Helix Formation and Stability in a Signal Sequence†

Martha D. Bruch, C. James McKnight, and Lila M. Gierasch*

Departments of Pharmacology and Biochemistry, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75235-9041

Received April 19, 1989; Revised Manuscript Received June 13, 1989

ABSTRACT: A detailed nuclear magnetic resonance analysis of the isolated LamB signal peptide (MMITLRKLPLAVAVAAGVMSAQAMA) under conditions defined by circular dichroism spectra to mimic the conformational distribution of this peptide in membranelike environments has provided a description of specific residue conformational preferences. This 25-residue long peptide in 20 mol % trifluoroethanol in water is in dynamic equilibrium between a helical and a more random conformation, and this equilibrium is shifted toward the more random structure as the temperature is raised. Part of the molecule, residues 10–18, exists in a stable helix at all temperatures studied (5, 25, and 50 °C). Propagation of the helix through the C-terminal end occurs at 25 °C, while the temperature must be lowered to 5 °C to observe any significant population of a helical conformation in the N-terminal region. These results argue that the Pro and Gly residues, which flank the helical segment, act to disfavor helix propagation on their N- or C-terminal sides, respectively. The influence of the Pro residue is stronger than that of the Gly. Furthermore, the most stable part of the helix in this signal peptide under the conditions studied is the hydrophobic core, which is the hallmark of functional signal sequences.

Signal sequences represent good candidates for characterization of secondary structure formation in isolated fragments of polypeptide chain for several reasons. Signal sequences are relatively short (15–30 amino acids), highly hydrophobic sequences which perform several crucial cellular functions. [For reviews, see Briggs and Gierasch (1986), Rapoport (1986), Randall et al. (1987), and Gierasch (1989).] These N-terminal sequences constitute the most general requirement for protein export and are essential for selective targeting of nascent protein chains either to the endoplasmic reticulum, in eukaryotes, or to the cytoplasmic membrane, in prokaryotes. However, there is little primary sequence homology among signal sequences despite the performance of similar, critical cellular functions. Therefore, signal sequences represent challenging cases for deciphering the relationship between the primary sequence and the conformations and interactions of the polypeptide chain. The formation of secondary structure in signal sequences also is important since the presence of the signal sequence may influence the folding of the mature protein in vivo (Park et al., 1988). In fact, some recent results emphasize coupling of folding and targeting (Randall & Hardy, 1986; Eilers & Schatz, 1988). Hence, characterization of the folding of isolated signal peptides may shed some light on protein folding in vivo.

The process by which a protein folds into a specific three-dimensional structure is not well understood. In one commonly proposed model of protein folding, the so-called framework model, local secondary structure formation precedes the formation of specific tertiary structure (Kim & Baldwin, 1982). The framework model suggests that insight into the process of protein folding can be gained by characterization of secondary structure formation in isolated protein fragments. Spectroscopic techniques such as circular dichroism (CD)[1] and two-dimensional nuclear magnetic resonance spectroscopy (2D NMR) are ideally suited for this purpose, and these techniques have been fruitfully applied to the determination of secondary structure in several protein fragments (Bierzynski et al., 1982; Dyson et al., 1988; Wright et al., 1988; Oas & Kim, 1988). In most cases, the secondary structure observed in these fragments was similar to that observed in the native protein. These results suggest that the behavior of isolated protein fragments can provide important insights into the folding of the intact protein.

In previous work from this and other laboratories, a strong preference for α-helical conformation was associated with signal sequences (Rosenblatt et al., 1980; Shinnar & Kaiser, 1984; Briggs & Gierasch, 1984; Batenburg et al., 1988a,b). We now report the detailed analysis of secondary structure formation in a synthetic peptide corresponding to the wild-type signal sequence of the *lamB* gene product (LamB, which is also known as the λ receptor or maltoporin) from *Escherichia coli*. This peptide consists of 25 amino acids with the sequence MMITLRKLPLAVAVAAGVMSAQAMA. Previous CD results indicated that this peptide adopts a similar conformation in an aqueous solution containing 20 mol % trifluoroethanol (TFE), and in membrane-mimetic environments such as sodium dodecyl sulfate micelles or phospholipid vesicles (Briggs, 1986; McKnight et al., 1989). Studies of other peptides in TFE suggest that it stabilizes helix, but not indiscriminately: Regions that have a high tendency to adopt helix become more helical, but regions that are helix-disfavoring are not induced to become helical. For example, a detailed study of the influence of TFE on the ribonuclease S-peptide argued that the TFE increased the helix content but that both the charge-dipole effect and the putative "helix stop signal" were still

† This research was supported by grants from the NIH (GM 34962), the NSF (DCB-8896144), and the Robert A. Welch Foundation.

* Author to whom correspondence should be addressed, at the Department of Pharmacology, University of Texas Southwestern Medical Center.

[1] Abbreviations: CD, circular dichroism; 2D, two dimensional; NMR, nuclear magnetic resonance; TFE, trifluoroethanol; TFA, trifluoroacetic acid; HPLC, high-pressure liquid chromatography; *t*-Boc, *tert*-butyloxycarbonyl; NOESY, two-dimensional nuclear Overhauser effect spectroscopy; NOE, nuclear Overhauser effect; TOCSY, total correlation spectroscopy.

0006-2960/89/0428-8554$01.50/0

operative (Nelson & Kallenbach, 1986, 1989). Therefore, we have chosen TFE/H_2O as a reasonable solvent to use for conformational analysis of this signal peptide. CD results show that this peptide is partially folded into a helical conformation, and 2D NMR results show that the helix can be localized to a specific part of the molecule. Furthermore, CD and NMR spectra as a function of temperature show that the partially folded conformation is in dynamic equilibrium with a more random conformation and that this equilibrium can be shifted toward the more random conformations by raising the temperature. By analysis of spectra at various temperatures, we have monitored initiation and propagation of helix formation spectroscopically. The most stable part of the helix is in the hydrophobic core of the signal peptide.

The LamB signal also allows a fundamental question regarding the structural requirements for helix formation to be addressed. This sequence contains one proline and one glycine, both of which have a low statistical probability of occurring in a helix (Chou & Fasman, 1974). While these residues are known from statistical analysis of protein structures to occur preferentially in nonhelical regions or at helix termini (Chou & Fasman, 1974), their influences on helix formation and stability are not well understood. Our results show that proline has a greater effect on helix stability than does glycine in this peptide.

Experimental Procedures

Synthesis. This synthesis of the LamB signal peptide was carried out by using standard solid-phase methods on an ABI 430A automated peptide synthesizer. *t*-Boc-protected amino acids were purchased from ABI. The peptide was cleaved from the resin by treatment with anhydrous HF in the presence of 10% anisole for 1 h at 0 °C. Crude peptide was purified by reverse-phase HPLC using a Vydac phenyl column (25 cm × 2 cm), with a water/acetonitrile gradient (0.1% TFA). Amino acid composition was confirmed by amino acid analysis on a Beckman 6300 amino acid analyzer. The sequence of the peptide was confirmed by liquid-phase sequencing on an ABI 477 protein sequencer. Details of synthetic procedures are reported elsewhere (Briggs & Gierasch, 1984; Briggs, 1986).

Circular Dichroism. All CD measurements were made on an AVIV Model 60DS spectrophotometer using a 5-mm cell at temperatures of 5, 25, and 50 °C. Temperature was regulated by a Hewlett-Packard 89100A temperature controller. The peptide concentration was determined by quantitative amino acid analysis. The solvent was 50% 5 mM Tris buffer (pH 7.3) and 50% trifluoroethanol (spectral grade, Aldrich) by volume, which is 20 mol % TFE. CD spectra also were obtained in 50% TFE and 50% phosphate buffer (pH 7.0, 4.0, 2.5) at 25 °C. All spectra were base-line-corrected and smoothed. Spectra were the average of three consecutive scans obtained with a total duration of 9 s at each nanometer from 250 to 190 nm.

Nuclear Magnetic Resonance. All NMR measurements were made on a Varian VXR 500 spectrometer operating at a proton frequency of 500 MHz. Peptide concentration in NMR samples was 2 mM. The solvent was 50% (by volume) trifluoroethanol-d_3 (Merck Isotopes) and 50% unbuffered water adjusted to a pH of ca. 4 to reduce the exchange rate of amide protons with water (which is equivalent to 20 mol % TFE). For all 2D NMR spectra, 256 free induction decays were obtained containing 2K complex points each, and the data were zero-filled to form a matrix of 1K by 1K real points. The sweep width was 6000 Hz in both dimensions, and 64 or 128 transients were accumulated per t_1 value, with a 1-s relaxation delay. The mixing time was 75 ms for the TOCSY spectra and 300 ms for the NOESY spectra. The water resonance was suppressed by preirradiation during the relaxation delay and during the mixing period in the NOESY spectra. All 2D data sets were collected in the phase-sensitive mode (States et al., 1982) and were analyzed in both unsymmetrized and symmetrized (Bauman et al., 1981) form. Symmetrized data sets are presented for clarity, and considerable care was taken to distinguish symmetrization artifacts from real cross-peaks in all 2D spectra. Data were processed on a Sun 4/260 computer, using software developed by Dennis Hare (Infinity Systems, Seattle, WA).

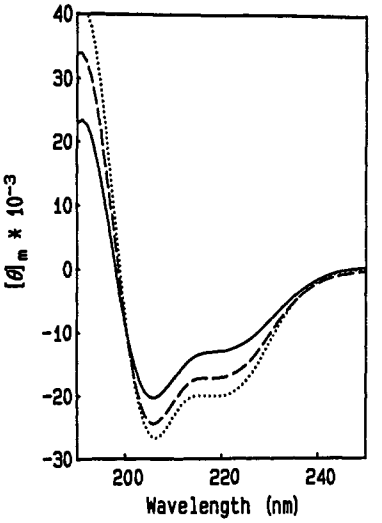

FIGURE 1: CD spectra of the LamB wild-type signal peptide in 20 mol % TFE as a function of temperature; peptide concentration 10 μM. Temperatures are 50 °C (solid line), 25 °C (dashed line), and 5 °C (dotted line).

Results

Helix Content by CD as a Function of Temperature. The CD spectra of the LamB signal peptide in 50/50 TFE/H_2O (by volume) at 5, 25, and 50 °C, shown in Figure 1, all exhibit characteristic double minima indicative of α-helical structure (Woody, 1985), and this implies that at least part of the peptide is in an α-helix at all temperatures. The spectra were fitted by computer to Greenfield–Fasman polylysine reference spectra (Greenfield & Fasman, 1969) to estimate the relative contributions of α-helix, β-sheet, and "random coil". This procedure gave values of 55% α-helix and 45% random coil at 25 °C. Similarly, if the molar ellipticity at 222 nm (θ_{222}) is used to calculate the percentage of α-helix for a 25-residue peptide (Chen et al., 1974), a value of 52% α-helix is obtained at 25 °C. When θ_{222} is used to estimate helix content, the amount of α-helix decreases to 39% at 50 °C but increases to 60% when the sample is cooled to 5 °C. These results suggest that the peptide is in dynamic equilibrium between an α-helical conformation and a more random structure and that this equilibrium is shifted toward the more random structure at higher temperatures. This interpretation is further supported by the isodichroic point at 201 nm which is indicative of interconversion between two states. It should be noted that CD spectra obtained in 50% TFE (by volume) and 50% phosphate buffer adjusted to pH values of 2.5, 4.0, and 7.0 are indistinguishable. This result demonstrates that pH does not have a significant influence on the conformation of the LamB signal peptide in TFE/H_2O. Furthermore, the molar ellipticity observed in the CD spectra of this peptide in 50/50 TFE and pH 4 phosphate buffer is invariant within experimental error over a range of concentrations from 0.4 μM to

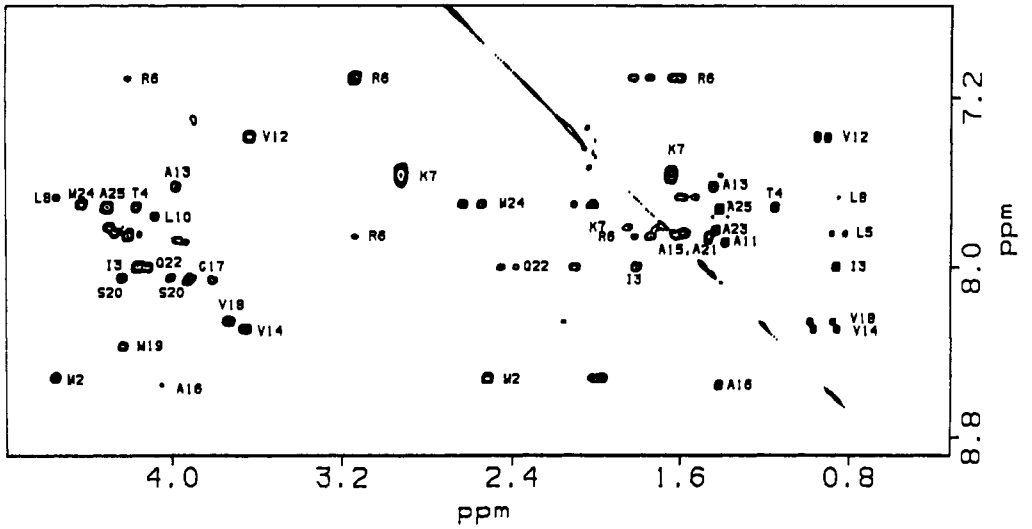

FIGURE 2: TOCSY spectrum of the LamB peptide in 20 mol % TFE at 25 °C. Peptide concentration 1 mM; mixing time 75 ms. Connectivities involving amide and aliphatic protons are shown.

2 mM. This result argues strongly that the peptide is monomeric under the conditions of these studies.

Conformational Analysis by 2D NMR. The circular dichroism results yield a description of the behavior of the peptide as a whole, but CD gives no insight into the conformational states of specific residues within the peptide. Residue-specific conformational information can be obtained from 2D NMR spectroscopy. Two-dimensional nuclear Overhauser effect spectroscopy (NOESY) yields a map of protons which are spatially close together, irrespective of bonding networks (Kumar et al., 1980; Macura et al., 1981). Although it is difficult to establish rigorously an upper limit on the interproton distance required to observe cross-peaks in the NOESY spectrum, a reasonable estimate is 4.5 Å under the conditions of our experiments. In a polypeptide chain that exists in a single conformational state, the type and pattern of the short distances (<4.5 Å) inferred from the NOESY spectrum indicate the presence of characteristic secondary structures. A regular α-helix is characterized by the distance between adjacent amide protons, $d_{NN}(i, i + 1)$, which is ~2.8 Å, the distance between the α and amide protons in neighboring residues, $d_{\alpha N}(i, i + 1)$, which is ~3.5 Å, and $d_{\alpha N}(i, i + 3)$, which is ~3.4 Å (Wuethrich et al., 1984). In addition, since $d_{\alpha N}(i, i + 4)$ and $d_{NN}(i, i + 2)$ are both approximately 4.2 Å, these interactions are anticipated in the NOESY spectrum of a regular, stable helix. By contrast, in extended structures, $d_{\alpha N}(i, i + 1)$ is ~2.2 Å, and $d_{NN}(i, i + 1)$ is ~4.3 Å. Since all other distances are greater than 4.5 Å, the dominant interactions expected for an extended structure are those between $\alpha(i)$ and NH($i + 1$). Another NMR parameter which can aid in identification of α-helical structure is the vicinal coupling constant between the amide proton and the α proton of each residue, $^3J_{HN\alpha}$. Residues in helical structures have small coupling constants ($^3J_{HN\alpha} < 5$ Hz), whereas residues in extended structures have significantly larger coupling constants ($^3J_{HN\alpha} > 8$ Hz) (Pardi et al., 1984). Therefore, observation of patterns of NOE connectivities and measurement of vicinal coupling constants allow secondary structure to be determined and localized to specific regions in the molecule. It is important to note that in the case of fast interconversion between more than one conformation, the observed parameters will be the population-weighted average of the parameters associated with the individual states.

The LamB signal peptide resonance assignments were obtained by the standard sequential assignment procedure (Billeter et al., 1982). Assignment of the set of lines associated with each type of amino acid was accomplished though total correlation spectroscopy (TOCSY) (Davis & Bax, 1985). Each amino acid has a characteristic pattern of cross-peaks observed in the TOCSY spectrum (Figure 2), so assignment as to type of amino acid can be made through pattern recognition of the TOCSY cross-peaks. Sequentially adjacent amino acids typically have some pair of protons which are spatially close, and the observed NOE, arising from these short distances between neighboring residues are used to make sequential assignments. The assignments are summarized in Table I.

Table I: Summary of NMR Proton Line Assignments for the LamB Signal Peptide in TFE/H_2O at 25 °C

	^{1}H chemical shift[a]				
residue	NH	Hα	Hβ	Hγ	others
Met-1		4.09	2.12	2.54	
Met-2	8.52	4.54	2.03, 1.97	2.52	
Ile-3	8.00	4.16	1.81	1.45, 1.15, 0.88	0.85
Thr-4	7.71	4.30	4.17	1.15	
Leu-5	7.84	4.27	1.59	1.59	0.89, 0.82
Arg-6	7.86	4.21	1.82, 1.74	1.64, 1.60	3.13
Lys-7	7.81	4.29	1.85, 1.71	1.43, 1.35	1.64, 2.92
Leu-8	7.68	4.54	1.60	1.52	0.85
Pro-9		4.37	2.35, 1.83	2.0	3.48
Leu-10	7.76	4.08	1.66	1.60	0.85, 0.90
Ala-11	7.88	3.94	1.39		
Val-12	7.39	3.64	2.04	0.96, 0.92	
Ala-13	7.63	3.98	1.44		
Val-14	8.30	3.66	2.04	0.97, 0.87	
Ala-15	7.87	3.98	1.46		
Ala-16	8.55	4.05	1.42		
Gly-17	8.06	3.92, 3.81			
Val-18	8.26	3.74	2.15	0.99, 0.89	
Met-19	8.38	4.23	2.11, 2.01	2.62, 2.54	
Ser-20	8.05	4.23	4.01, 3.91		
Ala-21	7.86	4.15	1.46		
Gln-22	8.00	4.11	2.09	2.45, 2.39	
Ala-23	7.83	4.20	1.43		
Met-24	7.71	4.42	2.11, 2.05	2.66, 2.55	
Ala-25	7.73	4.30	1.42		

[a] Chemical shifts are in ppm relative to 3.88 ppm for the CH_2 in TFE, and are ±0.02 ppm.

The NOESY spectrum of the LamB signal peptide at 25 °C (Figure 3) contains a large number of cross-peaks connecting amide protons. Assignment of the NH/NH NOEs was facilitated by knowledge from CD that the molecule has a high helical content. This overall structure implies that

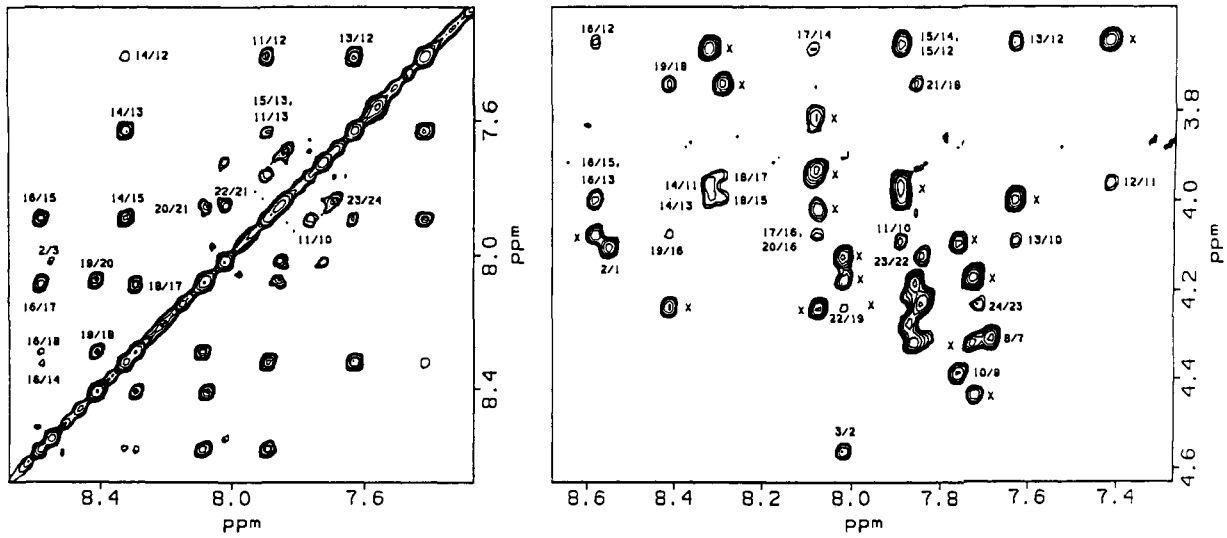

FIGURE 3: Expansions of the NOESY spectrum of the LamB signal peptide in 20 mol % TFE at 25 °C; peptide concentration 2 mM; mixing time 300 ms. On the left is the region containing NH to NH connectivities; on the right is the region containing NH connectivities to upfield protons. Intraresidue interactions are denoted by a "x".

interactions between sequentially close residues are much more likely than interactions between residues which are far apart in the sequence; this assumption was used in a few cases to resolve potential ambiguities from spectral overlap. All sequential NH(i)/NH(i + 1) interactions but one are observed for residues 10–24, which implies a significant population of a helical conformation for the residues in this region. Any NOE between M24 and A25 amide protons would not be observable due to the similarity of their chemical shifts. The α-helical conformation is confirmed by the large number of $\alpha(i)$/NH(i + 3) and $\alpha(i)$/NH(i + 4) interactions observed. There are probably additional medium-range interactions [$\alpha(i)$/NH(i + 3)] in the C-terminal region which cannot be observed because of spectral overlap. Further evidence of a high helical content is provided by observation of a network of weak NH(i)/NH(i + 2) interactions from residues 11 to 18. By contrast, there is no evidence for any of these interactions indicative of a α-helix for the N-terminal part of the peptide, residues 1–8. The entire network of $\alpha(i)$/NH(i + 1) interactions is observed in this region, and the large size of these cross-peaks, as well as the absence of other interactions to these amide protons, suggests that this part of the molecule is predominantly in an extended conformation. There is a weak NOE between the M2 and I3 amide protons, but this interproton distance may be sufficiently short (~4.3 Å) to see a weak NOE even in an extended structure. Due to the small chemical shift difference between resonances of the L5 and R6 amide protons, it cannot be determined whether or not an interaction exists between these protons. In addition, two other interactions between amide protons in the N-terminal region of the molecule would be difficult to observe because of partial spectral overlap. Nonetheless, the complete absence of any medium-range [e.g., $\alpha(i)$/NH(i + 3)] interactions among residues 1–8, most of which would be observable if present, argues strongly that there is no significant helical content in the N-terminal region of this peptide.

Since proline does not have an amide proton, the δ-methylene protons, which occupy a similar position to the amide proton in other residues, can be substituted for the amide proton in the interactions expected for a given type of secondary structure. In this peptide at 25 °C, a NOE is observed between Leu-8 NH and Pro-9 H^{δ}, and a very weak NOE is observed between Pro-9 H^{δ} and Leu-10 NH. These observations suggest that the helix may extend from residue 8 to the C-terminal end of the peptide.

Additional confirmation for the localization of the helix from residues 8–24 is obtained by measuring relative rates of chemical exchange of amide protons with D_2O. To determine the solvent accessibility of the various amide protons, the peptide sample was freshly prepared in a mixture of TFE-d_3 and D_2O (pH ~4), and proton spectra were obtained as a function of time. Those amide protons that remain after several hours in a peptide of this size represent amides which are inaccessible to the solvent, most often through intramolecular hydrogen bonding. Slow exchange of amide protons in residues 8–23 was indicated by the presence of these proton signals after 1 h in TFE/D_2O, while amide protons in the N-terminal exchange significantly with D_2O in 1 h. This result provides further evidence that residues 8–24 exist in an α-helix a substantial portion of the time. The NOE and exchange results are summarized schematically in Figure 4a.

Thus far, we have demonstrated that the region of the signal peptide from residues 8–24 has a high helical population while residues 1–8 have little, if any, helical content at 25 °C. However, there is a difference in the pattern of NOE interactions observed for the region from residues 10–18 and that from residues 19–24. Both NH(i)/NH(i + 2) and $\alpha(i)$/NH(i + 3) NOEs are seen in the former region, while only $\alpha(i)$/NH(i + 3) interactions are observed in the latter region. Since $d_{NN}(i, i+2)$ is longer than $d_{\alpha N}(i, i+3)$ (~4.2 Å versus ~3.4 Å, respectively), observation of NH(i)/NH(i + 2) interactions would require the helix to be more stable (i.e., a higher population of the helical conformation) than would observation of $\alpha(i)$/NH(i + 3) interactions. This suggests that the helix is more stable in the hydrophobic core (residues 10–18) than in the C-terminal part of the molecule (residues 19–24). This conclusion is further supported by the slower exchange rates observed for amide protons of residues 10–19 (still present after 3 h) relative to those observed in the C-terminal part of the peptide (completely exchanged by 3 h).

The conformational interpretation above is presented as though the cross-peaks observed in the NOESY spectrum arise from primary NOE interactions between two protons. In fact, it is unusual to observe NOEs corresponding to distances ≥3.5 Å in peptides of the size of the LamB signal sequence (Wright et al., 1988).[2] Thus, we must consider the possibility that the

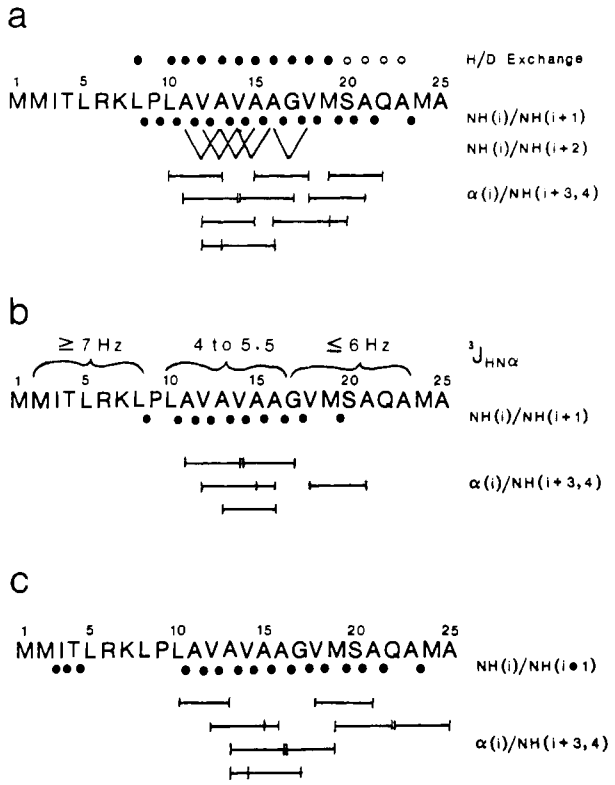

FIGURE 4: Summary of NMR data for the LamB signal peptide in 20 mol % TFE. (a) 25 °C. The top line summarizes amide exchange rates with D_2O. Closed circles indicate amide protons that remained for at least 3 h after addition of D_2O. Open circles indicate protons that remained after 1 h. Lines below the sequence indicate observed Overhauser effects between indicated protons. (b) 50 °C. The top line summarizes vicinal coupling constant data from Table II. Note that $^3J_{HN\alpha}$ of Val-14 is 5.8 (±0.2) Hz. Bottom lines summarize NOE data. (c) 5 °C NOE data. Note that the $\alpha(i)/NH(i + 3, 4)$ interactions listed for 5 °C are probable NOEs since extensive spectral overlap at this temperature precludes unambiguous assignment in some cases.

weaker interactions observed [$NH(i)/NH(i + 2)$ and $\alpha(i)/NH(i + 3, 4)$] arise from secondary handoffs where polarization is transmitted by spin diffusion through a third proton. Spin diffusion seems improbable between the α protons and NHs three or four residues ahead in the chain because there is no likely path (intermediate proton) for the transfer. The $NH(i)/NH(i + 2)$ interaction could readily arise from a secondary transfer of polarization subsequent to the $NH(i)/NH(i + 1)$. If this is the origin of the cross-peaks observed, our conclusions about proximity of the NHs are still valid, since the efficiency of the second transfer is also distance-dependent. Therefore, the observation of $NH(i)/NH(i + 2)$ NOEs selectively in the hydrophobic core implies that the helix is more stable in this region irrespective of the origin of these cross-peaks, i.e., whether they arise from primary polarization transfers, spin diffusion, or both.[3]

Temperature Effects on Helical Structure for Specific Residues of the LamB Signal Peptide. The amide region of the NOESY spectrum at 50 °C (Figure 5a) contains many of the same interactions observed in the 25 °C spectrum, but it is missing the NH/NH interactions previously observed in the C-terminal part of the peptide (residues 20–25). In addition, there are no medium-range interactions indicative of an α-helix in the C-terminal region at 50 °C. There are several reasons why the NOEs may be smaller at 50 °C relative to 25 °C. As a consequence of the lower solution viscosity at 50 °C, the molecular correlation time will be smaller, and this will result in decreased intensity of *all* cross-peaks relative to those at 25 °C. Hence, it is difficult to compare directly cross-peak intensities between NOESY spectra at the two temperatures. However, a shift in the conformational distribution from helix to random also can reduce the magnitude of the NOE at the higher temperature due both to an increased local mobility and to an increase in average interproton distances of conformationally diagnostic proton pairs. Hence, differences in cross-peak intensities for different parts of the molecule *within* the 50 °C NOESY data set will reflect differences in the conformational population, viz., the helix content. The pattern of NOEs observed at 50 °C (Figure 4b) shows that the helix is primarily localized to residues 10–21. The larger number of both $NH(i)/NH(i + 1)$ and $\alpha(i)/NH(i + 3)$ cross-peaks of residues within the hydrophobic core suggests that this region contains the highest helix content, which agrees well with the conclusions drawn from the 25 °C data.

Note that only the weakest NOE interactions [$NH(i)/NH(i + 2)$] have been lost from the region of the hydrophobic core in the 50 °C data set relative to the 25 °C data, whereas many of the interactions indicative of helix content are missing in the C-terminal region. The loss of the weak NOE cross-peaks in the hydrophobic core is likely to have occurred because of the reduced efficiency of both primary and secondary NOE transfers at the higher temperature, as discussed above, rather than because of a significant conformational redistribution, since all of the other diagnostic helical NOEs are still present. By contrast, there are only a few interactions indicative of a helix in the C-terminal region (residues 19–25) at 50 °C. Since a dense network of interactions was observed in this region at 25 °C, this suggests that the helix has been destabilized considerably in this region at the higher temperature (equilibrium shifted toward the random state).

Our conclusions regarding relative helix content in the different regions of the signal peptide at 50 °C are supported as well by the magnitudes of the amide coupling constants. The sharper resonance lines at 50 °C allow amide vicinal coupling constants to be measured quantitatively from 2D *J*-resolved spectroscopy (Aue et al., 1976; Nagayama et al., 1977) (Table II). In the N-terminal region (residues 1–8), all coupling constants are greater than 7 Hz, and these large values confirm the absence of helix in this region. By contrast, in the region from residues 10–16, most vicinal coupling constants are less than 5 Hz; those of Val-12 and Val-14, at

[2] Although often observed in proteins, the NOEs corresponding to distances longer than 3.5 Å are rarely seen in peptides. However, most of the detailed studies of peptides of similar length to the signal peptide have been carried out on aqueous solutions. The TFE/H_2O solvent mixture leads to substantially broader line widths than H_2O alone and is probably creating a distinctly different frequency of molecular tumbling. This may account for the higher efficiency of polarization transfer (protein-like). We do not know the origins of this effect, since the viscosity of TFE/water solutions is not significantly higher than that of water solutions. Immiscibility of the two solvents may give rise to domains of solvent around the peptide solutes.

[3] To ensure that NOEs potentially arising from spin diffusion at the mixing time of 300 ms do not result in erroneous conclusions, a NOESY spectrum was obtained using a mixing time of 100 ms. All cross-peaks were significantly less intense than in the 300-ms mixing time spectrum, and only the strongest interactions were large enough to be observed. Nonetheless, a dense network of $NH(i)/NH(i + 1)$ interactions occurs from residues 12 to 22. These results confirm the presence of a high population of helix in the C-terminal part of the molecule and demonstrate that any effects due to spin diffusion are not leading to erroneous conclusions regarding the location of the helix.

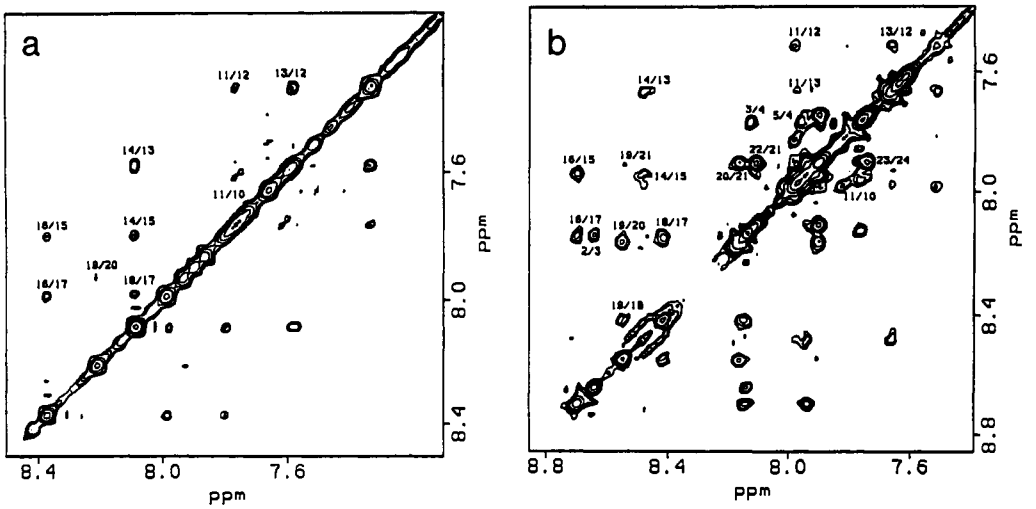

FIGURE 5: Expansions of the NH regions of the NOESY spectra of the LamB wild type in 20 mol % TFE; peptide concentration 2 mM; mixing time 300 ms. (a) 50 °C, (b) 5 °C.

Table II: Vicinal Amide Coupling Constants in TFE/H_2O at 50 °C

residue	chemical shift[a]	$^3J_{HN\alpha}$ (Hz)[b]
Met-2	8.43	7.3
Ile-3	7.89	7.5
Thr-4	7.68	7.8
Leu-5	7.74	7.0
Arg-6[c]	7.79	7.2
Lys-7	7.75	7.5
Leu-8	7.62	7.6
Leu-10	7.66	4.4
Ala-11[c]	7.79	4.9
Val-12	7.33	5.5
Ala-13	7.60	4.4
Val-14	8.11	5.8
Ala-15[d]	7.81	5.0
Ala-16	8.39	4.3
Gly-17[e]	8.00	5.8
Val-18	8.09	5.9
Met-19	8.21	5.7
Ser-20	7.94	5.8
Ala-21[d]	7.81	5.0
Gln-22	7.92	5.9
Ala-23	7.76	5.6
Met-24	7.68	7.8
Ala-25	7.69	6.6

[a] Estimated uncertainty on chemical shift is ±0.01 ppm. [b] Estimated uncertainty on coupling constants is ±0.2 Hz. [c] Arg-6 and Ala-11 are overlapped, but two components are present in the *J*-resolved spectrum with different coupling constants. Assignments of the two components cannot be made unambiguously. [d] Ala-15 and Ala-21 are overlapped, and only one component is observed in the *J*-resolved spectrum. Therefore, the coupling constants are assumed to be equal for these two residues. [e] A triplet was observed for Gly-17 which indicates that the coupling constants are the same to both α protons.

5.5 and 5.8 Hz, are still relatively small. This supports the notion of a nearly full population of α-helix in the hydrophobic core. Residues 17–23 have somewhat larger coupling constants (between 5.5 and 6 Hz); both the observations of larger coupling constants and fewer NOEs indicative of helix argue that there is a significant decrease in helix content in the C-terminal region of the signal peptide. By comparison, the data suggest that helix content is still high in the hydrophobic core. Therefore, we conclude that the decrease in helical content observed in the CD spectrum at 50 °C is primarily due to unraveling of the helix from the C-terminal end.

The increased helical content observed in the CD spectrum at 5 °C arises from an increase in the number of residues in a helical conformation as demonstrated by the summary of interactions observed in the NOESY spectrum at 5 °C (Figure 4c). At this temperature, the entire region from residues 10–25 exhibits the network of short-range [NH(i)/NH(i + 1)] and medium-range [$\alpha(i)$/NH(i + 3)] NOE interactions expected for an α-helix, as observed at 25 °C. There are also a few NH(i) to NH(i + 2) interactions observed, but these weak interactions are difficult to observe at 5 °C due to the broadness of the amide resonances. These results indicate that the entire region from residues 10–25 has a high population of an α-helical conformation. Moreover, the amide region of the NOESY spectrum at 5 °C (Figure 5b) contains cross-peaks between amide protons in the N-terminal region, most notably due to M2/I3, I3/T4, and T4/L5 interactions. These additional interactions suggest that the helix has started to propagate into the N-terminal region of this peptide at the lower temperature. However, no medium-range interactions indicative of a helix [$\alpha(i)$/NH(i + 3)] are observed in residues 1–8. These results imply that, in the N-terminal part of the peptide, the average distance $d_{NN}(i, i + 1)$ is sufficiently short to allow the observation of NH/NH NOEs but that, on the average, $d_{\alpha N}(i, i + 3)$ is too long for medium-range NOEs to be observed. These results suggest that the helical conformation for residues 1–8 is only marginally stable compared to the stable helix formed for residues 10–25. Lowering the temperature has shifted the equilibrium between helical and random structures so that the helical conformation is more favorable for the entire peptide.

Discussion

NMR analysis of the LamB signal peptide yields a picture of a sequence with a strong conformational preference within a particular region, but with a conformational equilibrium existing throughout the molecule. The central hydrophobic core of this signal sequence adopts an α-helical conformation that is stable to 50 °C. The C-terminal segment of this peptide (residues 18–25) shows a lower, but still significant tendency to take up a helical conformation. By comparison, the N-terminal region (residues 1–8) shows only a low population of helical conformation, even at lower temperatures. Flanking the hydrophobic core of this peptide are the proline residue at position 9 and the glycine residue at position 17. Clearly, these two conformationally influential residues are affecting the structure of the signal peptide. Both residues are believed to disfavor helix (Chou & Fasman, 1974). However, it is known that helices in globular proteins often begin with a Pro and, similarly, Gly residues often occur at the C-terminal end (Richardson & Richardson, 1988). Even in this short, isolated sequence containing Pro and Gly, we see analogous effects:

The Pro residue is included in the helix, but the propagation of the helix to the N-terminal side of the Pro is apparently disfavored. Only the residue immediately preceding Pro (Leu-8) shows distances typical of a helical state. The Gly residue, by contrast, appears to influence its C-terminal neighbors to take up helix less of the time. The effect of the Gly residue is significantly less than that of the Pro. Even at 50 °C the C-terminal region spends significant time in helix, while the N-terminal segment only shows the beginnings of helical structure at 5 °C.

This analysis illustrates the power of NMR to describe conformational behavior. Not only can one deduce the three-dimensional structure in cases where one conformational species is predominant, but also one can in favorable instances analyze a system in conformational equilibrium with rapid interconversion among states on the NMR time scale. The fact that NOEs and coupling constants average according to distinct functions and the availability of several distances, which can be used to set limits on populations in different states (as illustrated in the analysis above), combine to yield a plausible analysis if the nature of the interconverting states is known. Since polypeptide chains are restricted to regions of ϕ, ψ space that are quite well-defined, this approach should prove to be fruitful in protein folding studies. Our interpretation was further supported by deuterium exchange data, which can generally be used to test the consistency of conformational conclusions.

An interesting comparison can be made between the CD results and the NMR data presented so far: The NMR results show that at 25 °C, the hydrophobic core (residues 10–18), which comprises 36% of the molecule, is in a very stable helix, while another 32% of the molecule (residues 8, 9, and 19–24) has a significant, but smaller population of a helical conformation. In order to get an approximate estimate of the overall helical content of the molecule from the NMR data, we can assume that the hydrophobic core is helical all of the time while the region with significant, but lower helix content is helical 50% of the time. These admittedly rough calculations indicate from NMR a total helical population of ~52% [=36% + $^1/_2$(32%)], which is in surprisingly good agreement with the estimate of 52–55% helical content from CD spectra.

The present studies made use of aqueous trifluoroethanol as a solvent. While generally known to promote helix formation, this solvent is poorly understood in its influences on polypeptide structure. Results of detailed analyses on other peptides suggest that TFE stabilizes helix in regions that have a high intrinsic helical propensity (Nelson & Kallenbach, 1986, 1989; Merutka & Stellwagen, 1989). Furthermore, the LamB signal sequence displays similar overall content of helix in the chosen TFE/H_2O mixture as it does in SDS micelles or in phospholipid vesicles. Interestingly, our conformational analysis of the LamB wild-type signal peptide in transferred lipid monolayers using both CD and Fourier-transform infrared spectroscopy (Briggs et al., 1986; Cornell et al., 1989) led to a model for the structure of the lipid-inserted peptide that is virtually the same as that deduced from the present NMR analysis. Nonetheless, it is possible that the details of the residue preferences for helix may differ in the interfacial environments from those seen here. Thus, we intend to extend our NMR analysis to the micellar system to check whether the present findings are general.

Although the LamB signal sequence provides only one example, it is tempting to speculate that the observed high helical stability in the hydrophobic core will be general and has functional significance, since this region is the hallmark of functional signal sequences. Our previous results and current models for prokaryotic protein export would suggest that there are multiple roles for the signal sequence and, furthermore, that the helical character of the hydrophobic core may be required both for lipid interactions (Briggs & Gierasch, 1984; Briggs et al., 1985; McKnight et al., 1989) and for recognition by proteinaceous species such as SecA or PrlA (SecY) (Chen et al., 1987; A. Sgrignoli, L. L. Chen, P. C. Tai, and L. M. Gierasch, unpublished results).

In future work, we will examine the alterations in the conformational properties of mutated LamB signal sequences, we will test the generality of the present results by analysis of other signal sequences, and we will explore the influence of the mature sequence on the conformational behavior of the signal sequence.

ACKNOWLEDGMENTS

We thank Arthur Pardi for helpful discussion about NMR methods and Dennis Hare for making his software available and for giving us advice in its use. We appreciate the expert technical assistance of Sarah Stradley, Candace Millhouse, and Khuan Ng.

Registry No. MMITLRKLPLAVAVAAGVMSAQAMA, 97399-84-3.

REFERENCES

Aue, W. P., Karhan, J., & Ernst, R. R. (1976) *J. Chem. Phys. 64*, 4226–4227.

Batenburg, A. M., Brasseur, R., Ruysschaer, J.-M., van Scharrenburg, G. J. M., Slotboom, A. J., Demel, R. A., & de Kruijff, B. (1988a) *J. Biol. Chem. 263*, 4202–4207.

Batenburg, A. M., Demel, R. A., Verkleij, A. J., & de Kruijff, B. (1988b) *Biochemistry 27*, 5678–5685.

Bauman, R., Wider, G., Ernst, R. R. & Wuethrich, K. (1981) *J. Magn. Reson. 44*, 402–406.

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

Billeter, M., Braun, W., & Wuethrich, K. (1982) *J. Mol. Biol. 155*, 321–346.

Briggs, M. S. (1986) Ph.D. Dissertation, Yale University.

Briggs, M. S., & Gierasch, L. M. (1984) *Biochemistry 23*, 3111–3114.

Briggs, M. S., & Gierasch, L. M. (1986) *Adv. Protein Chem. 38*, 109–179.

Briggs, M. S., Gierasch, L. M., Zlotnick, A., Lear, J., & DeGrado, W. F. (1985) *Science 228*, 1096–1099.

Briggs, M. S., Cornell, D. G., Dluhy, R. A., & Gierasch, L. M. (1986) *Science 233*, 206–208.

Chen, L. L., Tai, P. C., Briggs, M. S., & Gierasch, L. M. (1987) *J. Biol. Chem. 262*, 1427–1429.

Chen, Y.-H., Yang, J. T., & Chan, K. H. (1974) *Biochemistry 13*, 3350–3359.

Chou, P. Y., & Fasman, G. D. (1974) *Biochemistry 13*, 211–245.

Cornell, D. G., Dluhy, R. A., Briggs, M. S., McKnight, C. J., & Gierasch, L. M. (1989) *Biochemistry 28*, 2789–2797.

Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2820–2821.

Dyson, H. J., Rance, M., Houghten, R. A., Wright, P. E., & Lerner, L. A. (1988) *J. Mol. Biol. 201*, 201–217.

Eilers, M., & Schatz, G. (1988) *Cell 52*, 481–483.

Gierasch, L. M. (1989) *Biochemistry 28*, 923–930.

Greenfield, N., & Fasman, G. D. (1969) *Biochemistry 8*, 4108–4115.

Kim, P. S., & Baldwin, R. L. (1982) *Annu. Rev. Biochem. 51*, 459–489.

Kumar, A., Ernst, R. R., & Wuethrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1–6.
Macura, S., Huang, Y., Suter, D., & Ernst, R. R. (1981) *J. Magn. Reson. 43*, 259–281.
McKnight, C. J., Briggs, M. S., & Gierasch, L. M. (1989) *J. Biol. Chem.* (in press).
Merutka, G., & Stellwagen, E. (1989) *Biochemistry 28*, 352–357.
Nagayama, K., Wuethrich, K., Bachmann, P., & Ernst, R. R. (1977) *Biochem. Biophys. Res. Commun. 78*, 99–105.
Nelson, J. W., & Kallenbach, N. R. (1986) *Proteins: Struct., Funct., Genet. 1*, 211–217.
Nelson, J. W., & Kallenbach, N. R. (1989) *Biochemistry 28*, 5256–5261.
Oas, T. G., & Kim, P. S. (1988) *Nature 336*, 42–48.
Pardi, A., Billeter, M., & Wuethrich, K. (1984) *J. Mol. Biol. 180*, 741–751.
Park, S., Liu, G., Topping, T. B., Cover, W. H., & Randall, L. L. (1988) *Science 239*, 1033–1035.
Randall, L. L., & Hardy, S. J. S. (1986) *Cell 46*, 921–928.
Randall, L. L., Hardy, S. J. S., & Thom, J. R. (1987) *Annu. Rev. Microbiol. 41*, 507–541.
Rapoport, T. A. (1986) *CRC Crit. Rev. Biochem. 20*, 73–137.
Richardson, J. S., & Richardson, D. C. (1988) *Science 240*, 1648–1652.
Rosenblatt, M., Beaudette, N. V., & Fasman, G. D. (1980) *Proc. Natl. Acad., Sci. U.S.A. 77*, 3983–3987.
Shinnar, A. E., & Kaiser, E. T. (1984) *J. Am. Chem. Soc. 106*, 5006–5007.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.
Woody, R. W. (1985) in *The Peptides* (Hruby, V. J., Ed.) Vol. 7, pp 15–114, Academic Press, Orlando, FL.
Wright, P. E., Dyson, H. J., & Lerner, R. A. (1988) *Biochemistry 27*, 7167–7175.
Wuethrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715–740.

A Thermodynamic Model for the Self-Association of Human Spectrin†

Michael Morris and G. B. Ralston*
Department of Biochemistry, University of Sydney, Sydney NSW 2006, Australia
Received October 27, 1988; Revised Manuscript Received May 22, 1989

ABSTRACT: The self-association of human spectrin at 28.8 °C in 0.11 M salt (pH 7.5) has been studied by means of sedimentation equilibrium. Coincidence of Ω function plots as a function of total spectrin concentration (0–2 g/L) indicated that equilibrium was achieved and that no significant concentration of solute was incapable of participating in the self-association reaction. On the basis of the root-mean-square deviation of the fits and the randomness of the residuals, the behavior can be described equally well, either by a cooperative isodesmic model, in which $K_{12} \approx 2 \times 10^6$ M^{-1} and all other $K \approx 10^6$ M^{-1}, or by an attenuated scheme in which $K_{(i-1)i} \approx (3.5 \times 10^6)/i$ M^{-1}. The returned values of the second virial coefficient, B, for both these models fall within the range calculated from the charge and Stokes radius of spectrin. A mechanism for spectrin self-association consistent with both schemes is proposed in which spectrin heterodimers undergo a reversible opening at the self-association interface. These open heterodimers then undergo indefinite self-association to form a series of open-chain oligomers in dynamic equilibrium with closed-loop oligomers.

Erythrocyte spectrin is capable of self-associating through the sequential addition of heterodimers to form tetramers (Ralston, 1978; Shotton et al., 1979) and higher oligomers (Morrow & Marchesi, 1981; Morrow et al., 1981; Morris & Ralston, 1984; Liu et al., 1984). The self-association of spectrin is critical for normal red cell shape, flexibility, and resistance to hemolysis (Palek & Lux, 1983; Mohandas et al., 1983; Elgsaeter et al., 1986). A detailed understanding of the thermodynamics of spectrin self-association is essential to understanding how small changes in the thermodynamic criteria governing the self-association in vitro may translate to large morphological changes in vivo.

Several studies (Morrow & Marchesi, 1981; Morrow et al., 1981; Shahbakhti & Gratzer, 1986) using nondenaturing gel electrophoresis have shown that the self-association of spectrin is probably of an indefinite type. These studies have indicated that the tetramer predominates in solution and that higher oligomers form only a small fraction, by weight, of the total spectrin at low solute concentrations. On the other hand, sedimentation equilibrium studies have shown that all steps in the self-association are fully reversible but that the proportion of larger oligomers was consistent with indefinite self-association models in which the values of the sequential equilibrium constants were not all identical but were all of the same magnitude ($\sim 10^6$ M^{-1}) (Morris & Ralston, 1984, 1985a).

In attempting to assess quantitatively the thermodynamic properties of a self-association reaction, it is of the utmost importance that (1) highly precise and accurate data be obtained at chemical equilibrium without perturbing the equilibrium and (2) an accurate and sensitive test be used to determine if chemical equilibrium has been achieved and to assess if there is any significant amount of protein incapable of participating in the self-association. The technique of sedimentation equilibrium meets both of these strict criteria (Teller, 1973; Milthorpe et al., 1975).

In the present paper, we confirm with the use of sedimentation equilibrium that the weight fraction of tetramer in vitro is much lower than that found in studies by other workers (Morrow & Marchesi, 1981; Morrow et al., 1981; Shahbakhti

† This work was supported by the Australian Research Council and by a Commonwealth postgraduate award to M.M.
* To whom correspondence should be addressed.

0006-2960/89/0428-8561$01.50/0