Biochem. J. (1981) **197**, 541–552
Printed in Great Britain

Primary structure of the ovine pituitary follitropin β-subunit

M. R. SAIRAM,* Nabil G. SEIDAH† and Michel CHRÉTIEN†
**Reproduction Research Laboratory and †the Protein and Pituitary Hormone Laboratory, Clinical Research Institute of Montreal, 110 Pine Avenue West, Montreal, Quebec H2W 1R7, Canada*

(*Received 6 January 1981/Accepted 22 April 1981*)

The complete amino acid sequence of the ovine pituitary follitropin β-subunit was established by studying the tryptic, chymotryptic and thermolytic peptides. The *N*-terminal sequence of the subunit was confirmed by subjecting the oxidized protein to Edman degradation in an automated sequenator. Automated Edman degradation of the reduced and alkylated (with iodo[^{14}C]acetamide) β-subunit indicated that most of the molecules used in the sequence studies had lost the *N*-terminal serine residue. This also confirmed the location of the first five half-cystine residues in the sequence. The proposed structure shows the presence of 111 amino acid residues with the two oligosaccharide moieties linked to asparagine residues located at positions 6 and 23. Heterogeneity occurs at both the termini of the polypeptide chain. Comparison of the sequence of β-subunit of the ovine hormone with that proposed for human follitropin β-subunit shows the absence of any deletions in the middle of the peptide chain. Of the 13 replacements, 11 residues can be explained on the basis of a single base change in the codon. The single tryptophan residue of the follitropin occupies an identical position in all the four species that have been studied. The region corresponding to residues 63–105 of the ovine β-subunit is highly conserved in all the species.

Among the glycoprotein hormones, all of which consist of two non-identical subunits (α and β), the β-subunit imparts specificity to the α–β complex. The comparison of amino acid sequences of the β-subunit of the three pituitary glycoproteins within a species offers valuable clues regarding structural features that may be necessary for hormone specificity, i.e. hormone–receptor interactions (Ward, 1978; Sairam, 1978). In the ovine species, lutropin (luteinizing hormone, 'LH') has been extensively studied with respect to its structure and function (Papkoff *et al.*, 1972; Ward *et al.*, 1972). However, relatively little information has been forthcoming on ovine follitropin (follicle-stimulating hormone, 'FSH') in similar respects. Papkoff (1972) reported the compositions of several peptides that he thought originated from the ovine follitropin subunits and compared them with the available data on lutropin and thyrotropin (thyroid-stimulating hormone, 'TSH'). More recently, investigations initiated in this laboratory sought to determine the primary structure of ovine follitropin. An improved procedure for the isolation of highly potent ovine hormone (Sairam, 1979*a*) and the preparation of subunits has been reported (Sairam, 1979*b*). In the preceding paper (Sairam, 1981), evidence for the identity of the ovine follitropin α- and lutropin α-subunits was presented. The present study provides data that permit us to propose the complete amino acid sequence of the ovine follitropin β-subunit. A comparison of these data with proposed sequences of follitropin β-subunit from other species shows remarkable degree of preservation of structural information.

Materials and methods

Preparation of β-subunit

The β-subunit was prepared from highly purified follitropin isolated from frozen ovine pituitaries (Sairam, 1979*b*). It effectively recombined with the α-subunit to regenerate biological and immunological activities.

Sequence determination

The strategy and methods employed for the determination of the amino acid sequence are described in the preceding paper (Sairam, 1981). A 55mg portion of reduced and aminoethylated β-subunit was used for trypsin digestion, whereas 6mg of the same preparation was utilized for obtaining the chymotryptic peptides. Additional

0306-3275/81/090541-12$01.50/1

methods employed for solving the structure of the β-subunit are given below.

Thermolysin digestion of oxidized β-subunit

A 6mg portion of performic-acid-oxidized β-subunit was digested with thermolysin (EC 3.4.24.4, Calbiochem–Boehringer, lot no. 901254, thrice-crystallized) in 1ml of 0.25M-NH_4HCO_3 at 37°C for 4h (50:1 ratio). The freeze-dried material was dissolved in 2M-acetic acid and subjected to two-dimensional paper chromatography–electrophoresis as described previously (Sairam, 1981). Peptides were detected after ninhydrin spray and eluted with NH_3 for further analysis.

Hydrazinolysis

Two different samples of performic-acid oxidized β-subunit were dried thoroughly over P_2O_5 for 48h. They were treated with 200μl of anhydrous hydrazine (taken from a newly opened bottle; Pierce Chemicals) in a sealed tube. Hydrazinolysis was performed for 16 and 24h at 80°C. The contents were then freeze-dried and the free amino acids were directly determined on the analyser.

Digestion with leucine aminopeptidase

The peptides were digested with pig kidney leucine aminopeptidase (EC 3.4.11.1, Calbiochem–Boehringer) in 0.1M-NH_4HCO_3 for 2–6h at 37°C. The contents were freeze-dried and the liberated amino acids were quantified on the analyser.

N-Terminal sequence by the automated sequenator

The *N*-terminal sequence of the reduced and *S*-aminoethylated β-subunit was determined by the dansyl (5-dimethylaminonaphthalene-1-sulphonyl)–Edman method (Sairam, 1981) as well as by using a Beckman automated sequenator. The entire reaction was carried out in the spinning cup by using 0.3M-Quadrol with Polybrene added as the carrier. The amino acid phenylthiohydantoin derivatives were identified by high-pressure liquid chromatography as described elsewhere (Seidah *et al.*, 1980).

Results

Terminal amino acids and sequence

In a previous study the *N*-terminal residue of the follitropin β-subunit was tentatively identified as aspartic acid (Sairam, 1979*b*). It was, however, noted that this result was not unequivocal. A re-examination of this with the preparations of the β-subunit employed for the current investigations for determination of sequence, again did not yield conclusive results, but traces of serine and cysteic acid were detectable by the dansyl chloride reaction.

By the dansyl–Edman technique the *N*-terminal sequence of $Cys(O_3H)$-Glx-Leu-Thr-Asx-[$Cys(O_3H)$ = cysteic acid] was identified for the oxidized subunit. A reduced and *S*-aminoethylated preparation on the automated sequenator showed the following sequence:

Xaa-Glu-Leu-Thr-Xaa-Ile-Thr-Ile-Xaa-Val-

An additional run with 3mg of the oxidized protein (0.2μM) provided further data on the sequence that extended up to residue 22:

Xaa-Glu-Leu-Thr-Xaa-Ile-Thr-Ile-Thr-Val-Glu-Xaa-Glu-Glu-Xaa-Xaa-Phe-Xaa-Ile-Xaa-Ile-.

In both experiments the residue designated Xaa at some positions could not be identified. The location of cysteine residues at some of these positions was confirmed by microsequence of reduced and iodo-[^{14}C]acetamide-treated protein (see Fig. 5 below). When the *C*-terminus was investigated by hydrazinolysis of the oxidized β-subunit, 0.44 residue of glutamic acid/molecule was found.

Sequence analysis of the β-subunit

The amino acid composition of the oxidized β-subunit after hydrolysis with 6M-HCl for 24–72h is shown in Table 1. The subunit entirely lacks methionine, thus the strategy of fragmentation of the protein with CNBr was not possible. Instead of this, information was sought from tryptic, chymotryptic and thermolytic peptides of the reduced and alkylated or oxidized β-subunit.

Tryptic peptides of aminoethylated β-subunit

It was considered desirable to perform tryptic digestion on the reduced and *S*-aminoethylated derivative of the β-subunit as potentially 12 sites of cleavage would be induced in addition to those occurring at the lysine and arginine residues present in the molecule.

Digestion of 55mg of the reduced and alkylated β-subunit with trypsin and fractionation on Sephadex G-50 columns yielded 11 fractions (Fig. 1). Analysis of the first fraction revealed that it contained some undigested subunit and it was not further investigated. Each of the other fractions was submitted to two-dimensional paper chromatography–electrophoresis for separating the constituent peptides. The scheme (Fig. 2) shows the origin of the different peptides isolated. A composite peptide map, constructed in Fig. 3, shows the relative migration of the different peptides. Thus a total of 27 peptides, three of which contained carbohydrate, were obtained. The presence of carbohydrate in the peptide was suggested by its mobility very close to the origin and confirmed by subsequent amino acid analysis. All were analysed for composition, homogeneity and subjected to sequence determination. Their compositions, along with end-group data, are given in Table 2. Fortu-

Table 1. *Amino acid composition of ovine follitropin β-subunit*

Results for the present study are means ± S.E.M. for duplicate determinations on three preparations of the subunit after hydrolysis in 6M-HCl for 24, 48 and 72h. Threonine, serine and tyrosine values have been corrected for destruction.

	Composition (residues/molecule of 15000 daltons)		
Amino acid	Previous study (Sairam, 1979*b*)	The present study	Sequence
Trp	1	1	1
Lys	6.2	6.2 ± 0.06	6
His	3	2.9 ± 0.07	3
Arg	5	4.79 ± 0.16	4
Asp	10.3	9.8 ± 0.23	10
Thr	11.0	11.0 ± 0.23	13
Ser	6.9	7.07 ± 0.1	9
Glu	9.9	10.1 ± 0.19	10
Pro	5.8	5.29 ± 0.26	5
Gly	5.5	5.29 ± 0.13	5
Ala	6.4	5.87 ± 0.32	5
½-Cys*	11.0	10.8 ± 0.4	12
Val	6.7	6.56 ± 0.1	7
Ile	5.3	4.95 ± 0.22	6
Leu	5.2	5.42 ± 0.11	5
Tyr	8.9	6.73 ± 1.08	7
Phe	3.1	3.03 ± 0.03	3

* Determined as cysteic acid on the performic-acid-oxidized protein.

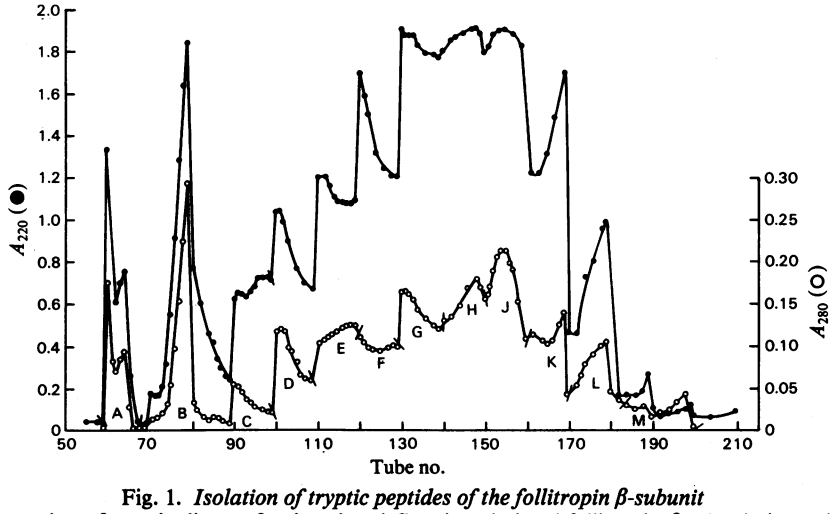

Fig. 1. *Isolation of tryptic peptides of the follitropin β-subunit*

Preliminary separation of tryptic digest of reduced and *S*-aminoethylated follitropin β-subunit (55mg) was achieved by chromatography on a column (2.5cm × 92cm) of Sephadex G-50 (Superfine grade) in 0.05M-NH_4HCO_3. The fraction size was 2.8ml/tube and the flow rate was 25ml/h. The different fractions were pooled as shown and freeze-dried. Fraction A contained some undigested subunit and fraction M consisted of salt. These were not further investigated.

nately all the peptides except T16 were obtained in a relatively pure form, including the two glycopeptides T2 and T5. The yield of the various peptides from 3.5μmol of the β-subunit varied from 0.06μmol for T10 to 1.1μmol for T3. Glycopeptide T3 is related to the glycopeptide T2, which has three fewer amino acids. Both of these were found in the B-fraction (Fig. 1). End-group determination on this

Table 2. *Tryptic peptides of S-aminoethylated β-subunit*

Peptide T22 is free *S*-aminoethylcysteine (Aec). The presence of carbohydrate (CHO) residues in peptides T2, T3 and T5 were not quantified. Yields of some peptides are the minimal calculated values found in the major fraction (see Figs. 1 and 2).

Amino acid	Peptide ... T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12
Lys		1.1 (1)	1.0 (1)								1.5 (2)	
His												
Arg									1.0 (1)	1.0 (1)	0.6 (1)	
Aec	0.8 (1)		0.7 (1)	1.1 (1)	0.8 (1)		0.9 (1)	1.0 (1)	0.8 (1)			1.0 (1)
Asp		1.0 (1)	1.0 (1)		1.0 (1)					1.0 (1)	2.4 (3)	
Thr		2.3 (3)	2.5 (3)		1.7 (2)			0.8 (1)	0.8 (1)		0.1	1.0 (1)
Ser	1.0 (1)	0.3		0.9 (1)	0.9 (1)						0.1	
Glu		2.2 (2)	3.4 (4)		0.4						1.2 (1)	
Pro											1.8 (2)	
Gly		0.2	0.38		0.2	1.0 (1)	1.0 (1)					
Ala		0.1			0.2	0.8 (1)	0.9 (1)				1.0 (1)	
Val		0.8 (1)	1.0 (1)							1.0 (1)	1.0 (1)	
Ile		2.2 (2)	2.1 (2)		2.0 (2)						0.8 (1)	
Leu		0.9 (1)	1.0 (1)		0.2					1.2 (1)	0.9 (1)	
Tyr						0.6 (1)	0.6 (1)	0.8 (1)	0.7 (1)	0.6 (1)	0.6 (1)	
Phe				1.0 (1)								
Trp†					(1)							
CHO		+	+		+							
N-Terminus	Ser	Glu	Glu	Ser	Ile	Ala	Ala	Tyr	(Aec)	Asp	Asp	Thr
C-Terminus	Aec	Lys	Aec	Aec	Aec	Tyr	Aec	Arg	Arg	Lys	Lys	Aec
No of residues	2	11	14	3	8	3	4	3	4	5	14	2
Yield (μmol)	0.1	0.4	1.1	0.25	0.8	0.3	0.3	0.5	0.1	0.06	0.5	0.4

* After further purification in the butanol/pyridine/acetic acid/water ('BPAW') system, this peptide was still not homogeneous.
† Value obtained by hydrolysis in presence of mercaptoacetic acid.

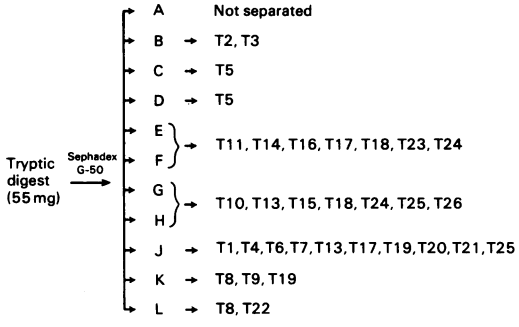

Fig. 2. *Scheme of fractionation of tryptic peptide fractions*

After preliminary separation on Sephadex G-50 (from Fig. 1), the different fractions were directly subjected to chromatography [butanol/acetic acid/water ('BAW') system, 18h] and electrophoresis (2kV, 1h, pH2.0). For peptide numbers, see Fig. 4. Some peptide spots were also found in the adjacent fractions. Each fraction was separated on two to four Whatman 3MM papers.

fraction revealed glutamic acid at the *N*-terminus, but the composition indicated some heterogeneity. High-voltage paper electrophoresis at pH2.0 resolved this fraction into two peptides, T2 and T3, both of which were similar, except for three additional residues in T3 (Table 2 and Fig. 4). In the composite peptide map the three glycopeptides are indicated as a large broad spot, as they normally tend to smear.

Chymotryptic peptides of aminoethylated β-subunit

Gel filtration of 6mg of chymotryptic digest on a Sephadex G-50 column (1.5cm × 110cm, 1.25ml/fraction) resulted in five fractions, four of which were further separated by paper chromatography–electrophoresis. The fractionation of the peptides in this digest was complicated by the unexpected occurrence of tryptic cleavages by the chymotrypsin preparation employed. Despite this, twelve useful peptides were isolated in a form amenable to further analysis (Table 3). End-group determination was performed on these peptides.

Thermolysin peptides

To obtain additional data that would be useful in ordering the tryptic peptides, thermolysin digestion of oxidized β-subunit was performed. In this instance the entire digest was used for two-dimensional paper chromatography and electrophoresis without preliminary separation on Sephadex G-50. Among the peptides separated, the composition and end-group analysis of seven peptides are given in Table 4.

Linear amino acid sequence

All the tryptic peptides isolated were subjected to dansyl–Edman degradation. The results obtained are

of follitropin

ascertained during analysis for composition by the amino acid analyser. The hexosamines liberated under the conditions were

Composition (residues/molecule)

T13	T14	T15	T16*	T17	T18	T19	T20	T21	T22a	T23	T24	T25	T26
1.2 (1)	1.1 (1)		0.7 (1)				1.0 (1)	1.0 (1)	1.0 (1)				
		1.6 (2)		1.6 (2)		1.0 (1)	0.2	1.0 (1)			0.15		
											0.94 (1)		0.8 (1)
		0.9 (1)	0.7 (1)		0.92 (1)	1.0 (1)		1.0 (1)	1.0 (1)	1.1 (1)	0.95 (1)	0.8 (1)	
		1.3 (1)	0.6 –	1.0 (1)						2.5 (3)	2.4 (3)	0.1	1.1 (1)
0.7 (1)	0.9 (1)	0.2	0.8 (1)		1.6 (2)					1.2 (1)	1.7 (2)		
	0.1	0.8 (1)	0.8 –	1.0 (1)						1.5 (2)	1.7 (2)	1.0 (1)	1.5 (2)
	1.7 (2)	0.2	1.6 (2)		1.2 (1)								1.2 (1)
			1.0 (1)		1.0 (1)							1.0 (1)	
0.2	0.15	0.7 (1)	1.2 (1)				1.0 (1)	0.9 (1)	0.8 (1)		0.3	1.6 (2)	
	0.1	0.9 (1)		1.5 (2)	1.2 (1)								
	2.0 (2)	1.7 (2)	2.2 (3)		1.0 (1)						1.0 (1)		
		0.7 (1)											1.0 (1)
	1.0 (1)	1.0 (1)	1.1 (1)	1.0 (1)								1.0 (1)	
	0.6 (1)	0.75 (1)	0.72 (1)	0.5 (1)	0.8 (1)							0.75 (1)	
1.0 (1)													1.0 (1)
Thr	Glu	Val	Glu	Ala	Thr	His	Gly	Gly	Gly	Asp	Asp	Gly	Ser
Lys	Lys	Tyr	Aec	Tyr	Aec	Aec	Lys	Lys	Aec	Aec	Arg	Aec	Arg
3	8	12	12	8	8	2	2	4	3	7	10	7	7
0.3	1	0.4	0.4	0.3	0.7	0.3	0.5	0.3	0.4	0.2	0.9	1	1

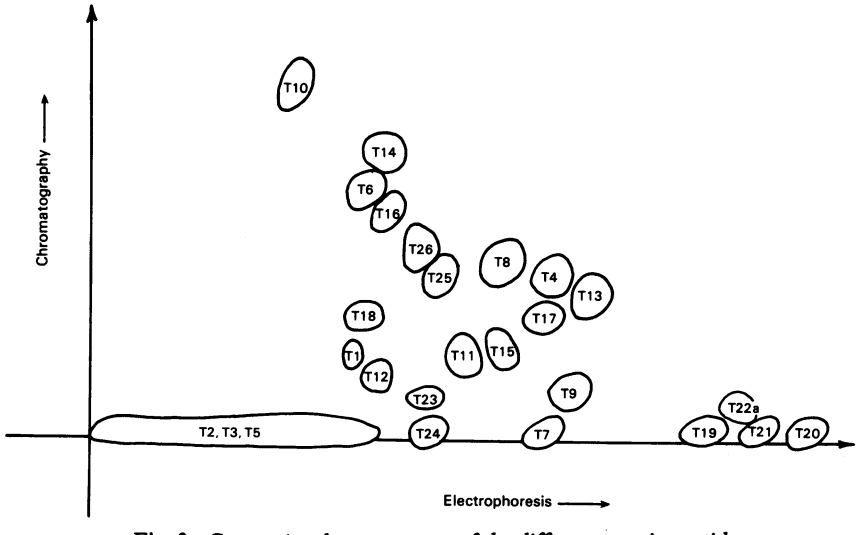

Fig. 3. *Composite chromatogram of the different tryptic peptides*

The map was reconstructed from the location of various spots revealed during fractionation on paper by chromatography (butanol/acetic acid/water ('BAW') system, 18h] and high-voltage electrophoresis (2kV, 1h, pH 2.0). The peptides were revealed by spraying the dry paper with 0.1% ninhydrin in ethanol and drying in a stream of air at room temperature.

shown in Table 5. These peptides were designated T1 to T26 in order of appearance in the proposed sequence (Fig. 4). Data obtained with chymotryptic and thermolytic peptides provided the overlap for all tryptic peptides except in two regions near the *C*-terminus, between T22 and T23 and between T24 and T25. These peptides were placed in sequence on the basis of homology with the reported sequences

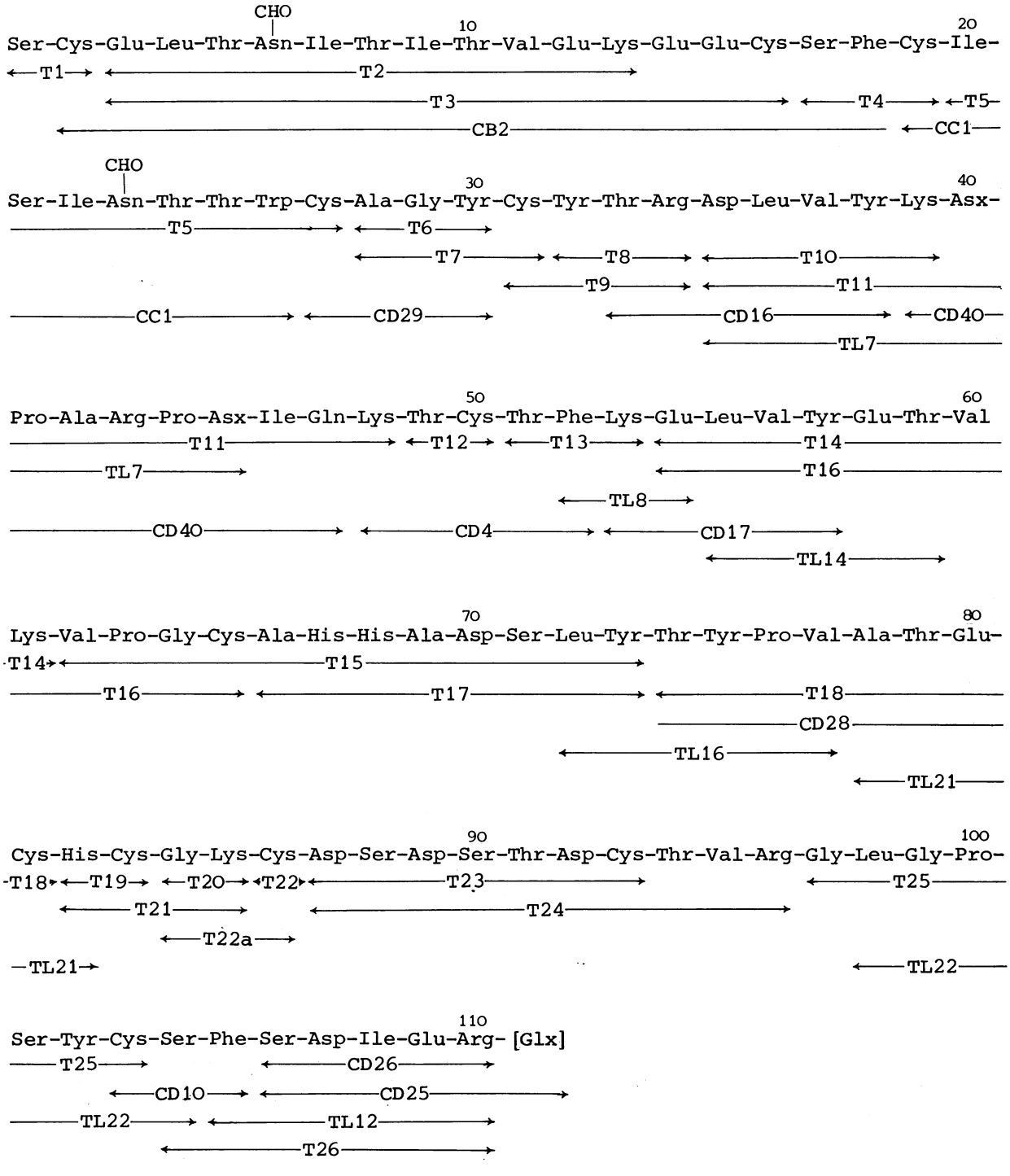

Fig. 4. *Amino acid sequence of ovine follitropin β-subunit*

T, tryptic peptides; C, chymotryptic peptides; TL, thermolysin peptides. In the tryptic and chymotryptic peptides, cysteine residues were identified as aminoethylcysteine; CHO represents an oligosaccharide moiety. The placement of peptide T24 in that position is based on homology with other known sequences (see Fig. 6).

Table 3. *Amino acid composition of chymotryptic peptides of reduced and aminoethylated follitropin β-subunit*
The differences in mobility of peptides CD25 and CD26 and the indicated composition appears to suggest the presence of an extra glutamic acid/glutamine residue in the latter. This additional glutamic acid/glutamine residue has been placed at the *C*-terminus of peptide CD25 and the β-subunit. The dansylated residue in parentheses was not identified. Abbreviation used: Aec, *S*-aminoethylcysteine.

Peptide no.	Composition	Total residues	End groups	
			N-Terminus	*C*-Terminus
CB2*	$Lys_1Aec_{1.8}Asp_{1.2}Thr_{2.6}Ser_{0.9}Glu_4Val_{1.0}Ile_{1.7}Leu_{0.9}Phe_{0.8}$	17	(Aec)	Phe
CC1*	$Aec_{1.0}Asp_{1.3}Thr_{1.8}Ser_{1.0}Ile_{2.0}Trp_1$	8	(Aec)	Trp
CD29	$Aec_{1.0}Ala_{1.0}Gly_{1.0}Tyr_{0.7}$	4	(Aec)	Tyr
CD16	$Arg_{1.0}Asp_{1.3}Thr_{0.7}Val_{1.0}Leu_{1.0}Tyr_{0.6}$	6	Thr	Tyr
CD40	$Lys_{0.9}Arg_{1.0}Asp_{2.0}Glu_{0.9}Pro_{2.0}Ala_{0.9}Ile_{0.6}$	9	Lys	Gln
CD4	$Lys_{1.0}Aec_{1.0}Thr_{1.3}Phe_{1.0}$	5	Lys	Phe
CD17	$Lys_{0.9}Glu_{1.0}Val_{1.0}Leu_{0.9}Tyr_{0.6}$	5	Lys	Tyr
CD28	$Aec_{1.8}His_{0.8}Thr_{1.5}Glu_{1.2}Pro_{1.0}Ala_{1.0}Val_{1.0}Tyr_{0.7}$	10	Thr	Aec
CD10	$Aec_{0.8}Ser_{0.9}Phe_{1.0}$	3	(Aec)	Phe
CD25	$Arg_{1.0}Asp_{1.0}Ser_{0.6}Glu_{1.5}Ile_{1.2}$	5–6	Ser	
CD26	$Arg_{1.0}Asp_{1.2}Ser_{0.6}Glu_{1.3}Ile_{1.2}$	5	Ser	Arg

* These peptides contained carbohydrate as revealed by the amino acid analyser. The presence of tryptophan peptide in CC1 was ascertained by digestion with carboxypeptidase A followed by amino acid analysis.

Table 4. *Amino acid composition of thermolytic peptides of performic-acid-oxidized follitropin β-subunit*
Among the 25 peptide spots observed on the chromatogram and eluted for the analysis, the composition of seven are shown below. The others are omitted as they merely confirmed the composition of the tryptic and/or chymotryptic peptides. Abbreviation used: $Cys(O_3H)$, cysteic acid.

Peptide no.	Composition	Total no. of residues	Termini	
			N	*C*
TL7	$Lys_{1.0}Arg_{1.0}Asp_{2.3}Pro_{1.9}Ala_{1.0}Val_{1.0}Leu_{0.7}Tyr_{0.8}$	11	Asp	(Asx)
TL8	$Lys_{0.8}Glu_{1.0}Phe_{0.8}$	3	Phe	Glu
TL14	$Thr_{1.0}Glu_{1.0}Val_{1.2}Leu_{0.8}Tyr_{1.0}$	5	Leu	Thr
TL16	$Thr_{1.0}Pro_{1.0}Val_{1.0}Leu_{1.2}Tyr_{1.5}$	6	Leu	Val
TL21	$His_{0.8}Cys(O_3H)_{1.0}Thr_{1.0}Glu_{0.9}Ala_{0.9}$	6	Ala	His
TL22	$Cys(O_3H)_{0.7}Ser_{1.6}Pro_{1.2}Gly_{1.1}Leu_{1.0}Tyr_{0.8}$	7	Leu	Ser
TL12	$Arg_{1.0}Asp_{1.3}Ser_{1.0}Glu_{1.0}Ile_{1.0}Phe_{0.6}$	6	Phe	Arg

of β-subunit from other species (see Fig. 6 below). The proposed amino acid sequence of the ovine follitropin β-subunit is depicted in Fig. 4. In order to confirm the position of the half-cystine residues in the *N*-terminal portion of the molecule, the following experiment was done. The β-subunit was reduced with β-mercaptoethanol in 8 M-urea/0.1 M-Tris/HCl, pH 8.5, and alkylated with ^{14}C-labelled iodoacetamide (NEN Ltd., MA, U.S.A.) The labelled subunit was subjected to automated sequence analysis, and the appearance of radioactivity in the phenylthiohydantoin derivative (Seidah *et al.*, 1980) at each step of degradation was determined by liquid-scintillation counting. Highest radioactivity (Fig. 5) was found in the first cycle itself, with substantially lower, but significant, amounts at the subsequent cycles (3, 15, 18, 26 and 30). It was concluded that a half-cystine residue was located at these positions in the β-subunit. From the sequence shown in Fig. 4, it may be seen that radioactivity of half-cystine at the first cycle corresponds to the half-cystine residue located at position 2. Such a result would be possible only if we assume that the preparation analysed in this experiment was missing the first residue of serine. Appearance of radioactivity at the third cycle also indicates that at least a small number of the subunit molecules had an additional amino acid before serine. We have not identified this, and a tripeptide Xaa-Ser-Aec (Aec = aminoethylcysteine) was not isolated from the tryptic digests. These data would strongly suggest *N*-terminal heterogeneity, as found for the α-subunit of ovine lutropin (Papkoff *et al.*, 1972; Ward *et al.*, 1972) and follitropin [the preceding paper (Sairam, 1981)].

Discussion

The primary structure of the ovine follitropin β-subunit, as derived from our data, and reliance on

Table 5. *Determination of the amino acid sequence of the tryptic peptides of reduced and aminoethylated follitropin β-subunit*

Key to arrows: →, *dansyl–Edman reaction;* ⇀, *digestion with leucine aminopeptidase;* ←, digestion with carboxypeptidases A or B or both. Peptide T22 is free *S*-aminoethylcysteine (Aec). CHO is carbohydrate.

Peptide no.	Sequence
T1	Ser-Aec
T2	Glu-Leu-Thr-Asn(CHO)-Ile-Thr-Ile-(Thr)-Val-Glu-Lys
T3	Glx-Leu-Thr-Asn-Ile-(Thr,Ile,Thr,Val,Glu,Lys) Glu-Glu-Aec
T4	Ser-Phe-Aec
T5	Ile-Ser-Ile-Asn(CHO)-Thr-Thr-Trp-Aec
T6	Ala-Gly-Tyr
T7	Ala-Gly-Tyr-(Aec)
T8	Tyr-Thr-Arg
T9	(Aec)-Tyr-Thr-Arg
T10	Asx-Leu-Val-Tyr-Lys
T11	Asp-Leu-Val-Tyr-Lys-Asx-Pro-Ala-Arg-Pro-Asx-Ile-Gln-Lys
T12	Thr-Aec
T13	Thr-Phe-Lys
T14	Glu-Leu-Val-Tyr-Glu-Thr-Val-Lys
T15	Val-Pro-Gly-Aec-Ala-His-His-Ala-Asp-Ser-Leu-Tyr
T16	Glx-Leu-Val-(Tyr,Glx,Thr,Val,Lys,Val,Pro,Gly)-Aec
T17	Ala-His-His-Ala-Asp-Ser-Leu-Tyr
T18	Thr-Tyr-Pro-Val-Ala-Thr-Glu-Aec
T19	His-Aec
T20	Gly-Lys
T21	His-(Aec,Gly) Lys
T22a	Gly-(Lys,Aec)
T23	Asp-Ser-Asp-Ser-Thr-Asp-Aec
T24	Asx-Ser-Asx-Ser-(Thr,Asx,Aec)-Thr-Val-Arg
T25	Gly-Leu-Gly-Pro-Ser-Tyr-Aec
T26	Ser-Phe-Ser-Asp-Ile-Glu-Arg

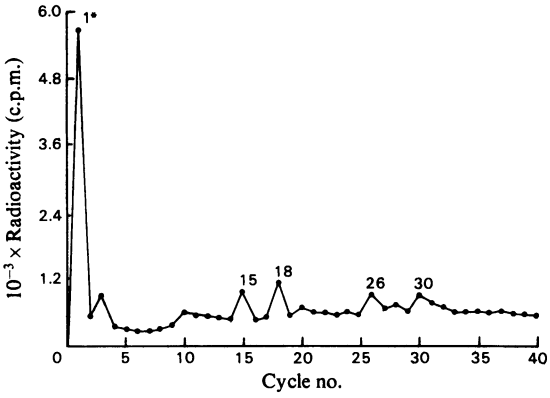

Fig. 5. *Identification of cysteine residues at the N-terminus of the β-subunit*

A freeze-dried preparation of 250μg of follitropin β-subunit was reduced by β-mercaptoethanol and alkylated with iodo[^{14}C]acetamide. After removal of salts, the labelled subunit with the ½-Cys residues containing radioactivity was subjected to Edman degradation in the automated sequenator. At each cycle, a sample of the amino acid phenylthiohydantoin derivative was verified for the presence of radioactivity by counting in a liquid-scintillation counter. The numbers near the peaks indicate the cycle number in which significant radioactivity was found. Cycle numbers 1, 15, 18, 26 and 30 correspond to residue positions 2, 16, 19, 27 and 31 in our proposed sequence (see Fig. 4). Presence of radioactivity in the third cycle would indicate a sequence Xaa-Ser-Cys- at the *N*-terminus of some molecules of the β-subunit.

homology with other species (see Fig. 6) shows 111 residues, with serine at the *N*-terminus and glutamic acid at the *C*-terminus. The total number of each amino acid residue derived from the sequence data agrees well with the composition data (Table 1). The two carbohydrate moieties are presumed to be linked to asparagine residues located at positions 6 and 23. Peptide T2 must be located at the *N*-terminus as shown by sequence data on the oxidized as well as the reduced and *S*-aminoethylated β-subunit. The reason for inability to identify, in the sequenator, the residue at position corresponding to asparagine-6 was owing to glycosylation, since the glycosylated asparagine would not be extracted by the butyl chloride solvent and would remain in the sequencer cup. However, manual dansyl–Edman data on the oxidized subunit and the peptide T2 clearly showed aspartic acid/asparagine at this position. Evidence for the *N*-glycosidic linkage of the oligosaccharide moiety in the second glycopeptide, T5, was based on digestion of the peptide with leucine aminopeptidase and carboxypeptidase, which liberated free serine and threonine respectively. If these were involved in an *O*-glycosidic linkage, they would not have been detectable by amino acid analysis. The peptide T1 was positioned at the *N*-terminus as it could not be accommodated at any other place in the sequence. A lack of this repeating sequence in the molecule (see Fig. 4) also makes this positioning of T1 at the *N*-terminus unequivocal. Additional confidence in this is provided by sequence homology, by comparing the structure of the ovine follitropin β-subunit with similar data published for other species (Fig. 6).

The chymotryptic and thermolytic peptides provided the necessary data to determine the order of the tryptic peptides (Fig. 4). The Tyr^{74}–Thr^{75} bond was highly susceptible to chymotryptic activity present in the trypsin used for digestion. A similar experience has been noted in our previous study with human thyrotropin β-subunit (Sairam & Li, 1977), which has a sequence Tyr-Phe in this region. In contrast with this, the Lys^{39}–Asx^{40} bond was not fully cleaved by trypsin; likewise tryptic cleavage at some of the *S*-aminoethylcysteine residues was also incomplete.

We could not isolate overlap peptides to unequivocally position peptide T23. All the other tryptic peptides were positioned in sequence by using overlap peptides obtained from either the chymotryptic or thermolytic digest. The link between peptides T25 and T26 was unequivocal by the isolation of chymotryptic peptide CD 10. In addition, the composition of a tryptic peptide isolated from oxidized β-subunit perfectly accounted for the region Gly^{97}–Arg^{110}. The placement of a glutamic acid residue at position 111 must be regarded as tentative. This was based on the data from hydrazinolysis of the oxidized β-subunit, which showed this amino acid residue at the *C*-terminus. In addition, the only difference between chymotryptic peptides CD25 and CD26 was the presence of an extra glutamic acid residue in the latter. The failure to detect arginine at the *C*-terminus after hydrazinolysis could be due to decomposition during the reaction (Narita *et al.*, 1975). It is conceivable that there could also be heterogeneity at the *C*-terminus of the β-subunit. This is not unusual among these hormones (Papkoff *et al.*, 1972; Ward *et al.*, 1972).

The single tryptophan residue present in ovine follitropin (Sairam, 1979*a*) is located in its β-subunit (Table 1). This was shown to be present in the tryptic glycopeptide T5 as determined by amino acid analysis with *p*-toluenesulphonic acid hydrolysis. Further evidence came from digestion of the peptide with carboxypeptidases A and B. During a 2h digestion, *S*-aminoethylcysteine, tryptophan and threonine were released quantitatively in that order.

Although quite useful in pointing out sequence heterogeneity at the *N*-terminus of the molecule, the use of automated sequenator to deduce the struc-

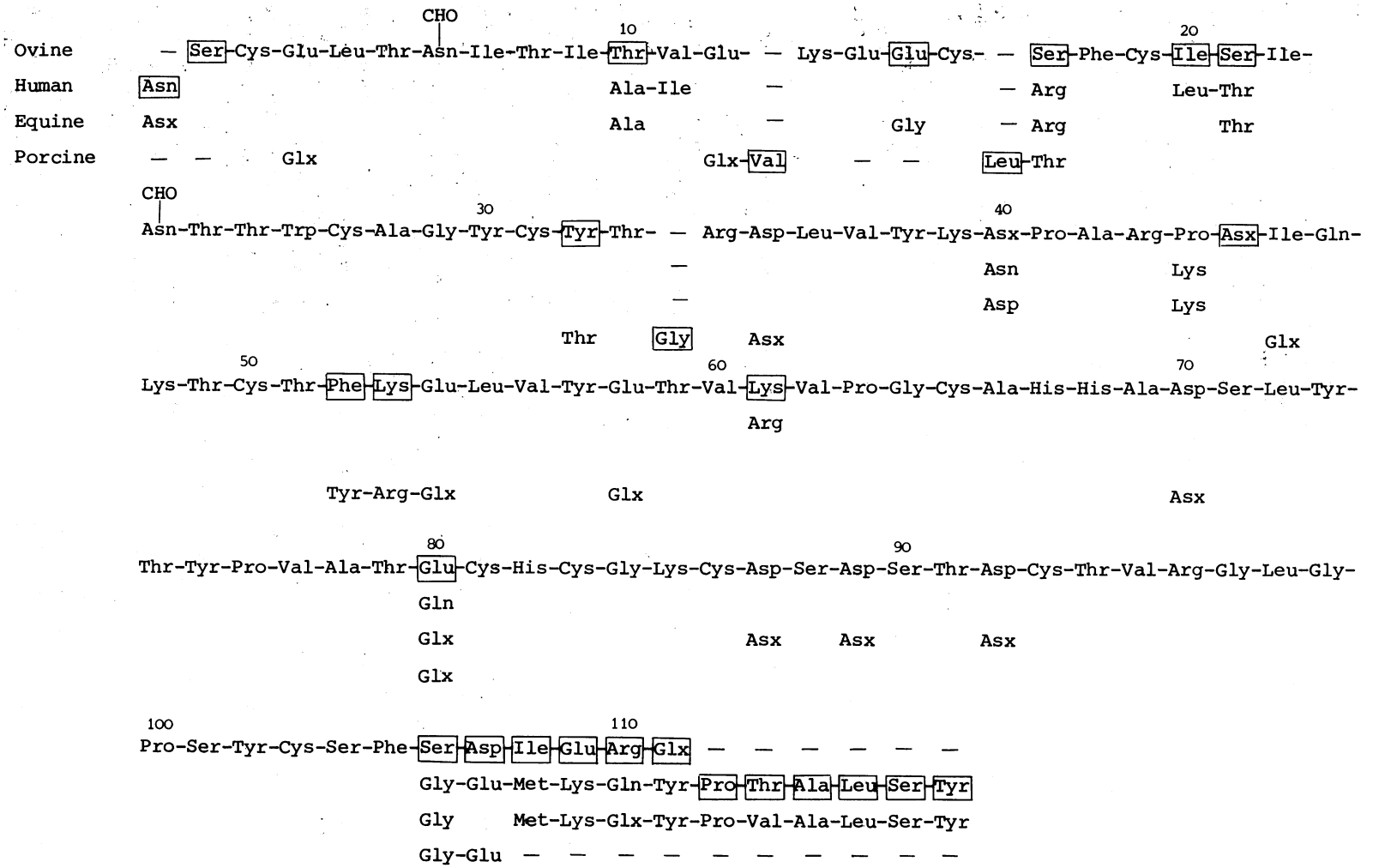

Fig. 6. *Comparison of the amino acid sequence of β-subunit of follitropin from different species*

The numbering of residues is based on the ovine β-subunit sequence from the present study (see Fig. 4). The boxed-in residues show the position of difference in one or the other sequence. –, denotes absence of an amino acid in the corresponding position. CHO denotes an oligosaccharide moiety. Data for human β-subunit is from Saxena & Rathnam (1976), equine is from Fujuki *et al.* (1978) and porcine is from Clossett *et al.* (1978).

ture of the β-subunit had limitations. The repetitive yield of the amino acid phenylthiohydantoin derivatives decreased markedly after the first step (see Fig. 5) of Edman degradation. The reasons for this are not clear. It occurred when both the performic acid-oxidized or the reduced and *S*-aminoethylated β-subunit were used for degradation. Perhaps the bulky carbohydrate groupings present in close proximity to the *N*-terminus were in some way responsible for this drastic decrease in yield. Thus it was decided to solve the primary structure by utilizing our previous strategy (Sairam & Li, 1977) of reliance on enzymically degraded peptides. In this we were fortunate because of the ease with which all the tryptic peptides of the reduced and *S*-aminoethylated protein (Table 2 and Fig. 3), including the two glycopeptides, could be separated rather easily by the direct two-dimensional chromatography–electrophoresis after preliminary separation on Sephadex G-50. In the entire study, the sequence data was gathered from experiments using approx. 75 mg (5 μmol) of the β-subunit. We were unable to clarify the nature of the Asx residues located at positions 40 and 45 in the sequence.

The amino acid sequence of the β-subunit of follitropin from human (Saxena & Rathnam, 1976), porcine (Closset *et al.*, 1978) and equine (Fujiki *et al.*, 1978) pituitaries have been reported. These data are compared with our proposed sequence for ovine follitropin β-subunit in Fig. 6. Comparing it to the human β-subunit, we find 13 residue changes if we assume that the placement of the unidentified amides/acids in our sequence proposal (Fig. 4) are the same as found in human (Saxena & Rathnam, 1976). Most of the replacements can be accounted for on the basis of a single base change in the codon. It should be noted that there are several discrepancies in the data reported on human follitropin β-subunit from two laboratories (Saxena & Rathnam, 1976; Shome & Parlow, 1974). For comparison in Fig. 6, we have selected the data (Saxena & Rathnam, 1976) that is more compatible with ovine hormone data and does not require the introduction of gaps for aligning the two sequences. The *N*-terminal asparagine residue and the additional *C*-terminal residues in the human β-subunit appear to be lacking in the ovine hormone β-subunit preparations examined. The species differences in these areas may be because of heterogeneity. In any case, they are apparently not essential for activity of the hormone. The porcine β-subunit is different from all the three other sequences in that it consists of several extra residues in the region 12–34.

The position of all the 12 half-cystine residues are highly conserved. The two carbohydrate moieties present in the three β-subunits are linked to asparagine residues at identical positions. This has not been identified in the porcine hormone (Closset *et al.*, 1978). The sequences Asn-Ile-Thr- and Asn-Thr-Thr- at the point of carbohydrate attachment are also invariant. The single tryptophan residue occupies an identical position. Apparently this is an inaccessible region of the hormone molecule that is shielded from reactivity with reagents such as *N*-bromosuccinimide. This treatment is reported to have no effect on the biological activity of human and porcine follitropin (Giudice *et al.*, 1978). The comparison of the four β-sequences also reveals that the second half of the molecule (residues 63–106) is remarkably resistant to species variations (assuming that Asx and Glx are same). This region might perhaps be important for interaction with the α-subunit or with the receptor or both. This constant region of the β-subunit contains several important amino acid residues with functional reactive side chains, namely tyrosine, histidine, lysine, arginine, aspartic acid and glutamic acid.

Availability of the complete amino acid sequence of ovine follitropin β-subunits permits a comparison with the sequence of ovine lutropin β-subunit and bovine thyrotropin β-subunit (see Ward, 1978). As ovine thyrotropin β-subunit has not been isolated and sequenced, data available on bovine thyrotropin β-subunit has been used. The sequences of follitropin β-subunit and lutropin β-subunit can be linearly aligned without any gaps, whereas follitropin and thyrotropin β-subunits require a gap of two amino acid residues for maximum fit. Between follitropin β-subunit and lutropin β-subunit, 41 residues are identical; and in follitropin β-subunit and thyrotropin β-subunit, 40 residues occupy positions of identity; among these are included the 12 identical cysteine residue positions.

Thus the studies on ovine follitropin β-subunit reinforce the structural homologies among the β-subunits first observed with all the three glycoprotein hormones of the human pituitary.

The assistance of Mrs. J. Sairam and Mr. R. Carrière is greatly appreciated. This study was supported by grants MT5475 (to M. R. S.) and PG-2 (to N. G. S. and M. C.) of the Medical Research Council of Canada. This paper is the ninth in a series of studies on pituitary follitropin.

References

Closset, J., Rogister, G. M., Hennen, G. & Strosberg, A. H. (1978) *Eur. J. Biochem.* **86**, 115–120

Fujiki, Y., Rathnam, P. & Saxena, B. B. (1978) *J. Biol. Chem.* **253**, 5363–5368

Giudice, L. C., Pierce, J. G., Cheng, K. W., Whitley, R. & Ryan, R. J. (1978) *Biochem. Biophys. Res. Commun.* **81**, 725–733

Narita, K., Matsuo, H. & Nakajima, T. (1975) in *Protein Sequence Determination* (Needelman, N., ed.), pp. 69–91, Springer-Verlag, New York

Papkoff, H. (1972) *Int. Cong. Ser. Endocrinol.* no. 273, pp. 596–600, Excerpta Medica, Amsterdam

Papkoff, H., Sairam, M. R., Farmer, S. W. & Li, C. H. (1972) *Recent Prog. Horm. Res.* **29**, 563–588

Sairam, M. R. (1978) *J. Sci. Ind. Res.* **37**, 86–109

Sairam, M. R. (1979*a*) *Arch. Biochem. Biophys.* **194**, 63–70

Sairam, M. R. (1979*b*) *Arch. Biochem. Biophys.* **194**, 71–78

Sairam, M. R. (1981) *Biochem. J.* **197**, 535–539

Sairam, M. R. & Li, C. H. (1977) *Can. J. Biochem.* **55**, 755–760

Saxena, B. B. & Rathnam, P. (1976) *J. Biol. Chem.* **253**, 993–1005

Seidah, N. G., Benjannet, S., Routhier, R., DeSerres, G., Rochemont, G., Lis, M. & Chrétien, M. (1980) *Biochem. Biophys. Res. Commun.* **95**, 1417–1424

Shome, B. & Parlow, A. F. (1974) *J. Clin. Endocrinol. Metab.* **39**, 203–205

Ward, D. N. (1978) in *Structure and Function of the Gonadotropins* (McKerns, K. W., ed.), pp. 31–45, Plenum Press, New York

Ward, D. N., Reichert, L. E., Jr., Liu, W. K., Nahm, H. S., Hsia, J. Lamkin, N. W. & Jones, N. S. (1972) *Recent Prog. Horm. Res.* **29**, 533–561