Intermediates in the Limited Proteolytic Conversion of Procollagen to Collagen†

Jeffrey M. Davidson, Linda S. G. McEneany, and Paul Bornstein*

ABSTRACT: The conversion of chick bone procollagen to collagen proceeds in a stepwise fashion to produce a limited number of intermediates. Initial proteolytic cleavages remove NH_2-terminal nonhelical extensions and yield an intermediate which remains disulfide-bonded via COOH-terminal extensions. Subsequent stepwise scission of one or two chains of the triple-stranded molecule in its COOH-terminal domain produces intermediates which can only be distinguished after dissociation of the noncovalently bonded α chains. A final cleavage in this region produces the collagen molecule and a disulfide-bonded triple-stranded fragment which represents the COOH-terminal domain. In all likelihood the endopeptidases which effect cleavage in the NH_2- and COOH-terminal regions differ. More than two enzymes may be required for conversion of procollagen to collagen if the nonhelical domains are not released in an en bloc fashion.

The conversion of procollagen to collagen has been reported to occur by scission of nonhelical polypeptides in the NH_2-terminal region of each constituent proα chain (see Bornstein, 1974 for a review). The enzymatic activity responsible for this proteolysis was termed procollagen peptidase (Lapière et al., 1971; Bornstein et al., 1972; Kohn et al., 1974). However, preliminary studies (Bornstein et al., 1975) led us to conclude that disulfide-bonded procollagen, which is larger than the substrate used for the initial characterization of the conversion process, was not converted to collagen by preparations containing procollagen peptidase. We therefore set out to determine whether the conversion of procollagen proceeded by single or multiple enzymatic cleavages in individual proα chains of the precursor. If conversion were multi-step, then chains intermediate in size between proα and α chains should have been generated during the course of collagen biosynthesis. Indeed, several laboratories, including our own, have recently presented preliminary evidence for the existence of intermediates in procollagen conversion (Goldberg et al., 1975; Davidson and Bornstein, 1975; Fessler et al., 1975; Tanzer et al., 1975).

We describe here the results of a study of the course of limited proteolysis of procollagen, as it occurs in organ culture of embryonic chick cranial bone. As part of this study we isolated and partially characterized a class of intermediates which contained intact COOH-terminal nonhelical extensions and which thus demonstrate that there are at least two distinct sites of proteolytic cleavage in each proα chain. In addition, recent information on the extent and nature of the COOH-terminal nonhelical domain in procollagen (Tanzer et al., 1974; Murphy et al., 1975; Byers et al., 1975) has permitted the construction of a model of procollagen conversion which reconciles the data reported from a number of laboratories.

Materials and Methods

Preparation of Procollagen. Cranial bones from 17-day-old chick embryos were placed into Dulbecco-Vogt modified Eagle medium which lacked the amino acid to be used as a radioactive label. The conditions of incubation have been described (Monson and Bornstein, 1973; Monson et al., 1975). The bones were preincubated in a saturated oxygen atmosphere at 37° for 20–60 min to deplete endogenous amino acid pools. Radioactive labeling was performed with L-[2,3-^{3}H]proline (New England Nuclear, 35 Ci/mmol), L-[U-^{14}C]proline (New England Nuclear, 200 Ci/mol), L-[G-^{3}H]tryptophan (New England Nuclear, 7.7 Ci/mmol), or L-[3,5-^{3}H]tyrosine (Amersham, 53 Ci/mmol). In pulse chase experiments with [^{3}H]proline, chase media were supplemented with 12.5 m*M* L-proline.

Proteins were extracted and fractionated by an adaptation of the method of Monson et al. (1975). At the end of the labeling period, the medium was decanted, and the bones were quickly rinsed with ice-cold water and then suspended in cold extraction buffer consisting of 1 M NaCl, 0.05 M Tris-HCl (pH 7.5), 25 m*M* EDTA,[1] 10 m*M* MalNEt, and 1 m*M* Dip-F in an approximate ratio of 1 ml per bone. Homogenization was carried out for 30 sec in a Virtis homogenizer and the homogenate was immediately clarified by centrifugation at 38000 *g* at 2° for 10 min.

Collagenous proteins were precipitated by addition of solid NaCl to the supernatant to a final concentration of 20% (w/v) and the suspension was stirred in the cold for 60 min. The precipitate was collected by centrifugation at 38000 *g* for 10 min and reextracted in an equal volume of extraction buffer for 1 hr at 4°. The insoluble material from this extraction was removed by centrifugation for 20 min at 38000 *g* and the supernatant was stored at −20° for further processing or directly precipitated with Cl_3CCOOH (see below).

DEAE-Cellulose Chromatography. Procollagen preparations extracted from bone were precipitated with 10% Cl_3CCOOH in the presence of 5 μg/ml of pepstatin at 0° for 30 min and insoluble material was collected by centrifu-

† From the Departments of Biochemistry and Medicine, University of Washington, Seattle, Washington 98195. *Received May 29, 1975.* This work was supported by National Institutes of Health Grants AM 11248 and DE 02600. J.M.D. is the recipient of a Research Fellowship (GM-04041) from the National Institutes of Health.

[1] Abbreviations used are: EDTA, ethylenediaminetetraacetic acid; MalNEt, *N*-ethylmaleimide; Dip-F, diisopropyl fluorophosphate; $PhCH_2SO_2F$, phenylmethanesulfonyl fluoride; DATD, *N,N'*-diallytartardiamide.

gation at 38000 *g* for 10 min. The pellet was extracted with 0.25 *M* Tris-HCl (pH 7.5) containing 1 *M* NaCl for 60 min at 4° and clarified by centrifugation. The soluble procollagen fraction was dialyzed against three changes (20 volumes each) of 6 *M* urea and 40 m*M* Tris-HCl containing 0.4 m*M* $PhCH_2SO_2F$, diluted with an equal volume of 6 *M* urea and applied to a column (2.5 × 20 cm) of microgranular DEAE-cellulose (Whatman DE52) previously equilibrated with 2 *M* urea and 0.02 *M* Tris-HCl (pH 7.5) containing 0.2 m*M* $PhCH_2SO_2F$ at 4°. This procedure was required in order to solubilize the majority of radioactive protein. Elution was performed with 100 ml of the equilibration buffer followed by a linear gradient from 0 to 0.2 *M* NaCl over a total volume of 2 l. A flow rate of 150 ml/hr was used.

Analytical Dodecyl Sulfate Polyacrylamide Gel Electrophoresis. Gel electrophoresis was carried out on 5% acrylamide gels (13 × 0.6 cm) as described by Goldberg et al. (1972) except that an equimolar concentration of DATD was substituted for methylenebisacrylamide as a cross-linking agent (Anker, 1970). Samples were prepared in one of two ways: (1) precipitation with 10% Cl_3CCOOH in the presence of 5 μg/ml of pepstatin at 0° for 30 min, followed by centrifugation at 11000 *g* for 20 min at 2° and dehydration of the pellet with successive ethanol, ethanol–ether, and ether washes; (2) dialysis against 1 m*M* ammonium bicarbonate (pH 7.6) followed by lyophilization.

Iodoacetamide (100 m*M*) or dithiothreitol (50 m*M*) was added to samples to achieve oxidizing or reducing conditions, respectively, and dansylated collagen components were added as internal reference standards. Electrophoresis was performed for 4 hr at 15 mA/tube and the gels were subsequently sliced at 1-mm intervals, digested at room temperature with 0.5 ml of 2% periodic acid (Anker, 1970), and counted in 10 ml of 25% Triton X-114 (Sigma) containing 0.3% Omnifluor (New England Nuclear) in xylene.

Slab Gel Dodecyl Sulfate Polyacrylamide Gel Electrophoresis. Samples were electrophoresed in 5% bisacrylamide slabs using an apparatus described by Studier (1973). A discontinuous dodecyl sulfate buffer system (Laemmli, 1970), with 0.5 *M* urea added to all buffers, was employed and samples were treated as described for analytical gels. After electrophoresis, gel slabs were permeated with dimethyl sulfoxide and 2,5-diphenyloxazole for scintillation autography, as described by Bonner and Laskey (1974), dried on Whatman 3MM paper under vacuum and heat, and exposed to RP Royal X-Omat film (Kodak) at −70°.

Immunological Studies. Antibodies to procollagen and acid-extracted proα1 were prepared as described by Nist et al. (1975). The ability of these antibodies to react with procollagen and with intermediates in the conversion to collagen was assayed by the double antibody radioimmune procedure described by von der Mark et al. (1973).

Results

Pulse Chase Experiments. Cranial bones were labeled with [^{3}H]proline for 18 min, transferred to medium with an excess of unlabeled proline, and incubated further for 15, 30, or 45 min. The collagenous proteins were then extracted and the samples analyzed by electrophoresis in dodecyl sulfate acrylamide gels. After an 18-min pulse (Figure 1A) virtually all the radioactivity in an unreduced sample was found in a single, high molecular weight peak with a molecular weight in excess of 400000. After reduction, two bands with apparent molecular weights of approximately 150000

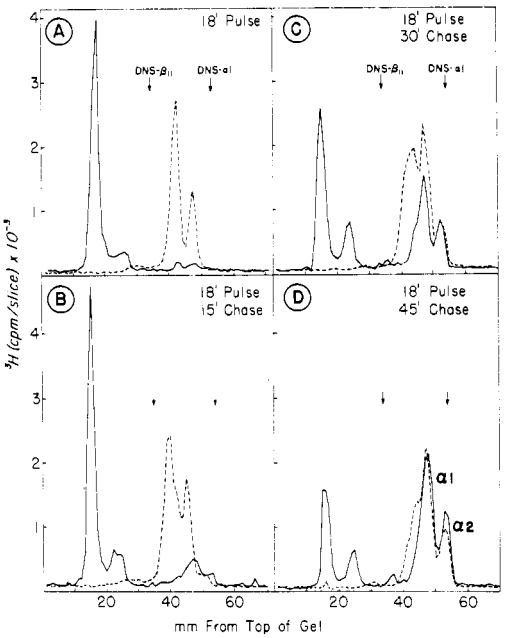

FIGURE 1: Dodecyl sulfate acrylamide gel electrophoresis of procollagen and its conversion products, labeled in pulse-chase experiments, before (—) and after (- - -) reduction with dithiothreitol. Four batches of cranial bones were labeled with [^{3}H]proline for 18 min and either extracted immediately or chased with an excess of unlabeled proline for the indicated times. The position of α1 and α2 are indicated in the solid trace in panel D. The positions of migration of dansylated (DNS) β_{11} and α1, used as internal standards, are indicated by arrows. DNS α1 comigrates with α2 since dansylation increases the mobility of chains in dodecyl sulfate containing buffers.

were obtained. This pattern is consistent with previously published electrophoretograms of procollagen and its component proα1 and proα2 chains (Monson and Bornstein, 1973; Fessler and Fessler, 1974; Goldberg et al., 1975).

With increasing time of chase, the electrophoretic data demonstrated the presence of intermediates in the conversion process (Figure 1). As expected, there was a gradual diminution in the relative amount of intact procollagen and a corresponding increase of two lower molecular weight nonreducible components whose mobilities coincided with those of α1 and α2 chains. In addition, with longer chase periods two high molecular weight reducible components appeared; one migrated between intact procollagen and dansylated β_{11}, the second migrated slightly faster than the dansylated β_{11} component. Bones labeled continuously for 2 hr also contained these intermediates (see below). In the absence of reduction, only a small amount of radioactivity was detected between the positions of proα1 and α1.

Reduction of the samples revealed a more complex pattern. The reduced samples indicated a shift of radioactivity with time from the region of proα to that of α chains (Figure 1). This shift was characterized initially by a broadening of the proα1 region and a concurrent increase of radioactivity in the proα2 region. We concluded from these experiments that conversion of procollagen to collagen was likely to be a multistep process in which some intermediates contained disulfide bonds.

Long-Term Labeling Studies. When cranial bones were continuously labeled for 2 hr with both [^{3}H]tryptophan, an amino acid limited in its distribution to regions of procollagen unique to the precursor, and [^{14}C]proline, a distinctive electrophoretic pattern of extracted collagenous proteins was obtained (Figure 2). Prior to reduction, three high mo-

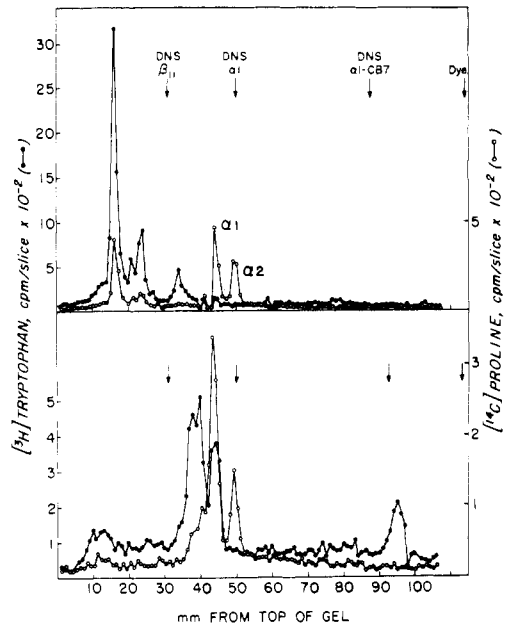

FIGURE 2: Dodecyl sulfate acrylamide gel electrophoresis of collagenous proteins extracted from cranial bones labeled for 2 hr with [^{3}H]tryptophan and [^{14}C]proline. Upper panel, before reduction; lower panel, after reduction. The arrows indicate the position of migration of dansylated β components, α chains, α1-CB7, and the dye marker. α1 and α2 are indicated on the upper panel. Since no radioactivity migrated beyond 60 mm in the unreduced sample, subsequent electrophoretograms of unreduced material (Figures 4–6) were not analyzed beyond this point.

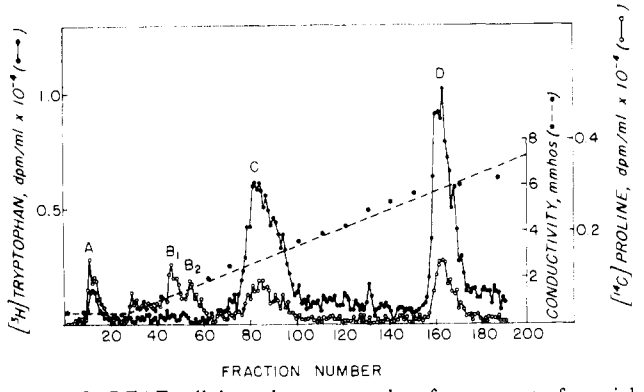

FIGURE 3: DEAE-cellulose chromatography of an extract of cranial bone labeled for 2 hr with [^{3}H]tryptophan and [^{14}C]proline. See Materials and Methods for details.

lecular weight components similar to those seen in the 45-min chase (Figure 1D) were resolved; an inverse correlation between the size of these components and the ratio of [^{3}H]tryptophan to [^{14}C]proline was noted. [^{14}C]Proline-labeled α1 and α2 chains which lacked tryptophan were also observed. After reduction the pattern of radioactivity was complex. Incompletely resolved tryptophan-containing chains migrated in positions ranging from that of proα1 to a position coinciding with the position of α1 (Figure 2, lower panel). An additional tryptophan-containing peptide (migrating near DNS α1-CB7) with a mobility corresponding to that of the reduced form of the procollagen-derived fragment isolated from cranial bone culture medium (Murphy et al., 1975) was also observed.

When an extract of cranial bones, labeled for 2 hr with [^{3}H]tryptophan and [^{14}C]proline, was chromatographed on DEAE-cellulose, four major components were resolved

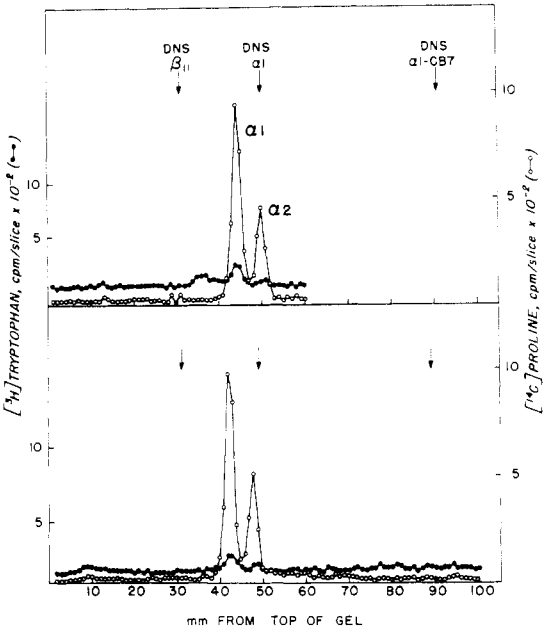

FIGURE 4: Dodecyl sulfate acrylamide gel electrophoresis of peak A, Figure 3. Upper panel, before reduction; lower panel, after reduction. The arrows indicate the positions of migration of dansylated β_{11}, α1, and α1-CB7. The positions of α1 and α2 are indicated on the upper panel. Note that DNS α1 comigrates with α2.

(Figure 3). Two fractions (A and B) were enriched in proline relative to tryptophan. Each of the peaks was isolated, concentrated, and examined by dodecyl sulfate gel electrophoresis. As seen in Figure 4, the pattern for peak A revealed [^{14}C]proline-labeled α1 and α2 chains. No change in pattern occurred after reduction and no lower molecular weight peptides were released. No significant differences in the electrophoretic patterns of radioactive chains were noted with peaks A, B_1, and B_2. The reasons for the chromatographic separation of these peaks (Figure 3) are not known but the findings were consistent.

Electrophoresis of peak C (see Figure 3) prior to reduction revealed three tryptophan-rich components as well as [^{14}C]proline-labeled α1 and α2 chains (Figure 5, upper panel). The former correspond in position of migration to the high molecular weight fractions noted in the starting material but with a diminished proportion of the largest species (see Figure 2, upper panel). As before, there was an inverse correlation between the mobilities of these high molecular weight components and the ratio of [^{3}H]tryptophan to [^{14}C]proline. In addition to the [^{14}C]proline-labeled α1 and α2 chains released by denaturation, reduction of peak C yielded two tryptophan-containing chains with mobilities similar to those of proα chains and a low molecular weight chain which migrated ahead of dansylated α1-CB7 (Figure 5, lower panel). The latter again corresponded to a similar peptide noted following reduction of the original extract (see Figure 2, lower panel).

Electrophoresis of peak D, examined before and after reduction (Figure 6), yielded patterns identical with those of intact procollagen (Monson and Bornstein, 1973; Murphy et al., 1975; see Figure 1A). Prior to reduction nearly all the radioactivity was in a high molecular weight peak; following reduction, bands migrating in the positions of proα1 and proα2 were noted. Only trace quantities of labeled material were seen in the region of the dansylated α1-CB7 marker. The resolution of these dodecyl sulfate acrylamide-

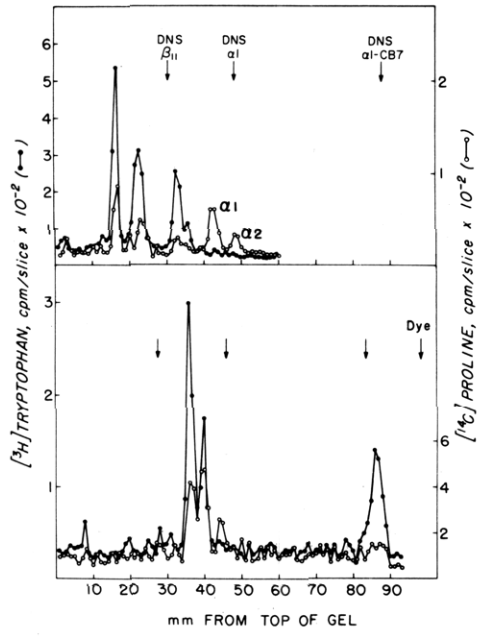

FIGURE 5: Dodecyl sulfate acrylamide gel electrophoresis of peak C, Figure 3. Upper panel, before reduction; lower panel, after reduction. The arrows indicate the positions of migration of dansylated β_{11}, $\alpha 1$, and $\alpha 1$-CB7. The positions of $\alpha 1$ and $\alpha 2$ are indicated on the upper panel. Note that DNS $\alpha 1$ comigrates with $\alpha 2$.

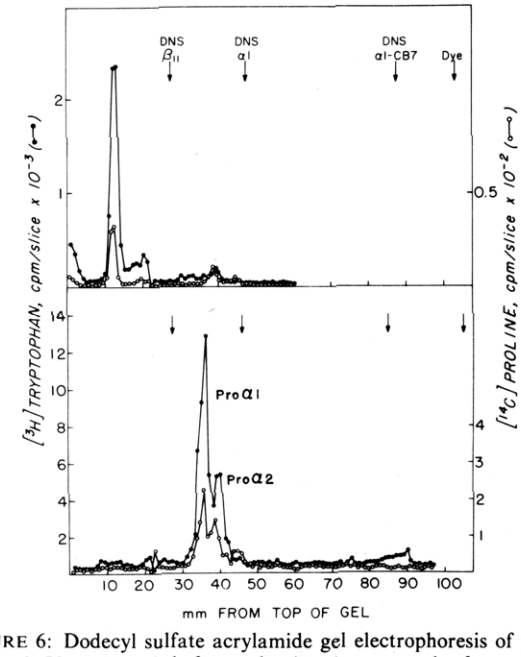

FIGURE 6: Dodecyl sulfate acrylamide gel electrophoresis of peak D, Figure 3. Upper panel, before reduction; lower panel, after reduction. The positions of pro$\alpha 1$ and pro$\alpha 2$ are indicated on the lower panel. The arrows indicate the positions of migration of dansylated β_{11}, $\alpha 1$, and $\alpha 1$-CB7.

DATD disc gels was insufficient to reproducibly distinguish relatively small molecular weight differences which would be expected as a result of cleavage of nonhelical peptide extensions from pro$\alpha 1$ and pro$\alpha 2$ (compare with Figure 5, lower panel). However, resolution was achieved by slab gel electrophoresis (see below).

In another experiment, cranial bones were labeled for 2 hr with [^{3}H]tyrosine. Since tyrosine exists both in collagen and in the precursor-specific extensions of procollagen (Murphy et al., 1975; Monson et al., 1975), procollagen as

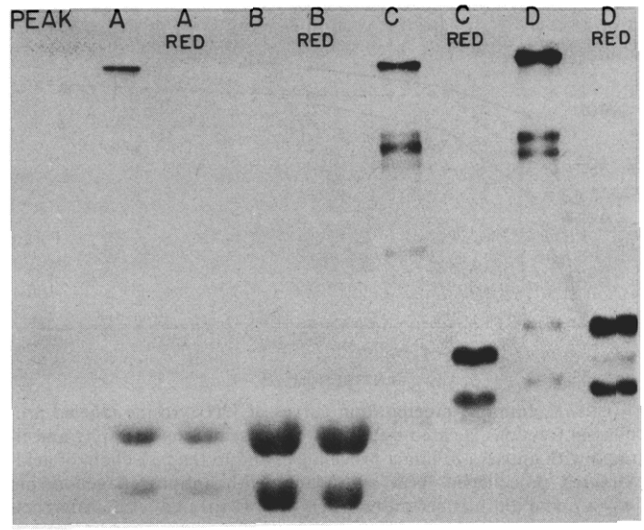

FIGURE 7: Composite fluorescent autoradiograms of [^{3}H]tyrosine-labeled collagenous proteins obtained by DEAE-cellulose chromatography and electrophoresed on dodecyl sulfate acrylamide slab gels. Approximately 4000 dpm were applied to each slot of the gel. The autoradiogram was exposed for 48 hr at −70°. In this experiment the low molecular weight fragment obtained following reduction of peak C and the dye marker were electrophoresed off the bottom of the gel. RED, reduced.

well as its derivatives were potentially identifiable in this material. Collagenous proteins were extracted and subjected to partial fractionation by DEAE-cellulose chromatography under conditions identical with those used to generate the data for the double label experiment described above. Peaks A, B, C, and D were then analyzed by electrophoresis and radiofluorography on acrylamide slab gels (Figure 7).

Peaks A and B contained only α chains (compare with Figure 4). The nature of the high molecular weight band in peak A which appeared to enter the gel following reduction is unknown. The small differences in mobility of α chains in peaks A and B are probably a function of differences in sample load. Peak C, before reduction, contained three high molecular weight bands and two faint bands corresponding to $\alpha 1$ and $\alpha 2$ (compare with Figure 5). Following reduction, two bands which migrated more rapidly than proα chains (as seen in reduced peak D material), but more slowly than α chains, were noted. Peak D consisted largely of a high molecular weight band which after reduction yielded pro$\alpha 1$ and pro$\alpha 2$ chains. However, bands of intermediate mobility as well as pro$\alpha 1$ and pro$\alpha 2$ bands were seen in the unreduced material. The proα chains and apparent dimers observed in peak D prior to reduction may reflect a fraction of native procollagen which had not yet become entirely disulfide bonded at the time of extraction (Fessler and Fessler, 1974; Harwood et al., 1975).

Immunologic Studies. Antisera against intact procollagen or the p$\alpha 1$ chain[2] of acid-extracted procollagen were reacted with material in peaks C and D. The radioimmunoassays (Figure 8) indicated that antibodies to intact procollagen reacted with both peak C and peak D material whereas antibodies to acid-extracted p$\alpha 1$ chain reacted only with peak D material. Since recent studies indicate that acid-extracted procollagen contains NH_2-terminal exten-

[2] In accord with a recently proposed nomenclature (Martin et al., 1975) chains intermediate in length between proα and α chains are termed pα chains.

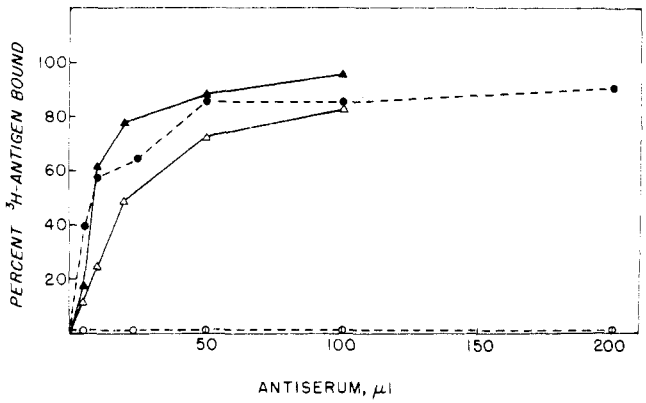

FIGURE 8: Immune precipitation curves of [^{3}H]tyrosine labeled procollagen fractions isolated by DEAE-cellulose chromatography and titrated with antisera to intact procollagen and to the pα1 chain of acid-extracted procollagen: peak D titrated with antiprocollagen serum (▲—▲), and antipα1 serum (△—△); peak C titrated with antiprocollagen serum (●- - -●), and antipα1 serum (○- - -○). Peaks C and D were obtained from a chromatogram similar to that illustrated in Figure 3. Antiprocollagen serum was diluted 1:50 and antipα1 serum was diluted to 1:10. Points represent the average of duplicate assays using about 3000 dpm of antigen.

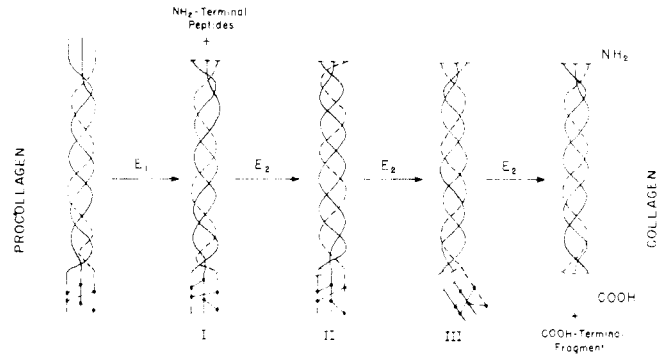

FIGURE 9: A proposed model of conversion of procollagen to collagen. Intact procollagen with a molecular weight of approximately 450000, and consisting of two proα1 chains (—) and one proα2 chain (- - -), contains both NH_2- and COOH-terminal nonhelical extensions. This molecule is first cleaved by enzyme E_1 to a 390000-dalton trimer (intermediate I) lacking NH_2-terminal peptides. Sites of cleavage are indicated by bars perpendicular to each chain. Subsequent stepwise scission in a disulfide-bonded COOH-terminal domain by enzyme E_2 produces intermediates II and III and eventually yields a 285000 dalton collagen molecule with release of a 105000-dalton disulfide-bonded fragment. This model presumes cleavage of a minimum of six peptide bonds per procollagen molecule. Interchain disulfide bonds are indicated (●—●).

sions but lacks a large COOH-terminal nonhelical domain (Byers et al., 1975) these results are consistent with the supposition that procollagen intermediates in peak C have lost NH_2-terminal nonhelical extensions.

Discussion

We interpret the data reported in this paper in the following manner. The procollagen molecule contains both NH_2- and COOH-terminal nonhelical domains which are not present in collagen (Tanzer et al., 1974; Murphy et al., 1975; Byers et al., 1975). Both of these regions must therefore be excised during conversion of procollagen to collagen. We suggest that, in the normal conversion process in bone, cleavage occurs first in the NH_2-terminal region of procollagen (Figure 9) because non-disulfide-bonded chains larger than α chains have not been identified in pulse chase or long term label experiments. Since interchain disulfide bonds occur exclusively in the COOH-terminal domain (Byers et al., 1975), prior removal of this region would have given rise to such chains. On the other hand, a class of disulfide-bonded intermediates of high molecular weight but differing in elution from DEAE-cellulose from procollagen was identified (peak C, Figure 3). This observation is consistent with the suggestion that such intermediates (I, II, and III, Figure 9) lack NH_2-terminal but retain COOH-terminal extensions.

We have not identified the NH_2-terminal peptides released during the physiological conversion of procollagen and therefore cannot ascertain the number of cleavages in each chain required to effect release of this domain. However, the finding that the enzyme procollagen peptidase produces en bloc excision of peptides from dermatosparactic procollagen (Kohn et al., 1974) suggests that only a single cleavage may be involved. Very possibly intermediates with cleavage of only one or two NH_2-terminal extension peptides exist.

The intermediate produced by removal of the NH_2-terminal extensions of procollagen remains disulfide bonded via COOH-terminal extensions (intermediate I, Figure 9; Byers et al., 1975) but elutes earlier than procollagen from DEAE-cellulose (peak C in Figure 3). The immunologic data reported in this work also support the existence of an intermediate which lacks NH_2-terminal extensions. The absence of NH_2-terminal extensions from an intermediate in the conversion of procollagen to collagen is further substantiated by the observation that purified tadpole collagenase cleaved peak C into an NH_2-terminal fragment similar in size to that obtained from collagen and a COOH-terminal fragment which was equivalent in size to that obtained from intact procollagen.[3]

The cleavage of the COOH-terminal domain appears to occur in a stepwise fashion with scission in first one, then in a second, and finally in a third chain of the intermediate (Figure 9). The order of cleavage will require further clarification. Peak C obtained by DEAE-cellulose chromatography (Figure 3) contained all three forms of the intermediate (I, II, and III, Figure 9). This is to be expected since the ion-exchange properties of the undenatured proteins are not likely to be very different. We suggest that intermediates with scission of one or two chains represent the high molecular weight forms migrating on either side of the dansylated β_{11} marker (Figure 5, upper panel) and also seen in slot C (Figure 7). Such intermediates would, after reduction, give rise to a mixture of three chain types: (a) pα chains intermediate in size between proα and α chains and lacking NH_2-terminal extensions, (b) α chains, and (c) the COOH-terminal extension peptide. These peptides are shown in Figure 5, lower panel, and in Figure 7, slot C, Red. The heterogeneity of the intermediate band (Figure 7, peak C) may reflect the presence of components with either two pα1 or one pα1 and one pα2 chains.

As scission of two chains results in a molecule with two noncovalently bonded α chains, denaturation would lead to a component with one pα chain linked by disulfide bonds to two COOH-terminal extension peptides (intermediate III, Figure 9). Since α chains are rich in proline but lack tryptophan, their progressive loss following denaturation accounts for the increasing [^{3}H]tryptophan to [^{14}C]proline ratio of

[3] Davidson, J. M., McEneany, L. S. G., and Bornstein, P., manuscript in preparation.

components with increasing mobilities on dodecyl sulfate gels (see Figures 2 and 5). Intermediates I, II, and III have been isolated and characterized further following preparative and two-dimensional acrylamide gel electrophoresis.[3]

The final cleavage at the COOH-terminus of the procollagen derivative yields collagen and a disulfide-linked COOH-terminal extension fragment (Figure 9). This fragment has recently been isolated from the medium of cultured bone and characterized (Murphy et al., 1975). Presumably, the similar but somewhat smaller fragments isolated by bacterial collagenase digestion of procollagen secreted by fibroblasts in culture (Sherr et al., 1973; Dehm et al., 1974) are also derived from the COOH-terminal domain.

Our best estimate of the molecular weight of an intact proα chain is approximately 150000, resulting from the addition of an NH_2-terminal extension of 20000 and a COOH-terminal extension of 35000 to an α chain of 95000 molecular weight (Monson et al., 1975; von der Mark and Bornstein, 1973; Murphy et al., 1975). The molecular weight of procollagen would therefore be 450000 and that of the intermediate, 390000. Intermediates which have sustained one or two cleavages in the COOH-terminal region would yield (following dissociation of the noncovalently linked α chains) components with molecular weights of approximately 295000 and 200000, respectively. These molecular weights are roughly consistent with the mobility of the bands observed in dodecyl sulfate gels (Figures 5 and 7).

Although it is possible that the same enzymatic activity is responsible for cleavage of both NH_2- and COOH-terminal extensions from procollagen, current evidence favors the existence of two separate enzymes. In dermatosparaxis, the procollagen derivative which accumulates in skin lacks interchain disulfide bonds (Lenaers et al., 1971) but contains NH_2-terminal extensions (Stark et al., 1971). We therefore conclude that, like acid-extracted procollagen, this derivative lacks the COOH-terminal domain. Presumably, the enzymatic activity, procollagen peptidase, which is defective in dermatosparaxis (Kohn et al., 1974), is responsible for cleavage of the NH_2-terminal extensions. Under these circumstances the enzyme or enzymes which cleave the COOH-terminal extensions may do so even though normally, in bone, the removal of the COOH-terminal domain follows cleavage of the NH_2-terminal extensions. However, if the procollagen derivatives identified in normal bovine and rat skin (Veis et al., 1973; Byers et al., 1974) and in the medium of cultured human skin fibroblasts (Lichtenstein et al., 1975) represent physiological intermediates, procollagen conversion in skin may first occur in the COOH- and then in the NH_2-terminal region, or may occur randomly.

Stepwise scission of procollagen chains has also been proposed by Goldberg and his associates (Goldberg and Sherr, 1973; Goldberg et al., 1975) and indeed in many respects our data are consistent with those published by these workers. It also seems likely that the disulfide-linked triple-stranded intermediate identified by Fessler and his associates (Fessler et al., 1975) is identical with intermediate I (Figure 9). Our model differs from that proposed by Goldberg in that (a) proteolytic steps at both ends of the procollagen molecule are invoked, (b) the disulfide-bonded extensions are placed in the COOH-terminal domain, and (c) the initial intermediate produced contains disulfide bonds linking all three chains. The enzymatic activity identified by Goldberg et al. (1975) in the medium of cultured 3T6 cells may therefore be responsible for removal of the COOH-terminal domain.

Much of the difficulty in establishing this scheme for conversion of procollagen to collagen may be ascribed to the presence of two different chains (proα1 and proα2) in Type I procollagen. We have recently performed analogous experiments with embryonic chick sternal cartilage in culture, a tissue which synthesizes a collagen containing three identical chains (Miller and Matukas, 1969). The results, which are simpler to interpret, indicate a similar process of conversion for this type II procollagen.[3]

Added in Proof

Subsequent to the submission of this manuscript, another group (Morris et al., 1975) has independently published evidence for at least one procollagen intermediate.

Acknowledgment

We thank Felicia Arguelles for her skillful technical assistance. Carol Brand assisted in the early phases of this work. Pepstatin was obtained through the generosity of Dr. T. Aoyagi, Institute of Microbial Chemistry, Tokyo, Japan.

References

Anker, H. S. (1970), *FEBS Lett. 7,* 293

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

Bornstein, P. (1974), *Annu. Rev. Biochem. 143,* 467.

Bornstein, P., Davidson, J. M., and Monson, J. M. (1975), *Cold Spring Harbor Symp.* (in press).

Bornstein, P., Ehrlich, H. P., and Wyke, A. W. (1972), *Science 175,* 544.

Byers, P. H., Click, E. M., Harper, E., and Bornstein, P. (1975), *Proc. Natl. Acad. Sci. U.S.A. 72,* 3009.

Byers, P. H., McKenney, K. H., Lichtenstein, J. R., and Martin, G. R. (1974), *Biochemistry 13,* 5243.

Davidson, J. M., and Bornstein, P. (1975), *Fed. Proc., Fed. Am. Soc. Exp. Biol. 35,* 562.

Dehm, P., Olsen, B. R., and Prockop, D. J. (1974), *Eur. J. Biochem. 46,* 107.

Fessler, L. I., and Fessler, J. H. (1974), *J. Biol. Chem. 249,* 7637.

Fessler, L. I., Morris, N. P., and Fessler, J. H. (1975), *Fed. Proc., Fed. Am. Soc. Exp. Biol. 34,* 562.

Goldberg, B., Epstein, Jr., E. H., and Sherr, C. J. (1972), *Proc. Natl. Acad. Sci. U.S.A. 69,* 3655.

Goldberg, B., and Sherr, C. J. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 361.

Goldberg, B., Taubman, M. B., and Radin, A. (1975), *Cell 4,* 45.

Harwood, R., Bhalla, A. K., Grant, M. E., and Jackson, D. S. (1975), *Biochem. J. 148,* 129.

Kohn, L. D., Isersky, C., Zupnik, J., Lenaers, A., Lee, G., and Lapière, C. M. (1974), *Proc. Natl. Acad. Sci. U.S.A. 71,* 40.

Laemmli, U. K. (1970), *Nature (London) 227,* 680.

Lapière, C. M., Lenaers, A., and Kohn, L. D. (1971), *Proc. Natl. Acad. Sci. U.S.A. 68,* 3054.

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

Lichtenstein, J. R., Byers, P. H., Smith, B. D., and Martin, G. R. (1975), *Biochemistry 14,* 1589.

Martin, G. R., Byers, P. H., and Piez, K. A. (1975), *Adv. Enzymol. Relat. Areas Mol. Biol. 42,* 167.

Miller, E. J., and Matukas, V. J. (1969), *Proc. Natl. Acad. Sci. U.S.A. 64,* 1264.

Monson, J. M., and Bornstein, P. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 3521.
Monson, J. M., Click, E. M., and Bornstein, P. (1975), *Biochemistry 14,* 4088.
Morris, N. P., Fessler, L. I., Weinstock, A., and Fessler, J. H. (1975), *J. Biol. Chem. 250,* 5719.
Murphy, W. H., von der Mark, K., McEneany, L. S. G., and Bornstein, P. (1975), *Biochemistry 14,* 3243.
Nist, C., von der Mark, K., Hay, E., Olson, B. R. Bornstein, P., Ross, R., and Dehm, P. (1975), *J. Cell Biol. 65,* 75.
Sherr, C. J., Taubman, M. B., and Goldberg, B. (1973), *J. Biol. Chem. 248,* 7033.
Stark, M., Lenaers, A., Lapière, C. M., and Kühn, K. (1971), *FEBS Lett. 18,* 225.
Studier, F. W. (1973), *J. Mol. Biol. 79,* 237.
Tanzer, M. L., Church, R. L., Yeager, J. A., and Park, E. D. (1975), in Extracellular Matrix Influences on Gene Expression, Slavkin, H. C., and Greulich, R. C., Ed., New York, N.Y., Academic Press, p 785.
Tanzer, M. L., Church, R. L., Yeager, J. A., Wampler, D. E., and Park, E. D. (1974), *Proc. Natl. Acad. Sci. U.S.A. 71,* 3009.
Veis, A., Anesey, J., Yuan, L., and Levy, S. J. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 1464.
von der Mark, K., and Bornstein, P. (1973), *J. Biol. Chem. 248,* 2285.
von der Mark, K., Click, E. M., and Bornstein, P. (1973), *Arch. Biochem. Biophys. 156,* 356.

Modification of Arginine and Lysine in Proteins with 2,4-Pentanedione†

H. F. Gilbert, III, and Marion H. O'Leary*

ABSTRACT: Primary amines react with 2,4-pentanedione at pH 6–9 to form enamines, *N*-alkyl-4-amino-3-penten-2-ones. The latter compounds readily regenerate the primary amine at low pH or on treatment with hydroxylamine. Guanidine and substituted guanidines react with 2,4-pentanedione to form *N*-substituted 2-amino-4,6-dimethylpyrimidines at a rate which is lower by at least a factor of 20 than the rate of reaction of 2,4-pentanedione with primary amines. Selective modification of lysine and arginine side chains in proteins can readily be achieved with 2,4-pentanedione. Modification of lysine is favored by reaction at pH 7 or for short reaction times at pH 9. Selective modification of arginine is achieved by reaction with 2,4-pentanedione for long times at pH 9, followed by treatment of the protein with hydroxylamine. The extent of modification of lysine and arginine side chains can readily be measured spectrophotometrically. Modification of lysozyme with 2,4-pentanedione at pH 7 results in modification of 3.8 lysine residues and less than 0.4 arginine residue in 24 hr. Modification of lysozyme with 2,4-pentanedione at pH 9 results in modification of 4 lysine residues and 4.5 arginine residues in 100 hr. Treatment of this modified protein with hydroxylamine regenerated the modified lysine residues but caused no change in the modified arginine residues. One arginine residue seems to be essential for the catalytic activity of the enzyme.

In spite of the availability of a variety of reagents for the modification of lysine residues in proteins and a smaller number of reagents for the modification of arginine (Means and Feeney, 1971; Glazer, 1970; Cohen, 1968, 1970), the search for new reagents for the modification of these amino acids continues, with a view toward selective modification and ease of analysis of extent of modification.

The most common modifications of lysine residues in proteins are acylations with anhydrides and arylations with substituted nitrobenzenes (Means and Feeney, 1971). Neither type of modification is entirely specific, because a number of other protein nucleophiles are capable of reacting with these reagents. None of these reactions are at the same time readily reversible and readily quantitatable. In the present paper we show that 2,4-pentanedione reacts rapidly with the side chain amino groups of lysine residues in proteins. No other side chains are rapidly modified, and the extent of modification is readily quantitated spectrophotometrically. Modification can be reversed at neutral pH or by treatment with hydroxylamine.

The modification of the side chains of arginine residues in proteins is more difficult than the modification of other side chain functional groups. 1,2-Dicarbonyl compounds have commonly been used (Patthy and Smith, 1975; Yankeelov, 1972; Takahashi, 1968; Signor et al., 1971) but such modifications often lead to more than one product, produce irreversible modification of lysine side chains, and can be quantitated only by amino acid analysis. Reaction of guanidines with 1,3-dicarbonyl compounds produces pyrimidines in high yield, but in spite of the widespread use of this reaction for the synthesis of pyrimidines (Brown, 1962), it has been little used for the modification of proteins (Beller et al., 1968). In this paper we show that reaction of 2,4-pentanedione with arginine residues of proteins leads to formation of a stable pyrimidine whose presence can be detected spectrophotometrically. The modification of lysine side chains, which also occurs when proteins are treated with this reagent, can easily be reversed following protein modification by treatment with hydroxylamine. The modification of hen egg white lysozyme by 2,4-pentanedione is presented to illustrate the use of this reagent for modification of both lysine and arginine.

† From the Department of Chemistry, University of Wisconsin, Madison, Wisconsin 53706. *Received June 30, 1975.* Supported by Grant NS-07657 from the National Institutes of Health.