Amino Acid Sequence of Progesterone-Induced Rabbit Uteroglobin†

Heiwig Ponstingl,* Antonio Nieto,‡ and Miguel Beato§

ABSTRACT: Uteroglobin, a steroid-binding protein of the uterine secretion of the rabbit which is induced by progesterone, comprises two identical polypeptide chains of 70 amino acid residues linked by two disulfide bonds. The primary structure has been determined by using both automated and manual methods of Edman degradation. Overlapping peptides were isolated from tryptic, chymotryptic, and CNBr digests. The sequence is not homologous to any known protein except for a small acidic region (residues 22–29) resembling a sequence found in somatotropin. The C-terminal half is relatively basic. Implications for the secondary structure are discussed.

Uteroglobin is a small secretory protein of the rabbit uterus which reaches high concentrations during the preimplantation period of pregnancy and can be induced experimentally by the combined action of estradiol and progesterone (Beier, 1968; Krishnan and Daniel, 1967; Arthur and Daniel, 1972; Bullock and Conell, 1973). In addition, uteroglobin binds the inducing steroid, progesterone, with high affinity and specificity (Beato and Baier, 1975; Beato, 1976; Beato et al., 1977).

We have elucidated the primary structure of the protein as part of an attempt to gain insight into the molecular architecture of the steroid binding site. In a previous paper we have reported a procedure for the purification of uteroglobin from the uterine flushings of rabbits treated sequentially with estradiol and progesterone. We also showed that uteroglobin is composed of two polypeptide chains covalently bound by two disulfide bridges (Nieto et al., 1977). Here we present the complete amino acid sequence of uteroglobin, as determined by automated and manual Edman degradation of the intact polypeptide and of fragments obtained by tryptic, chymotryptic, and cyanogen bromide cleavage.

Materials and Methods

Isolation. Uteroglobin was purified as previously described (Nieto et al., 1977) starting with the uterine flushings of rabbits treated with estradiol and progesterone. The procedure involves gel filtration on Sephacryl S200 (Pharmacia) and ion-exchange chromatography on CM[1]-cellulose (CM-32, Whatman). The final preparation was homogeneous by the criteria of gel electrophoresis, Ouchterlony double diffusion, immunoelectrophoresis, and determination of the N-terminal residue by dansylation. Furthermore, it has been recently crystallized (Bühner and Beato, 1978). For sequence analysis, this material was reduced and carboxymethylated as previously described (Nieto et al., 1977).

Amino Acid Analysis. Samples were hydrolyzed in 6 M HCl containing 0.02% 2-mercaptoethanol for 24 and 72 h, respectively, at 110 °C in sealed tubes under nitrogen, evaporated, and analyzed by a single-column procedure on a Beckman Multichrom analyzer (Beckman Instruments, München).

Tryptophan was determined in a protein sample hydrolyzed in 4 M methanesulfonic acid containing 0.2% 3-(2-aminoethyl)indole for 24 h at 110 °C (Liu and Chang, 1971).

Enzymatic Digests. Trypsin, treated with (tosylamido-2-phenyl)ethyl chloromethyl ketone (Serva, Heidelberg); chymotrypsin (Merck, Darmstadt); aminopeptidase M (Boehringer, Mannheim); and carboxypeptidase Y (Boehringer, Mannheim) were used without further treatment.

Tryptic Digestion. Reduced and S-carboxymethylated uteroglobin (670 nmol) was dissolved in 2 mL of 0.1 M NH_4HCO_3 (pH 8.0) and incubated for 12 h at 37 °C with 0.5 mg of trypsin. The digest was frozen, lyophilized, and redissolved in the starting buffer for ion-exchange chromatography.

Chymotryptic Digestion. Reduced and S-carboxymethylated uteroglobin (800 nmol) was dissolved in 2 mL of 0.01 M NH_4HCO_3, adjusted to pH 8.2 with ammonia. One milligram of chymotrypsin was added and the mixture incubated for 8 h at 37 °C. Then the digest was frozen and lyophilized twice.

Digestion with Aminopeptidase M. About 20 nmol of peptide was lyophilized, dissolved in 1 mL of 0.01 M Tris-HCl (pH 7.5), and incubated with 0.1 mg of the enzyme suspension at 37 °C for 1 h. The reaction was terminated with 1 drop of 6 M HCl and evaporated to dryness, and the liberated amino acids were determined on the amino acid analyzer.

Digestion with Carboxypeptidase Y (Hayashi, 1976). Reduced and carboxymethylated uteroglobin (50 nmol) was dissolved in 1 mL of 6 M urea and 0.1 M pyridine acetate (pH 6.0) and incubated for 2 h with 10 μg of the enzyme. The mixture was brought to pH 2.2 with HCl and analyzed on the amino acid analyzer. Norleucine was used as an internal standard.

Cleavage with cyanogen bromide was performed as described previously (Nieto et al., 1977) in 70% formic acid for 24 h.

Maleylation. Ten milligrams of S-carboxymethylated uteroglobin was dissolved in 2 mL of 0.1 M borate (pH 9.0) and reacted with 30 mg of maleic anhydride given in six aliquots of 5 mg each at 30-min intervals. Total incubation time was 3 h at 2 °C and pH 9.0 was maintained by adding 1 M NaOH. The mixture was desalted on Bio-Gel P2 (Bio-Rad) in 0.2 M

† From the Institut für Zellforschung, Deutsches Krebsforschungszentrum, Heidelberg, Federal Republic of Germany, and the Institut für Physiologische Chemie der Philipps-Universität Marburg, Federal Republic of Germany. *Received May 2, 1978.* This work was supported by the Deutsche Forschungsgemeinschaft.

‡ Recipient of an EMBO fellowship. Present address: Institute Biology Desarollo, CSIC, Velázquez 144, Madrid-6, Spain.

§ Institut für Physiologische Chemie, Philipps-Universität Marburg, F.R.G.

[1] Abbreviations used are: T, tryptic peptide; CB, cyanogen bromide peptide; TM, tryptic maleylated peptide; CM-Cys, carboxymethylcysteine; quadrol, *N,N,N',N'*-tetrakis(2-hydroxypropyl)ethylenediamine; CM, carboxymethyl; Tris-HCl, 2-amino-2-hydroxymethyl-1,3-propanediol hydrochloride.

0006-2960/78/0417-3908$01.00/0

TABLE I: Amino Acid Composition of Tryptic Peptides Isolated from Rabbit Uteroglobin.[a]

	T1	T1a[e]	T1b[e]	T2	T3	T4	T5	T6	T7	T8	T9	T10	total[e]	protein[f]
Cys[b]				1.01 (1)	1.00 (1)								2	1.96
Asp	1.23 (1)	1.13 (1)				0.99 (1)	1.95 (2)	1.15 (1)	1.03 (1)				6	6.06
Thr	1.89 (2)		2.01 (2)				0.96 (1)	1.78 (2)		1.00 (1)			6	5.89
Ser	1.10 (1)	1.08 (1)		0.98 (1)				2.72 (3)					5	5.80
Glu	1.13 (1)	1.00 (1)				1.05 (1)	1.92 (2)	1.82 (2)	0.93 (1)	0.96 (1)			8	8.16
Pro	1.09 (1)	0.99 (1)		1.00 (1)	1.09 (1)		1.03 (1)	1.08 (1)					5	4.66
Gly					0.92 (1)	1.03 (1)		1.00 (1)					3	3.41
Ala						0.94 (1)		1.24 (1)					2	2.14
Val	0.82 (1)	0.84 (1)						1.06 (1)				0.93[d] (1)	3	2.91
Met				0.89 (1)		1.62 (2)	1.07 (1)		1.08 (1)				5	4.09
Ile					0.99 (1)			0.98 (1)	0.96 (1)			0.85[d] (1)	4	3.77
Leu	1.88 (2)	1.96 (2)		1.01 (1)				3.81[d] (4)		1.03 (1)			8	8.17
Tyr								0.98 (1)					1	0.91
Phe							0.98 (1)	0.99 (1)					2	2.01
His								1.06 (1)					1	1.03
Lys						0.99 (1)	1.10 (1)	0.98 (1)	1.00 (1)	1.01 (1)	1.00 (1)	1.21[d] (1)	7	7.17
Arg	0.86 (1)		0.99 (1)		0.99 (1)								2	1.93
total res	10	7	3	5	5	7	9	21	5	4	1	3	70	
yield[g]	30	40	67	80	81	98	104	90	85	92	64	85	M_r = 7900	
N-term res	Val	Val	Thr	Ser	Gly	Asx	Glx	Phe	Glx	Leu	Lys	Ile		Gly
R_f[c]	0.49	0.56	0.25	0.58	0.37	0.11	0.21	0.86	0.20	0.40	0.27	0.50		

[a] Presented as molar ratios. [b] Measured as carboxymethylcysteine. [c] Thin-layer chromatography on cellulose plates in butanol–pyridine–acetic acid–H_2O, 90:60:18:72 (v/v). [d] After 72-h hydrolysis. [e] T1a and T1b are the products of a side reaction in the tryptic digestion and are not accounted for in the total amino acid composition. [f] From Nieto et al. (1977), recalculated for 70 amino acid residues. Uteroglobin is devoid of tryptophan. [g] In nanomoles.

TABLE II: Composition of Overlapping Chymotryptic Peptides of Rabbit Uteroglobin.[a]

	Ch1	Ch2	Ch3	Ch4	Ch5	Ch6.	Ch7	Ch8	Ch9	Ch10
Cys[b]				0.77 (1)			0.95 (1)			
Asp			1.03 (1)		1.12 (1)	3.05 (3)		1.15 (1)		
Thr	0.96 (1)	0.99 (1)				1.25 (1)		1.97 (2)	1.12 (1)	
Ser	1.97 (2)	1.12 (1)	0.96 (1)						0.87 (1)	
Glu		0.89 (1)	1.05 (1)		1.15 (1)	3.24 (3)		1.13 (1)	1.23 (1)	
Pro	0.93 (1)		0.99 (1)			1.10 (1)	1.00 (1)		0.91 (1)	
Gly	1.04 (1)					0.89 (1)	1.27 (1)			
Ala					0.90 (1)	0.91 (1)				
Val					0.95 (1)				0.75 (1)	0.98[d] (1)
Met				1.23 (1)		1.78 (2)				
Ile					0.92 (1)		1.05 (1)	0.81 (1)	0.70 (1)	
Leu	1.12 (1)	1.01 (1)	0.96 (1)		2.02 (2)				1.53 (2)	0.99[d] (1)
Tyr	0.76 (1)									
Phe						0.89 (1)	1.02 (1)			
His					0.95 (1)					
Lys						1.93 (2)			2.18 (3)	2.04[d] (2)
Arg							0.98 (1)	0.94 (1)		
total res	7	4	5	2	8	15	6	6	11	4
N-term res	Leu	Glx	Asx	CM-Cys	Ala	Lys	Gly	Thr	Lys	Lys
R_f[c]	0.78	0.70	0.55	0.73	0.89	0.61	0.82	0.77	0.88	0.60

[a] Presented in molar ratios. [b] Measured as carboxymethylcysteine. [c] Chromatographic system as in Table I. [d] After 72-h hydrolysis.

ammonium bicarbonate (pH 8.0) and lyophilized. The product was treated with trypsin as above.

Peptide Purification. Operating conditions for gel and ion-exchange chromatography are listed with the respective figures. With tryptic and chymotryptic peptides, an aliquot of each fraction was monitored with ninhydrin following alkaline hydrolysis. Purity of peptides was checked by thin-layer chromatography on cellulose plates (TLC plastic sheets, 0.1-mm cellulose layer, Merck, Darmstadt) in 1-butanol–pyridine–acetic acid–H_2O (90:60:18:72). Some peptides were preparatively purified using this system. Peptide spots were detected by spraying with ninhydrin (2 mg/mL in acetone) and developing at 100 °C.

Sequence Analysis. Reagents for sequence analysis were obtained from Beckman Instruments (Palo Alto, Calif.) and Pierce Chemical Co. (supplied by Günter Karl OHG, Geisenheim, F.R.G.). Polyamide sheets A 1700 were obtained from Schleicher and Schüll, Dassel, F.R.G., and cut into pieces of 5 × 5 cm.

Automated sequencing was performed on a Beckman 890C-sequencer using 0.1 M quadrol in program no. 122974 with single cleavage (adapted from Brauer et al., 1975). Thiazolinone derivatives were converted to the corresponding phenylthiohydantoins in 1 M HCl and identified by thin-layer chromatography (Edman and Henschen, 1975); nonpolar

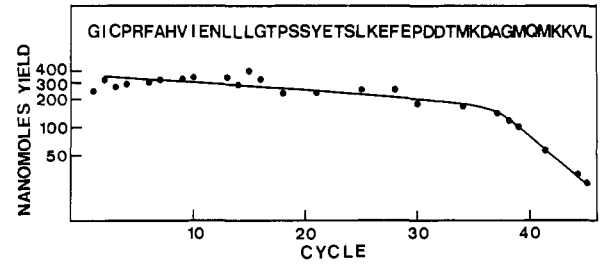


FIGURE 1: Automated Edman degradation of rabbit uteroglobin. All phenylthiohydantoins were identified by thin-layer chromatography; only nonpolar derivatives were quantitated by gas chromatography and corrected for overlap.

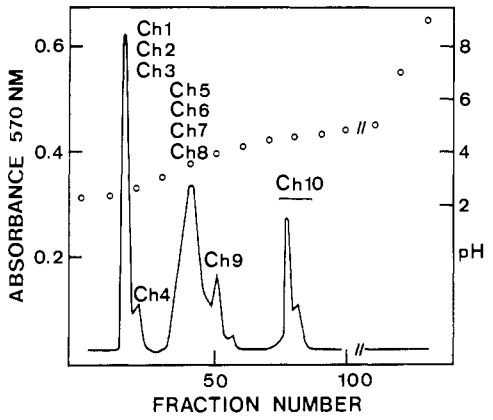


FIGURE 2: Chromatography of chymotryptic peptides on Dowex 50X2, column 0.6 × 16 cm, 45 °C. Linear gradient of 250 mL of 0.05 M pyridine formate (pH 2.4) and 250 mL 2.0 M pyridine acetate (pH 5.3), both buffers 30% (v/v) in 1-propanol. Fractions of 3 mL. Coeluting peptides were further purified by thin-layer chromatography.

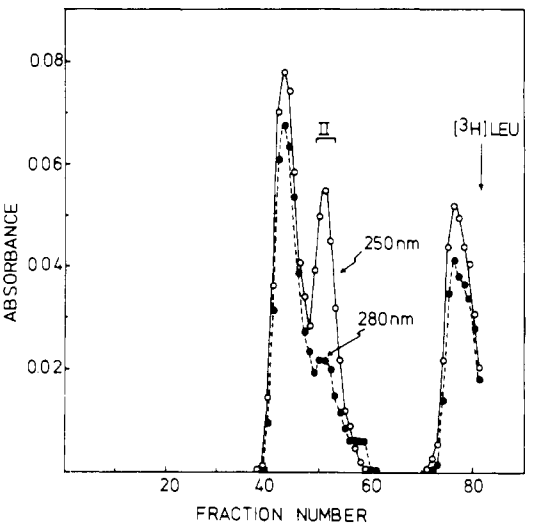


FIGURE 3: Gel chromatography of maleylated tryptic fragments on Sephadex G-25 superfine: column 1.5 × 150 cm. The tryptic digest (2 mL) was adjusted to pH 3.0 with 1 M acetic acid and incubated at 40 °C for 40 h. The sample was then frozen, lyophilized, resuspended in 1 mL of 1 M acetic acid, and applied to the column equilibrated with 1 M acetic acid. The column was eluted at a flow rate of 10 mL/h and 3-mL fractions were collected. Tritiated leucine was added as internal marker. Fraction II is the peptide used in sequence studies.

TABLE III: Composition of Overlapping Cyanogen Bromide (CB) Peptides and a Maleylated Tryptic (TM) Peptide of Rabbit Uteroglobin.[a]

	CB1	CB2	CB3	CB4	TM
Cys[b]	0.93 (1)				0.86 (1)
Asp	3.08 (3)	2.19 (2)	1.00 (1)		0.99 (1)
Thr	2.94 (3)	2.02 (2)			1.03 (1)
Ser	3.12 (3)	1.12 (1)			1.07 (1)
Glu	4.01 (4)	2.18 (2)		0.96 (1)	2.09 (2)
Pro	3.03 (3)	0.92 (1)			0.90 (1)
Gly	2.12 (2)		1.10 (1)		
Ala	1.22 (1)		0.99 (1)		
Val	0.85 (1)	1.00[d] (1)			1.07 (1)[d]
Met	0.81 (1)[e]	0.78[e] (1)	0.74[e] (1)	1.04[e] (1)	1.85 (2)
Ile	1.76 (2)	0.91[d] (1)			1.90 (2)[d]
Leu	3.87 (4)	1.81[d] (2)			2.05 (2)[d]
Tyr	0.92 (1)				
Phe	1.96 (2)				
His	0.96 (1)				
Lys	0.95 (1)	2.12[d] (2)	0.90 (1)		2.79 (3)[d]
Arg	0.99 (1)	0.86 (1)			
total res	34	16	5	2	17
N-term res	Gly	Lys	Lys	Glx	Glx
R_f[c]	0.71	0.76	0.82	0.72	nd[f]

[a] Presented in molar ratios. [b] Measured as cysteic acid in CB and as carboxymethylcysteine in TM. [c] Chromatographic system as in Table I. [d] After 72-h hydrolysis. [e] Measured as homoserine + homoserine lactone. [f] nd, not determined.

residues were quantitated by gas chromatography (Pisano, 1975) using methyl arachate as internal standard. Manual Edman degradation plus dansylation was performed with peptides as described by Hartley (1970). Dansylated amino acids were identified by two-dimensional thin-layer chromatography on polyamide sheets 5 × 5 cm (Woods and Wang, 1967), run in benzene–acetic acid (9:1, v/v) in the first direction and in water–formic acid (200:3, v/v) in the second direction followed by ethyl acetate–methanol–acetic acid (20:1:1, v/v) in the second direction for unequivocal separation of amino acids with a low R_f. Identification was based on an identical R_f of the sample spot and one spot of the cochromatographed standard mixture. Amides in side chains were assigned by digestion of the peptide with aminopeptidase M and amino acid analysis of the digest or by thin-layer identification of the degraded phenylthiohydantoin.

Results

Two thirds of the sequence was established using the automated sequenator method. Two runs with 400 nmol of carboxymethylated uteroglobin were performed giving qualitatively identical results. However, in the first run overlap increased considerably at each of the three proline residues in the N-terminal region. Therefore, in the second run (shown in Figure 1) triple cleavage with heptafluorobutyric acid and extraction with butyl chloride were used instead of the routine single cleavage in cycles where proline was expected. Thus, overlap was kept moderate amounting to 17% in cycle 38. On adding the overlap to the previous cycle, the repetitive yield was 98% up to residue 36. Thereafter it declined rapidly. The break in the curve of Figure 1 coincides with the loss of acidic side chains.

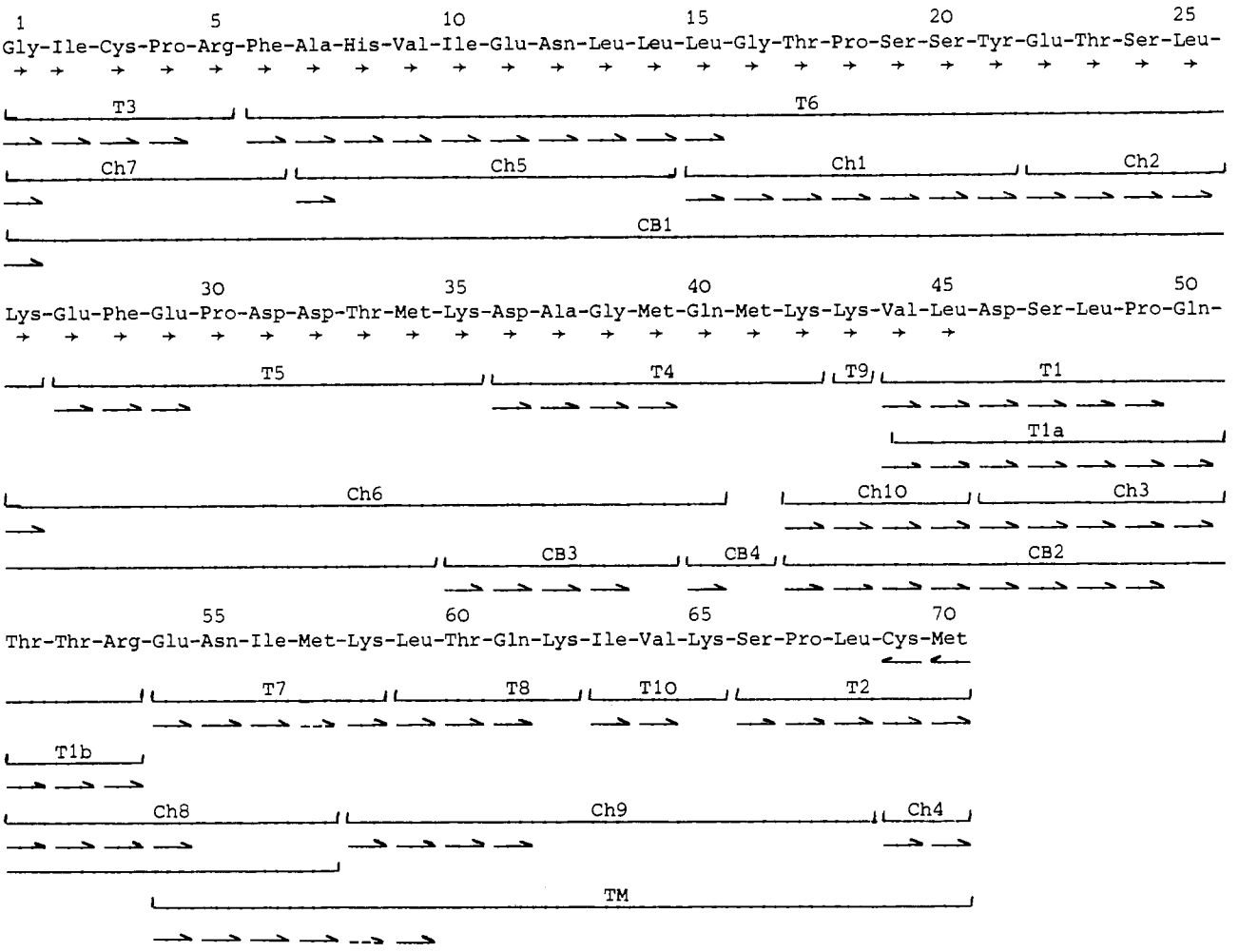
Since this sequence was largely confirmed with independent methods using peptides, we did not attempt to quantitate polar phenylthiohydantoins by hydrolysis to the free amino acid. Only apolar residues were determined quantitatively by gas chromatography.

Digestion of the polypeptide with carboxypeptidase Y gave 74% Met and 4% CM-Cys, suggesting a C-terminal sequence of Cys–Met.

Tryptic Peptides. By chromatography on Dowex-50, ten tryptic peptides were isolated (Beato and Nieto, 1976; composition given in Table I) which together account for the amino acid composition of the protein. Uteroglobin is devoid of tryptophan, as tested by amino acid analysis following hydrolysis in methanesulfonic acid, ultraviolet spectrum, and Ehrlich test with the peptides (Nieto et al., 1977). T1 was obtained in relatively low yield, apparently because a side reaction gave a partial split resulting in peptides T1a and T1b, which were also isolated. The alignment of most of these peptides was evident from terminal sequence data, except for peptides T7, T8, and T10 (Table IV).

Chymotryptic Peptides. Cleavage with chymotrypsin and chromatography of the digest on Dowex-50 followed by thin-layer separation yielded ten chymotryptic fragments (Figure 2, Table II). Some of them were used to confirm sequences

TABLE IV: Amino Acid Sequence of Rabbit Uteroglobin.[a]



[a] Abbreviations used: (→) determined by automated Edman degradation (see Figure 1); (⇀) determined by manual Edman degradation and dansyl technique as described in the text; (⇢) residue degraded but not identified; (←) determined by digestion with carboxypeptidase Y as described in the text.

already established. Ch8 proves that T7 must be N-terminally linked to T1 (Table IV). Ch9 established the order T8-T10-T2. However, since it overlaps with T7 only by a single N-terminal Lys, more evidence was desired to show that no tryptic peptide was missing between T7 and T8.

Cyanogen Bromide and Maleylated Tryptic Peptides. From the CNBr cleavage, four fragments were purified by gel filtration (Nieto et al., 1977; composition given in Table III). CB2 gives additional evidence for the ordering T9-T1-T7, but the expected C-terminal peptide was not obtained in sufficient purity for sequence studies. Therefore, uteroglobin was maleylated and split at the two Arg residues with trypsin, and the C-terminal piece was isolated by gel filtration (Figure 3; composition Table III). It establishes the overlap T7-T8 and proves that no tryptic peptide is missing. In case of successive identical residues (Leu 13-15, Ser 19-20, Asp 31-32, Lys 42-43, Thr 51-52), sequence determination is supported by the amino acid composition of all peptides comprising the region in question. The complete sequence of rabbit uteroglobin is shown in Table IV.

Assignment of Side-Chain Amides. Up to residue 40 this was done by thin-layer chromatography of the respective phenylthiohydantoins. In addition, Ch2 was digested with aminopeptidase M, yielding a composition of 0.99 Thr, 1.07 Ser, 0.97 Glu, and 1.01 Leu. The same procedure gave 1.00 Asp, 1.01 Gln, 0.89 Ser, 1.29 Pro, and 0.82 Leu for peptide Ch3. Hence, position 46 is Asp and position 50 Gln. T7 yielded 1.08 Asn, 1.08 Glu, 0.96 Met, 0.94 Ile, and 1.02 Lys, proving Glu-54 and Asn-55. Similarly, T8 gave 0.90 Thr, 0.69 Gln, 0.61 Glu, 0.86 Leu, and 0.93 Lys, characterizing a partially deamidated Gln in position 61.

Discussion

The amino acid sequence of uteroglobin was determined as a step in an attempt to elucidate the configuration of a steroid binding site. The protein has a molecular weight of about 15 000 before (Nieto et al., 1977) and 7900 after reduction of disulfide bonds. Sequence data show that it comprises two identical chains of 70 residues. Both cysteines participate in interchain disulfide bonds.

Noticeable features of the sequence are a highly polar region in positions 22-36 and a regular spacing of basic side chains in positions 53-65. While the sequence in a computer search turned out to be otherwise unrelated to known proteins, the polar region Glu-Thr-Ser-Leu-Lys-Glu-Phe-Glu (residues 22-29) is homologous to a sequence Asp-Thr-Phe-Lys-Glu-Phe-Glu (residues 26-32) in ovine and bovine somatotropin (Li et al., 1973; Wallis, 1973).

Uteroglobin binds progesterone with high affinity, which is enhanced upon reduction of the disulfide bonds. UV difference spectra indicate that the only tyrosine in the chain, residue 21, participates in steroid binding (Beato et al., 1977).

The relatively low molecular weight and the fact that it has been crystallized make uteroglobin a suitable object for evaluating prediction methods for secondary structures. Using the method of Chou and Fasman (1974) with updated parameters (Chou and Fasman, 1977a,b), the following predictions were made.

Residues 3–6 should form a β-turn, followed by an α-helix of residues 6–15. Residues 17–20 again have a high potential for β-turn formation. Positions 25–29 would be a short helical region terminated by another β-turn in positions 29–32. Residues 34–45 could be helical, and the region 56–65 has about equal helical and β-sheet potential. A final β-turn is predicted for residues 66–69. Given a helix in position 56–65, the spacing of Lys in the sequence of this region results in a stack of three lysines constituting one side of the helix.

Acknowledgments

We are indebted to Behringwerke Marburg for providing the animals used for the purification of uteroglobin and to Schering AG, Berlin, for the gift of steroid hormones required for uteroglobin induction. We thank Jutta Arnemann, Alfons Egle, and Hannelore Thierberger for their excellent technical assistance.

References

Arthur, A. T., and Daniel, J. C. (1972), *Fertil. Steril. 23,* 115.

Beato, M. (1976), *J. Steroid Biochem. 7,* 327.

Beato, M., and Arnemann, J. (1975), *FEBS Lett. 58,* 126.

Beato, M., and Baier, R. (1975), *Biochim. Biophys. Acta 392,* 346.

Beato, M., and Nieto, A. (1976), *Eur. J. Biochem. 64,* 15.

Beato, M., Arnemann, J., and Voss, H.-J. (1977), *J. Steroid Biochem. 8,* 725.

Beier, H. M. (1968), *Biochim. Biophys. Acta 160,* 289.

Brauer, A. W., Margolies, M. N., and Haber, E. (1975), *Biochemistry 14,* 3029.

Bühner, M., and Beato, M. (1978), *J. Mol. Biol. 120,* 331.

Bullock, D. W., and Conell, K. M. (1973), *Biol. Reprod. 9,* 125.

Chou, P. Y., and Fasman, G. D. (1974), *Biochemistry 13,* 222.

Chou, P. Y., and Fasman, G. D. (1977a), *J. Mol. Biol. 115,* 135.

Chou, P. Y., and Fasman, G. D. (1977b), *Trends Biochem. Sci. 2,* 128.

Edman, P., and Henschen, A. (1975), in Protein Sequence Determination, Needleman, S. B., Ed., New York, N.Y., Springer-Verlag, p 232.

Hartley, B. S. (1970), *Biochem. J. 119,* 805.

Hayashi, R. (1976), *Methods Enzymol. 45,* 568.

Krishnan, R. S., and Daniel, J. C. (1967), *Science 158,* 490.

Li, C. H., Gordon, D., and Knorr, J. (1973), *Arch. Biochem. Biophys. 156,* 493.

Liu, T.-Y., and Chang, Y. J. (1971), *J. Biol. Chem. 246,* 2842.

Nieto, A., Ponstingl, H., and Beato, M. (1977), *Arch. Biochem. Biophys. 180,* 82.

Pisano, J. J. (1975), in Protein Sequence Determination, Needleman, S. B., Ed., New York, N.Y., Springer-Verlag, p 280.

Wallis, M. (1973), *FEBS Lett. 35,* 11.

Woods, K. R., and Wang, K.-T. (1967), *Biochim. Biophys. Acta 133,* 369.