A Comparison of 30-kDa and 10-kDa Hormone-containing Fragments of Bovine Thyroglobulin*

(Received for publication, September 28, 1987)

Joni D. Gregg, Carol Dziadik-Turner, James Rouse‡, James W. Hamilton‡, and Allen B. Rawitch§

From the Department of Biochemistry, University of Kansas Medical Center, Kansas City, Kansas 66103 and the

‡Calcium Research Laboratory, Veterans Administration Hospital, Kansas City, Missouri 64128

Studies have been carried out on reduced and silylated 19 S bovine thyroglobulin to characterize naturally occurring, iodine-rich fragments. In this report, the purification and properties of a 30-kDa, hormone-enriched polypeptide (TgE) are described and compared to that of a previously reported 10-kDa fragment (TgF). The amino acid sequence of TgF was found to overlap with that of TgE. In spite of its larger size, TgE contains only a single hormone bearing site. Both the 10- and 30-kDa fragments are derived from the NH$_2$-terminal end of the bovine thyroglobulin. These fragments contain the principal hormone-forming site at residue 5 of the thyroglobulin sequence and appear to be formed by cleavage of the parent polypeptide chain. The mechanism which generates these cleavages is not clear since the sequences surrounding the cleavage points which give rise to these peptides are quite different. These two fragments may be precursor and product in such a process. The amino acid sequence contained within TgE includes two putative sites for N-linked glycosylation. Since no glucosamine was observed and only small amounts of neutral sugar were detected, it appears that this part of the molecule is not extensively glycosylated.

Thyroglobulin (Tg), the principal protein product of thyroid secretory epithelium cells, is one of the largest glycoproteins known. Its mRNA has been isolated and translated in heterologous systems yielding a primary gene product of ~330 kDa (Vassart et al., 1975; Alvino et al., 1982). The mature Tg molecule is a dimer of two similar-sized subunits (Van der Walt et al., 1983) and serves as the matrix within which thyroid hormone synthesis occurs. Thyroid peroxidase catalyzes the iodination and coupling of specific tyrosine residues to produce thyroxine (T4) within the Tg sequence (Taurog et al., 1970; Lamas et al., 1974). Extensive transport and proteolytic processing is then required to release free hormone into the bloodstream.

Over the years, much information has been obtained about the role of Tg, in the control of metabolism and development. However, the detailed molecular events surrounding the in vivo synthesis of Tg, remain to be studied. Analysis of thyroglobulin’s structure and function has been difficult due to its large size, complex microheterogeneity and posttranslational modification, and its unique, iodine-containing residues. Recent efforts have supplied new information about the chain of events involved in Tg processing and T$_4$ formation. The naturally occurring peptide patterns of reduced and silylated Tgs, as seen on denaturing polyacrylamide gel electrophoresis, have been documented in several species (Kim et al., 1984; Chernoff and Rawitch, 1981; Rawitch et al. 1984a). In each species, peptides over a broad range of sizes (~330-10 kDa) were observed. A consistent feature of the peptide patterns is the presence of two discrete, hormone-rich small peptides (Dunn et al., 1981, 1982). The presence of the small bands has been correlated with hormone production (Dziadik-Turner et al., 1983, 1985; Lejeune et al., 1983; Marrig et al., 1984). The peptide of lowest molecular weight has been isolated and characterized from the Tgs of several species including cow (Chernoff and Rawitch, 1981), sheep and pig (Rawitch et al., 1984a), and human (Marrig et al., 1984).

The amino acid sequence surrounding T4 was first determined by isolating a tryptic fragment from the smallest peptide in bovine Tg (Rawitch et al., 1983). The identical sequence has been found in the Tg of sheep and pigs (Rawitch et al., 1984a), humans (Lejeune et al., 1983; Dziadik-Turner et al., 1985), dogs (Gregg, 1985), and rats. The sequence was determined to be NH$_2$-Asn-Ile-Phe-Glu-To-$T_4$-Gln-Val-Asp-Ala-Gln-Pro-Leu-Arg-Pro-Cys-Glu-Leu-Gln-Arg-COOH. Examinations of the thyroglobulin cDNA sequences have placed the coding for this peptide at the extreme 5' end of the human, bovine, and rat mRNAs. (Malthiey and Lissitzky, 1985; Merckens et al., 1985a; Di Lauro et al., 1985).

In the present work, we have isolated and characterized a 30-kDa peptide of bovine Tg (TgE) and compared it with the previously studied 10-kDa fragment, TgF (Chernoff and Rawitch, 1981; Rawitch et al., 1984a, 1984b).

MATERIALS AND METHODS

Preparation of Tg—Tissue was prepared following the procedure of Chernoff and Rawitch (1981). Bovine thyroid glands were obtained from a slaughterhouse and were kept frozen until use. The thyroid tissue was sliced (1-2 mm) and extracted in 0.15 M saline, pH 7.0, for 12 h. The extract was subjected to ammonium sulfate fractionation, and the protein insoluble at 45% saturation but soluble at 37% saturation was designated crude Tg. After extensive dialysis against 0.1% ammonium bicarbonate, the preparation was lyophilized and stored at ~20 °C.

Agarose Chromatography—Approximately 1.0 g of crude thyroglobulin was applied to a 5.0 × 100-cm column of Bio-Gel A-5m (Bio-Rad). The protein was eluted from the column using 0.1 M ammonium bicarbonate. Fractions corresponding to 19 S thyroglobulin were lyophilized and stored at ~20 °C.

Reduction and Alkylation—Reduction and alkylation was performed as described earlier (Chernoff and Rawitch, 1981). Pure 19 S...
samples were incubated at 90°C for 20 min in the dark. The sample was either directly applied to a 4% SDS column or was extensively dialyzed against 0.1% ammonium bicarbonate and lyophilized.

**Urea CL-4B Chromatography**—Reduced and alkylated 19 S thyroglobulin (200 mg) was applied to a 5 x 100-cm column of Sepharose CL-4B (Bio-Rad) equilibrated with 0.1 M sodium phosphate, 6 M urea, pH 7.0. The sample was eluted with the same buffer, and 12-ml fractions were collected. The column profile was monitored at 280 nm and was divided into six pools designated TgA, TgB, TgC, TgE, and TgF. These pools were extensively dialyzed against 0.1% ammonium bicarbonate. Approximately 40 mg of either crude TgE or TgF was dissolved in the same buffer, applied to the column, and eluted at 4°C. Fractions (2-ml) were collected, monitored for absorbance at 280 nm, pooled, and lyophilized. Samples were stored at -20°C.

**Tryptic Digestion**—Protein samples (5 mg) were dissolved in 1 ml of 1.0% ammonium bicarbonate. Tosylphenylalanine chloromethyl ketone-treated trypsin (Sigma and Worthington) corresponding to 5% weight of the protein sample was dissolved in 0.001 M HCl. One-half of the trypsin was added, and the digest was held at 37°C for 30 min. After 2.5 h, the second half of the enzyme was added. The digestion was stopped at 5 h by freezing and lyophilizing the sample.

**Peptide Mapping by Reverse-phase HPLC**—Peptide maps using reverse-phase HPLC were prepared from trypptic digests of reduced and alkylated 19 S thyroglobulin, TgE, and TgF. A Beckman HPLC chromatograph Model 842 with a Beckman 165 detector and Vydac 25 cm, 300-A pore size) was used. A linear gradient from 0 to 35% acetonitrile in 70 min with a flow rate of 1 ml/min was used to separate the tryptic fragments. The starting buffer consisted of 95 parts of 0.1% ammonium bicarbonate and 5 parts of acetonitrile (v:v). HPLC profiles were monitored at 230, 325, and 350 nm to detect peptide bonds, iodinated peptides, and thyroxine-containing peptides, respectively.

**Amino Acid Analysis**—Protein samples were hydrolyzed in constant boiling HCl (5.7 N) for 22 h at 110°C under vacuum. The amino acid analyses were carried out on a Beckman Model 121 MB automated amino acid analyzer.

**Peptide Sequencing**—Automated Edman degradation was performed using a Beckman 880C protein-peptide sequenator with the 0.1% ammonium bicarbonate buffer. Approximately 200 ng of the protein sample was dissolved in 0.001 M HCl. The tryptic digest was further purified on a Sephadex G-100 column run on an S-200 column. The peaks were pooled as shown in the figure and lyophilized. Pool 1 contained higher molecular weight material and was not further processed. Pool 2 contained TgE. Fragment TgF, when present, eluted in fractions following TgE. Before use in peptide mapping, TgE and TgF were verified as homogenous using 10% polyacrylamide gel electrophoresis in the presence of 0.1% SDS and 6 M urea (Fig. 3). The S-200 column chromatography successfully separated fragments TgE and TgF from each other. Special care was taken to be sure all detectable TgF was absent in samples designated "pure TgF," and vice versa.

**Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis with 6 M Urea**—Aurea—sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis was performed in a vertical gel apparatus (Protein A or Mini-Protein, Bio-Rad). The separating gel was 10% acrylamide (acrylamide:bisacrylamide, 30:1), 0.23 M sodium phosphate, pH 7.2, 6 M urea, and 0.1% SDS. The stacking gel was 4% acrylamide (same stock solution as above), 0.16 M sodium phosphate, pH 7.2, 6 M urea, and 0.1% SDS. Protein samples were dissolved in a buffer consisting of 10 mm sodium phosphate, pH 7.2, 7 M urea, and 1% SDS. The samples were incubated at 90°C for 10 min before loading on the gel. The gel was run at constant voltage until the bromophenol blue tracking dye had migrated to approximately 1 cm from the bottom of the gel. The electrode vessel buffer consisted of 0.1 M sodium phosphate, 0.25% SDS, pH 7.2. Gels were fixed in a solution of 50% methanol and 10% acetic acid. The slabs were stained overnight in a solution of 0.1% Coomassie Brilliant Blue (Bio-Rad), 0.1% cupric sulfite, 50% methanol, and 10% acetic acid. The fixing solution was also used to destain the slabs.

**RESULTS**

Crude Tg was obtained from bovine thyroids by saline extraction and ammonium sulfate precipitation. Chromatography on a 4% agarose column yielded pure 19 S Tg. After reduction and alkylation of the 19 S Tg, the resulting peptides were fractionated on a CL-4B column equilibrated with 6 M urea, 0.1 M sodium phosphate, pH 7.0 (Fig. 1). The pools were designated TgA, TgB, TgC, TgD, TgE, and TgF in order of decreasing molecular weight. Each pool was dialyzed against 0.1% ammonium bicarbonate, lyophilized, and stored frozen. Typical CL-4B profiles showed peaks of approximately equal size for TgE and TgF. Yields in dry weight were also similar for TgE and TgF, at about 3-4 mg of each for every 200-1 mg sample of crude Tg processed. Since the mass of TgF (~30 kDa) is three times that of TgF, on a molar basis more TgF was recovered than TgE.

Samples of TgE and TgF were passed over a Sephacryl S-200 column for further purification. A profile of the TgE pool from the CL-4B column run on an S-200 column is shown in Fig. 2. The peaks were pooled as shown in the figure and lyophilized. Pool 1 contained higher molecular weight material and was not further processed. Pool 2 contained TgE. Fragment TgF, when present, eluted in fractions following TgE. Before use in peptide mapping, TgE and TgF were verified as homogenous using 10% polyacrylamide gel electrophoresis in the presence of 0.1% SDS and 6 M urea (Fig. 3). The S-200 column chromatography successfully separated fragments TgE and TgF from each other. Special care was taken to be sure all detectable TgF was absent in samples designated "pure TgF," and vice versa.

The amino acid sequence of nascent bovine thyroglobulin has been determined by sequencing its cDNA (Mercken et al., 1985a). It can be concluded from several pieces of data that the NH2-terminal region (residues 1-80) predicted from the cDNA corresponds to fragment TgF. Previous characterization of TgF showed it to be a 10-kDa fragment enriched in...
5192

30-kDa Hormone-containing Fragment in Bovine Thyroglobulin

Fig. 2. S-200 gel chromatography. Partially purified bovine TgE, obtained from a CL-4B column, was dissolved in 0.1% NH₄HCO₃ and applied to a Sephacryl S-200 column (150 × 1.5 cm) equilibrated with the same buffer. Fractions (2 ml) were collected at 4 °C and pooled as shown.

Fig. 3. SDS gel electrophoresis of purified TgE and TgF fractions. Coomassie Blue-stained minislab gel of purified TgE and TgF following S-200 chromatography. Lane 1 contained a mixture of molecular weight standards ranging from 14,300 to 200,000. Samples applied were: lane 2, TgF; lane 3, TgE. Sample loads were approximately 20 mg.

both iodine and thyroid hormone (Rawitch et al., 1983). TgF contains a single T₄-containing sequence near its NH₂ terminus. The amino acid composition of TgF corresponds very well to that of the first 80 residues predicted by the bovine cDNA for Tg (Table I). Moreover, the NH₂-terminal sequences established experimentally for both intact TgE and TgF were identical to that predicted for the NH₂-terminal 19 residues of Tg. A comparison of the amino acid composition of TgE and the composition of the NH₂-terminal 234 residues from the predicted sequence showed good agreement (Table I). The composition of TgE further suggested that all of the predicted peptides in TgF were also present in TgE and confirmed by amino acid composition. When the amino acid composition data enabled us to identify the elution positions of most of the tryptic peptides predicted from the bovine gene (see Table II). Peptides smaller than 6 residues were difficult to detect by their absorbance at 230 nm unless they contained an aromatic residue. All of the predicted peptides in TgF were identified except dipeptide T-2 and tetrapeptide T-3. The amino acid analysis of individual peaks was difficult to detect by their absorbance at 230 nm unless they contained an aromatic residue. The presence of the lysine residue does not significantly alter the retention time of this peptide.

Peptide mapping was performed to determine whether isolated proteins TgE and TgF had sequences in common. The proteins were digested with trypsin and the resulting fragments mapped on reverse-phase HPLC. When the column eluate from the digest of TgE was monitored at 230 nm, maps containing approximately 20 distinct peaks were obtained (Fig. 4a). The same experiment was performed on TgF (Fig. 4b). Careful study of the profiles showed that all of the peaks found on the map of TgF can also be located on the map of TgE. The individual peaks from the HPLC runs were collected, dried, hydrolyzed, and analyzed for amino acids. The amino acid composition of TgF was normalized to Asp equal to 8. The amino acid composition of TgE was normalized to Asp equal to 20.

The reverse-phase system was capable of resolving peptides differing by only an arginine residue. Peptide T-4b was the same sequence as T-4a but with an additional arginine residue at its NH₂ terminus. This single sequence change caused a 2-min difference in elution position.

The HPLC maps of TgE and TgF were also monitored at 325 nm to detect peptides containing diiodotyrosine or T₄ (Fig. 5). The 325-nm maps showed a prominent peak eluting at 50 min and several smaller peaks eluting from 44 to 49 min. The dominant peak in the TgE profile, retention time approximately 51 min, contained the 19-residue, hormonogenic fragment previously isolated from TgF. As in the 230-nm profiles (Fig. 4), all of the peaks seen in TgF were observed in TgE and confirmed by amino acid composition. When HPLC maps of TgE and TgF were monitored at 350 nm to detect T₄, the hormonogenic fragment was the only peak showing significant absorbance.

Intact peptide TgE was subjected to 30 rounds of Edman degradation and the PTH-derivatives identified by HPLC. The sequence determined corresponded exactly to that predicted by the cDNA as the amino terminus of intact, mature thyroglobulin. Peptide TgE was also subjected to carboxypep-

| TABLE I | Amino acid compositions of TgF and TgE |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| Amino acid      | TgF* predicted  | TgF* analysis   | TgF* predicted  | TgF* analysis   |
| Cys             | 7               | 6               | 17              | 16              |
| Asp             | 7               | 8               | 20              | 20              |
| Thr             | 1               | 2               | 8               | 6               |
| Ser             | 4               | 6               | 20              | 17              |
| Glu             | 15              | 15              | 32              | 40              |
| Pro             | 6               | 6               | 16              | 16              |
| Gly             | 6               | 6               | 16              | 17              |
| Ala             | 7               | 7               | 14              | 15              |
| Val             | 5               | 6               | 14              | 14              |
| Met             | 0               | 1               | 3               | 3               |
| Ile             | 1               | 2               | 5               | 6               |
| Leu             | 5               | 6               | 18              | 17              |
| Tyr             | 2               | 2               | 9               | 6               |
| Phe             | 4               | 3               | 11              | 9               |
| Lys             | 2               | 3               | 4               | 5               |
| His             | 0               | 0               | 2               | 2               |
| Arg             | 7               | 6               | 23              | 18              |
| Trp             | 1               | *               | 2               | *               |
| Molecular weight| 8,955           | 8,835           | 25,740          | 25,080          |

* Amino acid composition based on cDNA sequence in Fig. 6.
+ The amino acid composition of TgF was normalized to Asp equal to 8.
+ The amino acid composition of TgE was normalized to Asp equal to 20.
dissolved in a buffer of 0.1% NH₄HCO₃ and acetonitrile (95:5) and injected on a Vydac C-18 HPLC column (250 × 4.5 mm). The column was equilibrated in the same buffer. A linear gradient of 0–25% acetonitrile over 50 min was used to elute the tryptic fragments. The column effluent was monitored at 230 nm to detect peptide bonds. The peaks are numbered in order of elution. The peak at 2.5 min was present in every injection and contained no peptide material. Panel a, TgE; panel b, TgF.

The amino acid composition and molecular weight constructed for TgE from this predicted cleavage point agree very well with the experimentally determined values (Table I). The amino acid composition of TgE was found to be very similar to that found in TgF. No amino sugar was detected on standard amino acid analysis, but some neutral sugar was found (~5%). Like TgF, TgE was enriched in both iodine and thyronine. The HPLC tryptic peptide maps of peptides TgE and TgF were compared (Figs. 4 and 5). It was apparent that all of the peaks on the map of TgF were also contained on the maps of TgE. Individual peaks were collected and amino acid analysis performed. Peaks from TgE and TgF with the same elution time were compared to each other and to the composition of tryptic fragments predicted from cDNA sequence data (Mercken et al., 1985a). Each of the corresponding peaks from TgE and TgF had similar or identical compositions. The C-18 HPLC column was able to separate tryptic peptides T-4a and T-4b, sequences differing by only an Arg residue at the NH₂ terminus. Tryptic peptide T-17 eluted as HPLC peak 18 with a retention of approximately 49 min. Peptide T-17 contains 12 amino acids, four of which are phenylalanine residues. The hydrophobic nature of this peptide probably contributes to its long retention time. This is consistent with the observation that iodinated tyrosines and thyronines elute from reverse-phase columns with acetonitrile in order of increasing iodination, i.e., increasing hydrophobicity.

When HPLC peaks 16–19 were collected for amino acid analysis, their recovery was consistently low in relation to the magnitude of the detected peaks. Although HPLC peak 18 contained no iodinated residues, apparent absorbance at 325 nm was observed. HPLC peaks 17 and 18 were identified as tryptic fragments derived from TgE, but they were also detected, in reduced amounts, in TgF samples. We conclude that there were traces of TgE in our TgF samples. This led to those peptides with strong absorbance signals from the contaminating TgE being detected in the TgF peptide maps.

Several pieces of data were used to place the peptides TgE and TgF at the NH₂ terminus of thyroglobulin. The tryptic peptides mapped on HPLC gave excellent agreement with the predicted composition. Using the COOH-terminal determined in this study the overall amino acid compositions of TgF and TgE are very similar to those predicted from the gene. The COOH-terminal sequences determined by carboxypeptidase digestion were found to be Glu-Leu-Leu for TgE and Gln-Leu-Leu for TgF.

**DISCUSSION**

The results presented here characterize bovine TgE, a hormone-containing 30-kDa polypeptide from thyroglobulin. TgE was purified from bovine Tg by sequential gel filtration steps after reduction and alkylation. Its mass was determined on standard amino acid analysis, but some neutral sugar was found (~5%). Like TgF, TgE was enriched in both iodine and thyronine. The HPLC tryptic peptide maps of peptides TgE and TgF were compared (Figs. 4 and 5). It was apparent that all of the peaks on the map of TgF were also contained on the maps of TgE. Individual peaks were collected and amino acid analysis performed. Peaks from TgE and TgF with the same elution time were compared to each other and to the composition of tryptic fragments predicted from cDNA sequence data (Mercken et al., 1985a). Each of the corresponding peaks from TgE and TgF had similar or identical compositions. The C-18 HPLC column was able to separate tryptic peptides T-4a and T-4b, sequences differing by only an Arg residue at the NH₂ terminus. Tryptic peptide T-17 eluted as HPLC peak 18 with a retention of approximately 49 min. Peptide T-17 contains 12 amino acids, four of which are phenylalanine residues. The hydrophobic nature of this peptide probably contributes to its long retention time. This is consistent with the observation that iodinated tyrosines and thyronines elute from reverse-phase columns with acetonitrile in order of increasing iodination, i.e., increasing hydrophobicity.

When HPLC peaks 16–19 were collected for amino acid analysis, their recovery was consistently low in relation to the magnitude of the detected peaks. Although HPLC peak 18 contained no iodinated residues, apparent absorbance at 325 nm was observed. HPLC peaks 17 and 18 were identified as tryptic fragments derived from TgE, but they were also detected, in reduced amounts, in TgF samples. We conclude that there were traces of TgE in our TgF samples. This led to those peptides with strong absorbance signals from the contaminating TgE being detected in the TgF peptide maps.

Several pieces of data were used to place the peptides TgE and TgF at the NH₂ terminus of thyroglobulin. The tryptic peptides mapped on HPLC gave excellent agreement with the predicted composition. Using the COOH-terminal determined in this study the overall amino acid compositions of TgF and TgE are very similar to those predicted from the gene. The COOH-terminal sequences determined by carboxypeptidase digestion were found to be Glu-Leu-Leu for TgE and Gln-Leu-Leu for TgF.

C. Dziadik-Turner and A. Rawitch, unpublished data.
30-kDa Hormone-containing Fragment in Bovine Thyroglobulin

**TABLE II**

Summary of HPLC peaks, predicted peptides and amino acid analysis of individual peaks

| HPLC peak | c-DNA predicted trypsin peptide* | Amino acid Residue positions | Comments |
|-----------|----------------------------------|------------------------------|----------|
| 1         | T-6                              | 60-65                        | Smallest peptide identified.¹ |
| 2         | T-5                              | 47-59                        | Six homologous repeats centered around this sequence have been predicted in the amino half of bovine Tg (Merken et al., 1985b). |
| 3         | T-4b                             | 27-46                        | A consistently appearing, sharp peak that precedes a tall peak 4. Very difficult to collect free of contaminating peak 4. |
| 4         | T-4a                             | 26-46                        | Same as HPLC peak 5 but without an Arg residue. Differs from HPLC peak 4 by one additional amino acid (Arg). |
| 5         |                                   |                               | Could not be clearly identified as a cDNA predicted peptide. Not present in profiles of TgF. |
| 6         | T-15                             | 160-176                      | Contains 17 amino acids; no Tyr. |
| 7         | T-9                              | 118-133                      | Contains Cys-Trp-Cys-Val-Asp-Ala sequence also found in T-5. |
| 8         | T-19                             | 209-224                      | Composition almost identical to peak 10. |
| 9         | T-7                              | 66-81                        | Sequence contains 2 Tyr residues. When isolated as the C-terminal peptide in TgF, residue 81 (Lys) is absent. The presence of Lys does not alter the retention time. |
| 10        |                                   |                               | Composition very similar to T-7, peak 11. |
| 11        | T-20                             | 225-234, 235                 | The amino acids contained in peptide T-20 coelute with other amino acids. T-20 is the COOH-terminal peptide of TgE, as determined by carboxypeptidase Y digestion. |
| 12        | T-8a, b, c                       | 82-115, 116, 117             | Wide peak probably due to the close elution of three different peptides, each differing by 1 Arg residue (i.e. alternate trypsin cleavage sites). |
| 13        |                                   |                               | Not a clear match for any predicted tryptic peptide. |
| 14        | T-5                              | 82-116                       | Same composition as HPLC peak 14, but with low tyrosine. |
| 15        | T-8                              | 82-116                       | Very difficult to collect free of contaminating peptide peak. |
| 16        | T-17                             | 195-206                      | Long retention time probably due to 4 Phe residues. |
| 17        | T-1                              | 1-19                         | Hormone-containing peak; sequence conserved in Tg from several species. |

*Tryptic peptides are numbered from the NH₃ terminus of thyroglobulin. See Fig. 6.

¹In the absence of aromatic residues, peptides of less than six amino acids could not be detected using this method.

Leu-Gln for TgF. These sequences are located at residues 77-80 for TgF and 232-234 for TgE. It should be noted that the COOH-terminal sequences of TgF and TgE, as well as the additional neighboring amino acids near the cleavage sites which result in TgE and TgF, are quite different from each other (see Fig. 6). The sequences do not correspond to tryptic cleavages or to the cleavage pattern(s) of any of the common processing enzymes. Our data indicate that the TgF peptide is contained in the TgE peptide but cannot answer the question of whether TgF is formed through proteolytic processing of TgE or in a more direct process from the parent Tg chain. The complete sequencing of the bovine cDNA has revealed that the hormone-containing sequence found at the NH₃ terminus is not repeated (Mercken et al., 1985a). Thus, TgE and TgF are derived uniquely from the amino terminus of the thyroglobulin molecule, overlap within that sequence, and contain the primary T₄ synthesis site.

Whether TgE and TgF are formed from alternate cleavages of the initial Tg gene product or whether TgE is the precursor of TgF is not completely clear. Peptides TgE and TgF together account for a large fraction of the T₄ produced in bovine Tg. The very different COOH-terminal sequences of TgE and TgF imply that the same enzyme or chemical mechanism is probably not responsible for both cleavages. It has been suggested that clipping of the Tg protein may take place during iodination (Deme et al., 1975; Dunn et al., 1983). In addition, a CNBr fragment of Tg found only when T₄ is produced decreases in molecular weight after incubation with thyroid peroxidase (Dziadik-Turner et al., 1983). In neither case has a specific molecular mechanism for breaking peptides bonds been proposed. If TgE and TgF are precursor and product then one should be able to follow the incorporation of radiolabeled iodine during a pulse-chase experiment. Although this experiment has not been done using bovine Tg, it has been carried out with rabbit and human thyroglobulins (Dunn et al., 1981; Marriq et al., 1984). In both of these species, two low molecular weight, iodine-rich Tg fragments have been characterized. In both cases it was concluded that the iodine found initially in the larger of the two fragments was detected later in the smaller fragment.

Peptide TgE contains 7 tyrosine residues, in addition to the residue in tryptic fragment T-4 (residue 28). These occur at positions 89, 97, 130, 192, 215, and 217. HPLC maps made at 325 nm of tryptic digests of TgE show several peaks eluting from about 45 to 50 min, just prior to the hormone-containing peptide peak. When HPLC peak 17 was analyzed, its composition closely matched that predicted for T-8, except for the tyrosine value which was low. Peptide T-8 has 3 tyrosine residues; the composition of HPLC peak 17 showed only 2 tyrosine residues. The amino acid data, 325-nm absorbance, and long column retention time are all consistent with peptide T-8 containing at least one site for diiodotyrosine formation. As tyrosine residues are iodinated the peptide's retention time would be expected to increase and its tyrosine content on amino acid analysis decrease. It is unlikely that monoiodotyrosine alone is formed in this peptide since the extinction coefficient of monoiodotyrosine at 325 nm is too low to account for the observed absorbance (Dziadik-Turner et al., 1983).
30-kDa Hormone-containing Fragment in Bovine Thyroglobulin

and Rawitch, 1985). HPLC peak 16 also had a composition similar to T-8, but it was not as good a match as peak 17.

These observations illustrate the necessity of obtaining protein data along with DNA data in locating sites of iodination and/or modification in thyroglobulin. Although the cDNA data can yield a predicted amino acid sequence, they give no information relating to the amount or location of iodinated residues. Thus, both kinds of information are required to establish a clear picture of thyroglobulin's mature structure and the nature of its posttranslational changes.

Three hormone-containing sequences in addition to the 19-residue fragment discussed here have been reported in porcine thyroglobulin (Marriq et al., 1983). All three are located near the carboxy terminus of the intact Tg structure. However, hormone formation in the carboxy sites appears to contribute relatively little to the total T₃ production in normal animals in the absence of thyrotropin stimulation. This is consistent with the fact that only the amino-terminal T₃-containing peptide can be detected in goiter Tg at low iodine levels (Dziadik-Turner and Rawitch 1984). The other sites are most probably iodinated and coupled to some degree at higher iodine levels, after the primary site is filled. The location of the principal and secondary hormone-producing sites near the ends of the protein has implications for both hormone formation and proteolytic processing. The tyrosine used to form hormone must be accessible to thyroid peroxidase for both iodination and coupling. Furthermore, thyroglobulin containing T₃ can react with antibody specific for a T₃-containing epitope (Byfield et al., 1982, 1984). Clearly, hormone contained within the Tg sequence is exposed on the surface of the protein. Thus, a location near the chain terminus may be highly preferred as the coupling acceptor site. Thyroid hormone must be proteolytically cleaved from the Tg structure to become available as free, circulating triiodothyronine or T₄. Recent studies have shown that in hog thyroid lysosome extracts there are thiol proteases that can release T₃ from Tg (Yoshinari and Taurog, 1985). The rate of T₃ release was found to increase as the size of the thyroxine-containing polypeptide substrate decreased (Nakagawa and Ohtaki, 1985). In light of this, the formation of T₃ near a sequence terminus could be seen as facilitating its release. Thus, only one peptide bond near either terminus of the molecule need be broken to produce a small fragment.

The very similar amino acid compositions of fragments TgE and TgF invite examinations of their amino acid sequence for internal homology. A study of the cDNA sequence has outlined a 60-residue sequence that is repeated seven times within Tg (Mercken et al., 1985b). Data presented here indicates that three of these homologous duplications occur within the sequences of TgE.

Both TgE and TgF are naturally-occurring fragments of Tg, derived from the 330-kDa gene product via posttranslational processing. The in vivo production of these fragments may be a required step in the chain of events releasing free T₃ from the protein backbone. In this way hormone release can occur without requiring the complete degradation of thyroglobulin. The hormone contained in bovine TgE and TgF represents the major portion of the hormone synthesized in bovine thyroglobulin. While the specific cleavage mechanism which results in these polypeptide fragments in not yet

Fig. 5. HPLC maps of TgE and TgF iodotyrosine distribution. Samples and conditions were the same as in Fig. 4, except that the effluent was monitored at 325 nm to detect peptides containing iodinated tyrosine derivates. Panel a, TgE; panel b, TgF.
ACKNOWLEDGMENTS - We wish to thank Danny Snow for drafting the illustrations used in this manuscript. We also thank Kenneth Michael Zabel for his expert technical help with the thyroglobulin preparations.

REFERENCES

1. Abino, C. G., Tassi, V., Peterson, B. M., and Di Lauro, R. (1982) FEMS Lett. 137, 107-113
2. Byfield, P. G. H., Bond, A., Copping, S., and Himsworth, R. L. (1982) Biochem. J. 207, 471-478
3. Byfield, P. G. H., Clingan, D., and Himsworth, R. L. (1984) Biochem. J. 219, 405-410
4. Chernoff, S. B., and Rawitch, A. B. (1981) J. Biol. Chem. 256, 9425-9430
5. Di Lauro, R., Obici, S., Condiffe, S., Ursini, V. M., Musti, A., Moscatelli, C., and Avvedimento, V. E. (1985) Eur. J. Biochem. 148, 7-11
6. Demo, D., Fimiani, E., Pommier, J., and Nunez, J. (1975) Eur. J. Biochem. 51, 329-336
7. Downing, M. R., and Mann, K. G. (1976) Anal. Biochem. 74, 298-319
8. Dunn, J. T., Dunn, A. D., Heppner, D. G., Jr., and Kim, P. S. (1981) J. Biol. Chem. 256, 942-947
9. Dunn, J. T., Kim, P. S., and Dunn, A. D. (1982) J. Biol. Chem. 257, 85-94
10. Dunn, J. T., Kim, P. S., Dunn, A. D., Heppner, D. G., Jr., and Moore, R. C. (1983) J. Biol. Chem. 258, 9093-9098
11. Dziadik-Turner, C., and Rawitch, A. B. (1983) in The Enzymology of Posttranslational Modification of Proteins, Vol. 2 (Freedman, R. B., and Hawkins, H. C., eds) pp. 95-123, Academic Press, Orlando, FL
12. Dziadik-Turner, C., and Rawitch, A. B. (1984) Fed. Proc. 43, 1649
13. Dziadik-Turner, C., Chernoff, S., Taurog, A., and Rawitch, A. B. (1985) Arch. Biochem. Biophys. 222, 245-258
14. Dziadik-Turner, C., Rawitch, A. B., and Taurog, A. (1985) Fed. Proc. 44, 695
15. Gregg, J. D. (1985) The Study of Thyroid Hormone Formation in Bovine and Canine Thyroglobulins. Masters Thesis, Univ. of Kansas Medical Center
16. Kim, P. S., Dunn, J. T., and Kaiser, D. K. (1984) Endocrinology 114, 369-374
17. Lamas, L., Taurog, A., Salvatore, G., and Edelhoch, H. (1974) J. Biol. Chem. 249, 2372-2377
18. Lejeune, P.-J., Marriq, C., Rolland, M., and Lissitzky, S. (1983) FEBS Lett. 156, 77-82
19. Malthiery, Y., and Lissitzky, S. (1985) Prog. Endocr. Res. Ther. 2, 87-93
20. Marriq, C., Rolland, M., and Lissitzky, S. (1983) Biochem. Biophys. Res. Commun. 112, 206-213
21. Margit, C., Lejeune, P.-J., Malthiery, Y., Rolland, M., and Lissitzky, S. (1984) FEBS Lett. 175, 140-146
22. Mercken, L., Simons, M. J., Swillens, S., Massera, M., and Vassart, G. (1985a) Nature 316, 647-651
23. Mercken, L., Simons, M.-J., Swillens, S., Massera, M., and Vassart, G. (1985b) Eur. J. Biochem. 147, 69-64
24. Nakagawa, H., and Ohtaki, S. (1985) Endocrinology 116, 1433-1439
25. Rawitch, A. B., Chernoff, S. B., Litwer, M. R., Rouse, J. B., and Hamilton, J. W. (1983) J. Biol. Chem. 258, 2079-2083
26. Rawitch, A. B., Litwer, M. R., Gregg, J., Dziadik-Turner, C., Rouse, J., and Hamilton, J. W. (1984a) Biochem. Biophys. Res. Commun. 118, 423-429
27. Rawitch, A. B., Mercken, L., Hamilton, J. W., and Vassart, G. (1984b) Biochem. Biophys. Res. Commun. 119, 335-342
28. Taurog, A., Lothrop, M., and Estabrook, R. W. (1970) Arch. Biochem. Biophys. 139, 221-229
29. Van der Walt, B., Kotze, B., Edelhoch, H., and Van Jaarsveld, P. P. (1983) Biochim. Biophys. Acta 744, 90-98
30. Vassart, G., Refetoff, S., Brocas, H., Dinais, C., and Dumont, J. E. (1975) Proc. Natl. Acad. Sci. U. S. A. 72, 3833-3843
31. Yoshinari, M., and Taurog, A. (1985) Endocrinology 117, 1621-1628

Fig. 6. The amino acid sequence of the amino-terminal region of bovine thyroglobulin. The primary structure of the protein as predicted from cDNA studies of Mercken et al. (1985a). The amino terminus of the mature protein is preceded by a 19-residue leader sequence. Tryptic peptides T-1-T-20 are indicated. The heavy arrows indicate cleavage points which result in the hormone-rich 10- and 30-kDa fragments TgE and TgF. The principal hormone-forming tyrosine at residue 5 is boxed, and the putative N-glycosylation sites in this region are underlined.