Isolation and Amino Acid Sequences of Tropoelastin Peptides*

(Received for publication, August 22, 1972)

JUDITH ANN FOSTER,‡ EVELINE BRUENGER, WILLIAM R. GRAY, AND LAWRENCE B. SANDBERG§

From the Departments of Surgery, Pathology, and Biology, University of Utah, Salt Lake City, Utah 84112

SUMMARY

Seventeen peptides have been isolated from a tryptic digest of soluble elastin by ion exchange chromatography and gel filtration. Molecular weights range from approximately 1,700 to 10,500. Partial or complete amino acid sequences of the peptides were obtained and provide the first extensive data on the primary sequence of elastin. Several important findings stand out. First, the sequences obtained reveal an elastin primary structure quite distinct from collagen and possessing repeats of a tetrapeptide, Gly-Gly-Val-Pro, a pentapeptide, Pro-Gly-Val-Gly-Val, and a hexapeptide, Pro-Gly-Val-Gly-Val-Ala. Second, several of the peptides contain a partial hydroxylation of proline residues thereby definitely establishing hydroxyproline as a real constituent of the elastin chain.

Investigations into the primary structure of mature elastin have been greatly hindered because of its extreme insolubility and highly cross-linked nature. A major advance in elastin research came with the isolation of an elastin-like, soluble component from the aortas of copper-deficient swine (1, 2). Further purification and characterization of this soluble protein have led to the proposal that it indeed is the precursor to insoluble elastin and accordingly, it has been designated tropoelastin (3).

Soluble elastin, as well as the mature elastin, is rich in glycine (33%) and the nonpolar amino acids alanine, valine, and proline. It differs from the insoluble protein in its high content of lysine residues and lack of any cross-links.

Recently, we have reported on a strong clustering of alanine and lysine residues in the soluble protein, as evidenced by the structures of small tryptic peptides (4), and of the carboxy-terminal fragments of the large tryptic peptides (5). This clustering of alanine near the lysine residues of the molecule is in agreement with the isolation of alanine-enriched, cross-linked peptides from mature elastin (6-9). However, very little is yet known of the primary sequence of elastin outside of the cross-linked areas.

* This work was supported by grants from the Utah and American Heart Associations, the National Institutes of Health (FR 07092, HE 11963, HE 13218, and HE 12561), and the National Science Foundation (GB 20983).

‡ Recipient of National Institutes of Health Postdoctoral Fellowship AM 51051. Present address, Boston University School of Medicine, Department of Biochemistry, Boston, Mass. 02118.

§ Established Investigator of the American Heart Association.
polarizing a continuous sodium citrate gradient (pH 2.75 to 6.80; ACIDS AMINO ACIDS). Values are expressed as residues per peptide. A space indicates less than 0.2 residue. Residues in parentheses are fractional residues thought to be impurities.

Amino Acid Sequencing of Tryptic Peptides

Because of the large size and restricted composition of the peptides (see Table I), sequencing was carried out with automatic stepwise degradation (Beckman model 890 Sequencer) according to the procedure of Edman and Begg (12).

The extreme hydrophobicity of the peptides led us to use the procedure of Braunitzer et al. (13) for rendering peptides less extractable. However, instead of the Quadrol buffer system we used the volatile buffers based on dimethylallylamine to further minimize extraction losses. Despite these precautions the band of peptide was seen to move up the Sequencer cup as degradation proceeded, and yields dropped sharply after about 10 to 15 cycles. We countered this problem by using a "peptide drop" procedure of Bronzert (16). Columns (6 feet, 2 mm inside diameter) were conditioned on a Packard Series model 7400 Gas Chromatographic System, using a modification of the method of Pisano and Bronzert (16). Peaks were identified on a Packard Series Model 7400 Gas Chromatograph using a modification of the method of Pisano and Bronzert (16). Columns (6 feet, 2 mm inside diameter) were conditioned at 300 °C for 18 hours. Flow rate of the carrier gas (N2) was 40 cc per min.

Gas-Liquid Chromatography—PTH-amino acids were identified on a Packard Series model 7400 Gas Chromatographic System, using a modification of the method of Pisano and Bronzert (16). Columns (6 feet, 2 mm inside diameter) were packed with 60% Chromosorb WHP, 80 to 100 mesh, Supelco, Inc.) and conditioned at 300 °C for 18 hours. Flow rate of the carrier gas (N2) was 40 cc per min.

Amino Acid Analysis—The remaining of the PTH-amino acids was hydrolyzed to free amino acids by the procedure of Smithies et al. (15). By carrying out hydrolysis on the ethyl acetate-soluble fraction, rather than on the entire contents of the Sequencer tubes, we were able to minimize contamination of the samples by peptide material washed out of the cup. This was an especially bad problem with the hydrophobic peptides analyzed in this work. The aqueous fraction was occasionally

<table>
<thead>
<tr>
<th>PEPTIDES</th>
<th>T1</th>
<th>T2</th>
<th>T4</th>
<th>T6</th>
<th>T7</th>
<th>T9</th>
<th>T10</th>
<th>T11</th>
<th>T12</th>
<th>T13</th>
<th>T14</th>
<th>T15</th>
</tr>
</thead>
<tbody>
<tr>
<td>LYS</td>
<td>1.0</td>
<td>1.0</td>
<td>2.8</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.3</td>
<td>1.3</td>
<td>1.0</td>
<td>1.2</td>
<td>1.0</td>
<td>1.7</td>
</tr>
<tr>
<td>ARG</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
</tr>
<tr>
<td>HYP</td>
<td>2.4</td>
<td>0.7</td>
<td>0.7</td>
<td>0.7</td>
<td>0.7</td>
<td>0.7</td>
<td>0.7</td>
<td>0.7</td>
<td>0.7</td>
<td>0.7</td>
<td>0.7</td>
<td>0.7</td>
</tr>
<tr>
<td>ASX</td>
<td>1.2</td>
<td>2.8</td>
<td>1.2</td>
<td>2.8</td>
<td>1.2</td>
<td>2.8</td>
<td>1.2</td>
<td>2.8</td>
<td>1.2</td>
<td>2.8</td>
<td>1.2</td>
<td>2.8</td>
</tr>
<tr>
<td>THR</td>
<td>1.0</td>
<td>1.0</td>
<td>2.6</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>SER</td>
<td>0.6</td>
<td>0.9</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>GLX</td>
<td>1.0</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>PRO</td>
<td>10.7</td>
<td>9.3</td>
<td>12.1</td>
<td>1.7</td>
<td>2.6</td>
<td>1.0</td>
<td>5.1</td>
<td>2.8</td>
<td>1.9</td>
<td>3.3</td>
<td>4.1</td>
<td>2.4</td>
</tr>
<tr>
<td>GLY</td>
<td>21.6</td>
<td>20.2</td>
<td>34.7</td>
<td>15.4</td>
<td>23.9</td>
<td>4.4</td>
<td>14.9</td>
<td>13.1</td>
<td>10.8</td>
<td>7.1</td>
<td>8.3</td>
<td>15.2</td>
</tr>
<tr>
<td>ALA</td>
<td>9.1</td>
<td>12.8</td>
<td>33.1</td>
<td>8.8</td>
<td>14.8</td>
<td>8.4</td>
<td>12.9</td>
<td>9.2</td>
<td>7.6</td>
<td>6.8</td>
<td>0.8</td>
<td>2.7</td>
</tr>
<tr>
<td>VAL</td>
<td>16.2</td>
<td>10.0</td>
<td>16.1</td>
<td>9.1</td>
<td>9.4</td>
<td>1.1</td>
<td>3.5</td>
<td>7.9</td>
<td>3.0</td>
<td>5.0</td>
<td>6.1</td>
<td>3.4</td>
</tr>
<tr>
<td>ILE</td>
<td>3.8</td>
<td>5.2</td>
<td>1.0</td>
<td>3.4</td>
<td>0.9</td>
<td>0.6</td>
<td>(0.3)</td>
<td>1.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LEU</td>
<td>1.0</td>
<td>1.0</td>
<td>2.9</td>
<td>3.1</td>
<td>1.4</td>
<td>1.2</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>3.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>TYR</td>
<td>0.5</td>
<td>1.5</td>
<td>0.5</td>
<td>0.7</td>
<td>1.1</td>
<td>(0.2)</td>
<td>1.1</td>
<td>1.0</td>
<td>1.0</td>
<td>0.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHE</td>
<td>0.0</td>
<td>0.9</td>
<td>1.2</td>
<td>3.1</td>
<td>1.8</td>
<td>1.8</td>
<td>1.8</td>
<td>1.8</td>
<td>1.8</td>
<td>1.8</td>
<td>1.8</td>
<td>1.8</td>
</tr>
</tbody>
</table>

Values are expressed as residues per peptide. A space indicates less than 0.2 residue. Residues in parentheses are fractional residues thought to be impurities.

Each of the 15 peptide fractions isolated (Fig. 1) was further purified by gel filtration and various ion exchange systems with pyridine-acetate buffers.
RESULTS

Separation and Purification of Tryptic Peptides—The chromatographic distribution of the tryptic digest is given in Fig. 1. The peaks obtained were numbered consecutively in order of their elution and are referred to accordingly. Recovery based on the amount of lysine applied to the column (18.15 pmol) and the amount recovered in the various fractions (14.54 pmol) represents 80.7%.

Further purification of these peptide fractions led to the isolation of the peptides listed in Table I. Other peptides isolated were not obtained in sufficient purity to warrant sequence analysis at this stage; work on these is still in progress.

Figures in Table I are the mean of at least two analyses corrected for contamination by neighboring peptides (up to 5% on a molar basis). In general, peptides suspected of having a higher level of contamination than this were subjected to further purification. An exception to this was peptides T9b and T9c. One of these (T9c) was obtained pure, and was analyzed and sequenced as such. The other (T9b) was obtained only admixed with T9c. As there was insufficient material to allow further attempts at purification we subjected the mixture of two peptides to analysis, and deduced the sequence of T9b by subtracting that of T9c.

The compositions given in Table I are subject to some reservations, because of the large sizes of the peptides, their unusual amino acid compositions, and the effects this has on allowing for even low levels of cross-contamination. The reservations apply especially to the mole ratios of amino acids present in large amounts (particularly glycine and alanine). As sequence work continues, most of the uncertainties should be resolved; they do not affect the basic conclusions of this paper.

Sequence Determinations of Purified Peptides—Peptides T1, T2, T4, T6, T7a, T7b, T9a, T9b, T9c, T12, T13a, T14a, T14b, T14c, and T15 have been either partially or completely sequenced. Fig. 2 contains the sequences of all of the peptides with the unsequenced portion of each peptide included in parentheses.

Fig. 2. Amino acid sequence of the purified tryptic peptides. Those residues not sequenced in each peptide are included in parentheses at the end of the completed sequences. Tentative assignments are given in separate parentheses, i.e. -(Gly)- means definitely not hydroxylated. In cases where a glycine residue is suspected at this position, but not definitely hydroxylated, the residue is left unmarked. (X) signifies the residue is unknown. Pro indicates a partially hydroxylated proline residue, while Pro means that it is definitely not hydroxylated. In cases where we are not positive concerning hydroxylation, the proline is left unmarked.
Peptides T13a and T14a are not included in the figure since they represent the small tryptic peptide Ala-Ala-Ala-Lys and Ala-Ala-Lys previously reported (4).

**Discussion**

Elastin and collagen are the two major proteins of connective tissue in vertebrates. Both proteins share a number of unusual features that have led to much speculation concerning their structural relationships. The present work clarifies this relationship considerably, by providing the first extensive evidence about the amino acid sequence of elastin. Several things stand out.

First, both elastin and collagen contain glycine as about one-third of their residues, and are also rich in proline. The glycines of collagen occur in a strictly repeating pattern, every 3rd residue over the main part of the chain; this is essential for forming the closely wound triple helical structure. We find in the present work that glycines in elastin are not subject to this constraint, so there is not the strict triplet repeat. Instead, we find repetitions of a tetrapeptide (Gly-Gly-Val-Pro; in peptide T15, and the NH₂ terminus of the protein³), of a pentapeptide (Pro-Gly-Val-Gly-Val; in peptide T1, and a related chymotryptic peptide⁴), and of a hexapeptide (Pro-Gly-Val-Gly-Val-Ala; in peptide T2).

Second, there is no strong constraint on the location of proline in collagen, while there appears to be one in elastin. In collagen proline occurs both preceding glycine (where it is frequently hydroxylated), and following glycine (where it is never hydroxylated). Inspection of the elastin sequences in Fig. 2 reveals that proline is usually followed by glycine (22/27), but rarely preceded by it (3/20).

Third, hydroxyproline is a real constituent of the elastin chain and not present because of contamination by collagen. Principal occurrences of hydroxyproline are indicated in Fig. 2 by underlining the proline residues involved. It will be seen that they are in the same local peptide environment as those in collagen (=Gly-Val-Pro), but not part of an extended collagen-like structure.

It is thus clear that elastin is a unique molecule, and should be regarded quite separately from collagen in terms of structure-function relationships. These will be discussed in other publications.

Peptides analyzed in the present work, and in a previous publication (4), account for about 730 residues of the chain, which should have a length of approximately 850 residues based on a molecular weight of 68,000 to 72,000 (3, 17). It seems very likely that peptide T15 represents the NH₂ terminus of the protein, as its NH₂-terminal sequence is the same as that of the whole molecule, and its amino acid composition closely resembles that of the part of the chain that has been sequenced directly.³ It is possible that some redundancy exists in the peptides analyzed, as some have more than 1 lysine residue. Where we have identified probable redundant peptides, we have allowed for it in arriving at the figure of 730 residues (e.g. T9a had the same NH₂-terminal sequence as T4 and therefore could be a fragment of it; however it was not considered separately here).

One thing the present work has not done is to establish the alignment of the peptides within the chain. It will be of great interest to learn the distribution of the small tryptic peptides, since these contain lysines that must become part of the cross-links of mature elastin. Extremes of cross-link distribution could occur, depending whether the small peptides are highly clustered, or uniformly dispersed along the chain; such distribution would greatly affect the elastic properties of the final product. Work is in progress now to try and establish this important facet of elastin structure.

**Acknowledgments**—We are grateful to M. A. McNulty and R. P. Meehan for their assistance in sequence determinations. We also thank Jeanne Hessling and Allayne Christiansen who typed this manuscript many times during the course of numerous revisions.

**References**


³ Unpublished results.
1 Two papers will be submitted shortly concerning a model for elastin structure and a sterochemical approach to elastin cross-linking based on the data from this paper.
Isolation and Amino Acid Sequences of Tropoelastin Peptides
Judith Ann Foster, Eveline Bruenger, William R. Gray and Lawrence B. Sandberg


Access the most updated version of this article at http://www.jbc.org/content/248/8/2876

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/248/8/2876.full.html#ref-list-1