Conformational Requirement of Signal Sequences Functioning in Yeast: Circular Dichroism and ^1H Nuclear Magnetic Resonance Studies of Synthetic Peptides

Yoshio Yamamoto,‡ Tadayasu Ohkubo, Atuko Kohara, Toshiki Tanaka, Toshiaki Tanaka, and Masakazu Kikuchi*
Protein Engineering Research Institute, 2-3, Furuedai 6-chome, Suita, Osaka 565, Japan
Received October 25, 1989; Revised Manuscript Received May 31, 1990

ABSTRACT: Recently, we have designed a series of simplified artificial signal sequences and have shown that a proline residue in the signal sequence plays an important role in the secretion of human lysozyme in yeast, presumably by altering the conformation of the signal sequence [Yamamoto, Y., Taniyama, Y., & Kikuchi, M. (1989) *Biochemistry 28*, 2728–2732]. To elucidate the conformational requirement of the signal sequence in more detail, functional and nonfunctional signal sequences connected to the N-terminal five residues of mature human lysozyme were chemically synthesized and their conformations in a lipophilic environment [aqueous trifluoroethanol (TFE) or sodium dodecyl sulfate micelles] analyzed by circular dichroism (CD) and ^1H nuclear magnetic resonance (NMR) spectroscopy. The helix content of the peptides, including functional (L8, CL10) and nonfunctional (L8PL, L8PG, L8PL2) signal sequences, was estimated from CD spectra to be 40–50% and 60–70%, respectively, indicating that the helical structure is more abundant in the nonfunctional signal sequences. Two-dimensional NMR analyses in 50% TFE/H_2O revealed that each peptide adopted a helical conformation throughout the sequence except for a few residues at the N- and C-termini. Furthermore, H–D exchange experiments indicated that the helical structure of the C-terminal region of the functional signal sequences (L8 and CL10) was less stable than that of the nonfunctional signal sequences (L8PL and L8PL2). On the basis of these results, a model was developed in which the functional signal sequence is inserted in the membrane with a helical conformation and the C-terminal helix unraveled in an extended conformational form through an interaction with the signal peptidase.

While the roles of signal sequences involved in protein secretion have been studied intensively (Briggs & Gierasch, 1986; Gierasch, 1989), the detailed molecular mechanism of their function is still poorly understood. Although the signal sequences have a common function, they are composed of a variety of amino acids, and their primary structures are rarely conserved. This implies that, in general, overall characteristics such as hydrophobicity and conformation are important factors for their function. The conformational roles of the signal sequences have been studied genetically (Emr & Silhavy, 1983) and physicochemically (Briggs & Gierasch, 1984). These studies have suggested that the hydrophobic segment adopts an α-helical conformation in the membrane and that the helix-forming potential is well correlated with their function. However, few systematic studies have been performed to determine the conformation–function relationship of the C-terminal region of the signal sequence.

More recently, we have designed a series of simplified artificial signal sequences and found that a proline residue at position −4 to −6 in the signal sequence plays an important role in the secretion of human lysozyme in yeast (Yamamoto et al., 1989). This study suggested the following points regarding the conformation of signal sequences. (1) The functional signal sequence L8 = Met-Arg-$(Leu)_8$-Pro-Leu-Ala-Ala-Leu-Gly adopts a conformation composed of an α-helix in the hydrophobic core [$(Leu)_8$] with an extended segment following the proline residue. (2) The nonfunctional signal sequence L8PL, wherein the proline residue in L8 is replaced with a leucine, assumes an α-helical conformation throughout the sequence. (3) The slightly functional signal sequence L8PG, in which the proline residue in L8 is replaced by a glycine, adopts a conformation similar to that of L8PL, but the helical structure is less stable than that of L8 due to the presence of the glycine residue. (4) The slightly functional signal sequence L8PL2, in which two leucines are inserted after the proline residue, adopts a conformation with two helices connected by the proline. It was therefore proposed that these conformational differences in the C-terminal region might affect the function of the signal sequences.

Our present approach was to examine the conformational requirement of the signal sequences in more detail by analyzing CD[1] and ^1H NMR spectra of chemically synthesized peptides in apolar environments (aqueous TFE or SDS micelles). These peptides included the functional and nonfunctional signal sequences described above, fused to the N-terminal five residues of mature human lysozyme (M5). In addition, the conformation of the functional signal sequence CL10 fused to M5, derived from the chicken lysozyme signal sequence by replacement of the cysteine at position −9 with a leucine (Yamamoto et al., 1987), was also examined in order to compare the conformations of the hydrophobic segments with different primary structures. On the basis of these results, a model is now proposed describing the conformational change of the signal sequence in the membrane.

* To whom correspondence should be addressed.
‡ Present address: Chemistry Research Laboratories, Takeda Chemical Industries, Ltd., 17-85, Juso-Honmachi 2-chome, Yodogawa-ku, Osaka 532, Japan.

[1] Abbreviations: CD, circular dichroism; 1D, one dimensional; 2D, two dimensional; NMR, nuclear magnetic resonance; TFE, trifluoroethanol; TFE-d_3, fully deuterated TFE; SDS, sodium dodecyl sulfate; DQF, double quantum filtered; COSY, correlation spectroscopy; TOCSY, total correlation spectroscopy; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; ROESY, rotating-frame nuclear Overhauser enhancement spectroscopy; M5, N-terminal five residues of mature human lysozyme; DSS, 2,2-dimethyl-2-silapentane-5-sulfonate.

0006-2960/90/0429-8998$02.50/0

Table I: Peptides Studied

	signal sequence-hLZM[a]	secretion (mg/L)[b]
L8-M5	MRLLLLLLLLPLAALG-KVFER	5
L8PL-M5	MRLLLLLLLLLLAALG-KVFER	<0.1
L8PG-M5	MRLLLLLLLLGLAALG-KVFER	0.4
L8PL2-M5	MRLLLLLLLLPLLLAALG-KVFER	0.1
CL10-M5	MRSLLILVLLFLPLAALG-KVFER	4

[a] hLZM, human lysozyme. [b] Data from Yamamoto et al. (1987, 1989).

Materials and Methods

Materials. TFE and TFE-d_3 were purchased from Wako Pure Chemical Industries and Cambridge Research Laboratories, respectively. Reagents for peptide synthesis were obtained from Applied Biosystems. All other reagents were of analytical grade.

Synthesis and Purification of Peptides. Peptides were prepared by solid-phase synthesis on an Applied Biosystems Model 430A automated peptide synthesizer. (Phenylacetamido)methyl-resin and *tert*-butoxycarbonyl amino acids were used for synthesis with a coupling mixture composed of *N*-methylpyrrolidone and dimethyl sulfoxide. For the segments consisting of consecutive leucine residues, a double coupling cycle was applied to ensure completion of the reaction. Synthesized peptides were cleaved from the resin by trifluoromethanesulfonic acid in the presence of trifluoroacetic acid and thioanisole, extracted with 50% acetic acid/water, and lyophilized. The crude peptides were purified by preparative reversed-phase HPLC on an ODS-H-4251 column (Senshu Kagaku) with 0.1% trifluoroacetic acid in a water/acetonitrile gradient as eluent at a flow rate of 3 mL/min. The amino acid sequences were confirmed by automated Edman degradation on an Applied Biosystems Model 477A peptide sequencer. The sequences of the peptides synthesized are listed in Table I.

Measurement of CD Spectra. CD spectra were recorded at 25 °C from 250 to 195 nm on a J-600 spectropolarimeter (Japan Spectroscopic) with a 2-mm path-length cuvette. A 0.05-mm path-length cuvette was used for samples at high concentrations. Each peptide was dissolved in 20 or 50% (v/v) TFE/H_2O or in 1% aqueous SDS, at a concentration of 0.2 mg/mL (about 0.07 mM) and a pH meter reading (pH*) of 4.7. Exact peptide concentrations were determined by quantitative amino acid analysis. All the CD spectra are presented as the mean residue ellipticity $[\theta]$ (degrees centimeter squared per decimole).

Measurement of NMR Spectra. For measurement of NMR spectra, each peptide was dissolved in 50% (v/v) TFE-d_3/H_2O to a concentration of 3 mM with pH* of 3.1. ^1H NMR spectra were recorded at 30 °C on a Bruker AM500 or AM600 spectrometer equipped with digital phase shifters and an ASPECT 3000 computer. The water resonance was suppressed by selective irradiation during the relaxation delay and, in the case of NOESY, ROESY, and TOCSY spectra, during the mixing time as well. Phase-sensitive double quantum filtered (DQF) COSY spectra (Aue et al., 1976; Rance et al., 1983) and TOCSY spectra (Braunschweiler & Ernst, 1983; Davis & Bax, 1985) with a mixing time of 40 ms were recorded to identify amino acid spin systems. NOESY (Jeener et al., 1979) and ROESY (Bax & Davis, 1985) spectra were recorded with mixing times of 300 and 220 ms, respectively. In the case of NOESY spectra, a 5% random variation in the mixing time was used to eliminate zero quantum coherence transfer. Typically, 64–128 transients composed of 4096 data points were collected for each of 512 increments in t_1 with a relaxation delay of 2 s between successive transients. All the data were multiplied by a shifted sine bell, zero filled to 4096 × 4096 data points, and Fourier transformed. Chemical shifts are expressed in ppm relative to 2,2-dimethyl-2-silapentane-5-sulfonate (DSS).

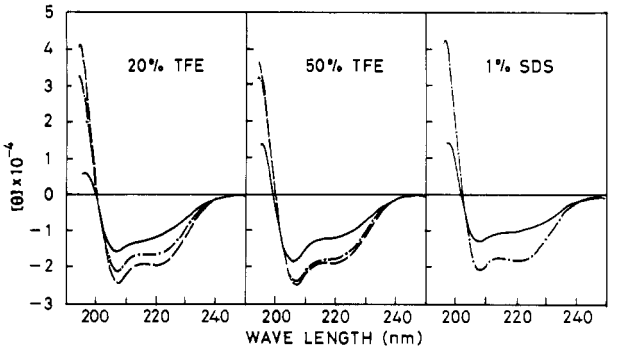

FIGURE 1: CD spectra of L8-M5 (—), L8PL-M5 (– –), and L8PG-M5 (–·–) in 20% TFE (left), 50% TFE (middle), and 1% SDS micelles (right). $[\theta]$ is the mean residue ellipticity in units of degrees centimeter squared per decimole.

Table II: Number of Helical Residues Estimated from CD Spectra

		no. of helical residues[a]		
peptide	no. of residues	20% TFE	50% TFE	1% SDS
L8-M5	21	8.2	9.7	6.1
L8PL-M5	21	14.5	14.3	ND[b]
L8PG-M5	21	12.0	13.7	12.0
L8PL2-M5	23	13.6	13.8	ND
CL10-M5	23	3.2	10.1	8.7

[a] Estimated by the method of Greenfield and Fasman (1969). [b] ND, not determined.

Results

Solubility of Peptides. All the peptides were practically insoluble in water and 10% TFE but readily soluble (>3 mM) in 20–100% TFE. L8-M5, CL10-M5, and L8PG-M5 were soluble in 1% SDS micelles at a concentration of 0.07 mM, but L8PL-M5 and L8PL2-M5 were significantly less soluble.

Measurement and Analysis of CD Spectra. The CD spectra of the five peptides (Table I) were measured in 20% TFE, 50% TFE, and 1% SDS micelles. However, the spectra of L8PL-M5 and L8PL2-M5 in 1% SDS micelles could not be obtained due to their insolubility. Two different TFE concentrations (20 and 50%) were used to examine their effect on the formation of α-helix, as TFE is known to promote helical conformation (Nelson & Kallenbach, 1986). Helix content was estimated by the method of Greenfield and Fasman (1969), which is based on the spectra of poly(L-lysine). Another method by Provencher and Glöckner (1981) gave similar results (data not shown).

The CD spectrum of L8-M5 in 20% TFE showed a minimum at 207 nm, with a shoulder around 222 nm (Figure 1), indicative of the presence of α-helical structure. Upon increasing the TFE concentration to 50%, the minimum at 222 nm became conspicuous (Figure 1), suggesting that the helical conformation is increased in 50% TFE. Secondary structure analysis revealed that about two residues were converted to a helical conformation upon increasing the TFE concentration from 20 to 50% (Table II). Further increase in the TFE concentration up to 100% did not change the spectra significantly (data not shown). Only a slight change in the spectra was observed in the pH* range of 2–9, indicating that charge effects on the conformation are negligible. Moreover, CD spectra of L8-M5 (and the other four peptides) in 50% TFE

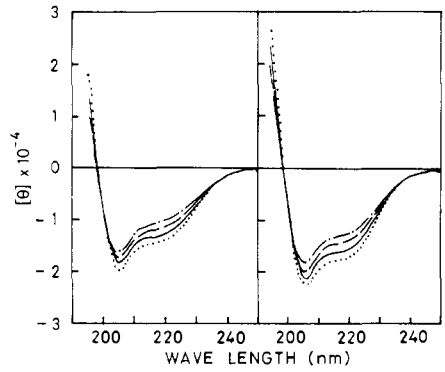

FIGURE 2: CD spectra of L8-M5 (left) and L8PL2-M5 (right) in 50% TFE as a function of temperature. Temperatures are 5 (···), 25 (—), 45 (– –), and 65 °C (–·–).

were nearly invariant over a range of peptide concentration from 70 μM to 3 mM (equal to the NMR condition), suggesting that the peptides are not aggregated under these conditions. In 1% SDS micelles, the CD spectrum was similar to that in aqueous TFE (Figure 1). The spectrum in 2% lysolecithin micelles, they being more closely related to membrane phospholipids, was also similar (data not shown).

CD spectra of L8-M5 in 50% TFE were also measured at 5, 45, and 65 °C as shown in Figure 2. They were similar in shape to the spectrum at 25 °C, but the ellipticity gradually decreased as temperature increased. Helix content at 5, 25, 45, and 65 °C was estimated to be 51, 46, 43, and 40%, respectively. This result suggests that the peptide is in dynamic equilibrium between a helical conformation and a more random structure and that this equilibrium is shifted toward the more random structure at higher temperature. Therefore, the helix content described in the text and Table II should be regarded as a value averaged over time. Similar results were obtained for the other peptides, indicating that they are also in conformational equilibrium. The temperature dependence of the CD spectra of L8PL2-M5, which has more helical structure than L8-M5 (see below), is also shown in Figure 2 for comparison.

The CD spectrum of L8PL-M5 in 20% TFE showed a more negative ellipticity than that of L8-M5, with two minima at 208 and 222 nm (Figure 1). The peptide contained six to seven more helical residues than L8-M5 (Table II). The spectra of L8PL-M5 in 50% TFE and 20% TFE did not differ significantly from each other. The spectrum of L8PG-M5 was intermediate between the spectra of L8-M5 and L8PL-M5 in 20% TFE and approached that of L8PL-M5 in 50% TFE, with one to two residues converted to a helical conformation. This result suggests that the glycine residue in L8PG-M5 is a much weaker helix breaker than the proline residue in L8-M5. In 1% SDS micelles, L8PG-M5 had secondary structures similar to those in aqueous TFE (Figure 1 and Table II).

The CD spectra of L8PL2-M5 were similar in 20 and 50% TFE (Figure 3). This peptide was estimated to have about five and four more helical residues than L8-M5 in 20% TFE and 50% TFE, respectively (Table II). These differences in the number of helical residues are significant, even considering that L8PL2-M5 has two more residues than L8-M5. Since the amino acid sequences of these two peptides differ in the region following the proline residue, the observed conformational difference can be assumed to reside in the C-terminal region of the signal sequences.

The CD spectrum of CL10-M5 in 20% TFE was significantly different from the other spectra (Figure 4) and looks rather similar to that of a typical spectrum of poly(L-lysine)

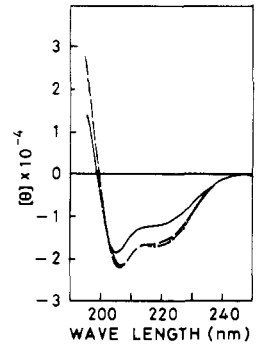

FIGURE 3: CD spectra of L8PL2-M5 in 20% TFE (–·–), L8PL2-M5 in 50% TFE (– –), and L8-M5 in 50% TFE (—).

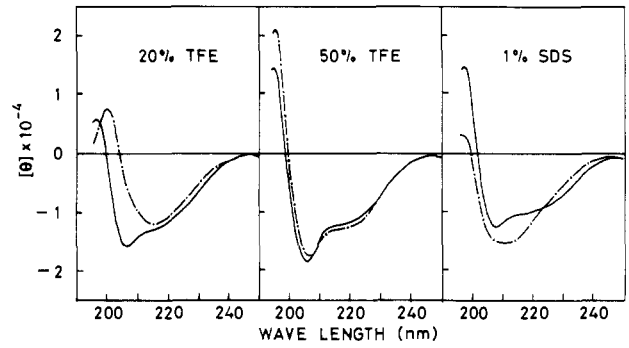

FIGURE 4: CD spectra of CL10-M5 (–·–) and L8-M5 (—) in 20% TFE (left), 50% TFE (middle), and 1% SDS micelles (right).

in a β-structure (Greenfield & Fasman, 1969). This conformational difference in CL10-M5 may be attributed to its weaker helix-forming potential in the hydrophobic segment, due to the presence of isoleucine, valine, and phenylalanine residues. In 50% TFE, the CD spectrum of CL10-M5 was similar to that of L8-M5, suggesting that formation of the α-helical conformation was promoted on increasing the TFE concentration. The spectrum of CL10-M5 in 1% SDS was somewhat different from that in 50% TFE, but the helix content was similar in both cases (Table II).

These results can be summarized as follows: (1) L8-M5 contains less helical conformation than L8PL-M5, due to the presence of the proline residue. (2) The glycine residue in L8PG-M5 can only partly substitute for the proline residue in destabilizing the helical conformation. (3) The insertion of two leucine residues after the proline residue in L8-M5 increases content of the helical conformation, probably in the C-terminal region of the signal sequence. (4) Idealization of the hydrophobic segment by arranging leucine residues alone actually increases its helix-forming potential. Results 1–3 are consistent with predictions previously reported on the conformation of signal sequences (Yamamoto et al., 1989). When these results are compared with the secretion data (Table I), an inverse correlation can be found between secretory capability and helix content of the peptides.

Measurement and Assignment of NMR Spectra. Although CD spectra can reveal overall characteristics of the secondary structure, no local information is available on the conformation of peptides. Therefore, analysis of 2D NMR spectra is essential to identify the secondary structure of each residue and to obtain a clearer picture of the conformation of the peptides. 2D NMR spectra of four of the five peptides (L8-M5, L8PL-M5, L8PL2-M5, and CL10-M5) were measured in 50% TFE/H_2O at a concentration of 3 mM. As the CD spectra at this concentration did not differ significantly from those shown in Figures 1–4, it is possible to compare these results with those obtained from 2D NMR. L8PG-M5 was not an-

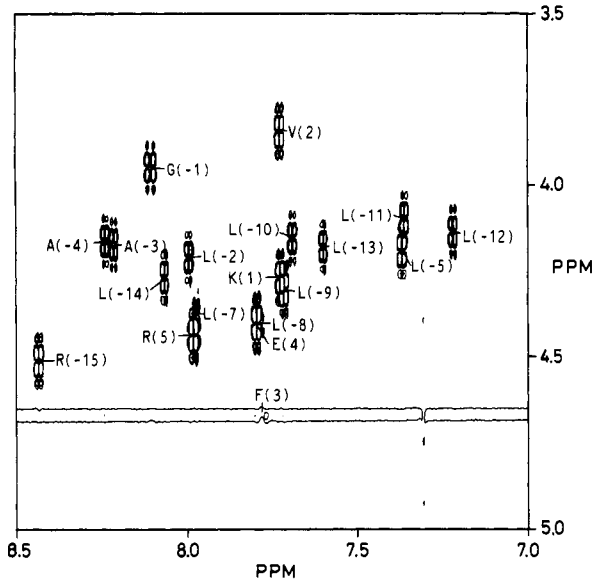

FIGURE 5: Fingerprint region of DQF-COSY spectrum of L8-M5 in 50% TFE-d_3/H_2O. The labels indicate backbone cross-peak assignments.

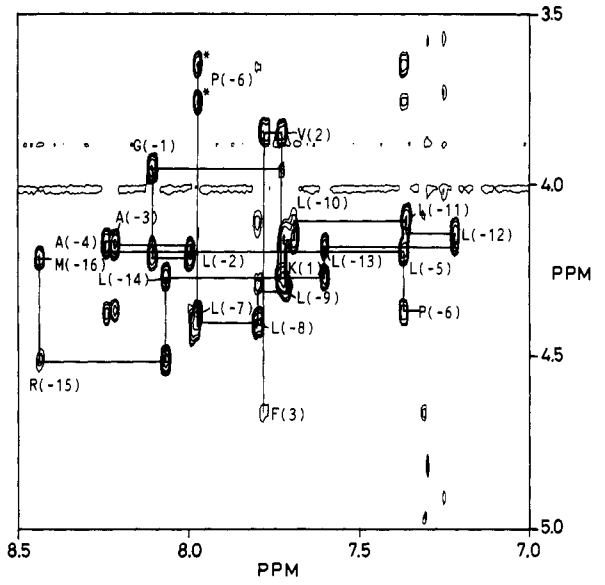

FIGURE 6: Fingerprint region of NOESY spectrum of L8-M5 in 50% TFE-d_3/H_2O with a mixing time of 300 ms. Sequential connectivities are indicated. The asterisks denote the cross-peaks between δCH of the proline and NH of the previous residue.

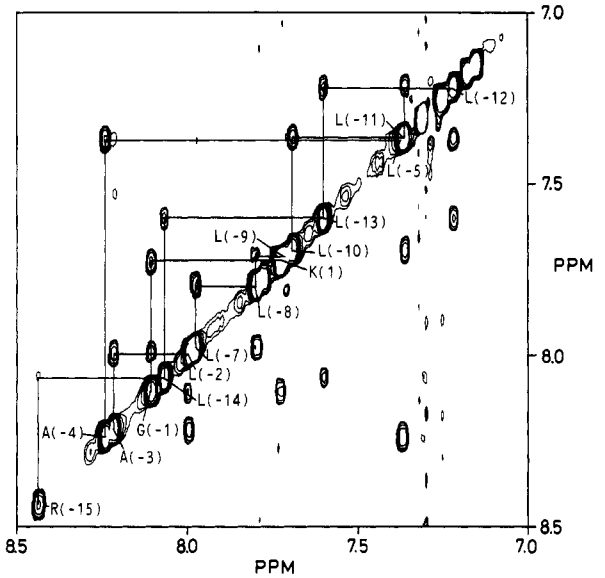

FIGURE 7: NH–NH region of NOESY spectrum of L8-M5 in 50% TFE-d_3/H_2O with a mixing time of 300 ms. Sequential connectivities are indicated.

alyzed by NMR, as it showed a CD spectrum similar to that of L8PL-M5 in 50% TFE (Figure 1) and was expected to adopt a conformation nearly identical with that of L8PL-M5.

First, 1D NMR spectra of these peptides were recorded. Each resonance peak was sharp enough to exclude the possibility of aggregation. In addition, the spectra contained no subpeaks that suggest the coexistence of two or more conformers. Thus, these peptides are assumed to adopt a single conformation or to be in rapid equilibrium between two or more conformers. However, the latter seems to be more reasonable, judging from the temperature dependence of the CD spectra mentioned above.

The strategy for carrying out the sequential assignments is as described by Wüthrich (1986). This involves identifying amino acid spin systems with DQF-COSY and TOCSY, followed by assignment of the spin systems to specific residues in the amino acid sequence by analysis of NOESY. We had expected that each of the leucine residues in the hydrophobic segment might be in a similar environment, resulting in degeneracy of the chemical shifts of these residues. Actually, the resonances of the main-chain protons were well separated in 2D spectra, as shown in the DQF-COSY spectrum of L8-M5 in the fingerprint region (Figure 5); 16 of the 20 expected cross-peaks were found separately. Two αCH of Gly(−1) were degenerated, and Leu(−8) and Glu(4) were overlapped. DQF-COSY spectra of the other peptides were also similar and well resolved (data not shown). Alanine, valine, isoleucine, phenylalanine, proline, and serine residues were easily identified from DQF-COSY spectra. Arginine and lysine residues were assigned by following connectivities along the side chain, both from αCH and from side-chain NH in TOCSY spectra (data not shown). Glutamic acid and methionine residues belong to the same spin system; however, the methionine residue, which is N-terminal, was identified by the absence of its NH resonance.

After the spin systems were assigned, sequential assignment was performed with NH(i)–NH(i+1) and αCH(i)–NH(i+1) connectivities in NOESY spectra, as exemplified in Figures 6 and 7. ROESY spectra of the peptides were nearly identical with the NOESY spectra, indicating that the observed NOE cross-peaks were not artifacts due to spin diffusion. At the proline residue, the loss of NH(i)–NH(i+1) connectivity is inevitable, but connectivities to δCH of the proline from NH of the previous residue were clearly seen in L8-M5, L8PL2-M5, and CL10-M5 (Figures 6 and 8). This also indicated that the proline residue was predominantly in the trans configuration. No evidence for a cis peptide bond was observed. Finally, all the resonances were assigned (Tables III–VI) except for those of γCH and δCH of the leucine residues, which were overlapped too much to assign individually.

Identification of Secondary Structures. The location of the secondary structure was deduced from the pattern of sequential and medium-range NOE connectivities observed in the NOESY spectra. It is generally assumed that a region with consecutive strong NH(i)–NH(i+1) and weak αCH(i)–NH(i+1) connectivities adopts a helical conformation (Wüthrich, 1986). Such regions were identified in both sides of the proline residue in L8-M5, L8PL2-M5, and CL10-M5 and over most of the sequence in L8PL-M5 (Figure 8). In addition, a number of medium-range connectivities, especially between αCH(i)

Table III: ^{1}H Chemical Shifts for L8-M5[a]

no.	residue	NH	αCH	βCH	γCH	others
−16	Met		4.21	2.23	2.62, 2.69	εCH 2.15
−15	Arg	8.44	4.51	1.85, 1.94	1.73	δCH 3.25; εNH 7.17
−14	Leu	8.07	4.27	1.68, 1.95		
−13	Leu	7.60	4.18	1.65		
−12	Leu	7.22	4.14	1.71		
−11	Leu	7.36	4.10	1.80		
−10	Leu	7.69	4.16	1.94		
−9	Leu	7.71	4.31	1.63, 1.93		
−8	Leu	7.80	4.40	1.88		
−7	Leu	7.97	4.38	1.66, 1.98		
−6	Pro		4.36	1.85, 2.33	1.98, 2.19	δCH 3.65, 3.75
−5	Leu	7.37	4.19	1.65, 1.96		
−4	Ala	8.24	4.16	1.53		
−3	Ala	8.21	4.17	1.51		
−2	Leu	7.98	4.21	1.68, 1.91		
−1	Gly	8.10	3.95			
1	Lys	7.72	4.27	1.96	1.54	δCH 1.74; εCH3.03; ζNH 7.58
2	Val	7.72	3.84	2.04	0.67, 0.92	
3	Phe	7.78	4.65	3.01, 3.33		aromatic CH 7.31
4	Glu	7.80	4.40	2.14, 2.23	2.55	
5	Arg	7.98	4.43	1.87, 2.01	1.73	δCH 3.26; εCH 7.17

[a] Chemical shifts are in ppm referenced to DSS.

Table IV: ^{1}H Chemical Shifts for L8PL-M5[a]

no.	residue	NH	αCH	βCH	γCH	others
−16	Met		4.21	2.24	2.61, 2.70	εCH 2.15
−15	Arg	8.41	4.53	1.86, 1.96	1.74	δCH 3.25; εNH 7.18
−14	Leu	8.13	4.25	1.70		
−13	Leu	7.62	4.17	1.67		
−12	Leu	7.15	4.18	1.79		
−11	Leu	7.34	4.13	1.70, 1.87		
−10	Leu	7.76	4.11	1.68, 1.84		
−9	Leu	7.86	4.17	1.83		
−8	Leu	8.20	4.10	2.03		
−7	Leu	8.27	4.14	1.58, 2.03		
−6	Leu	8.29	4.19	1.78, 1.91		
−5	Leu	8.70	4.10	1.50, 2.00		
−4	Ala	8.32	4.18	1.59		
−3	Ala	8.20	4.18	1.61		
−2	Leu	8.67	4.18	1.60, 1.93		
−1	Gly	8.20	3.95			
1	Lys	7.71	4.26	2.00	1.60	δCH 1.75; εCH 3.04; ζNH 7.58
2	Val	7.77	3.82	2.03	0.64, 0.93	
3	Phe	7.86	4.66	3.01, 3.33		aromatic CH 7.32
4	Glu	7.80	4.41	2.68, 2.74	2.56	
5	Arg	7.99	4.44	1.87, 2.01	1.74	δCH 3.26; εNH 7.15

[a] Chemical shifts are in ppm referenced to DSS.

and NH(i+3) and between αCH(i) and βCH(i+3), were observed in these regions (Figure 8). Although some of these connectivities were obscured by chemical shift degeneracy, other sequential and medium-range NOE connectivities support that these regions adopt helical conformations.

Furthermore, data from H–D exchange studies confirmed the location of the helical segments. We measured relative rates of chemical exchange of amide protons with D_2O. The peptide samples were freshly prepared in a mixture of TFE-d_3 and D_2O, and 1D NMR spectra were obtained as a function of time. As shown in Figure 8, amide protons in the helical regions were exchanged slowly and still observable after 3 h in 50% TFE/D_2O (see below for details). To obtain further data regarding the main-chain conformation, we attempted to measure spin–spin coupling constants, $^3J_{HN\alpha}$, from 1D NMR spectra. Although insufficient information was obtained due to overlap of the resonances, the $^3J_{HN\alpha}$ values determined were consistent with the above conformations (i.e., $^3J_{HN\alpha} < 5$ Hz for helical segments and $^3J_{HN\alpha} > 7$ Hz for the C-terminal region of the peptides). The helical regions identified from the NMR data are indicated at the bottom of Figure 8. Although it is not clear where the exact end of each helix is located, the helices are tentatively assumed to terminate at Gly(−1) in all the peptides.

These results show that each peptide adopts a helical conformation throughout the sequence except for a few residues at the N- and C-termini. Such a helical structure containing a proline residue has also been found in melittin (26 amino acid residues) in methanol (Bazzo et al., 1988) and melittin in dodecylphosphocholine micelles (Inagaki et al., 1989). These studies, in which the three-dimensional structure of melittin was determined by use of NOE connectivities and distance geometry calculation, have indicated that the helix is distorted in the vicinity of the proline residue with a bent angle of 120°–160°. The resemblance of the NOE connectivities between melittin and our peptides suggests that the latter also adopt a similar tertiary structure.

Helical segments identified on the basis of NOE connectivities (Figure 8) indicate that 13, 13, 15, and 15 helical residues are present in L8-M5, L8PL-M5, L8PL2–M5, and CL10-M5, respectively. The number of helical residues determined by NMR is consistent with that estimated by CD spectra (Table II) in L8PL-M5 and L8PL2-M5 but is overestimated in L8-M5 and CL10-M5. This discrepancy might

Table V: ^{1}H Chemical Shifts for L8PL2-M5[a]

no.	residue	NH	αCH	βCH	γCH	others
−18	Met		4.21	2.23	2.62, 2.69	εCH 2.15
−17	Arg	8.44	4.51	1.85, 1.95	1.73	δCH 3.25; εNH 7.18
−16	Leu	8.09	4.26	1.68		
−15	Leu	7.61	4.17	1.65, 1.95		
−14	Leu	7.19	4.14	1.71		
−13	Leu	7.36	4.09	1.77		
−12	Leu	7.71	4.13	1.86, 1.95		
−11	Leu	7.74	4.28	1.62, 1.95		
−10	Leu	7.85	4.44	1.67, 1.87		
−9	Leu	8.07	4.28	1.79, 1.95		
−8	Pro		4.25	1.81, 2.37	1.96, 2.21	δCH 3.64, 3.76
−7	Leu	7.30	4.22	1.75, 1.99		
−6	Leu	8.11	4.14	1.75, 1.86		
−5	Leu	8.63	4.10	1.46, 1.88		
−4	Ala	7.83	4.18	1.60		
−3	Ala	8.16	4.16	1.59		
−2	Leu	8.50	4.18	1.60, 1.92		
−1	Gly	8.14	3.94			
1	Lys	7.70	4.26	1.98, 2.01	1.58	δCH 1.75; εCH 3.04; ζNH 7.58
2	Val	7.76	3.82	2.04	0.65, 0.94	
3	Phe	7.85	4.66	3.01, 3.34		aromatic CH 7.32
4	Glu	7.80	4.40	2.17, 2.23	2.56	
5	Arg	7.99	4.44	1.87, 2.01	1.73	δCH 3.26; εNH 7.16

[a] Chemical shifts are in ppm referenced to DSS.

Table VI: ^{1}H Chemical Shifts for CL10-M5[a]

no.	residue	NH	αCH	βCH	γCH	others
−18	Met		4.20	2.24	2.66	εCH 2.15
−17	Arg	8.58	4.46	1.85	1.73	δCH 3.23; εNH 7.15
−16	Ser	7.98	4.50	3.84, 4.01		
−15	Leu	8.03	4.34	1.68, 1.73		
−14	Leu	7.68	4.11	1.66		
−13	Ile	7.33	3.88	1.96	0.99, 1.34, 1.60	δCH 0.93
−12	Leu	7.33	4.17	1.75, 1.93		
−11	Val	7.70	3.77	2.30	1.00, 1.08	
−10	Leu	7.98	4.17	1.93		
−9	Leu	7.94	4.27	1.35, 1.74	1.69	δCH 0.78, 0.86
−8	Phe	7.89	4.67	3.17, 3.33		aromatic CH 7.31
−7	Leu	8.21	4.41	1.74, 1.99		
−6	Pro		4.34	1.84, 2.32	1.97, 2.15	δCH 3.56, 3.79
−5	Leu	7.38	4.18	1.62, 1.91		
−4	Ala	8.14	4.12	1.47		
−3	Ala	8.21	4.16	1.50		
−2	Leu	7.98	4.20	1.68, 1.90		
−1	Gly	8.09	3.94			
1	Lys	7.72	4.26	1.95	1.53	δCH 1.74; εCH 3.03; ζNH 7.56
2	Val	7.72	3.84	2.03	0.66, 0.91	
3	Phe	7.77	4.66	3.01, 3.32		aromatic CH 7.31
4	Glu	7.79	4.40	2.13, 2.22	2.53	
5	Arg	7.99	4.42	1.86, 1.99	1.73	δCH 3.26; εNH 7.15

[a] Chemical shifts are in ppm referenced to DSS.

be attributed to the dynamic equilibrium between helical and nonhelical conformations. To clarify this issue, we compared the relative rates of chemical exchange of amide protons. The exchange rates were classified into four groups as shown in Figure 8. In all four peptides, the amide protons in the core of the hydrophobic segment were hardly exchanged at 18 h, indicating that this region forms a persistent helical structure. The amide protons in the N- and C-terminal (-KVFER) regions of the peptides were exchanged fast, consistent with a less helical conformation suggested by NOE connectivities. In the C-terminal region of the signal sequence (-LAALG-), however, a significant difference was observed between the functional (L8 and CL10) and nonfunctional (L8PL and L8PL2) signal sequences. In the case of L8PL-M5 and L8PL2-M5, the amide protons in this region were exchanged nearly as slowly as the protons in the core region. This suggests that a very stable helical structure is formed throughout the signal sequence except for a few residues at the N-terminus. On the other hand, the corresponding amide protons in L8-M5 and CL10-M5 were exchanged faster, indicating that the helical structure of this region is less stable (i.e., conformational equilibrium is somewhat shifted to a more random structure). In the case of CL10-M5, the helical structure immediately preceding the proline residue (-FL-) is also less stable than that in the other peptides (Figure 8D). Therefore, the overestimation by NMR of the helix content in L8-M5 and CL10-M5 may be attributed to the relative instability of the helical structure in these peptides. These results can also explain why L8PL2-M5 has significantly higher helix content than L8-M5 (Table II). In addition, NH(i)–NH(i+1) NOE connectivities were observed to the C-terminus in L8PL-M5 and L8PL2-M5 but not in L8-M5 and CL10-M5 (Figure 8), also in agreement with a less stable C-terminal structure of the latter peptides.

We have previously predicted that the C-terminal region of L8 is extended, and those of L8PL and L8PL2 are helical

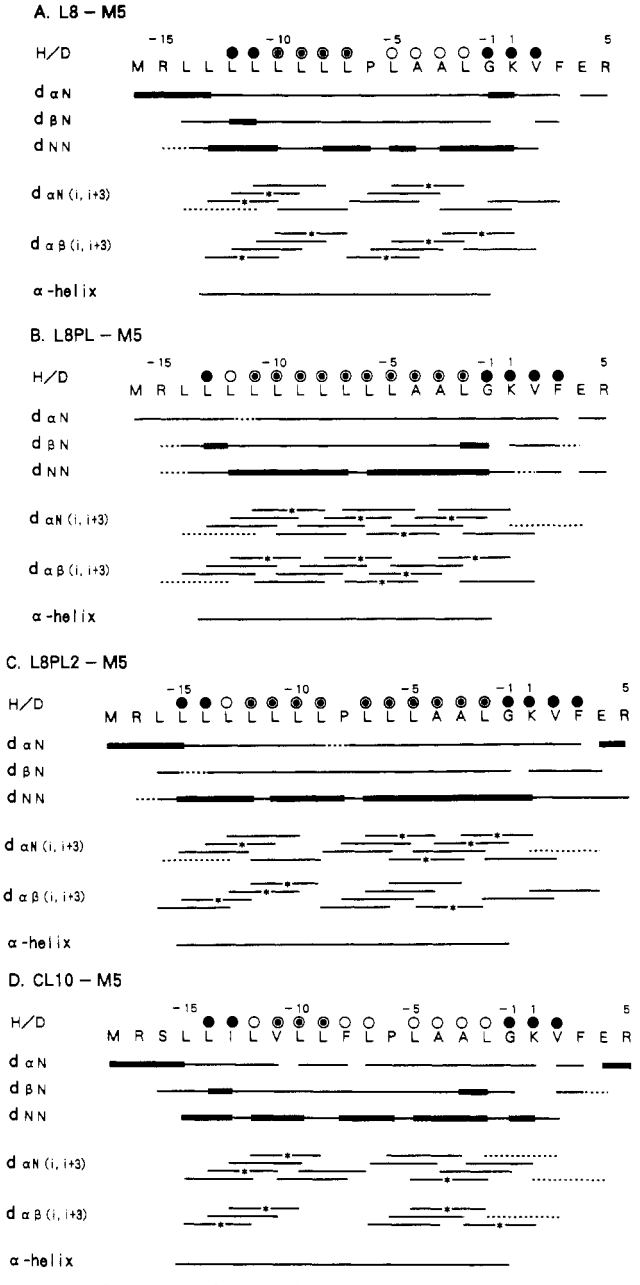

FIGURE 8: Summary of NMR data for L8-M5 (A), L8PL-M5 (B), L8PL2-M5 (C), and CL10-M5 (D). NOE connectivities are classified into three groups [strong (thick bars), medium (—), and weak (---)] according to the number of contours of the cross-peaks. Note that connectivities to δ-proton of the proline are also included. The connectivities marked with an asterisk (*) are rather ambiguous due to cross-peak overlap or chemical shift degeneracy. The top line summarizes amide exchange rates with D_2O. (⊙), (○), and (●) indicate amide protons that remained at 18 h, 3 h, and 30 min after addition of D_2O. Identified helical segments are also shown.

(Yamamoto et al., 1989). The prediction for L8 is not in agreement with the NMR structure described above, while the predicted conformations of L8PL and L8PL2 are consistent with those observed. However, if the conformational equilibrium in the C-terminal region of the signal sequence is considered, the present result would be compatible with the previous prediction (see also Discussion).

Discussion

It has been proposed that the helical conformation of the hydrophobic segment is necessary for the function of signal sequences (Briggs & Gierasch, 1984; Gierasch, 1989; McKnight et al., 1989). Our CD and NMR results have

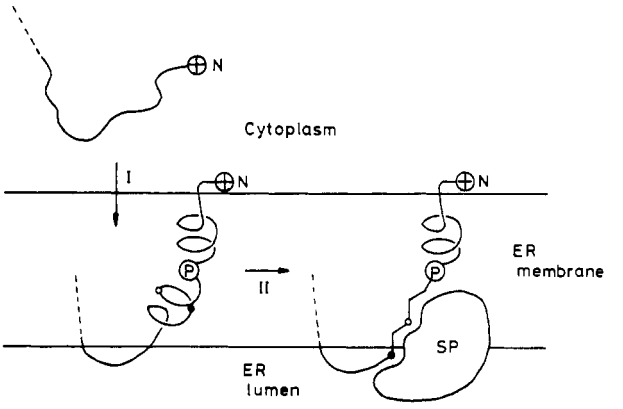

FIGURE 9: Proposed conformational change of functional signal sequence L8 in the membrane. SP and P denote the signal peptidase and the proline residue, respectively. The residues at positions −3 and −1 are indicated by (○) and (●), respectively. The bend of the helix is rather arbitrarily drawn according to the structure of melittin (Inagaki et al, 1989), and the helix might be nearly straight. Although the conformation of the signal sequence in water has not been determined due to its insolubility, it is probably random coiled, judging from the results of other signal sequences (Batenburg et al., 1988; Briggs & Gierasch, 1984; Shinnar & Kaiser, 1984).

indicated that signal sequence L8 predominantly adopts a helical conformation in apolar environment. The observed conformation of L8 is consistent with the above proposal and may represent one probable structure of the signal sequence in the membrane. On the other hand, our H–D exchange experiments have shown that the helical structure of the C-terminal region of L8 is less stable than that of the nonfunctional signal sequences, suggesting that the destabilization of the C-terminal helix has functional significance. Signal sequences are cleaved from mature proteins by the signal peptidase, which shows a specificity toward the residues at positions −3 and −1 [(−3,−1)-rule] (von Heijne, 1983, 1986). This rule suggests that the C-terminal region of the signal sequence adopts an extended conformation when bound to the signal peptidase. Moreover, a helical conformation seems unsuitable for interaction with the peptidase, inasmuch as the scissile bond is partially shielded in this conformation. These considerations led us to propose that L8 is inserted in the membrane with a helical conformation and the C-terminal helix becomes unraveled into an extended form upon interaction with the signal peptidase (Figure 9).

On the basis of this model, conformational properties of the other signal sequences were examined, in relation to their biological activities. CL10 adopts a conformation similar to that of L8 with a less stable C-terminal helix, being compatible with the data showing that CL10 is nearly as functional as L8 (Table I). L8PL, which has no proline residue in the sequence, will be inserted in the membrane as a straight helix. As this helix is more stable than that of L8, the conformational change described above could hardly take place, and the interaction with the signal peptidase would be inhibited. This might explain why the processing efficiency of L8PL is very low (Yamamoto & Kikuchi, 1989). L8PG may also adopt a conformation similar to that of L8PL. In this case, however, the substituted glycine residue destabilizes the helix to some extent (Figure 1 and Table II), and accordingly, some fraction of the molecules would be able to change in conformation to interact with the signal peptidase. This conformational feature of L8PG is consistent with its processing efficiency, which is low but significantly higher than that of L8PL (Yamamoto & Kikuchi, 1989). In L8PL2, two inserted leucine residues stabilize the α-helical conformation of the C-terminal region

compared with that of L8. This stabilization would also hinder the conformational change described above and can account for the low processing efficiency of L8PL2 (Yamamoto & Kikuchi, 1989). In sum, the experimental results can be well explained by our model, which implies that a stable helix in the hydrophobic segment is favorable but a too stable helix in the C-terminal region is unfavorable for the function of the signal sequence.

In the above discussion, it has been suggested that the conformational features of the C-terminal region of the signal sequences are responsible for the interaction with the signal peptidase. We have recently performed in vitro translocation experiments using rabbit reticulocyte lysate and dog pancreas microsomes and have found that precursor human lysozyme with L8PL or L8PL2 is translocated across the membrane but not processed to the mature form (Kohara et al., unpublished results). In addition, Nothwehr et al. (1989) have shown that the structural features of the C-terminal region of signal sequences are necessary for cleavage but not for targeting to and translocation across the microsomal membrane. These results are consistent with the interpretation discussed above.

A statistical analysis has shown that not just a few natural signal sequences have a proline residue at position −4, −5, or −6 (von Heijne, 1986). The model presented here can be applied to these signal sequences as well. However, it is also true that some signal sequences have no proline residue at these positions. In these cases, other helix-destabilizing amino acids such as glycine and serine are often found at positions −4 to −6 (von Heijne, 1986), and they may be substituted for the proline residue.

There have been several reports on the conformational analysis of synthetic signal peptides (Batenburg et al., 1988; Briggs & Gierasch, 1984; McKnight et al., 1989; Reddy & Nagaraj, 1985; Rosenblatt et al., 1980; Shinnar & Kaiser, 1984). In these studies, whole signal sequences or their partial fragments were used without the mature region of the secretory protein, except for one case (Rosenblatt et al., 1980). Such an approach would be valid for investigating the conformation of the hydrophobic segment itself, which is fairly distant from the C-terminal cleavage site. However, the loss of the N-terminal segment of mature protein will probably affect the conformation of the C-terminal region of the signal sequence. Moreover, the signal sequence is supposed to be attached to the mature protein when it is inserted in the membrane, as cleavage of the signal sequence represents a later step in protein translocation. For these reasons we used peptides in which the N-terminal five residues of mature human lysozyme were each connected to the C-terminus of the signal sequence. The five residues were chosen because mature human lysozyme has a cysteine residue at position 6, inclusion of which might complicate the interpretation of the results due to formation of a disulfide bond.

In the present study, aqueous TFE was used as the solvent to mimic the apolar or interfacial environment of the membrane. It remains unclear whether TFE can actually mimic the microenvironment to be encountered in the export process in vivo. However, the conformations of L8-M5 and L8PG-M5 in aqueous TFE turned out to be similar to their conformations in SDS micelles (Figure 1), which are more closely related to the membrane bilayers. In addition, a correlation was found between the secretory capability in vivo and conformational properties of the signal sequences in aqueous TFE. Such a correlation between the conformations in aqueous TFE and biological activities was also found in λ-receptor protein signal sequence and its mutants (Briggs & Gierasch, 1984). Thus, it seems reasonable as a first approximation to use the results obtained here to infer the conformational properties of the signal sequence in vivo. In the future work, we will examine the conformations of the peptides in lipid bilayers.

An interesting observation from our studies is the apparent difference in the properties (H–D exchange, NOEs) of the C-terminus as a function of the rest of the structure. The peptides studied have a common C-terminal sequence (LAALGKVFER), but its conformational properties are not identical among the peptides and are dependent on the residues preceding the sequence. This suggests the importance of context in the structure of a given sequence and may offer a clue to the problem of peptide folding.

During the preparation of the manuscript we found papers by Bruch et al. (1989) and Bruch and Gierasch (1990) that also describe CD and 2D NMR analyses of signal peptides in aqueous TFE solution. They have shown that the signal peptides are in dynamic equilibrium between a helical and a more random conformation and that the hydrophobic core has the most persistent helix conformation. These results are consistent with our data and may reflect general properties of functional signal sequences. They have also shown that the stability of the helix in the hydrophobic core correlates with in vivo function of different mutants of the LamB signal sequence. Our results, however, indicate that the formation of a stable helix is not sufficient for the function of signal sequences.

Acknowledgments

We thank Dr. M. Ikehara for advice and encouragement and Drs. H. Nakamura and W. Yoshikawa for discussion.

Supplementary Material Available

1D NMR spectra of L8-M5, L8PL-M5, L8PL2-M5, and CL10-M5 (4 pages). Ordering information is given on any current masthead page.

References

Aue, W. P., Bartholdi, E., & Ernst, R. R. (1976) *J. Chem. Phys. 64*, 2229–2246.

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

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 63*, 207–213.

Bazzo, R., Tappin, M. J., Pastore, A., Harvey, T. S., Carver, J. A., & Campbell, I. D. (1988) *Eur. J. Biochem. 173*, 139–146.

Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–528.

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

Briggs, M. S., & Gierasch, L. M. (1986) *Adv. Protein Chem. 38*, 109–180.

Bruch, M. D., & Gierasch, L. M. (1990) *J. Biol. Chem. 265*, 3851–3858.

Bruch, M. D., McKnight, C. J., & Gierasch, L. M. (1989) *Biochemistry 28*, 8554–8561.

Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2821–2822.

Emr, S. D., & Silhavy, T. J. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4599–4603.

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

Greenfield, N., & Fasman, G. D. (1969) *Biochemistry 8*, 4108–4116.

Inagaki, F., Shimada, I., Kawagishi, K., Hirano, H., Terasawa, I., Ikura, T., & Go, N. (1989) *Biochemistry 28*, 5985–5991.

Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546-4553.
McKnight, C. J., Briggs, M. S., & Gierasch, L. M. (1989) *J. Biol. Chem. 264*, 17293-17297.
Nelson, J. W., & Kallenbach, N. R. (1986) *Proteins 1*, 211-217.
Nothwehr, S. F., Folz, R. J., & Gordon, J. I. (1989) *J. Biol. Chem. 264*, 4642-4647.
Provencher, S. W., & Glöckner, J. (1981) *Biochemistry 20*, 33-37.
Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479-485.
Reddy, G. L., & Nagaraj, R. (1985) *Biochim. Biophys. Acta 831*, 340-346.
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.
von Heijne, G. (1983) *Eur. J. Biochem. 133*, 17-21.
von Heijne, G. (1986) *Nucleic Acids Res. 14*, 4683-4690.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Yamamoto, Y., & Kikuchi, M. (1989) *Eur. J. Biochem. 184*, 233-236.
Yamamoto, Y., Taniyama, Y., Kikuchi, M., & Ikehara, M. (1987) *Biochem. Biophys. Res. Commun. 149*, 431-436.
Yamamoto, Y., Taniyama, Y., & Kikuchi, M. (1989) *Biochemistry 28*, 2728-2732.

Reduction of the Potent DNA Polymerase III Holoenzyme 3′→5′ Exonuclease Activity by Template-Primer Analogues†

Mark A. Griep,‡ Jo Anna Reems, Mary A. Franden, and Charles S. McHenry*
Department of Biochemistry, Biophysics, and Genetics, University of Colorado Health Sciences Center, Denver, Colorado 80262
Received March 29, 1990; Revised Manuscript Received June 5, 1990

ABSTRACT: The DNA polymerase III holoenzyme of *Escherichia coli* contains a potent 3′→5′ exonuclease that removes the terminal nucleotide from a synthetic deoxyoligonucleotide primer with a half-life of approximately 2 s. Degradation of primers could not be effectively prevented by permitting the holoenzyme to "idle" at the primer terminus in the presence of limited deoxynucleoside triphosphates. To further characterize this exonuclease and to develop stable primers to facilitate experimental manipulations, we synthesized a series of twelve 25-mer oligonucleotides that differed only in the two 3′-terminal residues. The penultimate position contained either a CMP or a dCMP residue, while at the terminal position either AMP, dAMP, 2′,3′-dideoxyAMP, cordycepin (3′-dAMP), dAMPαS, or 2′,3′-dideoxyAMPαS was incorporated. No single change at either the 3′-penultimate or 3′-terminal positions resulted in a decrease in the exonuclease rate greater than 10-fold; however, combined changes at these two sites resulted in a strong synergistic effect. Placing a ribonucleotide at the penultimate position coupled by a phosphorothioate linkage to a terminal 2′,3′-dideoxynucleotide reduced the rate of exonucleolytic activity almost 30000-fold (half-life ~ 16 h). If only the ribonucleotide and phosphorothioate substitutions were made, a primer capable of being efficiently elongated was generated that exhibited a 500-fold increase in stability (half-life = 40 min). The elemental effect observed by substituting a nonbridging oxygen in the terminal phosphodiester bond for sulfur increased from 1.5 to 200 as other substitutions were made that decreased the exonuclease rate. This was consistent with a change in the rate-limiting step of the exonuclease reaction from a conformational change to the chemical step where the covalent bond is cleaved. At least part of this effect appears to be due to perturbations within the enzyme's active site and not solely due to changes in electrophilicity.

The DNA polymerase III holoenzyme is the major replicative polymerase of *Escherichia coli*. It contains a DNA polymerase III core plus auxiliary proteins that confer the special properties required of a true replicative complex. These properties include rapid polymerization rates, the ability to form a highly processive ATP-dependent clamp on the DNA template, the ability to interact with other replication proteins, and the apparent ability to function as an asymmetric dimer with distinguishable leading and lagging strand polymerases [for reviews, see Kornberg (1982) and McHenry (1988a,b)].

The DNA polymerase III core of holoenzyme[1] contains three subunits: α, ε, and θ of 130000, 27500, and 10000 daltons, respectively (McHenry & Crow, 1979). α, encoded

† This work was supported by Research Grants RO1 GM 35695 from the National Institutes of General Medical Sciences and MV348 from the American Cancer Society.
* Author to whom correspondence should be addressed.
‡ Supported in part by Postdoctoral Fellowship PF-3025 from the American Cancer Society.

[1] Abbreviations: SSB, *E. coli* single-stranded DNA-binding protein; ddATP, 2′,3′-dideoxyadenosine triphosphate; dATPαS, 2′-deoxyadenosine 5′-*O*-(1-thiotriphosphate); ddATPαS, 2′,3′-dideoxyadenosine 5′-*O*-(1-thiotriphosphate); HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; DTT, dithiothreitol; holoenzyme, *E. coli* DNA polymerase III holoenzyme; polIII′, *E. coli* DNA polymerase III′; PEI, poly(ethylenimine); 3′-[^{32}P]...NN, 25-mer primer radioactively labeled between nucleotides 24 and 25 at the 3′-end with either α-^{32}P or α-^{35}S in the phosphodiester or phosphorothioate linkage; "...", the 23 nucleotides preceding the two 3′-terminal nucleotides; NN, the 3′-penultimate and terminal nucleotide residues; ...dC, 24-mer 3′-terminated with 2′-deoxycytidine; ...rC, 24-mer 3′-terminated with cytidine; ...dCrA, 25-mer with a 3′-penultimate 2′-deoxycytidine and 3′-terminal adenosine; ...rCrA, 25-mer with a 3′-penultimate cytidine and 3′-terminal adenosine.

0006-2960/90/0429-9006$02.50/0