Journal of Cellular Biochemistry 33:65–75 (1987)
Proteases in Biological Control and Biotechnology 147–157

Role of Glandular Kallikreins as Growth Factor Processing Enzymes: Structural and Evolutionary Considerations

Paul J. Isackson, Joan C. Dunbar, and Ralph A. Bradshaw

Department of Biological Chemistry, College of Medicine, University of California, Irvine, California 92717

Hormones and growth factors are generally released from larger precursors by limited proteolysis. The causative agents remain poorly defined with respect to location and properties. One subset of proteases, the glandular kallikreins, have been implicated in a few cases, in part because of their specific association with mature forms of some hormones. However, limited distribution and low copy number in some species cast doubt on this hypothesis, and they may well play other physiological functions that remain to be elucidated.

Key words: **precursor, hormone, limited proteolysis, submandibular gland, prostate, nerve growth factor, epidermal growth factor**

Serine proteases act as regulatory agents in a broad range of biological activities including zymogen activation of digestive enzymes, blood clotting, and prohormone activation. Some of the serine proteases such as trypsin have general proteolytic activity within the context of their substrate recognition pattern, whereas many others have very limited substrate specificities as required for their narrow physiological functions. A subgroup of serine proteases of this latter type is the glandular kallikreins of the mouse submandibular gland. Glandular kallikreins are enzymes closely related to trypsin that specifically process kininogen to produce kallidin [1]. They have been most thoroughly studied from porcine pancreas, but they are also found abundantly in a number of other exocrine tissues and fluids, including the submandibular glands and urine [1]. The male mouse submandibular gland is one of the richest sources of

Abbreviations used: NGF, nerve growth factor; EGF, epidermal growth factor; EGF-BP, epidermal growth factor binding protein; pMK-1, mouse kallikrein protein-1; PTI, pancreatic trypsin inhibitor; BAPNA, benzoyl arginine p-nitroanilide; [^{3}H]DFP, ^{3}H-diisopropyl fluorophosphate.

Joan C. Dunbar's present address is Department of Physiology, College of Medicine, University of California, Irvine, CA 92717.

Received May 28, 1986, revised and accepted July 22, 1986.

© 1987 Alan R. Liss, Inc.

TABLE I. Known or Proposed Substrate Specificities of Glandular Kallikreins

Protease	Substrate
Kallikrein	Kininogen
Tonin	Synthetic renin substrate, angiotensinogen, angiotensin I [48]
γ-Renin	Synthetic renin substrate [49]
EGF-BP	(Epidermal growth factor)
γ-NGF	(Nerve growth factor)

enzymes of this type [2]. More recently, however, preparations of these kallikreins have been found to be heterogeneous and to contain a large number of very similar proteases with different substrate specificities [3].

The male mouse submandibular gland is also unusual in that it contains high levels of a number of bioactive proteins that are not present or are present at much lower levels in other mammals. It is the richest known source of nerve growth factor (NGF) and epidermal growth factor (EGF) and contains, in addition, renin, a thymocyte transformation factor, two mesenchymal growth factors, and a glucagon like peptide [4–6]. Mouse submandibular NGF and EGF are each associated with specific kallikreins that have been suggested to be involved in the processing of their precursors. Other kallikreins present, which comprise approximately 10% of the total protein produced by these glands, may be specific processing enzymes for the other bioactive peptides that are also abundant.

A general physiological role for glandular kallikreins as specific processing enzymes for hormone and growth factor precursors has been proposed [3]. This is based on the known specificity of porcine pancreatic kallikrein for kininogen, rat tonin for angiotensinogen, and the association of gamma-NGF with NGF and EGF-binding protein (EGF-BP) with EGF. Hormones that have been proposed to be processed by kallikreins are summarized in Table I.

There are substantial sequence data available for the glandular kallikreins. The partial or complete sequences of 18 kallikreins (11 of which are from the mouse submandibular gland) are shown in Table II. A comparison of these sequences along with the three-dimensional structures determined by crystallographic analyses available for porcine pancreatic kallikrein [7] and the closely related trypsin [8] should provide an excellent system for determining how processing enzymes specifically recognize hormone precursor substrates.

MOUSE α-NGF AND γ-NGF SUBUNITS

NGF is found in the male mouse submandibular gland as a high-molecular-weight complex designated 7S NGF that contains three different polypeptide chains: alpha, beta, and gamma [9,10]. The complex consists of a stable dimer of beta subunits, which possess the neurite growth-promoting activity, and two of each of the alpha and gamma subunits. The complex has an association constant of approximately 10^{10} M^{-1} and is sufficiently stable to remain associated during gel filtration, column chromatography, or native gel electrophoresis [11,12]. Each of the subunits has been partially or fully sequenced at the amino acid level [13–15], and full-length cDNAs of

each have been isolated and characterized [16–19]. The cDNA sequence of beta-NGF [16,17] predicts a precursor of 307 amino acids which must be processed at both the amino and carboxyl termini to produce the mature 118 amino acid beta subunit [13]. A peptide of 187 amino acids is located N-terminal to the mature beta subunit in the precursor and may require multiple processing steps, whereas only two amino acids, an Arg-Gly dipeptide, must be removed from the C terminus [16,17].

Amino acid sequence analysis of the gamma subunit demonstrated the protein to be a serine protease closely related to trypsin and strikingly similar to glandular kallikrein of the porcine pancreas [14] (Table II). Earlier studies by Greene et al [20] had identified this subunit as an esteropeptidase specific for substrates containing arginine residues and a lesser ability to cleave after lysine residues. This activity led to the prediction by Frazier et al [21] that gamma-NGF acts as a processing enzyme for beta NGF. Consistent with this hypothesis is the apparent association of the gamma subunit with the C-terminal arginine residue of beta-NGF. Enzymatic removal of the C-terminal arginine residue of beta-NGF eliminates the gamma-beta interaction [22]. Thus, the gamma subunit probably interacts with this arginine residue through its active site, particularly considering its known arginine preference, and may cleave the Arg-Gly peptide for the C terminal of beta-NGF. In such a model, it would subsequently remain tightly associated with the mature beta subunit in a manner somewhat similar to the trypsin-trypsin inhibitor interaction. The role of gamma-NGF in effecting the required N-terminal processing is more obscure. It has been shown by Berger and Shooter [23] to process cell-free extracts containing beta precursors to the mature protein, but other enzymes, such as trypsin, acted in a similar manner.

The alpha subunit of NGF, which was the last of the 7S subunits to be characterized, does not exhibit any known biological activity. Sequence analysis [15,19,24] provided the surprising result that the alpha subunit is also a member of the glandular kallikrein family with 78% sequence identity with gamma-NGF. The sequence, however, contains several novel amino acid changes relative to other serine proteases which render the protein catalytically inactive. An Arg→Gln substitution at amino acid −1 (Table II) prevents cleavage of the activation peptide which is normally necessary to achieve an active conformation. Five amino acids are deleted within the N-terminal region of the mature alpha subunit, including three residues of the Ile/Val-Ile/Val-Gly-Gly sequence that is absolutely conserved in other serine proteases. The highly conserved Gly-189 residue near the active site is replaced by a histidine residue, and an Asp→Tyr substitution at 185 results in a bulky side chain at the bottom of the substrate binding pocket which would prevent interaction with arginine or lysine substrate residues. Consistent with these radical sequence changes, the alpha subunit is catalytically inactive, and ^{3}H-diisopropyl fluorophosphate ([^{3}H]DFP) binds at a rate at least 1,000-fold lower than that of gamma NGF (D. Raben et al, unpublished observations).

The manner in which alpha-NGF interacts with beta-NGF has not been established. It is clear that it is not a processing enzyme, but it does stabilize the 7S NGF complex. It appears to interact only with the beta subunit, since the alpha and gamma subunits do not form a complex [12]. It will be interesting to determine which part of the beta subunit is recognized by alpha-NGF and whether it recognizes beta-NGF through the residues that are part of the active site in other serine proteases.

TABLE II. Comparison of Trypsin and Known Glandular Kallikrein Sequences*

rTP					M	S	A	L	L	I	L	A	L	V	G	A	A	V	A	F	P	L	E	D	D	D	
rPS	M	P	V	T	M	W	F	L	I	L	F	L	A	L	S	L	G	R	N	D	A	A	P	P	V	Q	S
rS2					M	W	L	Q	I	L	S	L	V	L	S	V	G	R	I	D	A	A	P	P	G	Q	S
rS3					M	W	F	L	I	L	F	L	A	L	S	L	G	Q	I	D	A	A	P	P	G	Q	S
mGK-1					M	W	F	L	I	L	F	L	A	L	S	L	G	G	I	D	A	A	P	P	V	Q	S
mEGFBP-B					M	W	F	L	I	L	F	P	A	L	S	L	G	G	I	D	A	A	P	P	L	Q	S
mγ-NGF					M	W	F	L	I	L	F	L	A	L	S	L	G	G	I	D	A	A	P	P	V	Q	S
mα-NGF					M	W	F	L	I	L	F	L	A	L	S	L	G	G	I	D	A	A	P	P	V	Q	S

		1								10										20							
rTP	K	I	V	G	G	Y	T	C	P	E	H	S	V	P	Y	Q	V	S	L	N	S	G	Y	H	F	-	C	G
pPK		I	I	G	G	R	E	C	E	K	N	S	H	P	W	Q	V	A	I	Y	H	Y	S	S	F	Q	C	G
gpPK		V	I	G	G	Q	E	C	A	R	D	S	H	P	W	Q	A	A	V	Y	Y	Y	S	D	I	K	C	G
rPS	R	V	V	G	G	Y	N	C	E	M	N	S	Q	P	W	Q	V	A	V	Y	Y	F	G	E	Y	L	C	G
rS2	R	I	V	G	G	Y	K	C	E	K	N	S	Q	P	W	Q	V	A	V	I	N	-	-	E	Y	L	C	G
rS3	R	V	V	G	G	Y	N	C	E	T	N	S	Q	P	W	Q	V	A	V	I	-	-	G	T	T	F	C	G
mγ-R		I	V	G	G	F	K	C	E	K	N	S	Q	P	W	Q	V	A	V	Y	Y	A	K	E	Y	L	C	G
mGK-1	R	I	V	G	G	F	K	C	E	K	N	S	Q	P	W	H	V	A	V	Y	R	Y	K	E	Y	I	C	G
mEGFBP-B	R	V	V	G	G	F	N	C	E	K	N	S	Q	P	W	Q	V	A	V	Y	Y	Q	K	E	H	I	C	G
mEGFBP-A		I	L	G	G	F	K	C	E	K	N	S	Q	P	W	Q	V	A	V	Y	Y	L	D	E	Y	L	C	G
mEGFBP-S		I	V	G	G	F	K	C	E	K	N	S	Q	P	W	H	V	A	V	Y	R	Y	X	E	Y	I	C	G
mγ-NGF	R	I	V	G	G	F	K	C	E	K	N	S	Q	P	W	H	V	A	V	Y	R	Y	T	Q	Y	L	C	G
mα-NGF	Q	V	-	-	-	-	D	C	E	-	N	S	Q	P	W	H	V	A	V	Y	R	F	N	K	Y	L	C	G

				30										40										50				
rTP	G	S	L	I	N	D	Q	W	V	V	S	A	A	H	C	Y	K	S	R	I	Q	V	R	L	G	E	H	N
pPK	G	V	L	V	N	P	K	W	V	L	T	A	A	H	C	K	N	D	N	Y	E	V	W	L	G	R	H	N
gpPK	G	V	L	V	D	P	Q	W	V	L	T	A	A	H	C	I	N	D	T	N	Q	V	K	L	G	R	H	N
rPS	G	V	L	I	D	P	S	W	V	I	T	A	A	H	C	A	T	D	N	Y	Q	V	W	L	G	R	N	N
rS2	G	V	L	I	D	P	S	W	V	I	T	A	A	H	C	Y	S	N	N	Y	Q	V	L	L	G	R	N	N
rS3	G	V	L	I	D	P	S	W	V	I	T	A	A	H	C	Y	S	K	N	Y	R	V	L	L	G	R	N	N
mγ-R	G	V	L	V	D	X	N	W	V	L	T	A	A															
mGK-1	G	V	L	L	D	A	N	W	V	L	T	A	A	H	C	Y	Y	E	K	N	N	V	W	L	G	K	N	N
mPI72							N	W	V	L	T	A	A	H	C	H	N	D	K	Y	K	V	W	L	G	K	N	N
mEGFBP-B	G	V	L	L	D	R	N	W	V	L	T	A	A	H	C	Y	V	D	Q	Y	E	V	W	L	G	K	N	K
mEGFBP-A	G	V	L	L	D	R	N	W	V	L	T	A	A															
mEGFBP-S	G	V	L																									
mγ-NGF	G	V	L	L	D	P	N	W	V	L	T	A	A	H	C	Y	D	D	N	Y	K	V	W	L	G	K	N	N
mα-NGF	G	V	L	L	D	R	N	W	V	L	T	A	A	H	C	Y	N	D	K	Y	Q	V	W	L	G	K	N	N

					60										70										80			
rTP	I	N	V	L	E	G	D	E	Q	F	I	N	A	A	K	I	I	K	H	P	N	Y	S	S	W	T	-	-
pPK	L	F	E	N	E	N	T	A	Q	F	F	G	V	T	A	D	F	P	H	P	G	F	N	L	S	-	-	-
gpPK	L	F	E	D	E	D	T	A	Q	H	F	L	V	S	Q	S	V	P	H	P	D	F	N	M	S	L	L	E
rPS	L	Y	E	D	E	P	F	A	Q	H	R	L	V	S	Q	S	F	P	H	P	G	F	N	Q	D	L	I	W
rS1																										L	I	W
rS2	L	F	K	D	E	P	F	A	Q	R	R	L	V	R	Q	S	F	R	H	P	D	Y	I	P	L	I	V	T
rS3	L	V	K	D	E	P	F	A	Q	R	R	L	V	S	Q	S	F	Q	H	P	D	Y	I	P	V	F	M	R
mGK-1	L	Y	Q	D	E	P	S	A	Q	H	R	L	V	S	K	S	F	L	H	P	C	Y	N	M	S	L	H	R
mPI72	F	L	E	D	E	P	S	A	Q	H	R	L	V	S	K	A	I	P	H	P	D	F	N	M	S	L	L	N
mEGFBP-B	L	F	Q	E	E	P	S	A	Q	H	R	L	V	S	K	S	F	P	H	P	G	F	N	M	S	L	L	M
mγ-NGF	L	F	K	D	E	P	S	A	Q	H	R	F	V	S	K	A	I	P	H	P	G	F	N	M	S	L	M	R
mα-NGF	F	L	E	D	E	P	S	D	Q	H	R	L	V	S	K	A	I	P	H	P	D	F	N	M	S	L	L	N

(continued)

TABLE II. (Continued)

							90									100										110		
rTP	-	-	-	-	-	-	-	-	-	L̂	N	N	D	I	M	L	I	K	L	S	S	P	V	K	L	N	A	R
pPK	-	-	-	-	A	D	G	K	D	Y	S	H	D	L	M	L	L	R	L	Q	S	P	A	K	I	T	D	A
gpPK	P	H	N	V	L	P	N	E	D	Y	S	H	D	L	M	L	L	R	L	N	Q	P	A	Q	I	T	D	S
rPS	N	H	T	R	Q	P	G	D	D	Y	S	N	D	L	M	L	L	H	L	S	Q	P	A	D	I	T	D	G
rS1	N	H	T	R	K	P	G		D	Y	S	N	D	L	M	L	L	H	L	S	E	P	A	D	I	T	D	G
rS2	N	D	T	E	Q	P	V	H	D	H	S	N	D	L	M	L	L	H	L	S	E	P	A	D	I	T	G	G
rS3	N	H	T	R	Q	R	A	Y	D	H	N	N	D	L	M	L	L	H	L	S	K	P	A	D	I	T	G	G
mGK-1	N	R	I	Q	N	P	Q	D	D	Y	S	Y	D	L	M	L	L	R	L	S	K	P	A	D	I	T	D	V
mPI9					E	P	E	D	D	E	S	N	D	L	M	L	L	R	L	S	E	P	A	D	I	T	D	A
mPI72	E	H																										
mPI196																				S	K	P	A	D	I	T	D	A
mPMK-1						P	E	Y	D	Y	S	N	D	L	M	L	L	R	L	S	K	P	A	D	I	T	D	V
mEGFBP-B	L	Q	T	T	P	P	G	A	D	F	S	N	D	L	M	L	L	R	L	S	K	P	A	D	I	T	D	V
mEGFBP-A					X	X	E	Y	D	Y	S	N	D	L	M	L	L	R	L	S	K	P	A					
mEGFBP-S					F	P	E	Y	D	Y	S	X	D	L	M	L	L	X	L	S	K	P	A					
mγ-NGF	K	H	I	R	F	L	E	Y	D	Y	S	N	D	L	M	L	L	R	L	S	K	P	A	D	I	T	D	T
mα-NGF	E	H	T	P	Q	P	E	D	D	Y	S	N	D	L	M	L	L	R	L	S	K	P	A	D	I	T	D	V

								120										130										
rTP	V	A	P	V	A	L	P	S	A	C	A	P	A	G	T	Q	C	L	I	S	G	W	G	N	T	L	S	N
pPK	V	K	V	L	E	L	P	T	Q	E	P	E	L	G	S	T	C	E	A	S	G	W	G	S	I	E	P	G
gpPK	V	Q	V	M	P	L	P	T	Q	E	V	Q	V	G	T	T	C	R	A	L	G	W	G	S	I	D	P	D
rPS	V	K	V	I	D	L	P	I	E	E	P	K	V	G	S	T	C	L	A	S	G	W	G	S	I	T	P	D
rS1	V	K	V	I	D	L	P	T	K	E	P	K	V	G	S	T	C	L	V	S	G	W	G	S	T	N	P	S
rS2	V	K	V	I	D	L	P	T	K	E	P	K	V	G	S	T	C	L	A	S	G	W	G	S	T	N	P	S
rS3	V	K	V	I	D	L	P	T	E	E	P	K	V	G	S	I	C	L	A	S	G	W	G	M	T	N	P	S
mGK-1	V	K	P	I	A	L	P	T	E	E	P	K	L	G	S	T	C	L	A	S	G	W	G	S	I	I	P	V
mPI9	V	K	P	I	A	L	P	T	E	E	P	K	L	G	S	T	C	L	V	S	G							
mPI196	V	K	P	I	A	L	P	T	K	E	P	K	P	G	S	T	C											
mPMK-1	V	K	P	I	A	L	P	T	E	E	P	K	L	G	S	T	C	L	A	S	G	W	G	S	I	T	P	T
mEGFBP-B	V	K	P	I	A	L	P	T	K	E	P	K	P	G	S	T	C	L	A	S	G	W	G	S	I	T	P	T
mγ-NGF	V	K	P	I	T	L	P	T	E	E	P	K	L	G	S	T	C	L	A	S	G	W	G	S	I	T	P	T
mα-NGF	V	K	P	I	T	L	P	T	E	E	P	K	L	G	S	T	C	L	A	S	G	W	G	S	T	T	P	I

	40									150										160								
rTP	-	-	G	V̂	N	N̂	P	D	L	L	Q	C	V	D	A	P	V	L	S	Q	A	D	C	E	A	A	Y	P
pPK	P	D	D	F	E	F	P	D	E	I	Q	C	V	Q	L	T	L	L	Q	N	T	F	C	A	H	A	H	P
gpPK	P	A	H	P	V	F	P	D	E	L	Q	C	V	G	L	E	I	L	P	S	K	N	C	D	D	A	H	I
rPS	-	-	G	L	E	L	S	D	D	L	Q	C	V	N	I	D	L	L	S	N	E	K	C	V	E	A	H	K
rS1	-	-	E	W	E	F	P	D	D	L	Q	C	V	N	I	H	L	L	S	N	E	K	C	I	K	A	Y	K
rS2	-	-	E	M	V	V	S	H	D	L	Q	C	V	N	I	H	L	L	S	N	E	K	C	I	E	T	Y	K
rS3	-	-	E	M	K	L	S	H	D	L	Q	C	V	N	I	H	L	L	S	N	E	K	C	I	E	T	Y	K
mγ-R				W	Q	K	P	D	D	L	Q	C	M	F	T	K	L	L	P	N	E	N	C	H	K	A	H	I
mGK-1	-	-	K	F	Q	Y	A	K	D	L	Q	C	V	N	L	K	L	L	P	N	E	D	C	D	K	A	Y	V
mGK-2					Q	N	A	K	D	L	Q	C	V	N	L	K	L	L	P	N	E	N	C	D	K	N	H	N
mPI72				Y	E	Y	P	D	E	L	Q	C	V	N	L	K	L	L	P	N	E	D	C	A	K	A	H	I
mPMK-1	-	-	R	W	Q	K	S	D	D	L	Q	C	V	F	Y	T	L	L	P	N	E	N	C	A	K	V	Y	L
mEGFBP-B	-	-	R	W	Q	K	S	D	D	L	Q	C	V	F	I	T	L	L	P	N	E	N	C	A	K	V	Y	L
mEGFBP-A				X	Q	N	A	D	D	L	Q	C	V		L		L	L	P	N	E	D	C	A	K			
mEGFBP-S				F	Q	N	A	K	D	L	Q	C	V	N	L	K	L	L	P	N	E	D	C	X	K	A	X	I
mγ-NGF	-	-	K	F	Q	F	T	D	D	L	Y	C	V	N	L	K	L	L	P	N	E	D	C	A	K	A	H	I
mα-NGF	-	-	K	F	K	Y	P	D	D	L	Q	C	V	N	L	K	L	L	P	N	E	D	C	D	K	A	H	K

(continued)

TABLE II. (Continued)

		170									180										190						
rTP	G	E	I	T	S	S	M	I	C	V	G	F	L	E	G	G	K	D	S	C	Q	G	D	S	G	G	P	V
pPK	B	K	V	T	E	S	M	L	C	A	G	Y	L	P	G	G	K	D	T	C	M	G	D	S	G	G	P	L
gpPK	A	N	V	T	G	T	M	L	C	A	G	D	L	A	G	G	K	D	T	C	V	G	D	S	G	G	P	L
rPS	E	E	V	T	D	L	M	L	C	A	G	E	M	D	G	G	K	D	T	C	K	G	D	S	G	G	P	L
rS1	E	K	V	T	D	L	M	L	C	A	G	E	L	E	G	G	K	D	T	C	R	G	D	S	G	G	P	L
rS2	D	N	V	T	D	V	M	L	C	A	G	E	M	E	G	G	K	D	T	C	A	G	D	S	G	G	P	L
rS3	N	I	E	T	D	V	T	L	C	A	G	E	M	D	G	G	K	D	T	C	T	G	D	S	G	G	P	L
mγ-R	L	K	V	T	D	L	M	L	X	X	I	E	M	X	E													
mGK-1	Q	K	V	T	D	V	M	L	C	A	G	V	K	G	G	G	K	D	T	C	K	G	D	S	G	G	P	L
mGK-2	K	K	V	T	D	V	M	V	C	V	G	E	M	D	G	G	K	D	T	C	V	G	D	S	G	G	P	L
mPI72	E	K	V	T	D	D	M	L	C																			
mPMK-1	Q	K	V	T	D	V	M	L	C	A	G	E	M	G	G	G	K	D	T	C	A	G	D	S	G	G	P	L
mEGFBP-B	Q	K	V	T	D	V	M	L	C	A	G	E	M	G	G	G	K	D	T	C	A	G	D	S	G	G	P	L
mEGFBP-A	Z																											
mEGFBP-S	E	K																										
mγ-NGF	E	K	V	T	D	A	M	L	C	A	G	E	M	D	G	G	K	D	T	C	K	G	D	S	G	G	P	L
mα-NGF	M	K	V	T	D	A	M	L	C	A	G	E	M	D	G	G	S	Y	T	C	E	H	D	S	G	G	P	L

				200										210										220				
rTP	V	C	N	G	Q	L	Q	G	I	V	S	W	G	Y	G	-	C	A	L	P	D	N	P	G	V	Y	T	K
pPK	I	C	N	G	M	W	Q	G	I	T	S	W	G	H	T	P	C	G	S	A	N	K	P	S	I	Y	T	K
gpPK	I	C	D	G	V	L	Q	G	L	T	S	W	G	D	S	P	C	G	V	A	X	D	P	S	L	Y	T	K
rPS	I	C	N	G	V	L	Q	G	I	T	S	W	G	F	N	P	C	G	E	P	K	K	P	G	I	Y	T	K
rS1	I	C	D	G	V	L	Q	G	I	T	S	W	G	S	V	P	C	G	E	P	N	K	P	G	I	Y	T	K
rS2	I	C	D	G	V	L	Q	G	I	T	S	G	G	A	T	P	C	A	K	P	K	T	P	A	I	Y	A	K
rS3	I	C	D	G	V	L	Q	G	L	T	S	G	G	A	T	P	C	A	K	P	K	T	P	A	I	Y	A	K
mGK-1	I	C	D	G	V	L	Q	G	L	T	S	W	G	Y	N	P	C	G	E	P	K	K	P	G	V	Y	T	K
mGK-2	I	C	D	G	V	L	Q	G	I	T	S	W	G	S	I	P	C	S	K	P	N	A	P	G	I	Y	T	K
mPMK-1	I	C	D	G	I	L	Q	G	T	T	S	N	G	P	E	P	C	G	K	P	G	V	P	A	I	Y	T	N
mEGFBP-B	I	C	D	G	I	L	Q	G	T	T	S	N	G	P	E	P	C	G	K	P	G	V	P	A	I	Y	T	N
mγ-NGF	I	C	D	G	V	L	Q	G	I	T	S	W	G	H	T	P	C	G	E	P	D	M	P	G	V	Y	T	K
mα-NGF	I	C	D	G	I	L	Q	G	I	T	S	W	G	P	E	P	C	G	E	P	T	E	P	S	V	Y	T	K

						230										
rTP	V	C	N	F	V	G	W	I	Q	D	T	I	A	A	N	
pPK	L	I	F	Y	L	D	W	I	N	B	T	I	T	E	N	P
gpPK	V	I	E	Y	R	E	W	I	E	R	T	M	A	D	N	P
rPS	L	I	K	F	T	P	W	I	K	E	V	M	K	E	N	P
rS1	L	I	K	F	T	S	W	I	K	E	V	M	K	E	N	P
rS2	L	I	K	F	T	S	W	I	K	K	V	M	K	E	N	P
rS3	L	I	K	F	T	S	W	I	K	K	V	M	K	E	N	P
mGK-1	L	I	K	F	T	S	W	I	K	D	T	L	A	Q	N	P
mGK-2	L	I	K	F	N	S	W	I	K	D	T	M	T	K	N	A
mPMK-1	L	I	K	F	N	S	W	I	K	D	T	M	M	K	N	A
mEGFBP-B	L	I	K	F	N	S	W	I	K	D	T	M	M	K	N	A
mγ-NGF	L	N	K	F	T	S	W	I	K	D	T	M	A	K	N	P
mα-NGF	L	I	K	F	S	S	W	I	R	E	T	M	A	N	N	P

*Boxes indicate residues highly conserved in mouse glandular kallikreins. Dashed lines designate surface loops as defined by Bode et al [7]. Capped residues (^) indicate porcine pancreatic kallikrein residues that form close contacts with bovine pancreatic trypsin inhibitor [39]. Sequences shown are for rTP, rat trypsinogen I [50]; pPK, porcine pancreatic kallikrein [51]; gpPK, guinea pig prostate kallikrein (J.C. Dunbar and R.A. Bradshaw, unpublished); rPS, rSl, rS2, and rS3, four cDNA clones from rat submaxillary gland [52] (rPS was identified as rat submaxillary kallikrein and rS2 as rat tonin); mγ-R, gamma-renin of the mouse submandibular gland [49]; mGK-1 and mGK-2, genomic clones of the mouse kallikrein family [3]; mPI9, mPI72, and mPI196, cDNA clones from a mouse submandibular gland library (P.J. Isackson and R.A. Bradshaw, unpublished); mPMK-1 = pMK-1 [33]; mEGFBP-A and mEGFBP-B, two putative mouse EGF-binding protein sequences [29-31]; mEGFBP-S, EGF-binding protein sequenced by Silverman [32] and P.J. Isackson et al (unpublished), and mγ-NGF and mα-NGF, cDNA sequences of mouse gamma-NGF [18] and alpha-NGF [19].

MOUSE EGF-BINDING PROTEIN

EGF, like NGF, is also found in high concentrations in the male mouse submandibular gland and is produced as a large precursor which must undergo processing to produce the mature hormone [25,26]. The precursor to EGF, however, is very large (1,217 amino acids) compared to the mature form of EGF (53 amino acids). The precursor may be a membrane-bound protein, as judged by a strongly hydrophobic sequence similar to other putative transmembrane segments located near the C terminal and, in addition to EGF, contains seven regions with substantial sequence homology to EGF [25,26]. The EGF sequence is flanked by single arginine residues which can be processed by trypsin-like serine proteases. However, the regions homologous to EGF within the precursor are not flanked by basic residues and either are not processed or are cleaved by proteases with different residue specificities.

The mature form of EGF is also present in the male mouse submandibular gland in the form of a high-molecular-weight complex. A dimer of EGF subunits is specifically bound to two subunits of a single kallikrein, designated EGF-BP [27,28]. Based on this association, EGF-BP has been postulated to act as the processing enzyme of the large pro-EGF [27]. EGF-BP has been purified and shown to specifically bind EGF in vitro [28]. EGF-BP does not recognize NGF in this type of assay, and, conversely, gamma NGF does not bind to EGF [28].

There are several inconsistent reports in the literature regarding the sequence of EGF-BP. Two different forms of EGF-BP (A and B) have been reported to exist by Anundi et al [29]. Partial amino acid sequences have been reported for both forms [29], and a complete cDNA sequence has been published for form B [30,31]. Contrary to this result, repeated preparations of EGF-BP, which consist of a single kallikrein, have a sequence that is very similar to form A of Anundi et al [29] but not to form B (as shown in Table II) [32] (P.J. Isackson et al, unpublished). Form B [29] is, however, identical in sequence to mouse kallikrein protein-1 (pMK-1) reported by Richards et al [33]. pMK-1 was the first mouse submandibular gland kallikrein sequenced at the cDNA level and is extremely abundant, comprising 1–5% of the total mRNA. Since Peterson and co-workers [29–31] did not demonstrate specific EGF complex formation with their two putative forms of EGF-BP, it is likely that Form B is a contaminating glandular kallikrein that does not form a complex with EGF. The relationship of form A to the actual EGF-BP will require further structural data. We are currently isolating the cDNA that corresponds to EGF-BP as purified by Shooter and co-workers. Characterization of this cDNA will be very useful for comparisons with other kallikreins with different binding specificities.

GUINEA PIG PROSTATE KALLIKREIN

A second source of NGF and EGF is the prostate of several higher vertebrates [34]. Of these, bovine and guinea pig prostate appear to have the highest concentrations, and NGF has been isolated and characterized from both species [35,36]. Guinea pig prostate EGF has also been isolated and partially sequenced [37]. In all cases, the hormones appear to be closely related to those isolated from the mouse submandibular gland. However, the prostate differs from this latter tissue in one significant aspect; it does not express the large number of kallikreins. In fact, initial characterization of the gel filtration elution profile of extracts of guinea pig prostate failed to reveal any

benzoyl arginine p-nitroanilide (BAPNA) activity, the usual assay for detecting gamma-NGF and EGF-BP [35, 37]. However, evidence for the presence of some gamma-NGF (or EGF-BP)-like material was also obtained. Northern analysis of guinea pig prostate mRNA using a cDNA probe corresponding to a portion of the gamma-NGF sequence revealed cross-reactive material (P.J. Isackson, unpublished experiments), and [^{3}H]DFP labeled a single protein in a soluble extract of the tissue [38]. Subsequent enzymatic analysis revealed that this enzyme utilized benzoyl arginine ethyl ester as substrate.

Homogeneous preparations of this guinea pig esteropeptidase were obtained using a combination of gel filtration and ion exchange chromatography [38]. The pure enzyme had an apparent molecular weight of 35,000 on SDS electrophoresis; deglycosylation by chemical or enzymatic means reduced the molecular mass to 26 and 30 kD, respectively. These observations suggest there are two carbohydrate sites on the single polypeptide backbone. The complete covalent structure of the enzyme (Table II) was determined from peptides derived from chemical and enzymatic cleavages (J.C. Dunbar and R.A. Bradshaw, unpublished). The protein shares 35% identity with rat trypsin but is over 60% homologous to the gamma-subunit of NGF and porcine pancreatic kallikrein. Thus on the basis of structural considerations, the guinea pig esteropeptidase can also be viewed as a member of the kallikrein subset. Several structural features of the molecule, including 10 half-cystines, the "kallikrein loop," and the carboxyl terminal sequence, are consistent with this identification.

Although it is structurally similar to gamma-NGF (and therefore presumably to EGF-BP), no direct evidence linking this enzyme with either of these growth factors in this tissue was obtained. Gel filtration profiles clearly revealed that the kallikrein did not form a detectable complex with either of the growth factors, and covalent cross-linking experiments failed to show evidence of any interaction. These observations do not, of course, eliminate a processing role for the kallikrein in releasing the mature factors from their precursor molecules. However, if this kallikrein does serve as a processing enzyme in this tissue, the situation is significantly different from the mouse submandibular gland in that a single enzyme would have to serve as a processing enzyme for both precursors, thereby lacking the specificity observed with gamma-NGF and EGF-BP, and no complex is formed between the kallikrein and either of the mature proteins.

KALLIKREIN SUBSTRATE SPECIFICITY

Mason et al [3] have identified a cluster of genes on chromsome 7 of the mouse genome that encode more than 25 homologous glandular kallikrein genes. The complexity of cDNA libraries prepared from male mouse submandibular glands [18,19,33] indicates that a substantial number of these genes are expressed. The amino acid coding sequences of these genes are highly conserved (70–90%) and, with the exception of alpha-NGF, are predicted to code for active serine proteases with a substrate specificity for arginine residues [3] similar to porcine pancreatic kallikrein and trypsin. The crystal structure of porcine pancreatic kallikrein has been determined by Bode et al [7] and found to be extremely similar to trypsin. Kallikrein has a number of surface loops, which are more extended than those of trypsin and form a compact structure that partially protects the substrate-binding pocket of the enzyme [7]. This structure presumably restricts the protein-binding specificity of kallikreins,

and amino acid changes within the surface loops are expected to determine the protein-binding specificity. The regions of sequence corresponding to surface loops, residues 18–26, 89–93, 133–147, and 159–169, are indicated in Table II. The regions of greatest sequence variation within the mouse kallikrein genes are within these surface loops and it has therefore been suggested that this is the basis for specific recognition of different protein substrates [3].

The crystal structure has also been determined for porcine pancreatic kallikrein complexed with pancreatic trypsin inhibitor (PTI) [39]. The complex formed is very similar in conformation to that of the individual structures of trypsin and PTI. Sixteen amino acids of kallikrein are involved in close contacts with PTI, and these are shown in Table II. Six of these—residues 24, 25, 26, 93, 143, and 145—are within surface loops. With the exception of Cys 26, all contact residues within the surface loops exhibit variability among the kallikreins that have been sequenced and may therefore contribute to determining substrate specificity. Since trypsin forms similar contacts with PTI [40], it seems likely that other residues must be responsible for determining more narrow kallikrein specificities, such as the ability of gamma-NGF to recognize beta-NGF. Current efforts in our laboratory are directed toward determining contact residues of alpha and gamma-NGF with beta-NGF to answer this question.

ROLE OF GLANDULAR KALLIKREINS

Glandular kallikreins present in the mouse submandibular gland apparently recognize and may specifically cleave bonds in precursors of important bioactive peptides. Nevertheless, several lines of evidence suggest that these glandular kallikreins are not the main physiological processors of these peptides. The strongest evidence is the lower genomic copy number of kallikreins in mammals other than the mouse. Although the mouse contains more than 25 kallikrein genes [3], the rat contains only 8–10 genes [41], and the hamster has been reported to have only 2–3 kallikrein genes [42]. These reduced numbers would obviously not provide enough specificity for processing enzymes.

The speculated role of gamma NGF- and EGF-BP as processing enzymes is based primarily on their association as high-molecular-weight complexes with beta-NGF and EGF. This is in fact rather unusual; glandular kallikrein, for example, does not remain bound to kallidin after processing kininogen. There are some similarities in the interaction of trypsin and kallikreins with trypsin inhibitors in that they remain bound as a complex. A covalent bond is formed at the active site between trypsin and trypsin inhibitor, unlike the case with gamma-NGF or EGF-BP; however, it has been shown that this bond does not contribute substantially to the stability of the complex [43]. The ability of gamma-NGF and EGF-BP to form stable complexes may indicate a different or additional role for these kallikreins other than processing enzymes. The existence of alpha-NGF, a catalytically inactive kallikrein that forms a complex with beta-NGF, also suggests some other function for these proteins. The stable complex formation provides an excellent experimental system for examining the features responsible for specific recognition of NGF and EGF.

Another argument against a general role as processing enzymes is the failure to localize specific glandular kallikreins at other sources of hormone synthesis. For example, alpha and gamma-NGF have not been identified anywhere other than in the mouse submandibular gland. This is a problem that has been technically difficult to

approach because of the extremely low levels of NGF in other sources and the high degree of similarity between the different kallikreins. Nevertheless, other sources where beta-NGF has been identified such as mouse L cells and rodent iris have failed to react with antisera to alpha and gamma-NGF [44, 45]. NGF of the guinea pig prostate is not associated with BAPNA activity, which is characteristic of gamma-NGF [35]. However, as noted above, the guinea pig prostate does contain at least one kallikrein that could be involved in precursor processing.

The availability of the cDNA probe for beta-NGF has recently allowed the detection of NGF mRNA in target tissues of the sympathetic nervous system where the levels of production are too low to be conclusively identified by immunoreactivity [46,47]. cDNA probes for alpha and gamma-NGF may similarly prove to be useful for identifying other sites of kallikrein synthesis. Oligonucleotide probes prepared against regions specific to alpha and gamma-NGF (eg, amino acid residues 82–88 or 142–147) (see Table II) will allow a determination of whether the three NGF subunits are coordinately expressed in target tissues of the sympathetic nervous system and in the mouse submandibular gland.

ACKNOWLEDGMENTS

The portions of this work arising from our laboratory were supported by USPHS research grants NS19964 and AM32465. During parts of this work, P.J.I. was supported by a USPHS postdoctoral fellowship NS07068.

REFERENCES

1. Schachter M: Pharmacol Rev 31:1, 1980.
2. Webster ME: In Erdos EG (ed): "Handbook of Experimental Pharmacology." Heidelberg: Springer-Verlag, 1970 Vol XVII, p 131.
3. Mason AJ, Evans BA, Cox DR, Shine J, Richards RI: Nature 303:300, 1983.
4. Cohen S, Taylor JM, Murakami K, Michelakis AM, Inagami T: Biochemistry 11:4286, 1972.
5. Barka T: J Histochem Cytochem 28:836, 1980.
6. Lawrence AM, Tan S, Hojvat S, Kirsteins L: Science 195:70, 1977.
7. Bode W, Chen Z, Bartels K, Kutzbach C, Schmidt-Kastner G, Bartunik H: J Mol Biol 164:237, 1983.
8. Bode W, Schwager P: J Mol Biol 98:693, 1975.
9. Varon S, Nomura J, Shooter EM: Biochemistry 7:1296, 1968.
10. Bradshaw RA: Annu Rev Biochem 47:191, 1978.
11. Varon S, Nomura J, Shooter EM: Biochemistry 6:2202, 1967.
12. Silverman RE, Bradshaw RA: J Neurosci Res 8:127, 1982.
13. Angeletti RH, Bradshaw RA: Proc Natl Acad Sci USA 68:2417, 1971.
14. Thomas KA, Baglan NC, Bradshaw RA: J Biol Chem 256:9156, 1981.
15. Isackson PJ, Bradshaw RA: J Biol Chem 259:5380, 1984.
16. Scott J, Selby M, Urdea M, Quiroga M, Bell GI, Rutter WJ: Nature 302:538, 1983.
17. Ullrich A, Gray A, Berman C, Dull TJ: Nature 303:821, 1983.
18. Ullrich A, Gray A, Wood WI, Hayflick J. Seeburg PH: DNA 3:387, 1984.
19. Isackson PJ, Ullrich A, Bradshaw RA: Biochemistry 23:5997, 1984.
20. Greene LA, Shooter EM, Varon S: Biochemistry 8:3735, 1969.
21. Frazier WA, Angeletti RH, Bradshaw RA: Science 176:482, 1972.
22. Moore JB Jr, Mobley WC, Shooter EM: Biochemistry 13:833, 1974.
23. Berger E, Shooter EM: Proc Natl Acad Sci USA 74:3647, 1977.
24. Ronne H, Anundi H, Rask L, Peterson PA: Biochemistry 23:1229, 1984.
25. Scott J, Urdea M, Quiroga M, Sanchez-Pescador R, Fong N, Selby M, Rutter WJ, Bell GI: Science 221:236, 1983.

26. Gray A, Dull T, Ullrich A: Nature 303:722, 1983.
27. Taylor JM, Mitchell WM, Cohen S: J Biol Chem 249:2188, 1974.
28. Server AC, Shooter EM: Adv Protein Chem 31:339, 1977.
29. Anundi H, Ronne H, Peterson PA, Rask L: Eur J Biochem 129:365, 1982.
30. Ronne H, Lundgren S, Severinsson L, Rask L, Peterson PA: EMBO J 2:1561, 1983.
31. Lundgren S, Ronne H, Rask L, Peterson PA: J Biol Chem 259:7780, 1984.
32. Silverman RE: PhD Thesis, Washington University, St. Louis, MO, 1977.
33. Richards RI, Catanzaro DF, Mason AJ, Morris BJ, Baxter JD, Shine J: J Biol Chem 257:2758, 1982.
34. Harper GP, Barde YA, Burnstock G, Carstairs JR, Dennison ME, Suda K, Vernon CA: Nature 279:160, 1979.
35. Ruben JS, Bradshaw RA: J Neurosci Res 6:457, 1981.
36. Harper GP, Glanville RW, Thoenen H: J Biol Chem 257:8541, 1982.
37. Ruben JS, Bradshaw RA: In Barnes DW, Sirbasku DA, Sata GH (eds): "Methods for Preparation of Media, Supplements and Substrate for Serum Free Animal Cell Culture." New York: Alan R. Liss, 1984, Vol 1, p 139.
38. Dunbar JC, Bradshaw RA: J Cell Biochem 29:309, 1985.
39. Chen Z, Bode W: J Mol Biol 164:283, 1983.
40. Bode W, Schwager P, Huber R: J Mol Biol 118:99, 1978.
41. Ashley PL, MacDonald RJ: Biochemistry 24:4520, 1985.
42. Howles PN, Dickinson DP, DiCaprio LL, Woodworth-Gutai M, Gross KW: Nucleic Acids Res 12:2791, 1984.
43. Vincent J-P, Peron-Renner M, Pudles J, Lazdunski M: Biochemistry 13:4205, 1974.
44. Pantazis NJ: Biochemistry 22:4264, 1983.
45. Murphy RA, Landis SC, Bernanke J, Simonoski K: Dev Biol 114:389, 1986.
46. Shelton DL, Reichardt LF: Proc Natl Acad Sci USA 81:7951, 1984.
47. Heumann R, Korsching S, Scott J. Thoenen H: EMBO J3:3183, 1984.
48. Lazure C, Leduc R, Seidah NG, Thibault G, Genest J, Chretian M: Nature 307:555, 1984.
49. Poe M, Wu JK, Florance JR, Rodkey JA, Bennett CD, Hoogsteen K: J Biol Chem 258:2209, 1983.
50. MacDonald RJ, Stary SJ, Swift GH: J Biol Chem 257:9724, 1982.
51. Tschesche H, Mair G, Godec G, Fiedler F, Ehret W, Hirschauer C, Lemon M, Fritz H, Schmidt-Kastner G, Kutzbach K: "Advances in Experimental Medicine and Biology, 110A, Kinins II. "In Fujii S, Moriya H, Suzuki T (eds): New York: Plenum, 1979, p 245.
52. Ashley PL, MacDonald RJ: Biochemistry 24:4512, 1985.