The Disulphide-Bonded Nature of Procollagen and the Role of the Extension Peptides in the Assembly of the Molecule

By RICHARD HARWOOD,*† ANTHONY H. MERRY,*§ DAVID E. WOOLLEY,†
MICHAEL E. GRANT* and DAVID S. JACKSON*

*Department of Medical Biochemistry, The Medical School, University of Manchester,
Oxford Road, Manchester M13 9PT, U.K., and †Department of Medicine,
University Hospital of South Manchester, Nell Lane, Manchester M20 8LR, U.K.

(Received 20 July 1976)

1. The molecular weights of chick tendon and cartilage procollagens, and their constituent polypeptides, were determined by gel filtration and gel electrophoresis. The values obtained are in good agreement and indicate that the mol. wts. of the secreted procollagens (types I and II) and their individual pro- α -chains are of the order of 405 000–445 000 and 137 000–145 000 respectively. 2. Digestion of tendon procollagen with human rheumatoid synovial collagenase gave products consistent with the presence of large non-helical peptide extensions at both *N*- and *C*-termini. Electrophoretic analysis gave apparent mol. wts. of 17 500 and 36 000 for the respective *N*- and *C*-terminal extensions of pro- α 1(I)- and pro- α 2-chains, and inter-chain disulphide bonds were restricted to the *C*-terminal location. 3. During the biosynthesis of procollagen by tendon and cartilage cells a close correlation was observed between the extent of inter-chain disulphide bonding and the proportion of procollagen polypeptides having a triple-helical conformation. These processes appeared to commence in the rough endoplasmic reticulum and be completed in the smooth endoplasmic reticulum, but the rate at which they occur in cartilage cells is markedly slower than that found in tendon cells. 4. When the intracellular [14 C]procollagen polypeptides present in the rough-endoplasmic-reticulum fractions of tendon and cartilage cells were analysed under non-reducing conditions on agarose/polyacrylamide composite gels, no significant pools of dimeric intermediates were detected. 5. In both cell types, inter-chain disulphide-bond formation occurred even when hydroxylation, and hence triple-helix formation, was inhibited. The presence of pro- α 1- and pro- α 2-components in a ratio of 2:1 in the disulphide-linked unhydroxylated procollagen isolated from tendon cells demonstrated that correct chain association occurs in the absence of hydroxylation. This observation is consistent with a model for the assembly of pro- γ 112-chains in which the recognition and selection of pro- α 1- and pro- α 2-chains in a 2:1 ratio are directed by the non-helical *C*-terminal extension peptides of tendon procollagen.

The very slow and incomplete renaturation *in vitro* of α -chains derived from extracellular collagen (for review see Kühn, 1969) led to the proposal by Speakman (1971) that the α -chains, as initially synthesized, might have peptide extensions. It was postulated that these extensions or 'registration peptides' might serve to bring the three α chains into a conformation that would allow rapid assembly of the triple helix. Subsequently, much evidence has accumulated to indicate the existence of a precursor of this type (called procollagen) during collagen biosynthesis in a wide variety of tissues (for reviews see Bornstein, 1974; Grant & Jackson, 1976). Best

characterized are the precursors of type I and type II collagens, but estimates of their mol. wts. have varied from approx. 350 000 to 600 000 (see Bornstein, 1974). This variation may in part be explained by the different analytical methods used, for most laboratories consider the types I and II procollagen molecules to comprise three pro- α chains linked by inter-chain disulphide bonds. Initially it was thought that these disulphide linkages were located in *N*-terminal extension peptides (Dehm *et al.*, 1972). However, subsequent studies on the synthesis of types I and II procollagens have demonstrated that these precursor molecules contain additional sequences at both *N*- and *C*-termini and that the pro- α -chains are covalently linked only between the *C*-terminal regions (Tanzer *et al.*, 1974; Byers *et al.*, 1975; Fessler *et al.*, 1975; Merry *et al.*, 1976). These extension peptides appear

† Present address: Department of Biochemistry, University College, P.O. Box 78, Cardiff CF1 1XL, U.K.

§ Present address: Regional Blood Transfusion Service, Roby Street, Manchester M1 3BP, U.K.

to play an important role in the association of the three polypeptide chains and the subsequent assembly of the collagen triple helix (for reviews see Grant & Jackson, 1976; Prockop *et al.*, 1976). In addition, these non-helical sequences confer solubility under physiological conditions on the procollagen molecule, and the formation of collagen fibrils requires their removal by extracellular peptidase action.

In the present paper the molecular sizes of the procollagens secreted by embryonic-chick tendon and cartilage cells have been examined by three different techniques, and C-terminal extensions, not previously detected in tendon procollagen preparations, have been partially characterized. The relationship between inter-chain disulphide bonding and triple-helix formation has been further investigated and the intracellular location of these events determined. Evidence is also presented that under conditions where hydroxylation, and hence helix formation, are inhibited (Prockop *et al.*, 1976) correct chain association occurs with subsequent formation of inter-chain disulphide bonds.

Experimental

Materials

Fertile eggs obtained from Mytholmroyd Hatcheries, Hebden Bridge, W. Yorks., U.K., were incubated for 17 days at 37°C in a moist atmosphere in a Westernette Egg Incubator (Western Incubators Ltd., Chelmsford, Essex, U.K.). [^{14}C]Proline (290 mCi/mmol), [^3H]proline (10 Ci/mmol) and [^3H]tryptophan (11 Ci/mmol) were purchased from The Radiochemical Centre, Amersham, Bucks., U.K. [^{35}S]Cystine (49 Ci/mmol) was obtained from New England Nuclear, Boston, MA, U.S.A. Iodoacetamide, phenylmethane sulphonyl fluoride, *N*-ethylmaleimide, β -aminopropionitrile fumarate, Coomassie Brilliant Blue and α -chymotrypsin (EC 3.4.21.1) were obtained from Sigma (London) Chemical Co., London S.W.6, U.K. Acrylamide, bisacrylamide and $\alpha\alpha'$ -bipyridyl were purchased from Koch-Light Laboratories Ltd., Colnbrook, Bucks., U.K., and Seakem agarose was obtained from Micro-Bio Laboratories Ltd., 46 Pembridge Road, London W11, U.K. Bio-Gel A-15m, A-5m and A-1.5m (all 200–400 mesh) were obtained from Bio-Rad Laboratories Ltd., Bromley, Kent, U.K. CNBr peptides were generously provided by Dr. C. A. Shuttleworth of this Department.

Tropocollagen was prepared by extraction of rat tail tendons with 0.2M-acetic acid followed by salt-precipitation procedures as described by Jackson & Cleary (1967). [^{14}C]Proline-labelled collagen α -chains were prepared by incubating approx. 200 cranial bones from 17-day chick embryos with 10 μCi of [^{14}C]proline for 3 h at 37°C in 50 ml of modified Krebs medium (Dehm & Prockop, 1971) containing

β -aminopropionitrile fumarate (64 $\mu\text{g}/\text{ml}$). The labelled calvaria were homogenized in 0.2M-acetic acid and extracted for 18 h at 4°C. The extract was exhaustively dialysed against 0.2M-acetic acid and the collagen precipitated by the addition of solid NaCl to 5% (w/v). The collagen was collected by centrifugation, redissolved in 0.1M-acetic acid and dialysed against 4 litres of 0.1M-acetic acid at 4°C before freeze-drying. All other reagents and materials were obtained from sources previously described (Harwood *et al.*, 1974b, 1975a).

Incubation of tendon and cartilage cells and isolation of procollagen

Cells were isolated from the leg tendons and sterna of 17-day chick embryos as described previously (Harwood *et al.*, 1974c, 1975a). Tendon cells were incubated in modified Krebs medium at a concentration of 10^7 cells/ml and cartilage cells were incubated at a similar concentration in the above medium supplemented with 10% (v/v) foetal calf serum. Cells were preincubated for 20 min at 37°C, then the appropriate radioactively labelled amino acid was added and the incubation continued for a further 2 h. Cells and medium were separated by centrifugation at 1800g for 10 min and the medium made 1.7M in phenylmethanesulphonyl fluoride, 10mM in *N*-ethylmaleimide and 25mM in EDTA and dialysed at 4°C overnight against 2 litres of 0.1M-Tris/HCl buffer (pH 7.4) containing 0.4M-NaCl and the proteinase inhibitors at the concentrations specified above. Procollagen in the medium was precipitated by slowly adding 176 mg of $(\text{NH}_4)_2\text{SO}_4/\text{ml}$ in the cold and then centrifuging the sample at 18000g for 1 h (Dehm & Prockop, 1973; Schofield *et al.*, 1974). The pellet was redissolved in the NaCl/Tris/HCl buffer containing the inhibitors and the $(\text{NH}_4)_2\text{SO}_4$ precipitation procedure repeated to yield a pellet containing radioactively labelled procollagen.

In experiments designed to examine the state of disulphide bonding in unhydroxylated procollagen (procollagen), cells were preincubated at 37°C for 45 min in modified Krebs medium containing 0.3 mM- $\alpha\alpha'$ -bipyridyl before the addition of [^{14}C]proline, and incubation was continued for a further 2 h. In an alternative procedure, hydroxylation of the procollagen was inhibited by conducting the cell incubations under an atmosphere of N_2 and taking the precautions described by Uitto & Prockop (1974a).

Subcellular fractionation of tendon and cartilage cells

The isolation of fractions of rough endoplasmic reticulum and smooth endoplasmic reticulum was achieved by using a discontinuous sucrose-density-gradient technique described previously (Harwood *et al.*, 1974b). It is noteworthy that Uitto *et al.* (1975) have confirmed that the application of these procedures to tendon cells affords an effective isolation of

fractions of smooth microsomal vesicles. However, the failure of these workers to obtain satisfactory fractions of rough endoplasmic reticulum is probably explained by their use of an homogenization buffer containing MgCl_2 , which has been shown to decrease the effectiveness of this separation (Dallner, 1967).

Characterization of [^{14}C]procollagen polypeptides

Gel-filtration on sodium dodecyl sulphate/agarose. The molecular size of procollagen polypeptides and the extent of aggregation resulting from inter-chain disulphide bonding was examined by gel filtration on sodium dodecyl sulphate/agarose before and after reduction with 2-mercaptoethanol. The samples were treated as described previously (Harwood *et al.*, 1975a) and then chromatographed at room temperature (20°C) on columns of either 6% agarose (Bio-Gel A-5m, 200–400 mesh) or 4% agarose (Bio-Gel A-15m, 200–400 mesh) both of which were equilibrated with 0.02M-Tris/HCl buffer, pH 7.4, containing 0.1% sodium dodecyl sulphate and 0.02% NaN_3 . The columns were 1.5 cm $\times$ approx. 95 cm, and fractions (2.4 ml) were collected at flow rates of approx. 5 ml/h. The total elution volume (V_t) for each chromatogram was established by the addition of 0.25 μCi of $^3\text{H}_2\text{O}$ to the sample before chromatography.

The ^{14}C in the fractions was assayed by taking 0.1 ml samples with 0.2 ml of 2M-HCl in the solvent system for liquid-scintillation counting as described by Harwood *et al.* (1974a). Hydroxy[^{14}C]proline in the fractions was assayed by the method of Juva & Prockop (1966).

Sucrose-density-gradient centrifugation under denaturing conditions. Procollagens secreted by tendon and cartilage cells were precipitated with $(\text{NH}_4)_2\text{SO}_4$ as described above and then redissolved in 0.1M-Tris/HCl buffer, pH 7.4, containing 0.4M-NaCl. The samples were each dialysed at 4°C against 500 ml of 0.04M-sodium acetate buffer, pH 4.5, containing 0.1M-NaCl, 0.1% Triton X-100 and 6M-urea which had been deionized with Amberlite MB1 ion-exchange resin (Hopkin and Williams Ltd., Chadwell Heath, Essex, U.K.). The samples were then denatured for 30 min at 40°C and portions (0.5 ml) of the procollagen solutions layered on 4.5 ml linear 5–20% (w/v) sucrose density gradients prepared in the acetate/urea buffer (Fessler *et al.*, 1973). These gradients were then centrifuged at 112000g (r_{av} 7.4 cm) for 36 h at 4°C in the 3 $\times$ 5 ml aluminium swing-out rotor of an MSE Superspeed 65 Ultracentrifuge. The gradient was removed from the centrifuge tube by upward displacement with 60% (w/v) sucrose and fractions (0.15 ml) were collected. To each fraction was added 1 ml of water and the radioactivity in the fractions determined by using the scintillant described previously (Harwood *et al.*, 1974a).

Electrophoresis on polyacrylamide/agarose composite gels. Initial characterization of the procollagen

polypeptides by polyacrylamide-gel electrophoresis was carried out on 5% (w/v) acrylamide gels by the method of Weber & Osborn (1969). This method is particularly appropriate for the separation of proteins of mol.wt. 30000–100000 and also permits a partial separation of radioactively labelled Type I collagen pro- α 1- and pro- α 2-chains having mol.wts. in excess of 120000 (Harwood *et al.*, 1975b). However, when ^{14}C -labelled proteins of higher molecular weight, e.g. unreduced procollagens, are to be characterized by gel electrophoresis, the above method had a limited application because of the restricted penetration and mobility of such species in gels of this composition.

Electrophoresis on polyacrylamide/agarose composite gels has been applied successfully to the analysis of high-molecular-weight macromolecular species such as bacterial polyribosomes and ribosomal subunits (Dahlberg *et al.*, 1969), RNA (Dingman & Peacock, 1968; Peacock & Dingman, 1968; Dingman *et al.*, 1970) and proteoglycans (McDevitt & Muir, 1971), and we here describe the application of this method to the analysis of types I and II procollagens and their constituent pro- α chains. The composite gels of 2.0% acrylamide and 0.5% agarose were prepared by a modification of the procedure described by Peacock & Dingman (1968). The agarose solution was prepared by stirring 0.2 g of Seakem agarose in 29 ml of 0.1M-sodium phosphate buffer, pH 7.4, to give a swollen agarose suspension, which was then refluxed at 100°C for 15 min and subsequently cooled to 40°C. A stock solution of 20% acrylamide monomer was prepared by dissolving 1 g of bisacrylamide and 19 g of acrylamide, which had been recrystallized from chloroform, in 100 ml of water (Loening, 1967). To 4 ml of this stock acrylamide solution was added 1.25 ml of water, 40 μl of $\text{NNN}'\text{N}'$ -tetramethyl-1,2-diaminoethane and 5.25 ml of 0.2M-sodium phosphate buffer, pH 7.4, containing 0.1% (w/v) sodium dodecyl sulphate, and the solution warmed to 35°C. The agarose and acrylamide solutions were mixed, the temperature was adjusted to 35°C and 0.4 ml of 1.6% (w/v) ammonium persulphate added. The completed gel solution was mixed well, poured rapidly into Perspex tubes (0.6 cm $\times$ 11.0 cm) and left to gel for 1 h at room temperature.

Samples containing ^{14}C -labelled protein were analysed on the composite gels before and after reduction with 2-mercaptoethanol. Treatment of the samples before electrophoresis was carried out as described for gel-filtration analysis except that the buffer used here was 0.1M-sodium phosphate buffer, pH 7.4, containing 0.1% sodium dodecyl sulphate. One drop of glycerol and 5 μl of tracking dye (0.05% Bromophenol Blue in water) were added to the unreduced and reduced samples (25–100 μl), which were then applied to the gels and electrophoresis was carried out at 10 mA/tube for up to 4 h at room temperature in 0.1M-sodium phosphate buffer,

pH 7.4, containing 0.1% dodecyl sulphate. The gels were stained for 2 h with Coomassie Brilliant Blue and destained in 7.5% acetic acid solution as described by Furthmayr & Timpl (1971). Analysis of the gels for radioactivity was carried out on 1 mm slices obtained by using a Mickle Gel Slicer (Joyce, Loeb and Co., Gateshead, U.K.). The slices were incubated at 50°C overnight in 0.5 ml of 30% (v/v) H_2O_2 and then 10 ml of scintillant [5 vol. of toluene containing 6 g of 2,5-diphenyloxazole/litre and 0.12 g of 1,4-bis-(5-phenyloxazol-2-yl)benzene/litre and 3 vol. of 2-methoxyethanol] was added and the sample counted for radioactivity in a Packard 2450 liquid-scintillation spectrometer.

Analysis of extension peptides of tendon procollagen

Studies to determine the size and the location of the extension peptides of procollagen were carried out by using a highly purified preparation of human rheumatoid synovial collagenase (Woolley *et al.*, 1975) and bacterial collagenase (Worthington Biochemicals, CLSPA grade; Cambrian Chemicals Ltd., Croydon, U.K.) further purified by the method of Peterkofsky & Diegelmann (1971). Tendon procollagen dual-labelled with [^{35}S]cystine and [3H]proline or labelled only with [3H]tryptophan was prepared as described above. The procollagen preparations were incubated with the two enzymes as described by Merry *et al.* (1976), except that the period of incubation with the human collagenase was decreased to 5 h because of the increased susceptibility of the type I procollagen to cleavage by this enzyme. The products of digestion were analysed under reducing and non-reducing conditions by sodium dodecyl sulphate/polyacrylamide-gel electrophoresis (Neville, 1971) and by gel filtration on 8% agarose (Bio-Gel A-1.5 m) (Merry *et al.*, 1976) equilibrated with 0.02 M-Tris/HCl, pH 7.4, containing 0.1% sodium dodecyl sulphate and 0.02% Na $_2$ S.

Results

Determination of the molecular size of the procollagens secreted by tendon and cartilage cells

Previous studies have demonstrated that matrix-free embryonic-chick tendon and cartilage cells secrete respectively type I and type II procollagens in which the pro- α -chains are linked by disulphide bonds. Reduction of these species yields pro- α -chains whose molecular weights when measured by sodium dodecyl sulphate/agarose gel filtration have been estimated to lie within the range 115 000–125 000 (Jimenez *et al.*, 1971; Dehm *et al.*, 1972; Dehm & Prockop, 1973; Harwood *et al.*, 1973, 1975a). These values were determined by gel-filtration chromatography on 6% agarose columns equilibrated with 0.1% sodium dodecyl sulphate, but this technique has limited resolution when applied to high-molecular-weight components. Consequently no satis-

factory values for the molecular size of the disulphide-linked trimeric procollagen molecules have been obtained, and on the basis of analyses carried out on other type I procollagens (for review see Grant & Jackson, 1976) the above values for the molecular weights of the individual pro- α -chains of tendon and cartilage procollagens may be underestimates. We have therefore re-examined the molecular sizes of

Fig. 1. Gel filtration on sodium dodecyl sulphate/agarose (Bio-Gel A15-m, 200–400 mesh) of unreduced and reduced [^{14}C]procollagen secreted by tendon and cartilage cells (a) Tendon cells (5×10^6) were incubated in 50 ml of modified Krebs medium at 37°C with 10 μ Ci of [^{14}C]proline. After an incubation period of 2 h, the cells were pelleted and the medium was dialysed against NaCl/Tris/HCl buffer containing proteinase inhibitors (see the text). [^{14}C]Procollagen in the non-diffusible material was precipitated with $(NH_4)_2SO_4$ added to 30% saturation and collected by centrifugation at 21 000g for 1 h. The precipitate was treated with 1% (w/v) sodium dodecyl sulphate in the presence of either 0.1 M-iodoacetamide or 5% (v/v) 2-mercaptoethanol. The sample was dialysed against 0.1% sodium dodecyl sulphate in 0.02 M-Tris/HCl buffer (pH 7.4) before co-chromatography with 5 mg of rat tail-tendon collagen which had been added to the sample before solubilization. The void volume of the column (V_0) as determined by the elution position of Blue Dextran was in fraction 23 (55 ml), and the total volume (V_t) as determined by the elution position of 3H_2O was in fraction 78 (187 ml). (b) Cartilage cells (4.5×10^6) were incubated in 45 ml of medium at 37°C with 10 μ Ci of [^{14}C]proline for 2 h. [^{14}C]Procollagen secreted into the medium was recovered and analysed as described above. ○, Unreduced sample; ●, reduced sample; ----, A₂₃₀ pattern of carrier rat tail-tendon collagen.

tendon and cartilage procollagen polypeptides by using the three separate techniques described below.

Gel-filtration chromatography on sodium dodecyl sulphate/agarose (Bio-Gel A-15m). Analysis of rat tail-tendon collagen on a column of 8% agarose equilibrated with 0.1% sodium dodecyl sulphate indicated that it was possible to resolve α -, β - and γ -components into three distinct peaks (Fig. 1a). Gel filtration of unreduced tendon procollagen indicated that the major peak of ^{14}C -labelled polypeptides was eluted in fraction 32 slightly ahead of carrier γ -chains (Fig. 1a). When chromatographed after reduction with 2-mercaptoethanol, the tendon procollagen polypeptides were eluted in a position intermediate between β - and α -chains (Fig. 1a). The sodium dodecyl sulphate/agarose column was calibrated with a number of collagenous [the mol.wt. of the standard α -chains was taken as 97500, the value calculated from the amino acid sequence of the $\alpha 1(\text{I})$ -chain reported by Hulmes *et al.* (1973)] and non-collagenous [bovine serum albumin (mol.wt. 68000), ovalbumin (mol.wt. 45000) and pepsin (mol.wt. 35000)] protein standards. A plot of the logarithm of molecular weight against the ratio of the elution volume to the total column volume was linear over a range of mol.wts. 35000–292000. As previously (Grant *et al.*, 1972), the polypeptide chains from extracellular collagen behave similarly to globular proteins when examined by gel filtration on sodium dodecyl sulphate/agarose, although they behave anomalously on sodium dodecyl sulphate/polyacrylamide-gel electrophoresis (Furthmayr & Timpl, 1971). Reference to the calibration curve indicated that the ^{14}C -labelled polypeptides obtained after reduction had mol.wts. of approx. 137000. Estimation of the molecular weight of the unreduced species by extrapolation of the calibration curve suggested this aggregate represented trimers of the reduced polypeptides, since their mol.wt. was approx. 445000. The absence of other peaks on reduction of the tendon [^{14}C]procollagen suggests that no other proteins are present in the aggregate species. This conclusion is supported by analyses of the hydroxy[^{14}C]proline contents of the unreduced and reduced peaks, which indicated that 42 and 43% respectively of the total ^{14}C in these peaks was present as hydroxy[^{14}C]proline.

Similar studies were conducted with the procollagen secreted by cartilage cells. Analysis of the molecular size of cartilage [^{14}C]procollagen recovered by $(\text{NH}_4)_2\text{SO}_4$ precipitation of medium in which cartilage cells had been incubated with [^{14}C]proline is presented in Fig. 1(b). The radioactivity was eluted slightly ahead of collagen γ chains when chromatographed under non-reducing conditions, but when the sample was reduced and alkylated before gel filtration, the polypeptides were eluted in a position corresponding to a mol.wt. of approx. 145000.

Analysis by gradient centrifugation under denaturing

conditions. The technique of ultracentrifugal velocity-sedimentation analysis under denaturing conditions has found wide application in the analysis of large macromolecules and has recently been used in studies of procollagen from embryonic-chick cranial bones (Fessler *et al.*, 1973, 1975; Fessler & Fessler, 1974; Morris *et al.*, 1975). When tendon [^{14}C]procollagen was reduced, alkylated and analysed on a 5–20% (w/v) sucrose gradient, the peak of radioactivity sedimented ahead of a standard of [^{14}C]proline-labelled α -chains of collagen run concurrently (Fig. 2a). A sample of unreduced tendon procollagen was also subjected to sedimentation analysis and found to

Fig. 2. Velocity-sedimentation analysis of unreduced and reduced [^{14}C]procollagen secreted by tendon and cartilage cells

In these experiments, 6.2×10^8 tendon cells and 5.4×10^8 cartilage cells were incubated with [^{14}C]proline under conditions described in the text. Tendon and cartilage [^{14}C]procollagens were partially purified by $(\text{NH}_4)_2\text{SO}_4$ precipitation, redissolved in 0.04M-sodium acetate buffer containing 6M-urea, 0.1M-NaCl and 0.1% Triton X-100, and dialysed against this buffer for 18 h at 4°C . The samples were centrifuged at $112000g_{\text{av}}$ ($r_{\text{av}} = 7.4\text{cm}$) for 36 h at 4°C on a linear gradient of 5–20% sucrose in the above buffer. Fractions (0.15 ml) were collected by upward displacement of the gradient with 60% sucrose and analysed for radioactivity (Harwood *et al.*, 1974a). A sample of [^{14}C]proline-labelled α chains (50000c.p.m.) was similarly treated and run concurrently in the ultracentrifuge. In experiments where the [^{14}C]procollagens were analysed under reducing conditions the partially purified procollagen precipitate was redissolved in buffer containing 5% mercaptoethanol. (a) Tendon procollagen, (b) cartilage procollagen. $\circ$, Unreduced sample; $\bullet$, reduced sample; ----, [^{14}C]proline-labelled α -chains. Sedimentation was from left to right as indicated by the arrow.

migrate with a sedimentation coefficient of approx. 4.4S.

Similar sedimentation studies were conducted with reduced and unreduced cartilage [^{14}C]procollagen (Fig. 2b), and values of 2.5S and 4.4S were obtained for the sedimentation coefficients of cartilage pro- α -chains and the intact procollagen respectively. These values determined for the sedimentation coefficients of tendon and cartilage procollagens and their individual pro- α -chains correspond well with those obtained for the unreduced and reduced forms of chick cranial bone procollagen (Fessler & Fessler, 1974).

Electrophoresis on sodium dodecyl sulphate/agarose/polyacrylamide composite gels. Sodium dodecyl sulphate/polyacrylamide-gel electrophoresis has been widely used in the analysis of collagenous molecules (Clark, 1976) and provides a means of separating dimers (β -chains) and trimers (γ -chains) from the monomeric α -chains, as well as separating the individual $\alpha 1$ - and $\alpha 2$ -chains. Penetration of the β - and γ -

Fig. 3. Electrophoretic analysis of unreduced and reduced [^{14}C]procollagen on agarose/polyacrylamide composite gels. Portions of [^{14}C]procollagen to which had been added rat tail-tendon collagen were treated with 0.1% sodium dodecyl sulphate in 0.1M-sodium phosphate buffer (pH 7.4) in the presence of either 0.1M-iodoacetamide or 5% mercaptoethanol. Samples (100 μl) containing approx. 15000 c.p.m. of labelled procollagen and up to 40 μg of carrier collagen were applied to 11 cm gels composed of 2% acrylamide and 0.5% agarose. Electrophoresis was carried out at 10mA/tube for 4h at room temperature. Gels were stained with Coomassie Brilliant Blue (Furthmayr & Timpl, 1971) and scanned at 620nm, or sliced and assayed for radioactivity. (a) Tendon procollagen, (b) cartilage procollagen. $\circ$, Unreduced sample; $\bullet$, reduced sample.

components into the gels is, however, somewhat restricted and in order to obtain a good separation of pro- $\alpha 1(\text{I})$ - from pro- $\alpha 2$ -chains and also to investigate the possible existence of dimeric intermediates (pro- β -chains) during the synthesis of procollagen (pro- γ -chains), we have modified the method of Dingman & Peacock (1968), using agarose/polyacrylamide composite gels. The electrophoretic behaviour of collagen $\alpha 1(\text{I})$ -components [rat tail-tendon $\gamma 111$ -chains (mol.wt. 292500), $\beta 11$ -chains (mol.wt. 195000) and $\alpha 1$ -chains (mol.wt. 97500) and the $\alpha 1$ -CB 6,7 CNBr peptide (mol.wt. 44000) from guinea-pig skin $\alpha 1$ -chains] and $\alpha 2$ -components [rat tail-tendon $\alpha 2$ -chains (mol.wt. 97500) and the $\alpha 2$ -CB 3,5 CNBr peptide (mol.wt. 60000) from rat skin $\alpha 2$ -chains] was examined and calibration curves were obtained. This technique affords increased penetration and greater resolution of high-molecular-weight collagen polypeptides, and the analysis of the distribution of radioactivity in gels of tendon [^{14}C]procollagen electrophoresed under non-reducing and reducing conditions is shown in Fig. 3(a). The unreduced procollagen migrates slower than do γ components of rat tail-tendon collagen, and by extrapolation from the calibration curve has an apparent mol.wt. of approx. 405000. When the reduced polypeptides are run in this system, two peaks of radioactivity occurring in the ratio of 2:1 and corresponding to pro- $\alpha 1(\text{I})$ - and pro- $\alpha 2$ -chains are observed intermediate between β -components and α -chains. By reference to the calibration curves for $\alpha 1(\text{I})$ - and $\alpha 2$ -derived molecules it can be calculated that the mol.wts. of these two species are approx. 140000.

Similar studies on cartilage [^{14}C]procollagen are reported in Fig. 3(b), from which it can be concluded that the unreduced molecule represents a trimer of pro- $\alpha 1(\text{II})$ molecules whose mol.wt. is approx. 140000 when calculated by this technique.

Identification and partial characterization of the extension peptides of tendon procollagen

Earlier studies on the procollagen secreted by embryonic-chick tendon cells indicated that this molecule differed from the collagen fibril monomer in having a higher molecular weight (Jimenez *et al.*, 1971), which was attributed to extension peptides at the *N*-terminus (Dehm *et al.*, 1972). However, subsequent studies on the procollagen synthesized by cultured fibroblasts (Tanzer *et al.*, 1974), embryonic-chick cranial bones (Byers *et al.*, 1975; Fessler *et al.*, 1975) and matrix-free cartilage cells (Merry *et al.*, 1976) have indicated that these precursors have both *N*- and *C*-terminal extensions. We have therefore reinvestigated the location of the extension peptides of tendon procollagen, taking care to minimize non-specific proteolysis during the partial purification of the procollagen.

When tendon procollagen dual-labelled with [^3H]-

Fig. 4. Electrophoretic analysis of dual-labelled tendon procollagen treated with human rheumatoid synovial collagenase

Some 1.34×10^9 tendon cells were labelled for 2 h with $100 \mu\text{Ci}$ of $[^3\text{H}]\text{proline}$ and $50 \mu\text{Ci}$ of $[^3\text{S}]\text{cysteine}$ and the secreted procollagen was isolated by $(\text{NH}_4)_2\text{SO}_4$ precipitation. (a) A portion of procollagen was denatured in sodium dodecyl sulphate, reduced with mercaptoethanol and a sample ($50 \mu\text{l}$, containing approx. 25000 c.p.m. of ^3H) applied to an 8 cm polyacrylamide (7.5%) gel and electrophoresis carried out at 2 mA/tube for 3 h (Neville, 1971). Gels were sliced and assayed for radioactivity as described in the text. (b) A portion of procollagen containing approx. 50000 c.p.m. of ^3H was digested for 5 h at 25°C with human collagenase as described by Merry *et al.* (1976) and the products were analysed as described above after denaturation in sodium dodecyl sulphate and alkylation with iodoacetamide. (c) Products of the human collagenase digest analysed after denaturation in sodium dodecyl sulphate and reduction with mercaptoethanol. The migration positions of collagen pro- $\alpha 1(\text{I})$ -, pro- $\alpha 2$ -, $\alpha 1(\text{I})$ - and $\alpha 2$ -chains and bovine serum albumin are labelled 1-5 respectively. The radioactive peaks representing the procollagen fragments resulting from digestion with human collagenase are designated A-E.

proline and $[^3\text{S}]\text{cysteine}$ (Fig. 4a) was digested with human collagenase and the products were analysed by gel electrophoresis under non-reducing conditions, three major radioactive peaks, A, B and C (Fig. 4b), were observed. After reduction, peaks B and C, migrating in the region of collagen α -chains, were unchanged, but the higher-molecular-weight peak A was absent and two new components, D and E (Fig. 4c), were observed to migrate well ahead of $\alpha 2$ -chains and in the region of bovine serum albumin (mol.wt. 68000). When the reduced products of this collagenase digest were analysed on 8% agarose equilibrated with 0.1% (w/v) sodium dodecyl sulphate, only two peaks were obtained. The higher-molecular-weight peak corresponding to components B and C was eluted with an apparent mol.wt. of 90000 and the second peak corresponding to components D and E chromatographed with an apparent mol.wt. of 61000.

The human collagenase used in the present study has been shown to cleave type I collagen at a specific locus three-quarters of the length of the molecule from the *N*-terminal end to yield two triple-helical moieties, TC^{A} and TC^{B} , comprising α -chain fragments of approx. mol.wt. 75000 and 25000 respectively (Woolley *et al.*, 1975). Since peaks D and E have molecular weights that are too small to represent the products *N*-terminal to the site of collagenase cleavage, they must represent the *C*-terminal fragments, pro- $\alpha 1(\text{I})^{\text{B}}$ and pro- $\alpha 2^{\text{B}}$, which under non-reducing conditions migrate as peak A, corresponding to a disulphide-linked trimer with the composition $[\text{pro-}\alpha 1(\text{I})^{\text{B}}]_2\text{pro-}\alpha 2^{\text{B}}$. Peaks B and C must therefore represent the *N*-terminal fragments, pro- $\alpha 1(\text{I})^{\text{A}}$ and pro- $\alpha 2^{\text{A}}$, which are present in a molar ratio of 2:1 as judged by the extent of incorporation of $[^3\text{H}]\text{proline}$. It is noteworthy that no incorporation of $[^3\text{S}]\text{cysteine}$ into pro- $\alpha 2^{\text{A}}$ molecules was observed (Fig. 4c) and that interchain disulphide bonding is restricted to the *C*-terminal domain of tendon procollagen.

The above results also suggest that the *C*-terminal extensions of tendon pro- α -chains are larger than those at the *N*-termini. These observations were confirmed by gel-electrophoretic analysis of the products obtained when tendon procollagen was digested with bacterial collagenase (Fig. 5). Under non-reducing conditions, two peaks of radioactivity were obtained corresponding to mol.wts. of approx. 100000 and 17500 (Fig. 5a). On reduction the high-molecular-weight species corresponding to the disulphide-linked *C*-terminal extension peptides migrated with a mol.wt. of 36000 (Fig. 5b). In similar studies on procollagen labelled with $[^3\text{H}]\text{tryptophan}$, it was found that this labelled precursor was present predominantly in the bacterial collagenase-resistant peak corresponding to the *C*-terminal extensions of tendon procollagen (Fig. 5b).

Fig. 5. Electrophoretic analysis of radioactively labelled tendon procollagen

(a) A portion of procollagen labelled with $[^3\text{H}]$ proline and $[^{35}\text{S}]$ cystine obtained as described in Fig. 6 was digested with purified bacterial collagenase (Merry *et al.*, 1976). The sample was denatured in sodium dodecyl sulphate, alkylated with iodoacetamide, applied to an 8cm polyacrylamide (11%) gel and electrophoresis carried out at 2mA/tube for 2h (Neville, 1971). Gels were sliced and assayed for radioactivity as described in the text. ●, $[^{35}\text{S}]$ Cystine; ○, $[^3\text{H}]$ proline. (b) $[^{35}\text{S}]$ Cystine-labelled procollagen prepared as described in Fig. 6 was digested with bacterial collagenase as above and analysed by gel electrophoresis after denaturation with sodium dodecyl sulphate and reduction with mercaptoethanol. Similar analyses were carried out on a preparation of $[^3\text{H}]$ tryptophan-labelled procollagen isolated from the medium of cells (1.29×10^9 cells) incubated for 2h with $100 \mu\text{Ci}$ of $[^3\text{H}]$ tryptophan. The migration positions of bovine serum albumin, ovalbumin and pancreatic ribonuclease are labelled 1, 2 and 3 respectively. ●, $[^{35}\text{S}]$ Cystine; ○, $[^3\text{H}]$ tryptophan.

Role of the C-terminal extension peptides in the assembly of procollagen

The above results, together with the data of Merry *et al.* (1976), indicate that the procollagen molecules secreted by tendon and cartilage cells are trimers of pro- α chains linked by inter-chain disulphide bonds between the C-terminal extension peptides. Previous studies have shown that the formation of inter-chain disulphide bonds occurs soon after the release of completed polypeptides from membrane-bound ribosomes (Harwood *et al.*, 1973, 1975a), and a close temporal relationship has been reported between the

Fig. 6. Electrophoretic analysis of unreduced $[^{14}\text{C}]$ procollagen polypeptides in submicrosomal fractions from tendon and cartilage cells

Some 1.27×10^9 tendon cells were labelled for 2h with $25 \mu\text{Ci}$ of $[^{14}\text{C}]$ proline and then rough- and smooth-endoplasmic-reticulum fractions isolated (Harwood *et al.*, 1974b). The fractions were solubilized by heating for 2min at 100°C in 1ml of 0.1% sodium dodecyl sulphate in 0.1M-sodium phosphate buffer (pH7.4) containing 0.1M-iodoacetamide. Samples (50 or $100 \mu\text{l}$) were applied to 11cm composite gels and electrophoresis was carried out as described in Fig. 4. In a parallel experiment 6.5×10^8 cartilage cells were labelled for 2h with $25 \mu\text{Ci}$ of $[^{14}\text{C}]$ proline. Rough- and smooth-endoplasmic-reticulum fractions were isolated and analysed as described above. Gels were analysed for radioactivity (see the text) and the profiles presented are: (a), rough-endoplasmic-reticulum fraction from tendon cells; (b), rough-endoplasmic-reticulum fraction from cartilage cells; (c), smooth-endoplasmic-reticulum fraction from tendon cells; (d), smooth-endoplasmic-reticulum fraction from cartilage cells. The migration positions of rat tail-tendon collagen standards are indicated.

processes of disulphide bonding and triple-helix formation (Schofield *et al.*, 1974; Uitto & Prockop, 1974b). We have further examined the formation of disulphide bonds between pro- α -chains in tendon and cartilage cells and studied their role in helix formation.

State of aggregation of procollagen polypeptides in subcellular fractions. The demonstration by gel filtration of disulphide-bonded aggregates in the rough-endoplasmic-reticulum fractions from tendon and cartilage cells (Harwood *et al.*, 1973, 1975a) prompted further analyses by gel electrophoresis on agarose/polyacrylamide composite gels. Tendon cells were labelled for 2 h with [14 C]proline and submicrosomal fractions were isolated as described previously (Harwood *et al.*, 1973, 1974b). Gel electrophoresis of unreduced polypeptides from the rough microsomal fraction of tendon cells yielded major peaks of radioactivity corresponding to pro- γ -, pro- α 1- and pro- α 2-chains (Fig. 6a). Two very minor peaks of radioactivity were also observed in the region where pro- β -chains might be expected to migrate. The finding that the ratio of [14 C]proline present in the peaks corresponding to pro- α 1- to pro- α 2-chains in this fraction was almost exactly 2:1 was significant, indicating that the synthesis of the constituent chains of type I procollagen occurs in the ratio appropriate to their assembly into the functional molecule.

Similar studies conducted with cartilage cells

indicated that the major species present in rough microsomal fractions from cells labelled under steady-state conditions were pro- γ -chains and pro- α 1(II)-chains (Fig. 6b). When the smooth-endoplasmic-reticulum fractions from both cell types were analysed on the composite gels the unreduced [14 C]-procollagens migrated predominantly as disulphide bonded trimers of pro- α chains (Figs. 6c and 6d).

Relationship between inter-chain disulphide bonding and triple-helix formation. The triple-helical nature of the procollagens secreted by tendon and cartilage cells was confirmed by an analysis of their resistance to digestion by α -chymotrypsin (Table 1). Determinations of the extents to which the intracellular procollagen polypeptides were triple-helical indicated that only approx. 28 and 56% of these polypeptides were resistant to proteolysis in cartilage and tendon cells respectively. These values were almost identical with the proportions of the polypeptides present intracellularly as disulphide-bonded aggregates (Table 1).

Subsequent studies investigated the extent of disulphide bonding and the degree of helicity of procollagen [14 C]polypeptides in the rough- and smooth-endoplasmic-reticulum fractions from tendon and cartilage cells which had been labelled for 2 h with [14 C]proline. A close correlation was again observed between resistance to chymotrypsin and the extent to which the polypeptides were linked by inter-chain disulphide bonds (Table 1). However, whereas

Table 1. Relationship between the assembly of the collagen triple helix and inter-chain disulphide bonding of pro- α chains during procollagen biosynthesis

Cells were labelled with [14 C]proline for 2 h and then the extent of inter-chain disulphide bonding and helicity were determined for the secreted [14 C]procollagen, total intracellular [14 C]procollagen polypeptides, and also [14 C]procollagen polypeptides in rough- and smooth-endoplasmic-reticulum fractions isolated as described in the text. Samples were divided into two portions, and one was treated with 1% sodium dodecyl sulphate and 0.1 M-iodoacetamide and subjected to gel filtration on a Bio-Gel A-5 m column equilibrated and eluted with 0.1% sodium dodecyl sulphate and 0.02 M-Tris/HCl buffer, pH 7.4 (Harwood *et al.*, 1975a). The extent of disulphide bonding was calculated as the ratio of hydroxy[14 C]proline eluted in the first major included peak, which has previously been demonstrated to comprise aggregates of pro- α chains (Harwood *et al.*, 1973, 1975a), to total hydroxy[14 C]proline in the chromatogram. The second portion of each sample was incubated in 0.1 M-Tris/HCl buffer, pH 7.4, containing 0.4 M-NaCl, with α -chymotrypsin (100 μ g/ml) at 25°C for 6 h. After this proteolytic digestion the sample was treated with 1% sodium dodecyl sulphate and 5% (w/v) mercaptoethanol and chromatographed on the above column. The degree of helicity was calculated as the ratio of hydroxy[14 C]proline recovered in the α -chain peak after digestion to total non-diffusible hydroxy[14 C]proline in the sample before digestion.

	Tendon		Cartilage	
	Extent of disulphide bonding (%)	Extent of α -chymotrypsin resistance (%)	Extent of disulphide bonding (%)	Extent of α -chymotrypsin resistance (%)
Medium	98	97	95	93
Cells	56	56	29	28
Rough endoplasmic reticulum	87	86	90	88
Smooth endoplasmic reticulum	45	44	25	22

approx. 90% of the procollagen polypeptides in the smooth-endoplasmic-reticulum fractions from both cell types were both disulphide bonded and helical, there was a marked difference between the extent of assembly of procollagen in the rough-endoplasmic-reticulum fractions. In the rough microsomal fraction obtained from cartilage cells, less than 25% of the procollagen polypeptides were triple helical, whereas the extent of helix formation in the same fraction from tendon cells was approximately double that of the cartilage cells. These results are consistent with earlier studies indicating that the assembly and secretion of cartilage procollagen proceeds at a slower rate than that of tendon procollagen (Dehm & Prockop, 1973; Uitto & Prockop, 1974b; Harwood *et al.*, 1975a).

Effect of the inhibition of hydroxylation on the assembly of procollagen. When tendon and cartilage cells are incubated for 2–4 h under conditions where hydroxylation of proline and lysine residues is inhibited, the cells accumulate unhydroxylated procollagen polypeptides (Jimenez *et al.*, 1973; Uitto & Prockop, 1974a). The molecular size and composition of such collagenous species was investigated by gel filtration on sodium dodecyl sulphate/agarose (Bio-Gel A-5m) followed by gel electrophoresis on composite gels.

Tendon and cartilage cells were preincubated for 45 min either in the presence of 0.3 mM- $\alpha\alpha'$ -bipyridyl or under an atmosphere of N_2 before labelling with [^{14}C]proline for 2 h at 37°C under these conditions. The cells were pelleted by centrifugation and treated with 0.1 M-iodoacetamide and 1% sodium dodecyl sulphate before chromatography on the sodium dodecyl sulphate/agarose column. With both cell types incubated under N_2 , two major peaks of radioactivity were observed to be eluted in positions corresponding to pro- γ - and pro- α -chains, but the proportion of the pro- γ -component in the tendon-cell preparation was markedly greater than that in the cartilage cells (Figs. 7a and 7b). Almost identical results were obtained when hydroxylation was inhibited by $\alpha\alpha'$ -bipyridyl and the unhydroxylated procollagen polypeptides were examined by gel filtration (results not shown). Thus, it can be concluded that the formation of inter-chain disulphide bonds between the C-termini of pro- α chains can occur in the absence of hydroxylation, and further studies were undertaken to determine the composition of these unhydroxylated aggregates.

Fractions (25–30) corresponding to unhydroxylated pro- γ components were pooled and concentrated to 3 ml by dialysis against polyethylene glycol. Analyses of these samples for hydroxyl [^{14}C]proline indicated that less than 0.5% of the incorporated [^{14}C]proline was hydroxylated. When portions of these concentrates were examined by gel electrophoresis after reduction with 2-mercaptoethanol, the tendon pre-

Fig. 7. Gel filtration on sodium dodecyl sulphate/agarose (Bio-Gel A-5m) of unreduced ^{14}C -labelled protein which accumulated in tendon and cartilage cells incubated under anaerobic conditions

(a) Some 1.5×10^8 tendon cells were incubated under N_2 for 60 min in 15 ml of modified Krebs medium containing 10% (v/v) foetal calf serum. [^{14}C]Proline (10 μ Ci) was added and incubation continued for a further 60 min under N_2 . Protein synthesis was then inhibited by the addition of cycloheximide to a final concentration of 100 μ g/ml and the sample was also made 1 mM in $\alpha\alpha'$ -bipyridyl to prevent hydroxylation occurring during processing. The cells were collected by centrifugation and immediately heated in 2 ml of 1% sodium dodecyl sulphate in 0.02 M-Tris/HCl buffer (pH 7.4) containing 0.1 M-iodoacetamide and 1 mM- $\alpha\alpha'$ -bipyridyl for 2 min at 100°C. The sample was dialysed against 0.1% sodium dodecyl sulphate in Tris/HCl buffer followed by gel filtration on a Bio-Gel A-5m column and fractions were collected and assayed for radioactivity (see Table 1 for details). (b) Some 2×10^8 cartilage cells were incubated for 60 min under N_2 with 10 μ Ci of [^{14}C]proline after pre-incubation under N_2 and the ^{14}C -labelled protein accumulating in the cells was analysed by gel filtration as described above.

paration was resolved into two radioactive peaks present in a ratio of 2:1 and migrating in the region of pro- α 1- and pro- α 2-chains of tendon procollagen (Fig. 8a). The unhydroxylated cartilage fraction yielded on reduction a single peak of radioactivity corresponding to unhydroxylated pro- α (II)-chains

Fig. 8. Electrophoretic analysis of unhydroxylated [^{14}C]-procollagen from tendon and cartilage cells labelled with [^{14}C]proline under anaerobic conditions

Fractions (25–30) from the experiments described in Figs. 7(a) and 7(b) were pooled and concentrated to 3 ml by dialysis against polyethylene glycol. The samples were reduced with mercaptoethanol and portions (100 μl) analysed on composite agarose/polyacrylamide gels (see the text). Simultaneous analyses of hydroxylated tendon and cartilage pro- α -chains were carried out and are indicated by the broken line. (a) Hydroxylated and unhydroxylated tendon pro- α -chains; (b) hydroxylated and unhydroxylated cartilage pro- α -chains.

(Fig. 8b). The peaks of radioactivity corresponding to unhydroxylated pro- α 1- and pro- α 2-chains migrate with slightly higher mobilities than the corresponding hydroxylated forms (Fig. 8a). This difference in mobility can be attributed to the lower molecular weights of the unhydroxylated tendon pro- α -chains, because of the absence of both hydroxyl groups on certain proline and lysine residues and of galactosyl and glucosylgalactosyl moieties attached to hydroxylysine residues. As might be expected, the difference in mobility of the hydroxylated and unhydroxylated cartilage pro- α 1(II)-chains is more pronounced (Fig. 8b), since the secondary modifications of hydroxylation and glycosylation are more extensive in cartilage procollagen (Dehm & Prockop, 1973) and a greater mass difference therefore exists between the hydroxylated and unhydroxylated forms. Similar results have been reported by Fessler & Fessler (1974) in studies of the hydroxylated and unhydroxylated forms of

type I procollagen from embryonic-chick cranial bones.

Discussion

The existence of procollagen has been well established by many laboratories using various cell systems (Bornstein, 1974; Martin *et al.*, 1975) and there is general agreement that the triple-helical procollagen molecule secreted by connective-tissue cells has a molecular weight greater than that of the collagen fibril monomer because of large non-helical peptide extensions (for review see Grant & Jackson, 1976). The size of the newly synthesized precursor has not been completely defined and it has been suggested that procollagen may be synthesized initially as a very large single polypeptide chain of molecular weight greater than 500 000 (Church *et al.*, 1971; Bankowski & Mitchell, 1973; Park *et al.*, 1975). However, most studies have indicated that procollagen has a mol.wt. of 350 000–400 000 and comprises three polypeptide chains (pro- α -chains) linked by inter-chain disulphide bonds (Bornstein, 1974; Prockop *et al.*, 1976). Pulse-labelling studies of the synthesis of type I procollagen have demonstrated that the synthesis of pro- α 1- and pro- α 2-chains is initiated simultaneously (Vuust & Piez, 1970, 1972; Byers *et al.*, 1975; Fessler *et al.*, 1975), and this finding, together with observations on the size of collagen-synthesizing polyribosomes and of procollagen mRNA (for review see Grant & Jackson, 1976), provide a considerable body of evidence in favour of procollagen biosynthesis being directed by monocistronic mRNA species for the individual pro- α -chains.

The studies presented here are also indicative of the separate synthesis of pro- α -chains, for gel filtration and gel electrophoresis of both secreted and intracellular procollagen from tendon and cartilage cells revealed only individual polypeptides of mol.wt. 137 000–145 000 when the analyses were conducted under denaturing and reducing conditions. The determinations on the individual pro- α -chains by these techniques are in good agreement, and although it is recognized that the presence of non-collagenous and collagenous sequences within the polypeptides will influence the accuracy of the values obtained (Furthmayr & Timpl, 1971; Bornstein, 1974) the resolution of the methods used is greater than that previously used in studies of tendon and cartilage procollagen (Jimenez *et al.*, 1971; Dehm & Prockop, 1973). Estimates of the molecular weights of the unreduced procollagens gave values in the range 415 000–445 000 and indicate the trimeric nature of the secreted forms of both type I and type II collagen precursors. The agreement between the values obtained here for the sedimentation coefficients of tendon and cartilage procollagens with those determined for [^{14}C]procollagen isolated from intact chick

cranial bones (Fessler & Fessler, 1974) suggests that precursors of similar nature and molecular size were being examined.

Earlier studies indicated that [^{14}C]cysteine and [^{14}C]tryptophan, amino acids not found in tendon collagen α -chains, were incorporated into tendon procollagen polypeptides (Dehm *et al.*, 1972). Electron microscopy of segment-long-spacing crystallites of these procollagen preparations revealed the presence of an *N*-terminal extension, which was presumed to be the location of the radioactively labelled amino acids (Dehm *et al.*, 1972). In view of the demonstration that type I procollagens secreted by cultured fibroblasts (Tanzer *et al.*, 1974) and chick cranial bones (Byers *et al.*, 1975; Fessler *et al.*, 1975) have both *N*- and *C*-terminal extensions, we have re-examined the procollagen secreted by isolated chick tendon cells. Experiments on the action of rheumatoid synovial collagenase on dual-labelled procollagen (Fig. 4) suggested the presence of extensions at both ends of the molecule, indicating that the electron-microscopy studies reported by Dehm *et al.* (1972) were carried out on a partially degraded form of procollagen. Further analyses of the fragments from the human collagenase digest revealed that inter-chain disulphide bonding was restricted to the *C*-terminal domain but that [^{35}S]cystine was also incorporated into the *N*-terminal extension of the pro- $\alpha 1(\text{I})$ -chain. Similar results have been reported by Fessler *et al.* (1975) for type I bone procollagen, and we have shown that chick cartilage pro- $\alpha 1(\text{II})$ -chains incorporate [^{35}S]cystine into both *N*- and *C*-terminal extension peptides (Merry *et al.*, 1976).

Analyses of the dual-labelled peptides isolated from digests of tendon procollagen with bacterial collagenase (Fig. 5a) indicated that the *N*- and *C*-terminal extensions had apparent mol. wts. of 17 500 and 36 000 respectively. It is noteworthy that [^3H]tryptophan was incorporated predominantly into the *C*-terminal extension peptides (Fig. 5b) and the molecular-weight value for these peptides is similar to those reported for the procollagens synthesized by cranial bones (Byers *et al.*, 1975; Fessler *et al.*, 1975) and sternal cartilage (Merry *et al.*, 1976) of embryonic chicks. The *N*-terminal extension appears to be larger than that of the pro- $\alpha 1(\text{II})$ -chain (Merry *et al.*, 1976), but its molecular weight is within the range reported for the type I procollagen of cranial bone (Byers *et al.*, 1975; Fessler *et al.*, 1975). No separation of the extension peptides of the individual pro- $\alpha 1(\text{I})$ - and pro- $\alpha 2$ -chains was observed, but the absence of [^{35}S]cystine from the pro- $\alpha 2^A$ fragment (Fig. 4c) suggests that the amino acid composition of the *N*-terminal peptides differ in the constituent pro- α -chains of tendon procollagen.

A role for the *C*-terminal peptides in aligning the procollagen polypeptides and promoting helix formation is suggested by the close correlation ob-

served between inter-chain disulphide bonding and helix formation during procollagen assembly in tendon (Schofield *et al.*, 1974) and cartilage cells (Uitto & Prockop, 1974b). However, because inter-chain disulphide bonds reside only at the *C*-terminus the synthesis of the individual polypeptides must be completed or virtually complete before covalent linkage can occur, and the suggestion that helix formation commences at the ribosomal level (Miller & Matukas, 1974) would seem improbable. Indeed in previous studies we have shown that disulphide bonding occurs after release of pro- α -chains from ribosomes (Harwood *et al.*, 1973, 1975a) and that nascent procollagen polypeptides are digested by a chymotrypsin/trypsin mixture at 4°C (Harwood *et al.*, 1974c), indicating that helix formation must also be a post-ribosomal event.

Inter-chain disulphide bonding has been shown to commence in the rough endoplasmic reticulum, and the experiments described in Table 1 confirm that this location is also the site at which triple-helix formation begins. The close correlation between disulphide bonding and helix formation is again observed, for in the rough and smooth microsomal fractions from tendon and cartilage cells the proportion of polypeptides resistant to proteolytic digestion was virtually identical with the proportion of pro- α -chains involved in disulphide linkage. In comparison with tendon cells under steady-state conditions, a greater proportion of the total intracellular procollagen polypeptides in cartilage cells are present as non-covalently linked pro- α -chains (Table 1). This difference appears to reflect the slower rate at which the disulphide-bonding process occurs in the rough endoplasmic reticulum of cartilage cells.

The mechanisms by which the pro- α -chains associate is unknown and although it is still possible that the *N*-terminal peptides play some role in the assembly of procollagen, covalent bond formation, as discussed above, occurs only between the *C*-terminal extension peptides. Application of the agarose/polyacrylamide composite gel system to the analysis of unreduced procollagen polypeptides present in subcellular fractions provides a means of investigating the possible existence of intermediates in the assembly of trimeric pro- γ -components. The analysis of [^{14}C]labelled polypeptides in rough microsomal fractions from tendon and cartilage cells revealed the presence of only small peaks of radioactivity in the pro- β regions of the gels. These results clearly demonstrate that during procollagen biosynthesis no significant pool of dimeric intermediates is present and, as suggested by Fessler & Fessler (1974), it appears that such an intermediate must be short-lived and that the nature of the interactions between the non-collagenous extensions requires that three of them interact virtually simultaneously.

Further insight into the association of the pro- α -

chains during procollagen assembly has been gained from studies with tendon and cartilage cells in which hydroxylation, and hence triple-helix formation, has been inhibited by either $\alpha\alpha'$ -bipyridyl or an atmosphere of N_2 . The synthesis of unhydroxylated pro- α -chains occurs under these conditions, and as has been reported (Schofield *et al.*, 1974; Uitto & Prockop, 1974b; Fessler & Fessler, 1974) inter-chain disulphide bonding can occur in the absence of hydroxylation. Analysis by gel electrophoresis of the disulphide-linked aggregate of unhydroxylated tendon pro- α -chains demonstrated that this species represented a trimer which on reduction yielded unhydroxylated pro- α 1- and pro- α 2-chains in a ratio of 2:1. The above results are supported by the studies of Fessler & Fessler (1974) and indicate that hydroxylation is not a prerequisite for correct chain association, and there is no evidence to support the suggestion that $\alpha\alpha'$ -bipyridyl influences the chain composition of type I unhydroxylated procollagen (Müller *et al.*, 1974).

The data here relating to the formation of hydroxylated and unhydroxylated tendon procollagen are consistent with a model for the assembly of pro- γ 112-chains in which the recognition and selection of pro- α 1- and pro- α 2-chains in a 2:1 ratio are directed by the non-collagenous extensions. Whether or not inter-chain disulphide bonding is an immediate consequence of chain association is not known, but the absence of significant amounts of dimeric intermediates indicates that when inter-chain disulphide bonding occurs between the three pro- α -chains it proceeds rapidly. The observations that the proportion of intracellular pro- α -chains in disulphide linkages varies among cell types (Grant & Jackson, 1976; Prockop *et al.*, 1976) indicates that differences in the timing of inter-chain disulphide bonding occur. How this process is controlled is not understood, but its importance is indicated by its correlation with triple-helix formation and the observations that the secondary modifications of hydroxylation (Berg & Prockop, 1973; Murphy & Rosenbloom, 1973; Ryhänen & Kivirikko, 1974) and glycosylation (Myllylä *et al.*, 1975; Risteli *et al.*, 1976) do not continue after assembly of the triple helix. Thus, by delaying the inter-chain disulphide bonding in the rough endoplasmic reticulum the pro- α -chains continue to be substrates for the hydroxylases and glycosyl transferases, and hence in cartilage cells it would be possible to explain the increased amounts of hydroxyproline, hydroxylysine and glycosylated hydroxylysine in the collagens secreted by these cells.

During the preparation of this manuscript Rosenbloom *et al.* (1976) reported experiments suggesting that chick tendon procollagen possesses C-terminal extensions that are required for normal assembly of the three individual pro- α -chains and for helix formation. A very recent paper by Olsen *et al.* (1976) also

provides evidence for N- and C-terminal extension peptides in the procollagens secreted by chick tendon and cartilage cells, and the demonstration that inter-chain disulphide bonds are located at the C-terminus agrees with the results presented here.

We thank Mrs. D. Harrott, Mrs. M. Kendrick, Miss J. Ross and Miss L. Seddon for skilled technical assistance. The support for this study by grants from the Science Research Council and Medical Research Council is gratefully acknowledged.

References

- Bankowski, E. & Mitchell, W. M. (1973) *Biophys. Chem.* **1**, 73–86
- Berg, R. A. & Prockop, D. J. (1973) *Biochemistry* **12**, 3395–3401
- Bornstein, P. (1974) *Annu. Rev. Biochem.* **43**, 567–604
- Byers, P. H., Click, E. M., Harper, E. & Bornstein, P. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 3009–3013
- Church, R. L., Pfeiffer, S. E. & Tanzer, M. L. (1971) *Proc. Natl. Acad. Sci. U.S.A.* **68**, 2638–2642
- Clark, C. C. (1976) in *The Methodology of Connective Tissue Research* (Hall, D. A., ed.), pp. 205–226, Joynson-Bruvvers Ltd., Oxford
- Dahlberg, A. E., Dingman, C. W. & Peacock, A. C. (1969) *J. Mol. Biol.* **41**, 139–147
- Dallner, G. (1967) *Acta Pathol. Microbiol. Scand. Suppl.* **166**, 1–94
- Dehm, P. & Prockop, D. J. (1971) *Biochim. Biophys. Acta* **240**, 358–369
- Dehm, P. & Prockop, D. J. (1973) *Eur. J. Biochem.* **35**, 159–166
- Dehm, P., Jimenez, S. A., Olsen, B. R. & Prockop, D. J. (1972) *Proc. Natl. Acad. Sci. U.S.A.* **69**, 60–64
- Dingman, C. W. & Peacock, A. C. (1968) *Biochemistry* **7**, 659–667
- Dingman, C. W., Kakefuda, T. & Aronow, A. (1970) *Biochim. Biophys. Acta* **224**, 114–127
- Fessler, J. H. & Fessler, L. I. (1974) *J. Biol. Chem.* **249**, 7637–7646
- Fessler, L. I., Burgeson, R. E., Morris, N. P. & Fessler, J. H. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 2993–2996
- Fessler, L. I., Morris, N. P. & Fessler, J. H. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 4905–4909
- Furthmayr, H. & Timpl, R. (1971) *Anal. Biochem.* **41**, 510–516
- Grant, M. E. & Jackson, D. S. (1976) *Essays Biochem.* **12**, 77–113
- Grant, M. E., Kefalides, N. A. & Prockop, D. J. (1972) *J. Biol. Chem.* **247**, 3545–3551
- Harwood, R., Grant, M. E. & Jackson, D. S. (1973) *Biochem. Biophys. Res. Commun.* **55**, 1188–1196
- Harwood, R., Grant, M. E. & Jackson, D. S. (1974a) *Biochem. J.* **142**, 641–651
- Harwood, R., Grant, M. E. & Jackson, D. S. (1974b) *Biochem. J.* **144**, 123–130
- Harwood, R., Grant, M. E. & Jackson, D. S. (1974c) *Biochem. Biophys. Res. Commun.* **59**, 947–954
- Harwood, R., Bhalla, A. K., Grant, M. E. & Jackson, D. S. (1975a) *Biochem. J.* **148**, 129–138

- Harwood, R., Grant, M. E. & Jackson, D. S. (1975b) *FEBS Lett.* **57**, 47–50
- Hulmes, D. J. S., Miller, A., Parry, D. A. D., Piez, K. A. & Woodhead-Galloway, J. (1973) *J. Mol. Biol.* **79**, 137–148
- Jackson, D. S. & Cleary, E. G. (1967) *Methods Biochem. Anal.* **15**, 25–76
- Jimenez, S. A., Dehm, P. & Prockop, D. J. (1971) *FEBS Lett.* **17**, 245–248
- Jimenez, S. A., Dehm, P., Olsen, B. R. & Prockop, D. J. (1973) *J. Biol. Chem.* **248**, 720–729
- Juva, K. & Prockop, D. J. (1966) *Anal. Biochem.* **15**, 77–83
- Kühn, K. (1969) *Essays Biochem.* **5**, 59–87
- Loening, U. E. (1967) *Biochem. J.* **102**, 251–257
- Martin, G. R., Byers, P. H. & Piez, K. A. (1975) *Adv. Enzymol.* **42**, 167–192
- McDevitt, C. & Muir, H. (1971) *Anal. Biochem.* **42**, 461–469
- Merry, A. H., Harwood, R., Woolley, D. E., Grant, M. E. & Jackson, D. S. (1976) *Biochem. Biophys. Res. Commun.* **71**, 83–90
- Miller, E. J. & Matukas, V. J. (1974) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **33**, 1197–1204
- Morris, N. P., Fessler, L. I., Weinstock, A. & Fessler, J. H. (1975) *J. Biol. Chem.* **250**, 5719–5726
- Müller, P. K., Meigel, W. N., Pontz, B. F. & Raisch, K. (1974) *Hoppe-Zeyler's Z. Physiol. Chem.* **355**, 985–996
- Murphy, L. & Rosenbloom, J. (1973) *Biochem. J.* **135**, 249–251
- Myllylä, R., Risteli, L. & Kivirikko, K. I. (1975) *Eur. J. Biochem.* **58**, 517–521
- Neville, D. M. (1971) *J. Biol. Chem.* **246**, 6328–6334
- Olsen, B. R., Hoffmann, H.-P. & Prockop, D. J. (1976) *Arch. Biochem. Biophys.* **175**, 341–350
- Park, E., Tanzer, M. L. & Church, R. L. (1975) *Biochem. Biophys. Res. Commun.* **63**, 1–10
- Peacock, A. C. & Dingman, C. W. (1968) *Biochemistry* **7**, 674–688
- Peterkofsky, B. & Diegelmann, R. F. (1971) *Biochemistry* **10**, 988–994
- Prockop, D. J., Berg, R. A., Kivirikko, K. I. & Uitto, J. (1976) in *Biochemistry of Collagen* (Ramachandran, G. N. & Reddi, A. H., eds.), Plenum Publishing Co., New York, in the press
- Risteli, L., Myllylä, R. & Kivirikko, K. I. (1976) *Biochem. J.* **155**, 145–153
- Rosenbloom, J., Endo, R. & Harsch, M. (1976) *J. Biol. Chem.* **251**, 2070–2076
- Ryhänen, L. & Kivirikko, K. I. (1974) *Biochim. Biophys. Acta* **343**, 129–137
- Schofield, J. D., Uitto, J. & Prockop, D. J. (1974) *Biochemistry* **13**, 1801–1806
- Speakman, P. T. (1971) *Nature (London)* **229**, 241–243
- Tanzer, M. L., Church, R. L., Yaeger, J. A., Wampler, D. E. & Park, E. (1974) *Proc. Natl. Acad. Sci. U.S.A.* **71**, 3009–3013
- Uitto, J. & Prockop, D. J. (1974a) *Eur. J. Biochem.* **43**, 221–230
- Uitto, J. & Prockop, D. J. (1974b) *Biochemistry* **13**, 4586–4591
- Uitto, J., Hoffmann, H.-P. & Prockop, D. J. (1975) *Science* **190**, 1202–1204
- Vuust, J. & Piez, K. A. (1970) *J. Biol. Chem.* **245**, 6201–6207
- Vuust, J. & Piez, K. A. (1972) *J. Biol. Chem.* **247**, 856–862
- Weber, K. & Osborn, M. (1969) *J. Biol. Chem.* **244**, 4406–4412
- Woolley, D. E., Lindberg, K. A., Glanville, R. W. & Evanson, J. M. (1975) *Eur. J. Biochem.* **50**, 437–444