Human Chorionic Gonadotropin

LINEAR AMINO ACID SEQUENCE OF THE α SUBUNIT*

(Received for publication, February 14, 1973)

RONALD BELLISARIO, ROBERT B. CARLSEN, AND OM P. BAHL†

From the Department of Biochemistry, Faculty of Health Sciences, State University of New York at Buffalo, Buffalo, New York 14214

SUMMARY

The linear amino acid sequence of the α subunit of human chorionic gonadotropin (hCG-α) has been derived from a study of the tryptic and cyanogen bromide peptides. The total number of amino acids varies from 89 to 92 because of NH₂-terminal heterogeneity probably caused by proteolysis of hCG. The molecular weight of hCG-α is estimated to be 14,900 as calculated from its chemical composition, approximately 10,200 for the protein and 4,700 for the carbohydrate part of the molecule. The carbohydrate moiety consists of 2 bulky carbohydrate units attached to the amide groups of asparagine residues 52 and 78 by N-glycosidic bonds. Assignment of the five disulfide bonds and some of the amide groups has not been made. There are no free sulfhydryl groups or tryptophan present in hCG-α.

The amino acid sequence of hCG-α exhibits considerable homology with the α subunits of luteinizing (LH) and thyroid-stimulating (TSH) hormones. The sequences of human LH-α and hCG-α differ only in a 2-residue inversion and a 3-residue deletion at the NH₂ terminus of hLH-α. There are 25 amino acid substitutions in hCG-α when compared with ovine or porcine LH-α or bovine TSH-α. The structure of the carbohydrate units of hCG-α appears to be different from that of the carbohydrate units in the α subunits of the other hormones.

Developments in the chemistry of human chorionic gonadotropin have been quite rapid. First, simple procedures suitable for large scale preparation of hCG were developed (1, 2). The hCG preparations obtained by these procedures possessed similar physicochemical properties. Subsequently, initial studies on the monosaccharide sequence of the carbohydrate moiety of native hCG, based on the sequential removal of monosaccharides by specific glycosidases, were reported (3). Initial studies based on NH₂- and COOH-terminal analyses and/or gel filtration of reduced and alkylated hCG suggested the presence of two identical chains (1, 2). However, polyacrylamide gel electrophoresis of S-carboxymethylated hCG indicated the presence of two dissimilar chains, α and β (then called A and B chains, respectively), which differed in their electrophoretic mobility and amino acid compositions (4). Further evidence for the nonidentical nature of the chains came from the observation that during the S-carboxamidomethylolation of hCG, the alkylated derivative of hCG α precipitated during dialysis (5). Whether the two chains were bonded noncovalently or by disulfide bonds remained unresolved. The dissimilarity of the subunits and the noncovalent nature of their attachment were unequivocally established by the separation of the α and β subunits from the native hormone on DEAE-Sephadex (5) and by their recombination (8). The reconstituted hCG was indistinguishable from native hCG in electrophoretic, immunological, and biological properties (7). This separation procedure was found suitable for the preparation of the subunits (5, 8) on a large scale. As a consequence, it became possible to initiate the work on their complete amino acid sequences (8). A preliminary report of the amino acid sequences of hCG-α and hCG-β has been made (10).

In this communication details of the isolation of the tryptic and cyanogen bromide peptides of hCG-α and their amino acid compositions and sequences are described. Comparisons between the amino acid sequences of hCG-α and the α subunits of bovine TSH (11) and human (12, 13), bovine (14, 15), ovine
MATERIALS AND METHODS

Trypsin (three times crystallized) and leucine aminopeptidase (treated with diisopropylphosphorofluoridate) were obtained from Worthington, chymotrypsin (3X crystallized) was supplied by Sigma, and thermolysin was purchased from Daiwa, Kasei, Osaka, Japan. All Sephadex preparations were from Pharmacia.

The hCG used in these studies was purified as reported earlier (1) from a crude commercial preparation, 2900 i.u., purchased from Organon, West Orange, N. J. The purified hCG was disassociated into subunits with 8 M urea and the subunits were separated by successive chromatography on DEAE-Sephadex and Sephadex G-100 (5, 9).

Paper Chromatography and Electrophoresis (19)—Paper electrophoresis was carried out in Varsol-cooled tanks at 60 to 70 volts per cm. The electrophotoetic buffer systems were 98% formic acid-acetic acid-water (25:37:980, v/v), pH 1.8; pyridine-acetic acid-water (20:20:900, v/v), pH 4.75; and pyridine-formic acid-acetic acid-water (25:87:890, v/v), pH 1.8; pyridine-water-triethylamine buffer (4:2.2:0.3), and 100 ~1 of 0.2% tryptamine at 110° in evacuated, sealed tubes. The oxidation of S-carboxymethylcysteine and the destruction of tyrosine were minimized by the addition of 2 ~1 of thioglycolic acid and 50 ~1 of a 5% solution of phenol to samples prior to hydrolysis (21). The tryptophan content of hCG-a was determined by hydrolysis with 3 N p-toluenesulfonic acid for 24 hours at 4°. Prior to use, the cyanogen bromide was treated with diisopropylphosphorofluoridate (sequential grade, Pierce Chemical Co., Rockford, Ill.). Suitable aliquots (10 ~m) were taken from the aqueous phase for dansylation and the remainder dried over NaOH and P2O5 at room temperature in vacuo prior to the next cycle.

Desialyization—A 2% solution of hCG-a in 0.025 N HCl was desialized at 80° for 1 hour. The desialized hCG-a was freed of sialic acid and salts by chromatography on coarse Sephadex G-25 with an elution buffer of 0.5% NH4HCO3. The solution of desialized hCG-a was concentrated by ultrafiltration with a UM-2 membrane (Amicon Corp.) and was lyophilized.

Reduction and S-Carboxamidomethylation and S-Aminomethylation of hCG-a—The S-carboxamidomethyl derivative of the desialized, reduced subunit was prepared as described in a succeeding paper (18).

S-Aminomethylation of reduced hCG-a was carried out essentially as described by Raftery and Cole (26). A solution of 107 mg of hCG-a in 20 ml of 8 M urea (ultrapure, Mann), containing 0.2% EDTA and 0.7 ~1 of 0.025 N Tris-HCl buffer, pH 8.5, was flushed with nitrogen, and 77 mg of dithiothreitol were added. The reaction mixture was left at room temperature for 1 hour and a total of 1.2 ml of ethyleneimine (Pierce Chemical Co.) was then added in three equal portions at 10-minute intervals. The S-aaminoethyl derivative was obtained free of salts with quantitative recovery by chromatography on coarse Sephadex G-25 in 0.5% NH4HCO3, followed by lyophilization.

Hydrolysis of Desialized, S-Carboxamidomethyl hCG-a with Trypsin—A solution of 110 mg of the desialized, S-carboxamidomethylated hCG-a in 15 ml of 0.1 M NH4HCO3 was treated with 200 ml of a trypsin solution (3 mg per ml). After 20 min and 40 min, 200 ~m aliquots of the enzyme were again added. The hydrolysis was stopped after 65 min by freeze drying the sample. The primary fractionation of the tryptic peptides was carried out by gel filtration as shown in Figs. 1 and 2. Further fractionation of the various Sephadex fractions was done by paper chromatography and/or paper electrophoresis.

Preparation of Cyanogen Bromide Cleavage Fragments of S-Aminoethyl hCG-a—The cleavage was carried out by treating 107 mg of S-aaminoethyl hCG-a with an equal weight of cyanogen bromide (50-fold excess over methionine) in 2.5 ml of 80% formic acid for 24 hours at 4°. Prior to use, the cyanogen bromide was sublimed and the formic acid was purified by fractional crystallization (27).

The reaction mixture was diluted with 40 ml of glass-distilled water and lyophilized. The resulting 130 mg of dried residue were dissolved in 5 ml of 1% proprionic acid and fractionated by gel filtration on Sephadex G-75 as shown in Fig. 4. The cyanogen bromide glycopeptides, found in Fraction IV of Fig. 4, were further separated from non-carbohydrate-containing peptides by countercurrent distribution as shown in Fig. 5. The solvent system of Howard and Pierce (21), formed by equilibrating equal volumes of 0.05 M p-toluenesulfonic acid and sec-butyl alcohol, was utilized. The peptide sample (usually 25 mg) was lyophilized and dissolved in 1.0 ml of 0.05 M p-toluenesulfonic acid saturated with sec-butyl alcohol. Twenty-five transfers utilizing 1 ml of each phase were carried out at room temperature in 12-ml glass-stoppered centrifuge tubes.

Upon completion, the two phases were broken by the addition of 0.1 ml of methyl alcohol, and aliquots (10 ~m) were taken from...
equilibrated with 0.05 M NH₄HCO₃ and were hydrolyzed with thermolysin, trypsin, or chymotrypsin (2 to 20% by weight) at 37°C for periods of 2 to 24 hours.

Determination of Amides—A 25-nmole sample of peptide was hydrolyzed with 12.5 μg of leucine aminopeptidase in 25 μl of 0.1 M Tris-HCl, and 0.0025 M MgCl₂, pH 8.8, for 40 hours at 37°C. The reaction mixture was analyzed for glutamine or asparagine as described in the following paper (18).

Nomenclature of Peptides—The major tryptic and cyanogen bromide peptides are numbered according to their sequence order; αT-1 and CNBr-1 are the respective NH₂-terminal peptides. The minor tryptic peptides are designated by adding the suffix of a lower case letter, such as αT-1a. Peptides obtained from a secondary enzymatic digest with thermolysin (Th), chymotrypsin (C), or trypsin (T) are designated by adding the respective suffixes C-, Th-, or T- to the designation of the major peptide, again numbering them in the order of their sequence. Example: αT-1, Th-1 is the NH₂-terminal thermolysin peptide of tryptic peptide αT-1.

RESULTS

Isolation and Sequencing of Tryptic Peptides from Reduced S-Carboxamidomethylated, Desialyzed hCG-α—The Sephadex G-25 elution pattern of the tryptic hydrolysate of the reduced S-carboxamidomethylated, desialylated hCG-α is shown in Fig. 1. Fraction 1 from the Sephadex G-25 column was further fractionated on Sephadex G-50 (Fig. 2) to effect a more complete separation of the tryptic glycopeptides from other large tryptic peptides. Each of the fractions from the Sephadex columns, shown in Figs. 1 and 2, was further fractionated by high voltage paper electrophoresis and/or paper chromatography as described for each tryptic peptide given below.

The amino acid compositions and yields of the tryptic peptides are given in Tables I and II. Residues sequenced by Edman degradation and dansylation are given with an arrow (−) over the residue in the text. Where they were established, the proper amide assignments are indicated in the peptide sequences.

Peptide αT-1 (Residues 1 to 28) Ala-Pro-Asx-Val-Glx-Asx-Cys(Ca)-Pro-Glx-Cys(Ca)-Thr-Leu-Glx-Glx-Asx-Pro-Phe and Peptide αT-1b (Residues 3 to 17) Asx-Val-Glx-Asx-Cys(Ca)-Pro-Glx-Cys(Ca)-Thr-Leu-Glx-GlxAx-Pro-Phe—These peptides presumably arose from a chymotryptic cleavage of peptide αT-1 at phenylalanine (residue 17) and were separated by high voltage electrophoresis at pH 4.7 of Fractions 2 and 3 (Fig. 2). Sequence analyses of these peptides showed that peptide αT-1a has the NH₂-terminal sequence Ala-Pro-Asx-Val-Glx, identical with...
### Table I

**Amino acid compositions of the major tryptic peptides of S-carboxamidomethyl hCG-α**

| Amino Acid | aT-1 | aT-2 | aT-3 | aT-4 | aT-5 | aT-6 | aT-7 | aT-8 | aT-9 | aT-10 | aT-11 | aT-12 | aT-13 | Total Residues |
|------------|------|------|------|------|------|------|------|------|------|------|------|------|------|---------------|
| Lysine     | 0.12 |      |      |      |      |      |      |      |      |      |      |      |      | 6              |
| Histidine  |      |      |      |      |      |      |      |      |      |      |      |      |      | 3              |
| Arginine   |      |      |      |      |      |      |      |      |      |      |      |      |      | 3              |
| Aspartic Acid | 2.60 (3) | 0.13 |      |      |      |      |      |      |      |      |      |      |      | 6              |
| Threonine  |      |      |      |      |      |      |      |      |      |      |      |      |      | 6              |
| Serine     | 0.81 (1) | 0.20 |      |      |      |      |      |      |      |      |      |      |      | 6              |
| Glutamic Acid | 6.42 (6) |      |      |      |      |      |      |      |      |      |      |      |      | 6              |
| Proline    | 4.65 (5) | 0.13 |      |      |      |      |      |      |      |      |      |      |      | 6              |
| Glycine    | 1.17 (1) | 1.09 (1) |      |      |      |      |      |      |      |      |      |      |      | 6              |
| Alanine    | 1.35 (2) |      |      |      |      |      |      |      |      |      |      |      |      | 6              |
| S-carboxymethylcysteine | 3.36 (3) | 0.98 (1) |      |      |      |      |      |      |      |      |      |      |      | 6              |
| Valine     | 0.80 (1) |      |      |      |      |      |      |      |      |      |      |      |      | 10             |
| Methionine |      |      |      |      |      |      |      |      |      |      |      |      |      | 10             |
| Isoleucine | 0.90 (1) |      |      |      |      |      |      |      |      |      |      |      |      | 10             |
| Leucine    | 2.12 (2) |      |      |      |      |      |      |      |      |      |      |      |      | 10             |
| Tyrosine   | 0.97 (1) |      |      |      |      |      |      |      |      |      |      |      |      | 10             |
| Phenylalanine | 1.99 (2) |      |      |      |      |      |      |      |      |      |      |      |      | 10             |
| Total Yield | 20    | 3    | 2    | 2    | 7    | 3    | 6    | 12   | 4    | 8    | 13   | 3    | 1    | 92             |

* No corrections were made for destruction of amino acids during hydrolysis.

* Represents the final yield with no corrections for manipulative losses.

* The low values for aspartic acid, proline, and alanine may be due to NH₂-terminal heterogeneity.

* Includes peptide aT-3, not isolated from the tryptic hydrolysates.
peptide αT-1, while peptide αT-1b has the NH$_2$-terminal se-
quence Asx-Val, with a faint background of Ala-Pro, indicating
contamination from peptide αT-1a. Thus, αT-1b lacks the
NH$_2$-terminal dipeptide Ala-Pro of αT-1a, substantiating the
NH$_2$-terminal heterogeneity of hCG-α. The total sequence of
peptide αT-1a was determined from thermolysin peptides ob-
tained from the NH$_2$-terminal cyanogen bromide peptide CNBr-1
shown in Table VII.

**Table II**

| Amino Acid | αT-1a | αT-1b | αT-1c | αT-2a | αT-2b | αT-6a | αT-6b | αT-7a | αT-7b | αT-9a | αT-10a | αT-12a | αT-17b |
|------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Lysine     | 0.94(1) | 1.00(1) | 1.14(1) | 0.16  |       |       |       |       |       |       |       |       |       |
| Histidine  | 0.16  |       |       |       |       |       |       |       |       |       |       |       |       |
| Arginine   |       |       |       |       |       |       |       |       |       |       |       |       |       |
| Aspartic Acid | 3.12(3) | 3.09(3) | 0.20  | 0.16  | 0.21  | 0.16  | 0.16  | 0.16  | 0.16  | 0.16  | 0.16  | 0.16  | 0.16  |
| Threonine  | 0.95(1) | 0.88(1) | 0.16  | 1.06(1) | 0.12  | 0.24  | 1.04(1) | 0.25  | 0.12  | 0.07(1) | 0.07(1) | 0.07(1) | 0.07(1) |
| Serine     | 0.23  | 0.13  | 0.92(1) |       |       |       |       |       |       |       |       |       |       |
| Glutamic Acid | 6.13(4) | 4.69(4) | 2.04(2) |       |       |       |       |       |       |       |       |       |       |
| Proline    | 2.79(3) | 2.27(2) | 2.14(2) | 0.08(1) | 0.15  | 0.37  |       |       |       |       |       |       |       |
| Glycine    | 0.26  | 0.15  | 1.08(1) |       |       |       |       |       |       |       |       |       |       |
| Alanine    | 0.82(1) | 0.25  | 1.02(1) | 0.02(1) | 0.22  | 0.55  | 0.34  |       |       |       |       |       |       |
| S-caryosymethylalanine | 2.04(2) | 1.87(2) | 0.81(2) |       |       |       |       |       |       |       |       |       |       |
| Valine     | 0.96(1) | 1.08(1) | 0.22  | 1.02(1) |       |       |       |       |       |       |       |       |       |
| Methionine | 1.00(1) |       |       |       |       |       |       |       |       |       |       |       |       |
| Isoleucine | 0.84(1) |       |       |       |       |       |       |       |       |       |       |       |       |
| Leucine    | 1.00(1) | 1.00(1) | 1.07(1) |       |       |       |       |       |       |       |       |       |       |
| Tyrosine   | 0.13  | 1.10(1) |       |       |       |       |       |       |       |       |       |       |       |
| Phenylalanine | 0.06(1) | 0.05(1) | 1.00(1) |       |       |       |       |       |       |       |       |       |       |

**Total Yields**

|          | 7.4  | 7.0  | 10.1 | 10.5 | 2.3  | 3.6  | 32  | 1.5  | 3.4  | 3.4  | 14.0 | 5.2  |

The peptide was hydrolyzed with thermolysin and yielded two
peptides, αT-5, Th-1 (Thr, 1.10(1); Pro, 2.29(2); Ala, 0.95(1);
Tyr, 0.45(1), and αT-5, Th-2 (Arg, 1.00(1); Leu, 1.08(1))
which were separated by paper electrophoresis at pH 1.9. The
sequence of αT-5, Th-1, determined by four steps of the Edman
procedure with dansylation detected a background of glutamic acid
appearing one step early had occurred between Cys(Cα)-Phosphoryl
hCG-α. The NH$_2$-terminal residue was isolated from the tryptic hydrosylate of S-carboxamidomethyl hCG-α. S-Carboxamidomethylglycine and phenyl-
alanine were isolated from Fraction 7 (Fig. 1) in low yield (4.7% and
1.4%, respectively), suggesting an anomalous tryptic cleav-
age had occurred between Cys(Cα)-Phos (peptide 32 to 33). The
positions of these residues were unequivocally assigned from the
sequence of peptides CNBr-2, Th-1 and CNBr-2, Th-2 (Table IX).

**Peptide αT-4 (Residues 34 to 35) Ser-Arg—Successive paper
electrophoresis at pH 1.9 and pH 4.7 of Fraction 7 (Fig. 1) yielded
peptide αT-4, which stained brown with ninhydrin. The
NH$_2$-terminal residue was identified as serine by the dansyl technique,
which established the sequence.

**Peptide αT-5 (Residues 36 to 42) Ala-Tyr-Pro-Thr-Pro-Leu-
Arg—Peptide αT-5 was isolated from a tryptic hydrolysate of
reduced and carboxamidomethylated hCG-α. The sequence of
the first 4 residues was determined by subtractive Edman degra-
dation.

Step 1: Arg, 1.00; Thr, 1.03; Pro, 1.70; Ala, 0.12; Leu, 1.01;
Tyr, 0.31
Step 2: Arg, 1.00; Thr, 1.01; Pro, 1.59; Ala, 0.07; Leu, 0.98;
Tyr, 0.08
Step 3: Arg, 1.00; Thr, 0.95; Pro, 1.99; Ala, 0.08; Leu, 1.01;
Tyr, 0.08
Step 4: Arg, 1.00; Thr, 0.32; Pro, 1.20; Ala, 0.07; Leu, 1.00;
Tyr, 0.00

The peptide was hydrolyzed with thermolysin and yielded two
peptides, αT-5, Th-1 (Thr, 1.10(1); Pro, 2.29(2); Ala, 0.95(1);
Tyr, 0.45(1), and αT-5, Th-2 (Arg, 1.00(1); Leu, 1.08(1))
which were separated by paper electrophoresis at pH 1.9. The
sequence of αT-5, Th-1, determined by four steps of the Edman
procedure with dansylation was Ala-Tyr-Pro-Thr-Pro. The
sequence of αT-5, Th-2 was obviously Leu-Arg since the parent peptide
αT-5 was a tryptic peptide.
identified by dansylation as serine, establishing the sequence of αT-6.

Peptides αT-6a and αT-6b were isolated from Fraction 5 (Fig. 1) by high voltage electrophoresis at pH 1.9. 

Peptide αT-7 (Residues 46 to 51) Thr-Met-Leu-Val-Gln-Lys; Peptide αT-7a (Residues 48 to 51) Leu-Val-Gln-Lys—Peptide αT-7 was separated from peptide αT-10 by subjecting Fractions 2 and 3 (Fig. 1) to high voltage electrophoresis at pH 1.9 followed by paper chromatography with Solvent 1. Paper chromatography separated peptide αT-7 into two peptides with identical amino acid compositions, presumably due to partial oxidation of the methionine at position 47.

Peptide αT-7a, resulting from a partial cleavage at methionine (residue 47), was isolated as a minor fraction during the purification of peptide αT-7 and was not studied further.

Peptide αT-8 (Residues 52 to 68) Asn(CHO) Val Thr Ser Glx Ser-Thr-Cys(Ca)-Cys(Ca)-Val-Ala-Lys—Glycopeptide αT-8 was contained in Fractions 2 and 3 (Fig. 2) along with glycopeptide αT-11 and a mixture of peptides from the NH₂ terminus. High voltage electrophoresis at pH 4.7 separated αT-8, a basic peptide, from the acidic NH₂-terminal peptides as well as from the slightly faster migrating basic glycopeptide αT-11. The sequence of peptide αT-8 was determined by Edman degradation with dansylan and appeared as a very faint fluorescent spot on the thin layer plates. In order to confirm the sequence of αT-8, 0.5 μmole of the peptide was hydrolyzed with thermolysin (0.1 mg) in 0.5 ml of 1% NH₄HCO₃ at 37° for 6 h. The hydrolysate was lyophilized, and two thermolysin peptides, αT-8-Th-1 and αT-8-Th-2, were separated by high voltage electrophoresis at pH 1.9. The amino acid compositions of the thermolysin peptides are given in Table III (see miniprint supplement p. 3). Sequence analysis showed peptide αT-8-Th-1 to be Asn(CHO)-Val-Thr-Glx-Ser-Thr-Cys(Ca)-Cys(Ca), and αT-8-Th-2 to be Val-Ala-Lys, corroborating the sequence of αT-8. These results are summarized in Table IV (see miniprint supplement p. 4).

Peptide αT-9 (Residues 64 to 67) Ser-Tyr-Asn-Ara; Peptide αT-9a (Residues 64 to 65) Ser-Tyr—Peptide αT-9 was isolated from the tryptic hydrolysate of S-carboxamidomethylated hCG, and its sequence was determined by two steps of subtractive Edman degradation.

Step 1: Asp, 1.03; Ser, 0.30; Tyr, 0.97
Step 2: Asp, 1.00; Ser, 0.32; Tyr, 0.48

Arginine was placed at the COOH terminus since αT-9 was a tryptic peptide.

Peptide αT-9a, contaminated with glycine (Table II), was obtained from the tryptic hydrolysate of S-carboxamidomethyl hCG-α by successive high voltage electrophoresis at pH 1.9 and paper chromatography with Solvent I of Fraction 7 (Fig. 1). The isolation of peptide αT-9a indicates that a chymotryptic-like cleavage had occurred at tyrosine (residue 65), although the corresponding peptide Asn-Arg was not isolated.

Peptide αT-10 (Residues 68 to 75) Val-Thr-Val-Met-Gly-Gly-Phe-Lys; Peptide αT-10a (Residues 68 to 71) Val-Thr-Val-Met—Successive high voltage electrophoresis at pH 1.9 and paper chromatography in Solvent I of Fractions 2 and 3 (Fig. 1) separated peptides αT-10 and αT-10a. Peptide αT-10 was obtained by paper chromatography in two fractions with identical amino acid compositions, presumably from partial oxidation of Met-71. Peptide αT-10a presumably arose from a chymotryptic-like cleavage at Met-71.

Peptide αT-11 (Residues 76 to 88) Val-Glx-Asn(CHO)-His-Thr-Ala-Cys(Ca)-His-Cys(Ca)-Ser-Thr-Cys(Ca)-Tyr—High voltage electrophoresis at pH 4.7 of Fractions 2 and 3 (Fig. 2) separated glycopeptides αT-8 and αT-11 as previously described. Unlike peptide αT-8, glycopeptide αT-11 smeared on paper electrophoresis giving a broad band. The NH₂-terminal sequence was determined to be Val-Glx. Peptide αT-11 (1.0 μmole) was hydrolyzed with 0.3 mg of thermolysin at 37° for 14 h in 0.7 ml of 0.5% NH₄HCO₃, which yielded seven peptides whose amino acid compositions are given in Table II. The hydrolysate was fractionated by gel filtration on Sephadex G-25 as shown in Fig. 3 (see miniprint supplement p. 1). Peptide αT-11-Th-1, obtained from Fraction 2, and peptide αT-11-Th-5 from Fraction 5 were found to be homogeneous as judged by paper electrophoresis and paper chromatography and were not purified further. The remaining thermolysin peptides from peptide αT-11-Th-2, αT-11-Th-2a, αT-11-Th-2b, αT-11-Th-3, and αT-11-Th-4 were isolated from Fraction 3 (Fig. 3) by high voltage electrophoresis at pH 1.9. Peptide αT-11-Th-3, which had NH₂-terminal serine, stained brown with ninhydrin. Peptide αT-11-Th-4, which also stained brown with ninhydrin, and peptide αT-11-Th-2b were impure and were not studied further. The amino acid compositions of the thermolysin peptides from peptide αT-11 are given in Table III. The sequences of the thermolysin peptides are shown in Table IV and their order was established with the chymotryptic peptides derived from peptide CNBr-4 (Table XIII).

Peptide αT-11 presumably arose from a chymotryptic-like cleavage at Tyr-88. However, the presence of the dipéptide αT-11-Th-5 (Tyr-Tyr) in the thermolysin hydrolysate of αT-11 indicates a partial chymotryptic-like cleavage at Tyr-89 must also have occurred during the tryptic hydrolysis of S-carboxamidomethyl hCG-α.

Peptide αT-12 (Residues 89 to 91) Tyr-His-Lys; Peptide αT-12a (Residue 89) Tyr; Peptide αT-12b (Residues 90 to 91) His-Lys—Peptide αT-12 was isolated from a tryptic hydrolysate of reduced and S-carboxamidomethylated hCG.

Peptides αT-12a and αT-12b were isolated after tryptic hydrolysis of S-carboxamidomethyl hCG-α, indicating that a chymotryptic-like cleavage at tyrosine (residue 89) had occurred. Peptide αT-12a was isolated from Fraction 4 (Fig. 1) by successive high voltage electrophoresis at pH 1.9 and paper chromatography using Solvent 1. In both cases, it migrated in the same position as a tyrosine standard.

Peptide αT-12b was isolated by high voltage electrophoresis at pH 1.9 of Fraction 5 (Fig. 1).

Peptide αT-13 (Residue 92) Ser—Peptide αT-13, which was serine, was isolated by high voltage electrophoresis at pH 1.9 of Fraction 5 (Fig. 1).

Isolation of Cyanogen Bromide Peptides from Reduced, S-Aminoethyl hCG-α—Amino acid analysis of the reduced S-aminoethyl hCG-α indicated that all of the cysteine residues were lost, with 84% recovery as S-a-monoethylcysteine. Although precise quantitation of S-a-monoethylcysteine was hindered due to incomplete separation from lysine on the 8-cm column, the results indicated that complete conversion of cysteine to S-a-monoethylcysteine had occurred.

Amino acid analysis of the mixture of cyanogen bromide reaction products showed 0.6 methionine residue remaining, indicating that 20% of the total methionine had not been converted to homoserine. Additional incubations with cyanogen bromide...
to 200-fold excess over methionine for 24 hours at 4°C resulted in no further conversion of methionine to homoserine.

The cyanogen bromide reaction products of S-aminoethyl hCG-α were fractionated by Sephadex G-75 column chromatography (Fig. 4). Fractions I to III were found to contain residual methionine when examined by amino acid analysis and most likely represent incompletely cleaved fragments or unreacted material. These fractions, comprising 33% by weight of the starting material, were not analyzed further.

Fraction IV (Fig. 4) contained peptide CNBr-1 and two glycopeptides, CNBr-3 and CNBr-4.

Fraction V contained peptide CNBr-2 (Fig. 4). This material was further purified by gel filtration through Sephadex G-25 to remove ultraviolet absorbing material and a small amount of CNBr-1 contamination. The amino acid composition of peptide CNBr-2, obtained in 31% yield from the starting material, is given in Table V.

Peptide CNBr-1 was separated from peptides CNBr-3 and CNBr-4 by countercurrent distribution as shown in Fig. 5. Amino acid analysis indicated that Fraction 1 of the countercurrent distribution system contained an approximately equimolar mixture of the glycopeptides, CNBr-3 and CNBr-4, each present in about 31% yield.

The glycopeptides were partially separated by gel filtration on Sephadex G-50 as shown in Fig. 6. CNBr-3 was obtained from Fraction 1 and CNBr-4 from Fraction 3. The middle fraction contained a mixture of the two glycopeptides and was repeatedly rechromatographed to effect further separation of the two glycopeptides. The compositions of CNBr-3 and CNBr-4 thus isolated are given in Table V.

**Table V**

Amino acid compositions of the cyanogen bromide peptides from S-aminoethyl hCG-α

| Amino Acid       | CNBr-1 | CNBr-2 | CNBr-3 | CNBr-4 | Total Residues | hCG-α |
|------------------|--------|--------|--------|--------|----------------|-------|
| Lysine           | 2.24 (2) | 2.06 (2) | 1.90 (2) | 0.39   | 3.00           | 5.84  |
| Histidine        | 2.01 (2) | 1.37 (1) | 1.11 (1) | 0.39   | 3.59           | 5.67  |
| Arginine         | 2.86 (3) | 2.30 (2) | 2.08 (2) | 6.50 (6) | 9.46           |
| Aspartic Acid    | 1.04 (1) | 1.94 (2) | 5.09 (5) | 2.23 (2) | 7.17           |
| Threonine        | 0.97 (1) | 1.92 (2) | 2.97 (3) | 2.18 (2) | 1.12 (1)       | 9.46  |
| Serine           | 1.37 (1) | 1.11 (1) | 1.06 (1) | 4.05 (5) | 7.02           |
| Glutamic Acid    | 1.07 (1) | 1.05 (1) | 1.26 (1) | 4.05 (5) | 7.02           |
| Proline          | 1.02 (1) | 1.04 (1) | 1.06 (1) | 4.05 (5) | 7.02           |
| Valine           | 1.13 (1) | 1.42 (2) | 1.62 (2) | 1.06 (1) | 7.02           |
| S-acetylarginine | 2.32 (3) | 1.42 (2) | 1.62 (2) | 1.06 (1) | 7.02           |
| Homoserine       | 0.83 (1) | 0.78 (1) | 0.96 (1) | 0.96 (1) | 1.96           |
| Alanine          | 0.92 (1) | 0.85 (1) | 0.96 (1) | 0.85 (1) | 1.96           |
| Tyrosine         | 1.92 (2) | 2.04 (2) | 2.04 (2) | 2.04 (2) | 4.08           |
| Phenylalanine    | 2.04 (2) | 0.98 (1) | 0.98 (1) | 0.98 (1) | 1.96           |
| Total Residues   | 29     | 18     | 24     | 21     | 92             |
| Yield            | 36%    | 31%    | 31%    | 31%    | 31%            |

* Assumed to be a contaminant; no evidence for glycine was found during the sequencing of peptide CNBr-1.
Fraction 2 from the countercurrent distribution (Fig. 5) contained peptide CNBr-1, which was further purified by gel filtration through Sephadex G-50 as shown in Fig. 7. The amino acid composition of peptide CNBr-1, obtained in 30% yield from Fraction 2 of the G-50 column, is given in Table V. Analysis of Fractions 1 (Fig. 7) revealed the presence of residual methionine and an amino acid composition similar to the sum of peptides CNBr-1 and CNBr-2. This material probably represents a peptide composed of CNBr-1 and CNBr-2 in which cleavage at Met-29 did not occur. It was not analyzed further.

Fraction 3 from the countercurrent distribution (Fig. 5) contained impure CNBr-1 in low yield and was not studied further.

Characterization of Cyanogen Bromide Peptides—hCG-α contains 3 methionine residues and yielded four peptides when cleaved with cyanogen bromide. Peptide CNBr-4 did not contain homoserine (Table V) and was therefore determined to be COOH-terminal. The positioning of the other three cyanogen bromide peptides, numbered according to their sequence order, was then established by means of the three methionine-containing tryptic peptides αT-2, αT-7, and αT-10.

NH₂-terminal sequences of the four cyanogen bromide peptides were determined by Edman degradation with dansylation; residues thus sequenced are shown in the text with a superscript arrow (→). The amide groups, wherever established, are indicated in the peptide sequences.

The complete sequences of the cyanogen bromide peptides were then determined by hydrolyzing the peptides with thermolysin, trypsin, or chymotrypsin and then sequencing the smaller secondary peptides. The ordering of these peptides was accomplished by means of NH₂-terminal analyses of the parent cyanogen bromide peptides and by the tryptic peptides previously described from reduced, carboxyamidomethylated, desialyzed hCG-α. These results, described below, are summarized in Tables V to XIII.

Peptide CNBr-1 (Residues 1 to 29) Ala-Pro-Asx-Val-Glx-Asx-Cys(Ae)-Pro-Gly-Cys(Ae)-Thr-Leu-Glx-Giz-Asx-Pro-Phe-Phe-Ser-Glx-Pro-Gly-Ala-Pro-Ile-Leu-Gly-Cys(Ae)-Hsr—The sequence of CNBr-1 was determined by hydrolysis with thermolysin and subsequent sequencing of the resulting thermolysin peptides. One micromole of peptide CNBr-1 was hydrolyzed with thermolysin (200 μg) in 1 ml of 1% NH₄HCO₃ at 37° for 3 hours. Six thermolysin peptides were separated by high voltage electrophoresis at pH 4.7. Peptides CNBr-1, Th-2 and CNBr-1, Th-5 were further purified by paper chromatography with Solvent I. The compositions of the thermolysin peptides are given in Table VI (see miniprint supplement p. 5). Their sequences are given in Table VII. Hydrolysis of peptide CNBr-1, Th-6 with leucine aminopeptidase indicated residue 27 to be glutamine. Three steps of Edman degradation with dansylation on peptide CNBr-1 showed the NH₂-terminal sequence Ala-Pro-Asx-Val. This information, plus the partial sequences and compositions of the NH₂-terminal tryptic peptides, was sufficient to establish the ordering of the thermolysin peptides from peptide CNBr-1 as shown in Table VII (see miniprint supplement p. 6).

Peptide CNBr-2 (Residues 30 to 47) Gly-Cys(Ae)-Cys(Ae)-Phe-Ser-Arg-Ala-Tyr-Pro-Thr-Pro-Leu-Arg-Ser-Lys-Lys-Thr-Hsr—The sequence of CNBr-2 was established by sequencing of its thermolysin peptides. An aliquot of 0.48 μmole of CNBr-2 was hydrolyzed with thermolysin (50 μg) in 0.25 ml of 1% NH₄HCO₃ at 37° for 2 hours. High voltage electrophoresis of the thermolysin hydrolysate at pH 1.9 yielded five peptides, one of which, CNBr-2, Th-1, stained gray-brown with ninhydrin. The compositions of the thermolysin peptides from peptide CNBr-2 are recorded in Table VIII (see miniprint supplement p. 7), and their sequences are given in Table IX (see miniprint supplement p. 8).

The sequence of the NH₂-terminal residues of peptide CNBr-2 was determined to be Gly-Cys(Ae)-Cys(Ae)-Phe-Ser-Arg-Ala-Tyr, and was sufficient to establish the positioning of peptides CNBr-2, Th-1, CNBr-2, Th-2, and CNBr-2, Th-3 as shown in Table IX. The positioning of peptide CNBr-2, Th-4 was determined with the overlapping tryptic peptide αT-5, which relegated the dipeptide CNBr-2, Th-5, Thr-Hsr, to the COOH-terminal position, as required for a cyanogen bromide peptide.
amino acid compositions of the five thermolysin peptides obtained from peptide CNBr-2 account for the composition of peptide CNBr-2, which precludes the possibility of residues between peptides CNBr-2, Th-4 and CNBr-2, Th-5.

Glycopeptide CNBr-3 (Residues 48 to 71) Leu-Val-Glu-Lys-Asn(CH0)-Val-Thr-Ser-Glu-Ser-Thr-Cys(Ac)-Cys(Ac)-Val-Ala-Lys-Ser-Tyr-Asn-Arg-Val-Thr-Ser at His—The sequence of amino acids in glycopeptide CNBr-3 was determined by sequencing the tryptic peptides obtained from it. An aliquot of 0.5 μmole of glycopeptide CNBr-3 was hydrolyzed with trypsin (75 μg) in 0.6 ml of 1% NH4HCO3 at 37°C for 1 hour. The initial fractionation of the six peptides thus obtained was carried out by Sephadex G-25 column chromatography as shown in Fig. 8a (see miniprint supplement p. 2). Fraction 1 contained the glycopeptide CNBr-3, T-2 which was further purified by high voltage electrophoresis at pH 4.7 and yielded three fractions due to incomplete hydrolysis at lysine (residue 51) and S-aminoethylcysteine (residue 59). Fraction 3 contained peptides CNBr-3, T-1, CNBr-3, T-3 (stained gray-brown with ninhydrin), CNBr-3, T-4, and CNBr-3, T-6, which were separated by high voltage electrophoresis at pH 1.9. Peptide CNBr-3, T-5, which stained brown with ninhydrin, was present in Fraction 5 and was subjected to high voltage electrophoresis at pH 1.9 to remove serine contamination. The compositions of the six tryptic peptides obtained from CNBr-3 are given in Table X (see miniprint supplement p. 9), and their sequences are given in Table XI (see miniprint supplement p. 10). The glutamine and asparagine residues at positions 50 and 66, respectively, were determined by hydrolysis of peptides CNBr-3, T-1 and CNBr-3, T-5 with leucine aminopeptidase. Edman degradation with dansylation of intact CNBr-4 established the NH2-terminal sequence Leu-Val-Glu-Lys-Asn(CH0)-Val-Thr, which established the positioning of peptides CNBr-3, T-1 and CNBr-3, T-2 as shown in Table XI. The position of peptide CNBr-3, T-3 was established with the overlapping tryptic peptide aT-8. Peptide CNBr-3, T-6 was placed at the COOH terminus because it contained homoserine. The only remaining peptide, CNBr-3, T-5, was positioned by difference, and its amino acid composition together with the compositions of the other tryptic peptides obtained from CNBr-3 account for all of the residues of peptide CNBr-3.

Glycopeptide CNBr-4 (Residues 72 to 92) Gly-Gly-Phe-Lys-Val-Glu-Lys-Thr-Ala-Cys(Ac)-His-Cys(Ac)-Ser-Thr-Cys(Ac)-Tyr-Tyr-His-Lys-Ser—The sequence of glycopeptide CNBr-4 was derived from the sequences of its chymotryptic peptides. A 0.5-μmole aliquot of CNBr-4 was hydrolyzed with chymotrypsin (100 μg) in 0.6 ml of 1% NH4HCO3 for 1 hour. The hydrolysate was fractionated by Sephadex G-25 column chromatography (Fig. 8b; see miniprint supplement p. 2). Glycopeptide CNBr-4, C-2 was in Fraction 2 and was further purified by high voltage electrophoresis at pH 4.7. Fraction 5 contained peptides CNBr-4, C-1, CNBr-4, C-3, and CNBr-4, C-4b which were separated by high voltage electrophoresis at pH 4.7. Peptide CNBr-4, C-1, which stained yellow with ninhydrin, was also present in Fraction 6 with peptide CNBr-4, C-4. These were separated by high voltage electrophoresis at pH 1.9. Fraction 7 contained peptide CNBr-4, C-4a, which was free tyrosine and was not purified further. The compositions of the six peptides obtained from chymotryptic hydrolysis of peptide CNBr-4 are given in Table XII (see miniprint supplement p. 11), and the sequences of the chymotryptic peptides of CNBr-4 are given in Table XIII (see miniprint supplement p. 12).

Edman degradation with dansylation of intact CNBr-4 established the following NH2-terminal sequence of the peptide: Gly-Gly-Phe-Lys-Val-Chx-Asn(CH0). This information, together with the overlapping thermolysin peptides obtained from the tryptic peptide aT-11, was sufficient to establish the sequence of peptide CNBr-4 as shown in Table XIII.

**DISCUSSION**

The linear amino acid sequence of hCG-α is shown in Fig. 9. The molecular weight of hCG-α is approximately 14,900, as computed from its chemical composition, 10,200 for the protein and 4,700 for the carbohydrate portion of the molecule. The total carbohydrate of hCG-α is contained in 2 bulky units attached by N-glycosidic bonds to asparagine residues 52 and 78 (Fig. 9). As shown for other glycoproteins (29), hCG-α has the invariant Asn-X-Ser/Thr sequence at the carbohydrate attachment sites. The total number of amino acids in hCG-α varies from 89 to 92 because of NH2-terminal heterogeneity discussed below. There are five disulfide bonds in hCG-α and there is no evidence for the presence of free sulfhydryl groups or tryptophan. The assignment of the disulfide bonds as well as some of the amide groups still remains to be completed. Attempts to determine the amide groups by leucine aminopeptidase, aminopeptidase M, or Pronase hydrolysis were not successful in the NH2-terminal region of the molecule because of the presence of clusters of adjacent acidic residues and several proline residues. Proline residues have been reported to form diketopiperazines (30) which might have resulted in incomplete hydrolysis. The Edman degradation procedure with direct identification of the phenylthiohydantoin derivatives should facilitate the determination of such amide groups. However, possible deamidation in the NH2-terminal region might still be a problem.

Several problems were encountered during the isolation and sequencing of the tryptic and cyanogen bromide peptides of hCG-α. Tryptic hydrolysis of reduced, S-carboxamidomethylated, desialylated hCG-α resulted in a large number of nonspecific partial cleavages in addition to the expected hydrolysis at arginyl and lysyl residues. Partial cleavages occurred at all 3 methionyl residues (positions 29, 47, and 71), at 3 of the 4 tyrosyl residues (positions 65, 88, and 89), and at S-carboxamidomethylesteryl residues 28 and 31. It appears that the trypsin preparation used in the present studies had some chymotryptic-like activity. The resulting chymotryptic-like cleavages yielded a multitude of minor peptides which complicated the fractionation of the tryptic hydrolysate.

Treatment of S-aminoethyl hCG-α with cyanogen bromide (50- to 200-fold excess over methionine residues) resulted in approximately 80% conversion of methionine to homoserine. The reasons for the incomplete reaction with cyanogen bromide are not clear, although Sairam et al. (13) have reported similar results with human LH-α. The two cyanogen bromide glycopeptides CNBr-3 and CNBr-4 proved difficult to separate by paper chromatography or electrophoresis. An adequate separation of these peptides in high yield was effected by Sephadex G-75 gel filtration as shown in Fig. 6.

Internally located S-aminoethylcysteine was consistently difficult to identify presumably due to destruction of its dansyl derivative during acid hydrolysis, as postulated by Gray (31). However, no problems were encountered in the identification of the dansyl S-aminoethylcysteine from the NH2-terminal S-aminoethylcysteiny1 residue in peptides CNBr-3, T-3 and CNBr-4, C-3.

The NH2 terminus of hCG-α exhibits heterogeneity. At least
Fig. 9. Linear amino acid sequence of hCG-α. The residues positioned by sequence methods are given with solid lines (---) in the arrows. The residues determined by composition and peptide α-T3 (not isolated, see text) are shown by broken lines (---) in the arrows.

Fig. 10. Comparison of the amino acid sequences of human and porcine LH-α and bovine TSH-α with hCG-α. Only residues not identical with hCG-α are shown. Bovine and ovine LH-α have identical amino acid sequences with TSH-α. The portions of the pLH-α sequence enclosed in parentheses were determined by composition only (17).
three populations of molecules were encountered in the present investigation, the molecules beginning with alanine, aspartic acid or asparagine, and valine. The source of this heterogeneity might be proteolysis of hCG during renal passage or bacterial proteases in the urine. It is improbable that these multiple forms of hCG-α are the products of independent genes. Both NH₂- and COOH-terminal heterogeneity have previously been noted in the α and β subunits of LH (14, 32) and TSH (11). This heterogeneity would suggest that the NH₂-terminal and COOH-terminal residues are not required for the biological function of these molecules.

Recently, an interesting picture of the interrelationships among various glycoprotein hormones has emerged. It has been shown that TSH (11), hCG (2, 3), LH (33, 34), and FSH (35) are all composed of two noncovalently bonded subunits designated α and β. The α subunits of hCG, LH, and TSH are interchangeable and the resulting recombinant hormones are indistinguishable from the native hormones in their biological, immunological, and electrophoretic properties (7). It is evident that the β subunits determine the specificity of the hormones. It is tempting to speculate that the α subunits might be responsible for the hormonal action; this would imply that functionally different hormones such as gonadotropins and TSH which share a common α subunit may have a similar mechanism of action. The interchangeability of the α subunits would also suggest that they must have a high degree of structural similarity. Indeed, this has been found to be the case. The α subunits of hCG, LH, and TSH show considerable homology (Fig. 10). Human LH-α, although the data are preliminary (13), differs from hCG-α only by a 2-residue inversion and a deletion of three amino acids at the NH₂ terminus of LH-α. Bovine TSH-α and ovine and bovine LH-α share identical amino acid sequences (14) and have only five amino acid substitutions when compared with porcine LH-α (17). Ovine, bovine, and porcine LH-α and bovine TSH-α have approximately 25 amino acid residue differences compared with human LH-α or hCG-α. About 80% of the differences occur near the NH₂ and COOH termini and only 20% in the interior of the molecule. Furthermore, the changes near the NH₂ terminus are more “radical”; hCG-α has acidic or neutral amino acids at positions 11, 13, 16, and 20 in place of lysyl residues at corresponding positions in ovine, bovine, or porcine LH-α. The amino acid differences of the α subunits of the various glycoprotein hormones are summarized in Table XIV.

Whereas there is a great deal of homology in the polypeptide chain of hCG-α and the α subunits of TSH and LH, the carbohydrate moieties appear to be rather different as reflected by their carbohydrate compositions (36). The α subunit of hCG contains α-galactose and sulfatic acid, whereas the α subunits of the other hormones contain little or none of these monosaccharides. Unlike the other α subunits, hCG-α lacks N-acetylgalactosamine. This amino sugar has been found to occur in nature at the reducing end of oligosaccharide chains in O-glycosidic linkage to the β-hydroxyl group of serine in mucins (37), at the nonreducing terminus in blood group substances (38), or in the interior of oligosaccharide chains in gangliosides (39), but generally not in any asparagine-linked carbohydrate units. The very presence of N-acetylgalactosamine in asparagine-linked carbohydrate units of LH and TSH would suggest a carbohydrate structure for these α subunits different from that of hCG-α. The α subunits of all of the hormones contain two asparagine-linked carbohydrate units at the same positions as shown in Fig. 9. Based on sequential degradation by specific glycosidases, a tentative structure (Structure II) of the carbohydrate units in hCG-α has been proposed earlier (3). Further details on the structure of the carbohydrate will be described in a subsequent communication. It is interesting to note that the α subunits are interchangeable despite differences in the carbohydrate units.

Acknowledgment—We wish to thank Mrs. Marguerite Warg for excellent technical help in amino acid analyses.

REFERENCES

1. Bahl, O. P. (1969) J. Biol. Chem. 244, 567-571
2. Bell, J. J., Canfield, R. E., and Sciarra, J. J. (1969) Endocrinology 84, 298-307
3. Bahl, O. P. (1969) J. Biol. Chem. 244, 575-583
4. Canfield, R. E., Agosto, G. M., and Bahl, J. J. (1970) in Gonadotropins and Ovarian Development (Butt, W. R., Crooke, A. C., and Hyle, M., eds) pp. 191-198, Livingston, London
5. Swaminathan, N., and Bahl, O. P. (1970) Biochem. Biophys. Res. Commun. 40, 422-427
6. Swaminathan, N., and Bahl, O. P. (1970) Fed. Proc. 29, 599
7. Pierre, J. G., Bahl, O. P., Cornnell, J. S., and Swaminathan, N. (1971) J. Biol. Chem. 246, 2321-2324
8. Morgan, F. J., and Canfield, R. E. (1971) Endocrinology 88, 1045-1054
9. Bahl, O. P. (1972) in Hormonal Proteins and Peptides (Li, C. H., ed) Academic Press, New York, in press
10. Bahl, O. P., Carlsson, R. B., Bellasario, R., and Swaminathan, N. (1972) Biochem. Biophys. Res. Commun. 48, 416-422
11. Liou, T.-H., and Pierce, J. G. (1971) J. Biol. Chem. 246, 806-815
12. Ingam, T., Murakami, K., Puett, D., Hartree, A. S., and Nureddin, A. (1972) Biochem. J. 126, 441-442
13. Sairam, M. R., Papkoff, H., and Li, C. H. (1972) Biochem. Biophys. Res. Commun. 45, 530-537
14. Liou, T.-K., Naum, H. S., Sweeney, C. M., Lamkin, W. M., Baker, H. N., and Ward, D. N. (1972) J. Biol. Chem. 247, 4351-4364
15. Pierce, J. G., Liou, T.-H., Carlsson, R. B., and Reim, T. (1971) J. Biol. Chem. 246, 566-572
16. Papkoff, H., Sairam, M. R., and Li, C. H. (1971) J. Amer. Chem. Soc. 93, 1531-1532
17. Maghuin-Rogister, G., Combaroux, Y., and Hennan, G. (1972) F. E. B. S. Lett. 26, 57-60
18. Carlsson, R. B., Bahl, O. P., and Swaminathan, N. (1973) J. Biol. Chem. 248, 6810-6827
19. Bahl, O. P., and Smith, E. L. (1965) J. Biol. Chem. 240, 3855-3863
20. Goffeney, P. F., and Pierce, J. G. (1966) Arch. Biochem. Biophys. 115, 277-290
21. Howard, S. M., and Pierce, J. G. (1968) J. Biol. Chem. 243, 6468-6476
22. Liu, T.-Y., and Chang, Y. H. (1971) J. Biol. Chem. 246, 2842-2848
23. Hofmann, T. (1964) Biochemistry 3, 356-364
24. Gray, W. R. (1967) Methods Enzymol. 11, 139-151, 469-475
25. Woods, K. R., and Wang, K.-T. (1967) Biochim. Biophys. Acta 133, 369-370
26. Raftery, M. A., and Cole, R. D. (1966) J. Biol. Chem. 241, 3457-3461
27. McMenemy, R. H., Dintzis, H. M., and Watson, F. (1971) J. Biol. Chem. 246, 4744-4750
28. Moore, S., and Stein, W. H. (1954) J. Biol. Chem. 211, 907-913
29. Spiro, R. G. (1970) Ann. Rev. Biochem. 39, 599-638
30. Light, A. (1972) Methods Enzymol. 25, 253-262
31. Gray, W. R. (1972) Methods Enzymol. 25, 191-138
32. Maguire-Rogister, G., and Hennen, G. (1972) P. E. B. S. Lett. 23, 225-229
33. Papkoff, H., Jr., and Ananth, T. S. (1967) Biochim. Biophys. Acta 147, 175-177
34. Reichert, L. E., Rasco, M. A., Ward, D. N., Niswender, G. D., and Midgley, A. R. (1969) J. Biol. Chem. 244, 510-5117
35. Papkoff, H., and Ekblad, M. (1970) Biochem. Biophys. Res. Commun. 40, 614-621
36. Bahl, O. P. (1972) in Gonadotropins (Saxena, B. B., Beling, C. G., and Gandy, H. M., eds) pp. 200-210, John Wiley, New York.
37. Gottschalk, A., Bhargava, A. S., and Murty, V. L. N. (1972) in Glycoproteins (Gottschalk, A., ed) 2nd Ed, Vol. 5, Part B, pp. 810-829, Elsevier, Amsterdam
38. Watkins, W. M. (1972) in Glycoproteins (Gottschalk, A., ed) 2nd Ed, Vol. 5, Part B, pp. 830-891, Elsevier, Amsterdam
39. Wiegandt, H., and Bücker, H. W. (1970) Eur. J. Biochem. 15, 287-292
**Table IV**

Sequence of the chemokine peptides obtained from the tryptic digests of α-4 and α-11.

| Peptide | Sequence |
|---------|----------|
| α-4     | GLY-LEU-VAL-VAL-GLY-VAL-GLY | GLY-LEU-VAL-VAL-GLY-VAL-GLY | GLY-LEU-VAL-VAL-GLY-VAL-GLY |
| α-11    | GLY-LEU-VAL-VAL-GLY-VAL-GLY | GLY-LEU-VAL-VAL-GLY-VAL-GLY | GLY-LEU-VAL-VAL-GLY-VAL-GLY |

**Table V**

Alanine acid composition of the chemokine peptides obtained from tryptic digests.

| Peptide | α-4 | α-11 |
|---------|-----|------|
| Alanine | 1.00 | 1.00 |
| Serine  | 0.00 | 0.00 |
| Threonine | 0.00 | 0.00 |

**Table VI**

Sequence and codons of the chemokine peptides from tryptic digests.

| Peptide | α-4 | α-11 |
|---------|-----|------|
| Sequence | GLY-LEU-VAL-VAL-GLY-VAL-GLY | GLY-LEU-VAL-VAL-GLY-VAL-GLY |
| Codons   | GCA  | GCA  |

**Figure**

Fig. 3. Gel filtration of the chemokine peptides on a Superose 12 column. The eluant was monitored at 280 nm using a fluorimeter. The flow rate was 0.5 ml/min, and the column temperature was kept at 25°C. The fractions were collected every 0.5 min.

**Figure**

Fig. 4. Gel filtration of the chemokine peptides on a Superose 12 column. The eluant was monitored at 280 nm using a fluorimeter. The flow rate was 0.5 ml/min, and the column temperature was kept at 25°C. The fractions were collected every 0.5 min.
### Table VI

| Amino Acid | Thr-2 | Thr-3 | Thr-4 | Thr-5 |
|------------|-------|-------|-------|-------|
| Lysine     | 0.14  | 0.11  | 0.16  | 0.12  |
| Histidine  | 0.96  | 1.02  | 1.12  | 1.22  |
| Arginine   | 0.15  | 0.20  | 0.25  | 0.28  |
| Threonine  | 0.61  | 0.71  | 0.82  | 0.91  |
| Serine     | 0.69  | 0.79  | 0.86  | 0.94  |
| Glutamic Acid | 0.40  | 0.45  | 0.47  | 0.51  |

| Amino Acid | Thr-2 | Thr-3 | Thr-4 | Thr-5 |
|------------|-------|-------|-------|-------|
| Proline    | 1.07  | 1.14  | 1.22  | 1.27  |
| Aspartic Acid | 0.86  | 0.95  | 1.05  | 1.13  |
| Alanine    | 0.76  | 0.86  | 0.95  | 1.04  |
| Glutamine  | 0.82  | 0.91  | 1.00  | 1.10  |
| Leucine    | 0.76  | 1.25  | 1.76  | 2.29  |
| Tyrosine   | 0.45  | 0.51  | 0.58  | 0.65  |
| Phenylalanine | 1.48  | 1.68  | 1.90  | 2.20  |

**Total Nucleotides**

| 2     | 3     | 4     | 5     |
|-------|-------|-------|-------|
| Total | 178   | 184   | 186   | 186   |

### Table VII

| Amino Acid | Thr-2 | Thr-3 | Thr-4 | Thr-5 |
|------------|-------|-------|-------|-------|
| Lysine     | 0.14  | 0.11  | 0.16  | 0.12  |
| Histidine  | 0.96  | 1.02  | 1.12  | 1.22  |
| Arginine   | 0.15  | 0.20  | 0.25  | 0.28  |
| Threonine  | 0.61  | 0.71  | 0.82  | 0.91  |
| Serine     | 0.69  | 0.79  | 0.86  | 0.94  |
| Glutamic Acid | 0.40  | 0.45  | 0.47  | 0.51  |

| Amino Acid | Thr-2 | Thr-3 | Thr-4 | Thr-5 |
|------------|-------|-------|-------|-------|
| Proline    | 1.07  | 1.14  | 1.22  | 1.27  |
| Aspartic Acid | 0.86  | 0.95  | 1.05  | 1.13  |
| Alanine    | 0.76  | 0.86  | 0.95  | 1.04  |
| Glutamine  | 0.82  | 0.91  | 1.00  | 1.10  |
| Leucine    | 0.76  | 1.25  | 1.76  | 2.29  |
| Tyrosine   | 0.45  | 0.51  | 0.58  | 0.65  |
| Phenylalanine | 1.48  | 1.68  | 1.90  | 2.20  |

**Total Nucleotides**

| 2     | 3     | 4     | 5     |
|-------|-------|-------|-------|
| Total | 178   | 184   | 186   | 186   |

### Table VIII

| Amino Acid | Thr-2 | Thr-3 | Thr-4 | Thr-5 |
|------------|-------|-------|-------|-------|
| Lysine     | 0.14  | 0.11  | 0.16  | 0.12  |
| Histidine  | 0.96  | 1.02  | 1.12  | 1.22  |
| Arginine   | 0.15  | 0.20  | 0.25  | 0.28  |
| Threonine  | 0.61  | 0.71  | 0.82  | 0.91  |
| Serine     | 0.69  | 0.79  | 0.86  | 0.94  |
| Glutamic Acid | 0.40  | 0.45  | 0.47  | 0.51  |

| Amino Acid | Thr-2 | Thr-3 | Thr-4 | Thr-5 |
|------------|-------|-------|-------|-------|
| Proline    | 1.07  | 1.14  | 1.22  | 1.27  |
| Aspartic Acid | 0.86  | 0.95  | 1.05  | 1.13  |
| Alanine    | 0.76  | 0.86  | 0.95  | 1.04  |
| Glutamine  | 0.82  | 0.91  | 1.00  | 1.10  |
| Leucine    | 0.76  | 1.25  | 1.76  | 2.29  |
| Tyrosine   | 0.45  | 0.51  | 0.58  | 0.65  |
| Phenylalanine | 1.48  | 1.68  | 1.90  | 2.20  |

**Total Nucleotides**

| 2     | 3     | 4     | 5     |
|-------|-------|-------|-------|
| Total | 178   | 184   | 186   | 186   |

---

[1] Results presented in sequence studies are given with bold type. Those in the sequence studies are consistent with results from other sources.

[2] *A* very faint d-glycine derivative was detected.

[3] Essential amino acids are given with bold type. Those in the sequence studies are consistent with results from other sources.

[4] Percentage of sequence studies are given with bold type. Those in the sequence studies are consistent with results from other sources.

---

[5] Essential amino acids are given with bold type. Those in the sequence studies are consistent with results from other sources.

[6] *A* very faint d-glycine derivative was detected.

---

[7] Essential amino acids are given with bold type. Those in the sequence studies are consistent with results from other sources.

[8] Percentage of sequence studies are given with bold type. Those in the sequence studies are consistent with results from other sources.

---

[9] *A* very faint d-glycine derivative was detected.
Human Chorionic Gonadotropin: LINEAR AMINO ACID SEQUENCE OF THE α SUBUNIT
Ronald Bellisario, Robert B. Carlsen and Om P. Bahl

*J. Biol. Chem.* 1973, 248:6796-6809.

Access the most updated version of this article at [http://www.jbc.org/content/248/19/6796](http://www.jbc.org/content/248/19/6796)

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

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

This article cites 0 references, 0 of which can be accessed free at [http://www.jbc.org/content/248/19/6796.full.html#ref-list-1](http://www.jbc.org/content/248/19/6796.full.html#ref-list-1)