N-Terminal Sequences of Uteroglobin and its Precursor

By M. ATGER,* J.-C. MERCIER,† G. HAZE,† F. FRIDLANSKY* and E. MILGROM*

*Groupe de Recherches sur la Biochimie Endocrinienne et la Reproduction (INSERM U 135), Faculté de Médecine Paris-Sud, 94270 Bicêtre, France, and †Laboratoire de Recherches sur les Protéines, INRA-CNRZ, 78350 Jouy en Josas, France

(Received 20 November 1978)

Translation of uteroglobin mRNA in wheat-germ extract has yielded a precursor protein (pre-uteroglobin) containing an N-terminal extension of 21 amino acid residues. The sequence of this extension and that of the 50 N-terminal amino acid residues of uteroglobin have been determined.

Uteroglobin (Beier, 1978) or blastokinin (Krishnan & Daniel, 1977) is a protein secreted by the endometrium of pregnant, pseudopregnant or progesterone-treated rabbits. It offers an interesting model for the study of the mechanism or action of steroid hormones, since it is a major gene product (over 50% of proteins in the endometrial secretion at day 5 of pregnancy) and since it has a relatively simple structure (two identical subunits each of about 8000 mol.wt.).

We have previously purified its messenger RNA (Atger & Milgrom, 1977) and shown that it is translated in the wheat-germ system as a precursor protein of about 10000 daltons. The existence of a precursor has also been demonstrated in other laboratories (Beato & Nieto, 1976; Bullock *et al.*, 1976). We now show that pre-uteroglobin has an extension of 21 amino acid residues in the N-terminal region. The amino acid sequence of this leader region and also that of the N-terminal part of uteroglobin have been determined.

Materials and Methods

Purification of uteroglobin

New Zealand rabbits, 5 days pregnant, were used. The collection of uterine fluid and purification to homogeneity of the protein by chromatography on Sephadex G-75 and DEAE-cellulose were previously described (Fridlansky & Milgrom, 1976).

Purification and translation of uteroglobin mRNA

RNA was extracted from polyribosomes of 5-day-pregnant-rabbit endometria. It was purified by poly(U)-Sephadex chromatography and two successive sucrose-gradient ultracentrifugations. It has been shown by various criteria that this preparation is over 90% pure (Atger & Milgrom, 1977).

The mRNA (1.5 µg/100 µl) was translated as previously described (Atger & Milgrom, 1977) in the presence of 0.1 mCi/ml of radioactive amino acid in a wheat-germ extract. At the end of the incubation

a 60 min/105000g centrifugation was performed. The supernatant was incubated at 37°C in 40 mM-EDTA, pH 7.3, and 80 µg/ml of ribonuclease A (Worthington, Freehold, NJ, U.S.A.). Radioactive uteroglobin was then immunoprecipitated by the double-antibody technique previously described (Atger & Milgrom, 1977).

Sources of materials

The following ³H-labelled amino acids were obtained from CEA-France, Gif-sur-Yvette, France: serine (30), glutamic acid (21), proline (22), glycine (26), alanine (50), valine (35), isoleucine (35), leucine (30), tyrosine (46), phenylalanine (29), lysine (33), histidine (46), tryptophan (4); and arginine (10). [³⁵S]Methionine (880) was purchased from the Radiochemical Centre, Amersham, Bucks, U.K. (the numbers in parentheses refer to specific radioactivities in Ci/mmol). Bovine β-lactoglobulin used as a carrier was kindly given by Dr. F. Addeo, Istituto Industrie Agrarie, Università Napoli, Italy.

Determination of the N-terminal sequence of uteroglobin

Uteroglobin was reduced and alkylated with iodoacetic acid by a slightly modified procedure of Crestfield *et al.* (1963) (6M-guanidine and dithioerythritol were used as dissociating and reducing reagent respectively) before sequence analysis. Automated Edman degradation was performed twice with a Beckman updated model 890B sequencer by using the Beckman dimethylbenzylamine program (no. 102473). Amino acid phenylthiohydantoin derivatives were identified by both t.l.c. (Cohen-Solal & Bernard, 1973) and g.l.c. (Pisano *et al.*, 1972), in the latter case after trimethylsilylation.

Determination of the N-terminal sequence of uteroglobin synthesized in vitro

The immunoprecipitates containing the uteroglobin translated and radiolabelled with one amino

acid at a time were analysed either individually or as a pool in a Beckman sequencer by using the 102473 program. Each sample was supplemented with 5 mg of bovine β -lactoglobulin carrier. The material released at each cycle of the Edman degradation was converted with 1 M-HCl at 80°C during 10 min, and the resulting amino acid phenylhydantoin derivatives analysed as follows. A portion ($\frac{1}{2}$) was gas-chromatographed to monitor the efficiency of the carrier degradation, in order to calculate the repetitive yield. Another portion ($\frac{1}{4}$) was assayed for radioactivity in toluene-based scintillation fluid in a Packard model 3375 liquid-scintillation counter. The remaining portion was used for identification of the labelled amino acid phenylhydantoin derivatives that were co-chromatographed with appropriate amino acid phenylhydantoin derivative carriers on silica gel (Merck F254). The spots were scraped off the plates and counted in the liquid-scintillation instrument. The aqueous phase was counted for radioactivity whenever a sample contained radiolabelled arginine or histidine.

Results

N-Terminal sequence of uteroglobin

Uteroglobin being composed of two subunits each of mol.wt. close to 8000 should contain about 70 amino acid residues per subunit. Automated Edman degradation of *S*-carboxymethylated uteroglobin through 52 cycles yielded the sequence of 50 residues (Fig. 1), thus providing the primary structure of two-thirds of the polypeptide chain. Two successive analyses were performed, giving the same results (not shown), and subsequent analyses of radioactive uteroglobin synthesized *in vitro* confirmed in all cases the proposed *N*-terminal sequence.

Amino acid sequence of the N-terminal extension of pre-uteroglobin

Amino acid-sequence analyses were performed on uteroglobin synthesized *in vitro* radiolabelled with one amino acid each time. Fig. 2 shows the amount of radioactivity recovered at each cycle of the automated Edman degradation carried out on 15 different samples. The partial *N*-terminal sequence of the product translated *in vitro*, deduced from these data, is clearly identical with that of authentic uteroglobin after cycle 21 (Fig. 1), thus demonstrating the occurrence of a precursor protein with a 21-residues-long *N*-terminal extension. In the case of arginine the release of radioactivity detected at the early sequencer cycles probably corresponded to contaminating free ^3H -labelled amino acid (as shown by repetitive yield measurements). In the pre-piece, residue 11, which has not been identified, should be one of the following

Fig. 1. *N*-Terminal sequences of uteroglobin and its precursor. Residues written in *italics* are those of the signal peptide. Others are those of authentic uteroglobin. Underlined amino acid residues of authentic uteroglobin were also identified by radiosequence analysis of pre-uteroglobin. Unidentified residues are marked ●. The upper numbers refer to the position of residues in the polypeptide chain of pre-uteroglobin. The lower ones refer to the position of residues in the signal peptide and authentic uteroglobin respectively. The residue occurring at position 11 in the pre-piece should be one of the following amino acids: aspartic acid, asparagine, glutamine, threonine or cysteine. Furthermore, one of these five amino acids might also be found at positions 6 and 16 (see the text).

Fig. 2. Sequence of immunoprecipitates of uteroglobin translation *in vitro* radiolabelled with one amino acid each time. Automated Edman degradation was performed either on separate ($[^3\text{H}]$ leucine; $[^3\text{H}]$ phenylalanine; $[^3\text{H}]$ serine; $[^3\text{H}]$ tryptophan; $[^{35}\text{S}]$ methionine) or pooled immunoprecipitates ($[^3\text{H}]$ isoleucine + $[^3\text{H}]$ alanine + $[^3\text{H}]$ glycine + $[^3\text{H}]$ arginine; $[^3\text{H}]$ lysine + $[^3\text{H}]$ proline + $[^3\text{H}]$ valine; $[^3\text{H}]$ glutamic acid + $[^3\text{H}]$ tyrosine + $[^3\text{H}]$ histidine). Radiosequence data have been corrected for background or any 'out-of-cycle' radioactivity. ---- Indicates the theoretical yields of radioactivity calculated from the repetitive yield (ranging from 94 to 96%) of the degradation of the bovine β -lactoglobulin carrier. Cycle zero represents the blank cycle. Amino acid panel labels have been underlined whenever the Figure refers to radioactivity recovered from silica-gel chromatograms. Sequence positions assigned to various amino acids are indicated by arrows. Sequence analysis of pre-uteroglobin radiolabelled with $[^3\text{H}]$ tryptophan gave no significant radioactivity above background (results not shown). $[^3\text{H}]$ Glycine immunoprecipitate was sequenced twice, and similar patterns of radioactivity were obtained in both cases. The panel labelled Leu* refers to direct sequence analysis of the translation product from purified uteroglobin mRNA (no immunoprecipitation was performed before Edman degradation).

amino acids: aspartic acid, asparagine, glutamine, threonine or cysteine. Furthermore, since the recovery of [^3H]phenylalanine and [^3H]glycine was rather low at positions 6 and 16 respectively, it is possible that one of the five preceding amino acids might also be found at these two positions.

Purity of the uteroglobin mRNA preparation

The purity of the uteroglobin-mRNA preparation used in the acellular synthesis of pre-uteroglobin has been supposed to be over 90%. To verify this assumption, the following experiment was performed. After acellular biosynthesis in the presence of radioactively labelled leucine, no specific immunoprecipitation was carried out, and the total protein preparation was analysed by Edman degradation. As may be seen in Fig. 2 (panel Leu*) the radioactive profile was similar to that observed after immunoprecipitation, showing that pre-uteroglobin is the only protein whose synthesis is observed in the presence of this mRNA preparation

Discussion

In common with many other secretory proteins, uteroglobin is synthesized in the form of a precursor protein that contains an *N*-terminal extension. This leader region consists of 21 amino acid residues and starts with a methionine residue that may be the initiator methionine residue. The last amino acid residue is alanine. Most of the amino acid residues have hydrophobic side chains. This *N*-terminal extension, as hypothesized by Milstein *et al.* (1972) and Blobel & Dobberstein (1975*a,b*), may play a role in the attachment of the protein to endoplasmic-reticulum membranes and secretion of the molecule. Owing to the small size of uteroglobin, the signal peptide contributes a major portion to the precursor protein (21 amino acids in a total of about 90 amino acids) differing from many proteins described up to date but resembling honey-bee (*Apis mellifera*) pre-promelittin (Suchanek *et al.*, 1978).

In the process of understanding the mechanism of action of progesterone on the regulation of uteroglobin synthesis, it becomes necessary to delineate structural and regulatory elements in uteroglobin messenger RNA and its precursor. In this respect the elucidation of pre-uteroglobin amino acid sequence is noteworthy. It also offers a good model for the study of steroid-hormone action at the post-transcriptional level, i.e. protein translation, processing and secretion.

We thank Dr. P. Gaye for helpful discussions and for the gift of many amino acid solutions. We thank Mr. P. Atger for his technical assistance. This work was supported by the Institut National de la Recherche Agronomique, Institut National de la Santé et de la Recherche Médicale, Délégation Générale à la Recherche Scientifique et Technique and la Fondation pour la Recherche Médicale Française.

References

- Atger, M. & Milgrom, E. (1977) *J. Biol. Chem.* **252**, 5412–5418
- Beato, M. & Nieto, A. (1976) *Eur. J. Biochem.* **64**, 15–25
- Beier, H. M. (1978) *Biochim. Biophys. Acta* **160**, 189–291
- Blobel, G. & Dobberstein, B. (1975*a*) *J. Cell Biol.* **67**, 835–851
- Blobel, G. & Dobberstein, B. (1975*b*) *J. Cell Biol.* **67**, 852–862
- Bullock, D. W., Woo, S. L. C. & O'Malley, B. W. (1976) *Biol. Reprod.* **15**, 435–443
- Cohen-Solal, M. C. & Bernard, J. L. (1973) *J. Chromatogr.* **80**, 140–143
- Crestfield, A. M., Moore, S. & Stein, W. H. (1963) *J. Biol. Chem.* **238**, 622–667
- Fridlansky, F. & Milgrom, E. (1976) *Endocrinology* **99**, 1244–1251
- Krishnan, R. S. & Daniel, J. C. (1967) *Science* **158**, 490–492
- Milstein, C., Browlee, G. C., Harrison, T. M. & Mathews, M. B. (1972) *Nature (London) New Biol.* **239**, 117–120
- Pisano, J. J., Bronzert, T. J. & Brewer, H., Jr. (1972) *Anal. Biochem.* **45**, 43–59
- Suchanek, G., Kreil, G. & Hermodson, M. A. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 701–704