Nath, N., Chandra, T. S., Suzuki, H., Carlos, A., & Seidel, J. C. (1982) *Biophys. J. 37*, 47a.
Okamoto, Y., & Sekine, T. (1981a) *J. Biochem. (Tokyo) 89*, 697–700.
Okamoto, Y., & Sekine, T. (1981b) *J. Biochem. (Tokyo) 90*, 833–842.
Onishi, H., & Watanabe, S. (1979) *J. Biochem. (Tokyo) 85*, 457–472.
Perrie, W. T., & Perry, S. V. (1970) *Biochem. J. 119*, 31–38.
Scholey, J. M., Smith, R. S., Drenckhahn, D., Groschel-Stewart, U., & Kendrick-Jones, J. K. (1982) *J. Biol. Chem. 257*, 7737–7745.
Seidel, J. C. (1980) *J. Biol. Chem. 255*, 4355–4361.
Sellers, J. R., & Harvey, E. V. (1984) *J. Biol. Chem. 259*, 14203–14207.
Sellers, J. R., Pato, M. D., & Adelstein, R. S. (1981) *J. Biol. Chem. 256*, 13137–13142.
Sellers, J. R., Eisenberg, E., & Adelstein, R. S. (1982) *J. Biol. Chem. 257*, 13880–13883.
Sellers, J. R., Harvey, E. V., & Soboeiro, M. F. (1985) *J. Cell Biol. 101*, 163a.
Sherry, J. M. F., Gorecka, A., Aksoy, M. O., Dabrowska, R., & Hartshorne, D. J. (1978) *Biochemistry 17*, 4411–4418.
Sobieszek, A., & Small, J. V. (1976) *J. Mol. Biol. 101*, 75–92.
Sobieszek, A., & Small, J. V. (1977) *J. Mol. Biol. 112*, 559–576.
Vu, N., & Wagner, P. D. (1986) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 45*, 1677.
Wagner, P. D., & Weeds, A. G. (1977) *J. Mol. Biol. 109*, 455–473.
Wagner, P. D., & George, J. N. (1986) *Biochemistry 25*, 913–918.
Wagner, P. D., & Vu, N. (1986) *J. Biol. Chem. 261*, 7778–7783.
Wagner, P. D., Vu, N., & George, J. N. (1985) *J. Biol. Chem. 260*, 8084–8089.
Walsh, M. P., & Hartshorne, D. J. (1982) in *Calcium and Cell Function* (Cheung, W. Y., Ed.) Vol. 3, pp 223–269, Academic Press, New York.
Weeds, A. G., & Pope, B. (1977) *J. Mol. Biol. 111*, 129–157.
Yamamoto, K., & Sekine, T. (1979) *J. Biochem. (Tokyo) 86*, 1855–1862.

Isolation of Products and Intermediates of Pancreatic Prosomatostatin Processing: Use of Fast Atom Bombardment Mass Spectrometry as an Aid in Analysis of Prohormone Processing†

P. C. Andrews and Jack E. Dixon*
Department of Biochemistry, Purdue University, West Lafayette, Indiana 47907
Received December 18, 1986; Revised Manuscript Received March 26, 1987

ABSTRACT: Major products and an intermediate in the proteolytic processing pathway of preprosomatostatin I from anglerfish (*Lophius americanus*) were purified and characterized. Proteolytic mapping by fast atom bombardment mass spectrometry was used to rapidly locate regions of the peptides whose masses deviated from those deduced from the cDNA sequence. Amino acid analysis and partial Edman sequencing were also used to confirm the structures. The protein structural data indicate a Glu for Gly substitution at position 83 of preprosomatostatin I (aPPSS-I, numbering from the initiator Met) relative to the cDNA sequence. Two of the peptides isolated, aPPSS-I (26–52) (7.5 nmol·g^{-1}) and aPPSS-I (26–92) (49.5 nmol·g^{-1}), define signal cleavage as occurring between Cys-25 and Ser-26. A partial sequence was obtained from fragment ions in the mass spectrum of a peptide corresponding to aPPSS-I (94–105) (58 nmol·g^{-1}). The 14-residue somatostatin [SS-14 corresponding to aPPSS-I (108–121)] has previously been isolated [Noe, B. D., Spiess, J., Rivier, J. E., & Vale, W. (1979) *Endocrinology (Baltimore) 105*, 1410–1415]. Taken together, these peptides suggest a pathway for prosomatostatin I processing in which the residues corresponding to SS-14 and the immediately preceding 14 residues are cleaved from the prohormone via endoproteolysis (order of cleavage not determined). The fragment aPPSS-I (94–105) was isolated in lower yield than SS-14 and may represent a secondary site of cleavage. Subsequent cleavage at arginine-53 results in the minor peptide aPPSS-I (26–52). The terminal basic amino acids generated by endoproteolytic processing were not found for any of the peptides isolated. The peptides described were identified as products of aPPSS-I processing in radiolabeling studies using intact anglerfish islets [Noe, B. D., Andrews, P. C., Dixon, J. E., & Spiess, J. (1986) *J. Cell Biol. 103*, 1205–1211].

Protein sequence data are now being deduced from cDNA sequences at rates that exceed the capacity of protein chemists to analyze posttranslational processing events. Proteins may undergo a large number of posttranslational modifications including glycosylation, phosphorylation, proteolytic processing, and disulfide bond formation among many others (Wold, 1981). At present, it is not possible to predict, with accuracy, which posttranslational modifications of a particular protein sequence might occur. Principal factors determining post-translational events include primary structure, but higher orders of structure also appear to be important (Oroszlan & Copeland, 1985). Tissue-specific expression of the processing enzyme is also important (Marshak & Yamada, 1984; Pra-

†This work was supported in part by a research development award from the American Diabetes Association and by Grants AM-18849 and AM-18024 from the National Institutes of Health. This is Journal Paper No. 10732 from the Purdue University Agricultural Experiment Station.

0006-2960/87/0426-4853$01.50/0

dayrol et al., 1980; Marshak et al., 1983). If the ultimate structure of a gene product is to be determined, it is presently necessary to purify the proteins or peptides in question and to determine their structures directly. Of the many proteins that undergo processing events, prohormones and precursors to neuropeptides represent a class of proteins exposed to a great variety of posttranslational modifications.

The peptide hormone somatostatin inhibits a wide range of biological functions including secretion of growth hormone, glucagon, insulin, secretin, and gastric acid (Koerker et al., 1974; Bloom et al., 1975; Konturek et al., 1976; Brazeau et al., 1973; Boden et al., 1975). In addition to its broad biological activities, somatostatin is widely distributed in tissues and has been reported in hypothalamus, extrahypothalamic brain, pancreas, gut, retina, and thyroid tissues (Parsons et al., 1976; Arimura et al., 1975; Brownstein et al., 1975; Spiess et al., 1979; Polak et al., 1975; Yamada et al., 1980) among other tissues. Two forms of somatostatin have been reported in mammals: the 14-residue somatostatin (SS-14),[1] which has an identical structure in all organisms from which it has been isolated (Spiess et al., 1979; Benoit et al., 1980; Plisetskaya et al., 1986; Andrews & Dixon, 1981; Schally et al., 1976; Noe et al., 1979a), and a 28-residue somatostatin (SS-28), which in rat and human appears to be derived from the same gene product as SS-14. The ratio of the two forms appears to change from one tissue to another, and they exhibit differential potencies for their various biological activities (Mandarino et al., 1981; Brown et al., 1981; Millar et al., 1983).

Early studies on the biosynthesis of somatostatin indicated the presence in rat pancreas of a 12.5-kilodalton prohormone (Patzelt et al., 1980). Two preprosomatostatins were observed on cell-free translation of mRNA from anglerfish islets (Goodman et al., 1980a; Shields, 1980). The existence of a potential third form has been deduced from cell-free translations (Warren & Shields, 1982, 1984). Pulse-chase studies using anglerfish islets (Noe et al., 1978a,b, 1979b) and rat islets (Patzelt et al., 1980) indicated that proteolytic conversion of the prohormone to somatostatin occurred and that no detectable intermediates were observed. The fate of the amino-terminal fragment, representing the major portion of the prohormone, was unknown. The lack of information on the structure of the prohormone limited the scope of early biosynthesis experiments.

The advent of recombinant DNA techniques resulted in the cloning of cDNA to the mRNA for two different preprosomatostatins from anglerfish endocrine pancreas (Hobart et al., 1980a; Goodman et al., 1980b, 1982a). Subsequently, the cDNA for somatostatins from a number of species were cloned, including two from catfish (Minth et al., 1982; Magazin et al., 1982) and one each from rat (Goodman et al., 1982b; Funckes et al., 1983) and human (Shen et al., 1982). The protein sequences deduced from the cDNA for preprosomatostatin from all species indicated that the portion corresponding to mature SS-14 occurs at the extreme carboxyl end of preprosomatostatin. SS-14 is preceded in prosomatostatin by two basic residues which appear to represent a recognition site for proteolytic processing. Similar sites are utilized during processing of other prohormones. A single arginine occurs 13 residues before the basic dipeptide processing site for SS-14 and is the cleavage site for generation of SS-28.

One of the two anglerfish preprosomatostatins (aPPSS-I) contains the sequence-invariant SS-14 at its carboxyl-terminal end. The second anglerfish preprosomatostatin (aPPSS-II) has a variant sequence in the carboxyl-terminal 14 residues and has been shown to be processed to a 28-residue somatostatin (Noe & Spiess, 1983) containing hydroxylysine (Spiess & Noe, 1985; Andrews et al., 1984). The existence of two different prosomatostatins in the same tissue that are separately processed to SS-14 and SS-28 provides an interesting model system for the study of prosomatostatin processing.

Although knowledge of preprohormone sequences from the cDNA has greatly augmented our understanding of prohormone processing, determination of the bonds cleaved in processing events requires isolation and structure elucidation of the products of processing. The analysis of such peptides is facilitated by fast atom bombardment mass spectrometry (FAB MS), which is used to determine the molecular weight and can provide partial sequence information. This approach has been used previously to partially analyze protein structures deduced from cDNA as a validation of the cDNA sequence (Gibson & Biemann, 1984; Beckner & Caprioli, 1985; Morris et al., 1983). This study highlights the utility of FAB MS for structure analysis of the products of prohormone processing. This paper describes the isolation and characterization of the peptides produced during processing of anglerfish preprosomatostatin I and suggests a pathway for preprosomatostatin I processing. The peptides described were used in a related study to demonstrate the biological relevance of these fragments to pancreatic prosomatostatin processing (Noe et al., 1986).

Materials and Methods

Materials. Decapsulated Brockman bodies (endocrine pancreas) from anglerfish (*Lophius americanus*) were obtained from Biofish Associates, Georgetown, MA. The tissue was obtained from fresh specimens and immediately frozen on dry ice. The tissue was stored at −80 °C until use. Reversed-phase HPLC columns were from SynChrom Corp., Lafayette, IN. Norleucine and trifluoroacetic acid were from Pierce Chemical Co. Acetonitrile was Fisher HPLC grade, and ribonuclease A was from Sigma Chemical Co. Bovine insulin was a gift from Eli Lilly and Co., Indianapolis, IN. Porcine elastase was the gift of M. Laskowski, Jr., and was prepared by M. Laskowski, Sr. *Staphylococcus aureus* V-8 protease was from Pierce, and porcine trypsin was obtained from Worthington and was further purified on an ovomucoid affinity column.

Tissue Extraction. Extracts were prepared by a modification of the method of Pettinga (1958) described by Andrews and Ronner (1985) except that 1.0 mM phenylmethanesulfonyl fluoride was present in the extraction buffers. Typical extractions used 5–10 g of tissue. Details of the purifications are included in the legends to Figures 1 and 2.

Amino Acid Composition, Gas-Phase Sequencing, and Proteolytic Mapping by Fast Atom Bombardment Mass Spectrometry. The large fragment, aPPSS-I (26–92) was hydrolyzed in vacuo in 6 N HCl for 24, 44, and 55 h at 110 °C. All other peptides were hydrolyzed for 24 h at 110 °C in 6 N HCl. The composition of the hydrolysate was determined on a Durrum D-500 amino acid analyzer. Values reported were extrapolated to zero time. Norleucine was added as a standard prior to hydrolysis.

All peptides were sequenced on an Applied Biosystems gas-phase sequenator using standard operating procedures.

Positive and negative ion fast atom bombardment mass spectra of proteolytic digests were obtained with a Kratos MS-50 mass spectrometer. The mass spectrometer was fitted

[1] Abbreviations: SS-14, the 14-residue somatostatin; SS-28, the 28-residue somatostatin; aPPSS-I, anglerfish preprosomatostatin I; aPPSS-II, anglerfish preprosomatostatin II; FAB MS, fast atom bombardment mass spectrometry; HPLC, high-performance liquid chromatography; DTE, dithioerythritol; DTT, dithiothreitol; SP, sulfopropyl.

with neither a postacceleration detector nor a data-processing system. The freeze-dried digest was dissolved in approximately 10 μL of 1 N formic acid. A volume containing 100–200 pmol of the digest was then applied to the probe tip and an approximately equal volume of 3:1 DTT/DTE added (Witten et al., 1984). Resolution was 5000 with the accelerating potential set at 5 kV. The mass determination was limited primarily by the accuracy with which the data could be read from the UV galvanometer tracing. Accuracy was estimated to be ±0.5 dalton. Xenon was the neutral particle source. CsI cluster ions were used for calibration. Peptides (1 nmol) were digested in 50 μL of 0.1 N *N*-ethylmorpholinium acetate, pH 8.0, using *S. aureus* V-8 protease, porcine elastase, or affinity-purified trypsin at a substrate to enzyme ratio of 20:1 (w/w). Incubation was at 37 °C for 2 h. Under the conditions of proteolysis described here *S. aureus* V-8 protease was observed to cleave after both Asp and Glu residues. The mass of intact aPPSS-I (26–92) was determined with a JEOL HX100HF double-focusing mass spectrometer having a mass range of 4500 mass units at 5-kV accelerating potential. Postacceleration ion detection was achieved by using an off-axis detector having a conversion dynode at −20 kV.

Manual Edman degradation was performed according to the method of Tarr (1977). In general, after a single cycle of Edman degradation, a significant improvement in ion intensity (3–5-fold) was apparent for most peptides in the digests. Data from the proteolytic digests were analyzed by using a program written for the IBM PC[2] that matches molecular ions to putative fragments within a protein sequence and which also calculates the expected masses for fragments derived from proteolytic digests. The masses reported in this paper are monoisotopic masses unless otherwise indicated.

Results

Peptide Purification. Generally, 5–10 g of decapsulated anglerfish islets was extracted as described under Materials and Methods, and the extract was applied to a G-50 superfine column (Figure 1A). Several absorbance maxima at 280 nm are apparent in the column profile. Many pancreatic hormones have previously been isolated from these fractions by using radioimmunoassay and various chemical methods for identification (see citations under Discussion). A major absorbance peak (fractions 94–100) corresponding to M_r 6500–7500 (calibration not shown; however, for purposes of comparison, insulin elutes in fractions 106–111) was observed (Figure 1A). In the preparation shown and in one other this represented a major absorbance peak. Three preparations contained reduced levels. It is not known whether the variability in this peak represents seasonal variation or differences in extraction, although every effort was made to ensure reproducible extraction conditions. This high molecular weight fraction was resolved into two peaks by reversed-phase HPLC (Figure 1B). The peptide having the earlier elution time would appear to correspond to aPPSS-I (26–92) by amino acid composition (Table I). However, the amino acid analysis indicated one less Gly than predicted from the cDNA sequence.

During the purification of the pancreatic hormones it was apparent that relatively few major peptides were present in the extract. All absorbance maxima from the G-50 column were further resolved chromatographically and their constituent peptides identified by application of amino acid analysis, Edman degradation, or FAB MS. Only two additional pools from the G-50 column were found to contain major fragments

[2] P. C. Andrews, unpublished program.

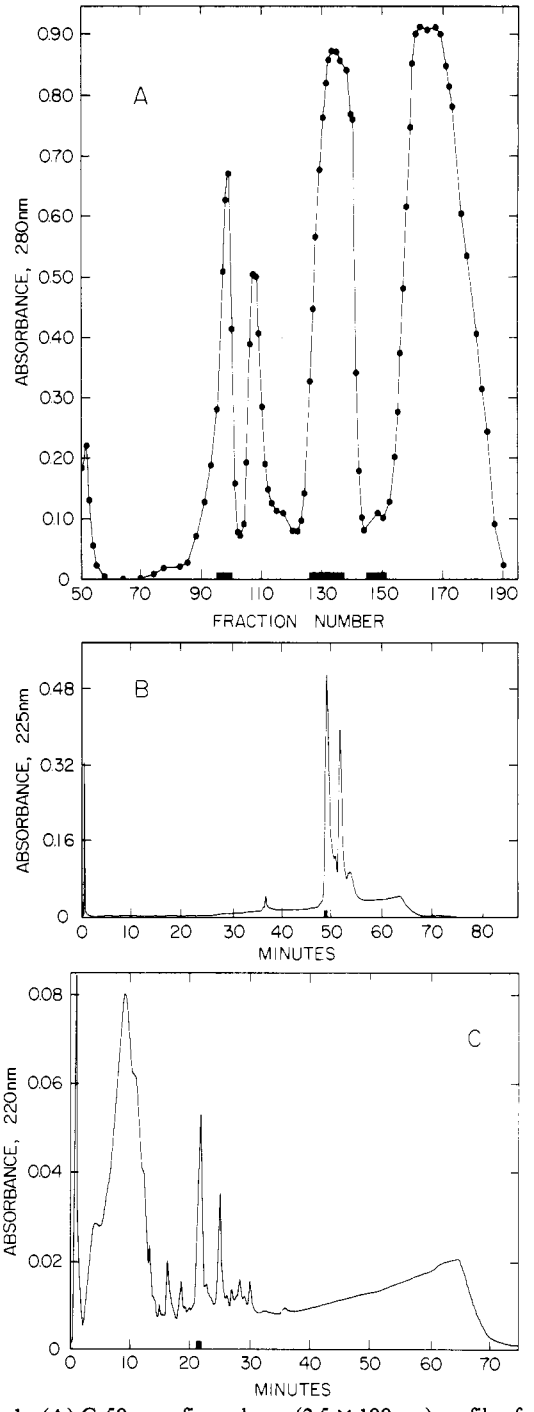

FIGURE 1: (A) G-50 superfine column (2.5 × 100 cm) profile of crude anglerfish islet extract representing 10 g of tissue. The column was calibrated with ribonuclease A (13 700 daltons), bovine insulin (5700 daltons), and carboxymethylated insulin B chain (3400 daltons) and A chain (2300 daltons). Peptides were eluted with 1 N formic acid. Absorbance was monitored at 280 nm, and 2.0-mL fractions were collected. The fractions indicated by the solid bar were combined separately. (B) Reversed-phase HPLC of the high-mass peak from the G-50 column indicated in A above. Fractions 94–100 from the G-50 column wer combined, freeze-dried, and dissolved in 0.1% trifluoroacetic acid (TFA) in water. The peptide was then applied to a 0.46 × 5.0 cm RPC8 HPLC column in 10% portions. The peptides were eluted from the column with a 60-min gradient to 40% acetonitrile at a flow rate of 1.0 mL·min^{-1}. Absorbance was monitored at 225 nm. (C) HPLC profile of fractions 145–155 from G-50 column (A) for a preparation from 8 g of tissue. The lowest size fraction indicated by a solid bar in (A) was freeze-dried, taken up in 0.1% TFA, and 1.5% of the total sample applied to a 0.46 × 5.0 cm reversed-phase HPLC column. The peptides were eluted with a 60-min gradient to 40% acetonitrile at a flow rate of 0.7 mL·min^{-1}. Both solvents contained 0.1% TFA. Absorbance was monitored at 110 nm, and the fraction corresponding to aPPSS-I (94–105) is indicated by a solid bar.

Table I

amino acid	anglerfish preprosomatostatin I (94–105) no. of residues	anglerfish preprosomatostatin I (94–105) residues from cDNA	anglerfish preprosomatostatin I (26–92) no. of residues	anglerfish preprosomatostatin I (26–92) residues from cDNA	anglerfish preprosomatostatin I (26–52) no. of residues	anglerfish preprosomatostatin I (26–52) residues from cDNA	anglerfish[b] preprosomatostatin I (54–92) no. of residues	anglerfish[b] preprosomatostatin I (54–92) residues from cDNA
Asp		0	6.9	7	2.1	2	5.0	5
Thr		0	0.9	1	0.8	1		0
Ser	1.0	1	5.0	5	2.5	3	2.1	2
Glu	1.1	1	13.3	13	2.9	3	10.8	10
Pro	2.4	2	2.7	3	2.0	2	2.0	2
Gly	2.3	2	3.9	5	2.0	2	2.2	3
Ala	3.0	3	7.0	7	1.0	1	6.0	6
Cys		0	0	0		0		0
Val		0	0	0		0		0
Met		0	1.3	1	0.8	1		0
Ile		0	0	0		0		0
Leu	2.0	2	13.8	14	4.9	5	8.4	9
Tyr		0	1.1	1	0.9	1		0
Phe		0	1.8	2	0.9	1	1.0	1
His		0	1.9	2	0.9	1	1.0	1
Lys		0	2.0	2	2.1	2		0
Arg	1.3	1	3.8	4	3.0	3		0
Trp	ND[a]	0	ND	0	ND	0	ND	0

[a] ND, not determined. [b] Tryptic fragment of aPPSS-I (26–92).

Table II: Tryptic Mapping by FAB MS

anglerfish preprosomatostatin I (26–92) residue no.	obsd $[M + H]^+$ (daltons)	theor $[M + H]^+$ (daltons)	post-Edman $[M + H]^+$ exptl	residue identified	anglerfish preprosomatostatin I (26–52) residue no.	obsd $[M + H]^+$ (daltons)	theor $[M + H]^+$ (daltons)
26–31	665.8	665.3	578.5	Ser	26–31	665.0	665.3
32–34	NO[a]	349.2	NO		32–34	NO	349.2
35–36	288.0	288.2	NO		35–36	288.0	288.2
37–41	651.8	651.4	538.5	Leu	37–41	651.0	651.4
42–48	792.6	792.4	764.6	Tyr	42–48	792.5	792.4
49–53	(648.9)[c]	(648.6)[c]	NO		49–52	NO	494.2
54–92	NA	4221.0	NM[b]				

[a] NO, not observed. [b] NM, not measured. [c] Observed in negative mode only—$[M - H]^-$.

of preprosomatostatin I. A fraction just preceding the salt volume of the G-50 column (Figure 1A, fractions 145–155) was collected and subjected to reversed-phase HPLC (Figure 1C). The amino acid composition of one peptide was found to be identical with that for aPPSS-I (94–105) (Table I).

Fractions from the G-50 column corresponding to M_r 3000–4000 (fractions 128–136) were combined and chromatographed on SP-Sephadex (Figure 2A). All major absorbance peaks eluting from the SP-Sephadex column were subjected to reversed-phase HPLC. One of these peptides (Figure 2B) had a tryptic map by FAB MS (Table II) and an amino acid composition (Table I) identical with that expected for aPPSS-I (26–52).

Peptide Structure Analysis. The amino acid composition of the peptide aPPSS-I (26–92) in Table I has a glycine value one residue lower than the composition deduced from the cDNA sequence. While the composition of aPPSS-I (26–52) was identical with that predicted from the cDNA sequence, the amino acid composition of the tryptic fragment aPPSS-I (54–92) contained one less Gly residue and one more Glx residue relative to that of the cDNA sequence. Five glycine residues are predicted from the originally reported cDNA sequence (Hobart et al., 1980a) to occur in aPPSS-I (26–92). The site of the Glx for Gly substitution suggested by amino acid analysis (Table I) was determined by FAB MS mapping of proteolytic digests of aPPSS-I (26–92). One of the advantages of this method is that the products of digestion need not be resolved chromatographically prior to analysis, decreasing the analysis time considerably. Fragment ions diagnostic of the amino acid sequence of 25 of the first 28 residues of aPPSS-I (26–92) (Table II) were obtained for 200 pmol of a tryptic digest. For example, a molecular ion $[M + H]^+$ for the amino-terminal fragment (residues 26–31) was observed at 665.8 daltons, consistent with the theoretical mass of 665.3 daltons. The sensitivity of the instrument was not sufficient in its present configuration to observe the tryptic fragment corresponding to aPPSS-I (54–92) at mass 4191.1 daltons. The only other fragment not observed was mass 348.2 daltons, corresponding to the tripeptide Asp-Ser-Lys (residues 32–34). The fragment corresponding to residues 49–53 was observed in the negative mode only ($[M - H]^- = 648.9$ daltons). The trifluoroacetic acid adduct of this peptide in the negative mode was of greater intensity than the molecular ion. The trifluoroacetic acid in the sample is a carryover from the HPLC solvents. The complete tryptic digest was subjected to a single cycle of Edman degradation (Table II) in order to confirm the mass assignments. The Edman-processed peptides whose molecular ions were observed gave the loss of mass expected for the correct amino-terminal residue. Lysine-containing peptides have an increment in mass of one phenyl isothiocyanate moiety (135.1 daltons) in addition to the loss of the amino-terminal residue. For example, the fragment corresponding to residues 42–48 of aPPSS-I decreased from 792.4 to 764.6 daltons after one cycle of Edman degradation. This corresponds to the loss of a single Tyr residue (163.1 daltons) and the gain of one phenyl isothiocyanate moiety (135.1 daltons) conjugated to the ϵ-amino group of Lys-48. The mass of the tryptic dipeptide (residues 35–36) was not observed after Edman degradation due to interference from the matrix and other low-mass components of the digest. All

Table III: Proteolytic Digests of aPPSS-I (26–92)

elastase digest					*S. aureus* V-8 protease digest				
residue no.	$[M + H]^+$ exptl (daltons)	$[M + H]^+$ theor (daltons)	$[M + H]^+$ post-Edman exptl	$[M + H]^+$ post-Edman theor	residue no.	$[M + H]^+$ exptl (daltons)	$[M + H]^+$ theor (daltons)	$[M + H]^+$ post-Edman theor	$[M + H]^+$ post-Edman exptl
58–82	2772.5	2772.3	2643.5	2643.3	33–50	2152.5	2152.2	2335.5	2335.8
63–79	1960.3	1959.9	1844.8	1844.9	59–70	1343.8	1343.7	1230.6	1231.3
26–38	1490.3	1490.8	1539.3	1538.9	59–68	1100.4	1100.6	987.5	987.9
39–47	1071.0	1070.6	957.8	957.5	51–58	878.7	878.4	747.4	747.8
39–46	983.5	983.5	870.8	870.4	76–81	690.6	690.3	576.3	576.7
34–38	642.8	642.4	514.3	NO[a]					
83–87	559.8	560.2	431.2	NO					

[a] NO, not observed.

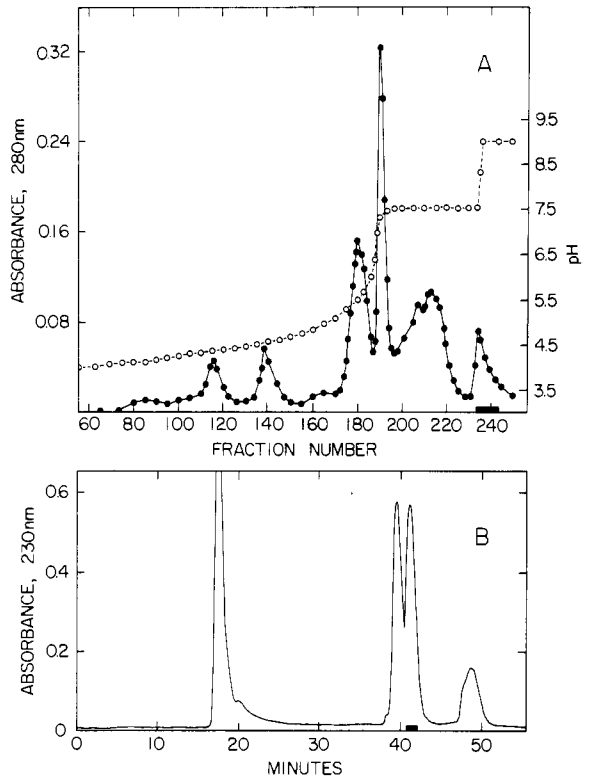

FIGURE 2: (A) SP-Sephadex column profile of M_r 3000–4000 absorbance peak from G-50 column of an 8-g tissue preparation. After the fraction from the G-50 column had been combined and diluted 10-fold and the pH adjusted to 3.0, it was applied to a SP-Sephadex column (1.0 × 15 cm), which had been previously equilibrated with 0.1 N ammonium formate, pH 3.5. The column was then washed with 50 mL of starting buffer and the peptides eluted with a 500-mL gradient to 0.1 N ammonium formate, pH 9.0. Absorbance was monitored at 280 nm. The pH gradient is indicated by the dashed line. (B) Reversed-phase HPLC elution profile of the peptide having an amino acid composition identical with that of aPPSS-I (26–52). The sample applied to the HPLC corresponded to the fractions indicated by a solid bar in (A) above (fractions 233–243). The fractions were freeze-dried, dissolved in 0.1% TFA in water, and then applied to a 1.0 × 25 cm reversed-phase HPLC column. The peptide was eluted with a 60-min gradient to 40% acetonitrile. Both solvents contained 0.1% TFA. Absorbance was monitored at 230 nm. The fraction of the peak collected is indicated by a solid bar.

mass assignments for the tryptic digest were consistent with the protein sequence deduced from the cDNA and indicated that neither Gly-29 nor Gly-46 of aPPSS-I is the site of substitution.

Two other proteolytic digests provided mass information spanning the remaining three glycyl residues. Digestion of aPPSS-I (26–92) by *S. aureus* V-8 protease or elastase (Table III) provided a series of overlapping peptides whose masses were consistent with the assignment of Gly-67 and Gly-82. The site of amino acid substitution was determined to be residue 83. The proteolytic fragment which contained residue 83 (83–87) was observed to have a molecular ion, $[M + H]^+$, of 560 daltons, indicating that the residue substituted for Gly-83 was Glu not Gln. The masses of the fragments derived from elastase and *S. aureus* V-8 protease were determined only in the positive mode. Those proteolytic fragments not observed in the elastase and V-8 protease digests may be inefficiently detected in the positive mode, poorly ionized, or, in the case of manual Edman degradation, lost during solvent extraction steps. The three proteolytic maps described here provided overlapping mass information for 62 of the 67 residues of aPPSS-I (26–92) (Figure 3). In agreement with the mass data from the proteolytic digests, the average mass of intact aPPSS-I (26–92) was found to be 7507.3 daltons (calculated average mass = 7507.1 daltons).[3]

Independent verification of the site and identity of the substitution was obtained by automated Edman degradation of the tryptic fragment, aPPSS-I (54–92) (Table IV, purification not shown). Intact aPPSS-I (26–92) was also sequenced (Table IV). Only PTH-Gly was observed at positions 29, 46, 67, and 82. The only residue apparent at position 83 was PTH-Glu. The amino acid composition of the tryptic fragment corresponding to aPPSS-I (54–92) confirmed that the carboxyl-terminal Arg (residue 93 of aPPSS-I) was absent (Table I).

An effort was made to isolate other fragments of aPPSS-I that might have arisen during processing (see Peptide Purification above). One of the peptides isolated (Figure 2B) had a tryptic FAB MS map consistent with an amino-terminal fragment of prosomatostatin I (Table II). Amino acid analysis (Table I) and automated Edman degradation (Table IV) confirmed this conclusion and indicated that the peptide isolated corresponded to aPPSS-I (26–52). The peptide corresponding to aPPSS-I (54–92) was not isolated. However, a peptide was isolated (Figure 1C) with a molecular ion (1138.2 daltons) consistent with the theoretical monoisotopic mass (1138.6 daltons) of aPPSS-I (94–105). Fragmentation ions Y_3'' through Y_{11}'' were observed [nomenclature is that of Roestorff and Fohlman (1984)] and provided sequence information for 8 of the 12 amino-terminal residues (Figure 4), consistent with the sequence of aPPSS-I (94–105). Further fragmentation ions observed represented simultaneous cleavage from the carboxyl and the amino termini. The fragment ions corresponded to $(B_{11}\ Y_{11}'')_{10}$, $(B_{11}\ Y_{10}'')_9$, $(B_{10}\ Y_{11}'')_9$, and $(B_9\ Y_{11}'')_8$, allowing the remaining carboxyl-terminal residues to be ordered. The amino acid composition of the peptide was also identical with that of aPPSS-I (94–105) (Table I). The yields of the aPPSS-I derived peptides were estimated by amino acid analysis to be aPPSS-I (26–52) 49.5 nmol·g^{-1},

[3] T. D. Lee and K. Legesse, personal communication; see Acknowledgments.

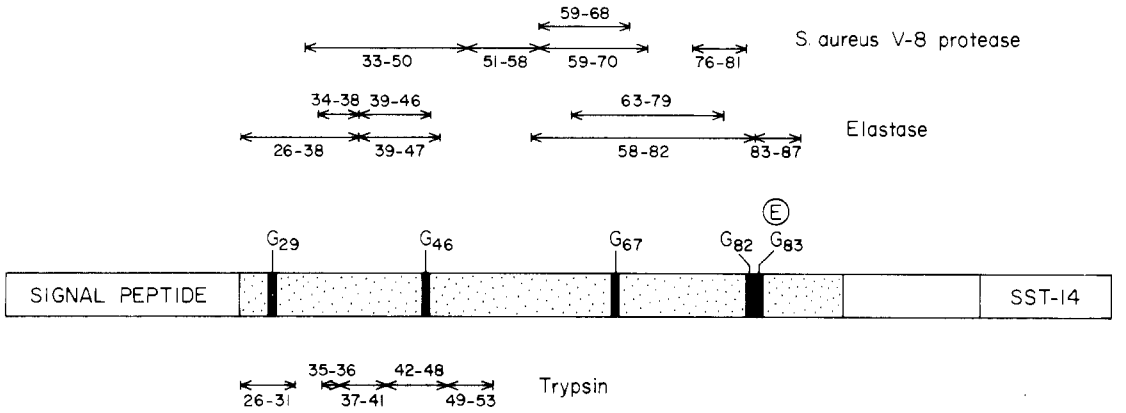

FIGURE 3: Summary of proteolytic maps of aPPSS-I (26–92) (stippled region). The arrows indicate the fragments for which molecular ions were observed for digestions with trypsin, elastase, and *S. aureus* V-8 protease. The positions of all five glycine residues deduced from the DNA are indicated by solid bars. The site of the Glu for Gly substitution is indicated (circled residue).

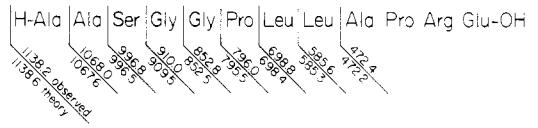

FIGURE 4: Summary of fragment ions observed for aPPSS-I (94–105), which provided partial sequence information. The upper value is the experimental value, the lower is the theoretical nominal mass.

aPPSS-I (26–92) 7.5 nmol·g^{-1}, and aPPSS-I (94–105) 58 nmol·g^{-1}.

Discussion

This paper identifies putative products and one intermediate of pancreatic prosomatostatin processing. The utility of FAB MS is demonstrated for mapping the in vivo proteolytic cleavage sites of prohormones and confirming the structures deduced from the cDNA sequences. Information provided by FAB MS is often complementary to data obtained by other methods of analysis and can be obtained without derivatizing the peptide to render it volatile. FAB MS can provide information unattainable or at least difficult to obtain by other methods. The information provided by FAB MS in this study was sufficient to indicate a discrepancy between the sequence of the peptide isolated from anglerfish endocrine pancreas and the original cDNA sequence and to identify the site and nature of the substitution. The discrepancy in the literature regarding the identity of residue 83 is now resolved in favor of glutamic acid. Confirmatory evidence obtained by classical methods (isolation of tryptic fragments and sequencing by gas-phase Edman degradation) required an equivalent amount of peptide and a considerably greater investment of time and effort.

The Glu for Gly substitution at position 83 of aPPSS-I may be explained by a single base change from the Gly codon, GGA, to the Glu codon, GAA. The protein sequence described here is consistent with the corrected anglerfish cDNA sequence of Goodman et al. (1982a) and with recent sequence analysis of the aPPSS-I gene.[4] Other than the change at residue 83, the data indicated that the peptide sequence was as deduced from the cDNA. No evidence for posttranslational modification of amino acid side chains was evident.

An interesting point regarding the structure of the peptides isolated in this study is that the terminal basic amino acids resulting from endoproteolysis are not present. All of the major peptides isolated from anglerfish islets have also been processed in this manner, including fragments of proglucagon (Andrews et al., 1986), prosomatostatin II (Andrews et al., 1984),[5] and propeptide Y (Andrews & Dixon, 1986; Andrews et al., 1985). The terminal basic residues would presumably have been removed by the potent basic-residue-specific exopeptidases reported to be present in secretory vesicles (Docherty & Hutton, 1983; Gainer et al., 1984).

The peptides described in this paper define proteolytic processing sites and allow a general scheme for preproso-

Table IV

preprosomatostatin I (26–92)				preprosomatostatin I (54–92)			
cycle	residue	pmol[a]	residue[b] from cDNA	cycle	residue	pmol[a]	residue[b] from cDNA
1	Ser	530	Ser	1	Ser	268	Ser
2	Phe	1065	Phe	2	Ala	827	Ala
3	Ala	860	Ala	3	Leu	568	Leu
4	Gly	645	Gly	4	Ala	518	Ala
5	Gln	758	Gln	5	Glu	353	Glu
6	Arg	150	Arg	6	Leu	556	Leu
7	Asp	250	Asp	7	Leu	482	Leu
8	Ser	111	Ser	8	Leu	505	Leu
9	Lys	532	Lys	9	Ser	83	Ser
10	Leu	450	Leu	10	Asp	121	Asp
11	Arg	48	Arg	11	Leu	360	Leu
12	Leu	407	Leu	12	Leu	375	Leu
13	Leu	416	Leu	13	Gln	172	Gln
14	Leu	503	Leu	14	Gly	188	Gly
15	His	104	His	15	Glu	93	Glu
16	Arg	53	Arg	16	Asn	206	Asn
17	Tyr	262	Tyr	17	Glu	149	Glu
18	Pro	221	Pro	18	Ala	181	Ala
19	Leu	277	Leu	19	Leu	98	Leu
20	Gln	215	Gln	20	Glu	74	Glu
21	Gly	143	Gly	21	Glu	121	Glu
22	Ser	*c*	Ser	22	Glu	102	Glu
23	Lys	129	Lys	23	Asn	81	Asn
24	Gln	99	Gln	24	Phe	96	Phe
25	Asp	48	Asp	25	Pro	57	Pro
26	Met	56	Met	26	Leu	70	Leu
27	Thr	203	Thr	27	Ala	49	Ala
28	Arg	*c*	Arg	28	Glu	27	Glu
				29	Gly	35	Gly
				30	Glu	27	Gly
				31	Pro	19	Pro
				32	Glu	17	Glu
				33	Asp	*c*	Asp
				34	Ala	39	Ala
				35	Xxx		His
				36	Ala	36	Ala

[a] Quantitation for PTH-Ser and PTH-Thr does not include dehydro forms. [b] The cDNA sequence is from Hobart et al. (1980a). [c] Observed but not quantitated.

[4] R. Crawford, P. Hobart, and W. J. Rutter, unpublished results.

[5] P. C. Andrews and J. E. Dixon, unpublished observations.

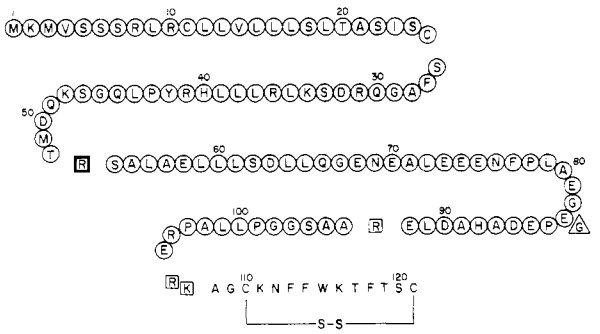

FIGURE 5: Sequence of anglerfish preprosomatostatin I from the corrected cDNA sequence (Goodman et al., 1982a) and the gene sequence.[4] Residues removed during processing are enclosed by squares. Putative processing sites are indicated by gaps. The somatostatin 14 sequence occurs in the last 14 residues (not circled). The site of the Glu for Gly substitution at residue 83 is indicated; the previously assigned Gly is enclosed by a triangle.

matostatin processing to be proposed (Figure 5). The largest fragment isolated, aPPSS-I (26–92), indicates that conversion to the prohormone involves removal of a 25-residue signal peptide via cleavage after Cys-25. Proteolytic processing at the basic dipeptide, -Arg-Lys- (residues 106 and 107), appears to occur. This signal-cleavage site is identical with that observed for processing of aPPSS-I by microsomes from canine pancreas in a cell-free translation system (Spiess & Noe, 1985). Radiolabeled sequencing of affinity-purified anglerfish prosomatostatin I also indicates this to be the site of signal cleavage (Noe et al., 1986). Cleavage at the single Arg-93 also occurs. This conclusion is consistent with previous pulse-chase studies in rat (Patzelt et al., 1980) and anglerfish (Warren & Shields, 1982, 1984), which indicate that somatostatin immunoreactivity is rapidly released from the prohormone with no other major detectable somatostatin immunoreactive species.

Cleavage of aPPSS at Arg-53 also occurs, generating aPPSS-I (26–52) and presumably aPPSS-I (54–92). Although no evidence for the latter peptide was obtained, the lack of basic residues in this region suggests that it may not be further processed.

Recent reports by Benoit et al. (1982, 1984) and others (Wu et al., 1983; Penman et al., 1983; Schmidt et al., 1985) describe the isolation from rat brain of several peptides derived from prosomatostatin. In addition to SS-14 and SS-28, the dodecapeptide analogous to aPPSS-I (94–105), the peptide analogous to aPPSS-I (26–105), and a shorter version of this peptide, which appear to be generated in rat brain via an anomalous cleavage between two leucine residues, were isolated. This anomalous cleavage occurs at a similar position to the cleavage of aPPSS-I (26–92) at Arg-53. Secondary structure calculations for rat and anglerfish SS-14 precursors using a modified Chou–Fasman calculation (Argos et al., 1983) indicate that while considerable differences exist in their primary structures, the secondary structure parameters exhibit significant similarities. The similar sites of cleavage for the rat and anglerfish SS-14 precursors suggest that some common structural feature might render these sites more susceptible to proteolysis. The brain is a heterogeneous tissue in which prosomatostatin is processed to SS-14 in some neurons and to SS-28 in others. It is not yet clear whether the peptides reported by Benoit et al. and others are products of alternative processing pathways in different cell types or if they represent products of a random processing pathway.

Because the peptides isolated in this study represent primarily the steady-state composition of the islets, it is not possible to comment on the relative order of cleavage at Arg-93 or the basic dipeptide (Arg-106, Lys-107), which would liberate SS-28 and SS-14, respectively. Nor is it possible to distinguish between a random and an ordered processing pathway unless it can be demonstrated that all intermediates have been isolated, not just the major ones. It is impossible to conclusively rule out the possibility that some of the peptides isolated might be proteolytic degradation products arising during the extraction or purification procedure. However, an effort was made to minimize this possibility through use of an acidic ethanol extraction procedure and the presence of a protease inhibitor. The collection method and initial storage conditions are other potential contributors to artifactual proteolysis. No peptides which had arisen as a result of cleavage at nonbasic residues were observed, nor were peptides with "ragged" ends observed, suggesting that random proteolysis had not occurred to an appreciable extent.

The biological significance of the peptides described in this paper has been the subject of a related study (Noe et al., 1986). Consistent with the present study, recent radiolabeling studies (Noe et al., 1986) indicate that aPPSS-I (26–52), aPPSS-I (26–92), and aPPSS-I (94–105) are products of anglerfish prosomatostatin processing in vivo. Additionally, no evidence for aPPSS-I (54–92) was observed. The significance of this latter observation is not known.

Of the five single Arg residues in aPPSS-I occurring at positions 31, 36, 41, 53, and 93, processing to any significant degree appears to occur at only two, Arg-53 and Arg-93. Processing at Arg-93 may occur more readily than at Arg-53, judging solely from the relative yields of aPPSS-I (26–92) (49.5 nmol·g^{-1}) and aPPSS-I (26–52) (7.5 nmol·g^{-1}). The parameters that allow a particular arginine residue to be a candidate for proteolytic processing are not yet clear. No obvious, consistent sequence homology occurs between basic processing sites in different prohormones. This lack of direct sequence homology suggests that higher order structure of the processing site may provide the necessary recognition parameters for the processing enzymes.

This study has provided the structures of major and minor products of prosomatostatin I processing in pancreas. Similar studies have identified the products of proglucagon processing (Andrews et al., 1986), propeptide Y processing (Andrews et al., 1985; Andrews & Dixon, 1986), and prosomatostatin II[5] and proinsulin[5] processing in anglerfish pancreas. The anglerfish is presently the only species for which all the major products of pancreatic prohormone processing have been determined. The structures of all the anglerfish prohormones are known from their cDNA sequences (Hobart et al., 1980a,b; Goodman et al., 1980b; Lund et al., 1982, 1983).[6] The hormones of the anglerfish pancreas may be the best-characterized pancreatic hormones in terms of processing. The relatively complete information available on anglerfish pancreatic prohormones provides a complete foundation for further studies addressing the roles of secondary and tertiary structure in determining prohormone processing.

ACKNOWLEDGMENTS

We thank Dr. M. Hermodson for access to the gas-phase sequenator and for helpful discussions. We also thank Dr. R. Yazdanparast for the fast atom bombardment mass spectra of proteolytic digests and Dr. D. Smith for helpful discussions. We especially thank Terry D. Lee and Kassu Legesse, Beckman Research Institute of the City of Hope, Division of Immunology, Duarte, CA, for determining the mass of intact

[6] P. Hobart et al., Pfizer Central Research, Groton, CT, personal communication regarding the pro-aPY cDNA sequence.

aPPSS-I (26–92) by FAB MS.

Registry No. SS-14, 38916-34-6; PPSS-I, 84068-95-1; aPPSS-I, 76975-09-2; aPPSS-I (94–105), 108868-00-4; aPPSS-I (26–92), 108868-02-6; aPPSS-I (26–52), 108868-01-5; aPPSS-I (54–92), 108868-03-7.

REFERENCES

Andrews, P. C., & Dixon, J. E. (1981) *J. Biol. Chem. 256*, 8267–8270.

Andrews, P. C., & Ronner, P. (1985) *J. Biol. Chem. 260*, 3910–3914.

Andrews, P. C., Hawke, D., Shively, J. E., & Dixon, J. E. (1984) *J. Biol. Chem. 259*, 15021–15024.

Andrews, P. C., Hawke, D., Shively, J. E., & Dixon, J. E. (1985) *Endocrinology (Baltimore) 116*, 2677–2681.

Andrews, P. C., Hawke, D. H., Lee, T. D., Legesse, K., Noe, B. D., & Shively, J. E. (1986) *J. Biol. Chem. 261*, 8128–8133.

Argos, P., Taylor, W. L., Minth, C. D., & Dixon, J. E. (1983) *J. Biol. Chem. 258*, 8788–8793.

Arimura, A., Sato, H., Dupont, A., Nishi, N., & Schally, A. V. (1975) *Science (Washington, D.C.) 189*, 1007–1009.

Beckner, C. F., & Caprioli, R. M. (1985) *Biomed. Mass Spectrom. 12*, 393–398.

Benoit, R., Böhlen, P., Brazeau, P., Ling, N., & Guillemin, R. (1980) *Endocrinology (Baltimore) 107*, 2127–2129.

Benoit, R., Ling, N., Alford, B., & Guillemin, R. (1982) *Biochem. Biophys. Res. Commun. 107*, 944–950.

Benoit, R., Böhlen, P., Esch, F., & Ling, N. (1984) *Brain Res. 311*, 23–29.

Bloom, S. R., Thorner, M. O., Mortimer, C. H., Besser, G. M., Hall, R., Gomez-Pan, A., Roy, V. M., Russell, R. C. G., Coy, D. H., Kastin, A. J., & Schally, A. V. (1975) *Lancet 2*, 1106–1109.

Boden, G., Sivitz, M., Owen, D., Essa-Koumar, N., & Landor, J. H. (1975) *Science (Washington, D.C.) 190*, 163–164.

Brazeau, P., Vale, W., Burgess, R., Ling, N., Butcher, M., Rivier, J., & Guillemin, R. (1973) *Science (Washington, D.C.) 179*, 77–79.

Brown, N., Rivier, J., & Vale, W. (1981) *Endocrinology (Baltimore) 108*, 2391–2393.

Brownstein, M., Arimura, A., Sato, H., Schally, A., & Kizer, J. (1975) *Endocrinology (Baltimore) 96*, 1456–1461.

Docherty, K., & Hutton, J. C. (1983) *FEBS Lett. 162*, 137–141.

Funckes, C. L., Minth, C. D., Deschenes, R., Magazin, M., Tavianini, M. A., Sheets, M., Collier, K., Weith, H. L., Aron, D. C., Roos, B. A., & Dixon, J. E. (1983) *J. Biol. Chem. 258*, 8781–8787.

Gainer, H., Russell, J. T., & Loh, Y. P. (1984) *FEBS Lett. 175*, 135–139.

Gibson, B. W., & Biemann, K. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 1956–1960.

Goodman, R. H., Lund, P. K., Jacobs, J. W., & Habener, J. F. (1980a) *J. Biol. Chem. 255*, 6549–6552.

Goodman, R. H., Jacobs, J. W., Chin, W. W., Lund, P. K., Dee, P. C., & Habener, J. F. (1980b) *Proc. Natl. Acad. Sci. U.S.A. 77*, 5869–5873.

Goodman, R. H., Jacobs, J. W., Chin, W. W., Lund, P. K., Dee, P. C., & Habener, J. F. (1982a) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1682.

Goodman, R. H., Jacobs, J. W., Dee, P. C., & Habener, J. F. (1982b) *J. Biol. Chem. 257*, 1156–1159.

Hobart, P., Crawford, R., Shen, L.-P., Pictet, R., & Rutter, W. J. (1980a) *Nature (London) 288*, 137–141.

Hobart, P. M., Shen, L.-P., Crawford, R., Pictet, R. L., & Rutter, W. J. (1980b) *Nature (London) 288*, 1360–1363.

Koerker, D. J., Ruch, W., Chideckel, E., Palmer, J., Goodner, C. J., Ensinck, J., & Gale, C. C. (1974) *Science (Washington, D.C.) 184*, 482–484.

Konturek, S. J., Jasler, J., Obtulowicz, W., Coy, D. H., & Schally, A. V. (1976) *J. Clin. Invest. 58*, 1–6.

Lund, P. K., Goodman, R. H., Dee, P. L., & Habener, J. F. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 345–349.

Lund, P. K., Goodman, R. H., Montminy, M. R., Dee, P. C., & Habener, J. F. (1983) *J. Biol. Chem. 258*, 3280–3284.

Magazin, M., Minth, C. D., Funckes, C. L., Deschenes, R., Tavianini, M., & Dixon, J. E. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 5152–5156.

Mandarino, L., Stenner, D., Blanchard, W., Nissen, S., & Gerich, J. (1981) *Nature (London) 291*, 76–77.

Marshak, D., Yamada, T. (1984) *Invest. Ophthalmol. Visual Sci. 25*, 112–115.

Marshak, D. W., Reeve, J. R., Shively, J. E., Hawke, D., Takami, M. S., & Yamada, T. (1983) *J. Neurochem. 41*, 601–606.

Millar, R. P., Klaff, L. J., Barron, J. L., Levitt, N. S., & Ling, N. (1983) *Clin. Endocrinol. (Oxford) 18*, 277–285.

Minth, C. D., Taylor, W. L., Magazin, M., Tavianini, M. A., Collier, K., Weith, H. L., & Dixon, J. E. (1982) *J. Biol. Chem. 257*, 10372–10377.

Morris, H. R., Panico, M., & Taylor, G. W. (1983) *Biochem. Biophys. Res. Commun. 117*, 299–305.

Noe, B. D., & Spiess, J. (1983) Program of the 65th Annual Meeting of the Endocrine Society, San Antonio, TX, June 8–10, p 85 (Abstr).

Noe, B. D., Fletcher, D. J., Bauer, G. E., Weir, G. C., & Patel, Y. (1978a) *Endocrinology (Baltimore) 102*, 1675–1685.

Noe, B. D., Weir, G. C., & Bauer, G. E. (1978b) *Metab. Clin. Exp. 27*, 1201–1205.

Noe, B. D., Spiess, J., Rivier, J. E., & Vale, W. (1979a) *Endocrinology (Baltimore) 105*, 1410–1415.

Noe, B. D., Fletcher, D. J., & Spiess, J. (1979b) *Diabetes 28*, 724–730.

Noe, B. D., Andrews, P. C., Dixon, J. E., & Spiess, J. (1986) *J. Cell Biol 103*, 1205–1211.

Oroszlan, S., & Copeland, T. D. (1985) *Curr. Top. Microbiol. Immunol. 115*, 221–233.

Parsons, J. A., Erlandsen, S. L., Hegre, O. D., McEvoy, R. C., & Elde, R. P. (1976) *Endocrinology (Baltimore) 96*, 1456–1461.

Patzelt, C., Tager, H. S., Carroll, R. J., & Steiner, D. F. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 2410–2414.

Penman, E., Wass, J. A. H., Butler, M. G., Penny, E. S., Price, J., Wu, P., & Rees, L. H. (1983) *Regul. Pept. 7*, 53–65.

Pettinga, C. W. (1958) *Biochem. Prep. 8*, 28–31.

Plisetskaya, E. M., Pollock, H. G., Rouse, J. B., Hamilton, J. W., Kimmel, S. R., Andrews, P. C., & Gorbman, A. (1986) *Gen. Comp. Endocrinol. 63*, 252–263.

Polak, J., Pearse, A. G. E., Grimelius, L., Bloom, S., & Arimua, A. (1975) *Lancet 1*, 1220–1222.

Pradayrol, L., Jornvall, H., Mutt, V., & Ribet, A. (1980) *FEBS Lett. 190*, 55–58.

Roepstorff, P., & Fohlman, J. (1984) *Biomed. Mass Spectrom. 11*, 601.

Schally, A. V., Dupont, A., Arimura, A., Redding, T., Nishi, U., Linthicum, G., & Schleisinger, D. (1976) *Biochemistry 15*, 509–514.

Schmidt, W. E., Mutt, V., Kratzin, H., Carlquist, M., Conlon, J. M., & Creutzfeldt, W. (1985) *FEBS Lett. 192*, 141–146.

Shen, L.-P., Pictet, R. L., & Rutter, W. J. (1982) *Proc. Natl.*

Acad. Sci. U.S.A. 79, 4575–4579.
Shields, D. (1980) *J. Biol. Chem. 255*, 11625–11628.
Spiess, J., & Noe, B. D. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 277–281.
Spiess, J., Rivier, J. E., Rodkey, J. A., Bennett, C. D., & Vale, W. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 2974–2978.
Tarr, G. E. (1977) *Methods Enzymol. Mol. Biol. 47*, 335–357.
Warren, T. G., & Shields, D. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 3729–3733.
Warren, J. G., & Shields, D. (1984) *Biochemistry 23*, 2684–2690.
Webster, T. A., Gibson, B. W., Keng, T., Biemann, K., & Schimmel, P. (1983) *J. Biol. Chem. 258*, 10637–10641.
Witten, J. L., Schaffer, M. H., O'Shea, M., Cook, J. C., Hemling, M. E., & Rinehart, K. L. (1984) *Biochem. Biophys. Res. Commun. 124*, 350–358.
Wold, F. (1981) *Annu. Rev. Biochem. 50*, 783–814.
Wu, P., Penman, E., Coy, D. H., & Rees, L. H. (1983) *Regul. Pept. 5*, 219–233.
Yamada, J., Marshak, D., Basinger, S., Walsh, J., Mosley, J., & Stell, W. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 1691–1695.

Role of Carbohydrate in the Function of Human Granulocyte–Macrophage Colony-Stimulating Factor†

Kenneth Kaushansky*

Division of Hematology, University of Washington, Seattle, Washington 98195

Patrick J. O'Hara, Charles E. Hart, John W. Forstrom, and Frederick S. Hagen

ZymoGenetics, Inc., Seattle, Washington 98103

Received August 12, 1986; Revised Manuscript Received February 17, 1987

ABSTRACT: cDNA clones for the human hematopoietic regulator granulocyte–macrophage colony-stimulating factor (hGM-CSF) were isolated from a λgt11 cDNA library prepared from RNA of COS cells transiently expressing the gene for hGM-CSF. As the RNA was a rich source of hGM-CSF mRNA, approximately 0.1% of the clones of this library contained hGM-CSF sequences. All of the clones analyzed were full length and were correctly processed. When subcloned into an expression vector and transfected into COS cells, the cDNA clones direct the synthesis of higher levels of the growth factor than the gene from which they were derived. The cDNA for native hGM-CSF was used to generate structural mutants which lack N-linked carbohydrate, O-linked carbohydrate, or both. Although the mutant proteins had differing specific activities, the nonglycosylated forms reproduce many, if not all, of the physiologic functions of authentic hGM-CSF. The role of carbohydrate in the secretion and function of hGM-CSF is discussed.

Colony-stimulating factors (CSFs) are acidic glycoproteins required for the survival, proliferation, and differentiation of hematopoietic progenitor cells in culture (Burgess & Metcalf, 1980). Functionally, the various CSFs are defined by the type of hematopoietic colony produced in semisolid culture. Hence, granulocyte–macrophage colony-stimulating factor (GM-CSF) stimulates the growth of progenitors which give rise to colonies containing granulocytes, macrophages, or a combination of both cell types (Wong et al., 1985). In addition to GM-CSF, granulocyte CSF (G-CSF or CSFβ), macrophage CSF (M-CSF or CSF-1), and multi-CSF (or IL-3) have been characterized and cloned from human sources (Souza et al., 1986; Kawasaki et al., 1985; Yang et al., 1986).

Recently, we obtained a genomic clone for GM-CSF from a human library in λ Charon 4A (Kaushansky et al., 1986). The gene is divided into four exons and three introns and contains the transcription- and translation-controlling elements typical of a eukaryotic structural gene. It is present in a single copy in the human genome and encodes a mature polypeptide of 127 amino acids. There are two N-linked glycosylation sites, and there are reported to be three O-linked glycosylation sites (S. C. Clark, personal communication) which account for approximately 6 kilodaltons (kDa) of the estimated 22 kDa of the mature growth factor.

Colony-stimulating factors are proteins of diverse physiologic function. We (Kaushansky et al., 1986) and others (Emerson et al., 1985) have found that recombinant hGM-CSF expressed in COS cells stimulates not only neutrophilic, eosinophilic, and monocyte–macrophage progenitor cells but also megakaryocyte colony-forming cells and, in the presence of erythropoietin, erythroid and mixed erythroid–nonerythroid colony-forming cells. Further, hGM-CSF has been shown to stimulate mature neutrophils to localize at sites of inflammation (Weisbart et al., 1985), mature eosinophils and monocytes to become activated and to enhance their killing of helminths (Handman & Burgess, 1979; Vadas et al., 1983), and mature monocytes and macrophages to enhance phagocytosis and tumor cell killing (Grabstein et al., 1986). In addition to these in vitro activities, recombinant hGM-CSF was recently demonstrated in primates to stimulate in vivo hematopoiesis (Donahue et al., 1986).

Despite the growing body of knowledge surrounding the in vitro and now in vivo physiology of hGM-CSF, little is known

† This work was supported in part by NIH Grants CA-31615 and AI22772.

* Address correspondence to this author. Recipient of Physician–Scientist Award AM-01304 from the National Institutes of Health.

0006-2960/87/0426-4861$01.50/0