Majors, J. (1977) Ph.D. Thesis, Harvard University.
McClure, W. R., Cech, C. L., & Johnston, D. E. (1978) *J. Biol. Chem. 253*, 8941–8948.
Miller, J. S., & Burgess, R. R. (1978) *Biochemistry 17*, 2064–2069.
Minkley, E. G., & Pribnow, D. (1973) *J. Mol. Biol. 77*, 255–277.
Pfeffer, S. R., Stahl, S. J., & Chamberlin, M. J. (1977) *J. Biol. Chem. 252*, 5403–5407.
Pribnow, D. (1975a) *J. Mol. Biol. 99*, 419–443.
Pribnow, D. (1975b) *Proc. Natl. Acad. Sci. U.S.A. 72*, 784–788.
Seeburg, P. H., Nüsslein, C., & Schaller, H. (1977) *Eur. J. Biochem. 74*, 107–113.
Stahl, S. J., & Chamberlin, M. J. (1977) *J. Mol. Biol. 112*, 577–601.
Williams, R. C., & Chamberlin, M. J. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 3740–3744.

Purification and Characterization of the Amino-Terminal Propeptide of Pro α1(I) Chains from Embryonic Chick Tendon Procollagen†

Donna M. Pesciotta, Mark H. Silkowitz, Peter P. Fietzek, Peter N. Graves, Richard A. Berg, and Bjørn R. Olsen*

ABSTRACT: A peptide with an apparent molecular weight of 23000 was isolated from the medium of cultured chick embryo tendons. Comparison of tryptic peptides derived from the medium peptide and from the amino-terminal, bacterial collagenase resistant portion of Type I procollagen indicated that the medium peptide represented the amino-terminal precursor-specific region of the pro α1(I) chain of procollagen. This conclusion was supported by the demonstration that antibodies against the medium peptide reacted with Type I procollagen in a radioimmune assay but did not react with a peptide derived from the carboxy-terminal propeptide of Type I procollagen. In addition, the reaction with Type I procollagen was inhibited with the purified amino-terminal, collagenase-resistant portion of pro α1(I) chains. Finally, amino acid sequencing demonstrated that the amino propeptide of dermatosparactic calf pN α1(I) chains and the medium peptide have similar amino-terminal sequences. Carbohydrate analysis established the presence of one residue of *N*-acetylglucosamine and a trace of mannose and galactose.

Collagen comprises a family of molecules with similar structural features. Each molecule consists of three helical polypeptide chains (α chains) held together by hydrogen bonds to form a triple helical structure. Collagens are synthesized as procollagens that are larger than collagens because of additional amino acid sequences at both the amino and carboxyl ends of the α chains of collagen molecules [for a review, see Fessler & Fessler (1978)]. During the extracellular conversion of procollagen to collagen, these propeptides are removed by procollagen proteases to yield the collagen molecule [see Fessler & Fessler (1978)].

A truncated form of procollagen with an amino propeptide but lacking the carboxyl propeptide has been found to accumulate in the skin of animals with dermatosparaxis (Lenaers et al., 1971; Furthmayr et al., 1972; Becker et al., 1976, 1977). This pN-collagen[1] has provided material for chemical, immunological, and physical studies on the amino propeptide of Type I procollagen (Becker et al., 1976; Rohde et al., 1976; Engel et al., 1977; Bruckner et al., 1978; Hörlein et al., 1978, 1979; Rohde & Timpl, 1979). These studies show that the amino-terminal propeptide of dermatosparactic calf and sheep pN α1(I) chains contains about 140 amino acid residues arranged in three structural domains. The amino-terminal portion of the peptide (~90 residues) contains a region which is resistant to bacterial collagenase. The central portion (~30 residues) is sensitive to bacterial collagenase, and this region is linked to a short, nonhelical sequence (~10 residues) which connects the propeptide to the collagen α1(I) chains (Becker et al., 1976). The complete amino acid sequence of the amino-terminal propeptide from dermatosparactic calf skin has been established (Hörlein et al., 1979). Also, the complete sequence of the collagenase-resistant domain of the dermatosparactic sheep peptide has been determined (Rohde et al., 1979). The sequence data for the calf and sheep dermatosparactic peptide indicate a highly conserved structure with only a few amino acid substitutions.

It has been assumed but not proven chemically that amino propeptides of Type I pN-collagen isolated from dermatosparactic skin are identical with amino propeptides of Type I procollagen secreted by cells in culture. Although it is clear that dermatosparactic pN-collagen is the product of normal processing at the carboxyl end of Type I procollagen, it is also possible that the amino-terminal propeptide is partially processed, so that truncated amino propeptides are produced. Here, we compare the structure of the pro α1(I) amino propeptide of Type I procollagen secreted by chick embryo fibroblast in suspension culture to that of calf and sheep dermatosparactic propeptides. Our data demonstrate that the propeptides of the dermatosparactic calf and sheep pN α1(I)

† From the Department of Biochemistry, College of Medicine and Dentistry of New Jersey–Rutgers Medical School, Piscataway, New Jersey 08854. *Received November 30, 1979.* Supported in part by Research Grants AM 21471 and GRS 279698 from the National Institutes of Health of the U.S. Public Health Service.

[1] Abbreviations used: BAPN, β-aminopropionitrile; NEM, *N*-ethylmaleimide; CM, carboxymethyl; DEAE, diethylaminoethyl; Tris, tris(hydroxymethyl)aminomethane; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; IgG, immunoglobulin G; Tos-Phe-CH_2Cl, L-1-tosylamido-2-phenylethyl chloromethyl ketone; pN, precursor molecule with an amino-terminal extension; $NaDodSO_4$, sodium dodecyl sulfate.

0006-2960/80/0419-2447$01.00/0

chains are similar to the amino propeptide of secreted chick procollagen. One striking difference, however, is that the first seven residues in the calf and sheep peptides appear to be absent from the chick pro α1(I) amino propeptide.

Materials and Methods

Preparation of the Medium Peptide. Leg tendons from 17-day-old chick embryos were incubated in Dulbecco's modified Eagle's medium as described previously (Olsen et al., 1977). A commercial mixture (1 μCi/mL) of ^{14}C-labeled amino acids (New England Nuclear Corp., L-alanine, L-arginine, L-aspartic acid, L-glutamic acid, glycine, L-histidine, L-isoleucine, L-leucine, L-lysine, L-phenylalanine, L-proline, L-serine, L-threonine, L-tyrosine, and L-valine) was added to the medium. After incubation for 24 h, the medium was separated from the tendons by centrifugation, protease inhibitors were added (Olsen et al., 1976), and the medium was dialyzed at 4 °C against 50 mM Tris-HCl, pH 8.6, at 25 °C, containing 2 M urea. The medium was chromatographed on a DEAE-cellulose column as described (Olsen et al., 1977). The fractions containing the last two peaks [see Olsen et al. (1977)] were pooled and desalted by chromatography on polyacrylamide (Bio-Gel P2; 200–400 mesh; Bio-Rad Laboratories) in 0.2 M ammonium bicarbonate and lyophilized. The lyophilized sample was dissolved in 10 mM sodium acetate, pH 4.2, containing 6 M urea and applied at room temperature to a 1.5 × 10 cm column of CM-cellulose (CM 52; Whatman, Inc.) equilibrated with the same buffer. The column was eluted with a 600-mL linear gradient of NaCl from 0 to 0.2 M in the sodium acetate–urea buffer. The flow rate was 60 mL/h and 2.5-mL fractions were collected. Appropriate fractions were pooled, desalted on a column of polyacrylamide (Bio-Gel P2) equilibrated with 0.2 M ammonium bicarbonate, and lyophilized.

Preparation of the Collagenase-Resistant Portion of the Medium Peptide. The medium peptide was reduced and alkylated as described (Olsen et al., 1977) except that the concentration of iodoacetic acid was 50 mM instead of 80 mM. The reduced and alkylated peptide was dialyzed against a 0.4 M NaCl, 0.1 M Tris-HCl buffer, pH 7.5, at 4 °C. For collagenase digestion, $CaCl_2$ and NEM were added to final concentrations of 5 mM and 2.5 mM, respectively. The sample was then incubated with 50 units/mL bacterial collagenase (Advance Biofactures Corp.) for 3 h at 37 °C. The reaction was stopped by dialysis against 0.1 M Tris-HCl buffer, pH 7.5, at 4 °C, containing 2 M urea.

The collagenase-resistant portion of the reduced and alkylated medium peptide was isolated by chromatography on a 1.5 × 5 cm DEAE-cellulose column equilibrated with 0.1 M Tris-HCl buffer, pH 7.5, at 25 °C, containing 2 M urea, and eluted with a 200-mL linear gradient of 0–0.3 M NaCl in the same buffer. The flow rate was 30 mL/h and 2-mL fractions were collected. In some experiments the unreduced medium peptide was digested with bacterial collagenase and chromatographed as described above.

Isolation of Pro α1(I) and Pro α2 Chains of Procollagen and Isolation of the Collagenase-Resistant, Amino-Terminal Region of Pro α1(I). [^{14}C]Procollagen was isolated from the medium of chick tendon fibroblasts incubated in suspension culture and purified by chromatography on DEAE-cellulose as described previously (Hoffmann et al., 1976).

The procollagen was dialyzed against 0.5 M Tris-HCl buffer, pH 8.1, 6 M urea, and 1.2 mM EDTA. DTT (Sigma Chemical Co.) was added to 50 mM final concentration, and the mixture was incubated under an atmosphere of N_2 for 4 h at room temperature. The protein was alkylated by adding iodoacetic acid to a final concentration of 110 mM, and the incubation was continued in the dark at room temperature for 1 h. The sample was dialyzed at 4 °C against 10 mM Tris-HCl buffer, pH 8.6, and 6 M urea, heated to 50 °C for 10 min, and applied to a 1.5 × 5 cm of column of DEAE-cellulose in the same buffer at room temperature. The column was eluted with a 200-mL linear salt gradient from 0 to 0.3 M NaCl. The flow rate was 20 mL/h and 2.5-mL fractions were collected.

The fractions containing pro α1(I) chains were pooled and dialyzed against 0.4 M NaCl and 0.1 M Tris-HCl buffer, pH 7.5. For collagenase digestion, NEM and $CaCl_2$ were added to concentrations of 5 and 2.5 mM, respectively. Bacterial collagenase (Advance Biofactures Corp.), 50 units/mL, was added and the sample was incubated at 37 °C for 3 h. After dialysis at 4 °C against 50 mM sodium acetate buffer, pH 4.2, it was applied to a 1.5 × 5 cm column of CM-cellulose. The column was eluted with a 240-mL linear salt gradient from 0 to 0.3 M NaCl. The flow rate was 25 mL/h and 3-mL fractions were collected. The amino-terminal, collagenase-resistant region of the pro α1(I) chain was eluted in the breakthrough volume of the CM-cellulose column, and it was further purified by chromatography on a 1.5 × 5 cm DEAE-cellulose column as described for the purification of the collagenase-resistant portion of the medium peptide (see above).

Intact procollagen isolated by chromatography on DEAE-cellulose (see above) was also treated with bacterial collagenase as previously described (Olsen et al., 1976). The digest was chromatographed in 0.2 M ammonium bicarbonate on a column of 8% agarose (Bio-Gel A 1.5m) as described (Olsen et al., 1976). The fractions containing the amino-terminal, collagenase-resistant region of procollagen were pooled and lyophilized.

Preparation of Tryptic Fragments of the Medium Peptide. To restrict the action of trypsin to arginyl residues, we modified the lysine residues in the medium peptide by reacting the peptide with phenyl isothiocyanate. For this treatment, ~5 mg of medium peptide, reduced and alkylated, was lyophilized and dissolved in 2.0 mL of dimethylallylamine–trifluoroacetic acid buffer in pyridine–water (Beckman), pH 9.3. After addition of 0.1 mL of phenyl isothiocyanate, the sample was incubated at 58 °C for 1 h under nitrogen. After extraction 5 times with 2 mL of benzene, the aqueous phase was lyophilized. The residue was dissolved in 2.0 mL of freshly made 0.2 M ammonium bicarbonate. Tos-Phe-CH_2Cl-treated trypsin (250 μg) (Worthington Biochemical Corp.) was added, and the sample was incubated at 37 °C for 4 h. The tryptic fragments were separated by chromatography on a 1.5 × 110 cm Sephadex G-50 column equilibrated with 0.2 M ammonium bicarbonate.

Tryptic peptides for fingerprinting were prepared from reduced and alkylated peptides by digestion with Tos-Phe-CH_2Cl-treated trypsin in 0.5 mL of 0.1 M ammonium bicarbonate at 37 °C for 4 h. The reaction was stopped by the addition of a few drops of 0.5 N acetic acid, and the samples were lyophilized several times. The peptides were separated according to the method of Bates et al. (1975) with the modifications described (Olsen et al., 1977).

Preparation of Antibodies and Radioimmune Assays. Antibodies to the medium peptide were prepared in rabbits by intradermal injections of purified medium peptide in Freund's complete adjuvant. The antisera were tested with direct binding radioimmune assays using a double-antibody method with sheep antirabbit IgG antiserum in the second precipitation step as described (Nist et al., 1975; Rohde et al.,

1976). The assays were performed in duplicate by adding the appropriate amount of antiserum to ^{14}C-labeled antigen (Olsen et al., 1977). For inhibition studies, the inhibitors were incubated with the antisera for 24 h at 4 °C before the antigen was added. The incubation was then continued for an additional 24 h. The antigen–antibody complexes were precipitated by the addition of sheep antirabbit IgG antiserum and incubated 24 h at 4 °C. The precipitate was isolated by centrifugation, washed, and counted as described (Olsen et al., 1977).

Removal of Pyroglutamic Acid with Calf Liver Pyroglutamyl Aminopeptidase. Lyophilized medium peptide (4 mg) was dissolved in 4 mL of 0.1 M sodium phosphate buffer, pH 8.0, 10 mM EDTA, and 5% glycerol and dialyzed overnight at 4 °C (Podell & Abraham, 1978). DTT was added to a final concentration of 5 mM. The sample was then incubated with 800 μg of pyroglutamyl aminopeptidase (Boehringer Mannheim) at 4 °C for 9 h. An additional 800 μg of pyroglutamyl aminopeptidase was then added, and the reaction was continued at 20 °C for 14 h. After incubation the peptide was desalted on a 2.0 × 56 cm column of polyacrylamide (Bio-Gel P2; 100–200 mesh) equilibrated and eluted with water adjusted to pH 8.5 with dilute ammonium hydroxide. The desalted peptide was concentrated and used for protein sequencing.

Protein Sequencing Techniques. Peptides were sequenced by automatic Edman degradation in a Beckman Model 890C protein–peptide sequencer. A modified Quadrol program was used. The PTH-amino acid derivatives were identified by high-pressure liquid chromatography (Waters Associates) with a methanol–water–acetic acid–acetone solvent system as described by Bhown et al. (1978) and by thin-layer chromatography on silica-coated 20 × 20 cm thin-layer plates using the P system (*tert*-butyl propionate) and H system (ethylene dichloride–acetic acid) of Edman (Fietzek & Rexrodt, 1975).

Other Assays and Procedures. Polyacrylamide slab gel electrophoresis in $NaDodSO_4$ (Bio-Rad Laboratories) was carried out on acrylamide slab gels as described (King & Laemmli, 1971; Olsen et al., 1977). Amino acid analyses were performed on peptides dialyzed against 0.1 N acetic acid and lyophilized. Hydrolysis was carried out in the presence of 6 N hydrochloric acid and 0.06 M 2-mercaptoethanol for 16 h at 116 °C under nitrogen. The analyses were performed on a JEOL Model JLC 6AH amino acid analyzer as described (Berg & Prockop, 1973) and on a Beckman Model 121 MB amino acid analyzer. Cysteine was assayed as (carboxymethyl)cysteine after reduction and alkylation of peptides. Hexosamines were determined by hydrolysis in 4 N hydrochloric acid for 12 h at 116 °C and analysis on the amino acid analyzer as described (Berg et al., 1979).

Neutral sugars were analyzed by the technique of Metz et al. (1971) as described (Olsen et al., 1977; Berg et al., 1979). Hydrolysis was under vacuum in 1 N hydrochloric acid at 100 °C for 9 h. Arabinose was added as an internal standard. The alditol acetate derivatives were analyzed with a Hewlett-Packard Model 5830A gas–liquid chromatograph by using a column packed with 3% SP-2340 on 100/120 Supelcoport (Supelco, Inc.).

Results

Purification of the Medium Peptide. Media from each tendon incubation were chromatographed on DEAE-cellulose in the presence of 2 M urea. The elution profile of the ^{14}C-labeled proteins is shown in Figure 1. As previously demonstrated (Olsen et al., 1977), the second, large peak contained the carboxyl propeptide of procollagen. For the present study

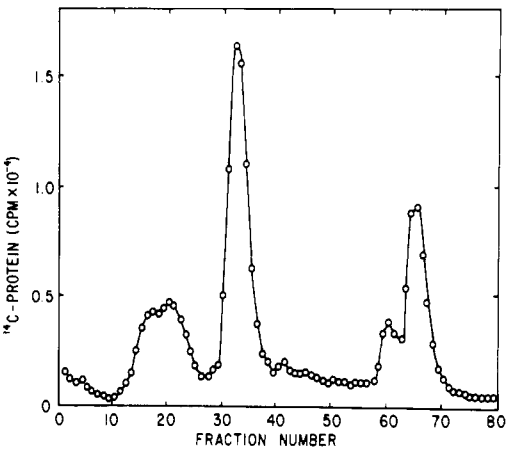

FIGURE 1: Elution profile of the medium peptides chromatographed on DEAE-cellulose. Leg tendons from ~30-dozen 17-day-old chick embryos were incubated as described under Materials and Methods. The medium was collected and chromatographed on a 2.5 × 10 cm column. The column was eluted with a linear salt gradient of 0–0.2 M NaCl in 50 mM Tris-HCl, pH 8.6, at 4 °C, containing 2 M urea. The flow rate was 120 mL/h and 6-mL fractions were collected. Fractions 58–69 were pooled for further analysis.

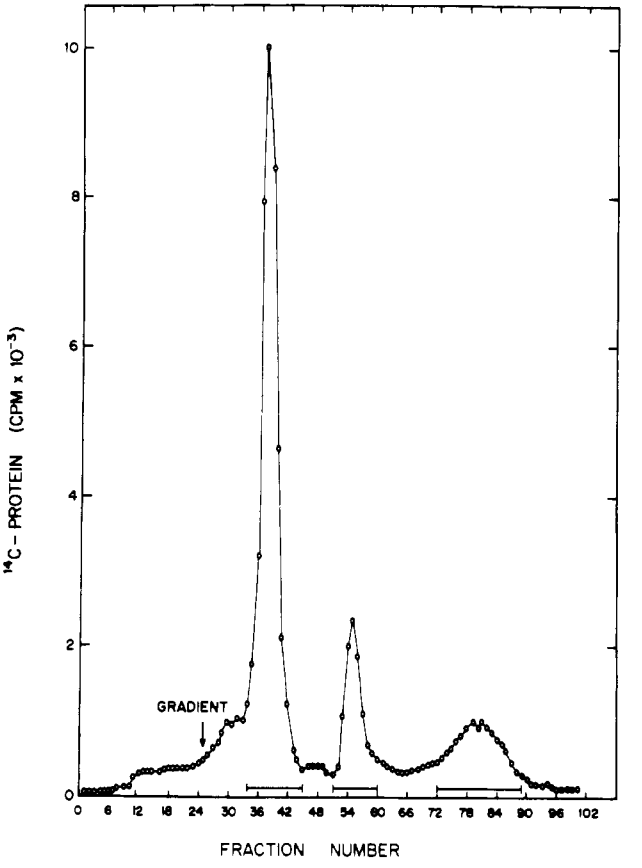

FIGURE 2: Purification of medium peptides by chromatography on CM-cellulose. Pooled material after DEAE-cellulose chromatography (see Figure 1) was chromatographed on CM-cellulose as described under Materials and Methods. Fractions pooled for further analysis are indicated by horizontal bars.

the last two peaks were pooled and chromatographed on a CM-cellulose column in the presence of 6 M urea (Figure 2). The elution profile from the CM-cellulose column contained three peaks of radioactive material. Polyacrylamide gel electrophoresis of material in the first peak gave a single band that migrated with a molecular weight of ~20 000 when compared with a set of globular protein standards (Figure 3). When the material in this peak was reduced, polyacrylamide

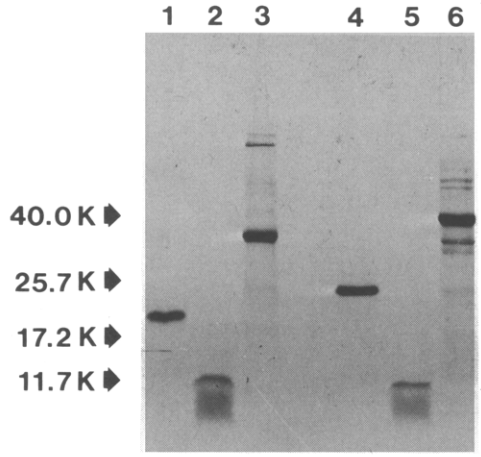

FIGURE 3: Polyacrylamide slab gel electrophoresis of peak fractions from the CM-cellulose column (Figure 2). Samples for electrophoresis were prepared by taking aliquots of the peak fractions and dialyzing them against a sample buffer consisting of 0.125 M Tris-HCl, pH 6.8, 2% $NaDodSO_4$, 10% glycerol, and 0.001% bromophenol blue (Eastman). Reduced samples were prepared by adding 5% mercaptoethanol and boiling for 3 min. Polyacrylamide slab gel electrophoresis was carried out as described (Olsen et al., 1977). The gel was stained for 1 h in 0.25% Coomassie Brilliant Blue R and 20% trichloroacetic acid and destained in 7.5% acetic acid and 15% methanol. Lane 1: fraction 38, unreduced. Lane 2: fraction 55, unreduced. Lane 3: fraction 80, unreduced. Lane 4: fraction 38, reduced. Lane 5: fraction 55, reduced. Lane 6: fraction 80, reduced. Numbers on the left indicate the molecular mass in kilodaltons of reduced globular protein standards run on the same gel.

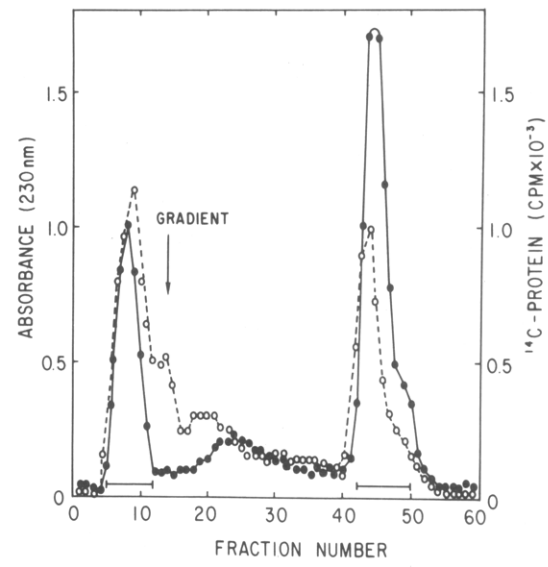

FIGURE 4: Chromatography of bacterial collagenase digested medium peptide on DEAE-cellulose. Medium peptide from the CM-cellulose column (Figure 2) was digested with bacterial collagenase and chromatographed on a 1.5 × 5 cm DEAE-cellulose column as described under Materials and Methods. Symbols: absorbance at 230 nm (○); ^{14}C-labeled protein (●).

gel electrophoresis gave a single band that migrated with an apparent molecular weight of ~23 000.

Gel electrophoresis of material contained in the second and third peaks showed bands with apparent molecular weights of about 12 000 and 36 000–40 000, respectively. The material in these peaks was not characterized further.

Isolation of a Collagenase-Resistant Portion from the Medium Peptide. Radioimmune assays (see below) indicated that the medium peptide was derived from the amino-terminal portion of procollagen. Since amino-terminal propeptides derived from dermatosparactic procollagen have been shown to contain a collagenase-sensitive region (see above), we examined the susceptibility of the medium peptide to bacterial collagenase. Material in the first, major peak from the CM-cellulose column (see Figure 2) was digested with bacterial collagenase, and the digest was chromatographed on a DEAE-cellulose column in 2 M urea (Figure 4). The elution profile showed two peaks of protein and radioactivity. The first peak contained a mixture of collagenase-digested peptides. Amino acid analysis indicated that these peptides had an amino acid composition similar to that of the helical portion of the collagen molecule (not shown). The second peak was examined by polyacrylamide slab gel electrophoresis and consisted of a major band migrating with an apparent molecular weight of 12 500 (Figure 5). This band comigrated with the amino-terminal, collagenase-resistant portion of procollagen isolated by gel filtration on a column of 8% agarose as described (Olsen et al., 1976). It also comigrated (Figure 5) with the amino-terminal, collagenase-resistant portion of the cyanogen bromide peptide isolated from sheep dermatosparactic pN α1(I) chains (Becker et al., 1976).

Isolation of Pro α1(I) Chains and Their Collagenase-Resistant Amino-Terminal Portion. Pro α1(I) chains were prepared from procollagen as described under Materials and Methods by chromatography on DEAE-cellulose (Figure 6). Two peaks of radioactive material were eluted from the column. Slab gel electrophoresis showed that pro α2 chains were

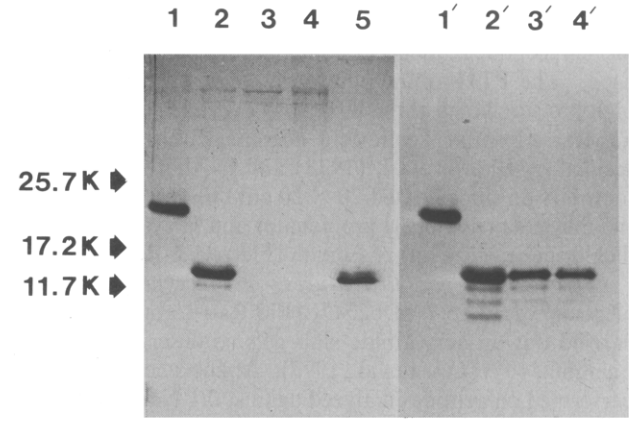

FIGURE 5: Polyacrylamide slab gel electrophoresis of collagenase-resistant regions of the medium peptide, pro α1(I) chain, procollagen, and dermatosparactic sheep pN α1(I) chain. The peptides were prepared for polyacrylamide slab gel electrophoresis and reduced as described in the legend of Figure 3. After photography of the stained gel, the gel was impregnated with PPO as described (Bonner & Laskey, 1974) and exposed to an X-ray film. Lane 1: medium peptide before collagenase digestion. Lane 2: medium peptide after collagenase digestion. Lane 3: amino-terminal, collagenase-resistant portion of pro α1(I) chain. Lane 4: amino-terminal, collagenase-resistant region of procollagen isolated by gel filtration (Olsen et al., 1976). Lane 5: collagenase-resistant region of pN α1(I) chain of dermatosparactic sheep [see Rohde et al. (1979)]. Lanes 1′, 2′, 3′, and 4′ represent the fluorograph of the gel. The peptide in lane 5 was not labeled and therefore did not show up in the fluorograph.

eluted in the first peak and pro α1(I) chains were eluted in the second peak. After collagenase digestion of the isolated pro α1(I) chains, the collagenase-resistant, amino-terminal portion was purified by a combination of CM-cellulose and DEAE-cellulose chromatography as described under Materials and Methods. Slab gel electrophoresis showed that the collagenase-resistant peptide derived from pro α1(I) chains comigrated with the collagenase-resistant portion of the medium peptide (Figure 5).

Further Characterization of the Medium Peptide. The identity of the medium peptide was established by immunological criteria, by mapping of tryptic peptides, and by amino

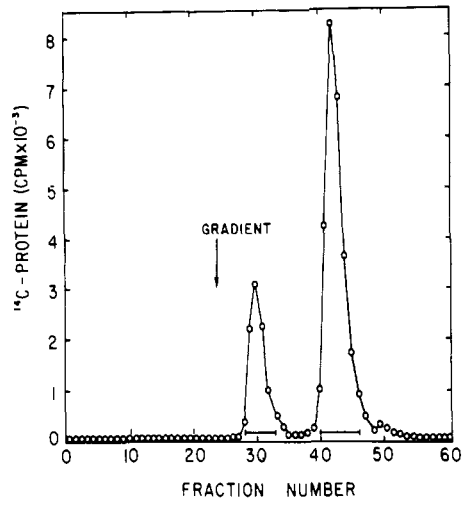

FIGURE 6: Isolation of pro α1(I) and pro α2 chains of Type I procollagen. Procollagen, isolated and purified as previously described (Hoffmann et al., 1976), was reduced and alkylated and chromatographed on a DEAE-cellulose column (see Materials and Methods). Fractions pooled for further analysis are indicated by horizontal bars.

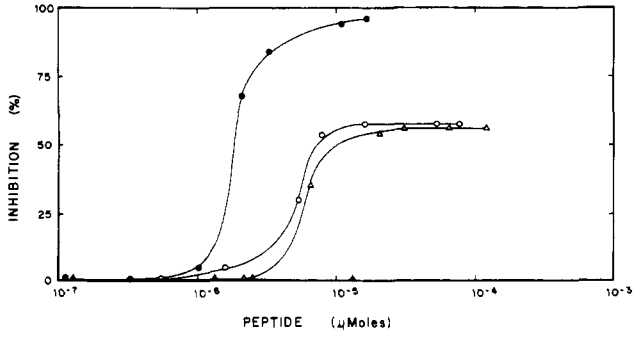

FIGURE 7: Inhibition assay with [^{14}C]procollagen and different peptide inhibitors. The assay involved precipitation of ^{14}C-labeled procollagen with a rabbit antiserum to the medium peptide. For the assay, sufficient antiserum was used to bind 80% of the [^{14}C]procollagen. Inhibitors used were the medium peptide (●), the collagenase-resistant portion of the medium peptide (○), the collagenase-resistant, amino-terminal region of procollagen (△), and the carboxyl propeptide (▲).

acid analysis. Since the antisera to the medium peptide were shown to precipitate procollagen, the antisera were used in a radioimmune assay to compare the ability of various peptides to inhibit the precipitation of procollagen. The medium peptide was 100% effective as an inhibitor in this assay (Figure 7). The collagenase-resistant portion of the medium peptide and the amino-terminal, collagenase-resistant region of procollagen inhibited the precipitation of procollagen up to 60%. When the peptides used as inhibitors in the assay were reduced and alkylated, they had no inhibiting activity. Also, the carboxy-terminal propeptide of procollagen (Olsen et al., 1977) showed no inhibiting activity in the assay.

For the tryptic fingerprinting, the collagenase-resistant portion of the medium peptide and the amino-terminal, collagenase-resistant region of pro α1(I) were digested with trypsin and analyzed by two-dimensional chromatography. The similarity of the fingerprints indicated that the medium peptide contained a portion identical with that of the collagenase-resistant, amino-terminal region of pro α1(I) chains.

Amino acid analysis of the medium peptide showed that its composition was similar to that of the precursor-specific cyanogen bromide fragment isolated from dermatosparactic calf and sheep pN α1(I) chains (Table I). It contained 4-hydroxyproline, suggesting collagen-like sequences within

Table I: Amino Acid and Carbohydrate Composition of Medium Peptide[a]

component	medium peptide	calf pN α1(I)-CBO.1[b]	sheep pN α1(I)-CBO.1[c]
amino acids			
4-Hyp	5	7	8
Asp	21	18	18
Thr	6	7	8
Ser	8	4	5
Glu	19	22	23
Pro	18	22	22
Gly	25	27	28
Ala	7	3	4
$^1/_2$-Cys	10	10	10
Val	8	11	12
Met		1	1
Ile	8	5	5
Leu	8	6	6
Tyr	3	3	3
Phe	3	1	1
Hyl			
Lys	4	4	5
His	1	1	1
Arg	5	5	5
carbohydrate			
N-acetylglucosamine	1		
mannose	<1		
galactose	<1		

[a] Given as residues per peptide rounded off to the nearest whole number. [b] Data from Hörlein et al. (1979). [c] Data from Becker et al. (1976).

the peptide, and relatively large amounts of aspartic and glutamic acid. Differences between the chick peptide and the dermatosparactic peptides were found in values for serine, isoleucine, leucine, tyrosine, and phenylalanine (Table I). Carbohydrate analysis showed the presence of 1 mol of *N*-acetylglucosamine per mol of peptide and only traces of other sugars (Table I).

Amino Acid Sequencing of the Medium Peptide. In a final series of experiments, the amino-terminal sequence of the medium peptide was compared with that of dermatosparactic pN α1(I) chains. When the medium peptide was subjected to stepwise Edman degradation without pretreatment with pyroglutamyl aminopeptidase, the yield in each degradation cycle was less than 5%. When the peptide was pretreated with pyroglutamyl aminopeptidase, the yield increased sixfold, indicating that the amino terminus of the peptide contains pyroglutamic acid. The amino-terminal sequence of the medium peptide was similar to a sequence determined for calf pN α1(I) chains (Figure 8). Alignment of residues 3–6 in the chick peptide with residues 10–13 in the calf peptide produced a good correspondence between residues 11–29 in the chick peptide and residues 18–36 in the calf peptide (Figure 8). We conclude, therefore, that the medium peptide is derived from the amino-terminal, precursor-specific region of pro α1(I) chains. For further comparison of the chick peptide to the dermatosparactic peptides, a tryptic fragment of the chick peptide was isolated and sequenced. When the medium peptide was treated with trypsin following the blockage of lysyl side chains with phenyl isothiocyanate and chromatographed on Sephadex G-50, four peaks containing tryptic fragments could be isolated (Figure 9). Amino acid analysis (not shown) showed that the fragment eluting in fractions 67–78 contained hydroxyproline, indicating that this fragment was derived from the collagenase-sensitive portion of the medium peptide. Amino acid sequence analysis showed that the fragment had a collagen-like sequence. On the basis of its identity with the calf

```
Chick:                                  pGlu-Gly-Glu-Glu-Asp-Ile-Gln-Thr-Gly-Ser-Cys-Val-Gln-
                                                    10                                    20
Calf:    pGlu-Glu-Glu-Gly-Gln-Glu-Glu-Gly-Gln-Glu-Glu-Asp-Ile-Pro-Pro-Val-Thr-Cys-Val-Gln

Chick:    Asp-Gly-Leu- X -Tyr-Asn-Asp- X -Asp-Val-Trp- X -Pro-Glu-Pro-Cys-
                                                30                                    40
Calf:    Asp-Gly-Leu-Arg-Tyr-His-Asp-Arg-Asp-Val-Trp-Lys-Pro-Val-Pro-Cys-Gln-Ile-Cys-Val-

Chick:              -Asp-Gly-Ile-Hyp-Gly-Gln-Hyp-Gly-Leu-Hyp-Gly-Pro-Hyp-Gly-Pro-Hyp-
                          110                                   120
Calf:          -Arg-Asp-Gly-Ile-Hyp-Gly-Gln-Pro-Gly-Leu-Hyp-Gly-Pro-Hyp-Gly-Pro-Hyp-
```

FIGURE 8: Partial amino acid sequence of the chick pro α1(I) amino-terminal propeptide. Data from Hörlein et al. (1979) for the calf dermatosparactic peptide are shown for comparison. The numbers indicate the residue number in the calf peptide. The amino-terminal pGlu residue was removed by incubating the medium peptide with pyroglutamyl aminopeptidase.

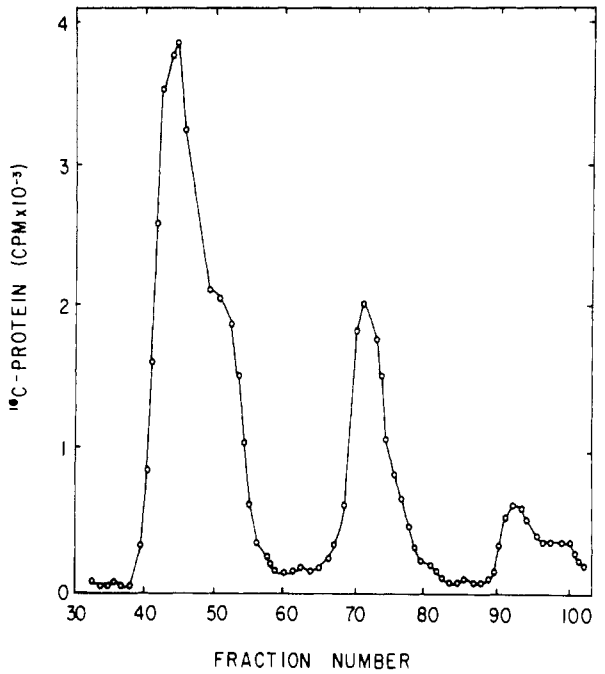

FIGURE 9: Elution profile of tryptic digest of the medium peptide after blockage of lysyl side chains with phenyl isothiocyanate. The fragment eluting in fractions 67–78 was used for amino acid sequencing. The other fragments were not characterized. The column was a 1.5 × 110 cm Sephadex G-50 (superfine) column equilibrated and eluted with 0.2 M ammonium bicarbonate. 1.9-mL fractions were collected. The eluate was assayed by scintillation counting.

dermatosparactic peptide, this tryptic fragment was tentatively placed at the carboxyl end of the medium propeptide (Figure 8). The other tryptic fragments were not characterized further.

Discussion

In a previous report (Olsen et al., 1977), we showed that a chick tendon organ culture system could be used as a convenient source of material for the isolation of relatively large amounts of the carboxyl propeptide of Type I procollagen. Here we report on the isolation and characterization of an amino-terminal propeptide of procollagen using the same system. The medium peptide purified here represented the amino-terminal propeptide of pro α1(I) chains based on several types of data. First, antibodies generated against the peptide reacted with procollagen in a direct binding radioimmune assay. Also, by use of an inhibition assay, the antibodies were found to cross-react with the amino-terminal, collagenase-resistant portion of procollagen but not with the carboxy-terminal propeptide. Second, bacterial collagenase digestion of the medium peptide produced a resistant portion with an apparent molecular weight of 12 500 that on polyacrylamide gel electrophoresis comigrated with the amino-terminal, collagenase-resistant portion of isolated pro α1(I) chains. Collagenase digestion of pro α2 chains did not produce amino-terminal, collagenase-resistant fragments that could be resolved by the gel electrophoresis technique used (unpublished data). Third, tryptic fingerprinting of the amino-terminal, collagenase-resistant portions of pro α1(I) chains and the medium peptide showed similar patterns. Finally, amino acid sequence analysis of the medium peptide showed sequence homology with the amino-terminal region isolated from dermatosparactic calf and sheep pN α1(I) chains.

In the sequence analysis of the medium peptide, we encountered one difficulty. In initial experiments the peptide appeared to have a blocked amino end, and the yields in the degradation cycles were lower than 5%. Pretreatment of the peptide with pyroglutamyl aminopeptidase increased the yields about sixfold. We conclude, therefore, that the blocked amino end was due to the presence of pyroglutamic acid. The sequence data obtained for the peptide showed that the chick peptide was similar to the dermatosparactic calf and sheep propeptides (Hörlein et al., 1979; Rohde et al., 1979). However, there are several notable differences between the calf dermatosparactic peptide and the chick peptide. First, there are several substitutions within the peptide. Second, the first seven residues, pGlu-Glu-Glu-Gly-Gln-Glu-Glu–, of the dermatosparactic peptide are not present in the chick peptide. These amino acid substitutions indicate a larger degree of variability in the amino-terminal propeptide region of pro α1(I) chains than anticipated on the basis of data obtained for the calf and the sheep dermatosparactic peptides (Rohde et al., 1979). The absence of the first seven residues in the chick peptide may represent a deletion of this region in the chick pro α1(I) gene as compared with that of calf and sheep. Evidence that pro α chains are synthesized as prepro α chains that are rapidly processed intracellularly to pro α chains (Graves et al., 1979; Palmiter et al., 1979) suggests the alternative explanation that the seven-residue sequence missing at the amino end of the chick propeptide may be contained within the "pre" sequence of the chick pro α1(I) chains. Replacement of the glycine residue in position 8 of the calf and sheep pro α1(I) chains with glutamine in the chick peptide may have created a new cleavage site for the enzyme responsible for the conversion of prepro α1 chains to pro α1 chains in the chick system.

The question remains whether the medium peptide as isolated here represents the complete amino propeptide of pro α1(I) chains in secreted chick procollagen or whether it is a smaller, partially processed fragment. To answer this question, we attempted to determine the amino-terminal sequence of intact pro α1(I) chains isolated by chromatography on DEAE-cellulose, but several experiments were unsuccessful,

apparently because of a blocked amino-terminal residue. Treatment of the pro α1(I) chain with pyroglutamyl aminopeptidase prior to sequencing gave equivocal results with several different residues being released in each sequencing cycle. This may have been due to the inability of the enzyme to remove the blocking group in the pro α1(I) chain, coupled with the presence of contaminating proteolytic activities in the enzyme preparation which degraded the pro α1(I) chain during the long incubation times. The inability to effectively remove the blocked amino-terminal residue may also be explained by the large size of the pro α1(I) chain (~150 000 daltons) or a limited solubility of the chain in the buffer used for incubation with pyroglutamyl aminopeptidase.

Although a direct proof based on amino acid sequencing is not available, several observations suggest that the medium peptide does indeed represent the intact amino-terminal pro α1(I) propeptide released from procollagen by the action of the procollagen N-protease. First, the collagenase-resistant portion of the medium peptide comigrates on $NaDodSO_4$ slab gels with the amino-terminal, collagenase-resistant region of isolated pro α1(I) chains. Second, the amino propeptide released from pro α1(I) chains when procollagen is incubated with partially purified N-protease (Tuderman et al., 1978) comigrates with the medium peptide isolated here (L. Tuderman and D. M. Pesciotta, unpublished observations). Third, the amino-terminal, collagenase-resistant portions of the medium peptide and of procollagen showed identical behavior as inhibitors in a radioimmune assay using antiserum against the medium peptide (Figure 7).

Recently, Rohde & Timpl (1979) showed that one of two sequential antigenic determinants is located at the amino end of dermatosparactic sheep pN α1(I) chains. The sequence of this antigenic determinant is pGlu-Glu-Glu-Gly-Gln-Glu-Glu–. The absence of this sequence in the chick peptide may explain previous observations that antibodies raised against dermatosparactic sheep pN α1(I) chains cross-react only to a limited extent with chick procollagen (R. Timpl and B. R. Olsen, unpublished observations).

The results of the inhibition assay indicate that most of the antibodies in the antiserum raised against the medium peptide are directed against conformational antigenic determinants in the peptide. After reduction and alkylation of the peptide, no inhibiting activity could be detected in the radioimmune assay. This agrees well with observations by Rohde et al. (1976), who found that most antisera to dermatosparactic sheep pN-collagen or the isolated amino propeptide showed high titers for the native peptide but weak binding with the reduced and alkylated peptide. The amino-terminal, collagenase-resistant portions of the medium peptide and of procollagen showed inhibiting activity in the assay (Figure 7), but the activity did not exceed 60% of that of the medium peptide. This is possibly due to a fraction of antibodies reacting with an antigenic determinant that is present in the medium peptide but removed by treatment with bacterial collagenase. The four- to fivefold difference in concentrations needed to inhibit the reaction with medium peptide and its collagenase-resistant portion may be due to an alteration of conformational determinants on collagenase treatment, leading to weaker binding by the antibody.

The apparent molecular weight of the intact medium peptide was found to be 23 000 by polyacrylamide slab gel electrophoresis in $NaDodSO_4$ when globular proteins were used as molecular weight standards, and the molecular weight of the collagenase-resistant portion of the peptide was found to be 12 500. However, given the abnormal behavior of collagen peptides on polyacrylamide gels, these numbers should be interpreted with caution.

Recent data suggest that the amino-terminal propeptide of procollagen has an important regulatory function and acts as a feedback inhibitor on procollagen synthesis in fibroblasts (Wiestner et al., 1979). The tendon culture system described here should be a convenient source of relatively large amounts of the physiologically processed amino-terminal pro α1(I) propeptide for future studies on its physiological role.

Acknowledgments

The authors express their sincere thanks to C. Condit for technical assistance. Thanks are also due to Dr. R. Timpl for the sheep dermatosparactic pN α1(I)CBO.1-col 1 peptide used.

References

Bates, D. L., Perham, R. N., & Coggins, J. R. (1975) *Anal. Biochem. 68*, 175.

Becker, U., Timpl, R., Helle, O., & Prockop, D. J. (1976) *Biochemistry 15*, 2853.

Becker, U., Helle, O., & Timpl, R. (1977) *FEBS Lett. 73*, 197.

Berg, R. A., & Prockop, D. J. (1973) *Biochemistry 12*, 3395.

Berg, R. A., Kedersha, N. L., & Guzman, N. A. (1979) *J. Biol. Chem. 254*, 3111.

Bhown, A. S., Mole, J. E., Weissinger, A., & Bennett, J. C. (1978) *J. Chromatogr. 148*, 532.

Bonner, W. M., & Laskey, R. A. (1974) *Eur. J. Biochem. 46*, 83.

Bruckner, P., Bächinger, H. P., Timpl, R., & Engel, J. (1978) *Eur. J. Biochem. 90*, 595.

Engel, J., Bruckner, P., Becker, U., Timpl, R., & Rutschmann, B. (1977) *Biochemistry 16*, 4026.

Fessler, J. H., & Fessler, L. I. (1978) *Annu. Rev. Biochem. 47*, 129.

Fietzek, P. P., & Rexrodt, F. W. (1975) *Eur. J. Biochem. 59*, 113.

Furthmayr, H., Timpl, R., Stark, M., Lapière, C. M., & Kühn, K. (1972) *FEBS Lett. 28*, 247.

Graves, P. N., Olsen, B. R., Fietzek, P. P., Monson, J. M., & Prockop, D. J. (1979) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 38*, 620.

Hoffmann, H.-P., Olsen, B. R., Chen, H.-T., & Prockop, D. J. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 4304.

Hörlein, D., Fietzek, P. P., & Kühn, K. (1978) *FEBS Lett. 89*, 279.

Hörlein, D., Fietzek, P. P., Wachter, E., Lapière, C. M., & Kühn, K. (1979) *Eur. J. Biochem. 99*, 31.

King, J., & Laemmli, U. K. (1971) *J. Mol. Biol. 62*, 465.

Lenaers, A., Ansay, M., Nusgens, B. V., & Lapière, C. M. (1971) *Eur. J. Biochem. 23*, 533.

Metz, J., Ebert, W., & Weicker, H. (1971) *Chromatographia 4*, 345.

Nist, C., von der Mark, K., Hay, E. D., Olsen, B. R., Bornstein, P., Ross, R., & Dehm, P. (1975) *J. Cell Biol. 65*, 75.

Olsen, B. R., Hoffmann, H.-P., & Prockop, D. J. (1976) *Arch. Biochem. Biophys. 175*, 341.

Olsen, B. R., Guzman, N. A., Engel, J., Condit, C., & Aase, S. (1977) *Biochemistry 16*, 3030.

Palmiter, R. D., Davidson, J. M., Gagnon, J., Rowe, D. W., & Bornstein, P. (1979) *J. Biol. Chem. 254*, 1433.

Podell, D. N., & Abraham, G. N. (1978) *Biochem. Biophys. Res. Commun. 81*, 176.

Rohde, H., & Timpl, R. (1979) *Biochem. J. 179*, 643.

Rohde, H., Nowack, H., Becker, U., & Timpl, R. (1976) *J. Immunol. Methods 11*, 135.
Rohde, H., Wachter, E., Richter, W. J., Bruckner, P., Helle, O., & Timpl, R. (1979) *Biochem. J. 179*, 631.
Tuderman, L., Kivirikko, K. I., & Prockop, D. J. (1978) *Biochemistry 17*, 2948.
Wiestner, M., Krieg, T., Hörlein, D., Glanville, R. W., Fietzek, P., & Müller, P. K. (1979) *J. Biol. Chem. 254*, 7016.

Structural Significance of the Amino-Terminal Residues of Sperm Whale Myoglobin†

Richard D. DiMarchi, George W. Neireiter,‡ William F. Heath, and Frank R. N. Gurd*

ABSTRACT: Following the development of a nondestructive synthetic procedure for rapid production of des-Val[1]-myoglobin in large quantities, the synthesis of a series of myoglobin derivatives varying in structure and charge in the NH_2-terminal region was accomplished. In comparison to the untreated myoglobin, the des-Val[1]-myoglobin was found to possess at low pH a decreased stability and an increased net positive charge in the pH range 5.5–8.5. While the elevated net positive charge was no longer apparent after removal of the second residue, the instability of the molecule was found to be sharply increased. Substitutions of the first residue, directed toward elucidating its structural importance, included glutamic acid, lysine, and glycine. Addition of any of the three amino acids to the des-Val[1]-myoglobin was found to restore much of the acid stability, with the [Gly[1]]myoglobin appearing nearly identical with the native molecule. All three semisynthetic myoglobins showed potentiometric titration curves characteristic of their respective, substituted residue. Carbamylation of the NH_2 terminal of myoglobin and des-Val[1]-myoglobin yielded two nearly identical molecules in terms of all physical properties examined. Consequently, it was concluded that the first residue primarily serves the function of maintaining the positively charged NH_2 terminus a certain distance away from the beginning of the A helix and from the charge pair interaction of Lys-133 with Glu-6. In addition, through physical measurements of the des-Val[1],Leu[2]-myoglobin prior and subsequent to carbamylation of the NH_2 terminus, it was apparent that the stabilization conferred on the des-Val[1]-myoglobin by the second residue was dependent to a large degree upon the hydrophobic interactions of its side chain.

Previous studies (Wang, 1977; Wang et al., 1978) have revealed that removal of the NH_2-terminal tetradecapeptide of myoglobin results in a molecule possessing less than half of the normal α-helical content and incapable of correctly positioning the heme. Which specific residue or residues are responsible for this loss of structure and through what mechanism they interact with those residues constituting the heme pocket are central to understanding the stabilizing forces of the molecule. Recent studies (DiMarchi et al., 1978b, 1979; Neireiter et al., 1979) have provided the necessary techniques for selective, consecutive removal and substitution of the NH_2-terminal residues of myoglobin by specific degradation and resynthesis. The application of these techniques can make apparent the structural significance of each amino acid residue and show the effects of various substitutions and modifications with respect to stability, conformation, and electrostatic and hydrophobic interactions.

Crystallographic studies of sperm whale aquoferrimyoglobin (Watson, 1969; Takano, 1977) have shown that the NH_2-terminal residues lie predominantly on the surface of the molecule with no direct interactions with those residues exhibiting short interatomic contacts (4.0 Å or less) with the heme. The NH_2-terminal valine residue is predicted to exhibit limited electrostatic interactions with the remaining polypeptide (Friend & Gurd, 1979a,b; Matthew et al., 1979), and both it and the second residue possess no regular, ordered secondary structure (Chou & Fasman, 1974; Takano, 1977). Consequently, the finding of only minor changes in the molecule following NH_2-terminal adduct formation was not surprising (Garner & Gurd, 1975; Gurd et al., 1977; DiMarchi et al., 1979). Nonetheless, the highly conserved nature of the first two residues suggests some type of specific structural importance; glycine is the only other NH_2-terminal residue observed among more than 50 myoglobin species sequenced, while leucine in the second position is invariant (Bogardt, 1978).

Artyukh et al. (1977, 1979) have recently implicated the deprotonation of the NH_2 terminus as a conformational trigger for the A helix and implied that this change is responsible for a perturbation in the spin equilibrium of the heme iron atom. The importance of these findings is emphasized by the known importance of the hemoglobin NH_2 terminals in regulating control of ligand binding (Kilmartin & Rossi-Bernardi, 1969; Benesch et al., 1969; Perutz, 1970; Garner et al., 1975; Matthew et al., 1977). Crystallographic analyses of the hydroxide, cyanide, azide, and fluoride complexes of sperm whale ferrimyoglobin have shown varying degrees of conformational changes in certain regions of the protein, including residues 62–75 of the E helix (Stryer et al., 1964; Watson & Chance, 1966; Bretscher, 1968; Schoenborn, 1969). While no connection had been drawn hitherto between conformational changes and the protonation state of the NH_2 terminus, it is interesting to note the additional minor alterations in several residues involved in important charge interactions neighboring the NH_2 terminus (Schoenborn, 1969; Friend & Gurd, 1979a,b).

The results of the present study clearly illustrate the delicate balance which exists in the myoglobin molecule between the structure of the NH_2-terminal residues and the conformation

† From the Department of Chemistry, Indiana University, Bloomington, Indiana 47405. *Received December 18, 1979.* This work was supported by U.S. Public Health Service Grants HL-14680 and HL-05556. This is the 120th paper in a series dealing with coordination complexes and catalytic properties of proteins and related substances. For the preceding paper see Gurd et al. (1980).
‡ Present address: Mead Johnson and Co., Evansville, IN 47721.

0006-2960/80/0419-2454$01.00/0