbodiimide and 1-hydroxybenzotriazole.

Each of the purified GCN4-lzK peptides was >95% pure by reversed-phase HPLC on a Vydac C18 column (4.6 × 250 mm, 5 μ, 300-Å pore) with acetonitrile gradients of 27–54% at 0.9–1.5%/min. All isosequential peptides have the same HPLC retention time and show no sign of heterogeneity when mixed (1:1) with pure, natural-abundance GCN4-lzK.

Electrospray mass spectrometry gave a molar mass for each peptide that was within 1 Da of its expected value. We find a sample standard deviation of 0.41 (N = 11) for all peptides (set includes duplicate measurement for one of the labeled peptides) and 0.34 (N = 10) if the natural abundance peptide is omitted. Further demonstration that the peptides are pure and isosequential is manifested by the invariance from peptide to peptide of the $^{13}C^{\alpha}$ chemical shifts for certain residues and, at a given concentration, of both local (NMR) and global (CD) thermal unfolding profiles (Lovett et al., 1996).

For the vast majority of measurements, peptides were dissolved in benign, near-neutral, saline buffer (100 mM NaCl, 50 mM sodium phosphate, pH 7.4); a few measurements substituted 150 mM NaCl for 100 mM, and showed no significant differences in either chemical shifts or fraction folded. These solutions contain no extraneous additives, such as chemical shift reagents or D_2O. However, a few measurements in D_2O showed only small variations in chemical shifts and virtually none in fraction folded, up to 16% D_2O. Two of the peptides were also studied in a solvent containing 6 M urea in addition to the saline buffer. Peptide concentrations were determined from the absorbance of Y17 at 275 nm; the extinction coefficients for GCN4-lzK are 1.40 cm^{-1} M^{-1} in benign buffer, and 1.54 cm^{-1} M^{-1} in denaturing media, with M in formality of peptide chains (Lovett et al., 1996).

$^{13}C^{\alpha}$ NMR

All ^{13}C NMR measurements were made with a Varian Unity Plus-500 spectrometer, operating at 125.703 MHz and equipped with a 10-mm probe and an Oxford Instruments temperature controller. Data collection, temperature measurement, calibration procedures, and external lock and referencing were as described earlier (Lovett et al., 1996). Recently Wishart and Sykes (1994) recommended 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS) as a reference standard. We find that the reference line of DSS depends somewhat on temperature, over the wide range employed in our study. Therefore, all primary chemical shifts used herein are referenced to external DSS at the appropriate temperature.

Primary chemical shift values (δ, ppm) were obtained in two ways: 1) as usual, from the maxima in the Fourier-transformed frequency-domain spectra; and 2) from a Bayesian analysis (Bretthorst, 1990) of the free induction decays. The two methods give indistinguishable values. Secondary chemical shifts ($\Delta\delta$) were obtained by the subtraction of extant random coil chemical shifts from each observed chemical shift, i.e., $\Delta\delta = \delta_{obs} - \delta_r$, where δ_r is 45.0, 52.2, 55.0, 58.0, and 62.2 ppm for G, A, L, Y, and V residues, respectively (table 9 in Wishart and Sykes, 1994). The chemical shifts given by Wishart and Sykes (1994) and by Spera and Bax (1991) for these residues agree quantitatively after the latter are re-referenced to DSS.

Integrated resonance intensities (areas) were determined both with standard software and by a Bayesian analysis (Bretthorst, 1990) of the free induction decays, the latter utilizing software provided by Dr. G. Larry Bretthorst. The results from the two methods do not differ significantly; nor does application of the saturation correction produce any appreciable change.

For our purposes, no distinction is made between the two slowly interconverting folded forms seen at several (but not all) of the sites examined (Lovett et al., 1996). The reported local value of the fraction folded therefore signifies the total number of peptide chains in which the given site is in some folded form, divided by the total number of chains.

Backbone and tyrosine CD

In many cases, it is useful to compare the local folded fraction with the fraction provided by CD at 222 nm. The latter is, of course, an average over all of the peptide bonds in the sample. To determine this backbone CD fraction, we employed the relation

$$f_{CD} = \{[\Theta] - [\Theta_U]\}/\{[\Theta_F] - [\Theta_U]\}$$

wherein $[\Theta]$ is the observed mean residue ellipticity at 222 nm, $[\Theta_F]$ is the corresponding value for a fully folded molecule, and $[\Theta_U]$ is the value for the fully unfolded molecule.

To account for the finite helix length (Chen et al., 1974; Holtzer et al., 1983), we determined the value of $[\Theta_F]$ from $[\Theta_F] = [\Theta_h^{\infty}][1 - (2.55/n)]$, where $[\Theta_h^{\infty}]$ is the value for an infinite α-helix, for which we used $-386°$ cm^2 $mmol^{-1}$; and n is the number of peptide bonds in the chain (34 in the case of the doubly capped GCN4-lzK). The result is $[\Theta_F] = -357°$ cm^2 $mmol^{-1}$. For $[\Theta_U]$, we used $-57°$ cm^2 $mmol^{-1}$. The latter is somewhat more negative than the value usually employed for randomly coiled peptides. However, this may reflect the somewhat nonrandom character of unfolded single chains of GCN4-lzK, as suggested by the observed chemical shift values (see below). This more negative $[\Theta_U]$ gives more realistic values of f_{CD} at high temperatures. Because we do not have NMR data for all sites, comparisons with backbone CD are necessarily heuristic only. We therefore make no attempt to evaluate the temperature dependence, if any, of $[\Theta_F]$ or $[\Theta_U]$. In any case, the observation (see below) that the linear, low-temperature region of CD at 222 nm is not simply baseline makes evaluation of these temperature dependences highly uncertain.

It is also useful to make comparisons between the NMR-observed fraction folded at residue Y17(b) and the CD in the tyrosine region, i.e., near 280 nm (Holtzer et al., 1996). Because the conformation-CD link in this spectral region is far less well understood than for the backbone region, we simply compare the NMR-derived fraction here with the ratio of the integrated tyrosine CD band intensity at temperature T to that at the lowest temperature, (i.e., 2°C).

RESULTS

$^{13}C^{\alpha}$ NMR chemical shifts

Fig. 1 summarizes the nine isosequential, labeled peptides that have been synthesized and studied so far. Overall, 12 chain sites become NMR-observable when these peptides are used. Some sites are labeled in more than one peptide. All heptad positions except c and f are represented. A few studies were also made of the natural-abundance peptide.

Because the unfolding equilibria of GCN4-lzK are concentration dependent (Lovett et al., 1996), a range of concentrations was employed. For the isotope-enriched samples, this range was 30–2640 μM, although not every peptide was examined over the entire range. No concentration dependence was seen for any chemical shift in any peptide. Some studies were made with natural-abundance samples at 8500 μM, and wherever corresponding reso-

nances could be identified in a natural abundance sample and a labeled peptide, the chemical shifts agreed. Thus the observed chemical shifts are doubtless valid up to a peptide concentration of ~8.5 mM and possibly higher. For this reason, no distinction need be made between chemical shifts determined at different concentrations or between those measured on labeled versus natural-abundance peptides.

The temperature dependence of these chemical shifts is, on the other hand, often observable, although rather small, and differs from site to site. A sense of the results may be gleaned from Fig. 2. Although it is not apparent from the data summary in Fig. 2, most of the chemical shift variation with temperature for folded species at a given site actually stems from the transition that occurs at most sites from the low- to moderate-temperature folded form rather than from inherent temperature dependence for a given folded form. Only positions V23(*a*), L19(*d*), and L26(*d*) do not show two folded forms. We do not yet know whether the nonappearance of two forms at some sites indicates that there is only one form there or whether the two forms have chemical shifts too similar, or too rapidly interconverting, to resolve. Further examination of Fig. 2 is deferred to the Discussion section, where its implications are examined, and an hypothesis is framed proposing a structural correlative for the appearance or nonappearance of the two folded forms in these spectra.

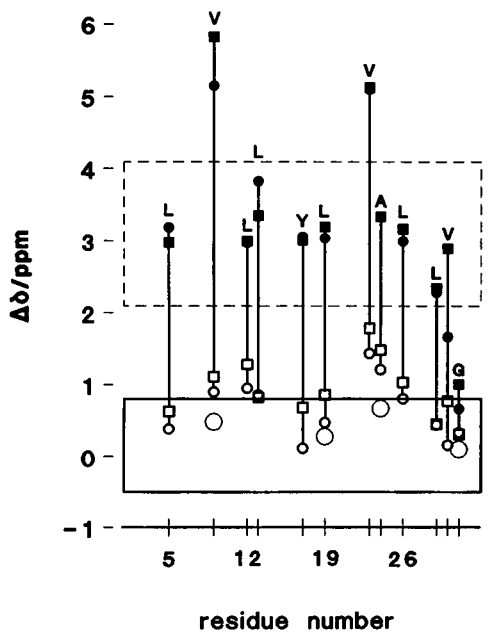

FIGURE 2 $^{13}C^{\alpha}$ secondary chemical shifts in GCN4-lzK versus residue number. Urea-unfolded forms: ○, 25°C. Thermally unfolded forms: ○, 73°C; □, 8.5°C (extrapolated for Y17(*b*), which shows no unfolded form at 8.5°C). Folded forms: ■, 8.5°C; ●, 58°C. The dashed-line box marks the established range for α-helical $^{13}C^{\alpha}$ (Spera and Bax, 1991). The solid-line box marks the established range for random $^{13}C^{\alpha}$ (Wishart and Sykes, 1994).

Local unfolding curves

Overall view

An overview of our data is given in Fig. 3 as fraction folded versus residue number for solutions, all at 300 μM as single chains, at several temperatures covering the entire accessible range. Data for G31(*b*) are not shown, not only because the strong overlap of resonances for folded and unfolded forms at this position make the area measurements less certain (Lovett et al., 1996), but also because the folded form's chemical shift is so atypical of α-helix (indeed is closer to random glycine's) as to raise a serious question as to what "folded" means at G31(*b*).

At the lowest temperature (8.5°C), virtually all of the molecules are two-stranded (Lovett et al., 1996). Nevertheless, Fig. 3 shows many idiosyncratic ups and downs in local structure content throughout the molecule. The overall "frown" shape indicates a lower fraction folded at the chain ends. However, there are also interior variations. For example, L19(*d*) is an interior low spot and L12(*d*) is only 80% folded, almost as low as the ends. Only Y17(*b*) appears to be 100% folded, even at this low temperature.

As temperature rises toward room temperature, a remarkable change occurs in the structure probability profile, as

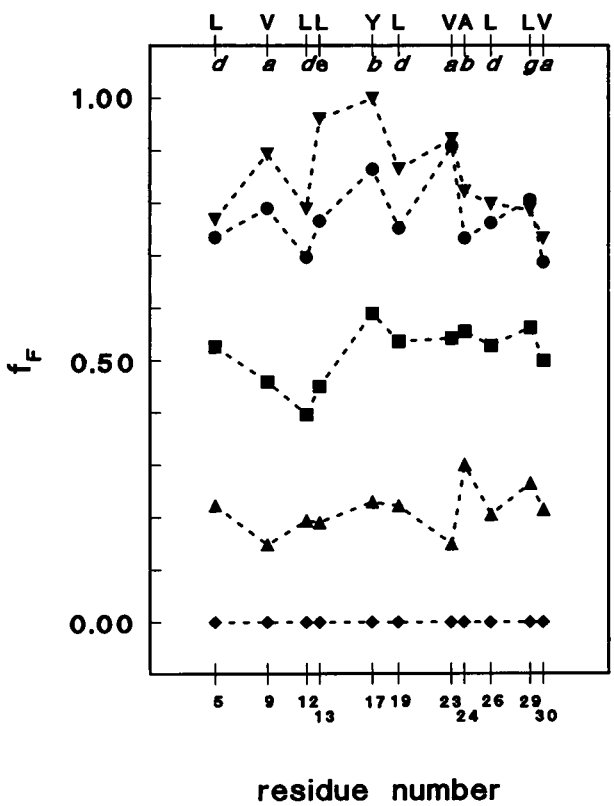

FIGURE 3 Local, NMR-derived fraction folded versus residue number for GCN4-lzK at 300 μM (in chains) and various temperatures. Amino acid type and heptad letter designations are at the top. Temperatures are (in °C): ▼, 8.5; ●, 27.9; ■, 46.9; ▲, 58.2; ◆, 73.0. Dashed lines merely connect the available data points; they do not necessarily represent behavior between points.

seen in Fig. 3. The structure content at the ends falls only slightly, whereas that at interior sites drops more significantly. For example, position Y17(*b*), completely folded at low temperature, drops to 85% when just above room temperature (27.9°C).

A further increase in temperature to 40–50°C at this peptide concentration leads to appreciable dissociation into unfolded chains, as is known from the concentration dependence of the backbone CD (Lovett et al., 1996). The CD-determined midpoint of the transition at 300 μM is in that temperature range. This is reflected in the marked drop in structure seen in Fig. 3 at all positions. As Fig. 3 also shows, at 46.9°C, the structure content is almost constant from Y17(*b*) to the C-terminus, but positions 12 and 13 have become the least structured of those measured. Further heating to 58.2°C not only markedly increases the population of single chains, as shown by the further drop in structure at all sites, but also leads to a different profile. Here the fraction folded is almost constant over positions 5–23, but A24(*b*), the only chain alanine, retains more structure than any other site. Finally, at 73.0°C, the only resonances seen are those for unfolded forms.

Positions L5(d), L26(d), and L29(g)

Fig. 4 shows thermal unfolding data for sites L5(*d*) and L29(*g*) for peptides at 300 μM. These two nearly terminal sites unfold rather similarly. The fraction folded, even at the lowest temperature, is only 80%. However, there is little change in structure content upon warming to room temperature and only a small drop from there to nearly 40°C. Above the latter temperature, a sharp drop is seen, reaching zero above 70°C.

Comparison with the folded fraction from CD is aided by the heavy solid curve in Fig. 4, which represents backbone CD data at the same concentration (Lovett et al., 1996). Thermal unfolding data for L26(*d*) (not shown) are distinguishable from, but rather like, those for L5(*d*) and L29(*g*), showing no new features.

These local unfolding curves show strong concentration dependence, as expected for a dissociating transition. This is shown in Fig. 5, which redisplays the data of Fig. 4 for L5(*d*) at 300 μM, along with data at 2640 μM. At the higher concentration, the folded fraction is essentially constant at nearly 80% up to 50°C, where the fraction at the lower concentration has fallen to ~50%. The corresponding curves from CD are given for comparison, showing once again that local differences from the overall molecular folded fraction characterize these transitions.

Positions L12(d), L13(e), L19(d), and A24(b)

Positions L12(*d*) and A24(*b*) show the greatest difference in local melting temperature determined so far. As Fig. 6 shows, the difference in temperature for 50% melting at 300 μM is ~8°C for L12(*d*) versus A24(*b*), the latter being more stable. These two sites also differ from those consid-

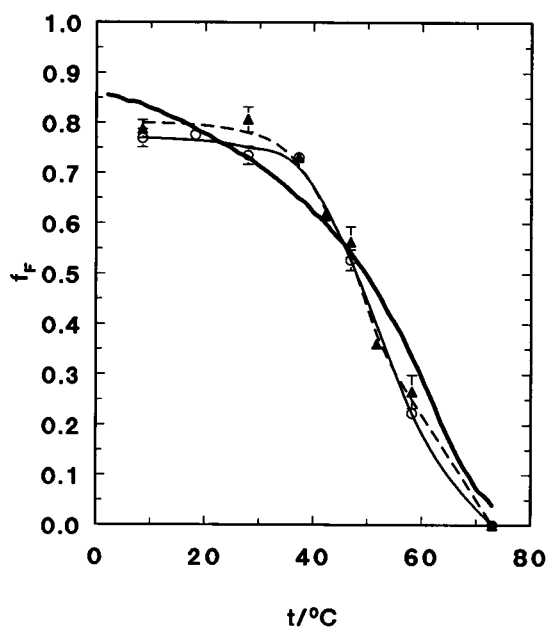

FIGURE 4 NMR-derived thermal unfolding curves for positions L5(*d*) (*open circles and thin solid curve*) and L29(*g*) (*filled triangles and thin dashed curve*) of GCN4-lzK at 300 μM (in chains). Thin curves merely guide the eye through the data. Error bars are standard error of the mean, and are given only where enough measurements exist to estimate errors. The thick solid line is for a folded fraction over the entire molecule at 300 μM, as determined from CD at 222 nm.

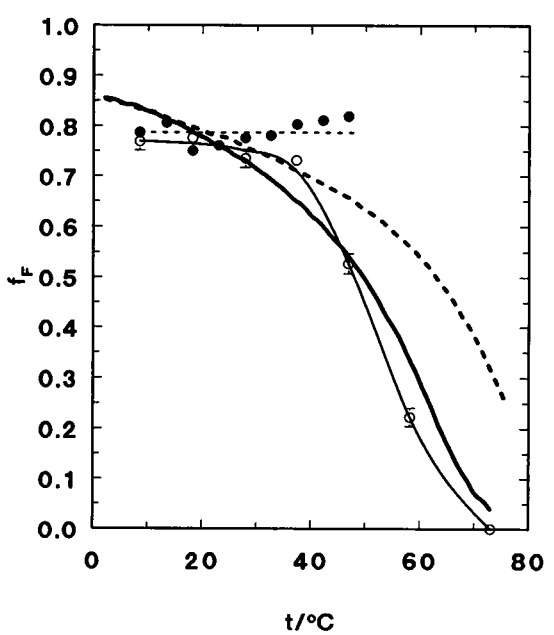

FIGURE 5 NMR-derived thermal unfolding curves for position L5(*d*) of GCN4-lzK at concentrations (in chains) of 300 μM (*open circles and thin solid curve*) and 2640 μM (*filled circles and thin dashed curve*). Thin curves merely guide the eye through the data. Error bars, as in Fig. 4. Thick lines are for a folded fraction for the entire molecule, as determined from CD at 222 nm, 300 μM (*solid*), and 2640 μM (*dashed*).

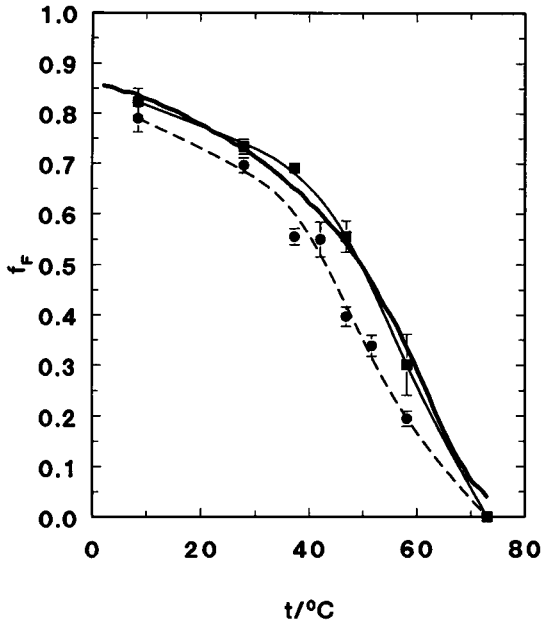

FIGURE 6 NMR-derived thermal unfolding curves for positions L12(*d*) (*filled circles and thin dashed curve*) and A24(*b*) (*filled squares and thin solid curve*) of GCN4-lzK at 300 μM. Thin curves merely guide the eye through the data. Error bars and thick (CD) curve as in Fig. 4.

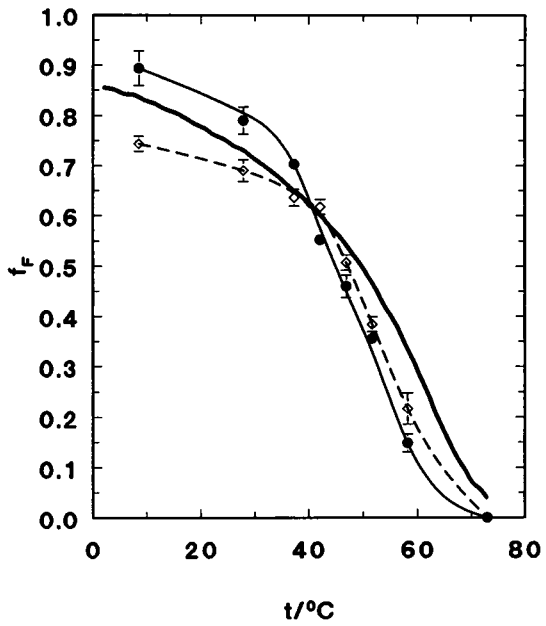

FIGURE 7 NMR-derived thermal unfolding curves for positions V9(*a*) (*filled circles and thin solid curve*) and V30(*a*) (*open diamonds and thin dashed curve*) of GCN4-lzK at 300 μM. Thin curves merely guide the eye through the data. Error bars and thick (CD) curve are as in Fig. 4.

ered above in that there is a marked change in structure content upon warming from the lowest temperatures to near room temperature. The decline in this range for L12(*d*) and A24(*b*) is comparable to that seen in CD at 222 nm, which is an average over the whole molecule. The unfolding curves for L13(*e*) and L19(*d*) (not shown) are rather similar to that for L12(*d*) and add no new features.

Positions V9(a), V23(a), and V30(a)

All three valines in the GCN4-lzK chain are located in interior, canonically hydrophobic *a* heptad positions. Their local thermal unfolding curves appear in Figs. 7 and 8. All three of these positions share with some others a modest but definite decline in structure fraction with temperature in the region below room temperature, followed by a steeper drop at higher values. Once again, the low temperature decline is commensurate with that seen in backbone CD, except, perhaps, for V30(*a*), where it is a bit shallower. Position V30(*a*) also differs from the other two in having a rather small structure content (75%), even at the lowest accessible temperature. At much higher temperatures, however, its structure fraction differs little from the others, indicating that reduced low-temperature structure content and thermal lability are not invariably coupled.

Fig. 8 displays NMR data for site V23(*a*) at two different concentrations of GCN4-lzK. The unfolding at 300 μM is like that of V9(*a*). The near-linear decline in the low- to room-temperature region persists at 2640 μM to at least 50°C, where the structure fraction is 75%, whereas the lower concentration (300 μM) is already in its more precipitous decline at that temperature, having fallen to 40%.

Position Y17(b)

The data in Fig. 9 for this position, the only tyrosine residue in the molecule, show that it is also the maximally struc-

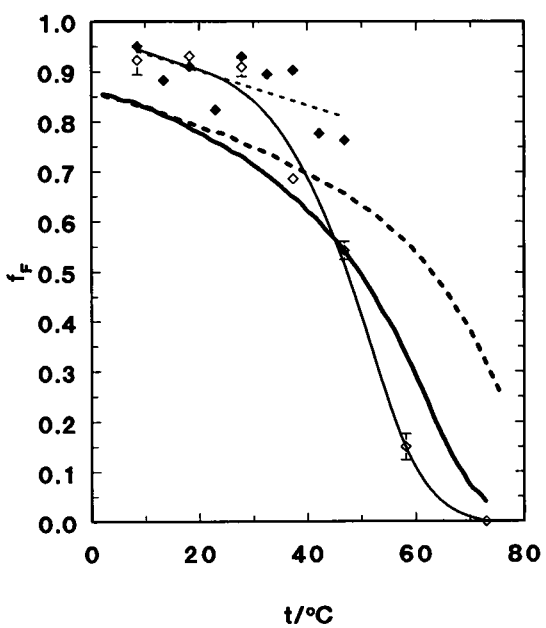

FIGURE 8 NMR-derived unfolding curves for position V23(*a*) of GCN4-lzK at 300 μM (*open diamonds and thin solid curve*) and 2640 μM (*filled diamonds and thin dashed curve*). Thin curves merely guide the eye through the data. Error bars and thick (CD) curves are as in Fig. 5.

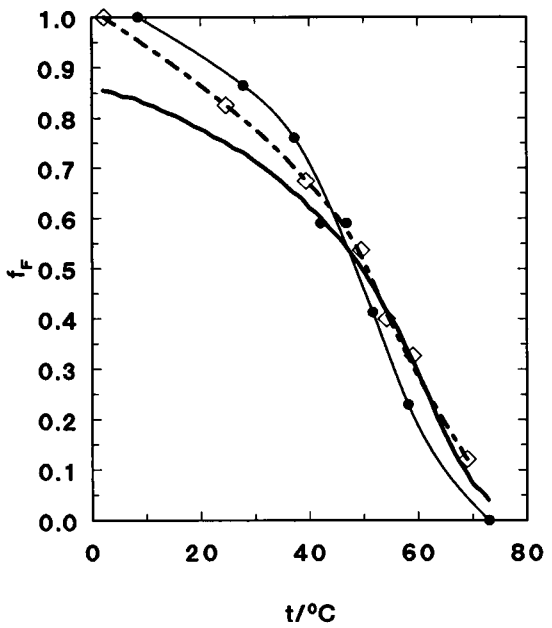

FIGURE 9 NMR-derived unfolding curves for position Y17(*b*) of GCN4-lzK at 300 μM (*filled circles and thin solid curve*). Two CD-derived curves for the same concentration are given for comparison. The thick curve is from backbone CD. Data from tyrosine CD (*open diamonds and thick dot-dashed curve*) were obtained as described in Materials and Methods. All curves merely guide the eye through the data.

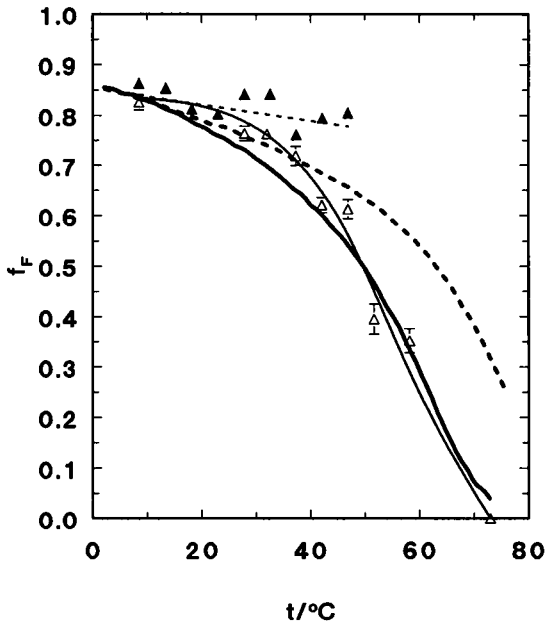

FIGURE 10 NMR-derived unfolding curves for position G31(*b*) of GCN4-lzK at 300 μM (*open triangles and thin solid line*) and 2640 μM (*filled triangles and thin dashed line*). Thick curves are backbone-CD derived curves, as in Fig. 5. All curves merely guide the eye through the data.

tured site, being 100% folded at the lowest temperature. Again, as at several other sites, there is a near-linear decline in structure fraction in the low- to room-temperature range, followed by a steeper drop at higher temperature.

Comparison of the local, NMR-derived values with CD unfolding curves is also made in Fig. 9. The backbone CD value (*thick solid curve*) for the folded fraction at low temperature is 85% at maximum, quite a bit lower than the maximum local value for this site. At higher *T*, these differences are much less. The estimate of structure fraction from tyrosine CD (*thick dot-dashed curve*), on the other hand, follows the NMR unfolding curve rather closely throughout. Because estimating structure from CD in the tyrosine region is in its infancy, the closeness of the values from tyrosine CD and NMR is encouraging.

Position G31(b)

Fig. 10 shows unfolding curves for this position at two vastly different concentrations of GCN4-lzK. Even at the lowest temperature, this site is only 85% folded. Moreover, it is not clear what "folded" means in this context, because the chemical shift for the separately observed resonance of the "folded" form at this site is closer to that of random glycine than to that generally assigned to helical glycine (Lovett et al., 1996). However, the unfolding curves themselves follow a pattern similar to that seen at other sites, in that an initially linear decline gives way to a steeper drop at higher temperatures.

DISCUSSION

Chemical shifts

Chemical shifts doubtless contain important information on local peptide conformations, but the absence of a readily implementable theory precludes quantitative extraction of this information. However, empirical methods can lead to some insights.

Such empirical examination is greatly aided by an enormous body of meticulously collected data on $^{13}C^{\alpha}$ shifts for α-helical and for randomly coiled residues that has been systematically catalogued (Spera and Bax, 1991; Wishart et al., 1991; Wishart and Sykes, 1994). The values for randomly coiled residues were obtained using both urea-unfolded proteins and small unstructured peptides (Wishart et al., 1991). Subtraction of these shifts for randomly coiled residues from those observed in our study, and plotting the difference versus chain position, yields the graph of Fig. 2. The lower (*solid line*) box defines the secondary chemical shift range for randomly coiled residues. It is not symmetrical about zero, because the range in the original data is not symmetrical (Wishart et al., 1991). The upper (*dashed line*) box defines the range expected for α-helical $^{13}C^{\alpha}$ secondary shifts (Spera and Bax, 1991). To summarize the temperature dependence for our data at each site, two data points are plotted for each local conformation, covering the total range observed for that conformation at that site. The substantial range shown for the folded form at some sites is a result not of large dependence of shift on temperature for a given folded form, but of the change from low- to higher-*T* folded forms at those sites.

Several conclusions can be drawn from Fig. 2. The urea-unfolded sites (*large open circles*) show shifts that are all within the expected range for randomly conformed residues. Thermally unfolded sites near the ends of the chain also show values within this expected range. However, many (not all) sites in the center of the chain show secondary shifts that are outside the expected range and in a direction suggesting that they retain some residual structure. Indeed, at a few sites (e.g., V23(a) and A24(b)), the secondary shift for the unfolded form is almost in the range expected for a helical residue! We conclude that although the urea-unfolded protein is probably close to randomly coiled, the thermally unfolded chain has a significant amount of residual local structure, although it is undoubtedly transient.

The secondary shifts for the folded forms are almost all within the range expected for α-helical $^{13}C^{\alpha}$. However, there are a few exceptions. Position G31(b) is remarkably atypical in that its low-temperature folded form has a shift almost in the range expected for unfolded glycine, and the higher-temperature folded form is actually in that range. We ascribe this to the position of G31 near the C-terminus, rather than to the atypicality of glycine itself, because secondary shifts for folded forms of L29(g) and V30(a) are also depressed. This conclusion is confirmed by our preliminary studies of an A24G mutant of GCN4-lzK bearing $^{13}C^{\alpha}$ labels at V9, L19, and G24. In benign saline buffer, the secondary shifts at V9(a) and L19(d) are virtually unchanged from those seen in GCN4-lzK. However, when a glycine is located at position 24, its secondary shift is well in the helix box of Fig. 2, unlike the shift when glycine is at position 31.

This and related questions can be clarified by studies of other mutants. Meanwhile, it does not seem likely that this atypicality of G31 is a result of participation in glycine C-terminal helix-capping, which requires that the sequence surrounding the glycine follow particular rules (Aurora et al., 1994). The sequence around G31 does not conform. However, very recently a new C-capping structure has been proposed, involving H-bonding between the amino side chain of a C-terminal lysine and the carbonyl of a leucine located four positions away (Esposito et al., 1997). In GCN4-lzK, such an interaction could link L29(g) and K33(d), possibly affecting the NMR spectra of their neighbors.

Secondary shifts for folded forms at V9(a) and V23(a) are also atypical: they are well above the expected helical range. At present, we have no explanation for these elevated values, nor is it clear whether the cause lies in the heptad a interior location of these valines or in valine itself. It is possible that V30(a) does not show such anomalous shifts simply because its near-C-terminal position overrides whatever is causing the anomaly.

Thus far we have no basis for assigning the two folded forms, F and F', observed at some sites to particular conformational variants; nor do we understand why some sites display these two folded forms and some do not. However, a hypothesis on the latter point does suggest itself. The positions showing only one folded form—L19(d), V23(a), and L26(d)—have in common the fact that each appears in an interior a or d position within a heptad whose e member participates in a canonical interchain salt bridge in GCN4-lz (O'Shea et al., 1991), and therefore probably in GCN4-lzK as well. The salt bridges in question are from K15(g) on one chain to E20(e) on the other, which brackets L19(d), and from E22(g) on one chain to K27(e) on the other, which brackets V23(a) and L26(d). All of our observations so far are consistent with the hypothesis that such salt-bridge bracketing suppresses the appearance of two folded forms in the $^{13}C^{\alpha}$-NMR spectrum. For example, there is a potential canonical interchain salt bridge spanning the distance from K1(g) to E6(e) in GCN4-lzK; however, the corresponding span in GCN4-lz, R1(g) to E6(e), is not salt-bridged (O'Shea et al., 1991), and indeed, the site L5(d) in GCN4-lzK shows both F and F' forms.

The secondary shifts of GCN4-lzK alone therefore suggest that 1) The urea-unfolded chain is randomly coiled. 2) The thermally unfolded chain includes some transient structure, except at the ends. 3) Near the C-terminus of the folded chains, the coiled-coil structure itself is somewhat atypical, either looser or influenced by end effects such as helix-dipole interactions, lack of H-bond partners, or capping. 4) The hypothesis that canonical, interchain salt bridges surrounding an interior a or d site inhibit the appearance there of two folded forms F and F' explains our observations so far.

Local unfolding curves

Our data on site-specific unfolding at 12 sites in the GCN4-lzK coiled coil provide a more detailed view of the population of conformational states in its thermal unfolding equilibria. In particular, the following five conclusions seem inescapable.

1. Contrary to evidence (for parent GCN4-lz) from calorimetry (Kenar et al., 1995) and other considerations (Sosnick et al., 1996), the molecular population is not two-state. The idiosyncratic local ups and downs shown in Fig. 3 for GCN4-lzK disagree sharply with any such model. This is demonstrated another way in Fig. 11, wherein the local melting temperatures are plotted versus residue number. For an all-or-none model, melting would be the same at every site; for a less extreme two-state case, it would be the same, except for a small number of residues that might be natively unfolded and would not change upon thermal denaturation. Neither result is seen in Figs. 3 or 11. Moreover, it seems doubtful that such sharp differences in the population of states could be explained by the slight differences in sequence between parent GCN4-lz and GCN4-lzK, especially in view of experiments on GCN4-lz that also point to a more complex array of states (Goodman and Kim, 1991; Lumb et al., 1994).

2. The molecule does not unfold simply by progressive fraying in from the ends. Although it seems clear from Fig.

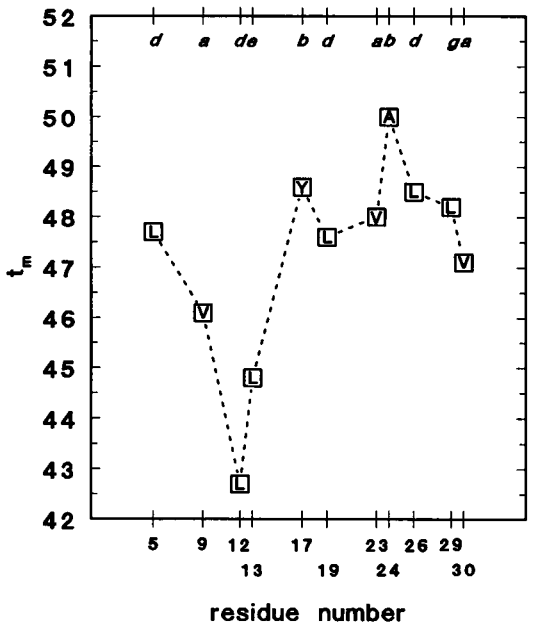

FIGURE 11 NMR-derived local melting temperatures for GCN4-lzK sites versus chain position. The ordinate gives the interpolated temperature at which the local folded fraction is 50%. Heptad letter designations are at the top. Dashed lines merely connect the available data points; they do not necessarily represent behavior between points.

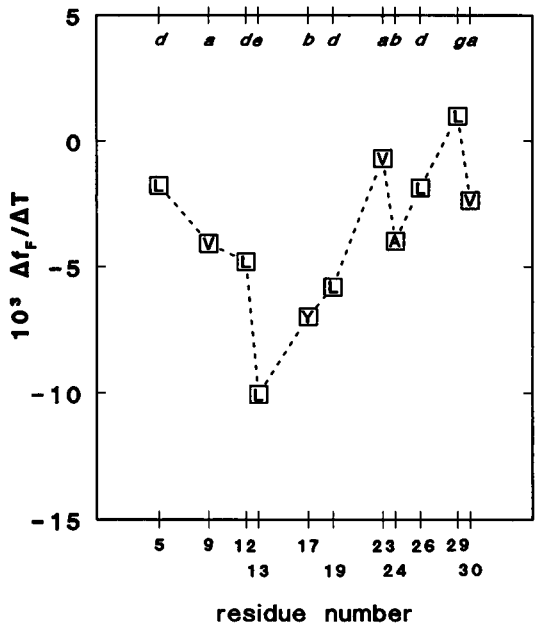

FIGURE 12 NMR-derived temperature coefficient of structural decline in the low- to room-temperature region, plotted as $\Delta f_F/\Delta T$, wherein $\Delta f_F = f_F(27.9°C) - f_F(8.5°C)$ and $\Delta T = 27.9-8.5$ versus the residue number. Heptad letter designations are at the top. Dashed lines merely connect the available data points; they do not necessarily represent behavior between points.

3 that the ends are indeed frayed, even at the lowest temperature, the overall picture conveyed in Figs. 3 and 11 is not a simple progression of unfolding in from the chain termini as temperature is raised, not even in the low- to room-temperature range, where dimers prevail (Fig. 3). Moreover, in several cases, the folded fraction and the melting temperature are higher for some end residues than for some more central ones (Fig. 11).

3. The least stable region of the molecule, which would be expected a priori to be at the C-terminal, in fact is somewhere in the region of residues 9–16. We cannot be more precise at present, because so far we only have data for residues 9, 12, and 13 in that region. Residue 17 is clearly excluded from this labile region, inasmuch as it is the highest residue in maximum structural content and second highest in melting temperature.

4. Heptad position does not dictate local stability. The highest local melting temperature measured so far is at A24(*b*), and the value at Y17(*b*) is also high, although the *b* heptad position has no known stabilizing role. The most obvious explanation is that it has something to do with the high helix propensity of alanine, but this leaves the stability of site Y17(*b*) unexplained. The presence of canonical interchain K15(*g*)–E20(*e*) and E22(*g*)–K27(*e*) salt bridges in this region is another possible cause.

5. These data clearly answer the question posed above concerning the conformational significance of the well-known change of backbone CD in the low- to room-temperature region. The rate of decline of local structured fraction from NMR is displayed in Fig. 12 as the low- to room-temperature regional slope versus residue number. As the figure shows, this slope varies from site to site, being steepest for the structurally rather weak site, L13(*e*). However, this rate does not correlate precisely with local stability, because L12(*d*) melts at an even lower temperature than L13(*e*), and the slope at Y17(*b*), the highest melting site, also declines steeply in the low- to room-temperature region.

These NMR data leave little doubt that conformational changes to less folded states occur in the region of temperature below that characterized by highly cooperative change. These structural changes must contribute to the observed linear change in backbone CD. Consequently, it is no longer possible to interpret that decline simply as the baseline in unfolding studies. As Fig. 9 demonstrates, the same can be said for the initial linear change of CD in the tyrosine spectral region. This is of some interest for coiled coils other than leucine zippers. For example, a study of the thermal unfolding of tropomyosin by CD measurements in the tyrosine region simply treats the similar initial linear change seen there as baseline for the folded form (Ishii et al., 1992). In view of Fig. 9, this procedure is questionable.

Our results, all at near-physiological pH and ionic strength, are not very comparable to those for acidic media. For example, one study of GCN4-lz via two-dimensional proton NMR (Oas et al., 1990) concludes that, at 20°C, 32 of the 33 residues are helical, a result quite different from what we find for GCN4-lzK. However, the proton study was carried out at pH 5, a very different solvent.

On the other hand, studies of H/D exchange in GCN4-lz (Goodman and Kim, 1991) employed near-physiological media, comparable to the solvent used here. A comparison of our lowest temperature (8.5°C) results with those from exchange (obtained at 6°C) is given in Fig. 13. As can be seen, there are regions of agreement and regions of disagreement. Both data sets indicate end fraying, for example. However, there are also regions of substantial disagreement, such as the region of residues 24–26.

In making such comparisons, however, one must recognize not only that the GCN4-lz and GCN4-lzK sequences differ at four sites, but also that the two studies differ significantly in other ways:

1. The H/D work used chains disulfide cross-linked via the cysteines of a supplementary CGG N-terminal sequence. Our peptides are non-cross-linked and have no supplementary residues. The cross-linked peptide used for exchange had a CD-observed melting temperature of 83°C. Thus the cross-linking certainly has a major effect on the conformational populations.

2. The H/D studies employed peptide concentrations of 6 mM; although some of our NMR studies are only at somewhat less than one-half that, the vast majority are at 1/20th that concentration. The resulting differences are doubtless quite significant.

3. The H/D exchange study was carried out in D_2O, ours in aqueous solutions at natural isotopic abundance. However, in our own limited studies using D_2O (not shown), the differences between these media in local unfolding at three sites—L12(*d*), V30(*a*), and G31(*b*)—were negligible. The small difference in temperature of the two studies (6°C versus 8.5°C) is probably also immaterial.

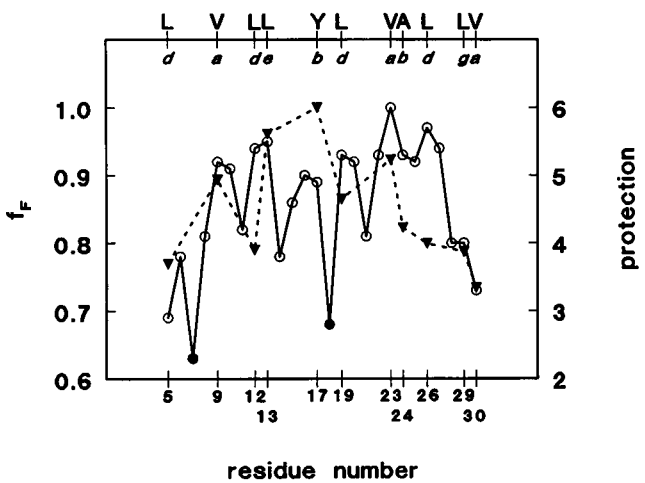

FIGURE 13 Comparison of position profile from NMR-derived local fraction folded for GCN4-lzK (*left ordinate scale, inverted filled triangles, and dashed line*) with that from H/D exchange-derived protection factors for GCN4-lz (*right ordinate scale, open circles, and solid line*). The fraction folded is from NMR data at 8.5°C and 300 μM. The protection factor ordinate is $-\log(k_{ex}/k_{int})$, wherein k_{ex} is the observed rate constant for exchange and k_{int} is the intrinsic value for that amino acid residue. Exchange data are from table 1 in Goodman and Kim (1991). Protection factors plotted as filled circles are upper limits only. Dashed lines simply connect successive data points; they do not necessarily represent behavior between points. The reader is reminded that GCN4-lzK is derived from parent GCN4-lz by R1K, H18K, R25K, and R33K mutations.

Although the local unfolding curves reported here provide a better picture of the transition than available hitherto, they have their limitations, which must be examined. Only the local fraction folded, i.e., the ratio of the number of chains in a locally structured conformation divided by the total number of chains, is obtained. While this is more detailed information than has hitherto been available, it is difficult to derive a precise structural model directly from it. The results certainly imply that both of the most popular models are too simplistic and that only a relatively complex population of states could explain the data. Consequently, it seems likely that only through simulations will the precise structural message in these data emerge.

The difficulty of extracting a precise local structure from local folded fractions is apparent, for example, in the case of the A24(*b*) site (Fig. 11), which has the highest local melting temperature determined so far. As Fig. 3 shows, this local superiority is most manifest at relatively high temperatures, where there is a significant population of single chains as well as dimeric molecules. Is the higher structural content at this site caused by extra local helix content in the separated chains (due, perhaps, to the high intrinsic helix propensity of alanine) or to the effect this local propensity has on the dimeric alanine sites? Our measurements so far cannot answer this question, although future studies with mutant chains may do so.

Two other features of the A24(*b*) site should be kept in mind. First, if the persistence of structure there is due to single-chain helix content, the question immediately arises as to why two separate resonances for structured and unstructured alanine are observed in NMR at all, because the helix-random transition in single chains requires only microseconds or less. Second, although the chemical shift of the folded form's resonance at the A24(*b*) site is typical of α-helix, that of the thermally unfolded form is quite atypical of random alanine, being closer to the helix's value than is usual (Fig. 2). The precise nature of the structures involved is not revealed by these data alone.

Finally, we consider briefly a question raised by the appearance of local instability near the middle of the chain. It is tempting, in view of Figs. 3 and 11, to conclude that many dimer molecules must have a "bubble" of unfolded residues in the region 9–16 between intact, coiled-coil terminal regions. Because such a structure is strongly destabilized by loop entropy, this seems paradoxical (Skolnick and Holtzer, 1986).

However, the paradox only arises because a simplistic view has been adopted. A heuristic example serves to demonstrate that intuition can mislead, where a complex population of states is involved. Consider a dimer made from two, 33-residue chains that consist of three equal-sized domains with only two equally populated molecular states, the two states being, specifically, 1) residues 1–22 random and residues 23–33 in coiled coil; and 2) residues 1–11

coiled coil and residues 12–33 random. The resulting overall population would show 50% coiled coil for both 1–11 and 23–33 terminal domains and zero structural content in the central 12–22 region. Yet no molecule in the population has a "bubble" of unfolded region bounded by coiled-coil domains. Thus even relatively simple populations can yield counterintuitive results (Skolnick and Holtzer, 1986).

It remains to be seen whether extension of the method described here to other sites in GCN4-lzK and to carefully designed mutant chains can help to further define the population of conformational states at equilibrium when a two-stranded coiled coil is thermally unfolded.

NOTE

A detailed summary of the temperature dependence of the chemical shifts at the various sites may be obtained from the corresponding author upon request.

We acknowledge important discussions with Dr. G. Larry Bretthorst concerning the integration of NMR data by Bayesian analysis.

This work was supported in part by grant GM 20064 from the Division of General Medical Sciences, U.S. Public Health Service, and a grant from the Muscular Dystrophy Association.

REFERENCES

Aurora, R., R. Srinivasan, and G. D. Rose. 1994. Rules for helix termination by glycine. *Science.* 264:1126–1130.

Bretthorst, G. L. 1990. Bayesian analysis. III. Applications to NMR signal detection, model selection, and parameter estimation. *J. Magn. Res.* 88:533–595.

Chen, Y., J. Yang, and K. Chau. 1974. Determination of the helix and β form of proteins in aqueous solution by circular dichroism. *Biochemistry.* 13:3350–3359.

Esposito, E., B. Dhanapal, P. Dumy, V. Varma, M. Mutter, and G. Bodenhausen. 1997. Lysine as helix C-capping residue in a synthetic peptide. *Biopolymers.* 41:27–35.

Goodman, E. M., and P. S. Kim. 1991. Periodicity of amide proton exchange rates in a coiled-coil leucine zipper peptide. *Biochemistry.* 30:11615–11620.

Holtzer, M. E., K. Adams, E. G. Lovett, and A. Holtzer. 1996. Tyrosines in two stranded coiled coils are CD active near 280 nm even in the absence of tyrosine-tyrosine interactions. *Biopolymers.* 38:669–671.

Holtzer, M. E., D. L. Crimmins, and A. Holtzer. 1995. Structural stability of short subsequences of the tropomyosin chain. *Biopolymers.* 35: 125–136.

Holtzer, M. E., and A. Holtzer. 1992. α-Helix to random coil transitions: interpretation of the CD in the region of linear temperature dependence. *Biopolymers.* 32:1589–1591.

Holtzer, M. E., A. Holtzer, and J. Skolnick. 1983. The α-helix-to-random-coil transition of two-chain, coiled coils. Theory and experiments for thermal denaturation of α-tropomyosin. *Macromolecules.* 16:173–180.

Holtzer, A., M. E. Holtzer, and J. Skolnick. 1990. Does the unfolding transition of two-chain, coiled-coil proteins involve a continuum of intermediates? *In* Protein Folding. L. M. Gierasch and J. King, editors. AAAS Books, Washington, DC. 177–190.

Hvidt, S., M. E. Rogers, and W. F. Harrington. 1985. Temperature-dependent optical rotatory dispersion properties of helical muscle proteins and homopolymers. *Biopolymers.* 24:1647–1662.

Ishii, Y., S. Hitchcock-DeGregori, K. Mabachi, and S. S. Lehrer. 1992. Unfolding domains of recombinant fusion $\alpha\alpha$-tropomyosin. *Protein Sci.* 1:1319–1325.

Kenar, K. T., B. Garcia-Moreno, and E. Freire. 1995. A calorimetric characterization of the salt dependence of the stability of the GCN4 leucine zipper. *Protein Sci.* 4:1934–1938.

Lehrer, S. S., and W. F. Stafford, III. 1991. Preferential assembly of the tropomyosin heterodimer: equilibrium studies. *Biochemistry.* 30: 5682–5688.

Lovett, E. G., D. A. d'Avignon, M. E. Holtzer, E. H. Braswell, D. Zhu, and A. Holtzer. 1996. Observation via one-dimensional $^{13}C^{\alpha}$ NMR of local conformational substates in thermal unfolding equilibria of a synthetic analog of the GCN4 leucine zipper. *Proc. Natl. Acad. Sci. USA.* 93: 1781–1785.

Lumb, K. J., C. M. Carr, and P. S. Kim. 1994. Subdomain folding of the coiled coil leucine zipper from the bZIP transcriptional activator GCN4. *Biochemistry.* 33:7361–7367.

McLachlan, A. D., and M. Stewart. 1975. Tropomyosin coiled-coil interactions. Evidence for an unstaggered structure. *J. Mol. Biol.* 98:293–304.

Miller, D. W., and K. Dill. 1995. A statistical mechanical model for hydrogen exchange in globular proteins. *Protein Sci.* 4:1860–1873.

Oas, T. G., L. P. McIntosh, E. K. O'Shea, F. W. Dahlquist, and P. S. Kim. 1990. Secondary structure of a leucine zipper determined by nuclear magnetic resonance spectroscopy. *Biochemistry.* 29:2891–2894.

O'Shea, E. K., J. D. Klemm, P. S. Kim, and T. Alber. 1991. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science.* 254:539–544.

O'Shea, E. K., R. Rutkowski, and P. S. Kim. 1989. Evidence that the leucine zipper is a coiled coil. *Science.* 243:538–542.

Privalov, P. L. 1982. Stability of proteins: proteins which do not present a single cooperative system. *Adv. Protein Chem.* 35:1–104.

Shalongo, W., L. Dugad, and E. Stellwagen. 1994. Distribution of helicity within the model peptide acetyl $(AAQAA)_3$ Amide. *J. Am. Chem. Soc.* 116:8288–8293.

Skolnick, J., and A. Holtzer. 1986. α-Helix-to-random-coil transitions of two-chain, coiled coils: a theoretical model for the "pretransition" in cysteine-190-cross-linked tropomyosin. *Biochemistry.* 25:6192–6202.

Sosnick, T. R., S. Jackson, R. R. Wilk, S. W. Englander, and W. F. DeGrado. 1996. The role of helix formation in the folding of a fully α-helical coiled coil. *Proteins.* 24:427–432.

Spera, S., and A. Bax. 1991. Empirical correlations between protein backbone conformation and Cα and Cβ ^{13}C nuclear magnetic resonance chemical shifts. *J. Am. Chem. Soc.* 113:5490–5492.

Wishart, D. S., and B. D. Sykes. 1994. Chemical shifts as a tool for structure determination. *Methods Enzymol.* 239:363–392.

Wishart, D. S., B. D. Sykes, and F. M. Richards. 1991. Relationships between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol.* 222:311–333.