Glucagon is a 29-amino acid peptide hormone that regulates blood glucose concentrations. It is a member of a family of structurally related hormones that includes in addition to glucagon, vasoactive intestinal peptide, gastrin inhibitory peptide, and secretin. Like other peptide hormones, glucagon is synthesized as a larger precursor. Previously, we reported the amino acid sequence of an anglerfish pre-proglucagon of apparent $M_r = 14,500$, derived from the sequence of cloned cDNAs. Now we have determined the nucleotide sequence of cDNAs encoding a separate anglerfish pre-proglucagon of apparent $M_r = 12,500$ and have derived the complete amino acid sequence of this second precursor. The configurations of the two pre-proglucagons are similar. Each pre-proglucagon is a polypeptide that contains two glucagon-related peptides arranged in tandem, a 29-amino acid glucagon sequence and a 34-amino acid sequence which shows homology to glucagon and the other members of the glucagon family. These two peptides are linked in the precursors by lysine-arginine and intervening pentapeptides. The glucagon sequences of 29 amino acids in the two precursors are closely homologous. Similarly, the 34-amino acid peptide sequences in the two precursors are highly homologous. Analyses of the genomic DNA prepared from the spleen of a single anglerfish show that these two pre-proglucagons are encoded by at least two separate genes. Analyses of the mRNAs in the anglerfish islets indicate that a single mRNA species encodes a $M_r = 14,500$ pre-proglucagon but suggest that two separate mRNAs contain coding sequences for the $M_r = 12,500$ pre-proglucagons. These studies indicate that in the anglerfish, glucagon is synthesized by way of the expression of at least two genes.

Glucagon is a hormone of 29 amino acids (molecular weight, 3,500) synthesized and secreted by "A" cells in the pancreatic islets (1) and involved in the regulation of plasma glucose concentrations (2). The hormone is a member of a family of structurally related hormones which includes gastrin inhibitory peptide, vasoactive intestinal peptide, secretin, and ghrelin (3). Analyses of labeled, newly synthesized proteins in intact islets have shown that glucagon is synthesized as a large precursor (4-6). Recently, we (7, 8) and Shields et al. (9) reported that islets of the anglerfish (Lophius americanus) contain poly(A) RNAs which direct the synthesis in cell-free systems of at least two precursors of glucagon (pre-proglucagons) of apparent molecular weights of 14,500 and 12,500. Hybridizations with islet poly(A) RNAs and cloned cDNAs prepared from total islet poly(A) RNAs demonstrated that the two pre-proglucagons are encoded by separate mRNAs (8). We previously described the cloning and nucleotide sequences of cDNAs encoding the $M_r = 14,500$ pre-proglucagon (10) and have now sequenced DNA encoding the $M_r = 12,500$ pre-proglucagon and here compare the nucleotide sequence and derived amino acid sequence of the precursor to those of the $M_r = 14,500$ pre-proglucagon. We also find by genomic blotting that the two precursors are encoded by at least two separate genes.

**EXPERIMENTAL PROCEDURES**

**Isolation of Recombinant Plasmids Containing cDNA Encoding Precursors of Anglerfish Islet Glucagons**—The details of the preparation of a cDNA library using the vector plasmid pBR322 and the host Escherichia coli (strain, 1776) and poly(A) RNA prepared from the islets of the anglerfish L. americanus have been described previously (11). Initial screening of this cDNA library by hybridization arrest and hybridization selection and cell-free translations provided two cloned recombinant plasmids containing cDNAs encoding the two pre-proglucagon precursors of $M_r = 14,500$ and 12,500 (8). These two recombinant plasmids, labeled by nick translation (12) with $[^{32}P]$dCTP, were used to screen by hybridization (13) 1,800 bacterial clones containing recombinant plasmids prepared from the anglerfish poly(A) RNA. Thereby, we identified 31 and 11 bacterial colonies containing coding sequences for anglerfish islet pre-proglucagon I and II, respectively.

**Determination of the Nucleotide Sequences of cDNAs Encoding the Two Pancreatic Islet Pre-proglucagons**—Plasmids were isolated from several of the clones described above using the cleared lysate technique (14) followed by two successive centrifugations in cesium chloride/ethidium bromide gradients. Two recombinant plasmids encoding the smaller and three encoding the larger of the two separate pre-proglucagons were selected for nucleotide sequence analysis. Complete nucleotide sequences of both the sense (coding) and nonsense (noncoding) strands were determined by the chemical sequencing method of Maxam and Gilbert (15). We have previously reported the nucleotide sequence of the cDNA encoding the anglerfish islet pre-proglucagon I (10).

**Analyses of Pre-proglucagon Coding Sequences in Genomic DNA**—Aliquots of DNA prepared from the spleen of a single anglerfish (16) were digested individually to completion with the restriction endonucleases EcoRI, BamHI, and HindIII. Ten micrograms of the digested DNA were separated by electrophoresis on 0.8% agarose gels and transferred to nitrocellulose (17). Duplicate nitrocellulose filters prepared from the same gel were hybridized individually with recombinant plasmids containing cDNAs encoding anglerfish pre-proglucagon I and pre-proglucagon II that had been nick translated with $[^{32}P]$dCTP (12). The conditions of hybridization were 42 °C for 18 h in the presence of 5× SSC (1× SSC is 0.15 M NaCl, 0.015 M Na citrate, pH 7.0), 50 mM sodium phosphate, pH 7.0, 50% formamide, and 10% dextran sulfate. Autoradiograms (72-h exposure) were prepared from the nitrocellulose filters following hybridization.

**Analyses of mRNAs Encoding Pre-proglucagons by Filter Hybridization (Northern Blot Analysis)**—Aliquots of the poly(A) RNA used for the preparation of the cDNA libraries were analyzed by electrophoresis on 1% agarose gels. The RNAs were transferred from
the gels to Gene Screen (New England Nuclear) by capillary transfer (17). Duplicate blots of the RNA on Gene Screen were hybridized individually to 32P-labeled cDNAs encoding the pancreatic pre-proglucagons I and II. Hybridizations were for 16 h at 42 °C in the presence of 6 × SSC, 50 mM Tris, pH 7.5, 0.1% sodium dodecyl sulfate, 10 × Denhardt's reagent (1 × Denhardt's reagent is 0.02% Ficoll 400, 0.02% bovine serum albumin, 0.02% polyvinylpyrrolidone 40), 50% formamide, and 0.01% sonicated denatured salmon sperm DNA.

RESULTS AND DISCUSSION

The sequences of the cDNAs encoding the M1 = 14,500 (AFG I') and M2 = 12,500 (AFG II) pre-proglucagons reveal coding sequences of 124 (AFG I) and 122 (AFG II) codons, beginning with ATG initiator methionine codons and flanked by 5' and 3' untranslated regions (Fig. 1). The overall configuration of the two pre-proglucagons is similar. The glucagon sequences reside in the midregion of each precursor with peptide extensions at the NH2 and COOH termini. At the NH2 terminus of each precursor, following the initiator methionine, are sequences of predominantly hydrophobic amino acids characteristic of signal or leader sequences found in the precursors of other secreted hormones and proteins (18, 19). Following the putative signal sequences in the two precursors (Fig. 1) are, in succession, an NH2-peptide or prosequence, a Lys-Arg, a 29-amino acid glucagon sequence, a Lys-Arg, a tetra- or pentapeptide, a Lys-Arg, and a 34-amino acid COOH-terminal peptide which is highly homologous to glucagon and to the other hormones of the glucagon family (Fig. 2). The paired basic residues flanking the peptides are characteristic of sites in prohormones that are cleaved in the post-translational formation of secreted hormones (20). Thus, post-translational processing of the proglucagons at these sites would result in the formation of four peptides from each precursor: an NH2-peptide, a 29-amino acid glucagon, a short intervening peptide, and a 34-amino acid glucagon-related COOH peptide.

The 29-amino acid glucagon sequences in the precursors are highly homologous to each other (89%) and to mammalian glucagon (Fig. 2). Likewise, the 34-amino acid glucagon-related COOH peptides in the precursors are highly homologous to each other (74%) (Fig. 1). These homologies indicate evolutionary pressure to conserve these sequences and are consistent with a specific biologic role of the four glucagon-related peptides. Predictions of the biologic role of the glucagon-related COOH-terminal peptides are speculative at this time. However, a gastric inhibitory peptide-like immunoreactive has recently been demonstrated in rat pancreatic A cells in the same secretory granules as glucagon (21, 22). The anglerfish glucagon-related COOH peptides show considerable homology with mammalian gastric inhibitory peptide (Fig. 2) and may represent the counterpart of the mammalian gastric inhibitory peptide-like immunoreactive in the fish. On the other hand, the glucagon-related COOH peptide of the AFG II pre-proglucagon has the sequence Ala-Gly-Arg-Gly-Arg at the COOH terminus of the glucagon precursor (Fig. 2). This finding demonstrates that the two pre-proglucagon-encoding mRNAs are transcribed from different genes in a single individual. Surprisingly, despite an overall 74% homology in the nucleotide sequences of the two cDNAs, we observed no cross-hybridization of the cDNAs with the genomic fragments under these stringent (42 °C) conditions of hybridization (Fig. 1). In preliminary studies, however, we observed weak cross-hybridization of the two cDNAs to the corresponding genomic restriction fragments under conditions of lower stringency of hybridization (37 °C). Although our findings demonstrate that there are at least two different glucagon genes in the anglerfish, they do not exclude the possibility that there are additional glucagon genes which do not hybridize to either the AFG I or AFG II cDNAs.

To determine the sizes and the complexity of the messenger RNAs encoding the two pre-proglucagons, we analyzed anglerfish islet poly(A) RNA by electrophoresis on agarose gels, followed by transfer of the RNA to filters and hybridization with the two 32P-labeled glucagon cDNAs (Northern blot analysis) (Fig. 4). By this procedure, a single mRNA species of 650 bases was identified by hybridization with the probe containing the coding sequence for anglerfish glucagon I. Two separate bands representing messenger RNAs of 630 and 670 bases were detected using the probe encoding the anglerfish glucagon II. As was found with the genomic blotting, there was no significant cross-hybridization between the two glucagon-cDNA hybridization probes. Thus, three separate mRNAs encoding the pre-proglucagons appear to exist. The two separate mRNAs encoding the anglerfish pre-proglucagon may arise from separate genes encoding this particular glucagon precursor, by utilization of differential splicing mechanisms during the maturation of a single pre-mRNA to the mature mRNAs, or by variations in lengths of poly(A) tracts at the 3' end of the mRNAs.

Our observations indicate that fish islet pre-proglucagons are polyproteins (proteins containing more than one peptide in the same precursor) as are precursors to several other hormones (23-25). Synthesis of polyprotein hormonal precursors may allow coordinate synthesis of multiple bioactive hormones in the same cell. Alternatively, tissue-specific cleavages of polyproteins may produce different biologically active peptides from the same precursor, as demonstrated for the pro-angiogenin (25). Based on our data on the configuration of the fish glucagon precursors and immunologic evidence about the mammalian glucagon precursors and products, it is tempting to speculate on the potential biologic importance of the processing of mammalian glucagon
pre-proglucagon cDNAs

Pre-proglucagon cDNAs

Fig. 1. Nucleotide and corresponding deduced amino acid sequences of cDNAs encoding two anglerfish pre-proglucagons (AFGI I (10) and AFGI I). The cDNA sequences shown were obtained by complete sequencing of both strands of each of five separate cDNAs. Dots between the nucleotide sequences indicate identical nucleotides (two single codon gaps shown as dashes were introduced into the AFGI I sequence to maximize homology). The amino acid sequences of the two precursors, derived from the nucleotide sequences of the cDNAs, are shown above the nucleotide sequences. Boxed regions indicate regions of identity of the amino acid sequences in the two precursors. Arrows represent predicted prohormone cleavage sites (heavy arrows predict sites of cleavage by trypsin-like activity; light arrows predict sites of cleavage by carboxypeptidase B-like activity). Horizontal lines delineate peptides formed by potential cleavages at the arrowed sites. The line indicating the NHz-terminal peptide is dashed to indicate uncertainty as to the site of cleavage of the signal sequence. The underlined DNA sequences AATAAA and AATTAAA at the 3' ends of AFGI I and AFGI I cDNAs, respectively, are characteristic of sites involved in addition of the poly(A) tract to eukaryotic mRNAs (31, 32).

precursors. Mammalian glucagon precursors in islets (5, 6, 26), intestine (27, 28), and brain (29) appear to show a similar configuration to the fish precursors with glucagon in the midregion and peptide extensions at the NH2 and COOH terminals. Glucagon, an NH2-terminal fragment of the glucagon precursor, and a gastric inhibitory peptide-like immunoreac-
tant have all been reported to exist in mammalian pancreatic A cells (21, 22, 26, 30). These observations indicate that the mammalian counterpart of the fish glucagon precursor may be processed completely in mammalian pancreatic A cells to release multiple peptides: glucagon, the COOH-terminal peptide, an NH2-peptide, and the short intervening peptide. In contrast, intestinal “L” cells are reported to contain a gastric inhibitory peptide-like immunoreactant (21, 22, 30) and a large glicentin-like molecule (27, 28), consisting of glucagon covalently linked to a long NH2-terminal extension and a short COOH-terminal extension. This situation suggests that incomplete processing of the mammalian counterpart of the fish precursors in intestinal cells may release only the COOH-terminal peptide and the remaining prohormonal fragment.
Little is known about the biosynthesis of other members of the glucagon family of hormones, gastric inhibitory peptide, vasoactive intestinal peptide, secretin, and glicentin. Our observations, however, about the nature of the glucagon precursors raise the possibility that other members of the glucagon family may arise from similar polyprotein precursors.

Acknowledgments—We thank Jeanne Sullivan and Eileen Morris for help in the preparation of this manuscript and Rex Bentley for laboratory assistance.

REFERENCES

1. Unger, R. H., and Orci, L. (1981) *N. Engl. J. Med.* 304, 1518-1524
2. Falhmer, S. J., and Patert, G. J. (1972) *Handb. Physiol.* 1, 1-23
3. Dockray, G. J. (1977) *Gastroenterology* 73, 344-348
4. Noe, B. D., and Bauer, G. E. (1975) *Endocrinology* 97, 868-877
5. Patzelt, C., Tager, H. S., Carroll, R. J., and Steiner, D. F. (1979) *Nature (Lond.)* 279, 514-515
6. Tager, H. S., Patzelt, C., Assoian, R. K., Chan, S. J., Duguid, J. R., and Steiner, D. F. (1980) *Ann. N. Y. Acad. Sci.* 343, 133-147
7. Lund, P. K., Goodman, R. H., Jacobs, J. W., and Habener, J. F. (1980) *Diabetes* 29, 583-586
8. Lund, P. K., Goodman, R. H., and Habener, J. F. (1981) *J. Biol. Chem.* 256, 6515-6518
9. Shields, D., Warren, T. G., Roth, S. E., and Brenner, M. J. (1981) *Nature (Lond.)* 289, 511-514
10. Lund, P. K., Goodman, R. H., Dee, P. C., and Habener, J. F. (1982) *Proc. Natl. Acad. Sci. U. S. A.* 79, 345-349
11. Goodman, R. H., Jacobs, J. W., Chin, W. W., Lund, P. K., Dee, P. C., and Habener, J. F. (1980) *Proc. Natl. Acad. Sci. U. S. A.* 77, 5869-5876
12. Rigby, P. W. J., Diekmann, M., Rhodes, C., and Berg, P. (1977) *J. Mol. Biol.* 113, 237-251
13. Gronstein, M., and Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. U. S. A.* 72, 3961-3965
14. Clewell, D. B., and Helsinki, D. R. (1970) *Biochemistry* 9, 4286-4440
15. Maxam, A., and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U. S. A.*, 74, 560-564
16. Blin, N., and Stafford, D. W. (1976) *Nucleic Acids Res.* 3, 2303-2308
17. Southern, E. M. (1975) *J. Mol. Biol.* 98, 503-517
18. Blobel, G., and Dobberstein, B. (1975) *J. Cell Biol.* 67, 835-851
19. Blobel, G., Walter, P., Chang, C. N., Goldman, B., Erickson, A. H., and Lingappa, V. R. (1979) *Secretory Mechanisms: Society of Experimental Biologists Symposium* 33, (Hopkins, C. R., and Duncan, C. J., eds) pp. 9-36, Cambridge University Press, England
20. Steiner, D. F., Quinn, P. S., Chan, S. J., Marsh, J., and Tager, H. S. (1980) *Ann. N. Y. Acad. Sci.* 343, 1-16
21. Smith, P. H., Merchant, F. W., Johnson, D. G., Fujimoto, W. Y., and Williams, R. H. (1977) *Am. J. Anat.* 149, 585-590
22. Ahren, B., Hakanson, R., Lundquist, I., Sjoheder, K., and Sundler, F. (1981) *Acta Physiol. Scand.* 112, 233-242
23. Habener, J. F., Lund, P. K., Jacobs, J. W., Dee, P. C., and Goodman, R. H. (1980) in *Peptides: Synthesis, Structure, Function: Proceedings of the Seventh American Peptide Symposium* (Rich, D. H., and Gross, E., eds) pp. 457-469, Pierce Chemical Co., Rockford, Illinois
24. Land, H., Schultz, G., Schmähe, H., and Richter, D. (1982) *Nature (Lond.)* 295, 299-303
25. Roberts, J. L., Phillips, M., Rosa, P. A., and Herbert, E. (1978) *Biochemistry* 17, 3699-3618
26. Moody, A. J., Holst, J. J., Thin, L., and Lindkaer-Jensen, S. (1981) *Nature (Lond.)* 289, 514-515
27. Thim, L., and Moody, A. J. (1981) *Regul. Pept.* 2, 139-150

**FIG. 2.** Comparison of the amino acid sequences of the anglerfish glucagons and COOH-terminal glucagon-related peptides with the sequences of mammalian glucagon (34) and *gastric inhibitory peptide* (GIP) (33). Bars indicate identity of residues in the anglerfish peptides and mammalian glucagon. Circles around residues indicate identity of amino acids in the anglerfish COOH-terminal glucagon-related peptides and mammalian gastric inhibitory peptide, and squares indicate different amino acids which are conservative changes from gastric inhibitory peptide (34).

**FIG. 3.** Autoradiogram of Southern hybridization of anglerfish spleen genomic DNA with *32*P-labeled glucagon cDNA containing plasmids which contain the sequences AFG I and AFG II (Fig. 1). The DNA was subjected to complete digestion with restriction endonucleases *Bam*HI (B), *Eco*RI (E), and *Hind*III (H) (see "Experimental Procedures"). Duplicate filters were hybridized with *32*P-labeled glucagon cDNA probes, AFG I and AFG II. *Left*, AFG I, lanes B, E, and H represent fragments of genomic DNA in *Bam*HI, *Eco*RI, and *Hind*III restriction digests, respectively, which hybridized to *32*P-labeled AFG I cDNA. *Right*, AFG II, lanes B, E, and H represent fragments in the same restriction digests which hybridized to *32*P-labeled AFG II cDNA. *m* (left and right) represents molecular weight markers from *Eco*RI digests of *λ* DNA (arrows 1-5 point to DNA fragments of 7.5, 5.9, 5.5, 4.8, and 3.4 kilobases, respectively). Exposure of the autoradiogram was for 72 h.

**FIG. 4.** Autoradiogram of a Northern blot showing mRNAs in anglerfish islets that hybridize with *32*P-labeled cloned cDNAs, AFG I and AFG II. Aliquots of islet mRNA and DNA molecular weight markers were fractionated by electrophoresis on a 1.4% agarose gel and transferred to Gene Screen (New England Nuclear). The blots were hybridized to *32*P-labeled (nick-translated) cDNAs encoding the two pre-proglucagons. Lane AFG I, a single mRNA band of 650 nucleotides that hybridized to AFG I cDNA; lane AFG II, two mRNA bands of 670 and 630 nucleotides that hybridized to AFG II cDNA; lane M, DNA molecular weight markers. Exposure of the autoradiogram was 16 h.
Pre-proglucagon cDNAs

28. Tager, H. S., and Markese, J. (1979) J. Biol. Chem. 254, 2229-2233
29. Tager, H., Hohenboken, M., Markese, J., and Dinersteen, R. J. (1980) Proc. Natl. Acad. Sci. U. S. A. 77, 6229-6233
30. Alumets, J., Hakanson, R., O'Dorsisco, T., Sjoland, K., and Sundler, F. (1978) Histochemistry 58, 253-257
31. Proudfoot, N. J., and Brownlee, G. G. (1976) Nature (Lond.) 263, 211-214
32. Hobart, P., Crawford, R., Shen, L., Pictet, R., and Rutter, W. J. (1980) Nature (Lond.) 288, 137-141
33. Jornvall, H., Carlquist, M., Kwauk, S., Otte, S. C., McIntosh, C. H. S., Brown, J. C., and Mutt, V. (1981) FEBS Lett. 123, 295-298
34. Dayhoff, M. O. (1972) Atlas of Protein Sequence and Structure pp. 89-99, National Biomedical Research Foundation, Washington, D. C.
Anglerfish islet pre-proglucagon II. Nucleotide and corresponding amino acid sequence of the cDNA.
P K Lund, R H Goodman, M R Montminy, P C Dee and J F Habener

*J. Biol. Chem.* 1983, 258:3280-3284.

Access the most updated version of this article at http://www.jbc.org/content/258/5/3280.citation

Alerts:

* When this article is cited
* When a correction for this article is posted

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

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/258/5/3280.citation.full.html#ref-list-1