The Primary Structure of Bovine Thyrotropin

II. THE AMINO ACID SEQUENCES OF THE REDUCED, S-CARBOXYMETHYL $\alpha$ AND $\beta$ CHAINS*

Ta-Hsiu Liao and John G. Pierce

From the Department of Biological Chemistry, UCLA School of Medicine, Los Angeles, California 90024

 SUMMARY

The linear amino acid sequence of the two chains of bovine thyrotropin (TSH-$\alpha$ and TSH-$\beta$) have been determined except for assignment of some amides. The order of the tryptic peptides of TSH-$\beta$ (described in the preceding publication) was established by studies on the peptides obtained from tryptic hydrolysates of maleylated, reduced, S-carboxymethyl TSH-$\beta$, and from the oxidized chain. The preparations of TSH-$\beta$ exhibited heterogeneity at the COOH terminus; molecules of 113 residues terminate with the tripeptide sequence -Ser-Tyr-Met-COOH while others, of 112 residues terminate with -Ser-Tyr-COOH. Evidence for the heterogeneity of the NH$_2$ terminus of TSH-$\alpha$ was again found in the peptides from its maleylated derivative.

The molecular weight, calculated from the amino acid sequences and carbohydrate compositions, of TSH-$\alpha$ preparations is approximately 13,600, that of TSH-$\beta$ preparations, 14,700. TSH-$\alpha$ contains two oligosaccharide moieties, TSH-$\beta$ one. Both chains are extremely rich in intrachain disulfide bonds, five in TSH-$\alpha$ and six in TSH-$\beta$; their positions remain to be determined. There are no interchain disulfide bonds. The sequences of TSH-$\alpha$ and $\beta$ do not resemble each other superficially, but several very similar or identical sequences are found in TSH-$\beta$ and in the hormone-specific chain, CII, of luteinizing hormone, despite marked differences in their amino acid compositions. The sequence of the latter has been announced recently by Liu et al. (Res. Commun. Chem. Path. Pharmacol., 1, 463 (1970)). The similarities may be a reflection of similar sites on the chains of TSH-$\alpha$ and TSH-$\beta$.

 The complete amino acid sequences, excluding some assignment of amides, of the reduced S-carboxymethyl derivatives of both TSH-$\alpha$ and $\beta$ are given.

 EXPERIMENTAL PROCEDURE

Bovine TSH was prepared with the final purification step by countercurrent distribution, and the two chains were separated by gel filtration after their dissociation in 1 M propionic acid as described previously (2). Most of the analytical methods and materials used are described in the preceding paper (3). Mepactoacetic acid and phenol were added to samples before hydrolysis for amino acid analysis (3, 5) if S-carboxymethylcysteine, methionine, or tyrosine were suspected. Most of the Edman degradations described herein were carried out by the method described by Gray in which the successive NH$_2$-terminal amino acids emerge from Sephadex columns; TSH-$\beta$, the second of the two subunits; dansyl (DNS), 1-dimethylaminonaphthalene-5-sulfonyl; Cys(Cm), S-carboxymethylcysteine; MetO$_2$, methionine sulfoxide; L.H, luteinizing hormone.

* This investigation was supported by United States Public Health Service Grant CA-2290 from the National Cancer Institute.
of 0.1 m NaHCO₃ and 150 mg of maleic anhydride (Matheson, Coleman and Bell, East Rutherford, New Jersey) were added slowly. The pH was adjusted to 9.0 with 1 N NaOH and, after standing 15 min at room temperature, the reaction mixture was passed through a column (2 x 90 cm) of Sephadex G-25 (coarse) in the presence of 1% NH₄HCO₃. The protein was then recovered, salt-free, by freeze-drying for 48 hours.

Oxidation with Performic Acid—Performic acid was prepared by adding 1 volume of 30% H₂O₂ to 9 volumes of 99% formic acid and allowing the mixture to stand at room temperature for 1 hour. The solution was chilled in ice. One volume of methanol and solid phenol (1%, w/v) were then added. Immediately, 8 mg of TSH-β were dissolved in 0.5 ml of the solution and allowed to remain in the ice bath for 3 hours. The excess reagents and salts were removed by the passage through a Sephadex G-25 (coarse) column (1.2 x 40 cm) in the presence of 0.5% NH₄HCO₃ at 4°C. Although some bubbles formed in the column, complete separation of the protein from lower molecular weight materials was obtained. The protein was recovered by freeze-drying.

Hydrolysis with Trypsin, Chymotrypsin, Thermolysin, and Subtilisin—Chymotrypsin, and thermolysin, as well as the trypsin (treated with L-(1-tosylamido-2-phenyl)ethyl chloromethyl ketone), were obtained from Worthington. Subtilisin (Carlsberg) was a gift from Dr. R. J. DeLange. The derivatized chains or peptides from them were hydrolyzed in the presence of 0.126 or 0.06 M NH₄HCO₃ with no further adjustment of pH.

Assignment of Amides—Peptides were digested with aminopeptidase M (Henley and Company, New York). Approximately 30 mmoles of peptides were dissolved in 1% NH₄HCO₃ and 20 μl of aminopeptidase M (10,000 milliunits per ml of water) added. The time of hydrolysis, at 40°C, was 24 to 48 hours with the longer time used for the large peptides. The hydrolysates were then taken to dryness and analyzed on the amino acid analyzer. Since both glutamine and asparagine appeared at the serine position, the appearance of a new or larger peak at the serine position as well as the disappearance of aspartic acid or glutamic acid relative to that found after acid hydrolysis permitted identification of amides in several peptides.

Nomenclature—Peptides obtained from tryptic digestions of the maleylated chains were designated as T(m) and from oxidized material as T(ox). Chymotrypsin, thermolysin, and subtilisin peptides were designated as C-, Th-, and S-, respectively. Other designations are as in the preceding paper (3).

RESULTS

The partial amino acid sequence of TSH-α and the sequences of many of the tryptic peptides of TSH-β are described in the preceding paper (3). The deduction of the ordering of the peptides of TSH-β and of the complete sequences of the two chains is best followed by reference to these sequences (Figs. 1 and 2).

Fig. 1. The linear amino acid sequence of bovine TSH-α. ——— residues that were positioned by sequence studies on the peptides; ——— residues by composition only. The half-cystines were determined as the S-carboxymethyl derivative.
Fig. 2. The linear amino acid sequence of bovine TSH-II, residues that were positioned by sequence studies on the peptides, -- --, residues by composition only. The half-cystines were determined as the S-carboxymethyl derivative.

The major peptides necessary for establishing the order are shown.

Tryptic Peptides from Maleylated, Reduced, S-Carboxymethyl TSH-α—The separation on a column of Sephadex G-50 (fine) of the peptides from the tryptic hydrolysate of the maleylated derivative is shown in Fig. 3. Details of the hydrolysis and separation are given in the legend. Fractions I, II, and III were passed through the column a second time, Fractions I and III gave symmetrical elution profiles; Fraction II was partially separated into two components, Fraction II-1 and II-2 as shown in Fig. 4. Amino acid analysis showed that Fractions I, II-1, and II-2 required further purification which was achieved, with only partial success for Fraction II-2, by paper electrophoresis at pH 6.4. The peptides were eluted with 10% pyridine. The amino acid compositions of the above fractions together with those of the other peptides from the maleylated derivative are
given in Table I. Each has a composition quite different from the others; they are assigned as αT(m)-1, etc. Thus the composition of Peptide αT(m)-1 is that of Residues 1 through 39 (Fig. 1) confirming the data from cleavage with cyanogen bromide.

**Fig. 3.** Gel filtration of the tryptic hydrolysate of maleylated, reduced, S-carboxymethyl TSH-α. The derivative, 2.5 amoles, was hydrolyzed with trypsin (2.5 mg) in 5 ml of 0.06 M NH₄HCO₃ at 40° for 2 hours. The hydrolysate was placed on a column (2.5 X 200 cm) of Sephadex G-50 (fine) and was eluted at room temperature with 0.06 M NH₄HCO₃. The flow rate was 30 ml per hour and 4 ml samples were collected. ---, the absorbance at 280 mp; -----, absorbance at 220 mp; ·······, the absorbance at 530 mp of the color developed in the orcinol reaction; solid bars, the pooled fractions.

**Fig. 4.** Gel filtration pattern of Fraction II, Fig. 3. Conditions were the same as those in Fig. 3. The pooled fractions are indicated by solid bars.

### Table I

**Amino acid composition of tryptic peptides from maleylated TSH-α**

The values are given as molar ratios of the amino acids in this and subsequent tables. Hydrolysis time was 22 hours and no corrections for destruction during hydrolysis were made. The numbers in parentheses are the theoretical values. The yields of each peptide are uncorrected.

| Amino Acid       | αT(m)-1 | αT(m)-3 | αT(m)-4 | αT(m)-8 | Total | THSH-α |
|------------------|---------|---------|---------|---------|-------|--------|
| Lysine           | 6.05(4) | 4.08(6) | 4.62(5) | 1.00(1) | 10    | 10.3   |
| Methionine       |         |         |         |         | 1.9   |
| Arginine         | 0.94(1) | 0.91(2) | 0.94(2) | 0.95(1) | 3     | 3.2    |
| Asparagine       | 3.18(3) | 3.15(3) | 2.17(2) | 1.34(1) | 0     | 4.5    |
| Threonine        | 1.02(1) | 1.05(2) | 6.64(2) | 6.64(2) | 9     | 9.5    |
| Histidine        | 1.98(3) | 1.96(3) | 1.96(2) | 1.96(2) | 4     | 6.8    |
| Glutamic acid    | 4.98(3) | 4.89(4) | 1.38(1) | 1.38(3) | 8     | 8.1    |
| Proline          | 3.61(4) | 3.43(3) | 2.11(2) | 1.10(1) | 7     | 7.0    |
| Glycine          | 3.16(3) | 2.96(3) | 1.38(1) | 4     | 4.1    |
| Alanine          | 1.14(1) | 1.16(1) | 1.16(1) | 4.12(1) | 7     | 7.1    |
| S-carboxymethyl cysteine | 4.16(5) | 4.21(3) | 2.05(3) | 3.07(3) | 10    | 9.9    |
| Valine           |         | 3.64(4) | 0.96(1) | 5.1    |
| Methionine       | 1.96(2) | 1.78(2) | 1.90(2) | 3.9    |
| Isoleucine       | 1.06(1) | 1.00(1) | 0.97(1) | 2.0    |
| Leucine          | 1.09(1) | 1.11(1) | 1.07(1) | 2.1    |
| Tyrosine         | 2.02(2) | 2.01(2) | 6.06(1) | 3.3    |
| Phenylalanine    | 3.57(4) | 2.39(2) | 1.02(1) | 4.7    |
| Total residues   | 39      | 33      | 7       | 33      | 17    | 96     |
| Yield %          | 61.8    | 35.6    | 94.0    | 76.4    | 70.0  |

"The values for proline and phenylalanine were about 0.5 residue lower than the theoretical ones. This is probably due to the presence of a peptide which has the phenylalanyl-prolynl dipeptide missing from the NH₂ terminus of αT(m)-1.

"The higher values of aspartic acid, glutamic acid, proline, glycine, and phenylalanine are probably due to contamination of this peptide with αT(m)-1.

"The total is that of the primary peptides only.

"See Reference 3, Table I."
fragments agree perfectly with the order of the tryptic peptides established in the preceding paper (3). It should be noted αT-6, which is placed within αT(m)-3, gave a poor analysis for methionine ((3), Table I). The finding of 2 residues of methionine in αT(m)-3 confirms the ordering of αCNBr-3 and αCNBr-4 (3). The summation of the residues in the four primary tryptic peptides from the maleylated derivative is in agreement with the composition of the intact chain and the yields are close to theory. As in the case of the cyanogen bromide fragments, the compositions of the large maleylated peptides are sufficiently distinctive to permit unequivocal assignment of the smaller tryptic peptides within them.

Sequence of TSH-α—In the preceding paper (3), the complete sequences of Peptides αT-2, αT-4, αT-5, αT-7, αT-8, αT-9, and αT-11 are given and the sequence of Peptide αT-6 was deduced from the partial sequences of cyanogen bromide fragments.

**TABLE II**

Amino acid sequence of Peptide αT-1

In this and subsequent tables the symbol — indicates that the sequence was determined by the dansyl method after each step of the Edman degradation. A solid line under the sequence indicates the residues positioned by sequence studies; a dashed line signifies residues known to be present by composition. CMCys is S-carboxymethylcysteine.

| Peptide | Sequence | Cys | Cys | Cys | Ala | Pro | Thr | Lys | Lys | Tyr
|---------|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| αT-1    | Tyr, 1.00(1) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| αT-2    | Tyr, 1.00(1) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| αT-3    | Tyr, 1.00(1) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| αT-4    | Tyr, 1.00(1) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| αT-5    | Tyr, 1.00(1) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| αT-6    | Tyr, 1.00(1) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |

**TABLE III**

Amino acid sequence of Peptide αT-3

| Peptide | Sequence | Cys | Cys | Cys | Ala | Pro | Thr | Lys | Lys | Tyr
|---------|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| αT-1    | Tyr, 1.00(1) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| αT-2    | Tyr, 1.00(1) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| αT-3    | Tyr, 1.00(1) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| αT-4    | Tyr, 1.00(1) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| αT-5    | Tyr, 1.00(1) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |

αCNBr-3 and αCNBr-4. The remainder of the TSH-α sequence was determined by enzymatic digestions of the larger tryptic peptides from the reduced, S-carboxymethyl derivative, both before and after its maleylation.

Peptide αT-1—It was found that the peak which emerged from columns of P-10 directly after the glycopeptide fraction of a tryptic hydrolysate of both chains contained only Peptides αT-1 and βT 12 (5, Fig. 1). The two peptides were separated during electrophoresis on paper at pH 1.9. Peptide αT-1 (0.5 μmol) was hydrolyzed with chymotrypsin (50 μg) in 0.1 ml of 0.126 M NH₄HCO₃ at pH 8.0 and 40° for 4 hours. The reaction mixture was spotted on paper and the peptides were isolated after electrophoresis at pH 4.8. They appeared as C-2, C-1, and C-3 in the order of anode to cathode, and their compositions as well as sequences, determined by the dansyl-Edman, are shown in Table II. The order of the chymotryptic peptides is established because the entire tryptic peptide has NH₂-terminal phenylalanine. Peptide C-2 must have originated from Peptide αT-1a which is missing its NH₂-terminal Phe-Pro sequence.

Peptide αT-3—The sequence of the first 7 residues from the NH₂ terminus and of 7 residues from the COOH terminus of Peptide αT-3 was described in the previous paper (3). The NH₂-terminal sequence was obtained by the Edman degradation (3, Table IV) and the COOH-terminal sequence was deduced from a study of CNBr fragments of TSH (3, Table XIII). To complete the sequence, Peptide αT-3 (0.5 μmol) was hydrolyzed with thermolysin (100 μg) in 0.5 ml of 0.06 M NH₄HCO₃ at pH 8.0 and 40° for 4 hours. The reaction mixture was spotted on paper and the peptides were isolated after electrophoresis at pH 3.6 (pyridine-acetic acid-water, 3 : 30 : 867 v/v). They migrated, in the order of anode to cathode, as Th-6, Th-5, Th-1, Th-2, Th-3, Th-4, and Th-7. Their compositions and the sequence of Peptide αT-3 are shown in Table III. These data, along with the data of the hydrolysis of αCNBr-2 with carboxypeptidase A, complete the sequence of αT-3. The 2nd residue of proline is placed by difference at the COOH terminus of Th-3 and Th-4.
Table IV
Amino acid sequence of Peptide αT(m)-4
m-Lys is ε-N-maleyllysine. CMCys is S-carboxymethylcysteine.

| Sequence | CHO |
|----------|-----|
| Val-Glu-Asn-His-Thr-Glu-CMCys-His-Glu-CMCys-Ser-Thr-CMCys-Tyr-Tyr-His-m-Lys-Ser | C-1 |
| C-2 (yield, 100%) | C-3 (yield, 50%) |
| Tyr, 0.96(1); His, 1.06(1); Lys, 1.04(1); Ser, 0.95(1) | C-4 (yield, 24%) |
| CMCys, 2.32(2); Ser, 0.95(1) | CMCys, 2.88(3); Ser, 0.51(0) |
| 3.5 hr, pH 8.0, 40°: Tyr, 0.84; CMCys, 0.39; Thr, 0.34; Ser, 0.21 | 12 hr, pH 7.0, 40°: Glu, 0.21; Thr, 0.03; His, 0.00 |
| Ser-Thr-CMCys-Tyr | Ser-Thr-CMCys |
| C-5 (yield, 33%) | C-6 (yield, 15%) |
| CMCys, 1.98(2); His, 1.03(1) | CMCys, 1.98(2); His, 1.03(1) |
| CMCys, 2.88(3); Ser, 0.51(0) | CMCys, 2.32(2); Ser, 0.95(1) |
| Tyr, 0.84; CMCys, 0.39; Thr, 0.34; Ser, 0.21 | 12 hr, pH 7.0, 40°: Glu, 0.21; Thr, 0.03; His, 0.00 |
| Ser-Thr-CMCys-Tyr | Ser-Thr-CMCys |
| C-7 (yield, 2%) | C-8 (yield, 0%) |
| His, 1.02(1); CMCys, 0.98(1) | His, 0.93(1); CMCys, 0.93; Ser, 1.02(1); Thr, 1.11(1) |
| CMCys, 0.99(1); Ser, 1.02(1); Thr, 1.00(1) | CMCys, 0.99(1); Ser, 1.02(1); Thr, 1.00(1) |
| CMCys-Ser-Thr | CMCys-Ser-Thr |
| Tyr, 1.06(1) | Tyr, 1.06(1) |
| Tyr, 0.96(1); Glu, 1.62(2); Asp, 1.00(1); Thr, 0.03; His, 0.00 | CMCys, 2.88(3); Ser, 0.51(0) |
| Tyr, 0.96(1); Glu, 1.62(2); Asp, 1.00(1); Thr, 0.03; His, 0.00 | 12 hr, pH 7.0, 40°: Glu, 0.21; Thr, 0.03; His, 0.00 |
| CMCys-Ser-Thr | CMCys-Ser-Thr |
| Tyr, 1.06(1) | Tyr, 1.06(1) |
| Tyr, 0.96(1); Glu, 1.62(2); Asp, 1.00(1); Thr, 0.03; His, 0.00 | CMCys, 2.88(3); Ser, 0.51(0) |
| Tyr, 0.96(1); Glu, 1.62(2); Asp, 1.00(1); Thr, 0.03; His, 0.00 | 12 hr, pH 7.0, 40°: Glu, 0.21; Thr, 0.03; His, 0.00 |
| CMCys-Ser-Thr | CMCys-Ser-Thr |
| Tyr, 1.06(1) | Tyr, 1.06(1) |
| Tyr, 0.96(1); Glu, 1.62(2); Asp, 1.00(1); Thr, 0.03; His, 0.00 | CMCys, 2.88(3); Ser, 0.51(0) |

* No dansyl amino acid was identified.
Peptide \( \alpha T(m) \text{-} 4 \) — In the tryptic hydrolysates of reduced \( S \)-carboxymethyl TSH and TSH-\( \alpha \), Glycopeptide \( \alpha T \text{-} 10 \) cochromatographed with Glycopeptide \( \alpha T \text{-} 7 \) on Bio-Gel P-10 (5). Although they can be separated by paper electrophoresis at pH 6.4, the yields are low. The sequence, summarized in Table IV, was determined from Peptide \( \alpha T(m) \text{-} 4 \) which can be isolated easily by gel filtration (Fig. 3, Fraction III). The \( NH_2 \)-terminal sequence was determined by the dansyl-Edman technique. The first 3 residues agreed with previous data from Peptide \( \alpha T \text{-} 10 \) (3, Table V). The 4th residue, histidine, could not be identified by the dansyl-Edman. By the difference in amino acid composition between \( \alpha T(m) \text{-} 4 \) and \( \alpha T \text{-} 10 \) (3, Table I), and because both peptides have the same \( NH_2 \)-terminal sequence, serine is placed at the COOH terminus of \( \alpha T(m) \text{-} 4 \). Hydrolysis of Peptide \( \alpha T(m) \text{-} 4 \) (0.7 \( \mu \)mole) with chymotrypsin (210 \( \mu \)g) in 0.5 ml of 0.06 M \( NH_4 \text{HCO}_3 \) at pH 8.0 and 40° for 2 hours yielded three peptides and free tyrosine. They were separated on a column of Sephadex G-25 as shown in Fig. 5a. Fraction I was the glycopeptide C-1; Fraction II was a tripeptide C-4 and Fraction IV consisted of free tyrosine. Fraction III contained two components with the same composition (Tyr, His, Lys, Ser), one of which gave a light pink color upon reaction with ninhydrin-collidine rather than purple. The two were separated by electrophoresis at pH 6.4; the sequence of the darker staining material was determined.

The positioning of maleyllysine as the 2nd residue from the COOH terminus of \( \alpha T(m) \text{-} 4 \) was based on the specificity of tryptic cleavage and the hydrolysis of \( \alpha T \text{-} 10 \) with carboxypeptidases A and B (3, Table V). The 3rd and 4th residues were positioned on the basis of the compositions of C-3 and C-4 as

| Amino Acid                  | \( \beta T(m) \text{-} 1 \) | \( \beta T(m) \text{-} 2 \) | \( \beta T(m) \text{-} 3 \) | \( \beta T(m) \text{-} 3a \) | \( \beta T(m) \text{-} 3b \) |
|-----------------------------|------------------|--------------|----------------|----------------|----------------|
| Lysine                      |                  |              |                |                |                |
| Histidine                   | 0.91(1)          | 1.07(1)      | 1.20(1)        | 0.96(1)        | 1.42(1)        |
| Arginine                    | 1.06(1)          | 0.85(1)      | 0.85(1)        | 0.94(1)        | 0.96(1)        |
| Aspartic acid               | 1.69(1)          | 3.68(4)      | 3.85(4)        |                |                |
| Threonine                   | 0.96(1)          | 1.08(1)      | 1.20(1)        | 1.04(1)        | 1.06(1)        |
| Serine                      | 1.06(1)          | 1.06(1)      | 1.14(1)        | 1.06(1)        | 1.13(1)        |
| Glutamic acid               | 1.06(1)          | 1.06(1)      | 1.20(1)        | 1.14(1)        | 1.20(1)        |
| Proline                     | 0.95(1)          | 0.92(1)      | 0.92(1)        | 0.86(1)        | 0.92(1)        |
| Glycine                     | 2.03(2)          | 0.86(1)      | 1.00(1)        | 1.00(1)        | 0.86(1)        |
| Alanine                     | 1.01(1)          | 1.44(1)      | 1.00(1)        | 1.00(1)        | 1.13(1)        |
| S-carboxymethyl cysteine    | 0.89(1)          | 4.02(4)      | 0.97(1)        | 2.80(3)        | 2.80(3)        |
| Valine                      | 2.03(2)          | 0.86(1)      | 1.00(1)        | 1.00(1)        | 0.86(1)        |
| Methionine                  | 1.02(1)          | 2.18(2)      | 1.06(1)        | 1.02(1)        | 1.02(1)        |
| Isoleucine                  | 0.86(1)          | 1.00(1)      | 1.00(1)        | 1.00(1)        | 1.00(1)        |
| Leucine                     | 0.96(1)          | 0.86(1)      | 1.00(1)        | 1.00(1)        | 1.00(1)        |
| Tyrosine                    | 0.92(1)          | 0.92(1)      | 0.92(1)        | 0.92(1)        | 0.92(1)        |
| Phenylalanine               | 0.96(1)          | 0.96(1)      | 0.96(1)        | 0.96(1)        | 0.96(1)        |
| Total Residues              | 9                | 4             | 21              | 5              | 16              |
| Yield %                     | 86               | 72            | 48              | 48             | 44              |

* Theory is 4 residues if all chains terminate with Ser-Tyr or Ser-Tyr-Met. Note that this peptide (Fraction I, Fig. 3) is eluted
  * The partial residue of methionine is from the COOH terminus of a portion of the \( \beta \) chain (see text).
well as the NH$_2$-terminal tyrosine of C-3 (Table IV). Hydrolysis of C-1 with carboxypeptidase A allowed tyrosine to be placed at the 5th residue.

The remaining portion of Peptide C-1 (0.65 μmole) was then treated with thermolysin. Thermolysin (200 μg) in 0.5 ml of 0.06 M NH$_4$HCO$_3$, pH 8.0, at 40° for 22 hours did not cause significant hydrolysis. Another 500 μg of thermolysin were then added and the reaction was continued for 18 hours at 40°. The products were separated on Sephadex G-25 as shown in Fig. 5b. Fraction II contains Peptide Th-2 which was purified on paper by electrophoresis at pH 3.6. Its composition, as well as sequence, is shown in Table IV. Fraction I contains both Peptide Th-1 and undegraded C-1 as indicated by the partial residues of threonine, S-carboxymethylcysteine, serine, and tyrosine (Table IV). Further separation of these two peptides was not attempted and the fraction was treated with subtilisin (54 μg) in 0.5 ml of 0.06 M NH$_4$HCO$_3$, pH 8.0, at 40° for 4 hours. The reaction mixture was then fractionated on Sephadex G-25. The elution profile is shown in Fig. 5c. Fraction I shown in this figure contained glycopeptides with COOH-terminal heterogeneity as indicated by its amino acid composition. Fraction II showed five ninhydrin-positive spots upon paper electrophoresis at pH 3.6. Only four of these contained sufficient material for analysis. The material nearest the anode contained Peptides S-4 and S-5, which were then separated by chromatography in Solvent I (3). The second, third, and fourth spots were S-2, S-8, and S-7, respectively. Their compositions as well as some of their sequences are also shown in Table IV. The overlaps of these subtilisin peptides along with the thermolysin peptide, Th-2, gave another 6 residues of the sequence. In order to

### Table V

peptides from maleylated TSH-β

| Fraction IV | Fraction VII-3 | Fraction V | Fraction VIII-1 | Fraction VIII-2 | Total | TSH-β |
|-------------|----------------|------------|-----------------|-----------------|-------|-------|
| 1.99(2)     | 1.11(1)        | 4.72(5)    |                 |                 | 9     | 9.2   |
| 0.96(1)     | 1.23(1)        | 0.97(1)    | 0.86(1)         |                 | 3     | 2.8   |
| 3.08(3)     | 1.04(1)        | 1.01(1)    | 3.13(3)         |                 | 11    | 10.5  |
| 0.95(1)     | 0.95(1)        | 0.99(1)    |                 |                 | 5     | 4.6   |
| 1.09(1)     | 0.89(1)        | 1.08(1)    |                 | 2.15(2)         | 4     | 7.0   |
| 1.11(1)     | 1.08(1)        | 0.98(1)    |                 | 1.96(2)         | 7     | 7.0   |
| 1.07(1)     | 2.10(2)        | 0.97(1)    |                 | 1.16(1)         | 4     | 4.2   |
| 1.08(1)     |                | 2.08(2)    |                 |                 | 6     | 5.0   |
| 1.01(1)     | 0.8(1)         |            |                 |                 |       |       |
| 0.86(1)     | 0.8(1)         |            |                 |                 |       |       |
| 1.79(2)     | 1.01(1)        | 0.97(1)    |                 |                 | 6     | 5.5   |
| 0.94(1)     | 1.03(1)        | 0.96(1)    |                 |                 | 4     | 4.3   |
| 2.99(3)     | 0.87(1)        | 2.82(3)    |                 |                 | 6     | 6.1   |
| 2.01(2)     | 1.28(1)        | 0.23       |                 |                 | 4     | 4.2   |
| 1.07(1)     | 1.00(1)        | 1.07(1)    |                 |                 | 11    | 10.8  |
| 21          | 10             | 38         |                 |                 | 112   |       |
| 86          | 12             |            | 72              |                 |       |       |

partially overlapping βT(m)-3 which contains no serine.
confirm that the 4th residue from the NH₂-terminus was histidine, Fraction I (0.60 µmole) of Fig. 5e was further hydrolyzed with subtilisin (50 µg, 8 hours; another 50 µg, 9 hours) at 40°C in 0.5 ml of 0.06 M NH₄HCO₃, pH 8.0. The reaction mixture was separated on a column of Sephadex G-50 as shown in Fig. 5d. Apparently Sephadex G-50 did not give any fractionation of the glycopeptides which contained different numbers of amino acid residues. Fraction I consisted of glycopeptides S-1 and S-3 which were separated by paper electrophoresis at pH 3.6. Their compositions are shown in Table IV. Hydrolysis of Peptide S-1 with carboxypeptidases A and B showed a Thr-Glu sequence at the COOH terminus of S-1. By difference, histidine was placed NH₂-terminal to the Thr-Glu sequence. The position of the glutamic acid was also confirmed by the composition of dipeptide S-6 which was isolated from Fraction II of Fig. 5d after paper electrophoresis at pH 3.6. Peptides S-2 and S-7 were also found in this fraction. The complete amino acid sequence of αT(m)-4 was thus established.

Tryptic Peptides from Maleylated, Reduced, S-carboxymethyl TSH-β—The separation on a column of Sephadex G-50 (fine) of the peptides from the tryptic hydrolysate of derivatized TSH-β is shown in Fig. 6. Details are given in the legend. Amino acid analyses showed that, except for Fractions III, VII, and VIII, the fractions did not require further purification. Their compositions are given in Table V together with their assignments as βT(m)-1, etc., which are based on their compositions. Fraction III was further purified by paper electrophoresis at pH 6.4; two components were seen. Fraction III-1 was assigned as βT(m)-3b and the other had the same composition as Fraction I. Electrophoresis at the same pH revealed 9 ninhydrin-positive components to be present in Fraction VII but only three contained significant amounts of peptide material. Their compositions are also given in Table V and show that the three can be designated as βT(m)-1 (Residues 1 through 9, Fig. 2), βT(m)-3a (Residues 14 through 18) which arose from a partial split of a Tyr-Cys(Cm) bond and βT(m)-4a (Residues 46 through 55) which originated from a partial split of a Tyr-Ala bond. Peptide βT(m)-1 was the main component of Fraction VI which gave no ninhydrin color. This is in agreement with subsequent data that showed βT(m)-1 to be derived from the NH₂-terminal sequence of TSH-β. Its NH₂-terminal group was blocked by the maleyl group; this leaves no free NH₂ groups available for reaction with ninhydrin. However, Fraction VII-2, which also contained Peptide βT(m)-1, gave a purple ninhydrin stain. This indicates that Peptide βT(m)-1 in Fraction VII-2 must have had its α-NH₂-maleyl group removed by hydrolysis during isolation. It is important to note that βT(m)-1 contains 2 methionine residues which strengthens the previous preliminary assignments of the cyanogen bromide fragments (βCNBr-1 and βCNBr-2) of TSH-β (3).

Fraction VIII was found to contain four ninhydrin-positive components, three of them in significant amounts. These were assigned as βT(m)-2, βT(m)-5a (Residues 56 through 59) resulting from a partial split of a Tyr-Lys bond, and βT(m)-6 which also arises from a chymotryptic-like split at the Tyr-Phe bond between Residues 74 and 75. The basis of assignment was their compositions (Table V) and comparisons with the cyanogen bromide fragments of TSH-β and the tryptic peptides derived from them (3, Fig. 5). The summation of the numbers of residues in the primary peptide (Table V) again was in good agreement with the composition of the entire chain. As subsequently shown, the partial residue of methionine in Peptide βT(m)-7 is significant.

The composition of Peptide βT(m)-3 equals that of Peptides βT-3 plus βT-4. Peptide βT-4, the tryptic glycopeptide of this chain, must be at the COOH terminus because it contains the arginine, thus the order of these two peptides is established. This order was confirmed by examination of a chymotryptic hydrolysate of the cyanogen bromide fragment, βCNBr-3 as shown later in Table VII. When added to βT(m)-3, the compositions of βT(m)-2, and βT(m)-4 account for 46 residues, 10 through 55, and includes all 4 leucine residues of TSH-β and one of the three histidines. The other two histidines are in βCNBr-4. Thus these three tryptic peptides of the maleylated derivative constitute the major portion of the cyanogen bromide fragment βCNBr-3 (49 residues), and their order is subsequently shown by the overlaps of chymotryptic peptides from βCNBr-3.
The composition of Peptide $\beta T(m)$-5 (Residues 56 through 69), which contains a methionine, shows that Peptide $\beta T$-5 (the overlap tryptic peptide between the two large cyanogen bromide fragments, $\beta CNBr-3$ and $\beta CNBr-4$) is connected to Peptide $\beta T$-9. The reason for placing $\beta T$-9 at the COOH terminus of Peptide $\beta T(m)$-5 is because $\beta T$-9 has an arginine residue at its COOH terminus and is in cyanogen bromide fragment $\beta CNBr-4$ which has a portion of $\beta T$-8 at its NH$_2$-terminal sequence. $\beta T(m)$-5 also overlaps $\beta CNBr-3$ and $\beta CNBr-4$ and, based on the sequence of $\beta T$-8 (Asp-Phe-Met'-Tyr-Lys (3)), the 14 residues of $\beta T(m)$-5 less Asp-Phe-Met plus the compositions of $\beta T(m)$-6 and $\beta T(m)$-7 (a total of 55 residues) match the composition $\beta CNBr-4$ with exception of a serine, a tyrosine, and some methionine. These discrepancies are caused by heterogeneity at the COOH terminus of TSH-$\beta$ as described below. The order of $\beta T(m)$-5 and $\beta T(m)$-6 was subsequently determined from the amino acid composition of a chymotryptic peptide from $\beta CNBr-4$.

Ordering of Peptide within TSH-$\beta$—The data obtained, so far, from tryptic digests of reduced S-carboxymethyl TSH derivatives with or without maleylation, and from CNBr cleavage of reduced S-carboxymethyl TSH, are insufficient to order all the primary tryptic peptides of TSH-$\beta$. The difficulty is that there are many chymotryptic-like cleavages (Met-His at Residues 9 and 10, Tyr-Cys(Cm) at Residues 18 and 19, Tyr-Phe at Residues 74 and 75) which occurred during tryptic hydrolysis of the derivatives. The cyanogen bromide fragments are also relatively large; two of them ($\beta CNBr-3$ and $\beta CNBr-4$) account for 101 residues. The position of $\beta CNBr-1$ and $\beta CNBr-2$ relative to $\beta CNBr-3$ was obtained through the hydrolysis of oxidized TSH-$\beta$ with trypsin. Data concerning the heterogeneity of the COOH terminus of TSH-$\beta$ were also obtained from this hydrolysate. The ordering of tryptic peptides within the CNBr fragments was then studied by hydrolysis of $\beta CNBr-3$ and $\beta T(m)$-7 with chymotrypsin.

Peptide $\beta T(ox)$-1—Fig. 7 shows the fractionation of the tryptic hydrolysate of TSH oxidized with performic acid, and the desired peptide, which should have the proposed Met-His bond intact (3, Fig. 5) because of conversion of the methionine to its sulfone, was found in Fraction III. This fraction contained two peptides which gave positive Sakaguchi reactions (for arginine), they were separated by electrophoresis at pH 6.4. The less acidic peptide had the expected composition (Residues 1 through 13) as shown in Table VI. Table VI also shows the relationship of this peptide to CNBr fragments and tryptic peptides from TSH-$\beta$ with and without maleylation. The order of primary tryptic peptides present in $\beta T(ox)$-1 is based on the specificity of trypsin and detection of NH$_2$-terminal phenylalanine in intact $\beta T(ox)$-1.

Chymotryptic Peptides from Peptide $\beta CNBr-3$—The primary tryptic peptides of this fragment are $\beta T$-2, $\beta T$-3, $\beta T$-4, $\beta T$-5, $\beta T$-6, and $\beta T$-7 (3). To order them, 0.1 $\mu$ mole of Peptide $\beta CNBr-3$ (3, Fig. 3) was hydrolyzed with chymotrypsin (50 $\mu$g) in 0.06 M NH$_4$HCO$_3$, pH 8.0, and at 40° for 24 hours. The resulting peptides were separated by paper electrophoresis at pH 6.4. The order of migration from the anode to the cathode is C-2, C-6, C-3, C-1, C-7, C-4, and C-5 and their amino acid compositions were determined.
compositions are shown in Table VII. The summation of all the isolated chymotryptic peptides, except Peptide C-5 which resulted from a partial split of Peptide C-4, equals that of the tryptic peptides present in BCNBr-3, from which all alanine seen in the amino acid composition of BCNBr-3 was isolated from any chymotryptic and tryptic hydrolysates. From Table VII, it can be seen that the chymotryptic pep-
### Table IX

Ordering of chymotryptic peptides from Peptide $\beta T(m)$-7 in relation to tryptic peptides

| Peptide | Tryptic Peptides | Chymotryptic Peptides | Tryptic Peptides | Chymotryptic Peptides |
|---------|------------------|-----------------------|------------------|-----------------------|
| $\beta T$-10 | Phe, 1.00(1) | Ser, 1.71(2); Tyr, 1.79(2); Pro, 1.17(1); Val, 0.85(1); Ala, 1.07(1); Ile, 1.05(1); CMCys, 2.35(2); Lys, 1.95(2); Gly, 1.00(1); Asx, 2.04(2); Thr, 1.04(1) |
| $\beta T$-13 | Ser, 0.96(1); Asx, 2.09(2); CMCys, 0.77(1); Ile, 1.91(2); His, 0.88(1); Glu, 1.07(1); Ala, 1.08(1); Lys, 1.16(1); Thr, 1.21(1); Tyr, 0.91(1) | CMCys, 1.08(1); Thr, 0.95(1); Lys, 1.97(2); Pro, 1.14(1); Glu, 0.71(1); Ser, 1.02(1); Tyr, 1.16(1) |
| $\beta T$-14a | Met, 1.00(1) | Ser, 0.96(1); Tyr, 1.04(1) |

### Carboxypeptidases A + B

| Reaction | Peptide Composition |
|----------|---------------------|
| $\beta T$-10 | Phe, Ser, Tyr, Pro, Val, Ala, Ile, Ser, CMCys, Lys, CMCys, Gly, Lys, CMCys, Asx, Thr, Asx, Tyr, Ser, Asx |
| $\beta T$-13 | Ser, Tyr, Pro, Val, Ala, Ile, Ser, CMCys, Lys, CMCys, Gly, Lys, CMCys, Asx, Thr, Asx, Tyr, Ser, Asx |

### Tryptic peptides from oxidized TSH-B

| Reaction | Peptide Composition |
|----------|---------------------|
| $\beta T$(ox)-13a | Ser, 0.96(1); Tyr, 1.04(1) |
| Dansylation | DNS-Ser |
| $\beta T$(ox)-13 | Ser, 0.98(1); Tyr, 1.06(1); MetU_{2}, 0.95(1) |
| Dansyl-Edman | Ser-Tyr-MetU_{2} |

* Upon paper electrophoresis at pH 6.4, free phenylalanine and methionine appeared at the same position. Their presence was shown by analyzing the whole fraction, without acid hydrolysis, on an amino acid analyser.

**Notes:**
- Tryptic peptides overlap the chymotryptic peptides, except in one location (the Tyr-Cys(Cm) sequence, between $\beta T$-3 and $\beta T$-4 as well as C-1 and C-2). The overlap at this point was obtained by the composition of Peptide $\beta T(m)$-3 in which cleavage at the susceptible bond did not occur. The complete sequence of $\beta CNBr$-3 is thus established. It should be noted that this fragment showed no detectable NH$_2$-terminal residue; in most cases, we were not able to identify histidine as its dansyl derivative.
- Ordering of Peptides from $\beta CNBr$-4—Five primary tryptic peptides ($\beta T$-9, $\beta T$-10, $\beta T$-11, $\beta T$-12, and $\beta T$-13) were found in $\beta CNBr$-4 (3, Fig. 5). The position of $\beta T$-9 in the sequence has been established from the composition of $\beta T(m)$-5 (Table V).
and the position of βT-10 by the isolation of an arginine-containing peptide from a chymotryptic digest of βCNBr-4. These data and the positioning of βT-11, 12 and 14a (3, Table III) from βT(m)-7 are described below.

Peptide βCNBr-4-C-2—In the separation of the cyanogen bromide fragments of the entire hormone, βCNBr-4 chromatographed with αCNBr-4 (3, Fig. 3) and the recoveries after their separation on paper were low. Accordingly, the entire fraction (3, Fraction V, Fig. 3) was hydrolyzed with chymotrypsin (0.18 μmole treated with 150 μg of chymotrypsin in 0.126 M NH₄HCO₃, pH 8.0, at 40° for 2 hours). The resulting peptides were separated by electrophoresis at pH 6.4. Peptide βCNBr-4-C-2 was the only chymotryptic peptide to contain arginine. It should be noted that αCNBr-4 did not contain arginine. The peptide was, therefore, easily located by the Sakaguchi reaction. Staining with ninhydrin indicated that the peptide was overlapped by a second peptide which was removed by electrophoresis at pH 1.9. The amino acid composition of Peptide βCNBr-4-C-2 and its position in the sequence established the order of three primary tryptic peptides, PT(m)-5, PT(m)-6, and PT(m)-7, as shown in Table VIII. The other chymotryptic peptides of the hydrolysate were difficult to fractionate, and the remaining order of the primary tryptic peptides was deduced from a chymotryptic hydrolysate of oxidized TSH-β. Peptide βT(m)-13 was isolated from Fraction VI, Fig. 7 by paper electrophoresis at pH 1.9 and Peptide βT(m)-13a was the sole component of Fraction VII, Fig. 7. Their amino acid compositions together with their relationships to the COOH terminus of Peptide βT(m)-7 are also given in Table IX. These data agree with the results of the hydrolysis of reduced, S-carboxymethyl TSH-β with carboxypeptidase A, in which the amount of methionine released was much less than that of tyrosine (2, Table II). Only trace amounts of homoserine and its lactone were found in the analysis of βCNBr-4 (3) which indicates that most of this particular preparation did not terminate with methionine. The overlap of the Ser-Tyr with the adjacent lysine was shown by the composition of βT(m)-7-C-4.

Sequence of TSH-β—The sequences of βT-2, βT-3, βT-5, βT-6, βT-7, βT-8, βT-9, βT-10a and βT-10b, βT-11, and βT-13 have been described (3, Tables VI and VII). The sequence of the NH₂-terminal nonapeptide, which was not isolated (3), is known from the sequences of βCNBr-1 and βCNBr-2. The sequence of the COOH-terminal tripeptide is shown in Table IX. The remaining primary tryptic peptides (βT-4 and βT-12) were determined by further hydrolysis with proteolytic enzymes.

Peptide βT(m)-3b—Glycopeptide βT(m)-3b has the same composition as Peptide βT-4 and was used for completion of the sequence in this region. The sequence of 8 residues from the NH₂ terminus of Peptide βT-4 had been determined with the dansyl-Edman and 2 residues from the COOH terminus with carboxypeptidases A and B (3, Table VI). Four additional residues from the NH₂ terminus were determined with Peptide βT(m)-3b as shown in Table X. To confirm the sequence, 0.1 μmole of βT(m)-3b was hydrolyzed with thermolysin (100 μg) in 0.1 ml of 0.06 M NH₄HCO₃, pH 8.0, at 40° for 24 hours. The reaction mixture was spotted on paper and the resulting peptides were separated by electrophoresis at pH 6.4. They migrated in order from the anode as C-2, C-4, C-3, and C-1 plus C-5. Their compositions, which establish the order of the tryptic peptides, are shown in Table IX.

COOH Terminus of TSH-β—A partial residue of methionine had been found in Peptide βT(m)-7 (Fraction I, Table V). This was shown to result from heterogeneity at the COOH terminus. First, free methionine was obtained in low yield from the chymotryptic hydrolysate of βT-7. Second, both a dipeptide, Ser-Tyr(βT(ox)-13a) and a tripeptide Ser-Tyr-MetO(βT(ox)-13) were found in the tryptic hydrolysate of oxidized TSH-β. Table X. Amino acid sequence of βT(m)-3b

| Peptide | Amino acid sequence of βT(m)-3b |
|---------|--------------------------------|
| CMCyS  | S-carboxymethylcysteine.      |

| Sequence | CMCyS-Leu-Thr-Ile-Ala-Thr-Thr-Val-CMCyS-Ala-Gly-Tyr-CMCyS-Met-Thr-Arg |
|----------|-----------------------------------------------------------------------|
| Dns-Eman | Chemotryptic peptides                                                 |
| Thermolysin peptides |                             |
| Th-1 (yield, 21%) | CMCyS, 1.06(1); Leu, 9.98(1); Thr, 7.82(3); Ile, 1.00(1); Arg, 1.14(1) |
| Th-2 (yield, 82%) | Val, 0.50(1); CMCyS, 1.97(2); Ala, 0.54(1); Gly, 1.18(1); Tyr, 1.14(1) |
| Dns-Eman |                             |
| Th-3 (yield, 30%) | Met, 0.93(1); Thr, 1.02(1); Arg, 1.13(1) |

* See Reference 3, Table III.

* After Steps 1, 3, 6, and 0 determination of the NH₂ terminus was not attempted.
Assignment of Amide Groups—Most of the amides in TSH-α and TSH-β were assigned based on data obtained by enzymatic hydrolysis of peptides to free amino acids. The results are summarized in Table XII. Five amides in each chain have not been assigned. The peptides containing these unassigned residues were resistant to hydrolysis by aminopeptidase M.

In retrospect, a more efficient approach would have been to carry out all studies on the separated chains. It should be noted that while limited amounts of material and the lack of recognition of the subunit nature of TSH until recently (1, 2) have necessitated combining data from studies on fragments from the undissociated hormone and the separate chains, no peptides have been isolated whose composition or sequence is incompatible with the structures given in Figs. 1 and 2. The distribution of methionine and arginine residues in both chains together with the relatively low numbers of such residues as leucine, isoleucine, histidine, and glycine greatly facilitated ordering of small tryptic peptides within larger fragments, in many cases simply by determination of the latter's composition.

The most difficult portion of the sequence was the region in TSH-α (Residues 83 through 95) adjacent to the carbohydrate unit at Residue 92. Two of the five tyrosines, three of the half-cystines, and all three of the histidines of TSH-α are in this region. Much of it was relatively resistant to proteolysis and the histidine at Residue 83 could not be assigned unequivocally on the basis of the Edman degradation. The major problem encountered with the dansyl-Edman was with histidine. In most instances, when this residue was exposed as a new NH2 terminus, the dansyl-histidine then formed may have been partially destroyed during hydrolysis and was not detected after subsequent thin layer chromatography. After the next step of degradation, the resulting NH2-terminal residue was always detected without ambiguity.

Other technical problems encountered were the apparently complete cleavage by trypsin of the Met–His bond at positions 9 to 10 of TSH-β and the heterogeneity of the COOH terminus of the same chain. Both were overcome by isolation of the appropriate peptides from a tryptic hydrolysate of oxidized TSH-β which established both the Met–Met sequence and showed that some molecules must terminate with methionine rather than tyrosine. The dipeptide, Ser–Tyr, was the only original tryptic peptide (3) not recognized to be in any of the cyanogen bromide fragments; it should have been obtained from the tryptic hydrolysate of βCNBr-4 (3, Table IX). Only 2.5 μmoles of TSH were used in the cyanogen bromide work and there was some uncertainty in the composition of this large fragment, although it was obtained in excellent yield. The isolation of Ser–Tyr–Met–COOH from the oxidized chain also

**Table XI**

Amino acid sequence of Peptide βT-12

| Peptide used | Technique used | Residue assigned |
|--------------|----------------|-----------------|
| DNS-CMCys    | Amn, Glu       | 1 and 1         |
| 30           | Carboxypeptidase N | 18 and 19     |
| 77           | Carboxypeptidase A | 30             |
| 36           | Carboxypeptidase A | 65 and 50     |
| 56           | Carboxypeptidase M | 63             |

© We were unable to identify DNS-His at this step of Edman degradation.
explains the partial residue of methionine released by treatment of either intact TSH or TSH-β with carboxypeptidases (2, 10). The data do not preclude the possibility that some β chain molecules terminate at lysine 110. The NH₂-terminal and the COOH-terminal heterogeneity of TSH-α and -β, respectively is probably caused by proteolysis (see Reference 3), and it is likely that, in the structure of the intact hormone, these termini are exposed to the environment. The hormone specific chain of LH, CII, also shows heterogeneity at its COOH terminus (11). In both TSH-α and -β, partially blocked NH₂-terminal residues of phenylalanine may also be present (2).

The approximate molecular weights of the two chains, calculated from the sequences and carbohydrate compositions, are 13,600 (TSH-α) and 14,700 (TSH-β). Approximation is necessary because of the terminal heterogeneities of the protein portions, the oligosaccharide portions, which contain partial residues of fucose and galactose and incomplete assignment of amide groups. Completion of the latter, determination of the disulfide bonds and of the structure of the oligosaccharides remain. It is likely that unequivocal assignments of all amides will be difficult because the polymorphism of TSH seen in gel electrophoresis may result from partial amiation at some positions. A few peptides with the same composition but different electrophoretic mobilities were isolated although lack of material has not yet permitted detailed investigation of all such peptides (3). These include the two tryptic glycopeptides of TSH-α which each migrate as two spots during electrophoresis (5). With the exception of heterogeneity at the chain termini, no other evidence for amino acid substitutions, which would be a reflection of allotropes such as found in carboxypeptidase A (12, 13) and bovine β-lactoglobulin (14), has been found thus far. With respect to the oligosaccharides, it is not known whether the heterogeneity at their chain termini, clearly present because of the partial residues of fucose and galactose, extends further into their structures. All the galactose was found to be in one oligosaccharide (5), that assigned to Residue 82 in TSH-α.

TSH and LH are sulfur-rich proteins; the chains of TSH contain 5 and 6 disulfide bonds, respectively, in molecules whose amino acid portions have molecular weights of about 10,800 and 13,100 and they must be extremely highly cross-linked. There is no evidence for free sulfhydryl groups. In TSH-α, two pairs of half-cystine residues occupy adjacent positions, one other is separated only by a residue of histidine. The half-cystines of TSH-β are more scattered.

Other points of interest in the sequences are that the same type of linkage, Asn-X-Thr is present as found in many other glycoproteins (compositions have recently been published (16, 18)), if one assumes that the TSH oligosaccharides are attached to the expected asparagine residues, 56 and 82 in TSH-α, and 23 in TSH-β. Indirect supporting evidence has been found for two of these positions (5). In the third case, position 82 in TSH-α, the data of Fig. 5 and Table IV show that the linkage cannot be to serine 89 and amino acid analyses of the material not extracted (2) still shows asparagine present. It must have remained in the aqueous phase after formation of its anilinothiazoline because of linkage to the polar carbohydrate moiety. In one case, the residue between the asparagine-carbohydrate link and threonine is apolar, -Ile- at position 57, which is an exception to the recent proposal concerning the influence of apolar residues in this position on the synthesis of simple carbohydrate units containing only N-acetylgalactosamine and mannose. No instances of histidine adjacent to the carbohydrate linkage, as found in TSH-α, were noted in a recent large compilation (16). There are no tripeptides with sequences of Asn-X-Thr (Ser) in either chain, which are not bonded to carbohydrate.

In general there are no notable clusters of polar or nonpolar side chains except in the NH₂-terminal portion of TSH-β where five of the six amino acids at positions 10 through 15 are occupied by polar residues which are followed by two hydrophobic clusters on either side of the carbohydrate at position 23. An interesting feature of the sequence becomes apparent when TSH-β (the hormone-specific chain for TSH (2)) is compared with the hormone-specific chain of LH, CII. The sequence of ovine CII chain has been reported recently by Liu et al. (17). This glycoprotein, of 115 residues, has a single carbohydrate moiety as does TSH-β, and it is located near the NH₂ terminus at position 13; that of TSH-β is at position 23. The two hormone-specific subunits differ markedly in amino composition; CII has many more hydrophobic side chains, contains 3 times as much proline and 2 rather than 11 tyrosine residues. However, there may be considerable homology between the two, particularly if single base changes in the triplet code are considered. Three tetrapeptide sequences are identical, Ala-Gly-Tyr-Cys- (28 through 31, TSH-β; 33 through 36, CII), Val-Cys-Thr-Tyr (51 through 54, TSH-β; 55 through 58, CII), and Pro-Gly-Cys-Pro (65 through 68, TSH-β; 69 through 72, CII). Both tyrosines of the CII chain are in two of these sequences. A nearly identical sequence of 8 residues, Phe-Ser-Tyr-Pro-Val-Ala-Ile-Ser- (75 through 82 in TSH-β) and Val-Ser-Phe-Pro-Val-Ala-Leu-Ser (79 through 86, CII) is also found. It is a reasonable hypothesis that these regions in each hormone-specific subunit are major portions of the contact points in their binding with the common TSH-α-C1 type of chain. Thus these residues may be exposed to the environment in the same relative positions in the three-dimensional structures of TSH-β and CII. Furthermore, one of the few clearly polar pairs in CII is -Lys- Glu- at positions 20 to 21 which is 11 residues away from its first common tetrapeptide sequence; the same 2 residues are also found in TSH-β at positions 14 to 15 which is 12 residues away from its first common sequence. Clearly determination of the disulfide bonds of these subunits will contribute additional information. The sequence of TSH α does not appear on superficial examination, to be as closely related to either hormone-specific subunit as these are to each other.

Acknowledgments—We wish to express our particular appreciation to Miss Tiina Reimo for much technical assistance, to Mr. R. B. Carlsen for help including examination of the sequences for similarities, and to Dr. A. N. Glaser for encouragement and counsel. The advice and suggestions of Doctors R. J. DeLange and E. L. Smith are also gratefully acknowledged as are amino acid analyses carried out by Miss Dorothy McNall.

REFERENCES
1. PIERCE, J. G., AND LIAO, T.-H., Fed. Proc., 29, 600 (1970).
2. LIAO, T.-H., AND PIERCE, J. G., J. Biol. Chem., 245, 3275 (1970).
3. Shome, B., Liao, T.-H., Howard, S. M., and Pierce, J. G., J. Biol. Chem., 246, 833 (1971).
4. Butler, P. J. G., Harris, J. I., Hartley, B. C., and Leberman, R., Biochem. J., 112, 679 (1969).
5. Howard, S. M., and Pierce, J. G., J. Biol. Chem., 244, 6468 (1969).
6. Gray, W. R., in C. H. W. Hirs (Editor), Methods in enzymology, Vol. XI, Academic Press, New York, 1967, pp. 139, 469.
7. Woods, K. R., and Wang, K.-T., Biochim. Biophys. Acta, 133, 399 (1967).
8. Liao, T.-H., Hennen, G., Howard, S. M., Shome, B., and Pierce, J. G., J. Biol. Chem., 244, 6458 (1969).
9. Pierce, J. G., Liao, T.-H., Carlsen, R. B., and Reimo, T., J. Biol. Chem., 246, 866 (1971).
10. Shome, B., Brown, D. M., Howard, S. M., and Pierce, J. G., Arch. Biochem. Biophys., 126, 456 (1968).
11. Liu, W.-K., Sweeney, C. M., and Ward, D. N., Res. Commun. Chem. Path. Pharmacol., 1, 214 (1970).
12. Petra, P. H., Bradshaw, R. A., Walsh, K. A., and Neurath, H., Biochemistry, 6, 2762 (1969).
13. Neurath, H., Bradshaw, R. A., Petra, P. H., and Walsh, K. A., Phil. Trans. Roy. Soc. London Ser. B. Biol. Sci., 267, 159 (1970).
14. McKensié, H. A., Advan. Protein Chem., 22, 56 (1967).
15. Jackson, R. L., and Hirs, C. H. W., J. Biol. Chem., 245, 624 (1970).
16. Hunt, L. T., and Dayoff, M. O., Biochem. Biophys. Res. Commun., 39, 757 (1970).
17. Liu, W.-K., Sweeney, C. M., Nahm, H. S., Holcomb, G. N., and Ward, D. N., Res. Commun. Chem. Path. Pharmacol., 1, 463 (1970).
