Biochemistry

© Copyright 1984 by the American Chemical Society Volume 23, Number 14 July 3, 1984

Accelerated Publications

Exploring the Conformational Roles of Signal Sequences: Synthesis and Conformational Analysis of λ Receptor Protein Wild-Type and Mutant Signal Peptides†

Martha S. Briggs and Lila M. Gierasch*

ABSTRACT: Secretion of the *Escherichia coli* λ receptor protein (LamB protein) appears from genetic evidence to be correlated with the predicted tendency of its signal sequence to adopt an α-helical conformation [Emr, S. D., & Silhavy, T. J. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4599]. We have tested this hypothesis by synthesizing major portions of signal sequences from the wild-type and mutant LamB proteins and analyzing their conformations by circular dichroism. The wild-type signal sequence contains a seven-residue hydrophobic region flanked by a proline and a glycine. Chou–Fasman rules predict that this segment will adopt an α-helical conformation. An export-deficient mutant is missing four residues from this region; the helix-breaking glycine and proline are thus separated by only three residues, and an α helix is not predicted to form. In each of the export-restored revertants, either the glycine or the proline is replaced with a residue which promotes helix formation. The helix content of the synthetic signal sequence fragments on the basis of CD measurements supports the secondary structure hypothesis described above. The relative helicity in aqueous sodium dodecyl sulfate, lysolecithin, or trifluoroethanol is as follows: wild type > R2(Pro→Leu) > R1(Gly→Cys) >> deletion mutant.

All but one (Palmiter et al., 1978) of the known secreted proteins are initially synthesized as precursors possessing an amino-terminal extension, called a signal sequence, which is removed during or after secretion by a specific peptidase (Blobel, 1977; Wickner, 1979; Kreil, 1981; Silhavy et al., 1983). Although the exact amino acid sequences of the signal peptides are seldom conserved, they have some general features in common (Inouye & Halegoua, 1980), the most notable being the presence of a large proportion of hydrophobic amino acids clustered in a core region of about 10–15 amino acids. Chou–Fasman secondary structure predictions (Chou & Fasman, 1974a,b) often indicate that these segments will adopt an α-helical conformation. Changes in the length, polarity, or net charge of the signal sequence (Inouye et al., 1982; Bedouelle et al., 1980; Emr & Silhavy, 1980; 1983; Emr et al., 1980; Hortin & Boime, 1980) may result in precursor proteins which are synthesized normally but which are translocated and processed slowly or not at all. These "preproteins" accumulate within the cell and can be isolated and sequenced. Thus, although a functional signal sequence is necessary for protein secretion, its role in the process is unclear.

Many hypotheses concerning the mechanism of protein secretion have been proposed. The theory which is supported by the most evidence in eukaryotes is the "signal hypothesis" of Blobel and Dobberstein (Blobel & Dobberstein, 1975a,b; Blobel, 1977), which postulates the existence of a membrane-associated protein or complex that recognizes the signal sequence of the nascent polypeptide, binds to it and to the polysome, and forms a channel through which the growing protein is extruded. Other translocation hypotheses propose that insertion and secretion occur spontaneously, due to the thermodynamics of protein folding, and that no specific transport apparatus is necessary (Wickner, 1979; Inouye & Halegoua, 1980; Engelman & Steitz, 1981; von Heijne & Blomberg, 1979). There are many variants of this idea, the best known being Wickner's "membrane trigger hypothesis" (Wickner, 1979), but all hold that the function of the signal sequence is to influence the conformation of the nascent polypeptide or completed preprotein so that it will partition into the membrane. In either case, the signal sequence must interact with the membrane, and thus its conformation in the presence of the lipid bilayer must play an important role in protein secretion.

Emr & Silhavy (1983) have isolated a strain of *Escherichia*

† From the Department of Chemistry, University of Delaware, Newark, Delaware 19716 (L.M.G. and M.S.B.), and the Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06511 (M.S.B.). *Received March 30, 1984.* This work was supported by Grants GM27616 and GM29754 from the U.S. Public Health Service.

0006-2960/84/0423-3111$01.50/0 © 1984 American Chemical Society

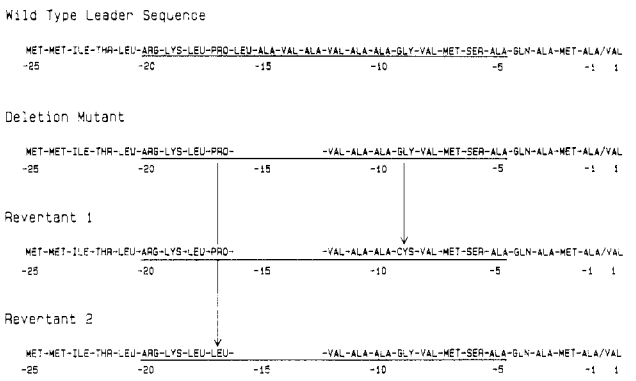

FIGURE 1: Amino acid sequences of the λ receptor protein signal sequences. A slash indicates the site of cleavage of the signal sequence from the mature protein by signal peptidase. The signal sequence is numbered from −1 at its carboxyl end, just before the cleavage site, to −25 at its amino terminus. Arrows indicate the point mutations mentioned in the text. The underlined segments are the portions that we have synthesized.

coli with a 12 base pair deletion (corresponding to 4 amino acids) in the region of the gene coding for the signal sequence of the λ receptor (LamB)[1] protein. This mutant synthesizes the LamB preprotein but is unable to export it to its normal location in the outer membrane. In addition, they found two pseudorevertants containing point mutations which lead to amino acid substitutions in the shortened signal sequence and which are able to export and process the LamB protein normally. The amino acid sequences of the wild-type and mutant LamB signal sequences are listed in Figure 1. The deletion mutant lacks the amino acid residues from the leucine located 16 residues before the cleavage site (position −16) to the alanine located 13 residues before the cleavage site (position −13). In one of the revertants, cysteine replaces the glycine at position −9, and in the other, leucine replaces the proline at position −17.

Emr & Silhavy (1983) report the result of Chou–Fasman calculations on all four of the signal peptide sequences. The rules predict that the wild-type signal sequence will adopt an α-helical conformation in the seven-residue hydrophobic region flanked by the helix-breaking glycine at position −9 and proline at position −17. The deletion mutant is missing four amino acids in this segment, making the helix nucleation region only three residues in length, which is shorter than the six-residue segment required by the calculation. The Chou–Fasman rules (Chou & Fasman, 1974a,b) predict no regular secondary structure in this region. The amino acid substitutions in each of the revertants lengthen the helix nucleation segment by changing one of the flanking helix-breaking residues to a helix-promoting residue, so each should adopt a helical conformation in this region. Thus, each revertant is predicted to regain the tendency to adopt a helical conformation in the central region of the signal. Emr and Silhavy proposed that it is this restoration of α-helical tendency in the revertants that accounts for their export functions and that the ability to form an α helix is essential for correct functioning of the signal sequences.

These predictions provide for the first time a testable hypothesis concerning the relationship between the secondary structure of the signal sequence and its biological activity. We have synthesized 12–16-residue portions of the wild-type and mutant LamB signal sequences and analyzed their conformations by CD both in bulk solvents and in micellar systems. We find that the wild-type peptide and two revertant signal peptides have a significantly greater tendency to adopt an α-helical conformation than does the deletion mutant.

Experimental Procedures

Materials

Side-chain-protected *N*-Boc-amino acids were purchased from Peninsula Laboratories; the purity was checked by melting point determination and TLC. Merrifield resin was purchased from Pierce Chemical Co. Solvents and reagents for solid-phase peptide synthesis were ACS reagent grade or spectroscopic grade. Solvents used for CD studies were spectroscopic grade. SDS was electrophoresis purity. L-α-Lysolecithin from egg yolk was obtained from Sigma Chemical Co. and contained primarily palmitic and stearic acids. Water was glass distilled and deionized.

Methods

Synthesis and Purification of Peptides. All of the peptides were prepared by using conventional solid-phase peptide synthesis techniques. Merrifield resin [(chloromethyl)polystyrene/1% divinylbenzene] was substituted with Boc-alanine using potassium fluoride in DMF at 50 °C overnight. The resin was defined, washed with methylene chloride, and dried in vacuo. Substitution levels were determined by a quantitative ninhydrin test (Sarin et al., 1981) and ranged from 0.43 to 0.81 mmol/g. The side-chain-protected Boc-amino acids were coupled as the hydroxybenzotriazole ester with DCC in DMF. The coupling reactions were monitored for completion by the ninhydrin test. Cleavage of the Boc groups was performed in 50% TFA in methylene chloride, followed by neutralization in 5% DIPEA in methylene chloride. The completed peptides were cleaved from the resin and deprotected in anhydrous HF (90%) in the presence of anisole (10%) for 1 h at 0 °C. The peptides were extracted from the resin with acetic acid/water (1:1), except for the wild-type peptide which was extracted with acetonitrile/acetic acid/water (2:1:1). The crude peptides were purified by semipreparative HPLC on a Hamilton PRP-1 column using 0.1% TFA in a water/acetonitrile gradient as eluent at a flow rate of 1.5 mL/min. Amino acid analyses were satisfactory. The amino acid sequences were confirmed by sequential Edman degradation.

CD Spectra. CD spectra were obtained on an upgraded Landis Associates Jasco CDSP spectropolarimeter in a 1-mm cell, at ambient temperature. See figure legends for sample concentrations. CD band intensities are expressed as the mean residue ellipticity, $[\theta]$ (in units of degrees centimeter squared per decimole).

Results

We have synthesized the underlined portions of the signal sequences listed in Figure 1. The wild-type segment is 16 residues long, extending from position −20 to position −5. The deletion mutant and the two revertant segments are each 12 residues long, as they are missing residues −16 to −13. The peptides consist of the proposed helical region of the hydrophobic core and its flanking proline and glycine. In addition, several residues on each end of this region are present to provide enough length for possible helix nucleation in the revertant peptides. All of the sequences include the basic residues from near the amino terminus, the −20 arginine and the −19 lysine, which may serve to position that end of the peptide near the negatively charged head groups of lipidlike environments (SDS or lysolecithin in micelles) and to enhance

[1] Abbreviations: LamB protein, λ receptor protein; CD, circular dichroism; Boc, *tert*-butyloxycarbonyl; TLC, thin-layer chromatography; SDS, sodium dodecyl sulfate; DMF, dimethylformamide; DCC, dicyclohexylcarbodiimide; TFA, trifluoroacetic acid; DIPEA, diisopropylethylamine; HPLC, high-pressure liquid chromatography; Me_2SO, dimethyl sulfoxide; TFE, trifluoroethanol.

Table I: Circular Dichroism Data for Signal Peptides[a,b]

solvent	peptide	CD extrema λ (nm)	[θ] (×10⁻³)	λ (nm)	[θ] (×10⁻³)	λ (nm)	[θ] (×10⁻³)
H_2O	WT[c]	222	−1.20	202	−13.87		
	DM	222	−0.69	206	−13.60	194	5.90
	R1	222	−1.51	202	−17.20	190	7.04
	R2	223	−2.03	204	−13.06		
40 mM SDS	WT	222	−7.63	207	−10.27	192	16.43
	DM	222	−2.42	202	−10.61		
	R1	226	−4.69	204	−10.84	193	7.61
	R2	219	−5.94	206	−13.36	190	9.54
25 mM lysolecithin	WT	222	−6.78	212	−9.04	193	6.77
	DM	222	−5.90	206	−15.36	191	6.73
	R1	222	−6.46	206	−13.44		
	R2	222	−8.00	206	−16.54		
20% TFE	WT	222	−5.37	206	−7.36	188	9.36
	DM	221	−2.07	204	−8.25	192	8.53
	R1	222	−3.90	202	−9.64		
	R2	222	−5.20	206	−8.87	188	9.18

[a] [θ] given in degrees centimeter squared per decimole. [b] Where a minimum occurred near 222 nm, its wavelength is noted. Where there was no minimum, the ellipticity at 222 nm is given. [c] Abbreviations: WT, wild-type signal peptide; DM, deletion mutant signal peptide; R1, revertant containing the Gly to Cys mutation; R2, revertant containing the Pro to Leu mutation.

the solubilities of the peptides in water.

All four of the peptides are soluble in water at fairly high concentration. The wild-type peptide is less water soluble than the other three but will dissolve at about 0.25 mM. The other peptides were soluble at about 2 mM or more. The wild-type and mutant-derived peptides also had different solubilities in organic solvents. The wild-type peptide was soluble in acetonitrile/water at all ratios of the two solvents; the other peptides were soluble only in mixtures containing less than 50% acetonitrile. All of the peptides are soluble in Me_2SO at about 1–2 mM.

CD spectra of the four peptides were obtained in two bulk solvents, water and aqueous TFE, and in micellar solutions of SDS and lysolecithin. Water and TFE were chosen to probe the peptides' response to varying bulk solvent polarity. An additional reason for our choice of TFE as a solvent is its helix-inducing property (Zahler & Niggli, 1977). Due to their amphiphilicity and anisotropic character, SDS micelles are more similar to a membrane bilayer than bulk solvents and have been used as a membrane-mimetic environment (Gierasch et al., 1982, 1983). The peptides were also studied in micelles of lysolecithin, which is more similar in structure to the membrane phospholipids than is SDS. CD extrema in bulk solvents and micellar media are listed in Table I.

The CD spectra were analyzed for contributions from the secondary structure by visual comparison of experimental data to the reference spectra of Greenfield & Fasman (1969). The fits of reference spectra to the experimental curves were good at wavelengths greater than 200 nm; below this value, the fits were generally poor. In addition, the ellipticity at 222 nm was used as a rough measure of the relative helicity. The same trends were observed in both cases. See Table I for the values of [θ] at 222 nm.

The CD spectra of the peptides in water and in aqueous SDS above the critical micelle concentration are shown in Figure 2. In water, all of the peptides have similar CD spectra, with a minimum at 202–206 nm, and little intensity at 222 nm. Comparison to reference spectra indicates that the peptides adopt a largely "random" conformation, with 4–8% α helix and 12–16% β sheet.

In 40 mM SDS, all of the peptides show increased intensity at 222 nm, and the wild-type and revertant peptides begin to have the double minimum characteristic of the α helix. Comparison to reference spectra indicates that the wild-type

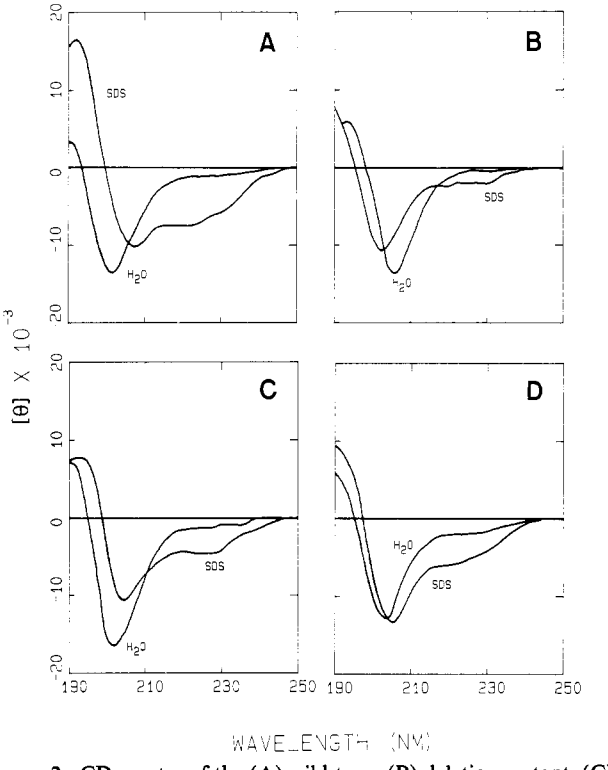

FIGURE 2: CD spectra of the (A) wild-type, (B) deletion mutant, (C) revertant 1, and (D) revertant 2 signal peptides in water and 40 mM aqueous SDS. The peptide concentrations were approximately 0.2 mg/mL. [θ] is expressed as degrees centimeter squared per decimole per residue.

and revertant peptides contain about 25% α structure, 15% β structure, and 60% aperiodic structure. The deletion mutant signal peptide has little change in structure between SDS and water. The CD spectra of the peptides in 20% TFE and 25 mM lysolecithin show similar trends to those in SDS.

Discussion

Our results support the proposal of Emr & Silhavy (1983) by showing that the nonfunctional deletion mutant signal peptide has much less tendency to form an α helix than do the functional wild-type or revertant peptides. Our conformational analysis shows that indeed the presence of proline and glycine separated by only three residues in the deletion mutant disrupts the helix-forming potential of the signal hy-

drophobic core. The revertants, in which one of these residues is replaced by a helix-promoting residue while maintaining the same length as the deletion mutant signal peptide, adopt a helical conformation in aqueous TFE or micellar solutions.

The site of the signal sequence–membrane interaction is uncertain. Signal peptides in eukaryotes interact with membrane-bound protein (Walter & Blobel, 1981); in in vitro prokaryotic systems, signal peptides interact with lipid vesicles in the absence of protein (Watts et al., 1981; Ohno-Iwashita & Wickner, 1983; Y. Ohno-Iwashita et al., personal communication). In vivo, it is indeed possible that a protein transport system exists in prokaryotes and that these signal sequences, too, interact with protein. In either case, however, the site of interaction will be highly hydrophobic. Finkelstein et al. (1983) have calculated that the signal peptide would probably adopt an α-helical conformation in a proteinaceous receptor site. Despite these uncertainties, Emr and Silhavy's proposal and the conclusion of this determination are not dependent on a particular model for protein secretion. The genetic results correlate helix-forming tendency with function. The conformational results characterize this tendency experimentally and give support to the Emr–Silhavy model. Moreover, it is very important to recognize that our results and those of Emr and Silhavy lead to a view that the intrinsic nature and conformation of the signal in the absence of the mature protein can be related to function. Thus, in this case, it is not necessary to invoke the influence of the rest of the protein to explain the secretion competency (or lack of it) of these mutants.

While there are intrinsic limitations in the approach described here, we find the results compelling. First, it is difficult to determine quantitatively the amount of periodic structure in such small peptides. The reference spectra (Greenfield & Fasman, 1969) which we used in this work are derived from spectra of polymers of lysine which are larger than our peptides. Other reference spectra are statistically derived from CD spectra of proteins of known structure (Provencher & Glockner, 1981; Chen et al., 1974). Small peptides are likely to be more flexible than proteins or large polypeptides. Thus, while CD studies of small peptides can indicate relative amounts of secondary structure and changes in conformation, the numerical values for structure content should be considered an approximation. Nonetheless, helix content is perhaps the one measurement that can be obtained with confidence. The reasonable percentages observed and the consistency of results in three helix-promoting environments argue for their validity. Also, we have not examined the entire signal peptides. Still, the expected trends in secondary structure were found. Extra residues were intentionally incorporated beyond the hydrophobic core. Syntheses of the entire signal sequences are in progress to check the correlations.

Despite great interest in the role of signal sequences in protein secretion, there have been few conformational studies involving signal sequences. Previous studies of signal sequences have determined the conformations of the M13 coat protein signal sequence (Shinnar & Kaiser, 1984) or the pre-proparathyroid signal sequence (Rosenblatt et al., 1980), both synthesized by solid-phase methods, in aqueous buffers and in the structure-promoting solvent hexafluoroisopropyl alcohol. The present study marks the first time conformation and function have been related in signal sequences; it is evident that the conformational preferences of these intriguing polypeptide sequences can be related to their functional roles.

Acknowledgments

We thank W. F. DeGrado and D. M. Engelman for helpful discussions during the course of this work, L. Parente and L. Gormley for the amino acid analyses, and R. Kutny for the peptide sequence analyses.

Registry No. Arg-Lys-Leu-Pro-Leu-Ala-Val-Ala-Val-Ala-Ala-Gly-Val-Met-Ser-Ala, 90220-08-9; Arg-Lys-Leu-Pro-Val-Ala-Ala-Gly-Val-Met-Ser-Ala, 90220-09-0; Arg-Lys-Leu-Pro-Val-Ala-Ala-Cys-Val-Met-Ser-Ala, 90220-10-3; Arg-Lys-Leu-Leu-Val-Ala-Ala-Gly-Val-Met-Ser-Ala, 90220-11-4; SDS, 151-21-3.

References

Bedouelle, H., Bassford, P. J., Fowler, A. V., Zabin, I., Beckwith, J., & Hofnung, M. (1980) *Nature (London) 285*, 78.

Blobel, G. (1977) in *International Cell Biology* (Brinkley, B. R., & Porter, K. R., Eds.) p 318, Rockefeller University Press, New York.

Blobel, G., & Dobberstein, B. (1975a) *J. Cell Biol. 67*, 835.

Blobel, G., & Dobberstein, B. (1975b) *J. Cell Biol. 67*, 852.

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

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

Chou, P. Y., & Fasman, G. D. (1974b) *Biochemistry 13*, 222.

Emr, S. D., & Silhavy, T. J. (1980) *J. Mol. Biol. 141*, 63.

Emr, S. D., & Silhavy, T. J. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4599.

Emr, S. D., Hedgpeth, J., Clement, J.-M., Silhavy, T. J., & Hofnung, M. (1980) *Nature (London) 285*, 82.

Engelman, D. M., & Steitz, T. A. (1981) *Cell (Cambridge, Mass.) 23*, 411.

Finkelstein, A. V., Bendzko, P., & Rapoport, T. A. (1983) *FEBS Lett. 161*, 176.

Gierasch, L. M., Lacy, J. E., Thompson, K. F., Rockwell, A. L., & Watnick, P. (1982) *Biophys. J. 37*, 275.

Gierasch, L. M., Lacy, J. E., Anderle, G., Lalancette, R., & Mendelsohn, R. (1983) *Biopolymers 22*, 381.

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

Hortin, G., & Boime, I. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 1356.

Inouye, M., & Halegoua, S. (1980) *CRC Crit. Rev. Biochem. 7*, 339.

Inouye, S., Soberon, X., Franceschini, T., Nakamura, K., Itakura, K., & Inouye, M. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 3438.

Kreil, G. (1981) *Annu. Rev. Biochem. 50*, 317.

Ohno-Iwashita, Y., & Wickner, W. (1983) *J. Biol. Chem. 258*, 1895.

Palmiter, R. D., Gagnon, J., & Walsh, K. A. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 94.

Provencher, S. W., & Glockner, J. (1981) *Biochemistry 20*, 33.

Rosenblatt, M., Beaudette, N. V., & Fasman, G. D. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 3983.

Sarin, V. K., Kent, S. B. H., Tam, J. P., & Merrifield, R. B. (1981) *Anal. Biochem. 117*, 147.

Shinnar, A., & Kaiser, E. T. (1984) *J. Am. Chem. Soc.* (in press).

Silhavy, T. J., Benson, S. A., & Emr, S. D. (1983) *Microbiol. Rev. 47*, 313.

von Heijne, G., & Blomberg, C. (1979) *Eur. J. Biochem. 97*, 175.

Walter, P., & Blobel, G. (1981) *J. Cell Biol. 91*, 551.

Watts, C., Silver, P., & Wickner, W. (1981) *Cell (Cambridge, Mass.) 25*, 347.

Wickner, W. (1979) *Annu. Rev. Biochem. 48*, 23.

Zahler, P., & Niggli, V. (1977) *Methods Membr. Biol. 8*, 1.