Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 345–349, January 1982
Biochemistry

Pancreatic preproglucagon cDNA contains two glucagon-related coding sequences arranged in tandem

(recombinant DNA/gastric inhibitory peptide)

P. KAY LUND, RICHARD H. GOODMAN, PHILLIP C. DEE, AND JOEL F. HABENER

Laboratory of Molecular Endocrinology, Massachusetts General Hospital, and Howard Hughes Medical Institute Laboratories, Harvard Medical School, Boston, Massachusetts 02114

Communicated by Alexander Rich, September 29, 1981

ABSTRACT **We have constructed and cloned in bacteria recombinant plasmids containing DNA complementary to the mRNA encoding a pancreatic preproglucagon (M_r 14,500), a product of cell-free translation of angler fish islet mRNAs shown previously by immunoprecipitation analyses to be a precursor of glucagon. cDNAs of 630, 180, and 120 base pairs were isolated and correspond to most of the mRNA for the preproglucagon (650 bases). The cDNAs contain a protein coding sequence of 372 nucleotides and 5′- and 3′-untranslated regions of 58 and 206 nucleotides, respectively. From the coding sequence of the cDNAs, we find that the sequence of glucagon, identical to mammalian glucagon in 20 of 29 positions, resides in the preproglucagon of 124 amino acids flanked by NH_2- and COOH-peptide extensions of 52 and 43 amino acids, respectively. The peptide extensions are linked to the glucagon by Lys-Arg sequences characteristic of the sites that are cleaved during the posttranslational processing of prohormones. Notable is the finding that, following the initial Lys-Arg sequence in the COOH-peptide extension is a pentapeptide, Ser-Gly-Val-Ala-Glu, followed by another Lys-Arg and a sequence of 34 residues that shows striking homology with glucagon and the other peptides of the glucagon family—gastric inhibitory peptide, vasoactive intestinal peptide, and secretin. Thus, the preproglucagon mRNA contains two glucagon-related coding sequences arranged in tandem. The finding of Lys-Arg sequences flanking the glucagon and glucagon-related sequences suggests that these two peptides and a pentapeptide are formed *in vivo* by posttranslational cleavages of a common precursor.**

Glucagon is a 29-amino acid peptide hormone ($M_r \approx 3500$) produced in the A cells of the pancreatic islets (1). The hormone belongs to a multigene family of structurally related peptides that include secretin, gastric inhibitory peptide, vasoactive intestinal peptide, and glicentin (2). These peptides variously regulate carbohydrate metabolism, gastrointestinal motility, and secretory processes (1, 3). In addition to their location in the islets and the gastrointestinal tract, substances that have glucagon-like immunoreactivity have been found in the central nervous system where they may be involved in neuroregulation (4). The principal recognized actions of pancreatic glucagon, however, are to promote glycogenolysis and gluconeogenesis, resulting in an elevation of blood sugar levels. In this regard, the actions of glucagon are counter regulatory to those of insulin and may contribute to the hyperglycemia that accompanies diabetes mellitus (5).

Interest has focused on studies of the biosynthesis of the glucagon-related peptides because of their widespread distribution in different tissues and their importance in the regulation of carbohydrate metabolism and possible functions as neuropeptides. Reports of analyses of newly labeled proteins in mammalian (6, 7), avian (8), and fish (9) islets indicate that pancreatic glucagon is synthesized in the form of a large precursor (8000–18,000 daltons). Structural analyses of glicentin, a 69-amino acid peptide isolated from porcine intestine, revealed the structure of glucagon with both NH_2- and COOH-terminal extensions (10). Earlier, Tager and Steiner (11) characterized a pancreatic proglucagon fragment that consists of glucagon with a COOH-terminal extension of eight amino acids.

Recently, we (12, 13) and Shields *et al.* (14) identified at least two preprohormonal forms of glucagon (M_r 14,500 and 12,500) among the products of the cell-free translation of mRNA prepared from angler fish islets. These precursors appear to be encoded by separate mRNAs (12). We have used the angler fish islet mRNAs as templates to prepare cDNAs and have constructed and cloned in bacteria recombinant plasmids containing these cDNAs. In this paper, we report the sequence of cDNAs that correspond to most (630 base pairs) of the mRNA (650 nucleotides) encoding the preproglucagon of M_r 14,500. The amino acid sequence was deduced from the nucleotide sequence. A surprising feature of the M_r 14,500 preproglucagon was the finding of two glucagon-related sequences arranged in tandem linked by two Lys-Arg sequences and an intervening pentapeptide. Inasmuch as short sequences of basic residues are typically found at sites of cleavage during posttranslational processing of prohormones (15), the structure of the preproglucagon suggests that, during protein synthesis *in vivo*, two glucagon-like peptides would be formed from a common precursor.

MATERIALS AND METHODS

DNA polymerase I (large fragment) and restriction endonucleases were from New England BioLabs. Lysozyme was from Sigma. Terminal transferase and polynucleotide kinase were from Boehringer Mannheim. Cordycepin 5′- [^{32}P]triphosphate (500 Ci/mmol; 1 Ci = 3.7×10^{10} becquerels) and [γ-^{32}P]ATP (2273 Ci/mmol) were from New England Nuclear. *Escherichia coli* strain χ1776 and plasmid pBR322 were provided by H. M. Kronenberg. Islets from the angler fish (*Lophius americanus*) were from Biofish (Gloucester, MA).

Construction and Cloning of Recombinant DNAs. Poly(A)$^+$RNA was extracted from the angler fish islets as described by Majzoub *et al.* (16). A cloned library of cDNAs representative of the total population of islet mRNAs was prepared as described (12, 17). cDNAs were inserted into the *Pst* I site of the vector plasmid pBR322 and replicated in host *E. coli* strain χ1776. Plasmids containing islet cDNAs were isolated by using the cleared lysate technique (18) followed by two successive centrifugations in CsCl/ethidium bromide gradients. Cloning was done under P-1/EK-1 conditions of containment in accord with National Institutes of Health guidelines involving recombinant DNA molecules.

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

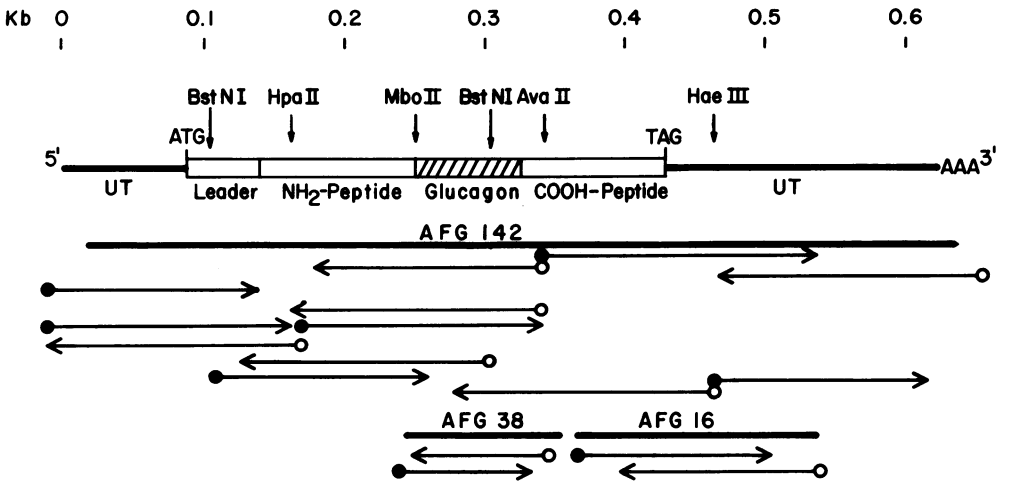

FIG. 1. Restriction endonuclease map and strategy for sequence analyses of cloned cDNAs encoding angler fish preproglucagon. Numbers at the top refer to nucleotide positions within the cDNA (kilobases). The overall map of the 630-base pair cDNA encoding most of the preproglucagon mRNA is shown with restriction endonuclease sites used to cleave it into fragments for nucleotide sequence analysis. Arrows denote directions and polarity [sense (●) or nonsense (○) strands] and length of cDNAs analyzed. Strands were labeled with ^{32}P at the 5′ or 3′ ends. All regions of cDNA AFG 142 were analyzed on both sense and nonsense strands. AFG 38 and AFG 16, cDNAs of 120 and 180 nucleotide base pairs, respectively; heavy lines, their regions of overlap with AFG 142.

Identification of cDNAs Coding for Preproglucagon. Bacterial colonies containing recombinant plasmids were screened by filter hybridization (19) to identify plasmids with cDNA inserts corresponding to the M_r 14,500 preproglucagon. A 120-base pair cDNA, which we have shown previously (12) to encode a portion of the M_r 14,500 preproglucagon, was used as a ^{32}P-labeled probe in the filter hybridizations. Colonies were further checked for the presence of cDNAs corresponding to the mRNA encoding the M_r 14,500 preproglucagon by hybridization-selection and cell-free translation experiments (12, 20).

Isolation and Nucleotide Sequence Analysis of Labeled cDNAs. Restriction fragments of DNA were labeled at the 5′ ends by using [^{32}P]ATP and polynucleotide kinase (21) or at the 3′ ends by using cordycepin [^{32}P]triphosphate and terminal transferase, as described by Goodman *et al.* (17). The fragments labeled at both ends were either cleaved by a second restriction endonuclease to generate fragments with one end labeled or were heated at 90°C in 30% dimethyl sulfoxide to separate the two labeled strands. DNAs with one end labeled were isolated by electrophoresis on 5% polyacrylamide gels containing Tris/borate/Na_2EDTA. The nucleotide sequences of end-labeled DNA fragments were determined by the chemical method of Maxam and Gilbert (22).

RESULTS

Two hundred bacterial colonies containing recombinant cDNAs prepared from angler fish islet poly(A)$^+$RNA were screened by hybridization with the ^{32}P-labeled cloned cDNA of 120 base pairs (AFG 38) previously shown by us to contain a coding sequence for glucagon. Of these 200 colonies, one colony, AFG 142 (630 base pairs), was identified by a positive hybridization. This cDNA and another cDNA, AFG 16 (180 base pairs), were shown to contain coding sequences for preproglucagon by demonstration that they each hybrid selected from total islet poly(A)$^+$RNA an mRNA that programs the synthesis of the preproglucagon of 14,500 daltons in a wheat germ cell-free system (data not shown).

Recombinant plasmids isolated from the three bacterial clones, AFG 16, 38, and 142, were digested with one or more of various restriction endonucleases, end labeled with ^{32}P, and analyzed by electrophoresis on 5% polyacrylamide gels containing Tris/borate/Na_2EDTA. The identification of sites

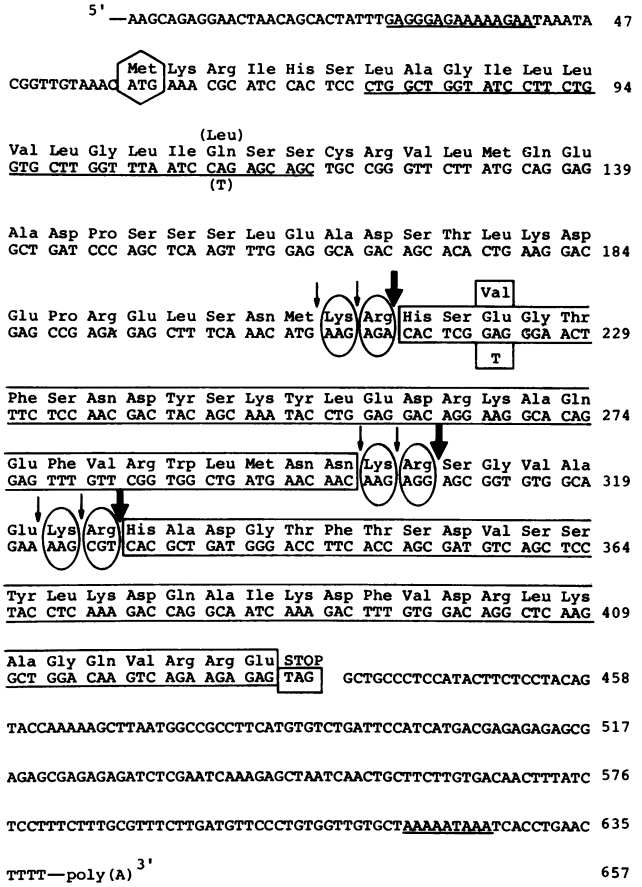

FIG. 2. Nucleotide sequence of preproglucagon mRNA determined from the nucleotide sequences of three cDNAs encoding a precursor of angler fish islet glucagon and amino acid sequence of the precursor. The 29-amino acid angler fish glucagon and the 34-amino acid glucagon-like COOH-terminal peptide, which is structurally homologous to hormones of the glucagon family, are indicated by boxed regions. Circled basic amino acids and arrows represent potential prohormone cleavage sites. The underlined sequence toward the 5′ end indicates consecutive hydrophobic amino acids characteristic of a leader or signal sequence. Initiator methionine and stop codons are indicated. The underlined sequence A-A-T-A-A-A at the 3′ end is characteristic of sites involved in addition of the poly(A) tract to eukaryotic mRNAs.

within the cDNAs that were cleaved by the restriction endonucleases *Ava* II, *Hpa* II, *Mbo* II, *Hae* III, and *Bst*NI proved useful in determining the nucleotide sequences (Fig. 1). One of the cloned angler fish islet cDNAs (AFG 142) was 630 base pairs long and contained the entire coding sequence (372 base pairs) for the M_r 14,500 preproglucagon and all of the 3′- and most of the 5′-untranslated sequences. Inasmuch as the mRNA encoding this preproglucagon has been shown to consist of ≈650 nucleotides, AFG 142 represents a nearly full-length copy of the preproglucagon mRNA (23).

The nucleotide sequence derived from the three cloned cDNAs and the corresponding 124 amino acid sequence of the M_r 14,500 preproglucagon obtained by decoding the longest open reading frame of 402 bases in the nucleotide sequence are shown in Fig. 2. The first ATG, encoding methionine, in the open reading frame is located 59 nucleotides from the 5′ end of the cDNA. Twenty to 30 nucleotides upstream from the ATG is a purine-rich sequence reminiscent of a Shine–Dalgarno sequence, believed to be necessary for the binding of the ribosome during translation (24). Five codons downstream from the initial methionine codon are 14 codons representing a sequence of predominantly hydrophobic amino acids characteristic of those found in the NH_2-terminal leader or signal sequences of precursors to secreted proteins (25). Following the last amino acid in the putative leader sequence, 30 amino acids appear before the first Lys-Arg sequence. This Lys-Arg sequence is followed by a sequence of 29 amino acids that are identical to those of mammalian glucagon in 20 of the 29 positions (3). The nucleotide sequences of cDNAs AFG 142 and AFG 38 are identical except for a single nucleotide; a thymidine-for-adenosine substitution that changes a valine to a glutamic acid residue at the third position in the glucagon sequence (Fig. 2).

Following the asparagine corresponding to the COOH-terminal residue of the glucagon sequence are, in succession, a Lys-Arg, a pentapeptide, another Lys-Arg, a sequence of 34 amino acids, and a termination codon, TAG. Remarkably, the COOH-terminal sequence of 34 amino acids in the preproglucagon shows striking homology to the sequence of glucagon preceding it as well as to the sequences of mammalian glucagon and other hormones of the glucagon family, including secretin, gastric inhibitory peptide, and vasoactive intestinal peptide (Fig. 3).

The stop codon is followed by a sequence of 206 nucleotides and a 3′-poly(A) tract characteristic of most eukaryotic mRNAs (26). The sequence A-A-T-A-A-A located 14 nucleotides upstream from the poly(A) tract is similar to sequences found in other eukaryotic mRNAs (27), which are thought to be involved in polyadenylation of the mRNA.

DISCUSSION

Surprisingly, the preproglucagon of 124 amino acids deduced from the nucleotide sequence of the cloned cDNAs contains two glucagon-related sequences arranged in tandem separated by a nonapeptide beginning and ending with the sequence Lys-Arg. Similar Lys-Arg sequences are sites that are cleaved during the posttranslational processing of prohormones (15). Because the primary structures of the various peptides in the glucagon family of hormones in the fish have not been determined, we are limited to comparing the structure of the fish preproglucagon with the known structures of mammalian glucagon and related peptides (Fig. 3). In making these comparisons, we believe it likely that the glucagon-related sequence located internally in the fish preproglucagon and flanked by the Lys-Arg sequences is fish glucagon. The fish sequence is the same length as mammalian glucagon (29 amino acids) and is identical to the sequence of mammalian glucagon in 20 of 29 positions. Eight of the nine substituted amino acids represent conservative changes with high probabilities of evolutionary acceptance in the protein structure (2). Moreover, analysis of the amino acid composition of glucagon isolated from angler fish islets has shown that it contains 29 amino acids (28). Our findings are also consistent with the earlier studies of Tager and Markese (7) and Patzelt *et al.* (6), who analyzed peptides prepared by enzymic digestion of newly synthesized proteins in rat islets and found

```
GLUCAGON-RELATED
COOH PEPTIDE          - Lys Arg His Ala Asp Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Lys Asp

FISH GLUCAGON         - —— —— —— Ser Glu —— —— —— Ser Asn —— Tyr —— Lys —— —— Glu ——

MAMMALIAN GLUCAGON              —— Ser Gln —— —— —— —— —— —— Tyr —— Lys —— —— Asp Ser

MAMMALIAN GIP                   Tyr —— Glu —— —— —— Ile —— —— Tyr —— Ile Ala Met Asp Lys

MAMMALIAN SECRETIN              —— Ser —— —— —— —— —— —— Glu Leu —— Arg Leu Arg Asp Ser

MAMMALIAN VIP                   —— Ser —— Ala Val —— —— Asp Asn Tyr Thr Arg Leu Arg —— Gln


                      - Gln Ala Ile Lys Asp Phe Val Asp Arg Leu Lys Ala Gly Gln Val Arg Arg Glu-COOH

                      - Arg Lys Ala Gln Glu —— —— Arg Trp —— Met Asn Asn COOH

                      - Arg Arg Ala Gln —— —— —— Gln Trp —— Met Asn Thr COOH

                      - Ile Arg Gln Gln —— —— —— Asn Trp —— Leu —— Gln —— Lys Gly Lys Lys

                      - Ala Arg Leu Gln Arg Leu Leu Gln Gly —— Val-CONH2

                      - Met —— Val —— Lys Tyr Leu Asn Ser Ile Leu Asn-CONH2
```

FIG. 3. Amino acid sequence of the 34-amino acid glucagon-related COOH-terminal peptide of angler fish preproglucagon compared with amino acid sequences of other hormones in the glucagon family—angler fish glucagon, mammalian glucagon, mammalian gastric inhibitory peptide (GIP), mammalian secretin, and mammalian vasoactive intestinal peptide (VIP). Lined positions, amino acids identical with those in the glucagon-related COOH-terminal peptide.

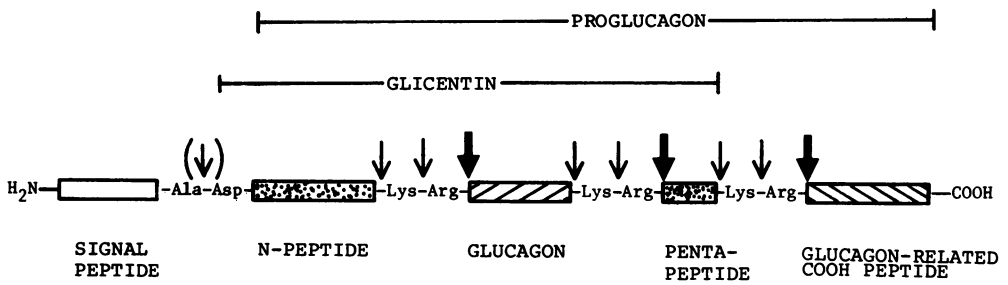

FIG. 4. Diagram of the structure of angler fish islet preproglucagon showing predicted sites of cleavages during posttranslational processing of the precursor *in vivo*. Heavy arrows, predicted sites of cleavages by trypsin-like activity; light arrows, sites cleaved by carboxypeptidase B-like activity; arrow in parentheses, putative (not proven) site of cleavage of the NH_2-terminal signal sequence. (See Fig. 2 for comparison.)

that glucagon resides within the sequence of a large precursor.

There is striking homology of the 34-amino acid glucagon-related sequence located at the COOH terminus of the preproglucagon with the sequences of the mammalian peptides glucagon, secretin, gastric inhibitory peptide, and vasoactive inhibitory peptide. This high degree of homology, coupled with the finding that the sequence is preceded by a Lys-Arg sequence, suggests that this glucagon-related COOH-terminal peptide is formed along with glucagon and an intervening pentapeptide during the posttranslational processing of preproglucagon *in vivo*. Perhaps, this COOH-terminal sequence most closely resembles that of mammalian gastric inhibitory peptide in that both the COOH-terminal sequence and gastric inhibitory peptide are longer than other members of the glucagon family of peptides. A further suggestion that this COOH-terminal sequence of the fish preproglucagon may be a fish gastric inhibitory peptide-like peptide is the report that gastric inhibitory peptide-like immunoreactivity was detected in both the pancreatic A cell and the intestinal L cell, which produce glucagon and glucagon-like immunoreactivity (29), respectively. At present, however, predictions of the biologic properties of the peptide remain speculative. The final identification of the COOH-terminal peptide will require specific biologic testing of peptide prepared by either chemical synthesis or extraction from fish islets.

In certain respects, the structure of angler fish preproglucagon is similar to the structure of glicentin, a polypeptide isolated from porcine intestine (10) (Fig. 4). Glicentin is a 69-amino acid peptide that contains the sequence of mammalian glucagon flanked at both NH_2 and COOH-termini by Lys-Arg and additional peptide extensions. It appears that, in the intestinal L cells, the Lys-Arg sequences are not cleaved, inasmuch as glucagon is not found in significant concentrations in intestinal extracts by using antisera specific for the COOH-terminus of the hormone (10). If islet preproglucagon and intestinal glicentin both result from the expression of the same gene, a possibility that is as yet unproven, then the single precursor would appear to be processed to form glucagon in one organ (islets) but not in another (intestine). In this regard, it is notable that Tager and Steiner (11) structurally characterized a "proglucagon fragment," isolated from commercial porcine glucagon (islets), that consists of the sequence of glucagon with only an eight-amino acid COOH-terminal peptide attached. Presumably, in the islets, the NH_2-terminal peptide extension was cleaved from the glucagon precursor and the cleavage of the COOH-terminal peptide was incomplete. It should be noted, however, that there is no discernible homology between fish preproglucagon and mammalian glicentin in the amino acids of the NH_2- and COOH-terminal peptide extensions. This may be a result of extensive mutational changes in the glucagon gene during the coevolution of fish and mammals. Alternatively, separate genes may encode intestinal glicentin and islet preproglucagon. Comparison of glucagon-related cDNAs derived from islets and intestine should elucidate the relationship between glicentin and glucagon.

There appear to be at least two angler fish islet preproglucagons (M_r 14,500 and 12,500) encoded by separate genes (12). The possibility of the existence of yet a third islet preproglucagon gene is raised as a result of our finding, in two cDNAs encoding the M_r 14,500 preproglucagon, a point mutation that changes a glutamic acid codon to a valine codon in residue three of glucagon. Because the mRNA used to prepare the cDNA library was extracted from islets obtained from several angler fish, we cannot exclude genetic polymorphism among individual fish as an explanation for the structural differences in the preproglucagons.

The finding of what appears to be two distinct hormones within a single precursor provides yet another example of a polyprotein—a protein from which multiple biologically important peptide fragments are released during posttranslational processing. Other hormonal polyproteins discovered recently include proopiomelanocortin (30), proenkephalin (31), propressophysin (32), and prosomatostatin (17, 33–35). Thus, it appears that, in addition to regulated expression of one or more genes within closely related multigene families, diversification of hormonal activities may be generated at the level of specific, and possibly regulated, posttranslational processing of the hormonal precursors.

We thank W. Chin, G. Heinrich, and H. Kronenberg for helpful advice and J. Sullivan for help in preparation of the manuscript. This work was supported in part by fellowships from the National Institutes of Health (R.H.G.) and the British Science Research Council (P.K.L.).

1. Unger, R. H. & Orci, L. (1981) *N. Engl. J. Med.* **304**, 1518–1524; 1575–1580.
2. Dayhoff, M. O. (1972) *Atlas of Protein Sequence and Structure* (Nat. Biomed. Res. Found., Washington, D.C.), Vol. 5.
3. Williams, R. H. & Porte, D., Jr. (1974) in *Textbook of Endocrinology*, ed. Williams, R. H. (Saunders, Philadelphia), 5th Ed., p. 502.
4. Loren, I., Alumets, J., Hakanson, R., Sundler, F. & Thorell, J. (1979) *Histochemistry* **61**, 335–344.
5. Unger, R. H. & Orci, L. (1975) *Lancet* **i**, 14–16.
6. Patzelt, C., Tager, H. S., Carroll, R. J. & Steiner, D. F. (1979) *Nature (London)* **282**, 260–266.
7. Tager, H. S. & Markese, J. (1979) *J. Biol. Chem.* **254**, 2229–2233.
8. Tung, A. K. (1973) *Horm. Metab. Res.* **5**, 416–424.
9. Noe, B. D. & Bauer, G. E. (1975) *Endocrinology* **97**, 868–877.
10. Thim, L. & Moody, A. J. (1981) *Regulatory Peptides* **2**, 139–150.
11. Tager, H. S. & Steiner, D. F. (1973) *Proc. Natl. Acad. Sci. USA* **70**, 2321–2325.
12. Lund, P. K., Goodman, R. H. & Habener, J. F. (1981) *J. Biol. Chem.* **256**, 6515–6518.
13. Lund, P. K., Goodman, R. H., Jacobs, J. W. & Habener, J. F. (1980) *Diabetes* **29**, 583–586.

14. Shields, D., Warren, T. G., Roth, S. E. & Brenner, M. J. (1981) *Nature (London)* **289,** 511–514.
15. Steiner, D. F., Quinn, P. S., Chan, S. J., Marsh, J. & Tager, H. S. (1980) *Ann. N.Y. Acad. Sci.* **343,** 1–16.
16. Majzoub, J. A., Kronenberg, H. M., Potts, J. T., Jr., & Habener, J. F. (1979) *J. Biol. Chem.* **254,** 7449–7455.
17. Goodman, R. H., Jacobs, J. W., Chin, W. W., Lund, P. K., Dee, P. C. & Habener, J. F. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 5869–5873.
18. Clewell, D. B. & Helsinki, D. R. (1970) *Biochemistry* **9,** 4428–4440.
19. Grunstein, M. & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 3961–3965.
20. Ricciardi, R. P., Miller, J. S. & Roberts, B. E. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 4927–4931.
21. Berkener, K. L. & Folk, W. R. (1977) *J. Biol. Chem.* **252,** 3176–3184.
22. Maxam, A. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 560–564.
23. Lund, P. K., Goodman, R. H. & Habener, J. F. (1981) *Biochem. Biophys. Res. Commun.* **100,** 1659–1666.
24. Shine, J. & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. USA* **71,** 1342–1346.
25. Blobel, G. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 1496–1500.
26. Lewin, G. (1975) *Cell* **4,** 11–20.
27. Proudfoot, N. J. & Brownlee, G. G. (1976) *Nature (London)* **263,** 211–214.
28. Trakatellis, A. E., Toda, K., Yamaji, K. & Gardiki-Kouidou, P. (1975) *Biochemistry* **14,** 1508–1512.
29. Larson, L.-I. & Moody, A. J. (1980) *J. Histochem. Cytochem.* **28,** 925–933.
30. Nakanishi, S., Inoue, A., Kita, T., Nakamura, M., Chang, A. C. Y., Cohen, S. N. & Numa, S. (1979) *Nature (London)* **278,** 423–427.
31. Lewis, R. V., Stern, A. S., Kimura, S., Rossier, J., Stein, S. & Udenfriend, S. (1980) *Science* **208,** 1459–1461.
32. Russell, J. F., Brownstein, M. J. & Gainer, H. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 6086–6090.
33. Esch, F., Bohlen, P., Ling, N., Benoit, R., Brazeau, P. & Guilleman, R. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 6827–6831.
34. Schally, A. W., Huang, W.-Y., Chang, R. C. C., Arimura, A., Redding, T. W., Millor, R. P., Hunkapillar, M. W. & Hood, L. E. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 4489–4493.
35. Hobart, P., Crawford, R., Shen, L., Pictet, R. & Rutter, W. J. (1980) *Nature (London)* **288,** 137–141.