Proc. Natl. Acad. Sci. USA
Vol. 75, No. 2, pp. 701–704, February 1978
Biochemistry

Amino acid sequence of honeybee prepromelittin synthesized *in vitro*

(precursor/secretory polypeptide/wheat germ extract)

GERDA SUCHANEK*, GÜNTHER KREIL*, AND MARK A. HERMODSON†

* Institute of Molecular Biology, Austrian Academy of Sciences, Billrothstrasse 11, 5020 Salzburg, Austria; and † Division of Medical Genetics, University of Washington, Seattle, Washington 98195

Communicated by Donald F. Steiner, November 14, 1977

ABSTRACT **Translation of melittin messenger RNA from queen bee venom glands in a cell-free system from wheat germ yielded prepromelittin. Sequence analysis of the labeled *in vitro* product was performed by automatic Edman degradation of the intact polypeptide as well as by analysis of some of its proteolytic fragments. Prepromelittin was shown to be composed of 70 amino acids, two of which have not been identified. The sequence of melittin is located in the COOH-terminal third of the polypeptide chain (residues 44–69). Prepromelittin starts with a very hydrophobic pre-region, probably 21 residues long, followed by a pro-part of unusual sequence, containing only alanine, proline, and acidic residues. At least three post-translational reactions are required to convert prepromelittin to melittin.**

Our studies on the biosynthesis of melittin have revealed the existence of two types of precursors. The first type, called promelittin, was detected in the venom glands of honeybees fed with radioactive amino acids (1) and also in frog oocytes injected with crude melittin mRNA from queen bee venom glands (2). When synthesized in frog oocytes, promelittin appears to be a stable end product, while in venom glands it is slowly converted to melittin. A second type of precursor with a molecular weight larger than that of promelittin was detected upon translation of melittin mRNA in cell-free systems (3, 4). This *in vitro* product, which was originally called protomelittin and has now been termed prepromelittin, has not been detectable in venom glands or frog oocytes, presumably because it is rapidly processed to promelittin. Evidence for such short-lived precursors has also been obtained for other secretory polypeptides from the translation of their corresponding mRNAs in cell-free systems. Examples are the precursors for different immunoglobulin light chains (5, 6), parathyroid hormone (7), insulin (8) and several other hormones (9–11), secretory enzymes (12, 13), and serum albumin (14). In all instances it was found that these precursors contain an extra sequence, 17–30 residues long, at the NH_2 terminus with a section rich in hydrophobic residues. Blobel and Dobberstein (15) have proposed that this "signal peptide" somehow allows the growing polypeptide chain to bind to and then penetrate the membrane of the endoplasmic reticulum. The nascent polypeptide is thus delivered into the cisternae of the endoplasmic reticulum and subsequently exported.

In a previous publication we presented the partial characterization of prepromelittin synthesized in a cell-free system derived from mammalian cells (3) and showed that the bulk of the extra residues present in this *in vitro* product are linked to the amino end of promelittin. More recently, prepromelittin has been obtained by translation of melittin mRNA in a cell-free system from wheat germ and its carboxyl-terminal sequence has been determined (4). In this paper it is shown that prepromelittin is composed of 70 amino acids and the complete sequence, except for two unknown residues, is presented.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Tritium-labeled amino acids and [^{35}S]methionine of highest specific activity currently available were obtained from The Radiochemical Centre (Amersham, England) or New England Nuclear (Dreieichenhain, W. Germany). Proteolytic enzymes were purchased from Serva (Heidelberg) or Sigma (St. Louis).

Preparation of Venom Gland RNA. Total RNA from queen bee venom glands was prepared as described (2, 16). For some experiments, the poly(A)-containing fraction isolated by chromatography on poly(dT)-cellulose was used. In this fraction melittin mRNA is the predominant species (16).

Protein Synthesis Assay. Melittin mRNA was translated in a cell-free system from wheat germ (17) as described earlier (4).

Product Analysis. Reaction mixtures from the translation assay were diluted with water and extracted with butanol (4). Denatured proteins and prepromelittin that accumulated at the interface were washed several times with butanol and water and then dried. For the sequenator runs, this fraction was further purified by chromatography over Sephadex G-50 in 9% formic acid. Prepromelittin eluted near the void volume and could thus be separated from radioactive contaminants of lower molecular weight.

Intact prepromelittin labeled with one radioactive amino acid at a time ([^{35}S]methionine or tritiated leucine, isoleucine, phenylalanine, tryptophan, valine, tyrosine, proline, or alanine) was degraded on a Beckman sequenator in the presence of 4 mg of hemoglobin as carrier using a peptide program (18). At each cycle of degradation, the radioactivity present in the organic extract was determined on an aliquot in a toluene-based scintillation fluid.

Peptic and chymotryptic fragments of prepromelittin were prepared and fractionated as described (1, 4). The proteolytic fragments were also analyzed by either automatic or manual Edman degradation. In the latter instances, fragments were repurified by paper electrophoresis at pH 4.8 or 1.8 after one to several steps. This gave additional information about the size and charge of the peptides.

RESULTS

Total RNA from queen bee venom glands or the poly(A)-containing fraction derived from it were translated in a cell-free system from wheat germ in the presence of one radioactive amino acid. The incorporation of [^{3}H]leucine into material precipitable by trichloroacetic acid was 15–60 times higher in the presence of venom gland RNA than in controls and the in-

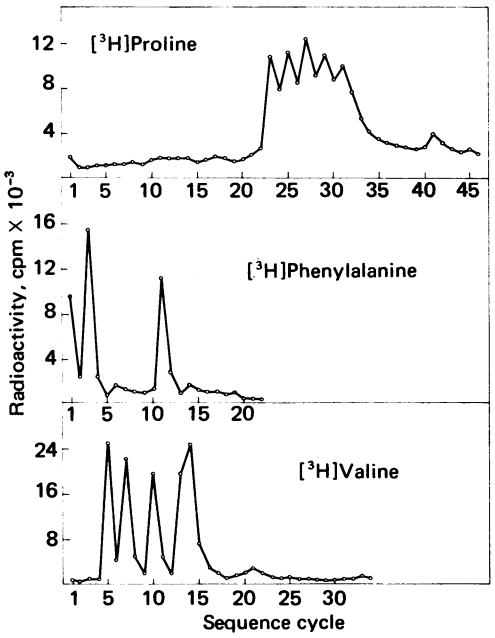

FIG. 1. Sequential Edman degradation of prepromelittin labeled with [³H]phenylalanine, [³H]valine, and [³H]proline, respectively. In the presence of carrier hemoglobin, radioactive prepromelittin was subjected to 35–50 cycles of degradation. The radioactivity released at each step is plotted against the cycle number.

terface material prepared by butanol extraction contained radioactive prepromelittin as the predominant labeled polypeptide.

The amino acid sequence of prepromelittin was determined by automatic Edman degradation of the intact polypeptide and by analysis of some of its proteolytic fragments. The first approach gave sequence information up to about residue 30 of prepromelittin. Sequencer runs of prepromelittin labeled with tritiated valine, phenylalanine, and proline, respectively, are shown as typical examples (Fig. 1). Severe overlap problems were inevitably encountered at the proline-rich segment, but the sequence is easily discernible. In experiments with radioactive phenylalanine, tyrosine, or tryptophan, labeled contaminants were found which gave high yields in the first one or two cycles of the degradation (Fig. 1 *middle*), and the radioactivity from these could not be washed out by cycling the sequenator without phenylisothiocyanate. However, after repeated chromatography on Sephadex G-50 the amount of these as yet unknown radioactive contaminants was considerably reduced. From these sequencer runs of prepromelittin labeled with radioactive phenylalanine, valine, proline (all shown in Fig. 1), methionine, leucine, tyrosine, isoleucine, tryptophan, and alanine (data not shown), residues could be assigned to the following positions: Met^{1}, Phe^{3}, Leu^{4}, Val^{5}, Val^{7}, Ala^{8}, Leu^{9}, Val^{10}, Phe^{11}, Met^{12}, Val^{13}, Val^{14}, Tyr^{15}, Ile^{16}, Tyr^{18}, Ile^{19}, Tyr^{20}, Ala^{21}, Ala^{22}, Pro^{23}, Pro^{25}, Pro^{27}, Pro^{29}, Pro^{31}, and Pro^{41}.

For the second type of approach, peptic fragments of prepromelittin were investigated first. One experiment with [³H]tyrosine is shown in Fig. 2. Peptic hydrolysis of prepromelittin yielded three fragments that contained tyrosine: an acidic fragment P-1 that migrated, upon paper electrophoresis at pH 4.8, about half as fast as glutamic acid and that had tyrosine in position 2, and two neutral tripeptides, P-2 and P-3, each possessing a COOH-terminal tyrosine residue. These latter fragments could be resolved by paper chromatography of the neutral fraction shown in Fig. 2. Experiments with different labeled amino acids showed that the acidic fragment P-1 con-

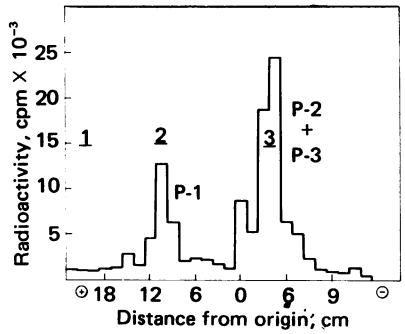

FIG. 2. Fractionation of the peptic digest of prepromelittin labeled with [³H]tyrosine. Interface material (see *Product Analysis*) was digested with pepsin (0.1 mg of enzyme, pH 1.5, 37°, 4 hr) and fractionated by high-voltage paper electrophoresis at pH 4.8 (1% pyridine/acetate buffer). The electropherogram was cut into 1.5-cm sections, and radioactivity of each was determined in a liquid scintillation counter (toluene-based scintillator). Radioactive peaks were eluted with 1% pyridine and used for further characterization. References: (*1*) glutamic acid, (*2*) methionyl-glutamic acid, and (*3*) leucine.

tained, besides tyrosine, the following amino acids: isoleucine, alanine, proline, and glutamic acid. The position of each of these labeled residues was determined by manual edman degradation (Table 1*A*).

In assays with [³⁵S]methionine, [³H]lysine, and [³H]phenylalanine, a basic fragment was detected in the peptic hydrolysates of prepromelittin, which had the sequence Met-Lys-Phe. The results from these experiments with peptic fragments are summarized in Table 1*A*.

After it became clear that the NH_2-terminal region contained three tyrosines and at least one phenylalanine (residues that are not found in promelittin), chymotryptic hydrolysates were also investigated (Table 1*B*). Most important for the further analysis of the amino acid sequence was the observation that these digests contained a large, acidic fragment of 29 residues, which started with residue 3 of P-1 (i.e., residue 21 of prepromelittin) and ended after the sixth residue of melittin. In different experiments this fragment was shown to contain the following amino acids: alanine, proline, glutamic acid, aspartic acid, isoleucine, valine, glycine, and leucine. The sequence of this fragment was determined in part with the Beckman sequenator and in part by manual Edman degradation. The results of au-

Table 1. Fragments characterized from enzymatic digests of prepromelittin labeled with different amino acids

(*A*) Peptic fragments	
P-1	Ile-Tyr-Ala-Ala-Pro-Glu-Pro-Glu-Pro-Ala-Pro-Glu-Pro-Glu-Ala-Glu
P-2	N-N-Tyr*
P-3	N-N-Tyr
P-4	Met-Lys-Phe
(*B*) Chymotryptic fragments	
C-1	Ala-Ala-Pro-Glu-Pro-Glu-Pro-Ala-Pro-Glu-Pro-Glu-Ala-Glu-Ala-Asp-Ala-Glu-Ala-Asp-Pro-Glu-Ala-X-Ile-X-Ala-(Val,Leu)
C-1a†	Ala-Ala-Pro-X-Pro . . .
C-1b†	Pro-Glu-Ala-Gly-Ile-Gly-Ala-Val-Leu
C-2	N-N-Tyr (identical to P-2)
C-3	N-N-N-Tyr
C-4	N-Phe
C-5	Met-Lys-Phe (identical to P-4)

* N stands for neutral amino acid.

† Hydrolysis in dilute HCl of C-1 yielded C-1a and C-1b, which were separated by paper electrophoresis at pH 4.8 (ref. 1)

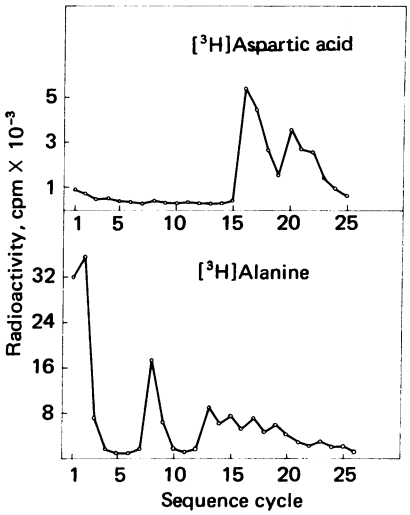

FIG. 3. Sequential Edman degradation of a chymotryptic fragment of prepromelittin. Interface material was digested with chymotrypsin (0.1 mg of enzyme, pH 8.5, 37°, 5 hr) and then fractionated by paper electrophoresis (see legend to Fig. 2). The acidic fragment C-1 (electrophoretic mobility relative to free glutamic acid was about 0.45) was eluted and analyzed by Edman degradation. Sequencer runs of fragment C-1 labeled with [³H]alanine and [³H]aspartic acid are shown as examples.

tomatic Edman degradation of this chymotryptic fragment are shown for the samples labeled with tritiated alanine and aspartic acid, respectively (Fig. 3). Again, large overlaps from the proline-rich segment were encountered, but the interpretation was nevertheless clear.

For the analysis of its COOH-terminal region, fragment C-1 was first cleaved in dilute HCl (10 mM HCl, 95°, 90 min). This treatment had previously been used to hydrolyze the Asp-Pro bond in honeybee promelittin (1). This procedure split fragment C-1 into two peptides C-1a and C-1b. In experiments with different labeled amino acids it could be shown that the sequence of C-1b was Pro-Glu-Ala-Gly-Ile-Gly-Ala-Val-Leu (Table 1*B*). This yielded the overlap into the sequence of melittin, which starts at the first glycine residue (Table 1*B*).

From the results presented in this communication together with the known sequences of melittin (19) and of the carboxyl end of prepromelittin (4), an almost complete amino acid sequence can now be presented. We conclude that prepromelittin is a polypeptide composed of 70 residues (Fig. 4) in which the sequence of melittin resides in the carboxyl-terminal region (residues 44–69 of prepromelittin). Only the identity of residues 6 and 17 in the NH_2-terminal part remains to be established.

DISCUSSION

The experiments presented in this paper demonstrate that prepromelittin, synthesized in a cell-free system from wheat germ programmed with melittin mRNA, is a polypeptide composed of 70 amino acids. It contains an amino-terminal methionine which is probably the initiating residue. In several instances it has been demonstrated that the NH_2-terminal methionine of presecretory polypeptides is the initiator (see e.g., ref. 20). As discussed previously (4) it is also quite likely that the carboxyl-terminal glycine represents the true COOH-end and is not an artifact of the cell-free system arising, for example, through premature chain termination. Accordingly, we suggest that prepromelittin is the primary translation product and that these 70 residues constitute the entire coding capacity of melittin mRNA.

From the amino acid sequence of prepromelittin three regions can be readily distinguished: a very hydrophobic amino-terminal part, a middle region which contains all the acidic and most of the proline residues, and, finally, the terminal region which contains the melittin sequence plus an extra glycine at the COOH-end.

In the earlier work with worker bees fed with radioactive amino acids, several species of promelittin were observed with six to nine amino acids linked to the amino end of melittin (1). The amino acid sequence proposed for worker bee promelittin several years ago (1) differs at three sites from the sequence now found for prepromelittin. For the former, the sequence was presented to be N-*Glu-Pro*-Glu-*Pro*-Asp-Pro-Glu-Ala-melittin (N stands for an unknown neutral amino acid), which contrasts with the sequence Ala-*Asp-Ala*-Glu-*Ala*-Asp-Pro-Glu-Ala determined in the present work (residues 35–43 of prepromelittin, differences are italicized). The alternating pattern of neutral and acidic residues was correctly established in this early work. However, it is now apparent that the fragments of worker bee promelittin containing labeled glutamic acid or proline (alanine was not incorporated in these *in vivo* experiments) were radiochemically impure, and this may have led to wrong assignments.

In order to clarify these discrepancies, we have, more recently, been analyzing queen bee promelittin, which can be obtained in larger quantities. The amino-terminal sequence of this peptide was found to be: Ala-Pro-X-Pro-X-Pro-Ala-Pro-X-Pro-X . . . (21). This indicates that queen bee promelittin starts at Ala^{22} of prepromelittin and contains a pro-part of 22 amino acids. The relationship between queen bee promelittin and the smaller species observed in worker bees is at present not clear. It is conceivable that the conversion of promelittin to melittin proceeds via intermediates, and some of these may be detectable in worker bee venom glands where venom production is much slower than in queen bees (22).

These recent results suggest that the pre-part of prepromelittin corresponds to the first 21 residues, at least 18 of which are hydrophobic. A similarly high lipophilicity appears to be characteristic for the NH_2-terminal extensions found in other presecretory polypeptides (6–14). According to the "signal hypothesis" (15), this region plays a crucial role in the attachment of the growing polypeptide chain to the membrane of the endoplasmic reticulum and its vectorial discharge into the lumen. However, apart from the common hydrophobic character there exists no discernible sequence homology between the pre-peptide of prepromelittin and those of other presecretory precursors.

10 20 30
NH_2·Met-Lys-Phe-Leu-Val- X -Val-Ala-Leu-Val-Phe-Met-Val-Val-Tyr-Ile- X -Tyr-Ile-Tyr-Ala-Ala-Pro-Glu-Pro-Glu-Pro-Ala-Pro-Glu-Pro-Glu-
40 50 60
Ala-Glu-Ala-Asp-Ala-Glu-Ala-Asp-Pro-Glu-Ala-*Gly-Ile-Gly-Ala-Val-Leu*-Lys-Val-Leu-Thr-Thr-Gly-Leu-Pro-Ala-Leu-Ile-Ser-Trp-Ile-Lys-
70
Arg-Lys-Arg-Gln-Gln-Gly·COOH

FIG. 4. Proposed amino acid sequence of prepromelittin. The two unidentified residues at positions 6 and 17 are designated by X. The sequence from residues 44 to 69 is that determined for melittin (19); the COOH-terminal sequence is taken from ref. 4. Residues in the amino-terminal part of melittin independently assigned by the present experiments are *italicized*.

The generation of melittin from prepromelittin must involve at least three post-translational reactions, presumably in the following temporal sequence: (*i*) cleavage of the Ala^{21}-Ala^{22} bond to remove the pre-peptide, (*ii*) formation of the COOH-terminal amide with concomitant loss of one glycine residue, and (*iii*) cleavage of the Ala^{43}-Gly^{44} bond to yield melittin. The two cleavage reactions may in turn involve one or more splits by endo- and/or exopeptidases.

The most interesting reaction appears to be the first one. The enzyme(s) catalyzing this hydrolysis is probably an essential component of the secretory pathway. That this enzyme is widely distributed is exemplified by the fact that translation of melittin mRNA in *Xenopus* oocytes yielded promelittin, while prepromelittin could not be detected (2). Also, correct cleavage of the pre-part from prepromelittin has been observed *in vitro* with subcellular fractions from rat liver (R. Kaschnitz and G. Kreil, unpublished experiments). A lack of species specificity has also been demonstrated in other cases, where such precursors were in *statu nascendi* correctly processed by microsomal membranes of diverse origin (23–25). This is all the more surprising since the amino acids adjacent to the susceptible bond vary considerably in different presecretory polypeptides. We are thus faced with the intriguing problem as to how an intracellular protease correctly cleaves different polypeptides at sites that have no similarity in primary structure. Since the processing of prepromelittin can be observed *in vitro*, the mechanism of this first cleavage should be amenable to further experimentation.

Note Added in Proof. Recently, melittin mRNA was translated in the cell-free system from wheat germ in the presence of 20 unlabeled amino acids and rabbit reticulocyte initiator tRNA charged with [^{35}S]methionine (the charged tRNA was a kind gift of D. Chung and E. W. Davie of the University of Washington, Seattle). In this experiment, 15,000 cpm were incorporated into the amino-terminal methionine and less than 1000 into the internal methionine (residue 12). This demonstrates that Met^1 of prepromelittin is the initiating residue.

We thank Mrs. Lela Frasel and Dr. Ingela Kindås-Mügge for the supply of venom gland RNA.

1. Kreil, G. (1973) *Eur. J. Biochem.* **33,** 558–566.
2. Kindås-Mügge, I., Lane, C. D. & Kreil, G. (1974) *J. Mol. Biol.* **87,** 451–462.
3. Suchanek, G., Kindås-Mügge, I., Kreil, G. & Schreier, M. H. (1975) *Eur. J. Biochem.* **60,** 309–315.
4. Suchanek, G. & Kreil, G. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 975–978.
5. Milstein, C., Brownlee, C. G., Harrison, R. M. & Matthews, M. B. (1972) *Nature New Biol.* **239,** 117–120.
6. Burstein, Y. & Schechter, I. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 716–720.
7. Kemper, B., Habener, J. F., Mulligan, R. C., Potts, J. T. & Rich, A. (1974) *Proc. Natl. Acad. Sci. USA* **71,** 3731–3735.
8. Chan, S. J., Keim, P. & Steiner, D. F. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 1964–1968.
9. Sussman, P. M., Tushinski, R. J. & Bancroft, F. C. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 29–33.
10. Maurer, R. A., Gorski, J. & McKean, D. J. (1977) *Biochem. J.* **161,** 189–192.
11. Boime, I., Boguslawski, S. & Caine, J. (1975) *Biochem. Biophys. Res. Commun.* **62,** 103–109.
12. Devillers-Thiery, A., Kindt, T., Scheele, G. & Blobel, G. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 5016–5020.
13. Thibodeau, S. N., Gagnon, J. & Palmiter, R. (1977) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **36,** Abstr. no. 2030.
14. Strauss, A. W., Donohue, A. M., Bennett, C. D., Rodkey, J. A. & Alberts, A. W. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 1358–1362.
15. Blobel, G. & Dobberstein, B. (1975) *J. Cell Biol.* **67,** 835–851.
16. Kindås-Mügge, I., Frasel, L. & Diggelmann, H. (1976) *J. Mol. Biol.* **105,** 177–181.
17. Marcu, K. & Dudock, B. (1974) *Nucleic Acids Res.* **1,** 1385–1397.
18. Hermodson, M. A., Schmer, G. & Kurachi, K. (1977) *J. Biol. Chem.*, **252,** 6276–6279.
19. Habermann, E. & Jentsch, J. (1967) *Hoppe-Seyler's Z. Physiol. Chem.* **348,** 37–50.
20. Kemper, B., Habener, J. F., Ernst, M. D., Potts, J. T. & Rich, A. (1976) *Biochemistry* **15,** 20–25.
21. Kreil, G. (1977) *11th FEBS Meeting (Copenhagen)*, Abstr. no. A6-2/L2.
22. Bachmayer, H., Kreil, G. & Suchanek, G. (1972) *J. Insect Physiol.* **18,** 1515–1521.
23. Habener, J. F., Kemper, B., Potts, J. T., Jr. & Rich, A. (1975) *Biochem. Biophys. Res. Commun.* **67,** 1114–1121.
24. Birken, S., Smith, D. L., Canfield, R. E. & Boime, I. (1977) *Biochem. Biophys. Res. Commun.* **74,** 106–112.
25. Shields, D. & Blobel, G. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 2059–2063.