# Rat pancreatic kallikrein mRNA: Nucleotide sequence and amino acid sequence of the encoded preproenzyme

(cDNA cloning/serine protease/signal peptide/activation peptide/zymogen)

Galvin H. Swift, Jean-Charles Dagorn<sup>\*</sup>, Patricia L. Ashley, Scott W. Cummings, and Raymond J. MacDonald

Division of Molecular Biology, Department of Biochemistry, University of Texas Health Science Center, 5323 Harry Hines Boulevard, Dallas, Texas 75235

Communicated by R. Estabrook, September 20, 1982

ABSTRACT We have cloned via recombinant DNA technology the mRNA sequence for rat pancreatic preprokallikrein. Four cloned overlapping double-stranded cDNAs gave a continuous mRNA sequence of 867 nucleotides beginning within the 5'-noncoding region and extending to the poly(A) tail. The mRNA sequence reveals that pancreatic kallikrein is synthesized as a prezymogen of 265 amino acids, including a proposed secretory prepeptide of 17 amino acids and a proposed activation peptide of 11 amino acids. The activation peptide, although similar in length, is distinct from those of the other classes of pancreatic serine proteases. The amino acid sequence of the predicted active form of the enzyme is closely related to the partial sequences obtained for other kallikrein-like serine proteases including rat submaxillary gland kallikrein, pig pancreatic and submaxillary gland kallikreins, the  $\gamma$  subunit of mouse nerve growth factor, and rat tonin. Key amino acid residues thought to be involved in the substrate-cleavage specificity of kallikreins are retained. Hybridization analysis showed relatively high levels of kallikrein mRNA in the rat pancreas, submaxillary and parotid glands, spleen, and kidney, indicating the active synthesis of kallikrein in these tissues.

Glandular kallikreins (EC 3.4.21.8) are members of a closely related subfamily of serine proteases that process polypeptide hormone precursors. Other members include the  $\gamma$  subunit of nerve growth factor (1), the epidermal growth factor-binding protein (2), and tonin (3). Characteristically, each has a muchmore limited substrate-cleavage specificity than other serine proteases such as trypsin, chymotrypsin, and elastase. These kallikrein-like proteases cleave at one or a very few peptide bonds in their natural substrates with a general, but not absolute, preference for residues that have positively charged side chains and a strong bias for arginine over lysine (4, 5). Glandular kallikreins liberate lysylbradykinin, a vasoactive decapeptide, by the selective cleavage of two peptide bonds in the precursor kininogen (for review, see ref. 4).

Kallikreins are found in many exocrine tissues, although their site(s) of synthesis remains unverified, and the nature and processing of presumed kallikrein precursors are unknown. Inactive precursor forms of kallikrein have been purified from rat (6) and porcine (7) pancreas. The precursors are acidic glycoproteins of apparent  $M_r$  37,000. Slow activation occurs spontaneously and is accelerated by catalytic amounts of trypsin. The activated enzymes are single polypeptides with slightly reduced molecular weight, suggesting the release of an activation peptide, which remains uncharacterized. Active kallikrein isolated from autolysed porcine pancreas consists of two polypeptide chains held together by disulfide bridges (8), indicative of further proteolysis.

We report the identification and sequence analysis of the cloned mRNA sequence for rat pancreatic kallikrein and describe the nature of the preproenzyme and the presence of kallikrein mRNA in a number of rat tissues.

# MATERIALS AND METHODS

**Preparation and Analysis of RNA.** RNA from rat tissue was prepared by the guanidine thiocyanate procedure of Chirgwin *et al.* (9). For mRNA blot analyses (10), RNAs resolved in agarose gels containing methylmercury hydroxide (11) were transferred to diazobenzyloxymethyl paper by electroblotting (12) and hybridized as described (13) to DNA of recombinant plasmid pcXP39 that had been labeled with <sup>32</sup>P by nick-translation (14). Polyadenylylated RNA was isolated from total pancreatic RNA by binding to oligo(dT)-cellulose (type 2, Collaborative Research, Waltham, MA) (15).

**Cloned Sequences of Pancreatic mRNAs.** The construction of a double-stranded cDNA (ds cDNA) library of approximately 1,000 recombinant clones that represents total rat pancreatic mRNA sequences has been described (16). The library was obtained by inserting ds cDNA synthesized from total pancreatic polyadenylylated RNA into pBR322 after addition of either *Hin*dIII recognition site decamers (17) or homopolymeric tails (18).

The 5' end of kallikrein mRNA was cloned specifically by a procedure to be described in detail elsewhere. Briefly, a DNA complementary to the 5' end of the mRNA was synthesized by extending a primer prepared from the ds cDNA insert of pcXP39 (see Fig. 1) by restriction endonuclease digestion and preparative polyacrylamide gel electrophoresis. A homopolymeric tail of deoxycytidine residues was added to the 3' end of the extended primer by using terminal deoxynucleotidyltransferase and the second strand was synthesized with reverse transcriptase by using oligo(dG) as primer (19). After addition of deoxycytidine tails, the ds cDNA was annealed to deoxyguano-sine-tailed pBR322 (19) and used to transform *Escherichia coli* RR1 (17). All experiments with bacteria containing recombinant plasmids were conducted according to the National Institutes of Health guidelines for recombinant DNA research.

Nucleic Acid Sequence Analysis. Nucleotide sequence determinations were carried out according to the protocols of Maxam and Gilbert (20).

### RESULTS

Identification of ds cDNA Clones for Kallikrein mRNA. Cloned ds cDNAs for rat pancreatic serine proteases in a large

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

Abbreviations: ds cDNA, double-stranded cDNA.

<sup>\*</sup> Present address: Unite de Reserches de Pathologie Digestive, Institute National de la Santé et de la Recherche Médicale, 13009-Marseille, France.

library of cloned ds cDNAs for the total pancreatic mRNA population were initially identified by selecting for recombinant plasmids bearing inserts for abundant mRNAs of the size expected (0.9–1.2 kilobases) to encode proteins the size of serine protease precursors (16). This screening approach led to the selection of clones encoding elastase I and II (16) and trypsin I and II (13) mRNAs. In subsequent studies, one recombinant plasmid, designated pcXP39, was chosen because it encoded a sequence of a less abundant mRNA about 1.1 kilobases long. Nucleotide sequence analysis of the ds cDNA insert of pcXP39 showed that the amino acid sequence of the encoded protein resembled porcine pancreatic kallikrein (see below).

When the entire pancreatic ds cDNA library was screened by hybridization to the ds cDNA insert of pcXP39, only one additional colony, pcXP3-81, was detected. The ds cDNA inserts of pcXP39 and pcXP3-81 encoded similar regions of the mRNA toward the 3' end (Fig. 1). To obtain the missing 5' end of the mRNA, we cloned cDNA synthesized from total pancreatic polyadenylylated RNA with reverse transcriptase using a restriction enzyme-generated fragment from the 5' end of the pcXP39 insert as a kallikrein-specific primer. Fourteen clones out of a total of 200 we obtained hybridized to a probe derived from the pcXP39 insert 5' to the fragment used to prime the cDNA synthesis; these, therefore, contained extended cDNA sequences. The sequences of the two clones (pcXP16 and pcXP130) with the longest inserts among the 14 were determined (Fig. 1). Both clones represent a fragment of the same mRNA as pcXP39 and pcX3-81, because each contained a 78nucleotide sequence identical to that of the corresponding region of pcXP39 5' to the primer.

The Nucleotide Sequence of Kallikrein mRNA. Fig. 1 illustrates the scheme for the determination of the kallikrein mRNA nucleotide sequence from four partial and overlapping ds cDNA



FIG. 1. The collection of ds cDNA clones and the strategy for sequence analysis of preprokallikrein mRNA. The thick horizontal lines represent the extent of each cloned ds cDNA; the numbers identify each pcXP recombinant plasmid. The position of the primer (F, *Fnu* 4H) derived from pcXP39 for cDNA synthesis is indicated. The horizontal arrows indicate the directions and lengths of the analysis runs; the starting restriction sites are noted (P, *Pst* I; R, *Eco*RI; B, *Bgl* II; H, *Hind*III). The restriction sites in parentheses at the end of each ds cDNA were generated by the cloning procedures. The positions of the catalytically active amino acid residues histidine, aspartate, and serine are noted.

clones. The entire length of the mRNA sequences [excluding the poly(A) tail] was determined by a minimum of two analyses, and 92% of the sequence was confirmed by analysis of both strands of the ds cDNA. One nucleotide discrepancy (in the codon for amino acid 142; see Fig. 2) was found among the four ds cDNA clones that included a total of 1,400 overlapping nucleotides. The nucleotide sequence and the encoded amino acid sequence (Fig. 2) are derived from all four ds cDNA clones. The

UAGGAAGGCGU AAGCUCAGCACCUGCUGCUCCUGC						-28 Met AUG	Pro CCU	Val GUU	Thr ACC	Met AUG	Trp UGG	Phe UUC	Leu CUG	-20 Ile AUC	Leu CUG	Phe UUC	Leu CUC	Ala GCC	Leu CUG	Ser UCC	Leu CUG	Gly GGA	Arg CGG	-10 Asn AAU	Asp GAU	A1a GCU	A1a GCA	Pro CCU	Pro CCC	Val GUC
Gln Se CAG UC	- r Ar U CG	1♥+1 g Val G GUU	Val GUU	G1y GGA	Gly GGA	Tyr UAU	Asn AAC	Cys UGU	Glu GAG	Met AUG	10 Asn AAU	Ser UCC	Gln CAA	Pro CCC	Trp UGG	Gln CAA	Va1 GUG	Ala GCU	Val GUG	Tyr UAC	20 Tyr UAC	Phe UUC	Gly GGC	Glu GAA	Tyr UAC	Leu CUA	Cys UGU	G1y GGG	Gly GGU	Val GUC
30 Leu I1 CUG AU	e As A GA	p Pro C CCC	Ser AGC	Trp UGG	Val GUG	Ile AUC	Thr ACA	Ala GCU	40 Ala GCU	★ His CAC	Cys UGC	A1a GCA	Thr ACC	Asp GAC	Asn AAU	Tyr UAC	G1n CAG	Val GUU	50 Trp UGG	Leu CUG	G1y GGC	Arg CGA	Asn AAC	Asn AAC	Leu CUA	Tyr UAU	G1u GAA	Asp GAU	60 Glu GAA	Pro CCC
Phe A1 UUU GC	a G1 U CA	n His G CAC	Arg CGG	Leu CUU	Val GUC	Ser AGC	70 Gln CAA	Ser AGC	Phe UUC	Pro CCU	His CAC	Pro CCC	G1y GGU	Phe UUC	Asn AAC	Gln CAG	80 Asp GAC	Leu CUC	Ile AUA	Trp UGG	Asn AAC	His CAC	Thr ACC	Arg CGA	G1n CAA	Pro CCU	90 Gly GGG	Asp GAC	Asp GAC	Tyr UAC
Ser As AGC AA	n As UGA	r p Leu U UUG	Met AUG	Leu CUG	100 Leu CUC	His CAC	Leu CUC	Ser AGC	Gln CAG	Pro CCU	Ala GCG	Asp GAC	Ile AUC	Thr ACA	110 Asp GAU	G1y GGU	Val GUG	Lys AAG	Val GUC	Ile AUC	Asp GAU	Leu CUG	Pro CCC	Ile AUU	120 G1u GAG	Glu GAG	Pro CCC	Lys AAG	Va 1 GUG	Gly GGG
Ser Th AGC AC	r Cy C UG	s Leu C CUU	130 A1a GCC	Ser UCG	Gly GGC	Trp UGG	Gly GGC	Ser AGC	Ile AUC	Thr ACA	Pro CCU	Asp GAC	140 Gly GGA	Leu UUG	Ala Glu GAA	Leu UUA	Ser AGU	Asp GAU	Asp GAU	Leu CUC	Gln CAG	Cys UGU	150 Val GUG	Asn AAC	Ile AUC	Asp GAU	Leu CUU	Leu CUG	Ser UCU	Asn AAU
Glu Ly GAG AA	16 s Cy G UG	0 s Val C GUC	Glu GAG	Ala GCA	His CAC	Lys AAA	Glu GAA	Glu GAG	Va1 GUG	Thr ACA	170 Asp GAU	Leu CUC	Met AUG	Leu CUG	Cys UGU	Ala GCA	Gly GGA	Glu GAG	Met AUG	Asp GAU	180 Gly GGG	Gly GGC	Lys AAA	Asp GAC	Thr ACU	Cys UGC	Lys AAG	G1y GGU	Asp GAC	★ Ser UCA
190 Gly Gl GGA GG	y Pr C CC	o Leu C CUC	Ile AUC	Cys UGU	Asn AAU	Gly GGU	Val GUG	Leu CUC	200 G1n CAA	Gly GGC	Ile AUC	Thr ACG	Ser UCC	Trp UGG	Gly GGC	Phe UUU	Asn AAC	Pro CCA	210 Cys UGC	Gly GGU	Glu GAA	Pro CCC	Lys AAG	Lys AAG	Pro CCA	Gly GGC	Ile AUC	Tyr UAC	220 Thr ACC	Lys AAA
Leu I1 CUU AU	e Ly U AA	's Phe IG UUC	Thr ACC	Pro CCC	Trp UGG	Ile AUA	230 Lys AAA	G1u GAA	Val GUU	Met AUG	Lys AAG	Glu GAA	Asn AAC	237 Pro CCC	sto UGA	p GUG	UCAC	ACAG	иссс	CUGG	ucuc	AAUA	<u>AA</u> AC	CCAC	CAUG	CAGC	AC-p	o]y(/	A)	

FIG. 2. Nucleotide sequence and amino acid translation of preprokallikrein mRNA. The deduced amino acid sequence is numbered sequentially from the amino terminus of the predicted active enzyme. The predicted cleavage points for the pre- and propeptides are indicated by arrows. The amino acid residues of the charge relay system are indicated by stars. The conserved sequence A-A-U-A-A in the 3'-noncoding region is underlined. The conserved nucleotide sequence near the 3' end of the eukaryotic 18S rRNA (21) is shown in italics at the position with the greatest base pairing within the 5'-noncoding region.



FIG. 3. Amino acid sequence comparison of kallikreins and related serine proteases. The numbering scheme starts at the amino terminus of the predicted active enzyme form of rat kallikrein. Chymotrypsinogen numbering is included for key amino acid positions to facilitate comparison with other serine proteases. The boxes delineate regions of amino acid sequence identical with that of rat pancreatic kallikrein. The large gaps in the rat submaxillary gland kallikrein (residues 79–87 and 129–237), mouse submaxillary gland kallikrein (residues 1–88), and rat tonin (residues 129–171 and 219–236) represent lack of sequence information and not length differences between the proteins. All other gaps are introduced to optimize sequence alignments. Data are presented for porcine pancreatic kallikrein (PPK; ref. 8 and W. Bode, personal communication); rat submaxillary gland kallikrein (MSK; ref. 22);  $\gamma$  subunit of nerve growth factor (NGF; ref. 25), rat tonin (3), and rat trypsin (Tryp; ref. 13).

cloned mRNA sequence comprises 867 nucleotides [plus a poly(A) tail] and includes twenty-four 5' - and forty-eight 3' -noncoding nucleotides. The 5' -noncoding region contains an octanucleotide sequence with the potential of forming seven base pairs with a conserved sequence near the 3' end of eukaryotic 18S rRNA (21). The 3' -noncoding region of the kallikrein mRNA contains the oligonucleotide sequence A-A-U-A-A-A that is present [sometimes as a slightly altered sequence (22)] 14–20 nucleotides from the poly(A) of eukaryotic mRNAs (23).

The Encoded Preproenzyme. The single continuous open reading frame prescribes a preproenzyme of 265 amino acids with a  $M_r$  of 29,227. The amino acids—histidine-41, aspartate-96, and serine-189—that constitute the charge-relay system of serine proteases and the conservation of adjacent amino acid sequences confirm that the encoded protein is a serine protease. Comparison of the amino acid sequence of the predicted active enzyme with that of other pancreatic serine proteases shows that the highest sequence identity (57%) is to porcine kallikrein (8). In contrast, the sequence is only 40% homologous with rat trypsin I (13) and 29% and 31% homologous with rat elastases I and II (16), respectively. In addition, the number of sequence gaps that need be introduced to maximize homology is lowest for porcine kallikrein.

The close relationship of the rat pancreatic kallikrein with the members of the kallikrein-like subfamily of serine proteases is shown in Fig. 3. Despite the lack of complete sequence information in some instances, the close relationships are indicated by the high levels of sequence identity and the small number of gaps required to optimize the alignments. The percent sequence identity with rat submaxillary gland kallikrein is 99 (with 118 residues compared). Similarly, the sequence identity with rat tonin is 80% (157 residues); with mouse submaxillary gland kallikrein, 71% (148 residues); with mouse  $\gamma$  subunit of nerve growth factor, 71% (237 residues); and with porcine pancreatic kallikrein, 57% (239 residues). Key amino acid residues that determine the characteristic cleavage preference of kallikrein-like serine proteases are preserved in the rat enzyme (see *Discussion*).

Based on the characteristic and highly conserved amino-terminal sequences of the active enzyme forms of simple serine proteases (e.g., see Fig. 3 for kallikrein-like proteases), the deduced rat kallikrein sequence has an additional 28 amino acids at the amino terminus that most likely represent the signal (pre) and activation (pro) peptides. We predict that the cleavage site between the pre- and propeptides is between glycine-12 and arginine-11 (Fig. 2), consistent with the preference of prepeptide cleavage after amino acids that have small side chains (26, 27). For the kallikrein precursor, this would define a signal peptide similar in length (17 amino acids) and amino acid composition to that of four other rat pancreatic serine proteases (13) and an activation peptide of 11 amino acids similar in nature, although distinct in sequence, to those of other rat serine proteases (13, 16).

**Expression of Kallikrein mRNA in Rat Tissues.** RNA from several rat tissues was tested for kallikrein mRNA sequences by hybridization with the cloned kallikrein cDNA (Fig. 4). The data show that kallikrein mRNA (or a closely related mRNA) is present at readily detectable levels in the submaxillary gland, parotid gland, spleen, and kidney, in addition to the pancreas, but not in lung. Kallikrein mRNA was also not detected in RNA from liver, intestine, and lacrimal gland (data not shown). The size of kallikrein mRNA from the pancreas is 1.1 kilobases (six determinations). The sizes of the cross-hybridizing RNAs from the other tissues differ measurably.

# DISCUSSION

The cloned mRNA sequences for pancreatic kallikrein encode a precursor kallikrein containing a predicted secretory prepeptide, an activation peptide, and an active form of kallikrein. The polypeptide molecular weights of the predicted prepro-, pro-(zymogen), and active forms of the enzyme are 29,227, 27,292, and 26,099, respectively. The most conclusive identification of the encoded serine protease is the sequence identity at 117 amino acid residues out of the 118 known residues of rat submaxillary gland kallikrein (Fig. 3). Indeed, this extensive homology indicates that the same kallikrein gene is expressed in the pancreas and submaxillary gland, and this result is supported by the sequence identity of the first 28 amino acids of porcine pancreatic and submaxillary gland kallikreins (8). The single residue difference between the rat enzymes may reflect the lack of genetic homogeneity of the Sprague–Dawley strain.

Amino acid residues required for either serine protease activity or for kallikrein-like cleavage specificity are retained in the rat pancreatic enzyme. All but one of the 29 invariant residues for the extended family of serine proteases listed by Young et al. (28), including the catalytic residues histidine, aspartate, and serine and their contiguous sequences, are preserved. The exception, residue 153 (aspartate instead of proline), is characteristically variable in the kallikrein family (Fig. 3). The presence of aspartate-183 at the sequence position equivalent to the bottom of the substrate binding pocket is characteristic for trypsin-like cleavage after arginyl and lysyl residues. The presence of glycines-206 and -217 is consistent with an open binding pocket to accommodate large amino acid side chains, distinct, for example, from the occluded binding pocket of elastase (29). X-ray crystallographic analysis by W. Bode (personal communication) shows that tyrosine-93 and tryptophan-205 form a hydrophobic sandwich to facilitate the binding of amino acid residues such as phenylalanine adjacent to the cleavage site (e.g., Phe-Arg-X) and thereby contribute to kallikrein specificity.

Porcine kallikrein isolated from autolysed pancreas consists of two polypeptide chains (8) generated by proteolytic cleavage within the extended kallikrein autolysis loop near amino acid 88 (Fig. 3). This "kallikrein autolysis loop" is on the surface of the enzyme (W. Bode, personal communication). The alignment of porcine and rat kallikreins shows an additional seven amino acids (residues 81–87) in this loop in rat and suggests that the porcine enzyme may also have additional residues that are removed after an initial cleavage. Lemon *et al.* (30) have suggested the presence of at least three additional amino acids (histidine, threonine, and lysine) in this region of the porcine enzyme prior to proteolytic processing. The presence of 11 additional amino acids in this external peptide loop relative to trypsin would be expected to extend it far into solution and readily explains its sensitivity to proteolytic cleavage.

Based on the primary structure of rat preprokallikrein and the properties of the purified porcine and murine enzymes, the proteolytic processing steps that occur during the maturation of pancreatic secretory kallikrein can be surmised. The initial processing event must be the removal of the secretory prepeptide (amino acids -28 through -12) that serves as the signal for vectorial transport (31) into the cisternae of the rough endoplasmic reticulum. After intracellular transport and secretion with the other pancreatic digestive (pro)enzymes, a trypsin-like cleavage between arginine -1 and valine-1 releases the 11amino acid activation peptide. The single polypeptide enzyme is fully active (4, 7). Further limited cleavage within the kallikrein autolysis loop, between arginine-87 and glutamine-88, may be accompanied by exoproteolytic removal of several amino acids from the newly generated carboxyl terminus. The twochain kallikrein retains enzymatic activity (4, 7, 8).

The subfamily of kallikrein-like enzymes with cleavage preference for arginyl residues shares many sequence characteristics. They have much greater sequence homology (together with fewer gaps in the optimized sequence alignments) with each other than with the other principal classes of simple serine proteases: trypsin, chymotrypsin, and elastase (3, 5, 24). They also share amino acid insertions at positions 25 and 208, length variations within the kallikrein autolysis loop, and an additional amino acid at the carboxyl terminus, generally proline. Amino acids involved in determining substrate-cleavage preference (aspartate-183, tyrosine-93, glycine-206, and residues that have



FIG. 4. Size and tissue distribution of rat kallikrein mRNA. Total RNAs prepared from the tissues indicated were resolved by electrophoresis in an agarose gel containing methylmercury hydroxide, transferred to diazobenzyloxymethyl paper, and hybridized to <sup>32</sup>P-labeled pcXP39 DNA. The lane marked standards contained yeast mitochondrial RNA and rabbit globin mRNAs as size standards. Also included on the gel were *Taq* I - and *Bst*NI-digested pBR322. The combined RNA and DNA standards were used to generate the size scale shown.

#### Biochemistry: Swift et al.

small side chains at 217) are conserved within this family.

The essentially identical amino acid sequences of rat pancreatic and submaxillary gland kallikreins imply that the same structural gene is expressed in these tissues. The limited sequence comparisons between porcine pancreatic, submaxillary gland, and urinary (i.e., kidney) kallikreins also indicates a single structural gene (8). These amino acid sequence comparisons and the mRNA hybridization data of Fig. 4 make it likely that the same gene is expressed in other tissues as well as in the pancreas. The slight differences in the mRNA length among the expressing tissues shown in Fig. A may be due to differences in the length of the poly(A) tract or of the 5'- and 3'-noncoding regions. Alternatively, two or more closely related genes may be expressed in these tissues. The presence of kallikrein mRNA in the kidney at a level equal to that of the pancreas supports the hypothesis (33, 34) that the source of urinary kallikrein is synthesis by the kidney. However, because the pancreatic, submaxillary gland, and kidney kallikreins may be identical, it is not possible to exclude a contribution by kidney filtration of circulating pancreatic or submaxillary gland kallikrein into the urine.

Kallikrein mRNA exhibits an intriguing pattern of tissue-specific regulation. This mRNA, probably derived from the same structural gene in at least two tissues, is regulated to relatively high, but variable, levels in a select number of tissues as part of the expression of their differentiative phenotype. The molecular mechanisms that determine the tissue-specific regulation of kallikrein are now amenable to analysis.

We thank Drs. C. Lazure for communicating sequence data prior to its publication, W. Swain and W. J. Rutter for providing part of the ds cDNA library, and H. Martinez for use of the computer programs for DNA sequence analysis. We also thank Ms. J. S. Stary for technical assistance and Ms. M. Rotondi for preparing the manuscript. This research was supported by Grant AM27430 from the National Institutes of Health and Grant PCM8006231 from the National Science Foundation. G.H.S. was a recipient of a Virginia Lazenby O'Hara Postdoctoral Fellowship. J.-C.D. was a recipient of a Fogarty International Center Research Fellowship.

- 1. Thomas, K. A., Baglan, N. C. & Bradshaw, R. A. (1981) J. Biol. Chem. 256, 9156-9166.
- Bothwell, M. A., Wilson, W. H. & Shooter, E. M. (1979) J. Biol. 2. Chem. 254, 7287-7294.
- Lazure, C., Seidah, N. G., Thibault, G., Boucher, R., Genest, 3 J. & Chretien, M. (1981) Nature (London) 292, 383-384.
- Fiedler, F. (1979) in Bradykinin, Kallidin and Kallikrein. Hand-4. book of Experimental Pharmacology, ed. Erdos, E. G. (Springer Verlag, Berlin), Vol. 25, Suppl., pp. 103-161.
- Fiedler, F. & Fritz, H. (1981) Hoppe-Seyler's Z. Physiol. Chem. 5. 362, 1171-1175.
- 6. Matsas, R., Proud, D., Nustad, K. & Bailey, G. S. (1981) Anal. Biochem. 113, 264-270.
- 7. Fiedler, F. & Gebhard, W. (1980) Hoppe-Seyler's Z. Physiol. Chem. 361, 1661-1671.

- Tschesche, H., Mair, G., Godec, G., Fiedler, F., Ehret, W. 8. Hirschauer, C., Lemon, M. & Fritz, H. (1979) Adv. Exp. Med. Biol. (Kinins II: Biochemistry Pathophysiology and Clinical Aspects) Vol. 120, Part A, pp. 245-260.
- Chirgwin, J. M., Przybyła, A. E., MacDonald, R. J. & Rutter, W. J. (1979) Biochem. 24, 5294-5299. 9.
- Alwine, J. C., Kemp, D. J. & Stark, E. R. (1977) Proc. Natl. Acad. 10 Sci. USA 74, 5350-5354.
- Bailey, J. M. & Davidson, N. (1976) Anal. Biochem. 70, 75-85. 11. Stellwag, E. J. & Dahlberg, A. E. (1980) Nucleic Acids Res. 8, 12.
- 299-317 MacDonald, R. J., Stary, S. J. & Swift, G. H. (1982) J. Biol. 13. Chem. 257, 9724-9732.
- Rigby, P. W. J., Dieckman, M., Rhodes, C. & Berg, P. (1977) J. Mol. Biol. 113, 237-251. 14
- Aviv, H. & Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1408-15. 1412.
- MacDonald, R. J., Swift, G. H., Quinto, C., Swain, W., Pictet, 16 R. L., Nikovits, W. & Rutter, W. J. (1982) Biochem. 21, 1453-1463
- Goodman, H. M. & MacDonald, R. J. (1980) Methods Enzymol. 17. 68, 75-90
- 18 Roychoudhury, R., Jay, E. & Wu, R. (1976) Nucleic Acids Res. 3, 101 - 116
- 19. Land, H., Grez, M., Hauser, H., Lindenmaier, W. & Schutz, G. (1981) Nucleic Acids Res. 9, 2251-2266. Maxam, A. & Gilbert, W. (1980) Methods Enzymol. 65, 499-560.
- 90
- Hagenbuchle, O., Santor, M., Steitz, J. A. & Mans, R. J. (1978) 21. Cell 13, 551-563.
- 22. MacDonald, R. J., Crerar, M. M., Swain, W. F., Pictet, R. L., Thomas, G. & Rutter, W. J. (1980) Nature (London) 287, 117-122
- Proudfoot, N. J. & Brownlee, G. G. (1976) Nature (London) 263, 23. 211-214.
- Richards, R. I., Catanzaro, D. F., Mason, A. J., Morris, B. J., Baxter, J. D. & Shine, J. (1982) J. Biol. Chem. 257, 2758-2761. 24.
- 25. Thomas, K. A., Baglon, N. C. & Bradshaw, R. A. (1981) J. Biol. Chem. 256, 9156-9166.
- Blobel, G., Walter, P., Change, C. N., Goldman, B. M., Erick-26 son, A. H. & Lingappa, U. R. (1979) Symp. Soc. Exp. Biol. 33, 9-36.
- 27. Carne, T. & Scheele, G. (1982) J. Biol. Chem. 257, 4133-4140.
- Young, C. L., Barker, W. C., Tomaselli, C. M. & Dayhoff, M. 28. O. (1978) in Atlas of Protein Sequence and Structure, ed. Dayhoff, M. O. (National Biomedical Research Foundation, Silver Spring, MD), Vol. 5, Suppl. 3, pp. 73–93. Shotton, D. M. & Watson, H. C. (1970) Nature (London) 225,
- 29. 802-806.
- 30. Lemmon, M., Fiedler, F., Forg-Brey, B., Hirschauer, C., Leysath, G. & Fritz, H. (1979) Biochem. J. 177, 159-168.
- Blobel, G. & Sabitini, D. (1971) in Biomembranes, ed. Manson, 31. L. A. (Plenum, New York), Vol. 2, pp. 193-195.
- Lazure, C., Seidah, N. G., Thibault, G., Genest, J. & Chretien, 32. M. (1981) in Peptides: Synthesis, Structure and Function. Proceedings of the VIIth American Peptide Symposium, eds. Rich, D. H. & Gross, M. R. (Pierce Chemical, Rockford, IL), pp. 517-520.
- 33. Nustad, K., Vaaje, K. & Pierce, J. V. (1975) Br. J. Pharmacol. 53, 229-234
- Roblero, J., Croxatto, H., Garcia, R., Corthorn, J. & DeVito, E. 34. (1976) Am. J. Physiol. 231, 1383-1389.