The Chemistry of the Collagen Cross-Links

A CONVENIENT SYNTHESIS OF THE REDUCTION PRODUCTS OF SEVERAL COLLAGEN CROSS-LINKS AND CROSS-LINK PRECURSORS

By NORMAN R. DAVIS and ALLEN J. BAILEY

*Department of Animal Husbandry, University of Bristol, Langford,
and Agricultural Research Council Meat Research Institute,
Langford, Bristol BS18 7DY, U.K.*

(Received 23 March 1972)

New synthetic routes to the reduction products of several collagen cross-links and cross-link precursors are described. By the use of these routes hydroxynorleucine, 5,6-dihydroxynorleucine, hydroxylysino-norleucine and hydroxylysino-hydroxynorleucine can be synthesized via one common synthetic intermediate. The synthetic routes provide a convenient source of these unusual amino acids, as well as confirming the structure of hydroxylysino-hydroxynorleucine and the other lysine-derived residues found in borohydride-reduced collagens.

Davis & Bailey (1971) described the synthesis of *N*^ε-(5-amino-5-carboxy-2-hydroxypentyl)-5-hydroxylysine (G'), which we showed to be identical with the reduction product of syndesine (G), an important intermolecular cross-link of collagen (Bailey *et al.*, 1969, 1970). This unambiguous chemical synthesis established the correct structure of syndesine as an aldimine-type cross-link, and confirmed the structure previously proposed by Mechanic & Tanzer (1970) based on mass-spectral evidence. In addition to syndesine, two other intermolecular cross-links (E, F) found in certain collagens have an aldimine-type structure derived from specific lysine residues and their oxidation products (B, C, D), which can be reduced to stable products (E', F', G') by treatment of collagen fibres with sodium borohydride (Bailey & Peach, 1968, 1971; Tanzer & Mechanic, 1970). The reduction products, although reasonably stable to acid hydrolysis, can be isolated only in small quantities, since they constitute less than one residue per thousand residues in collagen. In addition their purification is expensive, tedious and time-consuming. However, these compounds are useful for studies of the cross-linking mechanism of collagen fibres, both to provide pure standards and to furnish further confirmation of assigned structures.

Methods and Results

Analytical methods

Amino acid analyses were carried out with either a Beckman 120C or Locarte amino acid analyser. Purification of synthetic amino acids was carried out with the latter system, by using the usual citrate buffers (Spackman *et al.*, 1958). Thin layers of cellu-

lose and of silica gel on aluminium sheets (E. Merck A.G., Darmstadt, Germany) were used for the studies of amino acids. The sheets were developed with one of the following solvent systems: propan-2-ol-water (7:3, v/v), butan-1-ol-acetic acid-water (4:1:1, by vol.) or chloroform-methanol-aq. 17% (v/v) NH₃ soln. (2:2:1, by vol.). The solvent systems were employed to follow the reactions involved in these studies. The phenylhydantoins were detected with the aid of a u.v. lamp and a ninhydrin spray was used to detect amino acids.

Periodate oxidations. Each 0.1 μmol of compound to be oxidized was treated with 0.1 ml of 0.01 M-NaIO₄ at room temperature. After 1 h excess of NaIO₄ was destroyed by acidification with HCl and addition of 0.1 ml of 1 M-NaAsO₂. After 10 min, samples were removed for determination of formaldehyde and NH₃, as well as for treatment with NaB³H₄ and subsequent amino acid analysis and radioactivity determination. Tritium radioactivity was determined by adding 0.1 ml of each fraction from the amino acid analyser to 3 ml of Bray's (1960) solution.

Determination of formaldehyde. Portions of the oxidation mixture (1 ml) were distilled to dryness into chromotropic acid reagent from a specially constructed micro-distillation apparatus, and the formaldehyde was determined colorimetrically (Hay *et al.*, 1965).

Determination of ammonia. Ammonia was determined by distillation to dryness of portions (1 ml) (adjusted to pH 12 with 0.1 M-NaOH) of the oxidation mixture into 0.01 M-HCl (2 ml), and measurement of NH₄Cl by the ninhydrin method (Chibnall *et al.*, 1958). Ammonia blanks on the reagents and on unoxidized compounds were also run, and subtracted from the total NH₃, to yield estimates of NH₃ liberated by NaIO₄ oxidation.

Scheme 1. Cross-link precursors, cross-links and their reduced products derived from lysine

Reduction with NaB^3H_4 of periodate oxidation products. Portions of reaction mixture (1ml) were neutralized with 0.1M-NaOH and treated with excess (300 molar) of NaB^3H_4 in 0.1M-NaOH at room temperature. In general, the reduction mixture was acidified after 2h, desalted on Amberlite GC-120, and run on Locarte or Beckman amino acid analysers for product identification and purification.

Synthetic studies

Materials. All solvents used were reagent-grade products. Anhydrous ether and methanol were commercial products and were stored over Linde 5Å molecular sieves. Other solvents were dried by standing over Linde 5Å molecular sieves for at least 2 weeks before use. *m*-Chloroperbenzoic acid and *N*-bromosuccinimide were obtained from BDH Chemicals Ltd. (Poole, Dorset, U.K.). DL-Hydroxylysine (Koch-Light Laboratories Ltd., Colnbrook, Bucks., U.K.) was used without further purification to synthesize *N*^α-trifluoroacetylhydroxylysine methyl ester. Trifluoroacetic acid and its anhydride (BDH) were stored in glass bottles in a sealed desiccator containing P_2O_5 , and used without further purification. Diethylene glycol dimethyl ether (diglyme) was obtained from BDH Chemicals Ltd.

Preparation of trifluoroacetyl methyl esters. These volatile amino acid derivatives were prepared for mass spectral studies as described by Bailey *et al.* (1970).

Mass spectra. All mass spectra were recorded on an L.K.B. GC-MS instrument with direct insertion of the sample into the ion source.

Magnesium chelate. The magnesium chelate of 5-carboxy-3-phenylhydantoin (I) was synthesized exactly as described by Finkbeiner (1964).

5-(But-3-enyl)-3-phenylhydantoin (II). This compound was synthesized from 1-bromo-3-butene and the nucleophilic magnesium chelate (I, Scheme 2), under the conditions established by Finkbeiner (1964) for the alkylation of compound (I) with alkyl bromides. The product (II) was obtained as an impure yellow gum from ethyl acetate extracts of the acidified reaction mixture.

5-(3-Hydroxy-4-bromobutyl)-3-phenylhydantoin (V). Crude compound (II) (10mmol) was dissolved in 2.0ml of tetrahydrofuran and mixed with 0.6ml of water. Tetrahydrofuran was then added, dropwise, until one phase formed, and 1.8g (10mmol) of *N*-bromosuccinimide was added to the well-stirred solution, which was kept at room temperature for 5h.

Water was added to precipitate a crude yellow oil (V) and, when the oil had settled, the water was decanted. The oil was dissolved in a small volume of ethyl acetate, dried over MgSO_4 and kept at -20°C overnight to precipitate any 3-phenylhydantoin that might contaminate the product. Compound (V) was characterized by conversion into 5-hydroxylysine. This was effected by heating 1mmol of compound (V) with 7mmol of conc. aq. NH_3 soln. (sp.gr.0.88) in dioxan (1ml) overnight. Hydrolysis of the product with 2M-KOH (110°C , 18h) yielded 5-hydroxylysine, identified by ion-exchange chromatography and periodate oxidation. As expected of an alkyl bromide, compound (V) formed a precipitate of AgBr when heated with AgNO_3 in ethanol.

Scheme 2. Synthesis of 5-(but-3-enyl)-3-phenylhydantoin (II)

5-(3,4-Epoxybutyl)-3-phenylhydantoin (VI). *m*-Chloroperbenzoic acid (1.1 mmol) and compound (II) (1 mmol) were dissolved in 2 ml of anhydrous ether and left at room temperature for 5 h, with protection from moisture. Ethyl acetate (3 ml) was added, and the chlorobenzoic acid (and any remaining chloroperbenzoic acid) was extracted into excess of 2M-NaOH. The organic layer was quickly washed with NaCl-saturated water, and dried over MgSO₄. The epoxide (VI) (which yielded 5,6-dihydroxynorleucine on hydrolysis) was used as soon as possible for reaction with *N*^α-trifluoroacetylhydroxylysine methyl ester.

Like compound (V), the epoxide derivative was characterized by conversion into hydroxylysine by reaction with aq. NH₃ and subsequent hydrolysis with 2M-KOH.

N^α-Trifluoroacetylhydroxylysine methyl ester. The hydrochloride of this derivative was synthesized by the method of Bailey & Peach (1968) and dissolved in a small volume of water. The free base was formed by raising the pH to 11 and quickly evaporating the water *in vacuo*. The residue was repeatedly evaporated to dryness from methanol and then used as soon as possible for reaction with the 3-phenylhydantoin derivatives (V) or (VI).

N^ε-(2-Hydroxy-5-amino-5-carboxypentyl)-hydroxylysine (G', Scheme 1) (hydroxylysinohydroxynorleucine). The *N*^α-trifluoroacetylhydroxylysine methyl ester was allowed to react with the hydroxybromo derivative (V) or the epoxide (VI) in dimethylformamide at 110°C for 14 h to give crude compound (VII) (Scheme 3). All the dimethylformamide was removed *in vacuo* from crude compound (VII), and the residue was hydrolysed in 2M-KOH (18 h, 110°C) in sealed plastic bottles. The hydrolysates were desalted by addition of HClO₄ to give pH 2, the KClO₄ was removed by centrifugation, and compound G' was purified by chromatography on an extended basic column of the Locarte amino acid analyser. Compound G' was eluted as a double peak in the position of reduced syndesine, by using 0.35M-sodium citrate buffer, pH 5.25, and was desalted on Amberlite GC-120 columns. The purified material proved to be

too hygroscopic to give reproducible elemental analysis, but was characterized on the basis of chromatographic properties and periodate-degradation products.

The synthesized *N*^ε-(2-hydroxy-5-amino-5-carboxypentyl)-5-hydroxylysine (G') was chromatographically indistinguishable from reduced syndesine, isolated from borohydride-reduced collagen, on electrophoresis at pH 3.5 (pyridine-acetate buffer) with a Gamag HVE apparatus (Griffin and George Ltd., London, U.K.) and on t.l.c. with silica-gel sheets developed with butanol-acetic acid-water (*R*_F 0.04), with propan-2-ol-water (*R*_F 0.00) and with chloroform-methanol-water (*R*_F 0.82), as well as on cellulose sheets developed with the same three systems (*R*_F 0.02, 0.00 and 0.73 respectively).

Both reduced syndesine and compound G' yield the same degradation patterns on oxidation with periodate. That is, for each μmol of compound oxidized, 1.7–1.9 μmol of formaldehyde, 1 μmol of NH₃ and 1.5 μmol of [³H]proline are produced. This is precisely what is expected for a compound of this structure.

The trifluoroacetyl methyl ester derivatives of reduced syndesine and compound G' displayed identical mass spectra.

N^ε-(5-Amino-5-carboxypentyl)hydroxylysine (F') (hydroxylysinonorleucine). The synthetic route to this compound paralleled the corresponding routes to hydroxylysinohydroxynorleucine (G'). That is, *N*^α-trifluoroacetyl-lysine methyl ester was heated with compound (V) (or compound VI) in dimethylformamide at 110°C for 14 h, and the dimethylformamide was then removed *in vacuo*. The residue, which contained compound (VIII) (Scheme 3), was then heated with 2M-KOH for 18 h at 110°C, in a sealed plastic bottle, to produce compound F'. This was purified, after removal of potassium salts with HClO₄, by elution from an extended basic column of the Locarte amino acid analyser (cf. purification of compound G'). On analytical runs, compound F' emerged from the Locarte basic column peak in the same position as the hydroxylysinonorleucine from reduced collagen.

5-(4-Hydroxybutyl)-3-phenylhydantoin (III). This compound was prepared as a precursor of 6-hydroxynorleucine by hydroboration of the alkene (II) and

oxidation of the alkylborane intermediates, under the general conditions described by Zweifel & Brown (1963) for simple alkenes. The hydroboration was

Scheme 3. Synthesis of reduced collagen cross-links and their precursors

Method (a) $\text{BH}_3/\text{diglyme}$; method (b) $\text{H}_2\text{O}_2/\text{KOH}$, 1 h; method (c) performic acid, 3 h; method (d) 1 M-KOH, 100°C , 1 h; method (e) *N*-bromosuccinimide in tetrahydrofuran-water, 5 h; method (f) *m*-chloroperbenzoic acid in CHCl_3 , 5 h; method (g) *N* $^\alpha$ -trifluoroacetyl-lysine methyl ester in dimethylformamide, 100°C , 14 h; method (h) *N* $^\alpha$ -trifluoroacetylhydroxylysine methyl ester in dimethylformamide, 100°C , 14 h.

carried out with excess of diborane generated *in situ* from NaBH_4 and $\text{BF}_3 \cdot \text{Et}_2\text{O}$ in dry diglyme (diethylene glycol dimethyl ether), in a nitrogen atmosphere. After a reaction time of 1 h, water and KOH were added, and the alkylboranes were oxidized to the alcohol (III) by dropwise addition of 30% (v/v) H_2O_2 . After 2 h at 30°C, the product was precipitated from solution (as an oil) by addition of several volumes of water. The oil was extracted into ethyl acetate, and the solvent evaporated to leave crude compound (III).

6-Hydroxynorleucine (C'). This was obtained by hydrolysis of compound (III) in 2M-KOH (110°C, 18h), and purified after desalting by ion-exchange chromatography on the acid column of a Locarte amino acid analyser, where it emerged just before glycine.

5-(3,4-Dihydroxybutyl)-3-phenylhydantoin (IV). Acid hydrolysis of the epoxide (VI) with 30% (v/v) HClO_4 in tetrahydrofuran, for 18 h at room temperature, yielded the dihydroxy derivative (IV).

The compound was also synthesized by peroxidation of compound (II) with one equivalent of performic acid in formic acid and subsequent hydrolysis of the mixed α -hydroxy esters to the dihydroxy derivative (IV) (Scheme 3).

5,6-Dihydroxynorleucine (D'). The crude hydantoin derivative (IV) was hydrolysed with 2M-KOH (18h, 110°C) to yield 5,6-dihydroxynorleucine (D'), which was then purified by chromatography on the Locarte amino acid analyser.

An alternative procedure involved the prior hydrolysis of the phenylhydantoin derivative (II) and peroxidation of the resulting $\Delta^{5,6}$ -2-aminohexenoic acid to give 5,6-dihydroxynorleucine. The hydrolysis of compound (II) was done with $\text{Ba}(\text{OH})_2$, at 120°C for 1 h (Gaudry, 1948), and the barium salts were precipitated by treatment with $(\text{NH}_4)_2\text{CO}_3$. The filtrate was converted into 5,6-dihydroxynorleucine by treatment with excess of performic acid (5h, 20°C) and hydrolysis of the evaporated residues with 2M-KOH (1h, 100°C). The crude product was purified on the amino acid analyser. It was eluted from the column between the positions of serine and glutamic acid on the Locarte analyser and at the position of glutamic acid on the Beckman 120C analyser. The compound was eluted from both columns as a twin peak presumably owing to epimerization during hydrolysis, but rather surprisingly possessed a 'brown' ninhydrin colour (maximum colour on 440nm photometer).

Confirmation of the identity of 5,6-dihydroxynorleucine was obtained by oxidation-reduction analysis. Oxidation was carried out with 0.01M- NaIO_4 (1h, 20°C) and the products were reduced with NaB^3H_4 (300 molar excess) in 0.1M-NaOH. Analysis of the products on the amino acid analyser revealed the presence of radioactive proline as expected from a compound of this structure. Final confirmation was

obtained by mass-spectral analysis of the trifluoroacetyl derivative (Fowler *et al.*, 1970).

Discussion

Our initial aim in synthesizing hydroxylysino-hydroxynorleucine was to provide an unambiguous proof of the structure of the reduced form of one of the major cross-links stabilizing collagen fibres, and given the trivial name syndesine (Bailey *et al.*, 1969, 1970). The structure originally proposed on the basis of the mass-spectral data conflicted with the subsequent periodate-degradation studies. When reduced syndesine was treated with periodate, 2 μmol of formaldehyde and approx. 1 μmol of NH_3 were produced per μmol of the compound oxidized. Further, approx. 1.5 μmol of [^3H]proline was produced when the oxidation mixture was treated with NaB^3H_4 . These results suggested that the compound possessed a similar structure to hydroxylysino-hydroxynorleucine, previously isolated from reduced collagen (Bailey & Peach, 1968), but in this case derived from hydroxyallysine rather than allysine. The peroxidation studies were proved to be correct by chemical synthesis of hydroxylysino-hydroxynorleucine, which was shown to have properties identical with those of the reduced syndesine isolated from reduced collagen (Davis & Bailey, 1971).

The cross-link as it exists in the physiologically intact fibre must be the non-reduced form, i.e. an aldimine bond formed by the carbonyl group of hydroxyallysine with the ϵ - NH_2 group of hydroxylysine. The three reducible intermolecular cross-links of tendon collagen thus all appear to be of the aldimine-bond type although the nature of the third cross-link, designated fraction C (Bailey *et al.*, 1970), has not yet been established.

During the investigations into the various possible routes to synthesize these compounds we developed a common route whereby most of the reducible cross-links and their precursors could be synthesized.

It was decided that an effective synthetic strategy would be to prepare a suitable electrophilic derivative of 5-hydroxynorleucine capable of undergoing selective nucleophilic attack at C-6 by a N^{α} -protected derivative of hydroxylysine, particularly since it was apparent that derivatives of 2-amino-5-hexenoic acid could serve as intermediates to several other lysine-derived residues as well as reduced syndesine. The 3-phenylhydantoin derivative (II) of 2-amino-5-hexenoic acid was chosen for study because of the relative stability of the hydantoin ring to a wide variety of reaction conditions, the ease with which hydantoins are converted into the DL-amino acid derivatives, and the very convenient solubility properties displayed by phenylhydantoins. Further, the work of Finkbeiner (1964) suggested that

compound (II) should be readily obtainable from 1-bromo-3-butene and the nucleophilic chelate (I, Scheme 2). Indeed, compound (II) was readily obtained by this route, and served as a common intermediate for the synthesis of 6-hydroxynorleucine (C'), 5,6-dihydroxynorleucine (D') and hydroxylysinnorleucine (F'), as well as hydroxylysino-hydroxynorleucine (G').

All the amino acids obtained from compound (II) are necessarily racemic because alkylation of the optically inactive magnesium chelate (I) yields compound (II) as a racemic mixture. Moreover, alkaline hydrolysis of even optically active phenylhydantoins yield racemic amino acids. In view of the complex mixture of optical isomers expected from synthetic compound G', the chromatographic behaviour was surprisingly simple, single spots being obtained in all paper-chromatographic and electrophoretic systems, and a double peak on the extended ion-exchange columns.

The development of these synthetic routes makes these compounds available in relatively large quantities with far less effort than the same compounds isolated from NaB³H₄-reduced collagen. Needless to say, any suitably protected derivative of 2-amino-5-hexenoic acid could serve as a source of compounds C', D', F' and G' by using the general reactions outlined in Scheme 3.

Other routes to hydroxynorleucine (Gaudry, 1948), to dihydroxynorleucine (Fowler *et al.*, 1970; Mechanic & Tanzer, 1970) and to hydroxylysinnorleucine (Bailey & Peach, 1968) have been described previously. However, the availability of all three deriva-

tives, as well as reduced syndesine, from one easily obtained intermediate (II) is an especially attractive feature of the reaction schemes described here.

We thank Professor T. K. Ewer for his generous encouragement, Dr. S. M. Partridge, F.R.S., for his helpful advice, and the Wellcome Foundation for financial support to N. R. D.

References

- Bailey, A. J. & Peach, C. M. (1968) *Biochem. Biophys. Res. Commun.* **33**, 812
- Bailey, A. J. & Peach, C. M. (1971) *Biochem. J.* **121**, 257
- Bailey, A. J., Fowler, L. J. & Peach, C. M. (1969) *Biochem. Biophys. Res. Commun.* **35**, 663
- Bailey, A. J., Peach, C. M. & Fowler, L. J. (1970) *Biochem. J.* **117**, 819
- Bray, G. A. (1960) *Anal. Chem.* **1**, 279
- Chibnall, A. C., Manger, A. E. & Rees, M. W. (1958) *Biochem. J.* **68**, 111
- Davis, N. R. & Bailey, A. J. (1971) *Biochem. Biophys. Res. Commun.* **45**, 1416
- Finkbeiner, H. (1964) *J. Amer. Chem. Soc.* **86**, 961
- Fowler, L. J., Peach, C. M. & Bailey, A. J. (1970) *Biochem. Biophys. Res. Commun.* **41**, 251
- Gaudry, P. (1948) *Can. J. Res. Sect. B* **26**, 287
- Hay, S. W., Lewis, B. A. & Smith, F. (1965) *Methods Carbohyd. Chem.* **5**, 379
- Mechanic, G. & Tanzer, M. L. (1970) *Biochem. Biophys. Res. Commun.* **41**, 1597
- Spackman, D. H., Stein, W. H. & Moore, S. (1958) *Anal. Chem.* **30**, 1185
- Tanzer, M. L. & Mechanic, G. (1970) *Biochem. Biophys. Res. Commun.* **39**, 183
- Zweifel, G. & Brown, H. C. (1963) *Org. React.* **13**, 1