Nagel, R. L., Rothman, M. C., Bradley, T. B., Jr., and Ranney, H. M. (1965b), *J. Biol. Chem. 240*, PC4543.
Nagel, R. L., Wittenberg, J. B., and Ranney, H. M. (1965a), *Biochim. Biophys. Acta 100*, 286.
Ogawa, S., and McConnell, H. M. (1967), *Proc. Nat. Acad. Sci. U. S. 58*, 19.
Peacock, A. C., Pastewka, J. V., Reed, R. A., and Ness, A. T. (1970), *Biochemistry 9*, 2275.
Perutz, M. F. (1965), *J. Mol. Biol. 13*, 646.
Perutz, M. F., and Mazzarella, L. (1963), *Nature (London) 199*, 639.
Perutz, M. F., Muirhead, H., Mazzarella, L., Crowther, R. A., Greer, J., and Kilmartin, J. V. (1969), *Nature (London) 222*, 1240.
Platt, J. R. (1956), *in* Radiation Biology, Hollaender, A., Ed., Vol. III, New York, N. Y., McGraw-Hill, Chapter 2, p 71.
Polonovski, M., and Jayle, M. F. (1938), *C. R. Soc. Biol. 129*, 457.
Smith, M. J., and Beck, W. S. (1967), *Biochim. Biophys. Acta 147*, 324.
Tenhunen, R., Marver, H. S., and Schmid, R. (1969), *J. Biol. Chem. 244*, 6388.
Ueda, Y., Shiga, T., and Tyuma, I. (1970), *Biochim. Biophys. Acta 207*, 18.
Urry, D. W., and Pettegrew, J. W. (1967), *J. Amer. Chem. Soc. 89*, 5276.
Van Royen, A. H. H. (1950), Dissertation, Delft.
Willick, G. E., Schonbaum, G., and Kay, C. M. (1969), *Biochemistry 8*, 3729.
Willstätter, R., and Pollinger, A. (1923), *Hoppe-Seyler's Z. Physiol. Chem. 130*, 281.

Structural Features of Tropoelastin Related to the Sites of Cross-Links in Aortic Elastin*

L. B. Sandberg,† N. Weissman,‡ and W. R. Gray§

ABSTRACT: Tropoelastin is a soluble protein obtained from the aortas of copper-deficient swine. It is considered to be the precursor of insoluble elastin. Its isolation and purification has previously been described. When digested with trypsin, a number of small peptides are obtained. These have been isolated by ion-exchange chromatography and quantitated by amino acid analysis. Their amino acid sequence has been determined by the dansyl-Edman method and by mass spectrometry. Two of the peptides, Ala-Ala-Ala-Lys and Ala-Ala-Lys, appear to be repeated six times in the polypeptide chain. Because of this repeating structure, we propose that these peptides represent the areas of the chain involved in the formation of the desmosine and isodesmosine cross-links of insoluble elastin. We have proposed a linear folded model of tropoelastin which is compatible with these repeating structures.

We have recently isolated a protein from the aortas of copper-deficient swine which has elastin-like physical and chemical properties. The protein was given the name tropoelastin because it was considered to be the soluble precursor of insoluble elastin. Its molecular weight has been demonstrated to be approximately 68,000. Fingerprints of elastase-treated tropoelastin and insoluble elastin showed that there are many amino acid sequences common to both of these preparations. Edman degradation of tropoelastin has shown it to be a homogeneous protein with a unique N-terminal amino acid sequence (Sandberg *et al.*, 1969).

Tropoelastin does not have the desmosine cross-links present in insoluble elastin, but it does have a high content of lysine, more than sufficient for the production of the desmosine cross-links. During the process of cross-linking, the ϵ-amino groups of these lysine residues become altered by an enzyme, lysyl oxidase, to produce α-aminoadipic acid δ-semialdehyde (allysine) (Siegel and Martin, 1970; Pinnell and Martin, 1968). This latter substance is an intermediate in the formation of desmosine and other elastin cross-links (see review of Piez, 1968, and Miller *et al.*, 1967). About one-third of the lysine residues of tropoelastin is converted into desmosine cross-links. A large portion of the remainder of the lysine is also apparently oxidized to allysine and this then participates in cross-links which have not been completely elucidated. Some of these cross-links are of the aldol condensation type (Lent *et al.*, 1969). Lysinonorleucine and merodesmosine have also been identified as other possible cross-links (Franzblau *et al.*, 1969; Starcher *et al.*, 1967).

The present work is an evaluation of part of the structure of porcine aortic tropoelastin. It deals with portions of the molecule which are possibly destined to form the desmosine cross-links. These have been investigated as small lysine-containing peptides obtained by trypsin digestion of the tropoelastin molecule. Several peptides are present to the extent of two moles per mole of protein (68,000 molecular weight), suggesting some kind of duplicated structure. At present, it is not clear whether the duplication is an end-to-end

* From The Departments of Surgery and Pathology (L. B. S. and N. W.), University of Utah Medical Center, Salt Lake City, Utah 84112, and The Division of Biology, California Institute of Technology, Pasadena, California 91109 (W. R. G.). *Received July 16, 1970.* Supported by Grants from the American Heart Association and National Institutes of Health, HE 11963-02S1, HE 13218-01, and GM 06965.

† Established Investigator of the American Heart Association; to whom to address correspondence.
‡ Present address: Bio-Science Laboratories, Van Nuys, Calif. 91405.
§ Present address: Department of Biology, University of Utah.

genetic duplication, or merely that the 68,000 molecular weight is based on incomplete dissociation of subunits of 34,000 molecular weight. Two peptides are present in even larger amounts, more than 6 moles/mole, and we feel that these are likely to be intimately involved in the process of desmosine formation.

Experimental Section

Materials. Tropoelastin was prepared from the aortas of copper-deficient swine as previously described (Sandberg *et al.*, 1969). Trypsin (Worthington), TPCK treated, was utilized for all tryptic digestions. Aminex A-5 resin is a sulfonated polystyrene spherical resin with bead diameters ranging from 11 to 15 μ (Bio-Rad Corp.). DNS-amino[1] acid standards were purchased from the Calbiochem Corp. Thin-layer DNS-amino acid chromatography was performed on Eastman Chromatogram sheets 6061, silica gel, without fluorescent indicator.

Tryptic Digestion. Tryptic peptides were produced by first dissolving tropoelastin at a concentration of 10 mg/ml in 0.2 M ammonium bicarbonate (pH 8.0). Trypsin was then added to give an enzyme to substrate ratio 1:50. Digestion was allowed to proceed 4 hr at 37°. At the end of this time, the mixture was lyophilized.

Fingerprint and Amino Acid Analyses. Trypsin-digested tropoelastin (2 mg) was spotted on Whatman No. 3MM paper and chromatography was performed using 1-butanol–glacial acetic acid–water (3.4:1:5, v/v) for 18 hr. After drying, the strip containing the peptides was sewn into a second piece of Whatman paper (47 × 200 cm) and electrophoresis was performed in a Gilson Model DW high-voltage electrophorator using 1.64 M formic acid (pH 1.58) (Dreyer and Bynum, 1967), at 30° and 22 V/cm (3000 V) for 3 hr. Peptides were developed with a cadmium–ninhydrin reagent (Dreyer and Bynum, 1967). During this electrophoresis, fast-moving peptides migrated approximately 50 cm. A small sample of a standard amino acid mixture Pierce Chemical Company) was run concomitantly with the second-dimension electrophoresis. Amino acid analyses were performed as previously reported (Sandberg *et al.*, 1969).

Ion-Exchange Chromatography of Tryptic Peptides. Separation of peptides produced by trypsin cleavage was achieved by utilizing a 30 × 0.9 cm column of Aminex A-5 resin. A system of volatile buffers essentially as described by Schroeder *et al.* (1962) was used for development of the column so that the eluted peptides could be directly analyzed without desalting, thus assuring better recovery. A two-chamber gradient system was utilized, each chamber containing 900 ml of buffer. The mixing chamber contained 0.2 M pyridine adjusted to pH 3.01 with acetic acid. The reservoir contained 2.0 M pyridine adjusted to pH 5.0 with acetic acid. Flow rates were maintained at 40 ml/hr with a Milton-Roy piston pump. Column temperature was 30°. The trypsin digest of tropoelastin was dissolved in 1 ml of the starting buffer and applied to the column by the use of a Chromatronix sample valve (Chromatronix Corp.). Samples were collected in 5.0-ml quantities. Peptide peaks were located by removing a known amount from each sample tube collected and carrying out ninhydrin analyses with the use of a Technicon AutoAnalyzer (Schroeder *et al.*, 1962). Peptides were not hydrolyzed prior to ninhydrin analyses.

[1] Abbreviations used are: DNS, dimethylaminonaphthalenesulfonyl derivative; TPCK, L-(1-tosylamido-2-phenyl)ethyl chloromethyl ketone.

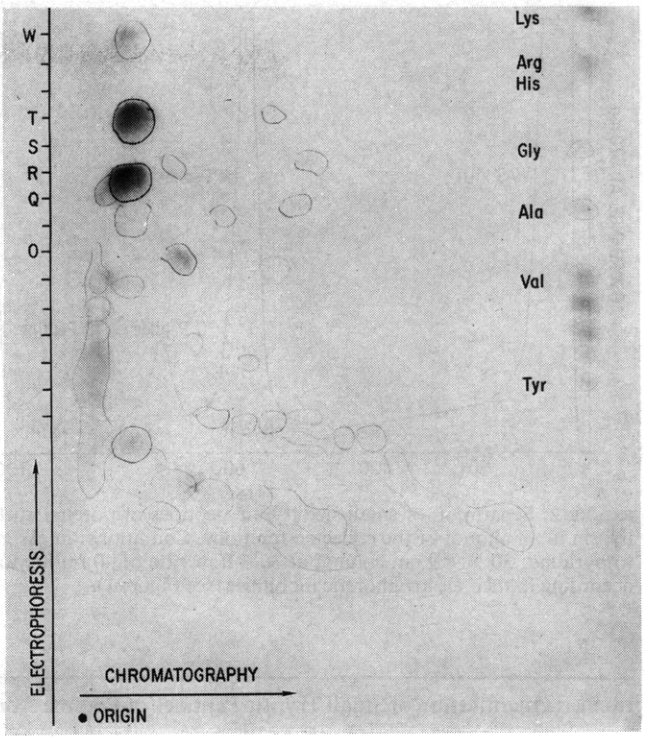

FIGURE 1: A fingerprint of 2 mg of trypsin-digested porcine aortic tropoelastin. Chromatography was performed first in a butanol–acetic acid–water system followed by electrophoresis in 1.64 M formic acid (pH 1.58). An amino acid standard was included in the electrophoresis to evaluate relative migration.

Sequence Analyses. The rapid sequence technique of Gray and Smith was utilized which employs a modified sequential Edman degradation plus a dansylation procedure (Gray and Smith, 1970). Dansylamino acids were identified initially by high-voltage electrophoresis (Gray, 1967). This gave good identification of all dansyl derivatives except DNS-serine, DNS-alanine, and DNS-proline. When these residues were encountered, a second step of identification was carried out by thin-layer chromatography using solvent system A of Morse and Horecker (1966). The spot in question was cut out of the electrophoresis paper and eluted with 0.1 N acetic acid. After drying *in vacuo*, the DNS-amino acid was dissolved in 50% pyridine and applied to the thin-layer chromatographic sheet.

Mass Spectrometry. The peptides (approximately 100 nmoles) were dissolved in 0.4 ml of 50% pyridine (v/v). Acetic anhydride (40 μl of an equimolar mixture of $(CH_3 \cdot CO)_2O$ and $(CD_3 \cdot CO)_2O$) was added, the addition being repeated after an interval of 5 min. The solutions were allowed to react for a further 30 min at 25°, and were then dried *in vacuo*. Permethylation was carried out using methylsulfinyl carbanion and methyl iodide (Hakomori, 1964; Thomas *et al.*, 1968; Gray and del Valle, 1970), the permethylated peptides being extracted into chloroform and dried *in vacuo*.

The samples were redissolved in 50 μl of chloroform, transferred to a direct insertion probe, and the solvent was removed in a stream of nitrogen. Spectra were recorded on an AEI Model MS9 mass spectrometer, using a resolving power of approximately 2000 (10% valley) and an ionizing voltage of 70 eV. Suitable spectra were obtained for R and T peptides at probe temperatures of 160 and 130°, respectively (see Results for nomenclature of peptides).

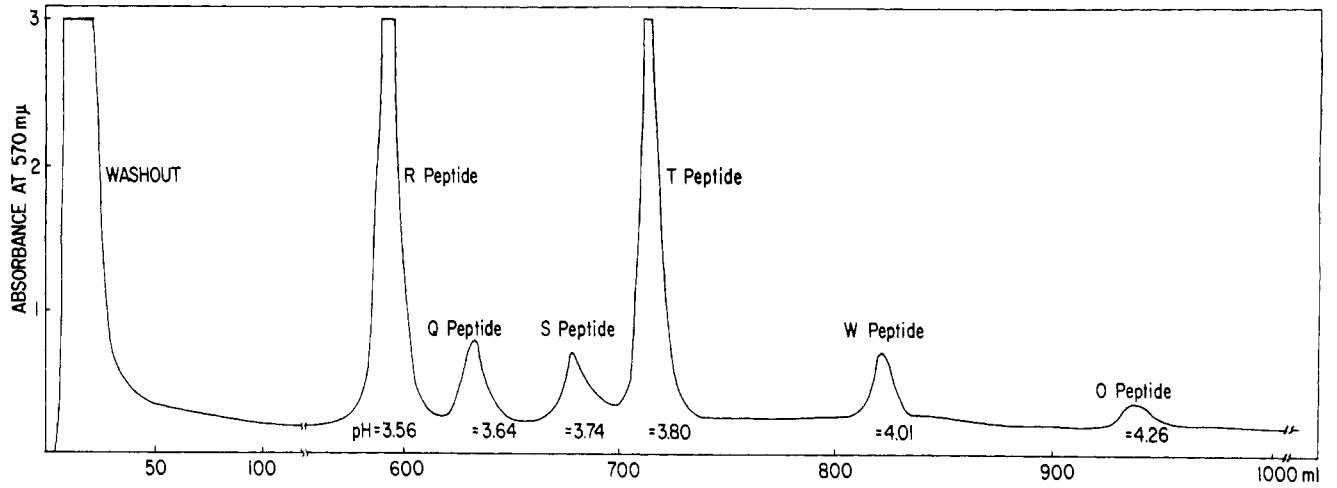

FIGURE 2: Separation of small-size tryptic peptides of porcine aortic tropoelastin on Aminex A-5 resin. Peptides were developed with ninhydrin in an aliquot of the collected fractions. Conditions: linear gradient of pyridine–acetate buffers (pH 3.01–5.00), 0.2–2.0 M with respect to pyridine; 30 × 0.9 cm column at 30°; flow rate of 40 ml/hr; total buffer volume in two chambers, 1800 ml. Peptides have been lettered according to their electrophoretic mobilities (see Figure 1).

TABLE I: Quantitation of Small Tryptic Peptides of Porcine Aortic Tropoelastin Resolved on Aminex A-5 Resin.

Peptide	Structure	μmoles of Lys Recov[a]	Moles of Peptide Recov/Mole of Protein (800 Residues)	
			78% Digestion	100% Digestion
Washout		12.72	19.5	15.2
R	Ala-Ala-Ala-Lys	4.18	6.4	5.0
Q	Ser-Ala-Lys	1.33	2.0	1.6
S	Ala-Pro-Gly-Lys	1.38	2.1	1.7
T	Ala-Ala-Lys	4.22	6.5	5.1
W	Ala-Lys	0.62	1.0	0.7
O	Tyr-Gly-Ala-Arg	1.29 (as Arg)	2.0	1.5
		24.45 = 78% of Lys in applied sample		

[a] As determined by amino acid analysis. Sample applied contained 31.19 μmoles of Lys.

Results

Fingerprint Pattern. Figure 1 represents a fingerprint of a tryptic digest of porcine aortic tropoelastin. The relative electrophoretic mobility of peptides has been designated by letters of the alphabet. Peptides which migrate in the region from O to W are those which are of prime interest in this study. Two very dark peptide spots are noted at positions R and T, and these have been named accordingly. From the appearance of this fingerprint, it seemed likely that the R and T spots represented peptides present at a higher molar concentration than the other peptides visualized on this fingerprint.

Quantitation on Aminex A-5 Resin. An attempt was made to quantitate the amount of peptides R and T, as well as other small peptides present in lesser amounts in the tryptic hydrolysate. These peptides were resolved from the total hydrolysate by ion-exchange chromatography on Aminex A-5 resin and identified by rerunning on electrophoresis. Figure 2 represents the separation achieved of six small peptides which occur in the tryptic hydrolysate. It is immediately apparent that R and T are present in much higher concentration than the other four peptides. Quantitation of all six peaks was accomplished by pooling every two tubes in the areas from peak R to peak T. This represented 21 pooled samples. Peaks W and O, because of much greater separation, were each pooled as one lot. Amino acid analysis was then carried out on the pooled samples. Corrections were made for minor contaminants after the amino acid sequences of the peptides had been determined. Contamination was estimated to be no greater than six per cent for any one peak. After the corrections, quantitation was determined on the basis of lysine content for peptides R, Q, S, T, and W. Peptide O was measured on the basis of arginine content.

Molar Ratios of Peptides. The results of the quantitation experiment are shown in Table I. Seventy-eight per cent of the 31.19 μmoles of lysine applied in the total digest was recovered from the ion-exchange column, including the lysine contained in the washout peak. The 6 small peptides contained 48%

of the recovered lysine and 37% of the lysine originally applied to the column.

Amino acid analyses of several tropoelastin preparations indicate an average lysine content of 37.4 moles/800 residues (the approximate molecular size as determined by gel electrophoresis in the presence of sodium dodecyl sulfate). The molar amounts of the various peptides are calculated in two ways in Table I. In column 5, it is assumed that 100% of the small peptides was recovered, and that all losses are of larger insoluble material; the figures given are thus the most conservative ones, and there is an absolute minimum of 5 moles/mole of peptides R and T. Column 4 lists values obtained on the assumption of equal recoveries of all peptides.

Sequential Degradation. The dansyl-Edman rapid sequence method for small peptides yielded the results shown in Table I, column 2. Peptides R and T represent a polyalanine-lysine sequence. It was felt that some ambiguity might result from determination of these two-peptide sequences by sequential degradation. Therefore, these two peptides were further analyzed by mass spectrometry.

Mass Spectra of R and T. Presence of the deuterium label given each peptide prior to mass spectrometry gave rise to characteristic doublets of equal-sized peaks at M and M + 3 amu for all fragments containing one acetyl group—and, therefore, having either the amino-terminal residue or the carboxyl-terminal lysine residue. Fragments containing both residues had two acetyl groups, and gave three peaks at M, M + 3, and M + 6 amu in the expected ratios of 1:2:1.

The spectra of the peptide derivatives were completely consistent with the amino acid sequences determined by sequential degradation. Peptide R gave three triplets corresponding to ions containing two acetyl groups: at *m/e* 527/530/533 (the molecular ion of fully acetylated and methylated Ala-Ala-Ala-Lys); at *m/e* 496/499/502 (loss of OCH_3 from the molecular ion); and *m/e* 470/473/476 (loss of CH_3—N=CH—CH_3 from the molecular ion). Analogous triplets were observed for peptide T, but were 85 mass units lighter, due to the lack of one alanine residue. In both cases, the three most intense doublets were *m/e* 100/103 (acetyl-Ala(NMe) — CO), 128/131 (acetyl-Ala(NMe)-), and 213/216 (acetyl-Ala(NMe)-Ala(NMe)-). A moderately strong doublet at *m/e* 298/301 (acetyl-Ala(NMe)-Ala(NMe)-Ala(NMe)-) was present in the spectrum of R, but not of T. Several other minor doublets were present also, almost all of which were fragments containing the carboxyl-terminal lysine residue. Spectra obtained from the other peptides also indicated the sequences given in Table I.

Discussion

Our results on the quantitation of the small tryptic peptides (Table I), indicate clearly the presence of multiple repeats of two of these peptides, R (Ala-Ala-Ala-Lys) and T (Ala-Ala-Lys). Three others were present at 2 moles/mole of protein, based on a measured molecular weight of 68,000 (Sandberg *et al.*, 1969). This raises the question of whether the molecular weight truly represents the smallest chemical subunit of the molecule, or whether there is a smaller unit of about 34,000 molecular weight. The value of 68,000 was based on gel electrophoresis in the presence of sodium dodecyl sulfate (Figures 1 and 2, Sandberg *et al.*, 1969); occasional preparations showed small amounts of a material having mol wt 35,000. It is possible that the purification procedure, involving reversible heat precipitation, results in the selective isolation of a dimerized molecule. This could easily arise by cross-linking due to traces of lysyl oxidase remaining in the copper deficient tissues; the lysine content of tropoelastin preparations varies somewhat around a mean of 37 residues/68,000 molecular weight, though none of the more common cross-links are detected. Only 1 mole of peptide W (Ala-Lys) was found in this experiment, which suggests that the higher molecular weight might be correct. It should be noted, however, that there are at least 15 moles of peptides ending in Ala-Lys, so that a small percentage of nontryptic cleavage could give rise to significant amounts of the dipeptide. Because of these uncertainties and the importance of the question, we are undertaking a thorough reevaluation of the molecular weight of the protein. If the higher value is confirmed, it would seem most likely that the molecule results from end-to-end duplication of a smaller unit at the genetic level.

Whichever molecular weight is correct, the peptides R and T present a very striking phenomenon. There is an absolute minimum of 5 moles of each per 68,000 (Table I, column 5), based on the most unfavorable assumptions concerning digestion and recovery of the peptides (100% both). If it is assumed that the yields were only 80%, which was the overall yield measured in the experiment, there are about 6.5 moles of each peptide. This figure could be even higher, since there is probably some loss due to oxidation by residual lysyl oxidase in the tissues. It is also very striking that these peptides are composed exclusively of alanine and lysine. Two recent reports (Keller *et al.*, 1969; Shimada *et al.*, 1969) have indicated that large desmosine-containing peptides from mature elastin are relatively rich in alanine.

Almost half of the total lysine residues of the tropoelastin molecule are released in the form of the small tryptic peptides, most of them as R and T. Practically all of these small peptides must be preceded by lysine also, because trypsin cleaves exclusively at lysine and arginine residues, and there are relatively few of the latter in the molecule. Thus it is clear that there are many sequences such as Lys-Ala-Ala-Lys and Lys-Ala-Ala-Ala-Lys in the molecule. If none of these are contiguous, all of the lysine residues could be accounted for in this way. It is possible, however, to construct a very economical model for the tropoelastin molecule based on the assumption that the peptides R and T are contiguous, and that the very basic sequences such as Lys-Ala-Ala-Lys-Ala-Ala-Ala-Lys are the principle source of the desmosine cross-links.

We envisage three or possibly four such sequences per subunit of 34,000 molecular weight (Figure 3A). For this discussion, it is not essential to know whether the polypeptide monomer has a molecular weight of 34,000 or 68,000; it is essential to note, however, that even if the lower figure turns out to be correct, there are still sufficient R and T peptides to provide three cross-linking points per chain. At least three must be present to generate the extended matrix characteristic of mature elastin unless a single cross-link can hold three chains together. Two simple cross-links would merely generate a very long linear molecule.

By having the desmosine precursor sequences very compact and well separated in the molecule, we allow ample room for the intervening peptide chain to fold into a globular conformation. A cross-linked network of such molecules would form a highly effective elastomer, as shown in Figure 3B. We feel that such a model has an advantage over those involving the distribution of cross-links over the surface of a sphere (Partridge, 1966). There is much greater flexibility in our model to allow the two chains to meet with their desmosine precursor sequences in the correct orientation

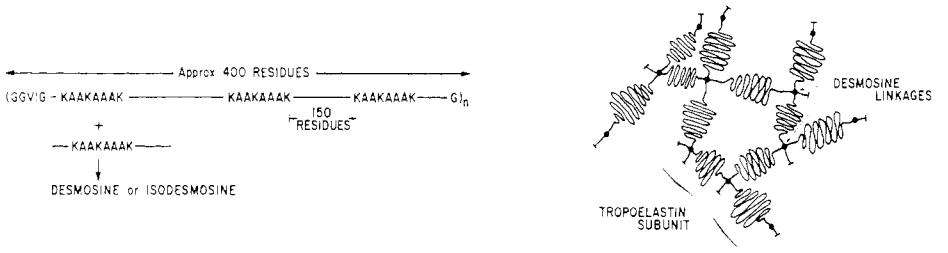

FIGURE 3: (A, left) A proposed model for a 400-residue polypeptide chain of porcine aortic tropoelastin. The sequences K-A-A-K-A-A-A-K are repeated at least three times in the chain. They are produced by peptides R and T occurring in juxtaposition. It is proposed that two such sequences occurring in adjacent chains and acted upon by a lysyl oxidase enzyme are necessary for the formation of desmosine or isodesmosine. An 800-residue model would consist of an end-to-end duplication of this model. (B, right) A proposed model for insoluble elastin containing the 400-residue subunits of tropoelastin proposed in part A. The folded globular regions of the polypeptide chain are presumed to be random coils due to the interactions of regions of the chain rich in amino acids with hydrophobic properties.

and still leave the cross-link areas accessible to enzymatic attack, or completion. It is of interest in this regard that a major product of elastase digestion of insoluble elastin is Ala-Ala-Ala, which we have postulated as one of the sequences bridging the "arms" of the desmosine complex in insoluble elastin (L. B. Sandberg, unpublished data, 1968).

Close proximity of several lysines in the peptide chain is obviously desirable when it is remembered that four lysine residues participate in the formation of each desmosine cross-link. Such a grouping might also be essential for activity of the lysyl oxidase. If each desmosine or isodesmosine link requires the participation of two RT (or TR) sequences from different chains, there is just sufficient R and T to account for the amount of these cross-links formed in mature elastin (16.1 moles of R and T per 1000 residues *vs.* 4.1 moles of desmosine and isodesmosine). In such a scheme it has to be assumed that only four of the six lysine residues are incorporated into the final desmosine structure; three of these are first converted into allysine, while the fourth is unoxidized and contributes its amino group to the pyridine nucleus. There are many different possibilities for which four of the six lysines become utilized and a random choice among several such combinations might very well result in the almost equimolar mixture of desmosine and isodesmosine which is found naturally.

In our hypothesis, desmosine and isodesmosine are the primary intermolecular cross-links responsible for the elastomeric properties of mature elastin. Most of the other cross-links require only two lysine residues for their formation. We have shown that nearly all lysines must be present in pairs separated by only two or three residues. It seems likely, therefore, that cross-links such as the aldol condensation product of two allysine residues will form intramolecularly from such a pair. Because they could span only a few residues of a single chain, our model would consider them as knots rather than bridges.

Acknowledgments

The expert technical assistance of Mrs. Ruth Ann Green, Mrs. Sandra Sweeney, and Mr. Ursino del Valle is greatly appreciated.

References

Dreyer, W. J., and Bynum, E. (1967), *Methods Enzymol. 11*, 32.

Franzblau, C., Faris, B., and Papaioannou, R. (1969), *Biochemistry 8*, 2833.

Gray, W. R. (1967), *Methods Enzymol. 11*, 139.

Gray, W. R., and del Valle, U. (1970), *Biochemistry 9*, 2134.

Gray, W. R., and Smith, J. F. (1970), *Anal. Biochem. 33*, 36.

Hakomori, S. (1964), *J. Biochem. (Tokyo) 55*, 205.

Keller, S., Levi, M. M., and Mandl, I. (1969), *Arch. Biochem. Biophys. 132*, 565.

Lent, R. W., Smith, B., Salcedo, L. L., Faris, B., and Franzblau, C. (1969), *Biochemistry 8*, 2837.

Miller, E. J., Pinnell, S. R., Martin, G. R., and Schiffman, E. (1967), *Biochem. Biophys. Res. Commun. 26*, 132.

Morse, D., and Horecker, B. L. (1966), *Anal. Biochem. 14*, 429.

Partridge, S. M. (1966), *Fed. Proc., Fed. Amer. Soc. Exp. Biol. 25*, 1023.

Piez, K. A. (1968), *Annu. Rev. Biochem. 37*, 547.

Pinnell, S. R., and Martin, G. R. (1968), *Proc. Nat. Acad. Sci. U. S. 61*, 708.

Sandberg, L. B., Weissman, N., and Smith, D. W. (1969), *Biochemistry 8*, 2940.

Schroeder, W. A., Jones, R. T., Cormick, J., and McCalla, K. (1962), *Anal. Chem. 34*, 1570.

Shimada, W., Bowman, A., Davis, N. R., and Anwar, R. A. (1969), *Biochem. Biophys. Res. Commun. 37*, 191.

Siegel, R. C., and Martin, G. R. (1970), *J. Biol. Chem. 245*, 1653.

Starcher, B. C., Partridge, S. M., and Elsden, D. F. (1967), *Biochemistry 6*, 2425.

Thomas, D. W., Das, B. C., Gero, S. D., and Lederer, E. (1968) *Biochem. Biophys. Res. Commun. 32*, 199.