"CHARACTERIZATION OF THE DROSOPHILA KLIP SERINE PROTEASE AND ITS GENE"

by

Joel Seth Hayflick

A DOCTORAL DISSERTATION

Presented to the Department of Cell Biology and Anatomy

and the Oregon Health Sciences University

School of Medicine

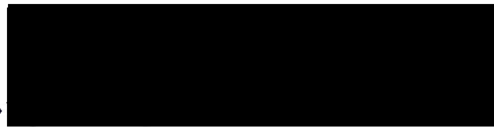
in partial fulfilment of

the requirements for the degree of

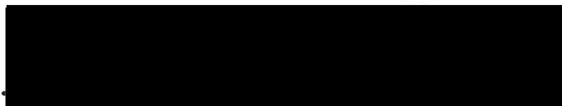
Doctor of Philosophy

© January 1992

APPROVED:



(Professor in Charge of Thesis)



(Chairman, Graduate Council)

TABLE OF CONTENTS

List of Tables....................viii

Acknowledgements....................ix

Abstract....................x

Introduction....................1

Materials and Methods....................21

Results....................46

Cloning and Characterization of dKLIP cDNAs....................46

RNA Analysis By Primer Extension....................57

Primary Structure Characterization of the Predicted dKLIP Polypeptides....................60

DKLIP-1 May Be Translocated Across Membranes....................70

DKLIP-1 is Proteolytically Active In Vivo....................73

Expression of dKLIP Transcripts in Adult Tissues....................77

The Expression of dKLIP Transcripts Are Developmentally Regulated....................82

Genomic Analysis of the dKLIP Gene....................90

DKLIP Protein Localization in Embryos....................96

Experiments Designed To Interfere With DKLIP Expression In Transgenic Flies....................107

Discussion....................110

The Predicted dKLIP-1 Polypeptide Has Unique Structural Features....110
DKLIP Transcript Expression is Developmentally Regulated.......116
Evidence for Complex Transcriptional Regulation for the dKLIP gene....118
Expression of dKLIP-1 in mammalian cells....120
Roles for Serine Proteases in Early Fly Development and Potential Substrates for the dKLIPs....121
Conclusions and Summary....127
References....133
Appendix
Sorting Signals For Proteins That Enter The Regulated Secretory Pathway....161
Introduction....161
Methods....166
Results....173
Discussion and Conclusions....194

List of Figures

Figure 1. Precursor structures for several representative proteins..3

Figure 2. Flow diagram depicting exocytic and lysosomal protein traffic..7

Figure 3. Processing pathway of pre-proinsulin in mammalian cells..8

Figure 4. Alignment of the amino acid sequences surrounding the catalytic triad residues in Kex2p, hfurin and hPC2..................48

Figure 5. Nucleotide sequence of the dKLIP-1 cDNA and deduced protein sequence..50

Figure 6. Nucleotide sequence of the dKLIP-2 cDNA and deduced protein sequence..54

Figure 7. Primer extension analysis of dKLIP transcripts............58

Figure 8. Structural features of the dKLIP-1 polypeptide with other Kex2p-like homologs..63

Figure 9. Selected amino acid sequence alignment analyses of dKLIP-1 with other Kex2-like homologs..............................65

Figure 10. Comparison of predicted drosophila Kex2-like proteins.....67

Figure 11. In Vitro translocation of dKLIP-1 across microsomal membranes..71

Figure 12. In Vivo proteolytic processing of pro-b-NGF by dKLIP-1...75
Figure 13. Northern blot analysis of head and body RNA from adult flies...78
Figure 14. In Situ hybridization to adult head and ovary sections...80
Figure 15. Temporal expression of dKLIP-1 transcripts during Drosophila development...84
Figure 16. In situ localization of dKLIP-1 transcripts in eight micron sections of embryos...86
Figure 16. In situ hybridization detection in a stage 16 whole mount embryo using an antisense digoxygenin-labeled probe............88
Figure 18. Cytological localization of the dKLIP locus on polytene chromosome preparations...91
Figure 19. Genomic Southern blot analysis of the dKLIP locus........93
Figure 20. Production of fusion proteins in Escherichia coli.........100
Figure 21. Antisera characterization using immunoprecipitation and western blot analysis...102
Figure 22. Immunocytochemical localization in whole embryos.........105
Figure 23. P-element plasmid DNA constructions used for embryo injections...109
Figure 24. Potential precursor protein substrates of the dKLIP endoproteases...125
Figure 25. Biosynthetic pathway for GnRH maturation..................165

Figure 26. Specificities of antisera against GnRH and its precursor used in this work..........168
Figure 27. Northern blot detection of hGnRH transcripts..........174
Figure 28. Processing of the GnRH precursor in two cell types..........177
Figure 29. Regulated release study of GnRH expressed from AtT-20 cells..........179
Figure 30. Schematic diagram depicting three fusion proteins that were expressed in VV infected cells..........183
Figure 31. Immunoprecipitation analysis of the fusion proteins..........186
Figure 32. Pulse-chase study of fusion protein GG1..........188
Figure 33. Pulse-chase study of fusion protein GG2..........190
Figure 34. Pulse-chase study of fusion protein GG3..........192

List of Tables

Table I. Percentage overall amino acid sequence conservation.........69

Table II. DKLIP Genomic DNA restriction fragment sizes...............95

Acknowledgements

I thank the following people for advice and support. In particular, Gary Thomas who provided a place to explore many directions, Laurel Thomas, Pat Bresnahan, Richard Leduc, Barb Thorne, Sean Molloy, Frank Quan, Xaio Li and John Adelman. Mike Forte kindly placed the fly lab at my disposal. Bill Wolfgang and Sarah Smolik-Utlaut were especially generous with their time. June Shiigi, Vicki Robertson and Paul Keating helped with figures, as well as software and darkroom advice.

I am also indebted to Louise who advised, consoled and helped push this endeavor forward and to Nate for constantly asking, "Why?"

Abstract

Endoproteolysis of precursor proteins is a common step required for the synthesis of many biologically active proteins in cells from all eukaryotes studied to date. The Kex2 precursor processing endoprotease of *Saccharomyces cerevisiae* is one of the best studied processing endoproteases. A family of newly discovered Kex2-like endoproteases have been implicated in the intracellular processing of protein precursors in metazoan cells. A strategy employing genetic manipulation of Kex2-like enzymes would allow progress in understanding the role(s) of this endoprotease family during complex processes such as development. Therefore this work describes the cloning and characterization of cDNAs encoding Kex2-like endoproteases from *Drosophila melanogaster*, an organism of intermediate complexity that is highly manipulable by genetics. Complimentary DNA sequences were cloned from a Drosophila library encoding two related polypeptides, dKLIP-1 and -2. The deduced proteins are structurally similar to the yeast Kex2 precursor protein endoprotease and its mammalian homologs (furin, PC2 and PC3 [also called PC1]) including the conserved Asp, His, and Ser residues characteristic of the catalytic triads found in the subtilisin protease family. When co-expressed in cultured cells with pro-β-NGF, dKLIP-1 greatly enhanced the endoproteolytic conversion of the precursor to mature β-NGF by cleavage at a Lys-Arg doublet. In adults, dKLIP-like

transcripts were detected in cortical regions of the CNS, fat body and female reproductive tissues, including developing oocytes. DKLIP-like immunoreactivity was localized to various tissues during early development including a population of macrophage-like cells surrounding the ventral nerve cord during late embryogenesis. Multiple transcripts and the characterization of dKLIP-1 and -2 cDNAs suggest that the regulation of RNA splicing may play an important role in the generation of endoprotease diversity. Cloning and partial characterization of the KLIP gene, which maps to the right arm of chromosome 3, position 96D1-2, is described. In addition, experiments designed to deplete or increase expression of the dKLIP proteases in transgenic fly strains are discussed.

Introduction

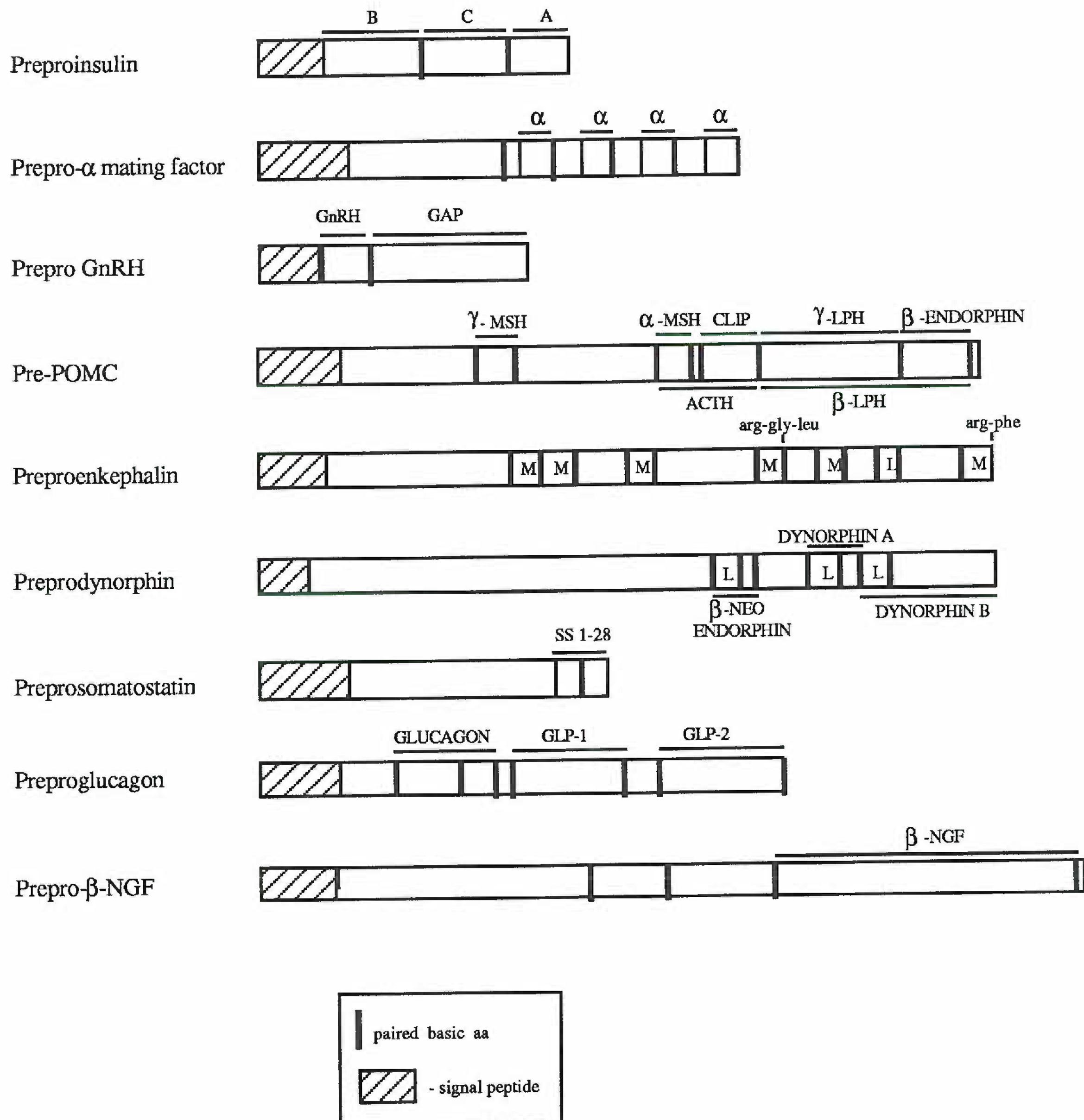
All eukaryotic cells synthesize proteins for expression on their exterior surfaces or export to the extracellular space. The biosynthetic pathway by which molecules are produced, stored and secreted has been elucidated using the insulin hormone as a model. Importantly, much of the subcellular pathway used for insulin maturation is followed by many eukaryotic proteins destined for the membrane or extracellular space and many aspects of this pathway seem to be conserved from yeast to man. Experimental systems that are tractable to genetic and biochemical studies (i.e. Saccharomyces cerevisiae) have provided much knowledge about the specific mechanisms involved in precursor protein synthesis and maturation. Based on the work presented in this dissertation as a foundation, Drosophila genetics may provide additional information regarding a specific group of endoproteases involved in precursor protein processing.

The Discovery of Precursor Protein Structure

Many proteins are initially synthesized as larger precursors from which polypeptide sequences are excised. The insulin precursor, discovered during the classic studies carried out by Donald Steiner and co-workers in the mid-1960's on the biosynthetic pathway of insulin, was the first precursor protein to be characterized (Steiner, Cunningham et al. 1967). In these studies a large insulin-like polypeptide (component

b) was isolated along with insulin from a human pancreatic islet tumor. Limited trypsin digestion of component *b* yielded insulin. It was determined that component *b* contained the B and A chains (in that order starting at the N-terminus) separated by an additional peptide sequence, termed the C chain. The entire B-C-A chain precursor was dubbed pro-insulin. With the revolution in recombinant DNA technology, in particular the techniques of cDNA cloning and DNA sequencing, a large number of precursor proteins requiring proteolytic maturation have been identified to date. From this database of cloned cDNAs several types of precursor proteins can be inferred. Some precursors contain a single peptide chain, in addition to the signal sequence, such as the honey bee toxin pro-mellitin (Sachanek, Kreil et al. 1978) or the component chains for a single peptide hormone such as insulin, or they can contain multiple peptide hormones like pro-opiomelanocortin (POMC) (Eipper and Mains 1980) (figure 1). Added to the differences in precursor structure is the generation of peptide diversity by tissue specific proteolytic processing of the same precursor polypeptide. For example, murine POMC is processed to one set of peptides in corticotrophs of the anterior pituitary lobe, but in neurointermediate lobe melanocytes the precursor is processed to a different set of peptides. Therefore, what was once believed to be unique to a few mammalian hormones has turned out to be a generalized mechanism for the biosynthesis of virtually every protein thus far identified that is destined for the plasma membrane or extracellular compartment of all eukaryotic examined.

Figure 1. Precursor structures for several representative proteins. The signal sequences are hatched boxes. Paired basic amino acid cleavage sites are shown as bold vertical lines.


Preproinsulin
B
C
A
Prepro-α mating factor
α
α
α
α
Prepro GnRH
GnRH
GAP
Pre-POMC
γ- MSH
α -MSH
CLIP
γ-LPH
β -ENDORPHIN
ACTH
β -LPH
Preproenkephalin
arg-gly-leu
arg-phe
M
M
M
M
M
L
M
Preprodynorphin
DYNORPHIN A
L
L
L
β -NEO ENDORPHIN
DYNORPHIN B
Preprosomatostatin
SS 1-28
Preproglucagon
GLUCAGON
GLP-1
GLP-2
Prepro-β-NGF
β -NGF
paired basic aa
- signal peptide

The Biosynthetic Pathway Of A Typical Secreted Protein: Human Insulin

Human insulin, like most proteins destined for export from or decoration on eukaryotic cells is co-translationally inserted and transited across the membrane of the rough endoplasmic reticulum (RER) (figure 2). Transit of nascent precursor pro-insulin molecules into the lumen of the RER depends on its signal sequence (or pre- sequence, hence the term pre-protein). The 24 amino acid signal sequence is removed co-translationally by signal peptidase (figure 3). Proper folding and disulfide bond formation of newly synthesized proinsulin occurs in the lumen of the RER. Several enzymes have been shown to catalyze or are implicated in these processes [i.e. protein disulphide isomerase (PDI) and immunoglobulin heavy-chain binding protein (Bip)] (for a recent review see (Gething and Sambrook 1992)). PDI catalyzes disulphide bond isomerization, while Bip is thought to modulate protein folding by an unknown mechanism. In addition, there is evidence that RER-Golgi transport of different newly synthesized proteins occurs at different rates suggesting that a selective process is taking place (Fitting and Kabat 1982). This process could be mediated by protein interactions that scan for properly folded nascent proteins. Vesicular transport then carries the proinsulin processing intermediate from the RER to the cis/medial/trans cisternae of the Golgi complex. During formation of proinsulin-containing secretory vesicles and/or transit to the cell surface, the A and B chains are liberated by cleavage at peptide bonds

on the C-terminal side of two sites that flank the C chain. These two sites are characterized by a pair of basic amino acids (Arg-Arg and Lys-Arg). Immuno-electronmicroscopy studies place the point in the subcellular pathway at which pro-insulin cleavage occurs as the trans-Golgi network (TGN) (Orci, Ravazzola et al. 1987b). However, biochemical studies of subcellular fractions from cells synthesizing insulin suggest that endoproteolysis of its precursor occurs in the secretory vesicles themselves (Davidson, Peshavaria et al. 1987). Following endoproteolysis, the final step in the maturation of insulin takes place in which carboxypeptidase H removes the pair of basic residues left at the COOH terminus of the B chain (Davidson and Hutton 1987).

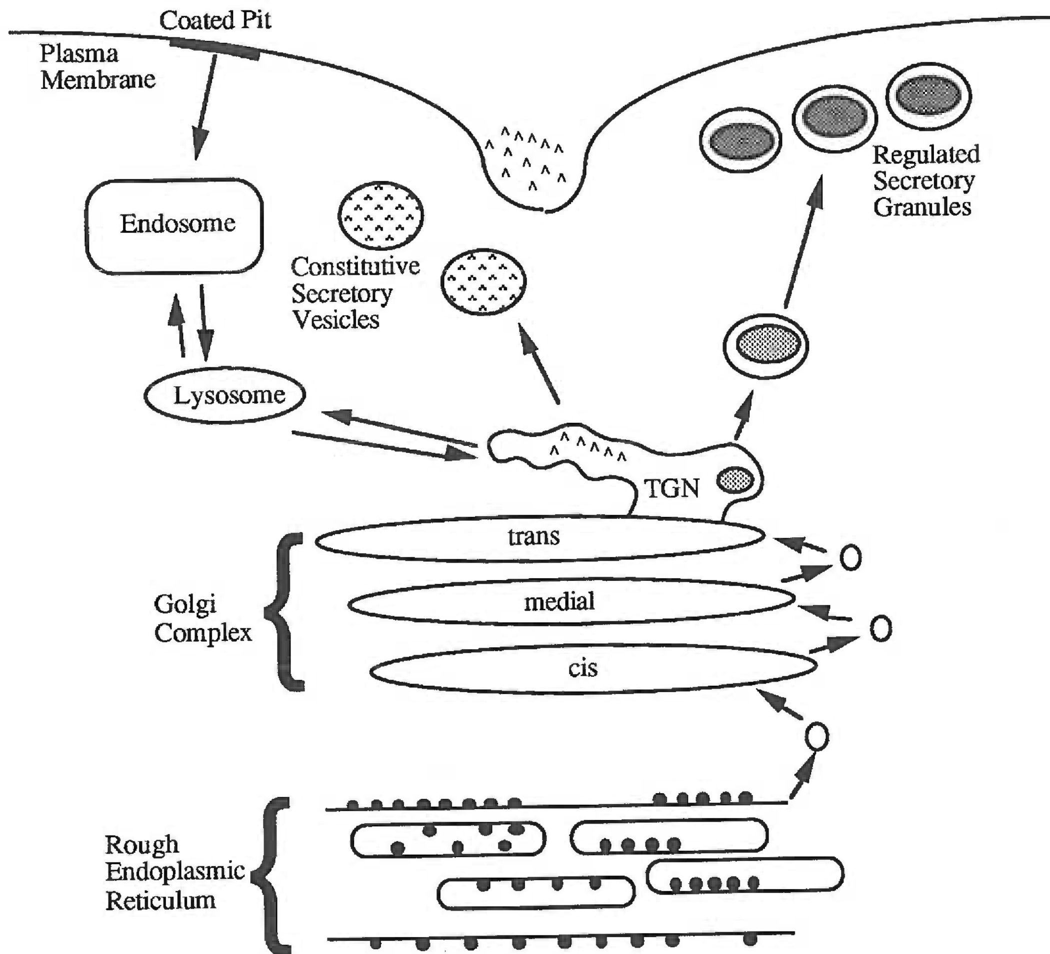
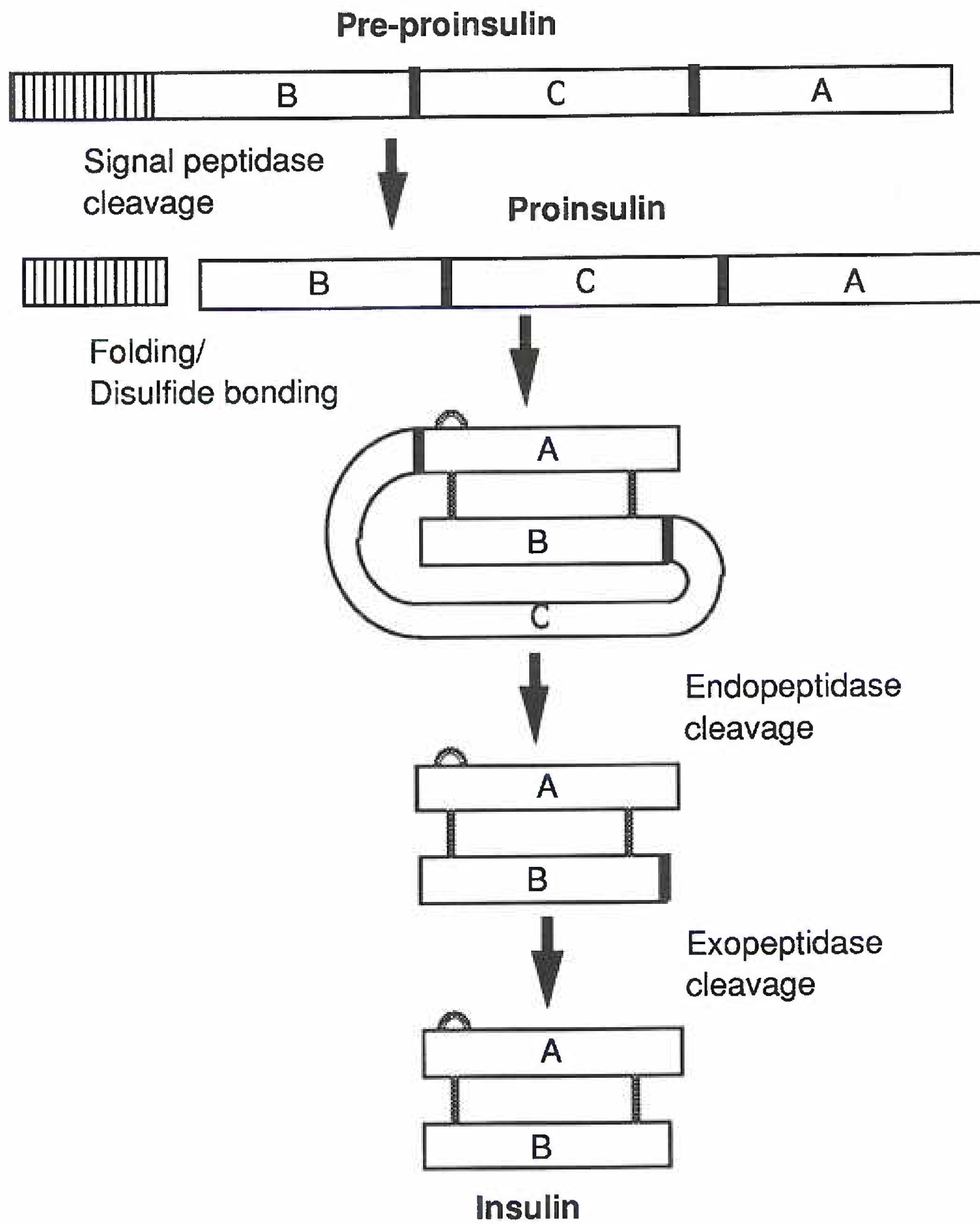


Figure 2. Flow diagram depicting exocytic and lysosomal protein traffic. Membrane-bound ribosomes are depicted as dark circles. Vesicular carriers are diagramed at the right. Each cisternae represents a separate compartment with distinct membrane components and internal environments. Constitutively-released secretory vesicle cargo is represented by carets. Condensing storage vesicle cargo is represented by increasing shades of gray. Lysosomes are shown as receiving contents by transport from both the TGN and the plasma membrane (coated pit) via an endosomal compartment.

Figure 3. Processing pathway of pre-proinsulin in mammalian cells. Synthesis and post-translational processing has been clarified by biochemical and immunochemical studies of the precursor and proteolytic activities in homologous and heterologous expression systems (Steiner, Cunningham et al. 1967) (Orci, Ravazzola et al. 1987a; Orci, Ravazzola et al. 1987b; Davidson and Hutton 1987; Davidson, Peshavaria et al. 1987; Davidson, Rhodes et al. 1988; Moore, Walker et al. 1983). The signal sequence of pre-proinsulin is marked with a striped box. Endoproteolytic cleavage sites flanking the C chain are marked by bold lines. Disulfide bonds are marked with gray lines.


Pre-proinsulin
B
C
A
Signal peptidase
cleavage
Proinsulin
B
C
A
Folding/
Disulfide bonding
A
B
C
Endopeptidase
cleavage
A
B
Exopeptidase
cleavage
A
B
Insulin

Sorting And Modification During Protein Biosynthesis

Many different proteins transit the RER/Golgi compartments destined for different fates. Possibly at the TGN/secretory vesicle interface within the β cell, the insulin molecule (or its precursor) must be sorted away from resident Golgi proteins, lysosomal hydrolases and constitutively released proteins, and segregated into storage vesicles to await regulated exocytosis in response to elevated blood glucose levels. There is some evidence for the existence of a "sortase" protein that actively recognizes some as yet unknown determinant common to proteins destined for the regulated secretory pathway such as growth hormone, prolactin or insulin (Chung, Walter et al. 1989). Regulated secretion is characterized by (i) vesicle membrane/plasma membrane fusion requiring marked increases in cytosolic free calcium concentrations as elicited by ligand/receptor interactions and second messenger activated processes, (ii) condensation of vesicular contents forming electron dense structures and (iii) relatively long vesicular half-lives on the order of 10 hours. There is evidence for a gradual, slow release (termed basal secretion) of regulated secretory vesicle contents especially in continuous cell lines in culture (Matsuuchi and Kelly 1991). Whether this phenomenon occurs in normal cells *in vivo* has yet to be reported. In contrast to the proposed sorting mechanism for regulated secretory vesicle contents, proteins destined for the surface and extracellular space are believed to transit the constitutive secretory pathway by a bulk flow or constitutive process.

Constitutively released secretory vesicles do not require increased intracellular calcium concentrations for plasma membrane fusion, nor do they undergo condensation of their contents. Their intracellular half-lives are short, about 10 minutes.

Many eukaryotic proteins receive a diverse array of modifications required for their bioactivity which typically occur as they are transported through the ER/Golgi/secretory vesicle compartments. For instance, amino terminal residues can be removed, a myriad of glycans may be added and modified, sulfation, amidation and methylation may occur. These post-translational modifications are catalyzed by specific enzymes. For example, peptidylglycine α-amidating monooxygenase (PAM) catalyzes the removal of a hydrogen atom from the C-terminal glycine residue of many peptide chains and hydrolysis of the resulting imino linkage to yield des-glycine peptide amide (Bradbury, Finnie et al. 1982). Galactosyl transferase catalyzes transfer of galactosyl residues to specific mannose sugars at the termini of processed glycan moieties. Tyrosylprotein sulfotransferase catalyzes the transfer of sulfate from 3'-phosphoadenosine 5'-phosphosulfate to tyrosine residues of proteins. PAM is a secretory vesicle resident protein while galactosyl tranferase and tyrosylprotein sulfotransferase are found in the trans-Golgi cisternae. Thus the sequential post-translational modification of a protein can be used to monitor its transit through the pathway.

Many proteins of invertebrate origin also carry post-translational modifications. These include, but are not limited to, amidation (van

Hofsten, Faye et al. 1985), glycosylation (Huber, Smith et al. 1990) and sulfation (Huttner 1987)). Since the types of modifications found on proteins of invertebrate origin are similar or identical to those found in vertebrate cells it is probable that the enzymatic machinery responsible is also highly related.

Storage And Biological Role Of Insulin

The role of fully mature insulin in the maintenance of serum glucose homeostasis is well understood. Equally well understood, is the role of endoproteolysis in the synthesis of bioactive insulin. Beta cells of the endocrine pancreas synthesize, store and release mature insulin. Regulated release of insulin comes in response to increased blood glucose levels. Insulin binding to high affinity receptors on hepatocytes inhibits glucose release. Whereas insulin binding to insulin receptors on adipocytes and muscle cells elicits glucose uptake by these cells. Glucose uptake via recruitment of specific transporter proteins is the result of a second messenger system cascade initiated by insulin binding to its receptor. Extracellular insulin binding activates the cytoplasmic tyrosine kinase domain of the receptor that then phosphorylates second messenger-generating components ultimately leading to plasma membrane fusion of vesicles containing glucose transporters. Glucose binds to the external face of the transporter leading to a conformational change which brings glucose to the cytoplasmic side of the membrane-bound transporter where it is released into the cytosol. Reduction in blood glucose levels reduces insulin

concentrations leading to internalization of glucose transporter molecules from the cell surface. In this way, insulin modulates blood glucose levels and maintains homeostasis.

Basic Residues Are Commonly Found At Precursor Protein Cleavage Sites

Comparison of proinsulin sequences from distantly related species reveals that pairs of basic amino acid residues flanking the C chains in these precursors are evolutionarily conserved. Because they are conserved, these residues probably play a critical role for endoproteolytic processing of proinsulin. Thorne and Thomas (Thorne, Caton et al. 1989; Thorne and Thomas 1990) provide evidence that the prohormone processing enzymes of the pancreatic b cell prefer cleavage sites in a model precursor protein consisting of paired basic amino acids (Lys-Arg and Arg-Arg). Since not all pairs of basic amino acids are processed, additional information may be required for processing at specific sites. For instance, higher order protein structure and the extent of carbohydrate addition may be crucial.

A wide variety of precursors possess single or multiple basic residues at which endoproteolysis occurs including prohormones, viral envelope glycoprotein precursors, pro-receptors, pro-growth factors and cellular adhesion protein precursors. From the database of precursors and their cleavage site sequences it has been proposed that there is a core set of enzymes that is capable of recognizing a few sites in a wide

variety of precursor proteins rather than a separate enzyme for each substrate (Thomas, Leduc et al. 1991).

Naturally occurring and specifically engineered mutations at the cleavage sites of a variety of precursors demonstrate the importance of specific amino acid residues for proteolysis to occur. For example, a Japanese patient suffering from insulin resistant (type A) diabetes was found to have a G-T transversion at codon 735 of her insulin pro-receptor gene. This resulted in alteration of the pro-receptor cleavage site separating the a and b chains of the mature receptor from Arg-Lys-Arg-Arg- to Arg-Lys-Arg-Ser (Yoshimasa, Seino et al. 1988). Cleavage site residues in this work are designated +1, +2, -1, -2, etc., depending on their position relative to the scissile bond. Thus the codon for the -1 position of the cleavage site in the patient's insulin pro-receptor gene described above was mutated to encode a Ser rather than an Arg residue. Systematic mutagenesis of the insulin pro-receptor codons at the -1 through -4 positions of the cleavage site by alanine substitution resulted in minimal precursor conversion of -1 and -4 alanine mutants, while -2 and -3 alanine mutants were efficiently processed (Yoshimasa, Paul et al. 1990). Similar studies of this sort in a wide variety of precursor polypeptides including influenza hemagglutinin (HA) (Kawaoka and Webster 1988) and gp160 of human immunodeficiency virus ((McCune, Rabin et al. 1988); (Freed, Myers et al. 1989)) led to similar conclusions regarding the importance of arginine residues at the -1 and -4 positions. Importantly, it seems

that an endogenous endoprotease efficiently cleaves similar sites comprised of an Arg-X-X-Arg sequence within many precursor proteins.

The cleavage site motifs discussed above (Lys-Arg, Arg-Arg, and Arg-X-X-Arg) are found in many precursor proteins synthesized from yeast to human cells. For example, cDNAs encoding an insulin-like proreceptor from Drosophila maintain the Arg-X-X-Arg cleavage site codons found in the human insulin proreceptor to be crucial for cleavage by the human endoprotease. Since the amino acid composition of the sites has been so well conserved it is likely that the endoproteases involved in catalyzing endoproteolysis are also related.

Identification Of The Endoproteases That Cleave Precursor Proteins At Basic Amino Acids

Initial approaches toward identification of the precursor protein endoproteases relied on large scale subcellular fractionation methods and biochemical characterization (Fletcher, Noe et al. 1980; Fletcher, Quigley et al. 1981; Loh and Gainer 1982). These attempts were problematic since other activities (notably lysosomal hydrolases) were difficult to separate from secretory vesicles, the presumed sites of endoproteolysis. One group led by John Hutton succeeded in isolating enzymatically pure preparations for two activities (type I and II) from transplantable rat insulinoma tissue that cleave pro-insulin at one or the other paired basic cleavage site (Davidson, Peshavaria et al. 1987; Davidson, Rhodes et al. 1988). Type I cleaves exclusively on the C-terminal side of the B-C junction (Arg-Arg) and type II cleaves

predominantly on the C-terminal side of the C-A junction (Lys-Arg) (Davidson, Rhodes et al. 1988). The enzymes are Ca^{++} dependent, acidic proteases that have yet to be fully characterized.

In *Saccharomyces cerevisiae* (baker's yeast), genetic and biochemical studies demonstrate that processing of the pro-alpha factor precursor requires Kex2p endoproteolytic activity (Fuller, Sterne et al. 1988). The product of the KEX2 locus (Kex2p) is a calcium dependent, Golgi membrane bound, subtilisin-like serine endoprotease that cleaves efficiently at the carboxyl side of Arg-Arg and Lys-Arg pairs (Fuller, Brake et al. 1989b; Mizuno, Nakamura et al. 1988; Mizuno, Nakamura et al. 1987). Since the cleavage sites in pro-alpha factor resemble those found in many metazoan precursors, Kex2p was tested for its ability to cleave in vivo mouse POMC (Thomas, Thorne et al. 1988b), human proalbumin (Bathurst, Brennan et al. 1987) and pro-protein C (Foster, Holly et al. 1991). These experiments where done by coexpressing enzyme and substrate together in cultured mammalian cells or mixing the two *in vitro*. Kex2p was found to properly cleave all of these substrates at pairs of basic residues.

Based upon the preceding results, it was hypothesized that the endoproteases in metazoans might share a high degree of amino acid and, by extension, nucleic acid similarity to Kex2p. To test this, metazoan genomic DNA or cDNA libraries were screened with the Kex2 DNA sequence yielding negative results. During a routine database screen in 1989, however, it was found that a sequence from the 5' flanking region of the

human fes/fps proto-oncogene showed a high degree of sequence similarity to Kex2 (Fuller, Brake et al. 1989a). Known as fur (fes/fps upstream region) (Roebroek, Schalken et al. 1986), this locus was cloned and found to encode a protein (furin) which has 48% amino acid identity with Kex2p in the catalytic domain (Fuller, Brake et al. 1989). Included in the conserved domain were the active site residues Asp, His and Ser. Like Kex2p, furin is a calcium dependent, membrane bound, subtilisin-like serine endoprotease (Bresnahan, Leduc et al. 1990; (Molloy, Bresnahan et al. submitted)). Coexpression of human furin with murine pro-β-nerve growth factor (Bresnahan, Leduc et al. 1990) or pro-von Willebrand factor (Wise, Barr et al. 1990) results in the proper cleavage of either substrate in the constitutive secretory pathway of cells in culture. Active site and proregion mutants demonstrate that furin, like bacterial subtilisin, is synthesized as a zymogen that is activated by an intramolecular endoproteolytic event to yield the active enzyme (Leduc, Molloy et al. submitted). Furin transcripts are found in all cell types and tissues so far investigated and presumed to be ubiquitous. Several other Kex2p-like endoproteases have also been identified. PC2 (prohormone convertase 2) and PC3 (also referred to as PC1) have been cloned and the distribution of their transcripts is far more limited (Seidah, Gaspar et al. 1990; Smeekens and Steiner 1990; Seidah, Marcinkiewicz et al. 1991; Smeekens, Avruch et al. 1991). Transcripts for these proteases are found to be limited to neuroendocrine and endocrine cell types. Preliminary results obtained

from antibody crossreactivity studies suggest that the type II endoprotease isolated from endocrine pancreatic tissues and found to cleave the pro-insulin molecule at a Lys-Arg site is PC2 (D. Bennett and J. Hutton, personal communication). Co-expression studies with PC2, PC3 and mPOMC demonstrate that each protease has unique cleavage site preferences and that the cleavage event occurs in the regulated secretory pathway of endocrine cells (Benjannet, Rondeau et al. 1991; Thomas, Leduc et al. 1991). Therefore, the hypothesis has been put forward (Bresnahan, Leduc et al. 1990; Thomas, Leduc et al. 1991) that furin cleaves its substrates (i.e. pro-growth factors and pro-envelope glycoproteins) in the constitutive exocytic pathway common to most cells, while PC2 and PC3 cleave their substrates (i.e. pro-hormones) in the regulated secretory pathway common to endocrine cells.

Research Goals and Rationale

The central issue in the study of the Kex2-like endoproteases is what are the endogenous substrates of the enzymes, as well as other protein/protein interactions. An additional area of interest is the process of development and how the Kex-like endoproteases are involved. An experimental system for addressing these questions, especially during developmental processes, is Drosophila genetics. Drosophila has for ninety years been the primary model system for geneticists from which many mutant lines have been characterized. If a mutant strain does not already exist, there are established techniques for generating and screening for mutants at a specific locus. Once mutations have been

identified, replacement with a wildtype gene or ones with specifically engineered alterations can be readily achieved. Additionally, developmental stages of Drosophila are well documented and easily studied making experiments for the study of enzyme/substrate interactions during these times tenable.

This dissertation documents the experimental results and conclusions that meet the following research goals:

I. As a foundation upon which to begin the genetic analysis of the substrate problem, a gene from *Drosophila melanogaster* (Drosophila Kex2-like endoprotease, dKLIP) was identified that encodes a protease with a high degree of sequence identity to the Kex-like proteases.

II. A detailed characterization of one of the proteases (dKLIP-1) and its mRNA was undertaken. Complimentary DNA cloning, sequence analysis, *in vitro* expression, as well as *in vivo* activity and localization are described.

III. A second cDNA encoding a related protease (dKLIP-2) was isolated and sequenced. Its relation to dKLIP-1 and the Kex2 homologs of other species is discussed.

IV. Production of transgenic Drosophila strains either over- or under-expressing dKLIP-1 was undertaken and is described. These experiments were designed to provide information regarding the developmental processes that the endoprotease may take part in and may shed light on potential substrates.

Materials and Methods

Fly Stocks and Reagents

Oregon R stocks were maintained at 25°C on 12hr light/dark schedules on a standard diet (Roberts 1986). Restriction enzymes, T7 RNA polymerase, SP6 RNA polymerase and Taq DNA polymerase were purchased from BRL, Inc. Oligonucleotides were purchased from Oligos Etc. Inc. The adult fly head gt11 library was a gift from P. Salvaterra (City of Hope, Duarte, CA). Oligo-dT cellulose was purchased from Boehringer Mannheim Biochemicals, Inc.

Cell Culture

BSC-40 cells, an African green monkey kidney epithelial cell line, and Hela cells, a human cervical epithelial cell line, were cultured in MEM (Gibco Laboratories) containing 10% heat inactivated fetal bovine serum (HyClone Laboratories).

Polymerase Chain Reaction and Library Screening

Two µg of lambda DNA from the library used below containing oligo dT primed adult head specific cDNAs were amplified with 100 pM each of two oligonucleotide pools. Pool 1, 5'-CA(C/T)GG(C/A/T)AC(C/G)CG(A/C/G)TG(T/C)GC(C/G)GG(A/T/C)GA(G/A); pool 2, 5'-CC(G/A)GC(G/A)GCCA(G/A)GGG(G/A)GCGG(A/C) (G/A)GC(G/A/T/C)GAGGT(G/T)CC. Twenty-five ml PCR reactions were performed using Taq polymerase with temperature and time parameters of 95°C, 1 min., 55°C, 1 min., 74°C, 1

min. for 30 cycles. The product was agarose gel purified and subcloned into pGEM7Zf+ (Promega Corp.). One hundred nanograms of PCR product was radio-labeled by the random primer method (Feinberg and Vogelstein 1983) and used to screen $5x10^5$ plaques of the lambda gt11 fly head library. Hybridization was performed in 50% formamide, 1% SDS, 1M NaCl, 10% dextran sulfate and 100 µg/ml sheared, denatured salmon sperm DNA at 37°C for 18 hrs. Filters were washed briefly in 2x SSC at room temperature and then in 0.1% SDS, 0.1x SSC at 65°C for 60 min. Positively hybridizing clones were purified and three phage isolates (1L, 3 and 5) were chosen for further study. The composite dKLIP-1 cDNA (see legend to Figure 5A) was ligated into pGEM7Zf+ and the resulting plasmid (pGEM7Zf+:dKLIP-1) was used for later studies. Clone 5 (dKLIP-2 cDNA) was also subcloned into pGEM for further analysis.

DNA Sequencing

DNA sequences were initially determined directly from double stranded plasmid subclones of lambda phages 3, 5, and 1L (Chen and Seeburg 1985). Fragments of lambda phages were subcloned into M13mp18 or mp19 vectors (Yanisch-Perron, Vieira et al. 1985) and sequenced using the dideoxynucleotide chain terminator method (Tabor and Richardson 1987) and Sequenase v2 (U.S. Biochemical). Reactions were run on 5% denaturing polyacrylamide wedge gels (Chen and Seeburg 1985). Sequences of overlapping restriction fragments were determined from both DNA strands.

Computer Analyses

DNA sequences were compiled and analyzed using the Genetics Computer Group Sequence Analysis Software Package v6.2, 1990 (Devereux, Haeberli et al. 1984) or the Intelligenetics, Inc. programs. The deduced protein sequence for dKLIP-1 was compared to the Kex2p homologs using the Bestfit and Genalign programs. Hydrophobicity analysis was performed using the Pepplot program.

RNA Isolation and Northern Blot Analyses

Fly heads were separated from bodies and stored at -70°C until use. Staged embryos were collected on apple juice agar plates and maintained at 25°C until harvest. One gram of whole flies or parts were crushed in liquid N_2. Total RNA was isolated as described (Cathala, Savouret et al. 1983). Ten μg of each poly-A+ selected RNA sample was fractionated by electrophoresis in 1% agarose (FMC), 6% formaldehyde, 0.02 M Hepes, 0.001 M EDTA, pH 7.8 gels and transferred to a nylon membrane (Hybond-N, Amersham). Hybridizations were carried out in 50% formamide, 5x SSC, 0.025 M $NaPO_4$, 5x Denhardt's solution, 0.005 M EDTA, 0.1% sodium pyrophosphate, 1% SDS, 100 μg/ml sheared, denatured salmon sperm DNA at 37°C for 18 hrs. Blots were washed in 2x SSC at room temperature and 0.1x SSC, 0.1% SDS at 65°C for 60 min. Filters were then exposed to X-OMAT AR film (Kodak).

***In Vitro* Expression**

RNA was synthesized *in vitro* essentially as described (Kreig and Melton 1987). Briefly, 10 µg of pGEM7Zf+:dKLIP-1 was linearized with XhoI endonuclease. Fifty µl RNA transcription reactions containing 1 µg of linearized plasmid DNA in 1x SP6 buffer, 0.01 M DTT, 50 µg/ml bovine serum albumin, 1 unit placental RNase inhibitor (Promega), 250 µM each ATP, CTP, UTP, 50 µM GTP, 250 µM GpppG and 50 units of SP6 RNA polymerase (BRL). Incubations were performed at 37°C for 5 hrs. RNA was extracted twice with phenol:chloroform, ethanol precipitated, and resuspended in 50 µl H_2O. Fifty µl *in vitro* translation reactions containing *in vitro* transcribed RNA, rabbit reticulocyte lysate (Promega) and [^{3}H]-leucine were set up as described by the manufacturer. After 30 min. at 30°C, 10% of the translation reaction was run on a 7% SDS-PAGE (Laemmli 1970) which was then processed for fluorography (Amplify, Amersham). The dried gel was exposed to X-OMAT film for 16 hrs at -70°C. For translation/translocation studies, canine pancreatic microsomes (Promega) were added to the reticulocyte reactions. Protease-digested reactions were adjusted to 0.01 M $CaCl_2$, 0.1 µg/ml proteinase K and detergent-disrupted reactions were brought to 0.1% Triton X-100. Following a 45 min. incubation on ice, the digestions were terminated by adjusting to 2 mM PMSF and analyzed by SDS-PAGE as described above.

Vaccinia Virus Recombinant Generation

Vaccinia virus (VV) strain WR was used biosafety level 2 containment precautions and infections were performed as described (Thomas, Thorne

et al. 1988a). Recombinant virus carrying the dKLIP-1 cDNA was constructed as follows. The 3.8 kb dKLIP-1 cDNA was ligated directionally into pZVneo. The plasmid pZVneo, contains the vaccinia thymidine kinase gene interrupted by a VV expression/selection cassette putting the expression of dKLIP-1 under the control of the constitutive p7.5 VV promoter. The Tn5 aminoglycoside transferase gene driven by the VV p11 promoter acts as a selectable marker by conferring neomycin resistance. The resulting plasmid (pZVneo:dKLIP-1) was cesium banded by equilibrium density centrifugation and used to introduce the cDNA into vaccinia virus by homologous recombination (Hruby, Thomas et al. 1986).

Hela cells were grown on 3.5 cm plates to 60-80% confluence in DMEM. Prior to transfection, 5 µg of pZVneo:dKLIP-1 plasmid DNA was diluted to 50 µl volume in sterile H_2O. Thirty µg of the cationic lipid DOTMA (Lipofectin™, Life Technologies, Inc.) was diluted in sterile H_2O. Plasmid DNA and DOTMA dilutions were gently combined in a polystyrene tube and allowed to form complexes by incubation at room temperature (RT) for 15 min. One plate of cells was trypsinized and counted. VV:wt stock was diluted in 0.5 ml PBS-MB (Phosphate-Buffered Saline with 1mM $MgCl_2$, 0.1% bovine serum albumin) to a multiplicity of infection (moi) of 1 plaque-forming unit (pfu) per cell. The Hela cell monolayers covering two 3.5 cm plates were rinsed briefly with 1ml warm PBS, 1mM $MgCl_2$ (PBS-M) which was aspirated and 0.5 ml/plate VV inoculum was added. The plates were gently rocked at 10 min intervals for a total of 30 min at RT. Aspirate the inoculum from the cells and rinse

the monolayer 2X with PBS-M. Add 1 ml of serum-free 202 medium and add dropwise the lipid-DNA complexes. After 3 hrs the lipid-DNA complexes were replaced with DMEM, 20% heat-inactivated fetal calf serum (FCS) and incubated for 24 hrs.

Crude stocks were prepared from the infected/transfected cells by scraping the cells from the plates prior to their lifting from the plastic. Cells were pelleted in a conical tube at 200xg for 5 min and after washing in 5 ml PBS-M, resuspended in 1 ml PBS-M. The cells were lysed by 3 cycles of snap freeze/warm thaws and the crude stock stored in sealed tubes at -70°C.

Assuming a titer of $1x10^7$ pfu/ml, serial dilutions of the crude stock were used to infect confluent monolayers of BSC-40 cells in 10 cm plates as described above. After 2 days in culture, 0.5 ml of 0.33% neutral red dye solution was added to the medium to stain the monolayer in which several hundred clearly defined plaques were visible and incubated at 37°C for 1 hr. The medium was removed and overlayed with one well marked, dry, circular nylon membrane (Colony/Plaque screen™, NEN-DuPont). The cell monolayer was transferred to the membrane by gently dabbing the membrane with a Kimwipe™ soaked in PBS-M. The nylon membrane was then lifted from the plate and placed cell-side up on a circular piece of blotting paper soaked in PBS-M. A dry, marked nitrocellulose membrane circle was then placed against the cells and pressed firmly together. Using a hand-held hole punch, holes were made around the membrane sandwich for later orientation of the membranes.

The membranes were separated, the nitrocellulose sheet stored at -70°C in a sealed dish and the nylon membrane processed for DNA hybridization as with any conventional plaque-lift, except after immobilization of the DNA to the filter, they were treated with Proteinase K for 30 min at 55°C. The filter was prehybridized in 1M NaCl, 10% dextran sulfate, 50% formamide, 1% SDS for 6 hours at 37°C and hybridized overnight in the same solution to a [^{32}P]-labeled dKLIP-1 cDNA probe. The filter was then washed in 2x SSC at RT, for 15 min and 0.1x SSC, 0.1% SDS at 55°C for 1 hr. and blotted to semi-dryness. An autoradiograph was then made from the filter and the location of the orientation holes marked on the film.

Plaques that hybridized to the cDNA probe were isolated from the nitrocellulose filter once it was thawed by lining up the orientation holes marked on the film with those on the filter and carefully cutting around the plaque with a ethanol-sterilized razor blade. The plaque was transferred to a sterile 1.5 ml Eppendorf tube containing 200 µl PBS-MB, vortexed, snap frozen and thawed 3 times and sonicated by bath sonication for 10 sec. Five and twenty microliters of this isolate were used to infect monolayers of BSC-40 cells for at least two further rounds of plaque purification. Once a pure population of positively-hybridizing plaques was attained the isolate was used to infect monolayers of BSC-40 cells for agar-overlay. The inoculum was removed and overlayed with 1x MEM, 5% FCS, 0.75% SeaPlaque™.agarose (FMC Bioproducts) which had been cooled to 48°C. The overlay was allowed to

cool at 37°C for 15 min and then incubated for 48 hrs at 37°C. A second overlay was prepared with 0.5% agarose, 0.05% neutral red, allowed to cool and returned to the incubator for 1 hr or until the dye had diffused sufficiently to stain the monolayer. Well isolated plaques were removed by boring through the overlay and scraping the infected cells gently with the small-bore end of a sterile Pasteur pipet. The agar plug was placed in 200 µl of PBS-MB and freeze/thawed 3 times to lyse infected cells. Half of the agar plug plaque isolate (100 µl) was used to infect BSC-40 cells grown in a single 1.6 cm well for preparation of a crude stock. After 48 hrs incubation, the medium was removed and 0.2 ml 0.25% trypsin in versine was added and incubated at 37°C for 15 min to dislodge the cells. Then 0.2 ml of PBS-MB, 0.3% FCS was added to terminate the trypsinization. This crude stock was freeze/thawed 3 times and used for the preparation of partially purified stock of virus. In addition, the identity of the crude stock was verified by passing 0.1 ml of the stock through a slot blot apparatus and probing for hybridizing dKLIP-1 sequences.

A partially purified stock of VV:dKLIP-1 was prepared by assuming the crude stock had a titer of 10^7 pfu/ml and infecting four 150 cm plates of BSC-40 cells at an moi of 0.005 pfu/cell. After 48 hours infection the monolayer was checked under the microscope to determine the extent of cytopathology. Prior to lifting from the substratum, the cells were harvested by scraping into the medium and transferred to 50 ml conical tubes and spun at 200xg for 5 min. The cell pellet was

resuspended in 10 ml PBS-M to rinse and pelleted again. The cell pellet was resuspended in 5 ml 0.01 M Tris pH 9 (4°C) and kept on ice throughout the rest of the prep. The cell suspension was lysed by passage through a chilled 25 ml glass dounce homogenizer (Kontes A pestle) with 25 strokes and transferred to a 15 ml conical bottom tube. The lysate was spun at 450 x g for 5 min in a refrigerated swinging bucket centrifuge. The supernatent was transferred to an SW28 tube (Ultraclear™, Beckman) and held on ice. The lysate pellet was resuspended in 5 ml of 0.01 M Tris pH 9 and spun at 450xg for 5 min. The lysate supernatents were pooled in the SW28 tube and 16 mls of sterile 36% (w/v) sucrose/0.01 M Tris pH 9 was layered underneath. A balance tube was set up and these were spun at 18,000 rpm for 80 min at 4°C. The pellet was resuspended in 1 ml Tris pH 9 and broken up using a pre-chilled 5 ml dounce homogenizer with a Teflon pestle. Aliquots of 50 µl were distributed to cryovials, snap frozen and stored at -70°C until use. The partially purified stock was titered by infecting monolayers of BSC-40 cells plated on 6 cm plates with serial dilutions of the stock. Expected titers were 10^{9-10} pfu/ml.

Expression Studies Of DKLIP-1

Co-expression studies of dKLIP-1 with pro-β-NGF were performed essentially as described (Bresnahan, Leduc et al. 1990). Briefly, parallel wells of BSC-40 cells were infected with partially purified preps of VV:NGF (m.o.i.= 2) and either VV:WT or VV:dKLIP-1 (m.o.i.= 5, total m.o.i.= 7). At 19 hr postinfection, the cells were starved for

methionine for 30 min. and then pulsed with 100 µCi [^{35}S]-Met,Cys (Express Label, NEN-Du Pont). After 3 hours, the medium was harvested and secreted products immunoprecipitated by overnight incubation with a rabbit anti-mouse-β-NGF serum (gift of S. Ojeda), following which immune complexes were collected using Protein-A Sepharose beads (Zymed Labs). The beads were washed twice with 1% NP40/1M NaCl and twice in RIPA buffer (Harlow and Lane 1988). Immune complexes were analyzed by SDS-PAGE in 12.5% gels. Fixed gels were then impregnated with fluor and exposed to X-ray film. Gels containing products to be subjected to Edman degradation were left unfixed.

Protein Transfer and Edman Degradation

Unstained SDS-PAGE gels were equilibrated in transfer buffer (0.192 mM glycine, 0.025 M Tris base, 0.15% MeOH) at RT for 30 min. A sheet of polyvinylidene difluoride (PVDF) membrane (Immobilon™, Millipore Corp.) was cut to match the size of the gel and wetted in MeOH, then equilibrated in transfer buffer for 15 min. A sandwich was constructed as suggested by the manufacturer of the electro-transfer apparatus (Hoefer Scientific) by placing the PVDF membrane and the gel between blotting paper. The transfer was done at 4°C overnight with 36 Volts at a current of 1 Ampere. After transfer the membrane was briefly rinsed in dH_2O and a fluorogen (EnHance, Nen-Dupont) was sprayed over the membrane surface. An autoradiograph was made by exposing the membrane to XO-mat film at -70°C. Membrane segments carrying bands for Edman degradation were excised from the rest of the membrane and placed

in an acid washed borosilicate glass test tube. To the dry membrane was added 80 µl dH_2O and 80 µl phenylisothiocyanate (PITC) in pyridine and stoppered under nitrogen. After 45 min incubation at 45°C, the liquid was lyophilized under vacuum in a Speed Vac™ (Savant Instruments) centrifuge until the membrane was completely dry. Then 80 µl anhydrous trifluoroacetic acid (TFA) was added, the tube stoppered under nitrogen and incubated at 45°C for 15 min. After drying the membrane in the Speed Vac™ centrifuge, 0.2 ml dH_2O was added to the membrane and the aqueous phase extracted 3 times with 0.3 ml ethyl acetate. The organic phases were pooled in a scintilation vial and mixed with 1 ml of Ecolume (NEN-DuPont) scintilation fluid. The aqueous phase was completely dried in the Speed Vac™ centrifuge prior to addition of dH_2O and PITC for the start of the next cycle of degradation chemistry.

In situ Hybridization

Sense and antisense RNA probes were generated by using the SP6 or T7 promoters of pGEM7Zf+:dKLIP-1 using α-[^{35}S]-thio-rUTP. Probes were hybridized to 8 mm tissue sections that had been mounted on microscope slides and fixed with 4% formaldehyde, then treated in 0.2 M HCl for 20 min., 2x SSC for 30 min. at 60°C, proteinase K at 37°C for 15 min., 4% formaldehyde for 20 min. and then acetylated and dried (Hafen and Levine 1986). After overnight hybridization at 50°C, the sections were rinsed briefly in 4x SSC and then incubated for 15 min. at 60°C in 50% formamide/0.3 M NaCl/0.03 M Tris, pH 7.5/1 mM EDTA/10 mM DTT. The slides were then washed in 4 liters of 2x SSC at 25°C for 30 min., 0.1 x SSC

(55°C, 10 min.) and 0.1x SSC (25°C, 10 min.). After drying the slides were coated with photographic emulsion (Kodak NTB2).

Embryos were collected and prepared by previously described methods (Mitchison and Sedat 1983). Sense and antisense transcripts were synthesized in the presence of digoxigenin-11-rUTP (Boehringer Mannheim). Following RNase-free DNAase treatment, the transcripts were subjected to limited hydrolysis in 0.1 M carbonate buffer, pH 10.2 (Angerer and Angerer 1981) yielding fragments of 100-125 nucleotides as determined by gel electrophoresis. Whole mount embryo hybridization and detection were done as previously described (Tautz and Pfeifle 1989). Embryonic stages are as described by Campos-Ortega and Hartenstein (1985).

Cytological Localization

In situ hybridization to polytene chromosomes was done by using a biotinylated probe essentially as described (Pardue 1986) except that the probe was made with biotin-21-dUTP (Clontech, Inc., Palo Alto, CA) and detected with avidin conjugated horseradish peroxidase (Vector Labs, Burlingame, CA).

Isolation of Drosophila Genomic DNA

Three grams of frozen flies (stored at -70°C) were ground to a fine powder under liquid nitrogen with a mortar and pestle and transferred to a small beaker. Twenty-five µl of extraction buffer (10 mM Tris HCl, 100mM EDTA and 0.5% sodium dodecyl sulfate (SDS)) were

slowly added, adjusted to 20 μg/ml pancreatic RNase and incubated at 37°C for 60 min.. In a 50 ml conical tube the RNase treated solution was adjusted to 100 μg/ml proteinase K and incubated at 50°C for 3 hrs. at which time the mixture was extracted 6x with 1 volume of TE pH 8 equilibrated phenol by slowly rotating the tube on a rotorack for 10 min. The mixture was then spun at 2,000 rpm for 3 min. for the first 4 extractions and then in an HB4 rotor at 5000 x gravity (g), 5 min. for the final 2 and the aqueous phase removed for one further extraction with one volume of $CHCl_3$. The aqueous phase was transferred to dialysis tubing and dialyzed against 4 liters of TE (10mM Tris HCl, pH 8, 1mM EDTA). The dialysis was carried out at room temperature with 3 changes made every 2 hrs. The final change was left overnight. DNA in the dialysate was concentrated by ethanol precipitation to 1 mg/ml for restriction analysis. Final yield of genomic DNA was 5 mg.

Southern Blot Analysis of Genomic DNA

Twenty five micrograms of DNA was digested with one of four different type II restriction enzymes (Eco RI, Bam HI, Pst I, Bgl II) (BRL). After 2.5 hours of incubation under the recommended conditions for digestion, the reactions were phenol/chloroform extracted. The volume of the aqueous phase was reduced to approximately 10 μl by butanol extraction and loaded into a 1% agarose, 1x TBE gel (Tris, borate, EDTA; (Maniatis, Fritsch et al. 1982)). After staining, the DNA was depurinated by treating the entire gel in 0.25 N HCl for 8 min at room temperature, rinsing in H_2O, denaturing (0.5 N NaOH, 1.5 M NaCl;

2x15 min.), neutralizing (3 M NaCl, 0.5 M Tris-HCl pH 7.5; 2x15 min.), and transfering to a nylon membrane (Hybond N, Amersham, Inc.) by capillary transfer using 20x SSC overnight. The next day the membrane was removed and the DNA crosslinked to it by exposure to UV radiation (1200 joules/cm^2). The gel was stained with EtBr to check for untransferred DNA and discarded. The membrane was incubated in prehybridization mix (1% SDS, 1 M NaCl, 10 % dextran sulfate, 100 µg/ml sonicated, denatured salmon genomic DNA) at 37°C for 6 hrs. One hundred nanograms of dKLIP-1 cDNA was labeled with [^{32}P] by the random primer method (Feinberg and Vogelstein 1983). The membrane was hybridized by adding the dKLIP-1 probe to a final specific activity of 2x10^6 cherenkov cpm/ml to the prehybridization mix and incubated for 36 hrs. at 37°C. The filter was rinsed in 4x SSC at room temperature for 15 min. and in 0.1x SSC, 0.1% SDS at 65°C for 2 hrs. Autoradiography was done by exposing the membrane to X0-mat film using an image intensifying screen at -70°C for up to five days.

Method For Analysis of DKLIP Transcripts By Primer Extension

The method used for this analysis follows closely one that is published (Kingston 1988). A 32 base single stranded DNA oligomer was designed using the sequence for the reverse complement at nucleotides 257-281 (including a 7 base extension for an Eco R1 cloning site at the 5' end) of the dKLIP-1 cDNA. Fifty nanograms of oligomer were 5' end labeled using 100 µCi [γ-^{32}P] ATP (NEN-Du Pont de Nemours, Wilmington, DE) and T4 polynucleotide kinase. Fifty thousand cpm of purified,

labeled oligomer were ethanol precipitated along with 5 mg of polyA+ selected RNA isolated from either Drosophila heads or 0-2 hr. postdeposition embryos. After suspension in hybridization buffer (1M NaCl, 0.16 M HEPES pH 7.5, 0.33 mM EDTA pH 8) the mixtures were denatured for 5 min. at 85°C and then immediately placed at 30°C overnight for primer annealing. The annealing reactions were ethanol precipitated and resuspended in 25 ml of 1x reverse transcriptase buffer, 0.56 mM dNTPs, 10 mM DTT, 50 units pancreatic RNase inhibitor, and 200 units murine Moloney leukemia virus reverse transcriptase minus RNase H activity (RT) (Superscript, BRL). The extension reactions were incubated at 42°C for 60 min. and after 100 units of RT were added, incubated further for 30 min.. The reactions were RNase treated, phenol/chloroform extracted and ethanol precipitated. After lyophilization, the pellets were resuspended in 1 ml each TE and formamide/dyes loading mix for sequencing gels. After denaturing at 100°C for 2 min. each sample was loaded in a single lane of a 4% denaturing wedge type sequencing gel along with a known sequence ladder as a molecular weight standard. The xylene cyanol dye was run to the bottom of the gel at which time it was dried and used to expose XO-mat film.

Isolation and Analysis of Genomic DNA Clones

Fifty thousand lambda phage clones (Charon 4 with ~20 kb partial Hae III inserts with Eco RI ends) containing Drosophila melanogaster genomic DNA were screened on LE392 host bacteria by standardized

procedures using nylon membrane lifts probed with full length [^{32}P] dKLIP-1 cDNA. Filters were hybridized in 50% formamide, 1 M NaCl, 1% SDS, 10% dextran sulfate and 100 μg/ml sheared, denatured salmon sperm DNA at 37°C. The probe was washed from the membranes in 2x SSC at room temp. and 0.1x SSC, 0.1% SDS at 65°C for 1-2 hrs. Exposures were made with XO-mat film. After several rounds of purification, the phage isolates were plated to check purity. Liquid lysates for DNA isolation were made by adsorbing host bacteria with phage at an moi of $2x10^{-3}$ pfu per cell in the presence of divalent metal ions and maltose and used to inoculate 30 ml of media. After 6 hours the cultures had uniformly lysed and the phage DNA prepared by published procedures (Maniatis, Fritsch et al. 1982). Yields were 15-25 μg of phage DNA per prep.

One quarter of each prep was digested with Eco RI endonuclease, electrophoresed in agarose gels and transferred to nylon membranes for probing with the labeled dKLIP-1 cDNA. Autoradiographs were made from the washed filter membranes.

Antibody Production

The method for production of the immunogen was chosen based upon the work of others in the field (Wise, Barr et al. 1990). This method consists of constructing a fusion protein between the active site region of the endoprotease and that of an efficiently expressed, inducible E. coli protein trp E (Dieckmann and Tzagoloff 1985). Two fusion protein encoding plasmids were constructed (figure 19B), (i) pathAct encoding 182 amino acids between the His and Ser residues of the active site

region fused to the carboxyl terminus of trp E and (ii) pathPro encoding 104 amino acids of the presumptive pro region fused to trp E.

All growth of plate and liquid cultures was done in the presence of 20 μg/ml L-tryptophan to repress the trp promoter. Three ml of Luria broth with or without 60 μg/ml carbenicillin were inoculated with one of two E. coli isolates containing pathAct (#3,12) or pathPro (#1,3) constructs or isolates harboring the path vector alone or no exogenous plasmid and grown overnight under tryptophan repression. The next day, 0.5 ml of each overnight culture was used to inoculate 25 mls of M9 medium containing 1% casamino acids and 1.3% tryptone with or without carbenicillin in a 125 ml flask . The cultures were grown at 37°C until the optical density (OD) $l{=}550nm$ reached 0.8-1. The cultures were induced to synthesize the fusion proteins by derepression of the trp operon with a 1:1000 dilution of, a 10 μg/ml indoleacrylic acid (Aldrich Chemical Co.) stock in 100% ethanol (Doolittle and Yanofsky 1968). Cells were harvested at 2, 8 and 20 hrs post induction and stored at -20°C until analyzed. Soluble and insoluble fractions of the 8 hr. induced pellets were prepared by a published method (Kleid, Yansura et al. 1981) with modifications. Briefly, the frozen pellets were resuspended in 0.1 ml of TEN buffer (50 mM Tris HCl (pH 7.5), 0.5mM EDTA, 0.3 M NaCl). Ten μl of 10 mg/ml egg white lysozyme in TEN buffer was added to the cell suspension and set on ice for 15 min. at which time 2 ml of 10% NP-40 was added and mixed. The lysates were sonicated on ice with a probe sonicator using a microtip probe and sonic output of 4 for 5 sec. The

lysates became noticeably less viscous after this treatment. They were spun to separate insoluble from soluble cellular fractions for 5 min. in a microfuge at 15,000 rpm and the supernatent (soluble fraction) transferred to a fresh tube for SDS-PAGE analysis. The pellet (insoluble fraction) was rinsed twice with 100 ml of TEN buffer by repeated trituration and pelleting. The final pellet was resuspended in 50 µl of SDS-PAGE loading buffer. Ten ml of each fraction were subjected to discontinous mini SDS-PAGE using a 5% stacking gel over a 10% separating gel. Each fusion protein was found predominantly in the insoluble fraction of the preparation (figure 19C). Relative molecular sizes were determined to be as expected (pathAct = 59 kD and pathPro = 43 kD).

The fusion protein pathAct was chosen as immunogen for the rabbits. Large scale insoluble fraction preps (150-300 ml) of pathACT3 were prepared as described above except that the insoluble pellet was resuspended in 0.8 ml phosphate buffered saline (PBS), 0.5% SDS. After addition of 100 ml of 5x loading buffer, heating at 60°C for 45 min. and 100°C for 15 min. the sample was loaded on a 3mm thick preparative discontinous SDS gel and the predominant 60 kD band excised from the Coomassie Brilliant Blue stained gel. The gel slice was equilibrated in 0.5x SDS running buffer for 30 min. and electroeluted in dialysis tubing in 0.5x running buffer at 200 volts for 3 hrs. and 55 volts for 7 hrs. Fifty-75% of the stained material was recovered in the eluate and was determined by gel analysis to have yielded ~1.5 mg of intact pathAct

fusion. Four volumes of ice cold acetone were added and placed on ice for 30 min.. The precipitate was pelleted at 10,000 x gravity. PathAct was found to be soluble in acetone so the cleared solution was lyophilized overnight. The powder was resuspended in 0.9% NaCl (normal saline) to 1 mg/ml.

Two 2.5 kg New Zealand white rabbits (arbitrarily given #88 and 89) were bled for preimmune serum. Five hundred micrograms (500 ml) of partially purified pathAct3 protein in normal saline was mixed with 2.75 ml of complete Freund adjuvant (BRL) and 1 ml of normal saline in two 5 ml syringes connected via a 3-way luer lock fitting. Each rabbit was injected at 10 intradermal sites along the back with 0.2 ml (25 µg fusion protein/site) of the immunogen mixture. The rabbits were boosted 4 times at 5 week intervals with 250 µg of fusion protein in incomplete Freund adjuvant. The boost was divided in half for intramuscular thigh injections (125 µg/site). Bleeds of 5-20 mls were taken by marginal ear vein 7-10 days after being boosted. After clotting for 1 hr. at room temp., the clots were "rung" with a Pasteur pipet and the clot allowed to shrink at 4°C overnight. The clot and serum were transferred to sterile Corex tubes and spun at 11,500 rpm for 10 min.. The serum was transferred and respun prior to aliquoting and quick freezing in liquid N_2 (LN_2). Serum aliquots were stored at -70°C until use and once thawed stored at 4°C.

Immunoprecipitation Of Heterologously Expressed Protein

Co-expression studies of dKLIP-1 with pro-β-NGF were performed as described above. Briefly, parallel wells of BSC-40 cells were infected with partially purified preps of VV:NGF (m.o.i.= 2) and either VV:wt or VV:dKLIP-1 (m.o.i.= 5, total m.o.i.= 7). At 19 hr postinfection, the cells were starved for methionine for 30 min. and then pulsed with 100 µCi [^{35}S]-Met,Cys (Express Label, NEN-Du Pont). After 3 hrs. incubation, the cells were rinsed with Tris-buffered saline (TBS) and scraped from the plates in 1 ml PBS, 1% Triton X-100. The lysed cells were quick frozen in LN_2 and thawed at 37°C three successive times and the debris removed by centrifugation. The supernatent (cell extract) was quick frozen and stored at -70°C until use.

One hundred microliters (1/10th) of the labeled cell extract from VV:wt/VV:NGF or VV:dKLIP-1/VV:NGF infections were incubated overnight with 0.5 µl of bleed #1 from rabbit #88 or #89 at 4°C overnight. The following day 15 µl of Sepharose 4CL beads coupled to protein A were added and the tubes set on a rotorack for 3 hours at 4°C. The beads were then pelleted and rinsed twice with a solution of 1M NaCl, 1% NP-40 and twice with RIPA buffer (1% deoxycholate, 1% Triton X-100, 1% SDS, 0.15 M NaCl, 50 mM Tris-HCl, pH 8) plus or minus 1M urea. The immune complexes were disrupted by addition of SDS-PAGE loading buffer and incubation at 100°C for 3 min. Each sample was loaded per well of a mini discontinous SDS-PAGE gel (3% stacking/8% separating). After the run the gel was stained with Coomassie brilliant blue, destained in 20%

2-propanol, 7% acetic acid, enhanced for fluorography and dried. Exposures of the dried gels were made at -70°C.

Western Blots for the Determination of Antibody Titer and Specificity

Serial dilutions of serum from each rabbit were titrated by western blot against dKLIP-1 expressed in mammalian cells. Ten microliters (1/100th of the total of 8.8 x 10^6 cells) of unlabeled cell extract (prepared as described above) from VV:dKLIP-1 or VV:wt infected BSC-40 cells was loaded per lane onto a mini discontinuous SDS-PAGE gel. The separated proteins were transferred to nitrocellulose (0.2μm, Schleicher and Schuell Co.) in carbonate buffer (10 mM $NaHCO_3$, 3mM Na_2CO_3, 10% MeOH, pH 9.9) and visualized with Ponceau S (Sigma Chemical Co.). The membrane was divided into strips and blocked in 1% nonfat dry milk (Carnation Inc.) dissolved in Tris-buffered saline with 0.1% Tween-20, 0.01% NaN_3 (TBST) for 30 min. at room temp. and transferred to serial dilutions of serum from bleed #3 taken from each rabbit in TBST. The final dilutions were 1:500, 1:1000, 1:2000, 1:4000 and 1:8000. After overnight incubation, the strips were rinsed 3 x 5 min. in TBST and then incubated in goat anti-rabbit IgG conjugated to alkaline phosphatase (Zymed) at 1:3000 in TBST for 1 hr.. After 2 x 5 min. TBST rinses and equilibration in alkaline phosphatase buffer (0.1 M Tris-HCl pH 9.5, 0.1M NaCl, 5 mM $MgCl_2$) the membrane strips were reacted with NBT and BCIP to generate a reaction product. The reaction was terminated by rinsing the membrane in 50 mM EDTA and dH_2O.

For antibody-blocking experiments, insoluble fractions from 5 ml cultures of wild type E. coli cells or those expressing the dKLIP-1 active site-TrpE fusion protein (pathAct) were prepared as described above and resuspended in 50 μl of PBS, 0.05% SDS (PBS-S). One μl of serum from the preimmune bleed or boost #3 from rabbits 88 and 89 was mixed with 9 μl of pathAct or wild type insoluble E. coli preparations or PBS-S and incubated at room temperature (rt) for 30 min. Cell extracts from mock, VV:wt or VV: dKLIP-1 infected BSC-40 cells prepared and fractionated by SDS-PAGE. The proteins were then transferred to nitrocellulose and blocked prior to antisera incubation. The treated antisera were diluted 200 fold (1:2000 final dilution) by addition of 1.99 ml TBST and allowed to incubate with pre-blocked membrane strips overnight at rt. Bound antibodies were detected using an alkaline-phosphate conjugated goat anti-rabbit IgG (Zymed) diluted to 1:5000.

Immunocytochemistry on Fixed Drosophila Embryos and Larval CNS Preparations

The general method for immunocytochemical staining of Drosophila embryos is published (Weischaus and Nusslein-Volhard 1986). Briefly, staged embryos were collected, dechorionated in 50% bleach and fixed in 4% paraformaldehyde for 15 min. The fixed embryos were collected, rinsed in PBT (phosphate buffered saline, 0.1% bovine serum albumin, 0.3% Triton X-100) and blocked in PBS, 10% horse serum for 4 hours, 4°C. All incubation steps were done using a rotorack to keep the embryos in suspension. Primary antisera was added in PBT and incubated overnight.

The embryos were washed in PBT with several changes over 1 hour and an affinity purified, biotinylated goat anti-rabbit secondary antibody (1:400) added in PBT for an overnight incubation. After washing the embryos, they were incubated in streptavidin conjugated horseradish peroxidase (HRP) (1:1000) for 30 min. HRP activity was determined by incubation in 1 mg/ml diaminobenzidine (DAB). The color reaction was terminated by rinsing in PBT several times. After the embryos had been dehydrated through a series of ethanol solutions (70, 95, 100%) they were cleared in cedarwood oil overnight. Stained embryos were mounted on glass slides and sealed under cover slips for microscopy.

Central nervous systems (brain, ring gland and ventral ganglion) were dissected from late 3rd instar larvae because at this stage the brain is relatively large and the ring gland is easily located. Dissections were done in Ringer's solution by pulling at the mouthparts and the tail simultaneously. Tissues were fixed for 1 hour, rinsed several times in PBT before blocking and incubated in primary and secondary antisera dilutions as described above. Detection was carried out by binding of an FITC-avidin conjugate (10 µg/ml) overnight and extensive washes in PBT. Tissues were aligned on glass coverslips between strips of double sided tape and two drops of antiquench solution (0.05 M CO_3, pH 8.5, 5 mg/ml p-phenylenediamine [an antioxidant], 50 % glycerol) added. A coverslip was overlaid, the intervening space backfilled with antiquench solution and sealed with nail polish.

Construction of the P-element vectors

The plasmid vector, pHSREM1 (Knipple and Marsella-Herrick 1988), was used for the initial subcloning of the dKLIP-1 cDNA. The cDNA was ligated in either the antisense (-) or sense (+) orientation between the heat shock protein promoter (hsp70) and terminator sequences of the vector. The promoter, -/+ oriented cDNA and terminator were excised as a single fragment and ligated into pW8 (Klemenz, Weber et al. 1987). The plasmid vector pW8 contains the white gene fused to the hsp70 promoter, rescuing the white^{-} phenotype (see Results). This construct pW8:dKLIP-1 (+) or (-) was grown up and supercoiled plasmid DNA equilibrium gradient purified. Cesium salt was removed and the plasmid DNA resuspended with injection buffer (5mM KCl, 0.1M Na_2PO_4, pH 6.8) after ethanol precipitation.

Embryo Injections

The methodology was described in detail (Pirrotta 1988). Drosophila strain (w; P [ry^{+}(D2-3)] (99B) was obtained from the Drosophila stock center (Robertson, Preston et al. 1988). This strain carries a defective P-element integrant which supplies P-factor (transposase) in *trans*. Bottles were set up with ~25-50 freshly eclosed male and female flies. All incubations were done at 25°C unless noted otherwise. Synchronized embryos were collected for successive 20 min. periods after an initial clearing period of 1-2 hours. After dechorionation and dessication, the embryos were injected with plasmid DNA at a concentration of 500 µg/ml. Injected embryos were transferred to apple juice, agar plates and incubated at 18°C until larvae hatched.

Male sterility is reduced if injected animals are allowed to recover at the reduced temperature. At hatching, 5-6 larvae were transferred to individual vials with food and upon pupation transferred individually into vials. When adults emerged they were mated with a w^- strain lacking the transposase activity thereby stabilizing any integrated P-elements in the f1 genome. The f1 generation was screened for red eye color.

Results

Cloning and Characterization of dKLIP cDNAs

A strategy for the identification of cDNAs encoding Kex2p-like endoproteases from Drosophila was developed by examining an alignment of the amino acid sequences of human furin (hfurin) (van den Ouweland, van Duijnhoven et al. 1990), human PC2 (hPC2) (Smeekens and Steiner 1990), and yeast Kex2p (Mizuno, Nakamura et al. 1988). Significant identity exists in the amino acid sequences that surround the residues forming the catalytic triad (Asp, His and Ser) with the greatest sequence conservation around the histidine and serine residues (figure 4). Based on this conservation and on Drosophila codon bias data (Wada, Aota et al. 1990), two degenerate oligonucleotide primer pools were designed for use in the polymerase chain reaction (PCR) (Saiki, Gelfand et al. 1988). The first primer pool was a sense 24mer that encoded eight amino acids beginning with the active site histidine residue. The second primer pool was an antisense 32mer corresponding to the coding region for the 11 amino acids surrounding the active site serine residue. A subset of these oligonucleotide primers was expected to anneal to Drosophila Kex2p-like cDNA sequences, producing an amplified product of ~542 bp. Indeed, PCR experiments using template DNA prepared from a Drosophila adult head cDNA library resulted in the isolation of a ~550 bp product. DNA sequence analysis demonstrated the presence of an open reading frame

(ORF) that shared significant homology to the predicted region of the Kex2p-like endoproteases.

```
                   *                *                    *
Kex2p ...171AAIVDDGLDYEN....212YHGTRCAGEI.....382GGTSAAAPLAAGV
hfurin...149VSILDDGIEKNH....193RHGTRCAGEV.....365TGTSASAPLAAGI
hPC2 ....138IGIMDDGIDYLH....207SHGTRCAGEV.....381SGTSAAAPEAAGV
Consensus      --I-DDG-----      -HGTRCAGE-        -GTSA-AP-AAG-
```

Figure 4. Alignment of the amino acid sequences surrounding the catalytic triad residues in Kex2p, hfurin and hPC2. The coding sequences surrounding the His and Ser residues showed greatest sequence identity and therefore, were used for the design of PCR primer pools. The amplified ~542 bp product predicted from PCR experiments was determined by adding the lengths of the primers (24 and 32 bases) to the sequence in Kex2p that separated their 3' ends (486 bases). Computer based alignment of polypeptide sequences for Kex2p (Mizuno, Nakamura et al. 1988), hfurin (van den Ouweland, van Duijnhoven et al. 1990) and hPC2 (Smeekens and Steiner 1990) was done using the Bestfit program (Devereux, Haeberli et al. 1984). Nonidentical residues are shown by dashes in the consensus (bottom) line. The asterisks denote aspartate, histidine and serine residues of the catalytic triad.

Seven independent phage clones were isolated to homogeneity from the adult head library using the labeled PCR product as a probe. Two of these, clones 3 and 1L contained cDNA inserts of 3.5 and 2.8 kb, respectively (figure 5A). Clone 1L contained an ORF with a methionine residue at nucleotide 232 and extended to the 3' end of the sequence. The nucleotide sequence surrounding this putative initiation codon (figure 5B, nuc. 225-235) was identical to the Drosophila consensus translation initiation site (ANN(C/A)A(A/C)(A/C)ATGN; (Cavener 1987)). In addition, this methionine residue was preceded by several in frame termination codons. Clone 3 initiated at nucleotide 240 of clone 1L and continued to a termination codon at position 3535. The protein product predicted from the two overlapping cDNA sequences was 1101 amino acids in length. A composite cDNA (hereafter referred to as the dKLIP-1 cDNA) of 3.78 kilobase pairs containing the complete ORF for dKLIP-1 was created (figure 5A).

Figure 5. Nucleotide sequence of the dKLIP-1 cDNA and deduced protein sequence.

A. Schematic diagram showing two overlapping cDNA clones, 1L and 3, isolated from an adult Drosophila head specific cDNA library. The approximate length of sequence determined from subcloned fragments is represented by arrows. The lower diagram of the panel shows the open reading frame (thick line) containing the subtilisin-like catalytic domain (stippled box) in relation to the composite cDNA. The composite dKLIP-1 cDNA was generated by ligating 784 bp from clone 1L to 2997 bp from clone 3 using the BamHI (B) restriction site.

B. The nucleotide sequence of the dKLIP-1 cDNA. An opa repetitive element (nuc. 257-347) (Wharton et al., 1985) is underlined between the nucleotide and amino acid sequences. Ten potential N-linked glycosylation sites (Asn-X-Ser/Thr) are indicated by bold underlines. Putative membrane spanning domains are boxed. Triangles represent potential paired basic residue autoproteolytic maturation sites; (*) Asp, His and Ser residues that form the catalytic triad, as well as the Asn residue that stabilizes the transition state complex with the substrate (Kraut, 1977); RGD sequence important for cellular adhesion are double underlined; (...) AUUUA motif important for conferring RNA instability. The polyadenylation addition site (nuc. 3755-3760) is underlined. The cysteine-rich region that shares conserved cysteine residues with hfurin is between amino acids 847-982 (see figure 9).

A

λ1L

λ3

B

1 kb

B

```
GACAACAACAAATTTAAATTGAGATTTTCCATATTATCAAAAGTTACAGGCGCCTATTTTGCAGTTATCAGTCTGAAGTCGGTGTGAAGTGCTATCCTGGACTTTTCGTTTTAGTCACCC  120
CTTGGCCTCATCTAACATCCACTTAATTGAAATAGAACATTTGCATCTCAGCCTCACAACCAATATCCCGAGCAACGTGCACAAAAACTATTGGAAAAATCCTAACCCACAATGAAAAAC  240
                                                                                                                 MetLysAsn    3

GACGTCGTGCGATGGAGCAGGCAGCCAACTAGCAACACCACTAACAGCAGCAGCAGCAGTAGAAGCGATAGCAACAGCACTCATAAGCACAGAAGCAAAAGCAACAAATTAAATGCTCGA  360
AspValValArgTrpSerArgGlnProThrSerAsnThrThrAsnSerSerSerSerSerArgSerAspSerAsnSerThrHisLysHisArgSerLysSerAsnLysLeuAsnAlaArg   43

CAATTAGGGTCAAATGCTGCCAGAAGTTGTCAGCAAAGATCCTCTGTGGCAACAACACTAGAAGATGAACAACAAACAATCATTGAATGTGATATAGGAAATTTCAATTTCGATTGCAAT  480
GlnLeuGlySerAsnAlaAlaArgSerCysGlnGlnArgSerSerValAlaThrThrLeuGluAspGluGlnGlnThrIleIleGluCysAspIleGlyAsnPheAsnPheAspCysAsn   83

TTATTCAAAACTAGTTTTTTAACGCAACACAAACAGAAACGTAGTGGCAAGAGCAGCAGCAAAAGTAAGAGCAACAGAAGTAGACCCCTAGCGAAAACGAAAGCGGTGTTTCTATTAGCT  600
LeuPheLysThrSerPheLeuThrGlnHisLysGlnLysArgSerGlyLysSerSerSerLysSerLysSerAsnArgSerArgProLeuAlaLysThrLysAlaValPheLeuLeuAla  123
                                          Δ
CTGCAATTTAGTGCCGTAGTTTTTTTATGTAATATTAATGTCGGTTTCGTGGCCGGAAGTGTGGCAACTGCGGCATCATCGGCAGGCGGCTCATCGCCGGCAGCTCCATCATCTGCGCCC  720
LeuGlnPheSerAlaValValPheLeuCysAsnIleAsnValGlyPheValAlaGlySerValAlaThrAlaAlaSerSerAlaGlyGlySerSerProAlaAlaProSerSerAlaPro  163

TCATCCCCGCCCACAGTTGCTGTACCACCGCCGCCGCCACCTTCGTCGGCACTCAAAGTGGATCCAAATGGTCAGTCACCAGTGCTGCCGCCCTACGTTCTCGATTATGAGACGGGGGGC  840
SerSerProProThrValAlaValProProProProProProSerSerAlaLeuLysValAspProAsnGlyGlnSerProValLeuProProTyrValLeuAspTyrGluThrGlyGly  203

AAGGCCAAGCTAACGCCAAACAATGGCAAGTTCGGCCAATCGGGCAGTTCCGGGAGCAATAACAACCACATCGTCGGACACTATACCCACACCTGGGCGGTGCACATACCAAACGGCGAT  960
LysAlaLysLeuThrProAsnAsnGlyLysPheGlyGlnSerGlySerSerGlySerAsnAsnAsnHisIleValGlyHisTyrThrHisThrTrpAlaValHisIleProAsnGlyAsp  243

AATGGCATGGCCGATGCGGTTGCCAAGGATCACGGATTCGTCAATTTGGGCAAGATCTTCGATGATCACTACCACTTCGCACATCACAAGGTCTCGAAGCGGTCGCTCTCCCCCGCCACG 1080
AsnGlyMetAlaAspAlaValAlaLysAspHisGlyPheValAsnLeuGlyLysIlePheAspAspHisTyrHisPheAlaHisHisLysValSerLysArgSerLeuSerProAlaThr  283
                                                                                                       Δ
CATCACCAAACTCGCCTGGATGACGACGATCGCGTCCACTGGGCGAAGCAGCAGCGGGCCAAGTCGCGATCCAAACGGGACTTTATCCGCATGCGACCCTCACGGACCTCCTCGCGAGCC 1200
HisHisGlnThrArgLeuAspAspAspAspArgValHisTrpAlaLysGlnGlnArgAlaLysSerArgSerLysArgAspPheIleArgMetArgProSerArgThrSerSerArgAla  323
                                                                                  Δ
ATGTCGATGGTGGACGCCATGTCCTTTAACGACTCCAAGTGGCCGCAGATGTGGTATCTGAATCGTGGTGGTGGCCTGGACATGAATGTGATACCCGCCTGGAAGATGGGCATAACCGGC 1320
MetSerMetValAspAlaMetSerPheAsnAspSerLysTrpProGlnMetTrpTyrLeuAsnArgGlyGlyGlyLeuAspMetAsnValIleProAlaTrpLysMetGlyIleThrGly  363

AAGGGCGTGGTGGTGACAATTCTGGATGATGGCCTGGAATCCGATCATCCGGACATACAGGATAACTACGATCCCAAAGCCTCGTACGATGTGAATAGCCACGACGACGATCCGATGCCG 1440
LysGlyValValValThrIleLeuAspAspGlyLeuGluSerAspHisProAspIleGlnAspAsnTyrAspProLysAlaSerTyrAspValAsnSerHisAspAspAspProMetPro  403
                          *
CATTACGATATGACGGACTCGAACCGCCATGGAACTCGCTGTGCCGGCGAGGTGGCAGCCACCGCCAACAATTCGTTCTGCGCGGTGGGTATTGCCTACGGCGCCAGTGTGGGCGGAGTC 1560
HisTyrAspMetThrAspSerAsnArgHisGlyThrArgCysAlaGlyGluValAlaAlaThrAlaAsnAsnSerPheCysAlaValGlyIleAlaTyrGlyAlaSerValGlyGlyVal  443
                            *
AGGATGCTGGACGGAGACGTCACGGATGCGGTTGAGGCACGGTCGCTGTCGCTGAATCCGCAGCACATTGACATATACAGTGCCTCCTGGGGACCCGATGACGATGGCAAGACGGTGGAC 1680
ArgMetLeuAspGlyAspValThrAspAlaValGluAlaArgSerLeuSerLeuAsnProGlnHisIleAspIleTyrSerAlaSerTrpGlyProAspAspAspGlyLysThrValAsp  483

GGACCCGGCGAACTGGCATCGCGCGCCTTTATCGAGGGCACAACTAAGGGACGCGGCGGCAAGGGCAGCATCTTCATATGGGCATCGGGCAATGGTGGGCGGGAGCAGGATAACTGCAAC 1800
GlyProGlyGluLeuAlaSerArgAlaPheIleGluGlyThrThrLysGlyArgGlyGlyLysGlySerIlePheIleTrpAlaSerGlyAsnGlyGlyArgGluGlnAspAsnCysAsn  523
                                                                                        *
TGCGACGGCTACACGAACTCCATCTGGACGCTGTCCATCTCCAGTGCCACGGAGGAGGGCCATGTGCCCTGGTACTCGGAGAAGTGCAGCTCCACGCTGGCCACCACCTACAGCAGCGGC 1920
CysAspGlyTyrThrAsnSerIleTrpThrLeuSerIleSerSerAlaThrGluGluGlyHisValProTrpTyrSerGluLysCysSerSerThrLeuAlaThrThrTyrSerSerGly  563

GGGCAGGGCGAGAAGCAGGTGGTCACCACGGACCTGCACCACTCGTGCACTGTCTCCCACACGGGCACCTCGGCGTCGGCCCCGCTCGCCGCTGGCATAGCCGCCCTGGTGCTGCAGTCC 2040
GlyGlnGlyGluLysGlnValValThrThrAspLeuHisHisSerCysThrValSerHisThrGlyThrSerAlaSerAlaProLeuAlaAlaGlyIleAlaAlaLeuValLeuGlnSer  603
                                                                              *
AACCAGAATCTCACCTGGCGCGATCTGCAGCACATTGTTGTGCGCACCGCCAAGCCGGCGAACCTTAAGGACCCCAGCTGGTCACGCAATGGGGTGGGGCGGCGGGTGAGCCACTCCTTT 2160
AsnGlnAsnLeuThrTrpArgAspLeuGlnHisIleValValArgThrAlaLysProAlaAsnLeuLysAspProSerTrpSerArgAsnGlyValGlyArgArgValSerHisSerPhe  643

GGCTACGGATTGATGGACGCCGCCGAGATGGTGCGCGTGGCCCGCAACTGGAAGGCGGTGCCGGAGCAGCAGCGGTGCGAGATTAACGCTCCCCATGTCGACAAGGTCATTCCACCTCGT 2280
GlyTyrGlyLeuMetAspAlaAlaGluMetValArgValAlaArgAsnTrpLysAlaValProGluGlnGlnArgCysGluIleAsnAlaProHisValAspLysValIleProProArg  683

ACCCATATCACCCTGCAACTGACGGTTAATCACTGTCGATCGGTCAATTACCTGGAGCACGTCCAGGCCAAGATTACGCTAACGTCGCAGAGACGAGGAGACATTCAGCTCTTTTTGAGG 2400
ThrHisIleThrLeuGlnLeuThrValAsnHisCysArgSerValAsnTyrLeuGluHisValGlnAlaLysIleThrLeuThrSerGlnArgArgGlyAspIleGlnLeuPheLeuArg  723

TCTCCCGCAAACACCAGTGTCACGCTCCTAACGCCTAGGATACATGACAACTCGTTCCGGATTCAATCAATGGCCCTTCATGTCTGTGCACACCTGGGGAGAGTCGCCGCAAGGAAAC 2520
SerProAlaAsnThrSerValThrLeuLeuThrProArgIleHisAspAsnSerSerGlyPheAsnGlnTrpProPheMetSerValHisThrTrpGlyGluSerProGlnGlyAsn  763

TGGCAGCTGGAGATCCACAACGAGGGTCGCTATATGGCACAAATCACACAATGGGATATGATATTCTACGGCACCGAAACGCCCGCCCAACCCGATGACGTGGCCAATCCCAGCCAGTCG 2640
TrpGlnLeuGluIleHisAsnGluGlyArgTyrMetAlaGlnIleThrGlnTrpAspMetIlePheTyrGlyThrGluThrProAlaGlnProAspAspValAlaAsnProSerGlnSer  803

AACCAGTTCAATCTGTACGGCAACGATATGGCCCACAATGACGTCGAGTACGATTCCACCGGCCAGTGGAGGAATATGCAGCAGGTGGGCGAGGTGGGCATGACCCGAGATCACAGCAAC 2760
AsnGlnPheAsnLeuTyrGlyAsnAspMetAlaHisAsnAspValGluTyrAspSerThrGlyGlnTrpArgAsnMetGlnGlnValGlyGluValGlyMetThrArgAspHisSerAsn  843

ACCGCCGCGTGCCTTAAGTGGAGCGATCGCAAGTGCTTAGAGTGCAATGACTCCGCCTATATGTTCGAGGACCAGTGCTACGATGTCTGCCCGGTGCACACATATCCATTGGACAAGTTC 2880
ThrAlaAlaCysLeuLysTrpSerAspArgLysCysLeuGluCysAsnAspSerAlaTyrMetPheGluAspGlnCysTyrAspValCysProValHisThrTyrProLeuAspLysPhe  883

CAAGCAGAGGAGGATGAGCAGGATGACGAGGTGACAAGGGGTCCGGTTAATCCCTACAGCAGCTCCCCCATGGATCATTCCCTGTTGATGTCCAACAGCTTGGATGACAAGCAAGATCCT 3000
GlnAlaGluGluAspGluGlnAspAspGluValThrArgGlyProValAsnProTyrSerSerSerProMetAspHisSerLeuLeuMetSerAsnSerLeuAspAspLysGlnAspPro  923

CTGCAGGCGGAGGATCGACGTCGTCGTAGTTCACTCACCCAACTGGTGGAGGTGCCTTCTCGGGTCTGTGCCGCCTGCGATCGGAGCTGCTTGGAGTGCTACGGTGCCCTGGCCTCCCAG 3120
LeuGlnAlaGluAspArgArgArgArgSerSerLeuThrGlnLeuValGluValProSerArgValCysAlaAlaCysAspArgSerCysLeuGluCysTyrGlyAlaLeuAlaSerGln  963

TGCAGCACCTGCTCCCCGGGCAGCCAACTGCGTAAGATCCTGAACGAGACCTTCTGCTATGCCTATGTGGTGCGTAGCACGGGAATGGCCAGTGTGGTGGACATTTCCAAAATGGACGAT 3240
CysSerThrCysSerProGlySerGlnLeuArgLysIleLeuAsnGluThrPheCysTyrAlaTyrValValArgSerThrGlyMetAlaSerValValAspIleSerLysMetAspAsp 1003

AGAGATACCCAGCAATATATGACTGGCACAATTGTGCTCCTTTTGGTTTCGGTGATTTTCACCCTGATGGGCGTAGCAGTGGCCGGTGGCATCGTGTATCATCGAAGGGCAATGGCTCGA 3360
ArgAspThrGlnGlnTyrMetThrGlyThrIleValLeuLeuLeuValSerValIlePheThrLeuMetGlyValAlaValAlaGlyGlyIleValTyrHisArgArgAlaMetAlaArg 1043

TCCAACGAACTATACTCCCGAGTTTCCCTAGTCCCCGGTGACGAAAGTGACTCTGACGAGGATGAACTGTTTACGGCCCACTTCCCAGCGAGAAAGAGTGGTGTTAATATATATCGCGAT 3480
SerAsnGluLeuTyrSerArgValSerLeuValProGlyAspGluSerAspSerAspGluAspGluLeuPheThrAlaHisPheProAlaArgLysSerGlyValAsnIleTyrArgAsp 1083

GAAGCGCCCAGCGAAAAGATATTCGAGGAAGATGAGATCAGCCACTTGGTACCCTAAATAATCAAAACAAGTTATTAGTTAGGTCATGCATTTAGCCATCCTCATAACCACCAGAAATAT 3600
                                                                                                   *****
GluAlaProSerGluLysIlePheGluGluAspGluIleSerHisLeuValPro                                                                    1101

ATCTATATGTTAAGCAACCGACAAACGCATTTTTGGTCTATATTTTGTTACATTCGACGGCATTACAGCTCCTTTATTTCAAGCCAAAATAGATTAAAGTATTAAATGTAAAAGAAAAAA 3720
AATCAAAACTAATTATTTGTAATATTTATGCCTTAATAAAATGATTTTAACTGAAAAAAAA
                    *****
                                                                                                                          3781
```

The dKLIP-1 cDNA contained several interesting features. Computer analysis was used to compare regions of nucleotide sequence similarity in the dKLIP-1 cDNA with sequences in the Genbank/EMBL database (Devereux, Haeberli et al. 1984). The FASTA program was used to compare the dKLIP-1 cDNA with 42,000 sequences using a word size of six and the twenty best comparisons were examined. These hits had greater than 65% sequence identity within a 64-96 base region of dKLIP-1. Most of the top twenty hits were in the region proximal to the putative initiator methionine (figure 5B, nuc. 258-347) that has a CAX_n trinucleotide repetitive sequence which closely matches similar repetitive elements found in several developmentally regulated mRNAs including the drosophila loci *dorsal* and *mastermind*. This sequence has been termed the opa element and is found in the genomes of many species including human (Wharton, Yedvobnick et al. 1985; Grabowski, Carney et al. 1991). The 245 nucleotide 3' untranslated region (UTR) contained several AUUUA repeats (figure 5B, nuc. 3570 and 3744), a motif which affects mRNA instability (Shaw and Kamen 1986), and a canonical polyadenylation addition sequence [AAUAAA, nuc. 3755, (Proudfoot and Brownlee 1976)] 19 bp from a poly-A sequence.

Of the seven cDNA clones isolated from the adult head specific library, all but one clone had dKLIP-1 sequences. The unique clone (#5) is a partial cDNA of 2182 bp that lacks a start codon and a polyadenylation addition sequence (figure 6). Clone 5 begins at nucleotide 745 of the dKLIP-1 cDNA and shares sequence identity through

nucleotide 2056 with dKLIP-1. From nucleotide 2057 through to the 3' end of clone 5 (figure 10B), the sequence shows complete divergence from dKLIP-1 but is contained in a newly described drosophila furin-like clone (Roebroek, Pauli et al. 1991). The divergent sequence was not found to be represented by sequence identity or conservation in any database surveyed. Interestingly, the point of divergence between clone 5 and dKLIP-1 is precisely at a consensus intron-exon splice junction (AG-GT) found at nucleotides 2055-2058 in the sequence of clone 5 (figure 6, 10B). The open reading frame continues to a termination codon at position 2164. This cDNA has been called dKLIP-2 and may represent an alternate transcript utilizing different exon sequences. Alternatively, this cDNA, which lacks a polyadenylation site and poly-A tract, could represent an RNA-processing intermediate.

Figure 6. Nucleotide sequence of the dKLIP-2 cDNA and deduced protein sequence. Potential autoproteolytic cleavage sites are marked with open triangles; asterisks, active site catalytic triad residues; double underline, integrin receptor binding site consensus sequence; bold nucleotides, consensus splice junction sequence where Dfurin1 has additional 21 nucs. (Roebroek, Pauli et al. 1991) (see text). Italicized lettering represents the potential transmembrane domain as determined by the Kyte and Doolittle hydropathy algorithm (data not shown). Underlined nucleotides are those common to Dfurin1, but not found in dKLIP-1 and underlined amino acids are predicted to be the cytoplasmic domain.

```
   1  CCACCGCCGCCGCCACCTTCGTCGGCACTCAAAGTGGATCCAAATGGTCAGTCACCAGTG
      ProProProProProProSerSerAlaLeuLysValAspProAsnGlyGlnSerProVal

  61  CTGCCGCCCTACGTTCTCGATTATGAGACGGGGGCAAGGCCAAGCTAACGCCAAACAAT
      LeuProProTyrValLeuAspTyrGluThrGlyGlyLysAlaLysLeuThrProAsnAsn

 121  GGCAAGTTCGGCCAATCGGGCAGTTCCGGGAGCAATAACAACCACATCGTCGGACACTAT
      GlyLysPheGlyGlnSerGlySerSerGlySerAsnAsnAsnHisIleValGlyHisTyr

 181  ACCCACACCTGGGCGGTGCACATACCAAACGGCGATAATGGCATGGCCGATGCGGTTGCC
      ThrHisThrTrpAlaValHisIleProAsnGlyAspAsnGlyMetAlaAspAlaValAla

 241  AAGGATCACGGATTCGTCAATTTGGGCAAGATCTTCGATGATCACTACCACTTCGCACAT
      LysAspHisGlyPheValAsnLeuGlyLysIlePheAspAspHisTyrHisPheAlaHis

 301  CACAAGGTCTCGAAGCGGTCGCTCTCCCCCGCCACGCATCACCAAACTCGCCTGGATGAC
      HisLysValSerLysArgSerLeuSerProAlaThrHisHisGlnThrArgLeuAspAsp
                       Δ
 361  GACGATCGCGTCCACTGGGCGAAGCAGCAGCGGGCCAAGTCGCGATCCAAACGGGACTTT
      AspAspArgValHisTrpAlaLysGlnGlnArgAlaLysSerArgSerLysArgAspPhe
                                                           Δ
 421  ATCCGCATGCGACCCTCACGGACCTCCTCGCGAGCCATGTCGATGGTGGACGCCATGTCC
      IleArgMetArgProSerArgThrSerSerArgAlaMetSerMetValAspAlaMetSer

 481  TTTAACGACTCCAAGTGGCCGCAGATGTGGTATCTGAATCGTGGTGGTGGCCTGGACATG
      PheAsnAspSerLysTrpProGlnMetTrpTyrLeuAsnArgGlyGlyGlyLeuAspMet

 541  AATGTGATACCCGCCTGGAAGATGGGCATAACCGGCAAGGGCGTGGTGGTGACAATTCTG
      AsnValIleProAlaTrpLysMetGlyIleThrGlyLysGlyValValValThrIleLeu

 601  GATGATGGCCTGGAATCCGATCATCCGGACATACAGGATAACTACGATCCCAAAGCCTCG
      AspAspGlyLeuGluSerAspHisProAspIleGlnAspAsnTyrAspProLysAlaSer
       *
 661  TACGATGTGAATAGCCACGACGACGATCCGATGCCGCATTACGATATGACGGACTCGAAC
      TyrAspValAsnSerHisAspAspAspProMetProHisTyrAspMetThrAspSerAsn

 721  CGCCATGGAACTCGCTGTGCCGGCGAGGTGGCAGCCACCGCCAACAATTCGTTCTGCGCG
      ArgHisGlyThrArgCysAlaGlyGluValAlaAlaThrAlaAsnAsnSerPheCysAla
          *
 781  GTGGGTATTGCCTACGGCGCCAGTGTGGGCGGAGTCAGGATGCTGGACGGAGACGTCACG
      ValGlyIleAlaTyrGlyAlaSerValGlyGlyValArgMetLeuAspGlyAspValThr

 841  GATGCGGTTGAGGCACGGTCGCTGTCGCTGAATCCGCAGCACATTGACATATACAGTGCC
      AspAlaValGluAlaArgSerLeuSerLeuAsnProGlnHisIleAspIleTyrSerAla

 901  TCCTGGGGACCCGATGACGATGGCAAGACGGTGGACGGACCCGGCGAACTGGCATCGCGC
      SerTrpGlyProAspAspAspGlyLysThrValAspGlyProGlyGluLeuAlaSerArg

 961  GCCTTTATCGAGGGCACAACTAAGGGACGCGGCGGCAAGGGCAGCATCTTCATATGGGCA
      AlaPheIleGluGlyThrThrLysGlyArgGlyGlyLysGlySerIlePheIleTrpAla

1021  TCGGGCAATGGTGGGCGGGAGCAGGATAACTGCAACTGCGACGGCTACACGAACTCCATC
      SerGlyAsnGlyGlyArgGluGlnAspAsnCysAsnCysAspGlyTyrThrAsnSerIle
             *
```

1081 TGGACGCTGTCCATCTCCAGTGCCACGGAGGAGGGCCATGTGCCCTGGTACTCGGAGAAG
TrpThrLeuSerIleSerSerAlaThrGluGluGlyHisValProTrpTyrSerGluLys

1141 TGCAGCTCCACGCTGGCCACCACCTACAGCAGCGGCGGGCAGGGCGAGAAGCAGGTGGTC
CysSerSerThrLeuAlaThrThrTyrSerSerGlyGlyGlnGlyGluLysGlnValVal

1201 ACCACGGACCTGCACCACTCGTGCACTGTCTCCCACACGGGCACCTCGGCGTCGGCCCCG
ThrThrAspLeuHisHisSerCysThrValSerHisThrGlyThrSerAlaSerAlaPro
*

1261 CTCGCCGCTGGCATAGCCGCCCTGGTGCTGCAGTCCAACCAGAATCTCACCTGGCGCGAT
LeuAlaAlaGlyIleAlaAlaLeuValLeuGlnSerAsnGlnAsnLeuThrTrpArgAsp

1321 CTGCAGCACATTGTTGTGCGCACCGCCAAGCCGGCGAACCTTAAGGACCCCAGCTGGTCA
LeuGlnHisIleValValArgThrAlaLysProAlaAsnLeuLysAspProSerTrpSer

1381 CGCAATGGGGTGGGGCGGCGGGTGAGCCACTCCTTTGGCTACGGATTGATGGACGCCGCC
ArgAsnGlyValGlyArgArgValSerHisSerPheGlyTyrGlyLeuMetAspAlaAla

1441 GAGATGGTGCGCGTGGCCCGCAACTGGAAGGCGGTGCCGGAGCAGCAGCGGTGCGAGATT
GluMetValArgValAlaArgAsnTrpLysAlaValProGluGlnGlnArgCysGluIle

1501 AACGCTCCCCATGTCGACAAGGTCATTCCACCTCGTACCCATATCACCCTGCAACTGACG
AsnAlaProHisValAspLysValIleProProArgThrHisIleThrLeuGlnLeuThr

1561 GTTAATCACTGTCGATCGGTCAATTACCTGGAGCACGTCCAGGCCAAGATTACGCTAACG
ValAsnHisCysArgSerValAsnTyrLeuGluHisValGlnAlaLysIleThrLeuThr

1621 TCGCAGAGACGAGGAGACATTCAGCTCTTTTTGAGGTCTCCCGCAAACACCAGTGTCACG
SerGlnArgArgGlyAspIleGlnLeuPheLeuArgSerProAlaAsnThrSerValThr

1681 CTCCTAACGCCTAGGATACATGACAACTCTCGTTCCGGATTCAATCAATGGCCCTTCATG
LeuLeuThrProArgIleHisAspAsnSerArgSerGlyPheAsnGlnTrpProPheMet

1741 TCTGTGCACACCTGGGGAGAGTCGCCGCAAGGAAACTGGCAGCTGGAGATCCACAACGAG
SerValHisThrTrpGlyGluSerProGlnGlyAsnTrpGlnLeuGluIleHisAsnGlu

1801 GGTCGCTATATGGCACAAATCACACAATGGGATATGATATTCTACGGCACCGAAACGCCC
GlyArgTyrMetAlaGlnIleThrGlnTrpAspMetIlePheTyrGlyThrGluThrPro

1861 GCCCAACCCGATGACGTGGCCAATCCCAGCCAGTCGAACCAGTTCAATCTGTACGGCAAC
AlaGlnProAspAspValAlaAsnProSerGlnSerAsnGlnPheAsnLeuTyrGlyAsn

1921 GATATGGCCCACAATGACGTCGAGTACGATTCCACCGGCCAGTGGAGGAATATGCAGC**AG**
AspMetAlaHisAsnAspValGluTyrAspSerThrGlyGlnTrpArgAsnMetGlnGln

1981 **GT**GGGCGAGGTGGGCATGACCCGAGATCACAGCAACACCGCCGCGTGCCTTAAGTGGAGC
ValGlyGluValGlyMetThrArgAspHisSerAsnThrAlaAlaCysLeuLysTrpSer

2041 GATCGCAAGTGCTTAGGTTTGTCCTTACTCTTTTTATGATCATGCAAGTCTTCTTTCTA
AspArgLys*CysLeuGlyLeuSerLeuLeuPhePheMetIleMetGlnValPhePheLeu*

2101 AACTTTAAACATGCCAACGACAACAACAACAAGAACAAAAACAACATTATCAAATGCATT
*AsnPhe*LysHisAlaAsnAspAsnAsnAsnLysAsnLysAsnAsnIleIleLysCysIle

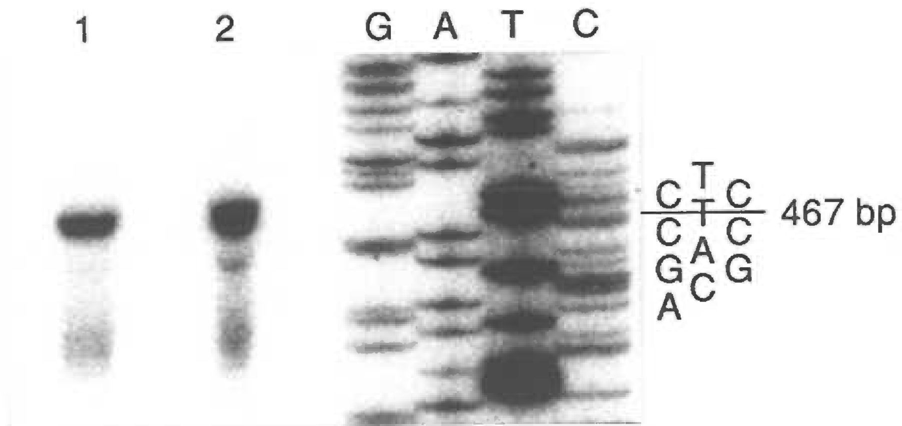
2161 AGATAATTTAGTAACAAATGAC
Arg

RNA Analysis By Primer Extension

To investigate how much (if any) 5' untranslated sequence was missing from the dKLIP-1 cDNA clone a primer extension analysis was performed. An antisense oligonucleotide was designed to anneal within the putative coding region (figure 5B, nucleotides 257-282) of the cDNA. The primer was predicted to produce an extension product of at least 282 nucleotides as determined from the primer annealing site to the 5' end of the dKLIP-1 cDNA (figure 5B). The 5' [^{32}P]-labeled oligonucleotide was annealled to poly-A+ RNA from fly heads and 0-2 hr. post-deposition embryos. The annealed primers were reverse transcribed to form an extension product which was sized alongside a sequence ladder by denaturing polyacrylamide gel electrophoresis and determined to be 467 nucleotides in length (figure 7). Therefore, the dKLIP-1 cDNA is missing 181 nucleotides of 5' untranslated sequence. The total size of the mRNA is calculated to be 3966 nucleotides, which closely matches in size the 4.0 kb transcript seen by Northern blot analysis (figure 11A and 13A).

Figure 7. Primer extension analysis of dKLIP transcripts.

Five micrograms of poly-A+ RNA from either Drosophila heads (lane 1) or 0-2 hr staged embryos (lane 2) was primed and extended from a 5'-end [^{32}P]-labeled 32mer as described (Kingston 1988). After RNase treatment, the denatured extension products were loaded on a 8M urea, 4% acrylamide sequencing gel and run until the xylene cyanol dye reached the bottom. A sequencing reaction using a known template primed with the lac Z 17mer was loaded alongside as a means of sizing the primer extension products.


1
2
G A T C
T
C T C
C A C
G C G
A
467 bp

Primary Structure Characterization of the Predicted dKLIP Polypeptides

Hydrophobicity analyses of the Kex2p-like endoproteases indicated that each has an amino terminal hydrophobic sequence (signal sequence) (figure 8 B). In the case of Kex2p, this sequence has been shown to target the protein into the secretory pathway where it becomes glycosylated (Fuller, Brake et al. 1986; Fuller, Brake et al. 1989a). A similar analysis of the predicted protein sequence of dKLIP-1 using a hydrophobicity algorithm (Kyte and Doolittle 1982) showed that, unlike the other Kex2p related proteins, dKLIP-1 lacks a signal sequence at its amino terminus. The hydrophobic domain nearest to the amino terminus is situated between residues 118-151 (figure 8 A,B). This analysis also indicated the presence of a second hydrophobic region near the carboxyl terminus in dKLIP-1 (residues 1006-1036), similar to Kex2p and furin, that is a putative transmembrane domain (TMD).

The predicted partial dKLIP-2 polypeptide maintains features found in dKLIP-1 including potential autocatalytic processing sites (figures 6 and 9), highly conserved catalytic domain and RGD sequence (residues 544-546). The carboxyl terminal region has a potential, although minimal in length (18 residue), transmembrane sequence of 18 hydrophobic amino acids (residues 684-702) which is followed closely by a region of polar residues. In contrast, however, the length of the presumed cytoplasmic domain of dKLIP-2 (20 residues) is almost 1/3rd as long as that of dKLIP-1 (51 residues) and hfurin (55 residues). If the

translation of the two dKLIP proteins is initiated from the same methionine residue then the predicted primary dKLIP-2 translation product is 892 amino acids (aa) which is closer to the size of Kex2p (814 aa) and furin (794 aa) than that of dKLIP-1 (1101 aa). The calculated pI of dKLIP-1 is 6.24 and its molecular weight comes to 121011.7 daltons, whereas the pI of dKLIP-2 is 6.65 and its molecular weight is 79653.83 daltons.

The percentages of overall amino acid sequence conservation between the dKLIPs and other members of the Kex2-like serine protease family are shown in Table I. Slightly higher percentages of sequence conservation were found in dKLIP-2 with respect to hfurin, hPC2 and mPC3 than with the related dKLIP-1 polypeptide sequence. The most conserved region of the dKLIPs are the subtilisin-like catalytic domains (figure 9A), suggesting functional and structural homology with other members of this family. In addition to highly conserved catalytic domains, Kex2p and its mammalian homologs have similar sizes, ranging from 638 residues for hPC2 to 814 residues for Kex2p (figure 8B). In contrast, the predicted length of dKLIP-1, 1101 residues, is significantly longer than any reported Kex2p related sequence. Alignment of the catalytic domains of each of the Kex2p-like homologs indicates that the longer size of dKLIP-1 is due to an extended N-terminus. The alignment also results in a shared positioning of two sets of paired basic residues in dKLIP-1 (Lys276-Arg277 and Lys308-Arg309) with a small cluster of paired basic amino acids in the other homologs that could be sites of autoproteolytic

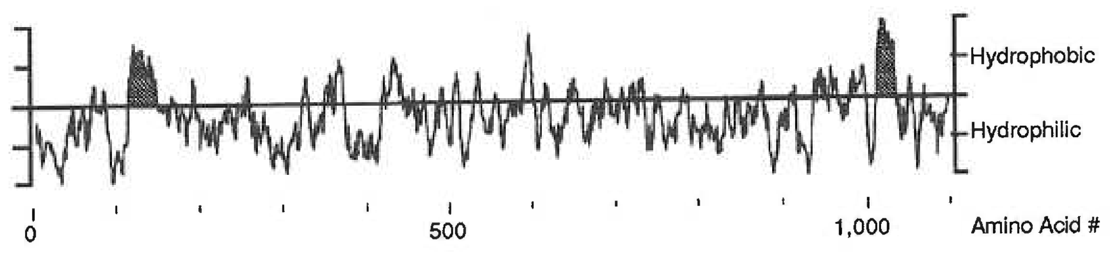
maturation; a property of subtilisin proteases (Power, Adams et al. 1986; Ikemura and Inouye 1988). Both of these potential cleavage sites are also present in the partial dKLIP-2 cDNA sequence. Another feature of the predicted dKLIP proteins that is conserved with PC2, PC3 and furin is an Arg-Gly-Asp (RGD) motif (figures 5B and 6). The RGD sequence, initially identified in fibronectin, is required for integrin binding and cellular adhesion (Pierschbacher and Ruoslahti 1984). Interestingly, the sequence of dKLIP-1 proximal to the predicted extracellular side of the TMD contains a cysteine-rich region as does hfurin (Roebroek, Shalken et al. 1990) (figure 9B). Within this region of the two proteins, twelve cysteine residues can be aligned with minimal gaps being introduced.

Figure 8. Structural features of the dKLIP-1 polypeptide with other Kex2p-like homologs.

A. Kyte and Doolittle hydropathy plot as performed by the Peplot program (Kyte and Doolittle 1982)(Devereux, Haeberli et al. 1984). The two hydrophobic regions that represent potential membrane spanning domains (res. 118-151 and 1006-1036) are filled with diagonal stripes.

B. Schematic diagram showing structural features and relative lengths of Drosophila KLIP-1, human furin, yeast Kex2p, human PC2 and mouse PC3 polypeptides. Ovals are potential N-linked glycosylation sites. Signal sequences are marked with vertically striped boxes. Putative TMDs are marked by diagonally striped boxes. Potential paired basic amino acid autoproteolytic cleavage sites (see Discussion) are marked as bold vertical bars. The subtilisin-like catalytic domains, with the catalytic triad residues (D=Asp, H=His, S=Ser), are shown by the stippled boxes.

A



B

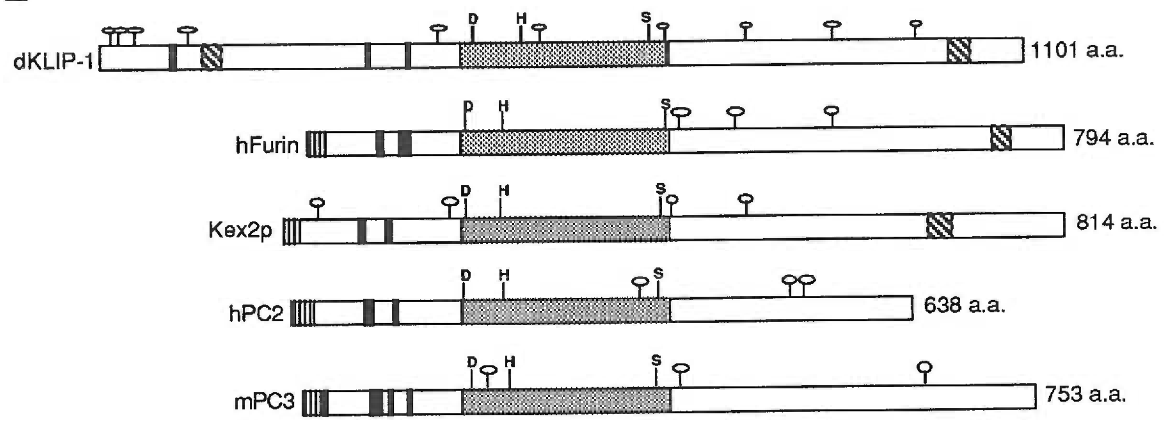


Figure 9. Selected amino acid sequence alignment analyses of dKLIP-1 with other Kex2-like homologs.

A. Amino acid alignment and consensus sequence of yeast Kex2p, dKLIP-1, hfurin, and hPC2 catalytic domains using the Bestfit program (Devereux et al., 1984). Residues important for substrate catalysis (*). The Asn residue involved in transition state hydrogen bonding, is conserved by dKLIP-1 but changed to an Asp residue in PC2 (Smeekens and Steiner, 1990).

B. Amino acid alignment using the Bestfit program (Devereux et al., 1984) showing the conserved alignment of cysteine residues (underlined) within the cysteine-rich region of dKLIP-1 and hfurin (Roebroek, Shalken et al. 1990; van den Ouweland, van Duijnhoven et al. 1990; Barr, Mason et al. 1991).

A

```
                                                             *
yKex2       134 NDPlFerQWhLvN----psfPGsDiNVldlWynniTGaGVVaaIvDDGldYeneDLkDNfcaEgSwDf
dKLIP-1     333 NDsKwPQmWYL-N-----rggGLDmNVipAWkmGiTGKGVVVtILDDGlesdHPDiqDNYDpkASYDv
hfurin      114 tDPKFPQQWYL------sGvtqrDLNVkaAWaqGyTGhGiVVsILDDGieknHPDLAgNYDpgASfDv
hPC2        122 NDPlFtkQWYLiNtgqadGtPGLDLNVaeAWelGyTGKGVtigImDDGidYlHPDLAsNYnaEASYDf

consensus       nDpkfpqqWyL--n----g-pglDlNV--aW--g-TGkGvvv-IlDDG--y-hpDladNyd-eaSyD-

                                  *
yKex2       198 NdNtnlPkPRls---DdyHGTRCAGEiAAkkgNNFCGVGVgYNAKisGiRiL-sGDiTtedEAaSLiy
dKLIP-1     395 NshDDDPmPhYdmtdsNRHGTRCAGEVAAtANNsFCaVGiAYgAsvGGvRMLD-GDVTDAVEARSLSL
hfurin      176 NdqDPDPqPRYTqmnDNRHGTRCAGEVAAvANNgvCGVGVAYNAriGGvRMLD-GeVTDAVEARSLgL
hPC2        190 ssNDPyPyPRYTddwfNsHGTRCAGEVsAaANNNiCGVGVAYNsKvaGiRMLDqpfmTDiiEAsSiSh

consensus       n--dpdP-Pryt---dnrHGTRCAGEvaA-anNnfCgVGvaYnak-gG-RmLd-gdvTdavEArSlsl

                                                                    *
yKex2       262 gldvnDIYScSWGPaDDGrhlqGPsDLvkkAlvkGVTeGRdsKGaIyVfASGNGGtrgDNCNyDGYTN
dKLIP-1     462 NPQHIDIYSASWGPdDDGKTVDGPgeLAsrAFieGtTKGRGGKGSIfiWASGNGGREqDNCNCDGYTN
hfurin      243 NPnHIhIYSASWGPeDDGKTVDGParLAeeAFfrGVsqGRGGlGSIfVWASGNGGREhDsCNCDGYTN
hPC2        258 mPQlIDIYSASWGPtDnGKTVDGPrDvtlqAmadGVnKGRGGKGSIyVWASGdGG-syDdCNCDGYas

consensus       npqhidIYSaSWGP-DdGktvdGP-dla--Af--GvtkGRggkGsI-vwASGnGGre-DnCNcDGYtn

                                                                           *
yKex2       330 SIysitIgaidhkdlhPpYSEgCSavmAvTYSSG-sG--Eyi-hssDinGrCsnSHgGTSAaAPLAAG
dKLIP-1     530 SIwTLSISSATeeGhVPWYSEkCSSTLATTYSSGgQG--EKQVVTTDLhhsCTvSHTGTSASAPLAAG
hfurin      311 SIyTLSISSATqfGnVPWYSEaCSSTLATTYSSGnQn--EKQiVTTDLrqkCTeSHTGTSASAPLAAG
hPC2        325 SmwTiSInSAindGrtalYdEsCSSTLAsTfSnGrkrnpEagVaTTDLyGnCTlrHsGTSAaAPeAAG

consensus       Si-t-sIssat--g-vpwYsE-CSstlAtTySsG-qg--Ekq-vttDl-g-Ct-sHtGTSA-APlAAG
```

B

```
dKLIP-1     851 CLKW-SDRKCLECNDSAYMFEDQCYDVCP
                |    |   |  |          |   ||
hfurin      592 CKTLTSSQACVVCEEGFSLHQKSCVQHCP

dKLIP-1     944 RVCAACDRSCLECYGALASQCSTC-SPGSQLRKILNETFC
                 ||| |  ||  | |     |  | |  | |     |  |
hfurin      639 SVCAPCHASCATCQGPALTDCLSCPSHAS-LDPV--EQTC
```

Figure 10. Comparison of predicted drosophila Kex2-like proteins.

A. The subtilisin-like catalytic domains are stippled. Active site residues are labeled (D=Asp, H=His, and S=Ser). Potential glycosylation sites are labeled with ovals. Transmembrane domains are bold cross-hatching. The darkened box in dKLIP-2 and Dfurin1 represents the sequence divergent to dKLIP-1. A 21 basepair sequence unique to Dfurin1 is lightly hatched.

B. Selected nucleotide sequence comparison of dKLIP-1, dKLIP-2 and dfurin1 showing the 21 basepair sequence unique to dfurin1 and the divergence point between dKLIP-2/dfurin1 and dKLIP-1 (Roebroek, Pauli et al. 1991; Hayflick, Wolfgang et al. 1992).

A

dKLIP-1 (1101 aa)

D H S

(A)n

dKLIP-2 partial sequence (722 aa)

D H S

dFUR-1 (898 aa)

D H S

(A)n

B

```
dKLIP-1   2701   GGCCAGTGGAGGAATATGCAGCAG---------------------GTGGGCGAGGTGGGC
                 ||||||||||||||||||||||||                     |||||||||||||||
dKLIP-2   1957   GGCCAGTGGAGGAATATGCAGCAG---------------------GTGGGCGAGGTGGGC
                 ||||||||||||||||||||||||                     |||||||||||||||
dfurin1   2554   GGCCAGTGGAGGAATATGCAGCAGCAATATCCTTTTCCTTTCCAGGTGGGCGAGGTGGGC

dKLIP-1   2761   ATGACCCGAGATCACAGCAACACCGCCGCGTGCCTTAAGTGGAGCGATCGCAAGTGCTTAG
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
dKLIP-2   1996   ATGACCCGAGATCACAGCAACACCGCCGCGTGCCTTAAGTGGAGCGATCGCAAGTGCTTAG
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
dfurin1   2634   ATGACCCGAGATCACAGCAACACCGCCGCGTGCCTTAAGTGGAGCGATCGCAAGTGCTTAG

dKLIP-1   2801   AG-TGCAATGACTCCGCCTATATGTTCGAGGACCAGTGCTACGATGTCTGCCCGG-TGCAC
                    ||   | ||||    | |||| ||  | |    | ||       ||       |||
dKLIP-2   2057   GTTTGTCCTTACTCTTTTT-TATGATCATGCAAGTCTTCTTTCTAAACTTTAAACATGCCA
                 ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
dfurin1   2685   GTTTGTCCTTACTCTTTTT-TATGATCATGCAAGTCTTCTTTCTAAACTTTAAACATGCCA
```

Table I. Percentage overall amino acid sequence conservation

Kex2p homolog	dKLIP-1	dKLIP-2
hFurin	48/65	51/67
mPC3	46/63	45/62
hPC2	44/62	44/61
Kex2p	29/52	32/54
Subtilisin	26/46	25/46

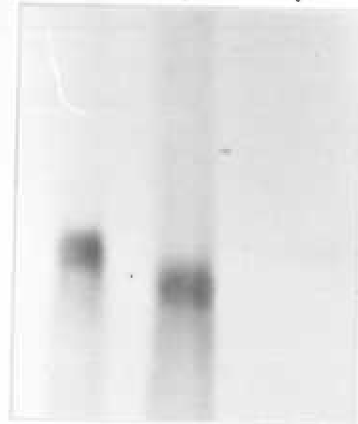
Data are presented as amino acid identities/similarities, the latter defined by evolutionary distance between the amino acids (Gribskov and Burgess 1986). This data was generated with a Sun 470 workstation and the Bestfit program (gap weight = 3.0, length weight = 0.1) from the Genetics Computer Group Sequence Analysis Software Package v6.2, 1990 (Devereux, Haeberli et al. 1984).

DKLIP-1 May Be Translocated Across Membranes

To determine whether the dKLIP-1 polypeptide, which appeared to lack an amino terminal signal sequence, could be translocated across microsomal membranes, a series of *in vitro* translation and translocation experiments were performed. Capped dKLIP-1 RNA was synthesized in vitro and translated in reticulocyte lysate reactions. SDS-PAGE analysis of the translation products showed a predominant protein product of approximately 140 kD (data not shown), in good agreement with the predicted size (131 kD) of the dKLIP-1 protein. To test whether the dKLIP-1 translation product was translocated across microsomal membranes, in vitro translation/translocation reactions were performed in the presence of membranes followed by protease treatment with or without detergent. As occurred in the absence of membranes, translation of the dKLIP-1 RNA in the presence of microsomal membranes also resulted in the synthesis of a 140 kD protein (figure 11, lane 1). Treatment of a replicate sample with proteinase K resulted in a 130 kD translation product that was protected from protease digestion (figure 11, lane 2). In contrast, when the translation/translocation reactions were protease treated in the presence of detergent to disrupt microsomal membranes, the protein was completely degraded (figure 11, lane 3). These results demonstrate that the dKLIP-1 cDNA encodes a 140 kD protein that can be translocated across microsomal membranes.

Figure 11. *In Vitro* translocation of dKLIP-1 across microsomal membranes.

Replicate tubes containing capped dKLIP-1 transcripts, rabbit reticulocyte lysate and [^{3}H]-leucine were incubated for 30 min at 30°C. Samples were then incubated in the absence (lane 1) or presence (lanes 2 and 3) of 0.1 mg/ml proteinase K and with 0.1% Triton X-100 (lane 3 only) for 45 min. on ice. SDS-PAGE analysis of the products and molecular weight standards was performed as described in Methods.


Triton – – +
Proteinase K – + +
140 kD
130 kD
1 2 3

DKLIP-1 is Proteolytically Active *In Vivo*

Previously both Kex2p and hfurin were shown to greatly enhance the endoproteolytic conversion of murine pro-β-nerve growth factor (pro-β-NGF) to mature 13 kD NGF *in vivo* (Bresnahan, Leduc et al. 1990). To assess the proteolytic activity of dKLIP-1, the protein was co-expressed with pro-β-NGF in cultured cells using vaccinia virus vectors and the conversion of the precursor to mature NGF was analyzed.

Replicate wells of BSC-40 cells were co-infected with a vaccinia virus recombinant expressing mouse pro-β-NGF (VV:NGF) and either wild-type vaccinia virus (VV:wt) or a vaccinia recombinant expressing dKLIP-1 (VV:dKLIP-1). Pulse-labeled products secreted into the medium were immunoprecipitated with an anti-β-NGF polyclonal antiserum and fractionated by SDS-PAGE. The results show that predominantly 40 kD pro-β-NGF was secreted from cells co-infected with VV:NGF and VV:wt (figure 12B, lane 1). However, when cells were co-infected with VV:NGF and VV:dKLIP-1, pro-β-NGF was completely converted to mature 13 kD NGF (figure 12B, lane 2). To determine the site of cleavage within the precursor, the 13 kD peptide generated from co-expression of dKLIP-1 and pro-β-NGF (figure 12B, lane 2) was transferred to a PVDF membrane and sequenced by Edman degradation (Matsudaira 1987). Release of [^{35}S] radioactivity was detected in the ninth cycle, which is consistent with the position of the first methionine within mature NGF formed by cleavage of the precursor at Lys^{119}-Arg^{120} (figure 12C). Thus, dKLIP-1

greatly enhanced conversion of the precursor to mature NGF peptide, by cleavage at a pair of basic amino acids.

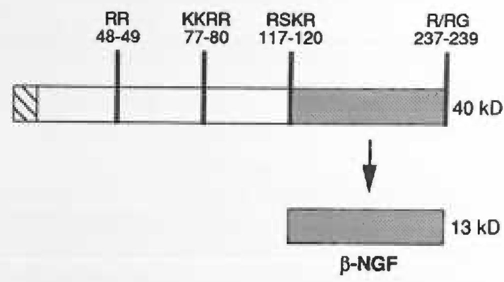
Figure 12. *In Vivo* proteolytic processing of pro-β-NGF by dKLIP-1.

A. Schematic diagram showing the precursor structure for the 40 kD pro-β-NGF (above) (Darling, Petrides et al. 1983) as expressed from the VV:NGF vector and its conversion to 13 kD β-NGF (below). The signal sequence is diagonally striped and potential paired basic amino acid cleavage sites are shown as bold vertical lines.

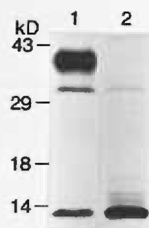
B. SDS-PAGE and fluorogram of immunoprecipitated products from pulse-labeled BSC-40 cells co-infected with (1) VV:NGF + VV:wt or (2) VV:NGF + VV:dKLIP-1. After a 30 min. methionine depletion, cells were metabolically labeled with [^{35}S]-methionine for 3 hrs. Equal volumes of media samples were then incubated with a β-NGF antiserum (Lara, McDonald et al. 1990) and the immunoprecipitated products analyzed as described in Methods. The glycosylated precursor migrates as a broad band of ~35-40 kD while mature β-NGF migrates at 13 kD. Products at 14, 16 and 17 kD are probably processing intermediates.

C. Sequential Edman degradation analysis of β-NGF from pulse-labeled BSC-40 cells co-infected with VV:NGF and VV:dKLIP-1. Cells were labeled for 3 hrs. with [^{35}S]-methionine and immunoprecipitated products separated by SDS-PAGE. Following transfer to a PVDF membrane (Immobilon, Millipore) and autoradiography, the 13 kD band was subjected to sequential rounds of Edman degradation from which aliquots were counted by liquid scintillation. The predicted N-terminal sequence of β-NGF is shown below.

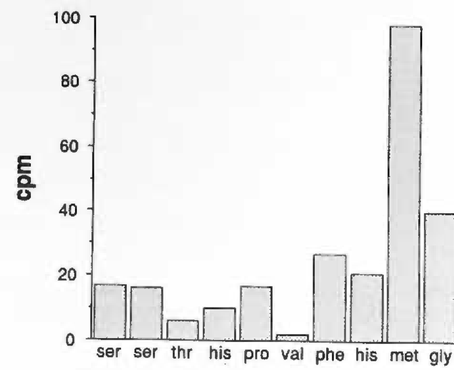
A.



B.



C.



Expression of dKLIP Transcripts in Adult Tissues

To determine the size of the dKLIP transcript(s), blots of poly-A+ RNA isolated from adult Drosophila heads and bodies were hybridized to a dKLIP-1 cDNA probe. Transcripts of 4.0, 4.5, and 6.8 kb were present in both heads and bodies at similar levels (figure 13 A). However, an 8.4 kb transcript was preferentially expressed in heads.

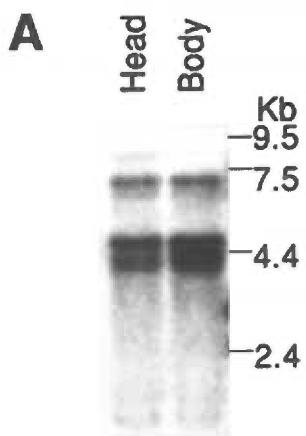
Determination of the anatomical distribution of dKLIP transcripts in adult tissues was performed on eight mm horizontal sections of male and female flies hybridized to [^{35}S]-labeled riboprobes. Antisense probes bound specifically to a number of tissues (figure 14 A,D,F), while control sense strand probes showed low, nonspecific signals (figure 14 C). In the head (figure 14 A,B), high levels of hybridization were present in all cortical regions of the brain containing neuronal cell bodies. In contrast, the neuropil, which contains cellular processes, showed only background levels of hybridization. A similar pattern was observed in neuronal cell bodies in the ventral ganglion (data not shown). Lower, but significant levels of hybridization were present in the fat body (figure 14 A,B).

Strikingly, maternally synthesized dKLIP mRNA was detected in nurse cells and oocytes (figure 14 D,E,F). In addition, follicle cells showed low levels of hybridization. The highest overall signals were observed in the oviduct (figure 14 D,E,F). These results demonstrate that dKLIP mRNAs are expressed at high levels in the adult CNS, fat body, and female reproductive tissues.

Figure 13. Northern blot analysis of head and body RNA from adult flies.

A. Polyadenylated RNA (10 μg per lane) from heads (lane 1) and bodies (lane 2) of both sexes, were isolated separately, blotted and hybridized with a [^{32}P]-labeled dKLIP-1 cDNA probe as described in Methods. Size markers are labeled in kb.

B. The same blot stripped and rehybridized with a [^{32}P]-labeled cDNA probe for the ribosomal protein, RP49 to quantitate relative RNA load per lane.


A
Head
Body
Kb
9.5
7.5
4.4
2.4

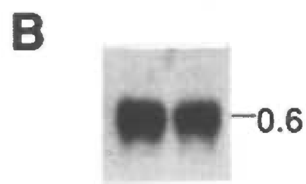
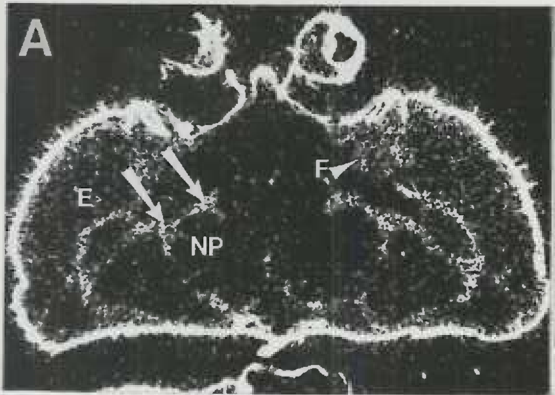
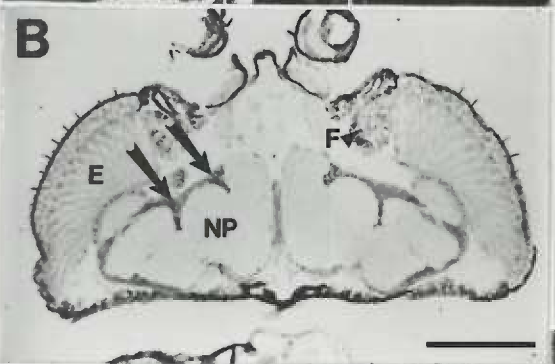
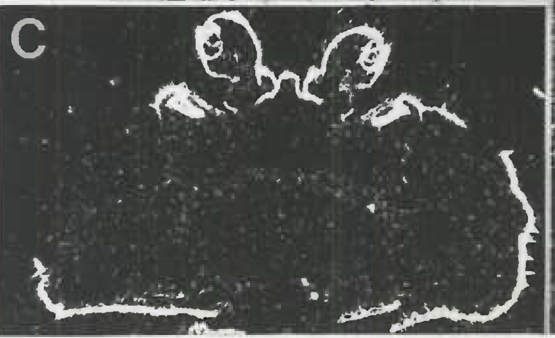
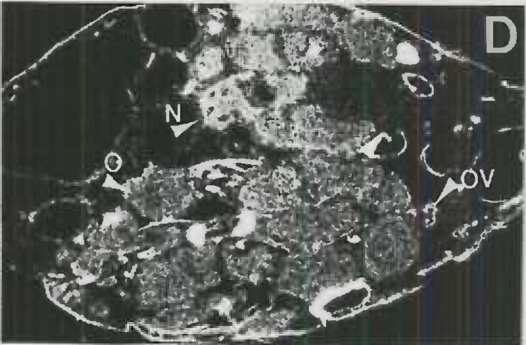

B
0.6

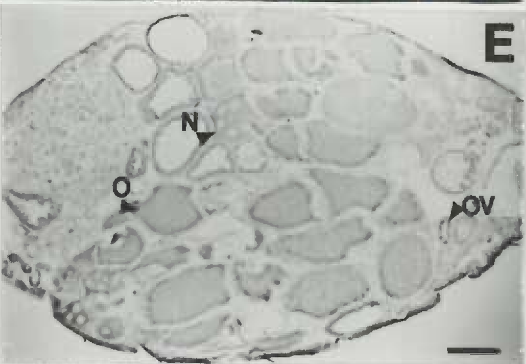
Figure 14. *In Situ* hybridization to adult head and ovary sections.
A and B. Darkfield and brightfield images, respectively, showing localization of dKLIP-1 mRNA using an antisense probe in a female head with high levels of hybridization in cortical brain regions (arrows), and the fatbody (F). The central neuropil (NP) of the brain has a background level of signal (compare with panel C). Low levels of hybridization in the eyes (E) are slightly above background and, therefore, are not a reliable indication of dKLIP-1 message. The high levels of signal outlining the head and abdomen results from non-specific binding of the probe to the extracellular cuticle and is a common artifact of the preparation. Anterior is up in panels A-C.
C. Darkfield image of an adjacent head section hybridized with a sense strand dKLIP-1 probe as a control shows low or background levels of hybridization.
D and E. Darkfield and brightfield images of a female abdomen hybridized with antisense probe showing localization of dKLIP-1 message to the ovaries including nurse cells (N), oocytes (O) and the oviduct (OV). Anterior is to the left in panels D-F.
F. High magnification image of the region surrounding the oviduct (OV) where the highest levels of expression are detected. Anterior is up in panels A-C. Anterior is to the left in panels D-F. Arrows, cell bodies of the brain; F, fat body; NP, neuropil; N, nurse cells; O, oocyte; OV, oviduct; bar = 200 μm

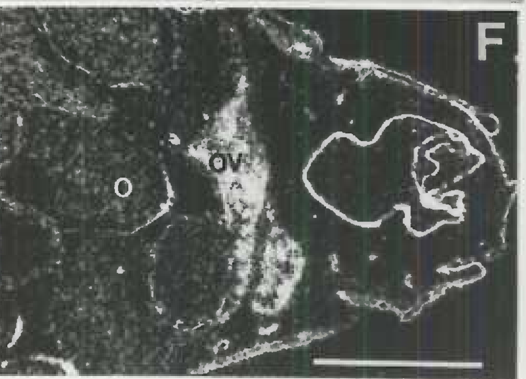

A
E
NP
F


B
E
NP
F


C


D
N
O
OV


E
N
O
OV


F
O
OV

The Expression of dKLIP Transcripts Are Developmentally Regulated

Based on the observation that dKLIP mRNA was detected in developing ooctyes, expression of the gene during Drosophila development was assessed by northern blot analysis. Poly-A+ RNA from embryonic, larval, and pupal stages was blotted and hybridized to a dKLIP-1 cDNA probe (figure 15A). A 4.0 kb RNA transcript was present at the earliest embryonic stages [0-2 hrs postfertilization (pf.)] (figure 15A, lane 1). From 2-10 hrs pf., the levels of all dKLIP transcripts were below detection. At 10-14 hrs pf., 4.5, 6.8 and 8.4 kb transcripts were evident with the 4.5 kb form being most abundant (figure 15A, lane 4). From 14 hrs pf. through pupal day 4, all dKLIP transcript levels were again below detection.

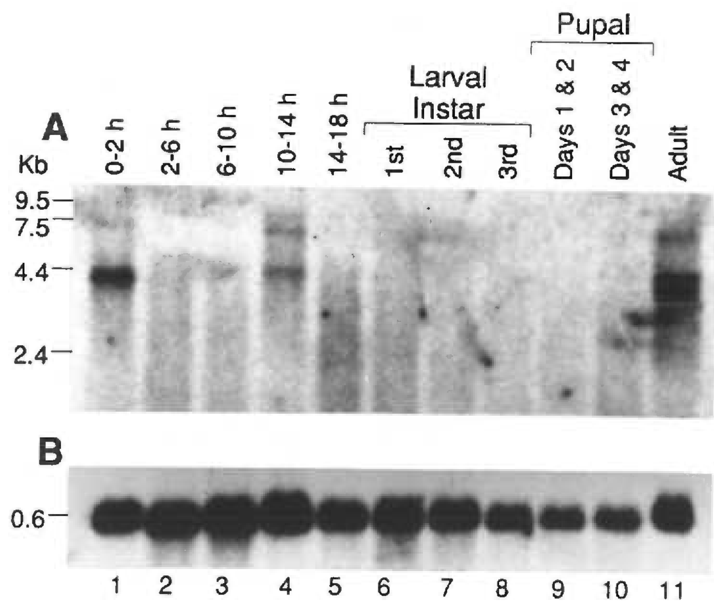
The spatial localization of dKLIP transcripts during embryogenesis was assessed by *in situ* hybridization of [^{35}S] and digoxygenin-labeled dKLIP-1 probes to embryonic tissue sections and whole mounts respectively. Hybridization was uniformly distributed in stages 1-4 (0-2.5 hr pf.) embryos (figure 16A,B and data not shown). The same photomicrograph shows a stage 10 embryo (5-7 hr pf.) which has undergone germband extension. Significantly reduced levels of hybridization were present at this stage as expected from Northern blot analysis (described above). This reduction in signal during embryogenesis was observed by stage 5 (cellular blastoderm formation, 2.5-3.25 hr pf.) (data not shown). Elevated levels of hybridization reappeared at stage 14 (11.5-

13 hr pf.) when signals were found in the anal pads, the hindgut, brain and ventral ganglia (figure 16C-F). Expression in the anal pads and hindgut was transient, disappearing by stage 16 (13-16 hr pf.). In contrast, expression in the CNS continued through stage 16 when it was found in whole mount preparations to be in the ventral ganglia in 12 sets of cell clusters located along the ventral midline and ventrolaterally (figure 17).

Figure 15. Temporal expression of dKLIP-1 transcripts during Drosophila development.

A. Polyadenylated RNA (10 μg per lane) isolated from indicated developmental stages grown at 25°C was blotted and hybridized with [^{32}P]-labeled dKLIP-1 cDNA probe. The adult RNA (lane 11) was a mixture of the RNA analyzed in the adult head/body blot (figure 11). Size standards are indicated in kb at the left of the panel.

B. Control hybridization using RP49 as an internal standard for quantitative RNA load per lane.



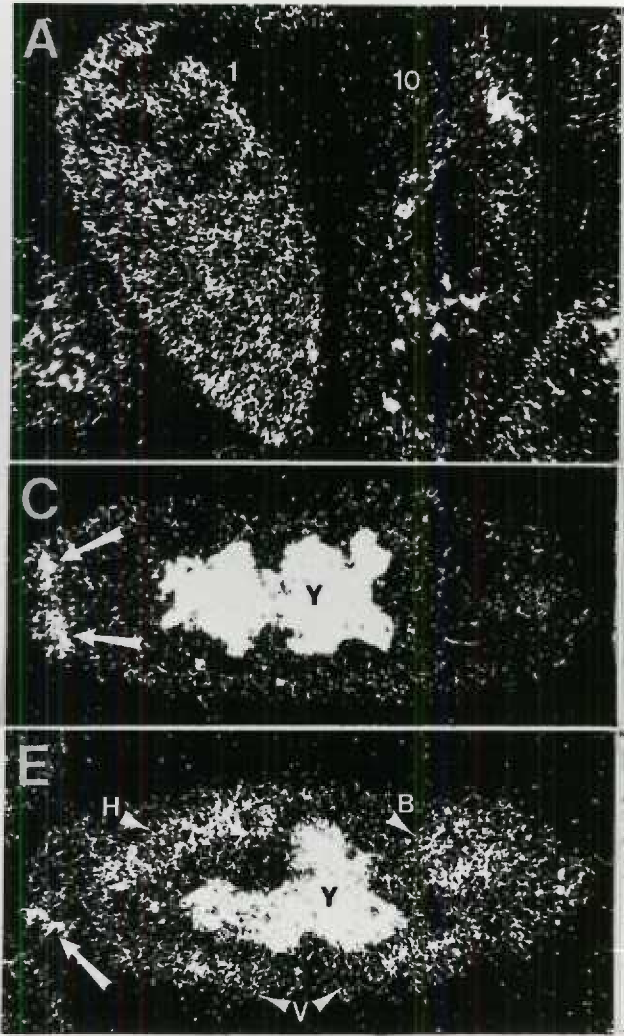
A
Kb
0-2 h
2-6 h
6-10 h
10-14 h
14-18 h
Larval
Instar
1st
2nd
3rd
Pupal
Days 1 & 2
Days 3 & 4
Adult
9.5
7.5
4.4
2.4
B
0.6
1
2
3
4
5
6
7
8
9
10
11

Figure 16. *In situ* localization of dKLIP-1 transcripts in eight micron sections of embryos.

A and B. Darkfield and brightfield images of an early nuclear cleavage (stage 1; ventral, right; anterior, down) embryo and a germband extended (stage 10; ventral, left; anterior, up) embryo. Note the elevated, signal in the stage 1 embryo compared to the background levels seen in the stage 10 embryo.

C and D. Darkfield and brightfield images of a horizontal section through a stage 14 embryo (anterior, right) showing signals located in the anal pads (arrows).

E and F. Darkfield and brightfield images of a parasagittal section through a stage 14 embryo (ventral, down; anterior, right) showing signals in the brain (B), hindgut (H), ventral ganglia (V) and anal pads (arrows). The central bright material in panels C and E is yolk (Y) and is an artifact of the preparation. In panels C-F, anterior is to the right. bar = 100 μm


A
1
10
C
Y
E
H
B
Y
V

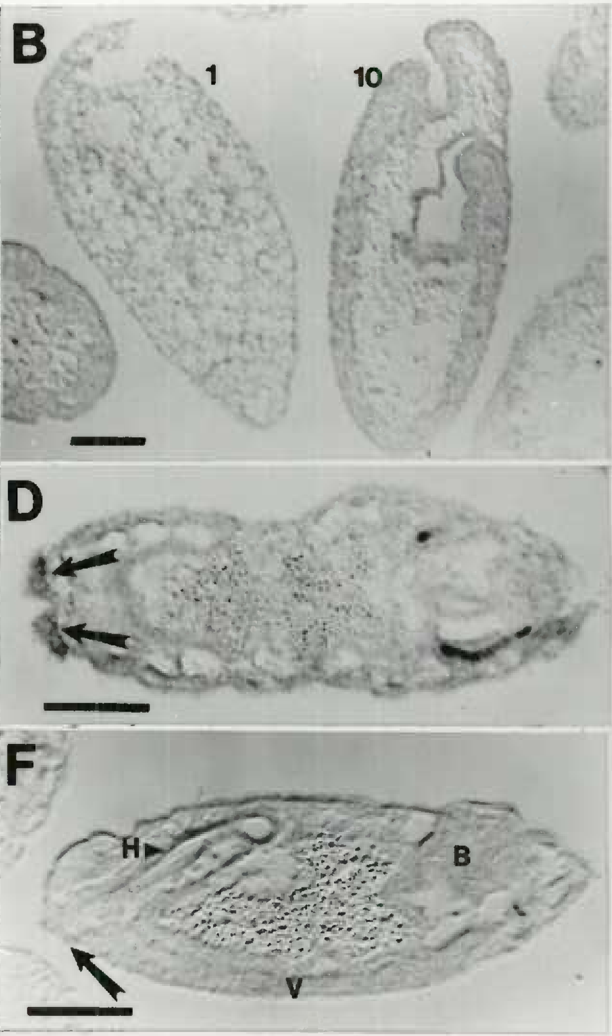
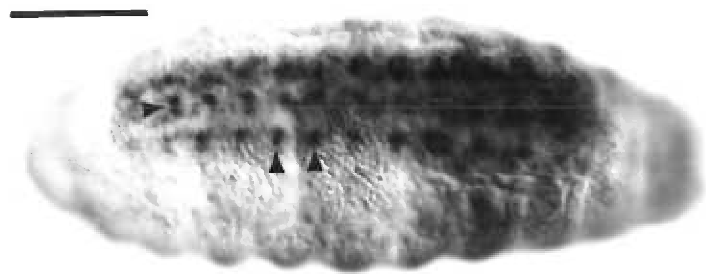

B
1
10
D
F
H
B
V

Figure 16. *In situ* hybridization detection in a stage 16 whole mount embryo using an antisense digoxygenin-labeled probe.
The signals are localized to three rows of cellbodies, two lateral (arrowheads) and one medial, in the ventral ganglia (anterior, left). bar = 100 μm.



Genomic Analysis of the dKLIP Gene

Northern blot analysis, as well as the isolation of cDNAs encoding divergent protein sequences from the dKLIP gene suggest that regulation at the transcriptional level may be critical for generating polypeptide diversity with the potential for different proteolytic activities. In conjunction with the cDNA clones, having the genomic organization in hand will allow identification of alternative exons. To begin to address this interesting problem the characterization and isolation of the dKLIP gene was initiated.

The location of the gene encoding dKLIP-1 was identified by *in situ* hybridization of a biotin-labeled dKLIP-1 cDNA probe to Drosophila salivary gland chromosomes. Hybridization was observed on the right arm of chromosome 3 at position 96D1-2 (figure 18).

The genomic Southern blot analysis (figure 19) shows several hybridizing bands from each lane (see Table II). The sizes of the bands add up to give a mean gene length of 28.75 (±3.22) kb. Since the dKLIP-1 cDNA used to probe the blot lacks 181 nucleotides of 5' untranslated sequence this estimate of gene length is understated.

Figure 18. Cytological localization of the dKLIP locus on polytene chromosome preparations. Larval polytene chromosomes were prepared and hybridized with a biotinylated dKLIP-1 cDNA probe essentially as described (Pardue 1986). The reaction product was determined to lie on the right arm of chromosome 3 at band position 96D 1-2.

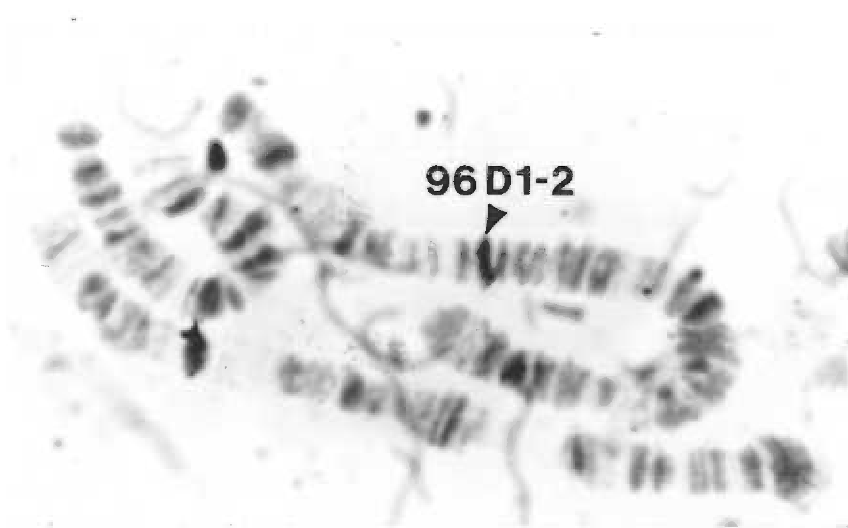

96 D1-2

Figure 19. Genomic Southern blot analysis of the dKLIP locus. Drosophila genomic DNA was digested with either Bgl II (lane 1), Pst I (lane 2), Bam HI (lane 3), or Eco RI (lane 4) and fractionated by electrophoretic mobility through a 1% agarose gel. The DNA was transferred and hybridized with a randomly-primed [^{32}P]-labeled dKLIP-1 cDNA probe. The blot was then washed and exposed for autoradiography. DNA size standards are designated in kilobase pairs.

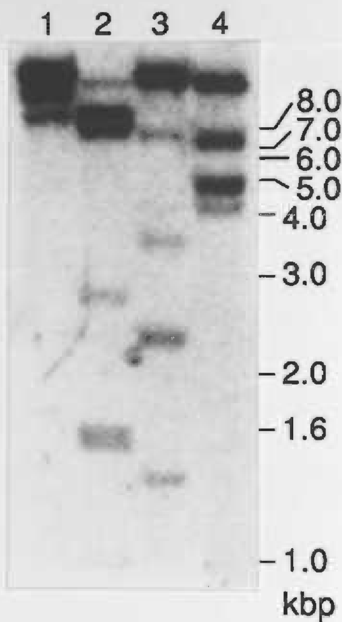

1
2
3
4
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.6
1.0
kbp

Table II. DKLIP Genomic DNA restriction fragment sizes

Restriction Endonuclease	Band Sizes (kilobase pairs)
Bgl II	7.8, 9, 13
Eco RI	4*, 4.4, 7.2, 11*
Bam HI	1.3, 2.3, 3.6, 5.9, 12
Pst I	1, 1.5, 1.6, 2.8, 7.5, 8, 11

The data are presented in kilobase pairs based on linear regression analysis of l Hind III DNA fragments. Bands in the Eco RI lane marked with an asterisk were observed during blot analysis of Eco RI digested genomic lambda phage clone DNAs.

Sixteen phages from the Drosophila genomic clone bank hybridized to the dKLIP-1 probe and were plaque purified and analyzed by restriction fragment analysis. Two classes of clones were recognized; those that shared (i) a 4 kb Eco RI fragment or (ii) a 10.5 kb Eco RI fragment (data not shown). The common fragments were found as anticipated, to be present in the genomic southern blot experimental results (asterisks in Table II).

DKLIP Protein Localization in Embryos

Localization of dKLIP proteins *in situ* is the next step in the progression of these studies. Isolation and characterization of dKLIP proteins from Drosophila extracts or cultured cells expressing these endoproteases will provide information regarding the molecular forms that the gene encodes, as well as their maturation pathways. In addition, if expression of dKLIP proteins were to be altered in future studies (possibly by P-element mediated insertion/deletion mutagenesis or construction of transgenic animals) one would want to make certain that protein levels and/or localization differed from wildtype animals. These experiments will require the production of specific antisera. This section describes the construction of fusion protein immunogens, the inoculation regimen used for raising an antibody response, and characterization of the resulting sera.

Fusion proteins were generated from plasmid DNA constructs using Escherichia coli trpE coding sequences fused to dKLIP coding sequences. These chimeric proteins contained either the active site region

(pathACT, using the PCR product initially amplified and cloned) or a region that spans the codons for the two potential paired basic residue cleavage sites believed to be critical for autoproteolytic maturation (pathPRO) (figure 20A,B). The expressed fusion proteins were found in the insoluble E. coli pellet fractions after lysis and migrated in SDS-PAGE experiments at the expected sizes (pathACT = 59 kD, pathPRO = 43 kD) (figure 20C). Fusion pathACT was purified from large scale preparations of insoluble E. coli pellet fractions, run on SDS-PAGE and electroeluted from gels. Fusion pathPRO was left for future antibody production and was not used as an immunogen.

Western blot and immunoprecipitation analyses demonstrated that the two rabbits immunized with pathACT produced antisera which recognize proteins of ~130 and ~160 kD expressed only from VV:dKLIP-1 infected BSC-40 cells (figure 21A,B). In addition, the two bands visualized on western blots could be competed from antibody binding by pretreating the immune sera with E. coli extracts expressing pathACT fusion protein (figure 21B, lanes 28-30 and 34-36). The molecular forms (2) and masses (130 and 160 kD) of dKLIP-1-like proteins seen by western blot and immunoprecipitation differ from the mass (140 kD) of the protein translated in *in vitro* expression studies. This could be due to several reasons including post-translational modifications (i.e. O-linked carbohydrates) and/or proteolytic processing activities not present *in vitro*. Alternatively, these antisera could be directed against epitopes related to the trpE portion of the fusion protein rather than the dKLIP-

1 active site and the protein products recognized by immune sera are unrelated to dKLIP-1. The fact that potentially two forms of dKLIP-1 were observed is consistent, however with studies of hfurin in which the smaller molecular mass form can be shown to be a proteolytic product of the larger form (Leduc, Molloy et al. submitted). Since the crude sera recognize multiple proteins in each analysis performed (i.e. a ~95 kD protein from VV:wt and VV:dKLIP-1 infected cells [figure 21A, lanes 3,7]), these sera cannot be considered specific until definitively proven so. This would require affinity purification of the sera by using purified dKLIP-1 protein and amino acid sequencing of the resulting immunoreactive proteins.

Proteins of 130 and 160 kD from drosophila extracts failed to be visualized, perhaps due to low abundance. Interestingly, a protein of 70 kD from adult drosophila body extracts did cross-react with serum #88. This is of interest because the Dfurin1 and putative dKLIP-2 polypeptides would be about 70 kD if (i) initiation of translation was at the same methionine residue used for dKLIP-1 and (ii) the putative pro region was cleaved from the rest of the protein at Lys308-Arg309.

Immune sera from each rabbit immunoprecipitated a 90 kD protein from mammalian cells transiently expressing hfurin (data not shown). Whether this protein is hfurin remains to be determined. In addition, serum #88 recognized a purified soluble form of hfurin (S. Molloy, personal communication). The other mammalian Kex2-like homologs (hPC2

and mPC3(PC1)) also appeared to be highly crossreactive with antisera #88 and #89 (S. Molloy, personal communication).

Results obtained from the incubation of embryo specimens with crude serum #88 must be interpreted with caution since the specificity for dKLIP-1 is questionable at this time. Immune serum #88 did show immunocytochemical staining of embryos that was greatest during cleavage stages 1-4 and during germ band shortening (stages 11-13). These embryonic stages also showed high levels of signal from *in situ* hybridization experiments (figure 16). The staining in early embryos was uniformly distributed over the surface with an anastomosing pattern (not shown). In stage 11-13 a population of cells located subepidermally, in the paraneural space (between the epidermis and retracting nerve cord) and along the entire ventrolateral axis, were observed to stain (figure 22A,B). During stages 15-16, granular staining was seen in the columnar epithelial cells of the esophagus and cells lining the midgut and hindgut (not shown). Both hindgut and nerve cord were found to express high levels of transcripts during these times by *in situ* hybridization. Whether these results are indicative of the presence of dKLIP-like proteins remains to be determined since the specificity of the antisera has not been definitively demonstrated.

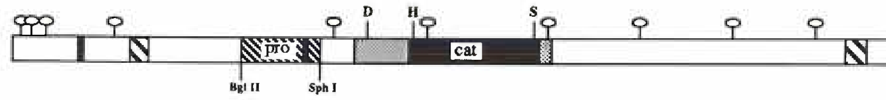
Figure 20. Production of fusion proteins in Escherichia coli.

A. Diagram of the domain structure of dKLIP-1 showing the putative pro region ("pro", diagonally striped box located between the Bgl II and Sph I sites) and subtilisin-like catalytic domain ("cat", black box).

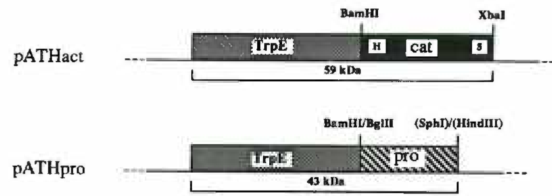
B. Diagram of the plasmid DNA constructs used to express trp E fusion proteins pathACT and pathPRO in E. coli.

C. Coomassie stained SDS polyacrylamide gel showing the relative sizes of the highly expressed fusion proteins. Molecular weight standards (lanes 1 and 14) are labeled in kilodaltons. Soluble (lanes 2,4,6,8,10,12) and insoluble (lanes 3,5,7,9,11,13) fractions were prepared from 5 ml induced cultures as described (Kleid, Yansura et al. 1981). Two isolates for each fusion protein (pathACT, lanes 4-7; pathPRO, lanes 8-11) were prepared. Control cells either without (lanes 1,2) or with path plasmid alone (lanes 12,13) were also prepared. The prominent bands at 59 kD and 43 kD migrate at the expected weights for pathACT and pathPRO, respectively.

A



B



C

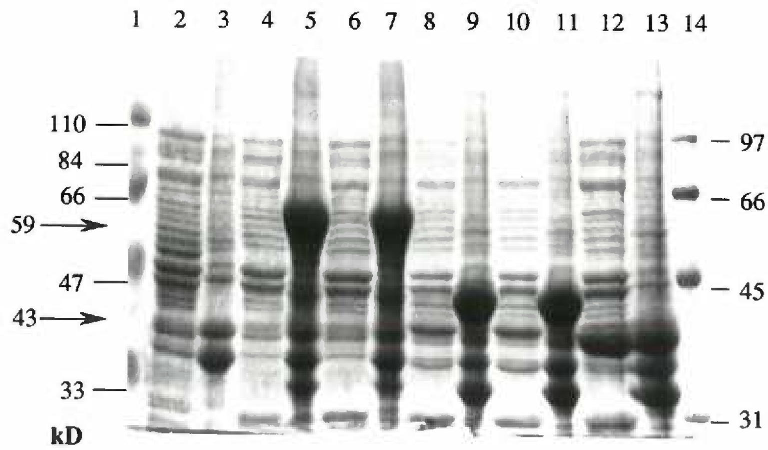
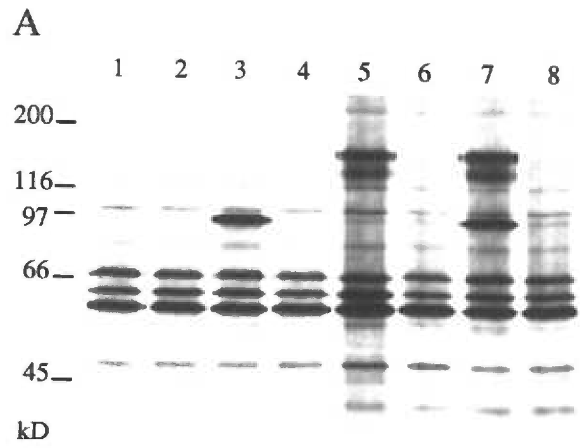


Figure 21. Antisera characterization using immunoprecipitation and western blot analysis.

A. Sera from the preimmune bleed (lanes 2,4,6,8) or boost #3 (lanes 1,3,5,7) from either rabbit 88 (lanes 3,4,7,8) or 89 (lanes 1,2,5,6) was incubated with BSC-40 cell extracts infected with either VV:wt/VV:NGF (lanes 1-4) or VV:dKLIP-1/VV:NGF (lanes 5-8). The infected cells were pulse-labeled with [^{35}S]-methionine prior to harvesting. The two bands with approximate molecular masses of 130 and 160 kD found in the cell extract from VV:dKLIP-1 infected cells incubated with either immune rabbit sera 89 (lane 5) or 88 (lane 7) are potentially two forms of dKLIP-1. The 95 kD band which is present in both VV:wt and VV:dKLIP-1 cell extracts incubated with bleed 3 from rabbit 88 (lanes 3,7) is possibly a vaccinia protein. The media from this experiment was used to isolate processed and unprocessed pro-β-NGF forms (see figure 11).

B. Western blot analysis. Soluble proteins from mock (lanes 1-12), VV:wt (lanes 13-24) or VV:dKLIP-1 (lanes 25-36) infected BSC-40 cell extracts were prepared and separated by SDS-PAGE as described in Methods. After transfer to nitrocellulose membranes, the lanes were divided, blocked and incubated with either preimmune serum 88 (lanes 1-3, 13-15, 25-27), preimmune serum 89 (lanes 7-9, 19-21, 31-33), immune serum 88, boost 3 (lanes 4-6, 16-18, 28-30) or immune serum 89, boost 3 (lanes 10-12, 22-24, 34-36). Sera was preincubated with lysates from either wildtype E. coli (lanes 2,5,8,11,14,17,20,23,26,29,32,35), pathACT expressing E. coli (lanes 1,4,7,10,13,16,19,22,25,28,31,34) or

lysate suspension buffer alone (lanes 3,6,9,12,15,18,21,24,27,30,33,36). Bound antibodies were detected with a goat anti-rabbit IgG conjugated to alkaline phosphatase. Two bands (arrowheads) of ~130 and ~160 kD are successfully competed from antibody recognition by preincubation of each immune sera with pathACT lysate only (compare lanes 28-30 and 34-36).


A
1 2 3 4 5 6 7 8
200
116
97
66
45
kD

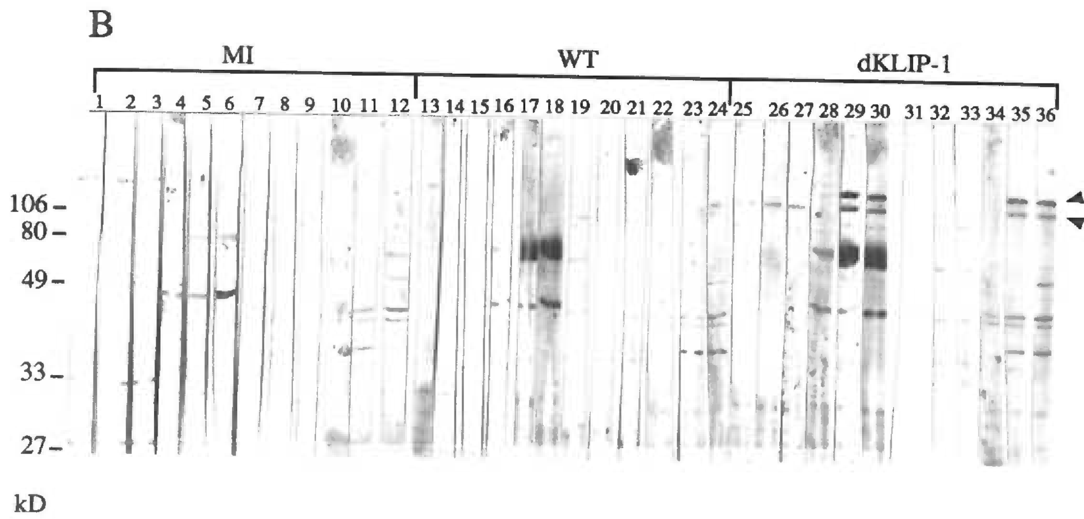
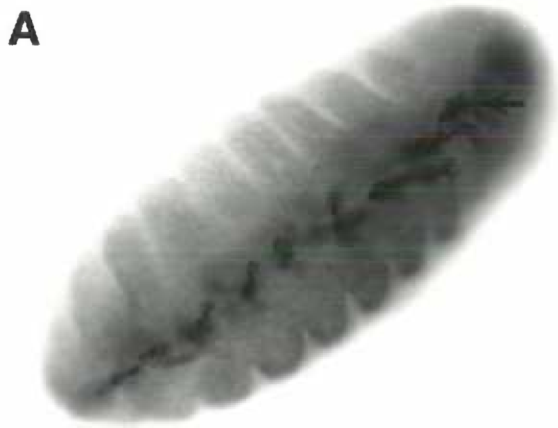

B
MI
WT
dKLIP-1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36
106
80
49
33
27
kD

Figure 22. Immunocytochemical localization in whole embryos.

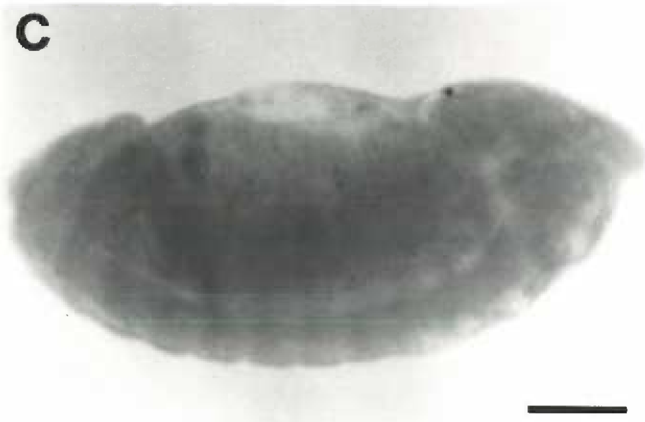
A. Stage 11 embryo stained with dKLIP (pathACT) antiserum as described in Methods. Cells along the ventral midline (arrow) show high levels of immunoreactivity.

B. Parasagittal view of stage 13 embryo showing prominent dKLIP immunoreactivity in cells (arrows) surrounding the ventral nerve cord (N), the brain (B) and into the clypeolabrum (C).

C. Parasagittal view of a stage 13 embryo incubated with pre-immune serum. In A-C, anterior is to the right. In B and C, ventral is down. bar=50 mm.


A

B
C
B
N


C

Experiments Designed To Interfere With DKLIP Expression In Transgenic Flies

As an initial means toward determining what developmental processes might involve the action of dKLIP-1, an experimental paradigm involving ectopic expression of sense and antisense transcripts was undertaken. This technique involves microinjection of a plasmid DNA P-element construct into stage 1 embryos that are expressing P-element transposase (Rubin and Spradling 1982). The activity of the transposase provided in *trans* allows integration of the P-element construct into the germline DNA of the embryonic pole (germ) cells. These flies are also deficient in an enzyme activity resulting in a white eyed (w) phenotype and by providing the enzyme coding sequences in the plasmid construct, flies can be selected for successful integration and inheritance of the P-element. This paradigm has been successfully used for the ectopic expression of a variety of proteins (Rubin 1988).

This experimental system was set up with the goal of using antisense mRNA to deplete the quantity of endogenous dKLIP transcripts thus reducing the level of endoprotease present. In this way, one might be able to generate a simple "knock-out" paradigm and ask the general question, "What developmental processes are perturbed if dKLIP expression is reduced or increased?" Antisense RNA transcripts or oligonucleotides have been successfully utilized to specifically inhibit the expression of particular target genes in several experimental systems. These RNA molecules anneal with target transcripts and prevent

(i) nuclear processing and/or transport of the complex to the cytoplasm (Munroe 1988) or (ii) ribosome binding and/or translation of mRNA in the cytoplasm (Ch'ng, Mulligan et al. 1989).

Two plasmid DNA constructions (figure 23) were generated for the purpose of either over- or under-expression of dKLIP proteins using P-element mediated germline insertion. These constructs carried the entire dKLIP-1 cDNA in either the sense or antisense orientation with respect to the heat-shock 70 (hsp70) promoter. The hsp70 promoter is a stress inducible promoter that is routinely used for ectopic expression of proteins in transgenic Drosophila (Pirrotta 1988). In practice the promoter is somewhat "leaky" such that enough transcription is always happening to allow for the production of the protein that complements the eye color defect in the recipient strain.

A total of 314 embryos were injected with buffer containing CsCl gradient purified plasmid DNA (213 with antisense and 101 with sense oriented cDNA). Twenty-four larvae hatched and 13 pupated. Twelve adult flies emerged (6 from antisense and 6 from sense strand construct injections) and were crossed. One hundred and seventy-four progeny were screened for red (or a shade from yellow to red) eye color. All progeny checked had the white eye phenotype. Thus none of the f1 progeny carried the integrated constructions.

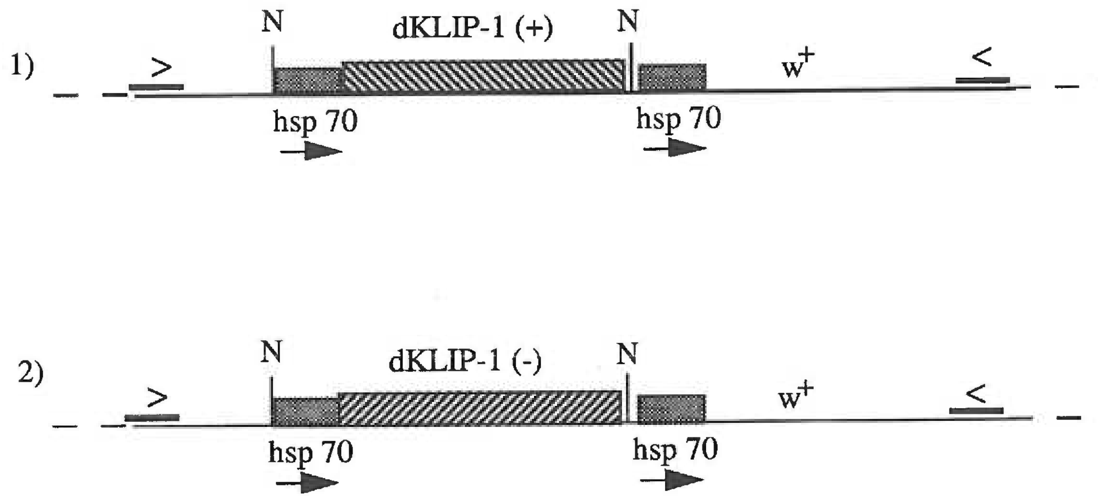


Figure 23. P-element plasmid DNA constructions used for embryo injections.

DKLIP1 cDNAs (striped boxes) and the hsp 70 promoter sequences (grey boxes) were ligated into the Not I (N) site of plasmid vector pW8 in either sense (1) or antisense (2) orientations. The white (w^+) gene is driven by its own hsp 70 promoter. P-element sequences are marked with >, < symbol.

Discussion

The Predicted dKLIP-1 Polypeptide Has Unique Structural Features

This dissertation documents the successful cloning of Drosophila homologs to the Kex2-like serine protease family and demonstrates the utility of this PCR-based strategy when cloning cDNAs from evolutionarily distant organisms. A similar strategy was employed by Sakanari and co-workers for the cloning of trypsin-like serine proteases and cysteine (thiol) proteases from several pathogenic organisms (Sakanari, Staunton et al. 1989). The nucleotide sequence of the dKLIP-1 cDNA (figure 5B) predicted an 1101 amino acid polypeptide containing the conserved Asp, His, and Ser catalytic triad residues characteristic of the subtilisin family of serine endoproteases. DKLIP-1 is structurally similar to a recently identified group of mammalian Kex2-like precursor processing endoproteases, furin, PC2, and PC3. Comparison of the predicted amino acid sequence of dKLIP-1 and -2 with the other Kex2p-like endoproteases revealed a high degree of sequence identity, particularly within the subtilisin-like catalytic domain (figures 8B, 9A and Table I). In addition, alignment of the catalytic domains resulted in a similar positioning of the C-terminal transmembrane domains (TMDs) in dKLIP-1, hfurin, and Kex2p, shared positions for a small cluster of paired basic amino acids, as well as a cysteine-rich region conserved in hfurin (figure 9B). In the case of

Kex2p, one of the paired basic sequences represents a proteolytic maturation site (Fuller, Brenner et al. 1991) which is conserved at nearly the identical position in each of the other endoproteases (figure 8B). Like the mammalian Kex2-like proteases, dKLIP-2 maintains at least one potential autoproteolytic processing site and the truncated C-terminal domain also includes a possible, although short, TMD.

A family of related serine proteases, the bacterial subtilisins, are synthesized as inactive pre-pro-enzymes. Subtilisins share the conserved Asp, His and Ser residues of the catalytic domains found in Kex2p, furin, PC2, and PC3 plus other regions of sequence conservation. A 29 amino acid signal sequence is required for protein secretion, while a 77 residue pro-sequence is required for proper protein folding of the remaining 275 amino acids (Wells, Ferrari et al. 1983). The pro region is removed by an intramolecular, autocatalytic event to release active subtilisin (Ikemura and Inouye 1988; Power, Adams et al. 1986). Similarly, the precursor of the Kex2p protease undergoes cleavage of its signal sequence and removal of its pro-peptide (Brenner and Fuller 1991; Wilcox and Fuller 1991). However, it has not been proven that Kex2p requires pro region cleavage for generation of its activity. Human furin is found in two forms (96 and 90 kD) and preliminary results demonstrate that the 96kD form is autocatalytically activated by cleavage on the carboxyl side of the sequence Arg-Thr-Lys-Arg^{186} to generate the 90 kD form (Leduc, Molloy et al. in preparation). Amino terminal sequencing of endogenous proteolytic activities suggests that

cleavage at a similar paired basic site in bovine PC2 and PC3, as well as anglerfish somatostatin pro-peptide converting enzyme takes place (Christie, Batchelor et al. 1991; Mackin, Noe et al. 1991). The dKLIPs also contain a potential autocatalytic processing site, -R-X-K/R-R-, within their N-terminal domain (figure 5B, residues 306-309; figure 6 residues 135-138) that can be aligned with similar sites in the other related proteases. It is therefore probable that the same activation mechanism takes place with respect to the dKLIPs.

There are, however, some marked differences between the structure of dKLIP-1 and the other Kex2p homologs. First, each of the previously reported homologs is similar in size. In contrast, the predicted translation product encoded by dKLIP-1 is 287 residues longer than Kex2p. Interestingly, the possible alternate translation products of the KLIP gene, dKLIP-2 and Dfurin1 (Roebroek, Pauli et al. 1991) are much closer in size to the mammalian homologs than is dKLIP-1. Alignment of the catalytic domain of dKLIP-1 with the catalytic domains of the other Kex2p-like endoproteases suggests that the larger size of dKLIP-1 is apparently due to an extended N-terminal sequence (figure 8B). Correspondingly, this amino-terminal extended domain of dKLIP-1 contains the least amount of sequence conservation compared with the other Kex2p-homologs. DKLIP-2 and Dfurin1 appear to have maintained the extended N-terminal domain of dKLIP-1, however they have much smaller domains on the C-terminal side of the catalytic domain (figure 10). Second, whereas each of the other Kex2p-like endoproteases possess an N-

terminal signal sequence to permit transit into the secretory pathway, dKLIP-1, Dfurin1 and probably dKLIP-2, all lack this domain. Nonetheless, when translated *in vitro*, dKLIP-1 appeared to be translocated across microsomal membranes (figure 11). In addition, co-expression of dKLIP-1 *in vivo* resulted in the complete conversion of pro-β-NGF at -Lys^{119}-Arg^{120}- in the secretory pathway (figure 12B,C). Possibly, the internal hydrophobic domain (residues 118-151 of dKLIP-1) unique to the dKLIPs and Dfurin1 facilitates their translocation into the lumen of the endoplasmic reticulum.

The results presented for the *in vitro* translocation experiments (figure 11) can be interpreted in an alternative manner, however. Protection of the 130 kD product could result from protein folding rather than membrane translocation, thus preventing exposure of sensitive sites to the protease. In the presence of Triton X-100, the 130 kD domain could be unfolded due to the denaturing effect of the detergent and exposing protease sensitive sites. These conclusions could be tested if the translation products were treated with protease K plus or minus Triton X-100 in the absence of membranes.

If dKLIP-1 is indeed translocated across membranes then several predictions can be made with respect to the orientation of the protein in the membrane. One hypothesis, the charge-difference rule (Hartmann, Rapoport et al. 1989) predicts membrane protein orientation based on charge differences, by subtracting the net charge of the 15 residues on the C-terminal side from that of the N-terminal side of a signal/anchor

sequence. Another model, the positive-inside rule (von Heijne 1986; von Heijne and Gavel 1988) predicts that positively charged amino acids are found on the cytoplasmic side of the membrane spanning segment. By either of these analyses, it could be predicted that both the N- and C-termini of dKLIP-1 are facing the cytoplasm forming a looping configuration with the catalytic domain disposed within the lumenal or extracellular space. However, the translocation data for dKLIP-1 suggests that only one of the termini faces the cytoplasm since ~10 kD is unprotected in the presence of proteinase K (figure 11). A size shift of greater than 21 kD is expected if both termini were facing the cytoplasm and were equally accessible to protease digestion, although the difference between 10 and 20 kD in this range of molecular masses is difficult to determine. The predicted topology of dKLIP-1 is with the C-terminus facing the cytoplasm and the rest of the protein within the lumen or extracellular space if only one terminus is anchored. If both termini are anchored however, the protein might take on a horseshoe-like configuration.

If either of the TMDs of dKLIP-1 are inserted in the membranes of microsomes then this suggests that the endoprotease could be localized to cellular membranes as well, possibly those of the same subcellular compartments that Kex2p and furin are found associated with. Kex2p and furin are Golgi membrane-associated enzymes (Bresnahan, Leduc et al. 1990; Redding, Holcomb et al. 1991). Intracellular retention of Kex2p requires both an intact cytoplasmic domain and a functional clathrin

heavy chain, suggesting an association between the C-terminus of Kex2p and clathrin in coated membranes (Fuller, Brake et al. 1989b; Payne and Schekman 1989). Clathrin association with the cytoplasmic domains of membrane proteins, such as receptors (and possibly including Kex2p and its homologs that possess TMDs) is thought to occur via an adaptor intermediate (Pearse and Robinson 1990). A recognition sequence involved with this interaction has not been reported, however that the cytoplasmic domain of Kex2p is required for proper localization suggests that one exists. The cytoplasmic domains of Kex2p and dKLIP-1 share some limited sequence conservation with proteins known to be concentrated in coated pits (i.e. man-6-pr and transferrin receptor).

An alternative idea regarding dKLIP localization holds that dKLIP proteases could be found on the external surface of cells. Indeed, the fact that the dKLIPs have a highly conserved integrin-receptor binding-site consensus sequence (Arg-Gly-Asp, dKLIP-1 residues 715-717) suggests that this might be true. This motif, found in the extracellular domain of fibronectin and other plasma membrane proteins, is required for cell-cell and cell-substratum attachment and its presence suggests that the dKLIPs may be involved in similar processes. Additional support for an extracellular role comes from studies of hfurin which, when expressed transiently in cells in culture, has been found to be active possibly at the level of the cell surface (S. Molloy and G. Thomas, unpublished observations). It is conceivable then, although far from proven, that the dKLIPs are involved in cellular attachment, via the Arg-Gly-Asp

site, and detachment, via their proteolytic activity. Interestingly, cell-associated proteases (i.e. plasminogen activator) are known to be important for the migration and invasiveness of neurite growth cones and neoplastic cells suggesting that proteolytic degradation of extracellular matrix components is critical for these processes (Monard 1988).

DKLIP Transcript Expression is Developmentally Regulated

In situ studies of adult fly sections detected dKLIP transcripts in ventral ganglia and brain as well as the fat body and hindgut (figure 14A,B). However, most striking was the detection of dKLIP transcripts in ovaries (figure 14D,E,F). Expression in developing oocytes and nurse cells demonstrated that these dKLIP transcripts are of maternal origin and implies a role for dKLIP in early embryogenesis, prior to zygotic transcription.

Detection of dKLIP transcripts in 0-2 hr embryos (figure 15A, lane 1 and Figure 16A,B) is associated with the expression of a variety of molecules involved in the determination of polarity and positional information in the embryo (Ingham 1988). Later in development, dKLIP transcripts are expressed in embryos at discrete times and restricted to specific cell populations. After germ band shortening (10-14 hrs) RNA levels were detected in the hindgut, ventral ganglia, brain and anal pads (figures 16C-F and 17). During this time the hindgut lengthens and the neuropil forms in the nervous system.

After stage 4, dKLIP transcripts appeared to be rapidly degraded (figure 15A, lanes 1,2). The loss of signal suggests that dKLIP-1 expression is developmentally regulated, in part, by rapid turn over of its RNA. DKLIP-1 and several mammalian mRNAs (e.g. GM-CSF, c-fos, and c-myc) possess a motif (AUUUA) repeated within their 3' UTRs (Shaw and Kamen 1986) (figure 5B). Mammalian transcripts possessing this motif are selectively and rapidly degraded.

In addition to the AUUUA sequence in the 3' UTR, the dKLIP-1 transcript contains an opa element (figure 5B, nucleotides 258-347), a repeated trinucleotide (CAXn) located in the non-conserved amino-terminal domain of the protein coding sequence (Wharton, Yedvobnick et al. 1985). Although the function of this element is unknown, opa sequences are present in several developmentally regulated Drosophila genes including notch, antennapedia, and dorsal (Grabowski, Carney et al. 1991).

The immunocytochemical analysis using crude sera that was not demonstrated to be dKLIP specific, produced a similar pattern of distribution that was seen from the *in situ* RNA localization studies. The finding that cross-reactive proteins were concentrated in the cells that surround the neuromeres of stage 11-13 embryos is of some interest. These cells are a population of phagocytes that are involved in scavenging dead cells from sites of tissue modification. Interestingly, a similar phagocyte population in imaginal disks is dependent upon the activity of dpp, a protein released from its precursor at an Arg-X-X-Arg

cleavage site (see below) (Bryant 1988). It is impossible to know at this time whether the proteins detected by the antibodies are dKLIP-1 or 2 (or other related proteins) since they both share the same active site region used in the construction of the fusion protein.

Further characterization of the antisera using dKLIP-1 affinity purification must be done to determine the specificity of the sera. Any results obtained from the antisera discussed in this thesis must therefore be cautiously interpreted. In particular, since dKLIP-1, -2 and Dfurin1 all share the same catalytic region used for the fusion protein immunogen, discrimination of these and other dKLIP-like proteins must rely on the production of antisera that specifically reacts with each protein uniquely.

Evidence for Complex Transcriptional Regulation for the dKLIP gene

The genomic DNA analysis documented here shows that the dKLIP gene is spread over approximately 29 kb on the right arm of chromosome three at position 96D1-2. These studies suggest that dKLIP is a single locus although localized gene duplication cannot be disproven from these data. The pattern of transcripts seen by northern blot experiments (figures 13 and 15), as well as the sequences of the variant cDNAs reported here and the newly reported Dfurin1 sequence (Roebroek, Pauli et al. 1991) suggest that this gene is transcribed into a complex set of RNA molecules. The possibility that dKLIP primary transcripts are differentially processed is supported by the following evidence. The

cDNAs for dKLIP-1 and -2 differ from each other at a consensus exon/intron splice junction sequence (AG/GT, figure 10). It is possible that this could be the result of incomplete RNA processing, at least in the case of dKLIP-2. Such intermediates are common from certain drosophila genes (i.e. voltage-gated K^+ channel, J. Adelman, personal communication) and could be a key point for gene regulation. Furthermore, the recent report describing the isolation of the Dfurin1 cDNA which contains sequence identical to dKLIP-2 except for an additional 21 nucleotides inserted at postion 1980 of the dKLIP-2 cDNA sequence (figure 10B) (Roebroek, Pauli et al. 1991). The additional coding sequence of 21 bp found in Dfurin1 is bounded by consensus splice junctions upstream of the dKLIP-1/2 divergence point (figure 10B). Importantly, the Dfurin1 cDNA clone has a polyadenylation site and poly-A tract, suggesting that it represents a fully processed transcript.

Additional support for the alternative RNA splicing phenomenon found in dKLIP cDNAs, in particular the diversity at their 3' ends, comes from a recently characterized *Caenorhabditis elegans* Kex2p homolog (Peters and Rose submitted). Several transcripts are generated by this gene (*Blister* -4, *bli*-4) which differ in their splicing at the C-terminal-encoding region resulting in synthesis of three Kex2p-like products from a single gene. An allele was found that deletes one of the alternative exons resulting in a transcript that deviated in its C-terminal coding sequence and the possible selective loss of function for one of the gene products. The isolation and sequence analysis of dKLIP

and Dfurin1 cDNA clones suggests that the alternative use of protein-coding exons is not unique to the C. elegans *bli*-4 gene. Taken together, these data support the hypothesis that a single gene for the dKLIPs is transcribed to yield RNAs which utilize alternative C-terminal protein coding exons and 3' untranslated regions.

Expression of dKLIP-1 in mammalian cells

Murine pro-β-NGF can serve as a substrate for dKLIP-1 using transient expression in cultured monkey cells and the site for cleavage within the NGF precursor is Arg-Ser-Lys-Arg (figure 12). Similarly, hfurin can also process the pro-β-NGF precursor (Bresnahan, Leduc et al. 1990; J. Hayflick, unpublished results). These results demonstrate that dKLIP-1 can potentiate the endogenous processing of a mouse growth factor precursor with an Arg-X-Lys/Arg-Arg cleavage site and that proteolysis occurs in a subcellular compartment that allows for mature NGF to be released constitutively. Formal proof that dKLIP-1 cleaves pro-β-NGF by direct protein interaction must await *in vitro* experiments using purified dKLIP-1 and pro-β-NGF. In support of a direct enzyme/substrate interaction are the results obtained in vitro with a purified, soluble, truncated form of hfurin and a purified substrate protein (Molloy, Bresnahan et al. 1991). Using this paradigm cleavage site preferences were explored. The truncated furin protein cleaved at sites with a single Arg or Lys-Arg pair if accompanied at the -4 position by an Arg residue. In addition, the inhibitor profile showed it to be a serine protease and to be strongly inhibited by zinc metal

ions, as anticipated. Whether dKLIP-1 will cleave similar sites within Drosophila precursor proteins and show the same inhibitor profile is uncertain at this time.

Roles for Serine Proteases in Early Fly Development and Potential Substrates for the dKLIPs

Numerous studies have demonstrated that trypsin-like activities and other proteins with homology to mammalian serine proteases have essential roles in fly development. During larval metamorphosis, for example, imaginal disk eversion appears to be regulated by a trypsin-like protease activity since trypsin treatment of disks in culture increased the rate of eversion (Pino-Heiss and Schubiger 1989). Conversely, trypsin inhibitors retarded this process. During early embryogenesis, dorsal-ventral axis formation is mediated by a biochemical pathway that includes the products of two genetic loci, *easter* and *snake* (DeLotto and Spierer 1986; Chasan and Anderson 1989). The products of both loci are structurally related to mammalian trypsin/chymotrypsin serine proteases, and include the His, Asp, and Ser residues forming the characteristic catalytic triad (as opposed to the Asp, His, and Ser order in the subtilisin-like triad) (Neurath 1984).

Unlike the trypsin/chymotrypsin family of endoproteases, the Kex2p-like subtilisin endoproteases specifically cleave substrates on the C-terminal side of paired basic amino acids (Lys-Arg, Arg-Arg) in the secretory pathway (Fuller, Brake et al. 1989b; Mizuno, Nakamura et al. 1988; Thomas, Thorne et al. 1988b; Thomas, Leduc et al. 1991).

Analysis of mutant substrates (Bentley, Ree et al. 1986; Yoshiro and Webster 1988; Yoshimasa, Seino et al. 1988; Yoshimasa, Paul et al. 1990) suggests that an additional feature of the cleavage site commonly found in many precursor proteins is an arginine located four residues (-4 Arg) N-terminal to the paired basic cleavage site (-Arg-X-Lys/Arg-Arg-). Indeed, studies in vitro indicate a preference by dKLIP-1 for synthetic peptide substrates containing a pair of basic residues and a -4 Arg (J. Hayflick and G. Thomas, unpublished results).

Several Drosophila precursor proteins possessing the -R-X-X-R- cleavage site motif, have been shown to be homologs of mammalian proteins (figure 24). In addition, they are known to be expressed during fly development and could therefore overlap spatially and temporally with dKLIP-1 expression. One of these, the product of the *decapentaplegic (dpp)* locus, is a protein with structural similarities and homology to the transforming growth factor beta (TGF-ß) superfamily (Padgett, St. Johnston et al. 1987). The proteins in this family contain a conserved bioactive sequence at their carboxyl terminus which, in most cases, is liberated by cleavage on the C-terminal side of an Arg-X-Lys/Arg-Arg or Arg-X-X-Arg site (Gentry, Lioubin et al. 1988; Lyons, Gentry et al. 1990; Mason, Hayflick et al. 1985; Pepinsky, Sinclair et al. 1988). The *dpp* product is known to be a secreted factor that has multiple roles during development. *Dpp* transcripts have been detected on the dorsal surface of stage 11 embryos and through dorsal closure in epidermal cells (St.Johnston and Gelbart 1987). Dpp is also expressed

in a subset of imaginal disk cells, without which these cells die (Bryant 1988; Posakony 1987).

Biochemical studies based on expression of decapentaplegic translation product (dpp) in Schneider 2 cells resulted in the cleavage of pro-DPP on the C-terminal side of a single basic residue with a -4 Arg (Panganiban, Rashka et al. 1990). A conserved paired basic residue site with a -4 Arg (-Arg-Asn-Lys-Arg-), proximal to the cleavage site in pro-dpp, was not utilized as it is in the structurally related pro-TGFb1 and the inhibin precursors (Derynck, Jarrett et al. 1985; Mason, Hayflick et al. 1985). It has not been reported if processing of pro-DPP in S2 cells mimics the endogenous processing in the developing embryo. Studies examining the ability of dKLIP-1 to process these precursors are currently underway.

In addition to dpp, a recently characterized Drosophila locus, *tolloid (tld)* (figure 24), encodes a protein that is closely related structurally and by sequence comparison to the bone morphogenic protein, BMP-1 (Shimell, Ferguson et al. 1991). *Tld* transcripts are expressed in early pre-cellularized embryos as well as in germ band extended embryos. Pro-tld has a signal sequence and a metalloprotease-like domain that is preceded by a tetra basic residue sequence that could be a potential cleavage site. Whether cleavage of the tld precursor occurs at this site remains to be reported.

A potential substrate protein in the receptor class is the Drosophila insulin-like pro-receptor (dILR) (figure 24). The transcript

has been found to be expressed predominantly in mid-embryonic development (Nishida, Hata et al. 1986). In mammals, the receptor is synthesized as a larger precursor from which the α and ß subunits are cleaved intracellularly at a paired basic amino acid site with a -4 Arg. The dILR shares considerable amino acid sequence conservation with the human insulin pro-receptor including the cleavage site, however its synthesis and structure specifically with regard to proteolytic processing has not been reported.

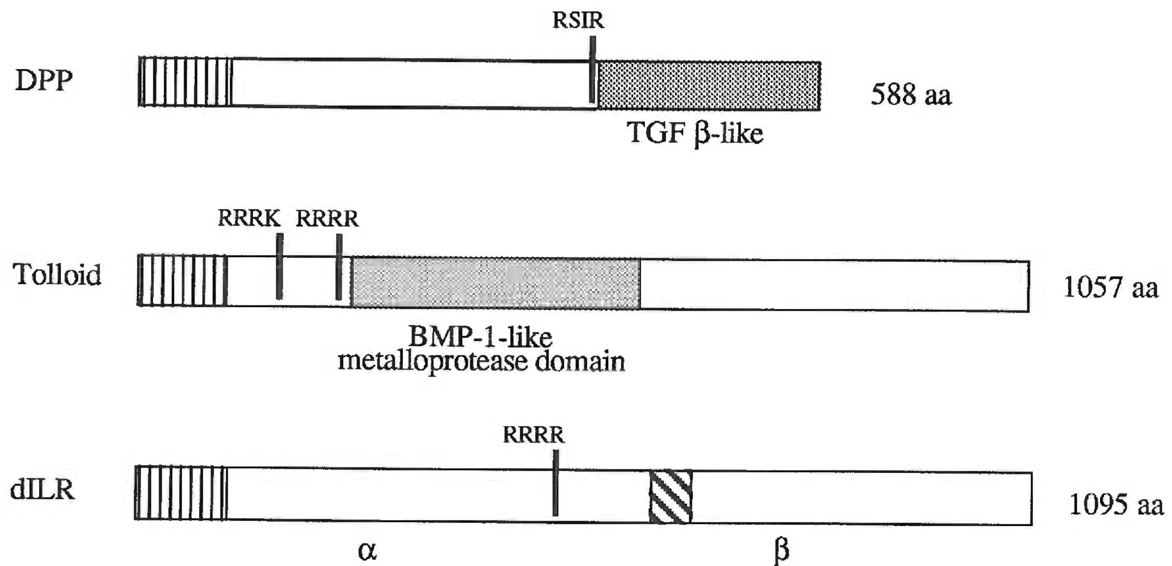


Figure 24. Potential precursor protein substrates of the dKLIP endoproteases.

Signal sequences are marked with vertically striped boxes. Putative or documented processing sites are marked with the bold line and the four residues preceding the cleaved peptide bond. Transmembrane domains are cross hatched. These data are from Panganiban, Rashka et al. 1990; Shimell, Ferguson et al. 1991 and Nishida, Hata et al. 1986.

Perhaps something can be learned from the genetic analysis of C. elegans *bli*-4 as to the potential substrates and function of dKLIP-1, as well as mutant phenotypes of the dKLIP gene. Disruption of the *bli*-4 gene results in localized separations of the adult-stage specific cuticle layer from the basement layer that become fluid filled blisters (Peters, McDowall et al. 1991). Five loci beside *bli*-4, can also be mutated leading to the development of blisters (Brenner 1974; Park and Horvitz 1986). In addition to the *bli*-4 gene product (blisterin) which is a Kex2p serine protease homolog (Peters and Rose submitted), an adult cuticle specific component, possibly a substrate or enzymatic cofactor, is required for blister formation (Peters, McDowall et al. 1991). Reports of the expression and proteolytic activity of blisterin have not been made at this time. Many mutations that affect worm morphology, including those known to encode structural proteins like collagens, also affect the blister phenotype (Cox, Laufer et al. 1980). The interaction of blisterin with proteins likely to be part of the extracellular matrix (ECM) suggests that one or more of the structural proteins of the cuticle requires endoproteolytic processing for proper function. Alternatively, the blister phenotype may result from mis-localization of blisterin enzymatic activity to the ECM and subsequent detrimental cleavage of structural proteins.

Conclusions and Summary

The study that is documented in this thesis is motivated by the development and use of genetic tools to understand the role of dKLIP-1 during the life cycle of Drosophila and to make generalizations about the roles of furin and the other Kex2p-like endoproteases in metazoans. This broad understanding will come from answering some of the following questions. What substrate(s) does dKLIP-1 interact with? When do these interactions occur during the development and life of the organism? Do the enzymes have functions beside catalyzing endoproteolysis?

A mutational analysis of the dKLIP locus could provide answers to these questions. Understanding the genomic organization of the dKLIP gene will be useful for interpreting results from P-element mediated insertional mutagenesis or radiation-induced translocation experiments. A detailed genomic analysis will also provide information regarding the relative locations of the alternative exons found in the dKLIP-1 and -2 cDNAs. Comparison of the Drosophila locus with the *blisterin* locus of C. elegans may allow identification of putative regions critical for regulation of splicing. *In situ* hybridization to embryo and adult tissues with dKLIP-1 or -2 exon-specific probes would reveal whether the regulation is tissue and/or developmentally specific. This analysis could be backed up with developmental and adult RNA blot studies using the same probes to characterize specific transcript lengths, as well as temporal regulation. Promoter studies could be initiated to identify

specific promoter elements. These elements could then be placed in P-element based plasmid constructs driving a reporter molecule (i.e. β-galactosidase) for temporal and spatial expression studies. The use of a sensitive reporter like β-galactosidase could reveal tissues that drive low levels of expression from the dKLIP promoter element. Such tissues are currently being overlooked by the relatively insensitive localization studies to date with either RNA hybridization or immunochemical methods.

The transgenic fly experimental system described above was set up with the goal of using antisense mRNA to deplete the quantity of endogenous dKLIP transcripts thus reducing the level of endoprotease present. In this way, one might be able to generate a simple "knock-out" paradigm and ask the general question, "What developmental processes are perturbed if dKLIP expression is reduced or increased?" Antisense RNA transcripts or oligonucleotides have been successfully utilized to specifically inhibit the expression of particular target genes in several experimental systems. These RNA molecules anneal with target transcripts and prevent (i) nuclear processing and/or transport of the complex to the cytoplasm (Munroe 1988) or (ii) ribosome binding and/or translation of mRNA in the cytoplasm (Ch'ng, Mulligan et al. 1989).

The transgenic animal section of the Results chapter details the methodology and results for the experiments designed to knock out dKLIP expression or cause overexpression of dKLIP-1 protein in transgenic

flies. None of the f1 progeny from the injected embryos showed the red eye color phenotype indicating none had the P-element construct integrated in the germ line. One reason may have been the relatively high death rate of injected embryos. This could have been due to the large needle taper generated during the needle pulling operation which created large openings in the dechorionated embryos and allowed leakage of substantial quantities of cytoplasm. In addition, the embryos floated on the hydrocarbon oil during the post-injection recovery period and appeared to dessicate. The technique of embryo microinjection is tricky and these experiments had batch to batch and day to day variations in several parameters. Finally, the plasmid vector pHSREM has inherent expression problems resulting in accumulation of transcripts synthesized from the hsp70 expression cassette but little protein translated (M. Forte, personal communication).

Experiments designed to properly "knockout" the dKLIP locus via homologous recombination may provide information regarding the developmental processes in which the dKLIP proteases participate (Gloor, Nassif et al. 1991). This method (gene replacement by gap repair) takes advantage of the double stranded gaps left as a result of P-element transposition and the repair of these gaps using a sister strand or another homologous strand as a template. By supplying a template sequence (i.e. an *in vitro* modified sequence) that has homology to the gap and flanking region, it is possible to "repair" the strand by incorporating the desired modification. The system presupposes,

however, that a P-element be inserted near the locus of interest. Such a fortuitous insertion adjacent to the dKLIP locus has not been reported but could be generated by existing methodology (Cooley, Kelley et al. 1988). This system is superior to homologous recombination in mice because loss of the KLIP locus in Drosophila should yield a clean result. The mouse genome, on the other hand, contains three highly related loci (furin, PC2 and PC3(PC1)) which could provide proteins with partially redundant functions thus diluting or minimizing the phenotypic effects of deleting one locus. A putative second Kex2-like locus in drosophila has been found which, if substantiated, weakens the argument for a "clean knock-out" result (W. van de Ven and M. Hoffmann, personal communication).

A powerful approach to study the roles of these newly characterized Drosophila proteases is to delete the function of the gene product by mutation of the dKLIP locus. Using mutagens (i.e. ethyl methanesulphonate or ionizing radiation) and the proper type of genetic screen one could turn up a conditional mutation in the dKLIP gene. This could be important for carrying viable lines if the complete loss of dKLIP function produces a lethal phenotype. Once a mutant has been identified a complementation test would be performed, thus ensuring that replacement of dKLIP function rescues the mutant under non-permissive conditions.

An alternative approach for the study of lethal mutations uses clonal or mosaic analysis. In this method, a mitotic recombination

event is induced using x-irradiation of embryos or larvae that are heterozygous for both the mutant dKLIP locus and a nearby marker mutation (Ashburner and Novitski 1976). The use of mutations that effect, for example bristle morphology or tissue color are common markers for localizing clones of cells. After chromosomal exchange via mitotic recombination and subsequent segregation at mitosis, a rare somatic cell that is homozygous for both the dKLIP and marker loci arises. The cell will give rise to a clone of daughter cells that will form a patch of marked tissue in the adult fly. A dKLIP mutation that causes a lethal phenotype would be used to ask what possible perturbations might be detected in cell clones (or tissues) that are homozygous for the lethal dKLIP mutation. Once mutated, one could re-introduce copies of the dKLIP locus that have specific mutations engineered into them such as deletion of dKLIP-1 specific exons or truncation of the coding frame befor the TMD. This could be used to ask what phenotype might emerge if dKLIP-1 protease function were lost or perturbed. In addition, the use of second site suppressor mutants (i.e. mutations mapping to loci other than dKLIP) that complement dKLIP loss-of-function mutants could begin to reveal what proteins interact with dKLIP.

A potential applied area that the work documented here may contribute to is that of agricultural pest control. The secretory pathway of insect cells is poorly understood and assumed to function as it does in mammalian cell which are studied more often. More complete

knowledge regarding the processes involved with insect precursor protein maturation could be valuable for the design of rational methods used to perturb polypeptide synthesis and bioactivity. Specific interruption of the biosynthetic pathway of proteins derived from larger precursors which are responsible for development or reproduction might be a viable way to prevent crop damage by specific insect pests.

Summary

In the Results chapter, cloning and characterization of cDNAs for two related Drosophila Kex2-like endoproteases were presented, allowing polypeptide sequence comparisons with other members of the serine protease family to be made. The Drosophila proteins were most closely related to human furin, PC2 and PC3(PC1) rather than to the yeast Kex2 or Bacillus subtilisin proteases. Co-expression of one of these proteins (dKLIP-1) in mammalian cells with a model precursor protein (mouse pre-pro-β-nerve growth factor) demonstrated endoproteolytic activity at a paired basic residue cleavage site within the precursor releasing a mature peptide. Spatial and temporal studies of dKLIP transcripts showed high levels of dKLIP expression in adult nervous and reproductive tissues. Embryos showed dKLIP expression in nervous system and gut tissues. Finally, a dKLIP "knockout" paradigm using antisense transcripts expressed ectopically in transgenic animals was presented. In summation, the work in this volume documents the establishment of a genetic experimental system for future studies regarding the role(s) of the Kex2-like serine protease family in metazoans.

References

Adelman, J. P., A. J. Mason, J. S. Hayflick and P. H. Seeburg. (1986). Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and prolactin release-inhibiting factor in human and rat." Proc. Natl. Acad. Sci. USA. **83**: 179-183.

Angerer, L. M. and R. C. Angerer. (1981). Detection of poly-A RNA in sea urchin eggs and embryos by quantitative in situ hybridization." Nucleic Acids Res. **9**: 2819-2840.

Ashburner, M. and E. Novitski. (1976). The genetics and biology of *Drosophila*. London, Academic Press.

Barr, P. J., O. B. Mason, K. E. Landsberg, P. A. Wong, M. C. Kiefer and A. J. Brake. (1991). "cDNA and gene structure for a human subtilisin-like protease with cleavage specificity for paired basic amino acid residues." DNA Cell Biol. **10**: 319-328.

Bathurst, I. C., S. O. Brennan, R. W. Carrell, L. S. Cousens, A. J. Brake and P. J. Barr. (1987). Yeast KEX2 protease has the properties of a human proalbumin converting enzyme." Science (Wash. DC). **235**: 348-350.

Benjannet, S., N. Rondeau, R. Day, M. Chretien and N. G. Seidah. (1991). PC1 and PC2 are proprotein convertases capable of cleaving proopiomelanocortin at distinct pairs of basic residues." Proc. Natl. Acad. Sci. USA. **88**: 3564-3568.

Bentley, A. K., D. J. G. Ree, C. Rizza and G. G. Brownlee. (1986). Defective propeptide processing of blood clotting factor IX caused by mutation of arginine to glutamine at position -4." Cell. **45**: 343-348.

Bole, D. G., L. M. Hendershot and J. F. Kearney. (1986). Posttranslational association of immunoglobulin heavy chain binding protein with nascent heavy chains in nonsecreting and secreting hybridomas." J. Cell. Biol. **102**: 1558-1566.

Bradbury, A. F., M. D. A. Finnie and D. G. Smyth. (1982). "Mechanism of C-terminal amide formation by pituitary enzymes." Nature. **298**: 686-688.

Brenner, C. and R. S. Fuller. (1991). Structural and enzymatic characterization of a purified prohormone processing enzyme: secreted, soluble Kex2 protease." Proc. Natl. Acad. Sci. USA. in press.

Brenner, S. (1974). The genetics of *Caenorhabditis elegans*." Genetics. **77**: 71-94.

Bresnahan, P. A., R. Leduc, L. Thomas, J. Thorner, H. L. Gibson, A. J. Brake, P. J. Barr and G. Thomas. (1990). Human *fur* gene encodes a yeast KEX2-like endoprotease that cleaves pro-β-NGF *in vivo*." J. Cell Biol. **111**: 2851-2859.

Bryant, P. J. (1988). Localized cell death caused by mutations in a *Drosophila* gene coding for a transforming growth factor-b homolog." Dev. Biol. **128**: 386-395.

Cathala, G., J.-F. Savouret, B. Mendez, B. L. West, M. Karin, J. A. Martial and J. D. Baxter. (1983). A method for isolation of intact, translationally active ribonucleic acid." DNA. **2**: 329-335.

Cavener, D. (1987). Comparison of the consensus sequence flanking the translational start sites in Drosophila and vertebrates." Nucleic Acids Res. **15**: 1353-1361.

Ch'ng, J. L. C., R. C. Mulligan, P. Schimmel and E. W. Holmes. (1989). Antisense RNA complementary to 3' coding and noncoding sequences of creatine kinase is a potent inhibitor of translation *in vivo*." Proc. Natl. Acad. Sci. USA. **86**: 10006-10010.

Chasan, R. and K. V. Anderson. (1989). The role of easter, an apparent serine protease, in organizing the dorsal-ventral pattern of Drosophila embryo." Cell. **56**: 391-400.

Chen, C., J. S. Bonifacino, L. C. Yuan and R. D. Klausner. (1988). Selective degradation of T cell antigen receptor chains retained in a pre-golgi compartment." J. Cell Biol. **107**: 2149-2161.

Chen, E. Y. and P. H. Seeburg. (1985). Supercoil sequencing: a fast and simple method for sequencing plasmid DNA." DNA. **4**: 165-170.

Christie, D. L., D. C. Batchelor and D. J. Palmer. (1991). Identification of kex2-related proteases in chromaffin granules by partial amino acid sequence analysis." J. Biol. Chem. **266**: 15679-15683.

Chung, K. N., P. Walter, G. W. Aponte and H.-P. Moore. (1989). Molecular sorting in the secretory pathway." Science (Wash. DC). **243**: 192-197.

Cooley, L., R. Kelley and A. Spradling. (1988). "Insertional mutagenesis of the Drosophila genome with single P elements." Science (Wash. D.C.). **239**: 1121-1128.

Cox, G. N., J. S. Laufer, M. Kusch and R. S. Edgar. (1980). Genetic and phenotypic characterization of roller mutants of *Caenorhabditis elegans*." Genetics. **95**: 317-339.

Darling, T. L. J., P. E. Petrides, P. Beguin Frey, P., E. M. Shooter, M. Selby and W. J. Rutter. (1983). The biosynthesis and processing of proteins in the mouse 7S nerve growth factor complex." Cold Spring Harbor Symp. Quant. Biol. **48**: 427-434.

Davidson, H. W. and J. C. Hutton. (1987). The insulin-secretory-granule carboxypeptidase H. Purification and demonstration of involvement in proinsulin processing." Biochem. J. **245**: 575-582.

Davidson, H. W., M. Peshavaria and J. C. Hutton. (1987). Proteolytic conversion of proinsulin to insulin." Biochem. J. **246**: 279-286.

Davidson, H. W., C. J. Rhodes and J. C. Hutton. (1988). Intraorganellar calcium and *p*H control proinsulin cleavage in the pancreatic b cell via two distinct site-specific endopeptidases." Nature. **333**: 93-96.

DeLotto, R. and P. Spierer. (1986). A gene required for the specification of dorsal-ventral pattern in Drosophila appears to encode a serine protease." Nature. **323**: 688-692.

Derynck, R., J. A. Jarrett, E. Y. Chen, D. H. Eaton, J. R. Bell, R. K. Assoian, A. B. Roberts, M. B. Sporn and D. V. Goeddel. (1985). Human transforming growth factor-ß complementary DNA sequence and expression in normal and transformed cells." Nature. **316**: 701-705.

Devereux, J., P. Haeberli and O. Smithies. (1984). A comprehensive set of sequence analysis programs for the VAX." Nucleic Acids Res. **12**: 387-395.

Dieckmann, C. L. and Tzagoloff. (1985). Assembly of the mitochondrial membrane system." J. Biol. Chem. **260**: 1513-1520.

Doolittle, W. F. and C. Yanofsky. (1968). Mutants of *Escherichia coli* with an altered tryptophanyl-transfer ribonucleic acid synthetase." J. Bacteriol. **95**: 1283-1294.

Eipper, B. A. and R. E. Mains. (1980). Structure and biosynthesis of pro-adrenocorticotropin/endorphin and related peptides." Endocr. Rev. **1**: 1-27.

Ellinwood, W. E., O. K. Ronnekleiv, M. J. Kelly and J. A. Resko. (1985). A new antiserum with conformational specificity for LHRH: usefulness for radioimmunoassay and immunocytochemistry." Peptides. **6**: 45-52.

Farquhar, M. G. (1985). Progress in unraveling pathways of Golgi traffic." Ann. Rev. Cell Biol. **1**: 447-488.

Feinberg, A. P. and B. Vogelstein. (1983). A technique for radiolabeling DNA restriction fragments to high specific activity." Anal. Biochem. **132**: 6-12.

Fitting, T. and D. Kabat. (1982). Evidence for a glycoprotein "signal" involved in transport between subcellular organelles." J. Biol. Chem. **257**: 14011-14017.

Fletcher, D. J., B. D. Noe, G. E. Bauer and J. P. Quigley. (1980). Characterization of the conversion of a somatostatin precursor to somatostatin by islet secretory granules." Diabetes. **29**: 593-599.

Fletcher, D. J., J. P. Quigley, G. E. Bauer and B. D. Noe. (1981). Characterization of pro-insulin and pro-glucagon-converting activities in isolated islet secretory granules." J. Cell Biol. **90**: 312-322.

Foster, D. C., R. D. Holly, C. A. Sprecher, K. M. Walker and A. A. Kumar. (1991). Endoproteolytic processing of the human protein C precursor by the yeast Kex2 endopeptidase co-expressed in mammalian cells." Biochem. **30**: 367-372.

Freed, E. O., D. J. Myers and R. Ridder. (1989). Mutational analysis of the cleavage sequence of the human immunodeficiency virus type 1 envelope glycoprotein precursor gp160." J. Virol. **63**: 4670-4675.

Fuller, R. S., A. Brake and J. Thorner. (1986). The *Saccharomyces cerevisiae* Kex2 gene, required for processing prepro-alpha-factor, encodes a calcium-dependent endopeptidase that cleaves after Lys-Arg and Arg-Arg sequences. Microbiology. Washington, DC, American Society for Microbiology.

Fuller, R. S., A. Brake and J. Thorner. (1989a). Intracellular targeting and structural conservation of a prohormone-processing endoprotease." Science (Wash. DC). **246**: 482-485.

Fuller, R. S., A. Brake and J. Thorner. (1989b). Yeast prohormone processing enzyme (Kex2 gene product) is a Ca^{2+}-dependent serine protease." Proc. Natl. Acad. Sci. USA. **86**: 1434-1438.

Fuller, R. S., C. Brenner, P. Gluschankof and C. A. Wilcox. (1991). The yeast prohormone-processing Kex2 protease, an enzyme with specificity for paired basic residues. Advances in Life Sciences. Berlin, Birkhauser Verlag.

Fuller, R. S., R. E. Sterne and J. Thorner. (1988). Enzymes required for yeast prohormone processing." Ann. Rev. Physiol. **50**: 345-362.

Gentry, L. E., M. N. Lioubin, A. F. Purchio and H. Marquardt. (1988). Molecular events in the processing of recombinant type 1 pre-pro-transforming growth factor beta to the mature polypeptide." Mol. Cell. Biol. **8**(10): 4162-4168.

Gething, M.-J. and J. Sambrook. (1992). "Protein folding in the cell." Nature. **355**: 33-45.

Gloor, G. B., N. A. Nassif, D. M. Johnson-Schlitz, C. R. Preston and W. R. Engels. (1991). "Targeted gene replacement in *Drosophila* via P element-induced gap repair." Science (Wash. D.C.). **253**: 1110-1117.

Grabowski, D. T., J. P. Carney and M. R. Kelley. (1991). A Drosophila gene containing the opa repetitive element is exclusively expressed in adult male abdomens." Nuc. Acids Res. **19**: 1709.

Greenwood, F. C., W. M. Hunter and J. S. Glover. (1963). The preparation of ^{125}I-labelled human growth hormone of high specific activity." Biochem. J. **89**: 114-123.

Gribskov, M. and R. R. Burgess. (1986). Sigma factors from *E. coli, B. subtilis*, phage SPO1, and T4 are homologous proteins." Nucleic Acids Res. **14**: 6745-6763.

Gumbiner, B. and R. B. Kelly. (1982). Two distinct intracellular pathways transport secretory and membrane glycoproteins to the surface of pituitary tumor cells." Cell. **28**: 51-59.

Hafen, E. and M. Levine. (1986). The localization of RNAs in Drosophila tissue sections by in situ hybridization. Drosophila: a practical approach. Oxford, IRL Press.

Harlow, E. and D. Lane. (1988). Antibodies: a laboratory manual. Cold Spring Harbor, CSH Press.

Hartmann, E., T. A. Rapoport and H. F. Lodish. (1989). Predicting the orientation of eucaryotic membrane- spanning proteins." Proc. Natl. Acad. Sci. USA. **86**: 5786-5790.

Hayflick, J. S., W. J. Wolfgang, M. A. Forte and G. Thomas. (1992). A unique Kex2-like endoprotease from *Drosophila melanogaster* is expressed in the central nervous system during early embryogenesis." J. Neurosci. in press:

Hruby, D. E., G. Thomas, E. Herbert and C. A. Franke. (1986). Use of vaccinia virus as a neuropeptide expression vector." Methods Enzymol. **124**: 295-309.

Huber, A., D. P. Smith, C. S. Zuker and R. Paulsen. (1990). Opsin of *Calliphora* peripheral photoreceptors R1-6." J. Biol. Chem. **265**: 17906-17910.

Huttner, W. B. (1987). Protein tyrosine sulfation." Trends Biochem. Sci. **12**: 361-363.

Ikemura, H. and M. Inouye. (1988). *In vitro* processing of pro-subtilisin produced in *Escherichia coli*." J. Biol. Chem. **263**: 12959-12963.

Ingham, P. W. (1988). The molecular genetics of embryonic pattern formation in Drosophila." Nature. **335**: 25-34.

Kawaoka, Y. and R. G. Webster. (1988). Sequence requirements for cleavage activation of influenza virus hemagglutinin expressed in mammalian cells." Proc. Natl. Acad. Sci. USA. **85**: 324-328.

Kazim, A. L. and M. Z. Atassi. (1982). Structurally inherent antigenic sites." Biochem. J. **203**: 201-208.

Kelly, M. J., J. Garrett, M. A. Bosch, C. E. Roselli, J. Douglass, J. P. Adelman and O. K. Ronnekleiv. (1989). Effects of ovariectomy on GnRH mRNA, pro-GnRH and GnRH levels in the preoptic hypothalamus of the female rat." Neuroendocrinology. **49**: 88-97.

Kingston, R. E. (1988). Analysis of RNA structure and synthesis. Current Protocols in Molecular Biology. New York, Greene Publishing and Wiley-Interscience.

Kleid, D. G., D. Yansura, B. Small, D. Dowbenko, D. M. Moore, M. J. Grubman, P. D. McKercher, D. O. Morgan, B. H. Robertson and H. L. Bachrach. (1981). Cloned viral protein vaccine for foot-and-mouth disease: responses in cattle and swine." Science (Wash. DC). **214**: 1125-1129.

Klemenz, R., U. Weber and W. J. Gehring. (1987). The white gene as a marker in a new P-element vector for gene transfer in *Drosophila*." Nucleic Acids Res. **15**: 3947-3959.

Knipple, D. C. and P. Marsella-Herrick. (1988). Versatile plasmid vectors for the construction, analysis and heat-inducible expression of hybrid genes in eucaryotic cells." Nucleic Acids Res. **16**: 7748.

Kozutsumi, Y., M. Segal, K. Normington, M.-J. Gething and J. Sambrook. (1988). Thr presence of malfolded proteins in the endoplasmic reticulum signals the induction of glucose regulated proteins." Nature. **332**: 462-464.

Kreig, P. A. and D. A. Melton. (1987). In vitro RNA synthesis with SP6 RNA polymerase." Methods Enzymol. **155**: 397-409.

Kyte, J. and R. F. Doolittle. (1982). A simple method for displaying the hydropathic character of a protein." J. Mol. Biol. **157**: 105-132.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature. **227**: 680-685.

Lara, H. E., J. K. McDonald and S. R. Ojeda. (1990). Involvement of nerve growth factor in female sexual development." Endocrinology. **126**: 364-375.

Leduc, R., S. Molloy, B. A. Thorne and G. Thomas. (submitted). "Autoproteolytic activation of the human furin precursor." J. Biol. Chem. :

Lippincott-Schwartz, J., J. S. Bonifacino, L. C. Yuan and R. D. Klausner. (1988). Degradation from the endoplasmic reticulum: disposing of newly synthesized proteins." Cell. **54**: 209-220.

Loh, Y. P. and H. Gainer. (1982). Characterization of pro-opiocortin converting activity in purified secretory granules from rat pituitary neurointermediate lobe." Proc. Natl. Acad. Sci. USA. **79**: 108-112.

Lyons, R. M., L. E. Gentry, A. F. Purchio and H. L. Moses. (1990). Mechanism of activation of latent recombinant transforming growth factor *b*1 by plasmin." J. Cell Biol. **110**: 1361-1367.

Mackett, M., G. L. Smith and B. Moss. (1984). General method for production and selection of infectious vaccinia virus recombinants expressing foreign genes." J. Virol. **49**: 857-864.

Mackin, R. B., B. D. Noe and J. Spiess. (1991). Identification of a somatostatin-14-generating propeptide converting enzyme as a member of the kex2/furin/PC family." Endocrinology. **129**(4): 2263-2265.

Maniatis, T., E. F. Fritsch and J. Sambrook. (1982). Molecular cloning: a laboratory manual. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory.

Mason, A. J., J. S. Hayflick, N. Ling, F. Esch, N. Ueno, S.-Y. Ying, R. Guillemin, H. Niall and P. H. Seeburg. (1985). Complimentary DNA sequences of ovarian follicular fluid inhibin show precursor structure and homology with transforming growth factor-ß." Nature. **318**: 659-663.

Matsudaira, P. (1987). Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes." J. Biol. Chem. **262**: 10035-10038.

Matsuuchi, L. and R. B. Kelly. (1991). Constitutive and basal secretion from the endocrine cell line, AtT-20." J. Cell Biol. **112**: 843-852.

McCune, J. M., L. B. Rabin, M. B. Feinberg, M. Lieberman, J. C. Kosek, G. R. Reyes and I. R. Weissman. (1988). Endoproteolytic cleavage of gp160 is required for the activation of human immunodeficiency virus." Cell. **53**: 55-67.

Mitchison, T. J. and J. Sedat. (1983). Localization of antigenic determinants in whole Drosophila embryos." Dev. Biol. **99**: 261-264.

Mizuno, K., T. Nakamura, T. Ohshima, S. Tanaka and H. Matsuo. (1988). Yeast Kex2 gene encodes an endopeptidase homologous to subtilisin-like serine proteases." Biochem. Biophys. Res. Commun. **156**: 246-254.

Mizuno, K., T. Nakamura, K. Takada, S. Sakakibara and H. Matsuo. (1987). A membrane-bound, calcium-dependent protease in yeast a-cell cleaving on the carboxyl side of paired basic residues." Biochem. Biophys. Res. Commun. **144**: 807-814.

Molloy, S. S., P. A. Bresnahan, S. H. Leppla, K. R. Klimpel and G. Thomas. "Human furin is a calcium-dependent serine endoprotease that recognizes the sequence Arg-X-X-Arg and efficiently cleaves anthrax toxin protective antigen." Proc. Natl. Acad. Sci. USA (submitted).

Monard, D. (1988). "Cell-derived proteases and protease inhibitors as regulators of neurite outgrowth." Trends Neurosci. **11**: 541-544.

Moore, H.-P. H., M. D. Walker, F. Lee and R. B. Kelly. (1983). Expressing a human proinsulin cDNA in a mouse ACTH-secreting cell: intracellular storage, proteolytic processing, and secretion on stimulation." Cell. **35**: 531-538.

Munroe, S. H. (1988). Antisense RNA inhibits splicing of pre-RNA *in vitro*." EMBO J. **7**: 2523-2532.

Neurath, H. (1984). Evolution of proteolytic enzymes." Science (Wash, DC). **224**: 350-357.

Nishida, Y., M. Hata, Y. Nishizuka, W. J. Rutter and Y. Ebina. (1986). Cloning of a Drosophila cDNA encoding a polypeptide similar to the human insulin receptor precursor." Biochem. Biophys. Res. Commun. **141**: 474-481.

Orci, L., M. Ravazzola, M. Amherdt, A. Perrelet, S. K. Powell, D. L. Quinn and H.-P. H. Moore. (1987a). The *trans*-most cisternae of the Golgi complex: A compartment for sorting of secretory and plasma membrane proteins." Cell. **51**: 1039-1051.

Orci, L., M. Ravazzola, M. Storch, R. G. W. Anderson, J. Vassalli and A. Perrelet. (1987b). Proteolytic maturation of insulin is a post-Golgi event which occurs in acidifying clathrin-coated secretory vesicles." Cell. **49**: 865-868.

Padgett, R. W., R. D. St. Johnston and W. M. Gelbart. (1987). A transcript from a Drosophila pattern gene predicts a protein homologous to the transforming growth factor-ß family." Nature. **325**: 81-84.

Panganiban, G. E. F., K. E. Rashka, M. D. Neitzel and M. Hoffmann. (1990). Biochemical characterization of the Drosophila dpp protein, a member of the transforming growth factor-ß family of growth factors." Mol. Cell. Biol. **10**: 2669-2677.

Pardue, M. L. (1986). *In situ* hybridization to polytene chromosomes. Drosophila: A Practical Approach. Oxford, IRL Press. p.111.

Park, E. C. and H. R. Horvitz. (1986). Mutations with dominant effects on the behavior and morphology of the nematode *Caenorhabditis elegans.*" Genetics. **113**: 821-852.

Payne, G. S. and R. Schekman. (1989). Clathrin: a role in the intracellular retention of a Golgi membrane protein." Science (Wash. DC). **245**: 1358-1365.

Pearse, B. M. F. and M. S. Robinson. (1990). Clathrin, adaptors, and sorting." Ann. Rev. Cell. Biol. **6**: 151-171.

Pepinsky, R. B., L. K. Sinclair, E. P. Chow, R. J. Mattaliano, T. F. Manganaro, P. K. Donahue and R. L. Cate. (1988). Proteolytic processing of mullerian inhibiting substance produces a transforming growth factor-b-like fragment." J. Biol. Chem. **263**: 18961-18964.

Peters, K., J. McDowall and A. M. Rose. (1991). Mutations in the *bli-4 (I)* locus of *Caenorhabditis elegans* disrupt both adult cuticle and early larval development." Genetics. **129**: 95-102.

Peters, K. and A. M. Rose. (1991). The bli-4 locus of Caenorhabditis elegans encodes developmentally essential kex2-like endoproteases related to mammalian prohormone convertases." Genes Devel. submitted

Pierschbacher, M. D. and E. Ruoslahti. (1984). Cell attachment activity of fibronectin can be duplicated by small fragments of the molecule." Nature. **309**: 30-33.

Pino-Heiss, S. and G. Schubiger. (1989). Extracellular protease production by Drosophila imaginal discs." Dev. Biol. **132**: 282-291.

Pirrotta, V. (1988). Vectors for P-mediated transformation in *Drosophila*. Vectors: A Survey of Molecular Cloning Vectors and Their Uses. Boston, Butterworth.

Posakony, L. M. (1987). The role of the DPP-C in the development of the imaginal discs in *Drosophila melanogaster*. Ph.D. Thesis, Harvard University.

Power, S. D., R. M. Adams and J. A. Wells. (1986). Secretion and autoproteolytic maturation of subtilisin." Proc. Natl. Acad. Sci. USA. **83**: 3096-3100.

Proudfoot, N. J. and G. G. Brownlee. (1976). 3' non-coding region sequences in eucaryotic messenger RNA." Nature. **263**: 211-214.

Redding, K., C. Holcomb and R. S. Fuller. (1991). Immunolocalization of Kex2 protease identifies a putative late golgi compartment in the yeast *Saccharomyces cerevisiae*." J. Cell Biol. **113**: 527-538.

Roberts, D. B. (1986). Basic Drosophila care and techniques. Drosophila: a practical approach. Oxford, IRL Press.

Robertson, H. M., C. R. Preston, R. W. Phillis, D. M. Johnson-Schlitz, W. K. Benz and W. R. Engels. (1988). A stable genomic source of P element transposase in *Drosophila melanogaster*." Genetics. **118**: 461-470.

Roebroek, A. J. M., I. G. L. Pauli, Y. Zhang and W. J. M. van de Ven. (1991). cDNA sequence of a *Drosophila melanogaster* gene, D*fur1*, encoding a protein structurally related to the subtilisin-like proprotein processing enzyme furin." FEBS Letters. **289**: 133-137.

Roebroek, A. J. M., J. A. Schalken, J. A. M. Leunissen, C. Onnekink, H. P. J. Bloemers and W. J. M. van de Ven. (1986). Evolutionary conserved close linkage of the c-fes/fps proto-oncogene and genetic sequences encoding a receptor-like protein." EMBO J. **5**: 2197-2202.

Rubin, B. S., J. C. King, R. P. Millar, P. H. Seeburg and A. Arimura. (1987). Processing of luteinizing hormone-releasing hormone precursor in rat neurons." Endocrinology. **121**: 305-309.

Rubin, G. and A. Spradling. (1982). Genetic transformation of *Drosophila* with transposable element vectors." Science (Wash. DC). **218**: 348-353.

Rubin, G. M. (1988). *Drosophila melanogaster* as an experimental organism." Science (Wash. DC). **240**: 1453-1459.

Sachanek, G., G. Kreil and M. A. Hermodson. (1978). Amino acid sequence of honeybee pre-promelittin synthesized *in vitro*." Proc. Natl. Acad. Sci. USA. **75**: 701-704.

Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis and H. A. Erlich. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase." Science (Wash. DC). **239**: 487-491.

Sakanari, J. A., C. E. Staunton, A. E. Eakin, C. S. Craik and J. H. McKerrow. (1989). Serine proteases from nematode and protozoan parasites: Isolation of sequence homologs using generic molecular probes." Proc. Natl. Acad. Sci. USA. **86**: 4863-4867.

Seeburg, P. H. and J. P. Adelman. (1984). Characterization of cDNA for precursor of human luteinizing hormone releasing hormone." Nature. **311**: 666-668.

Seidah, N. G., L. Gaspar, M. Marcinkiewicz, M. Mbikay and M. Chretien. (1990). cDNA sequence of two distinct pituitary proteins homologous to Kex2 and furin gene products: Tissue-specific mRNAs encoding candidates for pro-hormone processing proteinases." DNA. **9**: 415-424.

Seidah, N. G., M. Marcinkiewicz, S. Benjannet, L. Gaspar, G. Beaubien, M. G. Mattei, C. Lazure, M. Mbikay and M. Chretien. (1991). Cloning and primary sequence of a mouse candidate prohormone convertase PC1 homologous to PC2, furin and Kex2: Distinct chromosomal localization and messenger RNA distribution in brain and pituitary compared to PC2." Mol. Endo. **5**: 111-122.

Shaw, G. and R. Kamen. (1986). A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective messenger RNA degradation." Cell. **46**: 659-667.

Shimell, M. J., E. L. Ferguson, S. R. Childs and M. B. O'Connor. (1991). The *Drosophila* dorsal-ventral patterning gene *tolloid* is related to human bone morphogenetic protein 1." Cell. **67**: 469-481.

Simon, K., E. Perara and V. R. Lingappa. (1987). Translocation of globin fusion proteins across the endoplasmic reticulum membrane in *Xenopus laevis* oocytes." J. Cell Biol. **104**: 1165-1172.

Sly, W. S. and H. D. Fischer. (1982). The phosphomannosyl recognition system for intracellular and intercellular transport of lysosomal enzymes." J. Cell Biochem. **18**: 67-85.

Smeekens, S. P., A. S. Avruch, J. LaMendola, S. J. Chan and D. F. Steiner. (1991). Identification of a cDNA encoding a second putative prohormone convertase related to PC2 in AtT20 cells and islets of Langerhans." Proc. Natl. Acad. Sci. USA. **88**: 340-344.

Smeekens, S. P. and D. F. Steiner. (1990). Identification of a human insulinoma cDNA encoding a novel mammalian protein structurally related to the yeast processing protease Kex2." J. Biol. Chem. **265**: 2997-3000.

St.Johnston, R. D. and W. M. Gelbart. (1987). Decapentaplegic transcripts are localized along the dorsal-ventral axis of the *Drosophila* embryo." EMBO J. **6**: 2785-2791.

Steiner, D. F., D. Cunningham, L. Spigelman and B. Aten. (1967). Insulin biosynthesis: evidence for a precursor." Science (Wash. DC). **157**: 697-700.

Stoller, T. J. and D. Shields. (1989). The propeptide of preprosomatostatin mediates intracellular transport and secretion of α-globin from mammalian cells." J. Cell Biol. **108**: 1647-1655.

Tabor, S. and C. C. Richardson. (1987). DNA sequence analysis with a modified bacteriophage T7 DNA polymerase." Proc. Natl. Acad. Sci. USA. **84**: 4767-4771.

Tautz, D. and C. Pfeifle. (1989). A non-radioactive in situ hybridization method for the localization of specific RNAs in Drosophila embryos reveals translational control of the segmentation gene hunchback." Chromosoma. **98**: 81-85.

Thomas, G., E. Herbert and D. E. Hruby. (1986). Expression and cell type-specific processing of human preproenkephalin with a vaccinia recombinant." Science (Wash. DC). **232**: 1641-1643.

Thomas, G., B. A. Thorne and D. E. Hruby. (1988a). Gene transfer techniques to study neuropeptide processing." Ann. Rev. Physiol. **50**: 323-332.

Thomas, G., B. A. Thorne, L. Thomas, R. G. Allen, D. E. Hruby, R. Fuller and J. Thorner. (1988b). Yeast Kex2 endopeptidase correctly cleaves a neuroendocrine prohormone in mammalian cells." Science (Wash. DC). **241**: 226-230.

Thomas, L., R. Leduc, B. A. Thorne, S. P. Smeekens, D. F. Steiner and G. Thomas. (1991). Kex2-like endoproteases PC2 and PC3 accurately cleave a model prohormone in mammalian cells: Evidence for a common core of neuroendocrine processing enzymes." Proc. Natl. Acad. Sci. USA. **88**: 5297-5301.

Thorne, B. A., L. W. Caton and G. Thomas. (1989). Expression of mouse proopiomelanocortin in an insulinoma cell line." J. Biol. Chem. **264**: 3545-3552.

Thorne, B. A. and G. Thomas. (1990). An *in vivo* characterization of the cleavage site specificity of the insulin cell prohormone processing enzymes." J. Biol. Chem. **265**: 8436-8443.

Thorne, B. A., O. H. Viveros and G. Thomas. (1991). Expression and processing of mouse proopiomelanocortin in bovine adrenal chromaffin cells." J. Biol. Chem. **266**: 13607-13615.

van den Ouweland, A. M., H. L. P. van Duijnhoven, G. D. Keizer, L. C. J. Dorssers and W. J. M. van de Ven. (1990). Structural homology between the human fur gene product and the substilisin-like protease encoded by yeast Kex2." Nucleic Acids Res. **18**: 664.

van Hofsten, P., I. Faye, K. Kockum, J.-Y. Lee, K. G. Xanthopoulos, I. A. Boman, H. G. Boman, A. Engstrom, D. Andreu and R. B. Merrifield. (1985). Molecular cloning, cDNA sequencing, and chemical synthesis of cecropin B from *Hyalophora cecropia*." Proc. Natl. Acad. Sci. USA. **82**: 2240-2243.

von Heijne, G. (1986). The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology." EMBO J. **5**: 3021-3027.

von Heijne, G. and Y. Gavel. (1988). Topogenic signals in integral membrane proteins." Eur. J. Biochem. **174**: 671-678.

Wada, K., S. Aota, R. Tsuchiya, F. Ishibashi, T. Gojobori and T. Ikemura. (1990). Codon usage tabulated from the GenBank genetic sequence data." Nucleic Acids Res. **18**: 2367.

Weir, J. P., G. Bajszar and B. Moss. (1982). Mapping of the vaccinia virus thymidine kinase gene by marker rescue and by cell-free

translation of selected mRNA." Proc. Nat. Acad. Sci. USA. **79**: 1210-1214.

Weischaus, E. W. and C. Nusslein-Volhard. (1986). Looking at *Drosophila* embryos. Drosophila: a practical approach. Oxford, IRL Press, Ltd.

Wells, J. A., E. Ferrari, D. J. Henner, D. A. Estell and E. Y. Chen. (1983). Cloning, sequencing and secretion of *Bacillus amyloliquefaciens* subtilisin from *Bacillus subtilis*." Nucleic Acids Res. **11**: 7911-7925.

Wharton, K. A., B. Yedvobnick, V. G. Finnerty and S. Artavanis-Tsakonas. (1985). Opa: a novel family of transcribed repeats shared by the Notch locus and other developmentally regulated loci in D. melanogaster." Cell. **40**: 55-62.

Wilcox, C. A. and R. S. Fuller. (1991). Posttranslational processing of the prohormone-cleaving Kex2 protease in the *Saccharomyces cerevisiae* secretory pathway." J. Cell. Biol. **115**: 297-307.

Wise, R. J., P. J. Barr, P. A. Wong, M. C. Kiefer, A. J. Brake and R. J. Kaufman. (1990). Expression of a human proprotein processing enzyme: Correct cleavage of the von Willebrand factor precursor at a paired basic amino acid site." Proc. Natl. Acad. Sci. USA. **87**: 9378-9382.

Yanisch-Perron, C., J. Vieira and M. J. (1985). Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors." Gene. **33**: 103-121.

Yoshimasa, Y., J. Paul, J. Whittaker and D. F. Steiner. (1990). Effects of amino acid replacements within the tetrabasic cleavage site on the processing of the human insulin receptor precursor expressed in chinese hamster ovary cells." J. Biol. Chem. **265**: 17230-17237.

Yoshimasa, Y., S. Seino, J. Whittaker, T. Kakehi, A. Kosaki, H. Kuzuya, H. Imura, G. I. Bell and D. F. Steiner. (1988). Insulin-resistant diabetes due to a point mutation that prevents insulin proreceptor processing." Science (Wash. DC). **240**: 784-787.

Yoshiro, K. and R. G. Webster. (1988). Sequence requirements for cleavage activation of influenza virus hemagglutinin expressed in mammalian cells." Proc. Natl. Acad. Sci. USA. **85**: 324-328.

Yuen, L. and B. Moss. (1987). Oligonucleotide sequence signaling transcriptional termination of vaccinia virus early genes." Proc. Nat. Acad. Sci. (USA). **84**: 6417-6421.

Appendix: Sorting Signals For Proteins That Enter The Regulated Secretory Pathway: A Study Of Pro-Gonadotropin Releasing Hormone

Introduction

Many peptides released from secretory cells, such as peptide hormones are derived from larger polypeptide precursors. Pro-hormones are proteolytically processed in association with regulated secretory vesicles (RSV). In contrast, other pro-proteins, pro-growth factor precursors, for instance, are cleaved in association with constitutive secretory vesicles (CSV). Therefore the mechanism(s) for properly targeting a precursor protein to a compartment for endoproteolysis are critical for the continued maturation of the biologically active molecule. There is evidence that certain lysosomal hydrolases are targeted to the lysosome via a specific targeting signal which is recognized by a receptor (Farquhar 1985; Sly and Fischer 1982). In this case, phosphorylation of the 6-carbon of terminal mannose residues attached to these proteins is specifically bound by a mannose-6-phosphate receptor (man-6-pr) (Sly and Fischer 1982). The receptor-ligand complex is targeted to clathrin coated regions of the TGN from which lysosomal membranes form. In the acidic lysosomal environment the complex dissociates and the man-6pr is recycled to the TGN leaving behind the hydrolase. The studies described in this appendix were designed to test the hypothesis that a domain of a model prohormone

precursor (pro-gonadotropin-releasing hormone, pro-GnRH) targets the protein into RSV. A contiguous stretch of amino acid sequences could serve as a sorting sequence. To test whether this occurs, a series of fusion proteins consisting of proGnRH domains and a reporter protein were expressed. These constructs would have been used to localize a putative sorting signal when introduced into endocrine cells.

Several in-frame control constructs were made (see figure 30). These are: GG1 (signal sequence/globin), GG2 (full length prepro-hGnRH/globin) and GG3 (signal sequence/ globin/prohGnRH). The first construct (GG1) is designed to confirm the finding that globin lacks regulated secretory pathway sorting signals (Simon, Perara et al. 1987). GG2 and GG3 test the ability of proGnRH to properly target globin in a position independent fashion.

ProhGnRH was chosen for these studies for several reasons. 1) The 69 amino acid precursor is relatively small. Thus compared to other regulated secretory proteins (i.e. 211 aa mPOMC, 190 aa hGH), domain analysis for sorting will be greatly simplified. 2) ProhGnRH has relatively few paired basic amino acid cleavage sites (figure 25), making analysis of mature products less complicated. 3) Several well characterized domain specific antisera have been provided for these studies, making protein analysis simplified

The reasons for using alpha-globin as a reporter protein are: 1) addition of an N-terminal signal sequence results in its translocation across the RER membrane in AtT-20 cells (Simon, Perara et al. 1987).

The protein can therefore be targeted for the secretory pathway. 2) Moreover, translocated globin protein is then secreted (Simon, Perara et al. 1987, and V. Lingappa, personal communication). Thus globin appears to lack a sorting signal or get turned over in lysosomes. 3) The protein is relatively small (141 amino acids) when compared to other constitutively secreted proteins. This will reduce the chance of the reporter protein interfering with the sorting of proGnRH. 4) The full length coding sequence for alpha-globin is available for use (Steve Liebhaber personal communication). 5) Well characterized globin specific antisera is available (M.Z. Atassi, personal communication).

The GnRH prohormone contains a 23 amino acid signal sequence, the GnRH decapeptide, and a 56 amino acid peptide called GnRH-associated peptide (GAP) (Seeburg and Adelman 1984) (figure 25). The precursor is synthesized in the cell bodies of neurons in the medial basal hypothalamus and prior to anterograde vesicular transport the mature peptide is processed at a paired basic amino acid cleavage site (Lys-Arg) typical of other prohormones (see Thesis Introduction and figures 1 and 25). Portal blood supply transports the peptide from its site of release at axonal termini located adjacent to the median eminence to the anterior pituitary where it binds to receptors on gonadotrophs inducing follicle stimulating hormone and luteinizing hormone release. These peptide hormones then trigger the maturation and release of oocytes from ovarian follicles and secretion of estradiol for maintenance of the uterine lining until blastula implantation. GAP may inhibit prolactin

release from lactotrophs and effect gonadotropin release from gonadotrophs, however these are still in question.

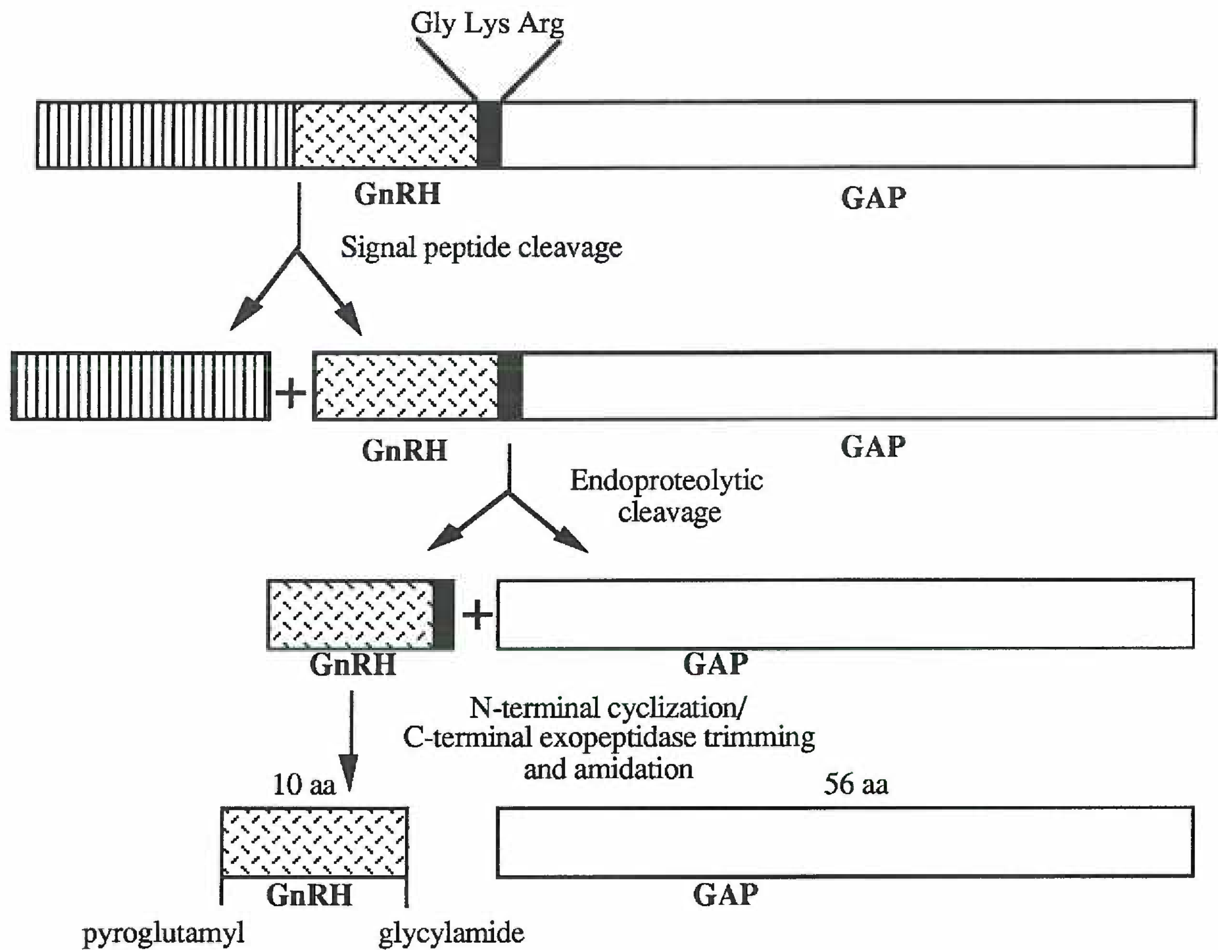


Figure 25. Biosynthetic pathway for GnRH maturation. The precursor structure was deduced from the cloned cDNA (Seeburg and Adelman 1984; Adelman, Mason et al. 1986). The signal sequence is denoted by a striped box. The endoproteolytic cleavage site is marked with a bold black line. Paired basic residue endoproteolysis and other modifications are thought to occur in secretory granules during transit along neuronal fibers and at the termini (Rubin, King et al. 1987).

Methods

Materials

Purified α-globin was generously given by Dr. Daniel Shih (OHSU). Dr. Karoly Nikolics (Genentech, Inc.) kindly provided antisera KN16 and synthetic GAP1-56 peptide. Dr. Martin Kelly (OHSU) kindly provided EL14 and ARK2 antisera and synthetic proGnRH 5-16 peptide. Dr. Steve Liebhaber (University of Pennsylvania) kindly provided the human α-globin cDNA in pSP64.

Cells

BSC-40, an African Green monkey kidney epithelial cell line, AtT20, a murine corticoptroph cell line, and Rin m5F, a rat pancreatic β insulinoma cell line were cultured as described (Thorne, Caton et al. 1989). Primary bovine adrenomedullary chromaffin cells (BAM cells) were isolated and cultured as described (Thorne, Viveros et al. 1991).

Antisera

EL14 antisera is specific for fully mature GnRH decapeptide (Ellinwood, Ronnekleiv et al. 1985). ARK2 is specific for rat pro-GnRH 6-16 spanning the cleavage site between GnRH and GAP and recognizes the human precursor as well (Kelly, Garrett et al. 1989). KN16 is specific for the midportion of human GAP (residues 20-43) (K. Nikolics, personal communication). The domains in GnRH and its precursor that interact

with these antisera are shown in figure 26. The anti-globin antiserum is specific for human hemoglobin (Kazim and Atassi 1982).

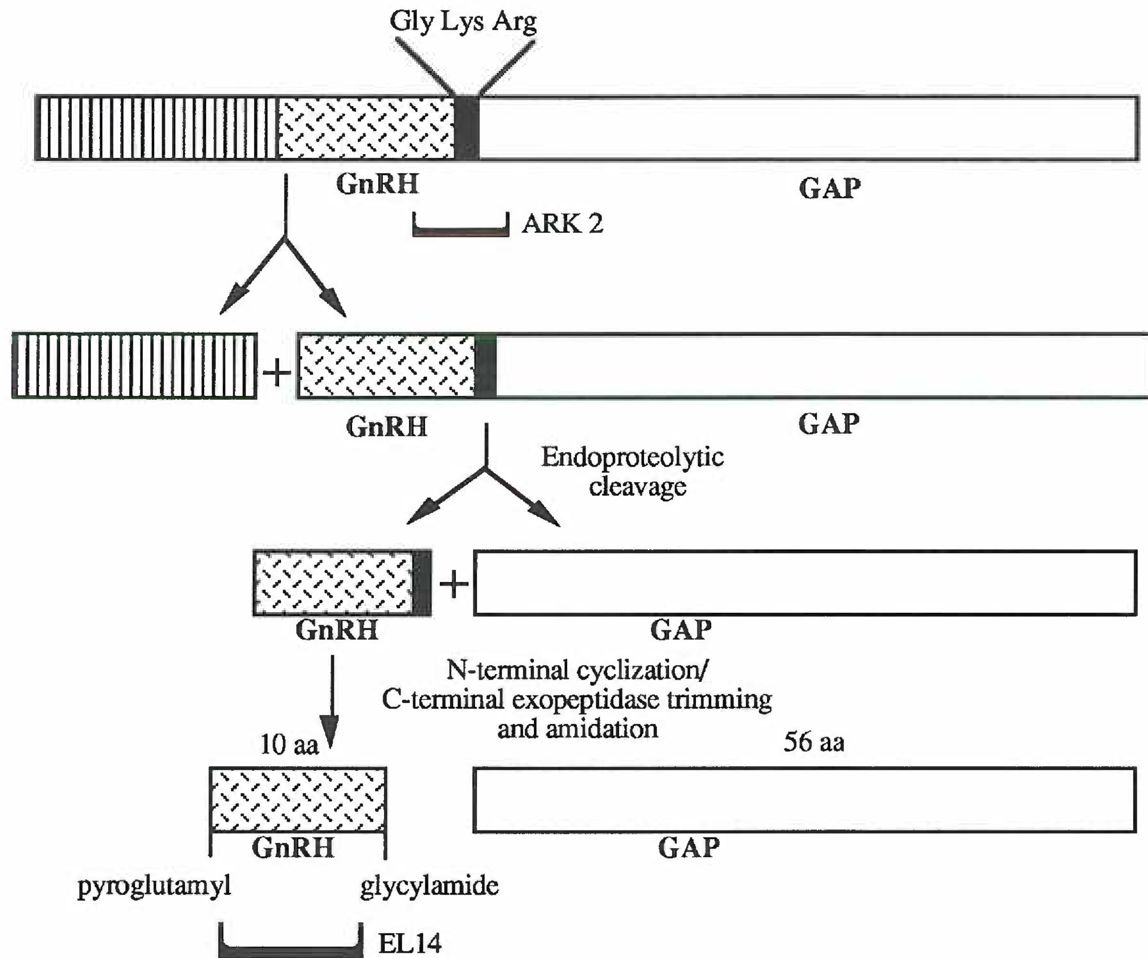


Figure 26. Specificities of antisera against GnRH and its precursor used in this work. ARK2 recognizes the endoproteolytic cleavage site (residues 6-16) and is specific for the precursor (Kelly, Garrett et al. 1989). EL14 is specific for the fully mature form of GnRH decapeptide (Ellinwood, Ronnekleiv et al. 1985). KN16 recognizes an epitope on the GAP portion of the molecule (K. Nikolics, personal communication).

Construction of VV:hGnRH and globin/pro-GnRH fusions

The full length hypothalamic cDNA for GnRH was isolated from M13mp19:HLH (Adelman, Mason et al. 1986) and ligated into pVV3 (Hruby, Thomas et al. 1986) to generate pVV3:GnRH. The plasmid contains the constitutive VV 7.5K promoter adjacent to a multiple cloning site (Yuen and Moss 1987). Flanking the promoter and cloning sites are genomic sequences from the viral thymidine kinase (tk) gene. Recombination of the plasmid sequences via these genomic flanking sequences generates a virions with a tk^- phenotype that can be selected for in the presence of 5-bromo-2'-deoxy-uridine (Mackett, Smith et al. 1984; Weir, Bajszar et al. 1982). Recombinant virus was plaque purified and grown on a large scale for production of working stocks. Stocks were stored at -70°C until use.

Each of the fusion proteins was constructed as described below. Fusion GG1 (signal/globin) was generated using a fragment containing the 7.5K promoter and most of the the signal sequence of proGnRH was isolated from pVV3:GnRH. Two complimentary oligomers were synthesized that completed the coding sequence for the GnRH signal peptide joined to the amino terminus of α-globin. A fragment containing the remainder of the coding sequence for globin was isolated from pSP64:α-globin. These three DNAs were ligated together with M13mp19 to form a GG1 intermediate construct. The junctional sequences of several resulting clones were checked by dideoxynucleotide sequencing. The entire fusion coding sequence from one of these clones, M13:GG1, was isolated and ligated to

pVV3 to generate pVV3:GG1. Fusion GG2 (pre-proGnRH/globin) was constructed using a fragment containing the entire coding sequence of pre-proGnRH was isolated from M13mp19:HLH and ligated to the full length coding sequence for α-globin isolated from pSP64:globin and M13mp18. Several clones were checked for accuracy of the junctional sequences to ensure that the reading frame was maintained through the fusion sequence. The entire GG2 fusion sequence was isolated from M13 and ligated into pVV3 to generate pVV3:GG2. Fusion GG3 (signal/globin/proGnRH) was made with a fragment containing the coding sequences for the carboxyl terminal domain of globin isolated from pSP64:globin. Two complimentary oligomers were synthesized that encoded the last amino acids of globin fused to the N-terminal coding sequences of proGnRH. A fragment containing the remaining coding sequences for proGnRH was isolated from M13mp19:HLH. A four-way ligation was set up with these DNA fragments and M13mp19 to form a GG3 intermediate construct. Several isolates were sequenced to check the junctional sequences. The coding sequences for globin back end/proGnRH were isolated from the M13:GG3 intermediate. A fragment from M13:GG1 that encoded the signal peptide fused to the front end of globin from M13:GG1 was also isolated and these two fragments were ligated together with pVV3 to form pVV3:GG3.

Radioimmunoassay Of Recombinant hGnRH And Its Precursor

ProGnRH 5-16 (containing tyrosine at position 1) and GnRH 1-10 were radiolabeled with [^{125}I] by the chloramine T method (Greenwood,

Hunter et al. 1963). Incorporated iodine was separated from unincorporated using Sep-pak C18 cartridges and a step gradient of acetonitrile. EL14 bound 25% of the labeled 1-10 trace at a dilution of 1:100,000 and a standard curve was generated using triplicate doses of 1-1000 pg of GnRH 1-10 which resulted in a sensitivity of 1 pg and an ED_{50} of 9 pg. ARK2 bound 25% of the labeled pro-GnRH 5-16 trace at a dilution of 1:2500 and a standard curve was generated using triplicate doses of 5-5000 pg of proGnRH 6-16 which resulted in a sensitivity of 35 pg and an ED_{50} of 140 pg. The intra-assay variance was 8% for EL14 and 10% for ARK2.

Metabolic Labeling And Immunoprecipitation

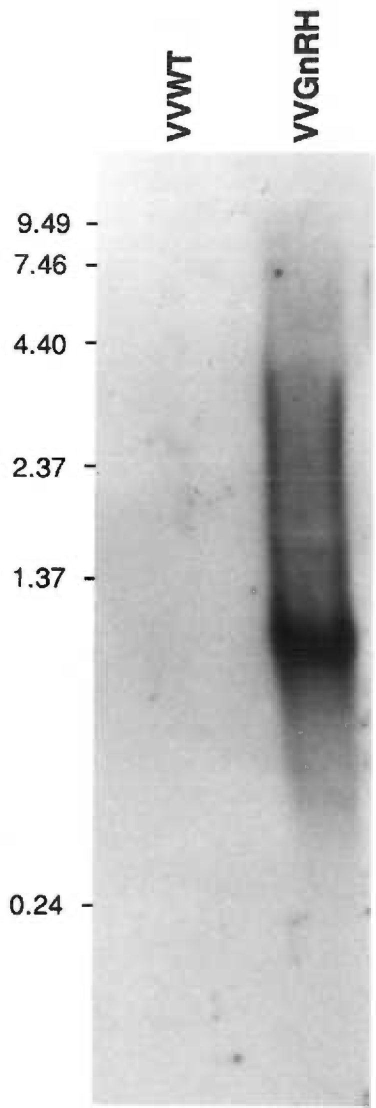
The method is described in Chapter 4.

Results

Characterization Of VV:hGnRH

To verify the proper insertion of the hGnRH cDNA into the virus and to determine that VV:hGnRH would drive transcription of hGnRH mRNA *in vivo*, a Northern blot of mRNA from VV:wt and VV:hGnRH infected BSC-40 cells was probed with hGnRH cDNA. A band of about 950 bases is seen only in the VV:hGnRH lane (figure 27). It is the expected size because initiation occurred at the 7.5K promoter cap site and the transcripts terminated at the tk polyadenylation site (Yuen and Moss 1987).

Figure 27. Northern blot detection of hGnRH transcripts. Parallel 10 cm plates of confluent BSC-40 cells were infected with VV:wt or. VV:GnRH at 5 pfu/cell. After 20 hrs incubation at 37°C, cytoplasmic RNA was isolated as described (Cathala, Savouret et al. 1983) and fractionated on a denaturing 1.5% agarose gel. After transferring to nitrocellulose, the RNA was hybridized with a randomly primed [^{32}P]-labeled hGnRH cDNA probe. After washing, the blot was exposed to x-ray film. The RNA molecules of the size standard are in kilobase pairs.


VVWT
VVGnRH
9.49
7.46
4.40
2.37
1.37
0.24

Expression And Secretion Of Recombinant hGnRH

To verify that the virus directs the expression of immunologically reactive hGnRH prohormone and mature GnRH decapeptide several cell lines were infected and assayed by RIA. Initially, BSC-40 cells were mock infected or infected with VV:wt or VV:hGnRH at a multiplicity of infection of 5 pfu/cell. Cell extracts were assayed for EL14 and ARK2 immunoreactivity (IR). VV:hGnRH infected samples were found to contain exclusively prohormone IR (2.4 pmoles/10^6 cells after 20 hrs., figure 28). Thus, consistent with results obtained from expression of mPOMC (Thorne, Caton et al. 1989) and hPE (Thomas, Herbert et al. 1986), BSC-40 cells are processing deficient for proGnRH.

Next, AtT-20 cells were mock infected or infected with VV:wt or VV:hGnRH. Cell extracts were assayed by radioimmunoassay (RIA) for EL14 and ARK2 IR. VV:GnRH cell extracts contained significant amounts of EL14 IR (5 pmoles/10^6 cells in 20 hrs.) and ARK2 IR (3 pmoles/10^6 cells in 20 hrs.) (figure 28). Thus AtT-20 cells are processing and modification proficient (i.e. proteolytic cleavage, N-terminal cyclization, C-terminal residue trimming and amidation) for proGnRH.

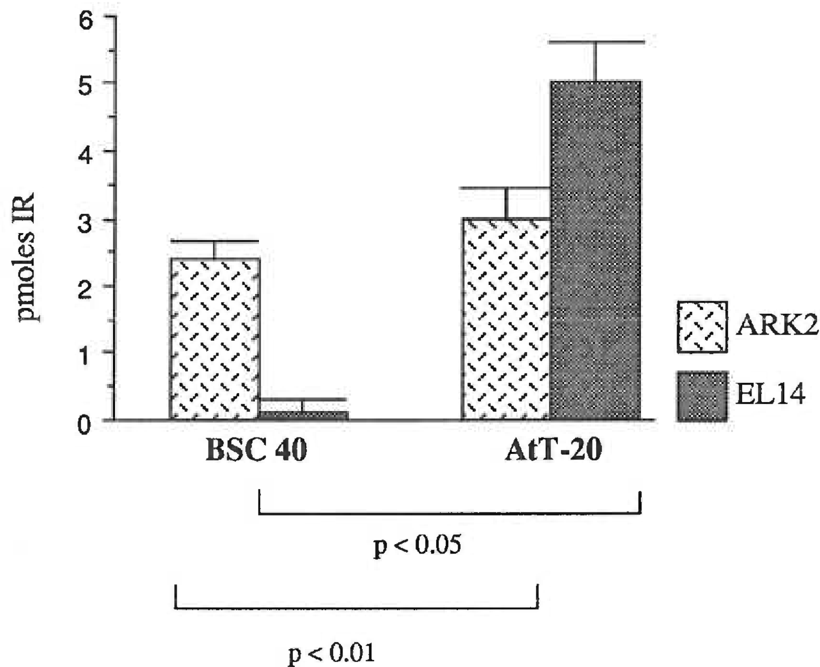


Figure 28. Processing of the GnRH precursor in two cell types. Parallel 10 cm plates of BSC-40 or AtT20 cell were infected with either VV:wt, VV:GnRH at 5 pfu/cell or mock infected. After 20 hrs incubation at 37°C, the cells were harvested by scraping into a solution of 1N HCl, 1% trifluoroacetic acid (TFA), 5% formic acid, and 1% NaCl (w/v) (M. Seger, personal communication). Following brief sonication on ice and centrifugation in a microfuge, the supernatants (cell extract) were stored frozen until the time of assay. Triplicate aliquots of cell extracts were lyophilized and assayed for EL14 and ARK 2 IR by RIA. Mock and VV:wt infected cells contained no measurable immunoreactivity. The data from three separate experiments (n=3) was analyzed using the Student t test.

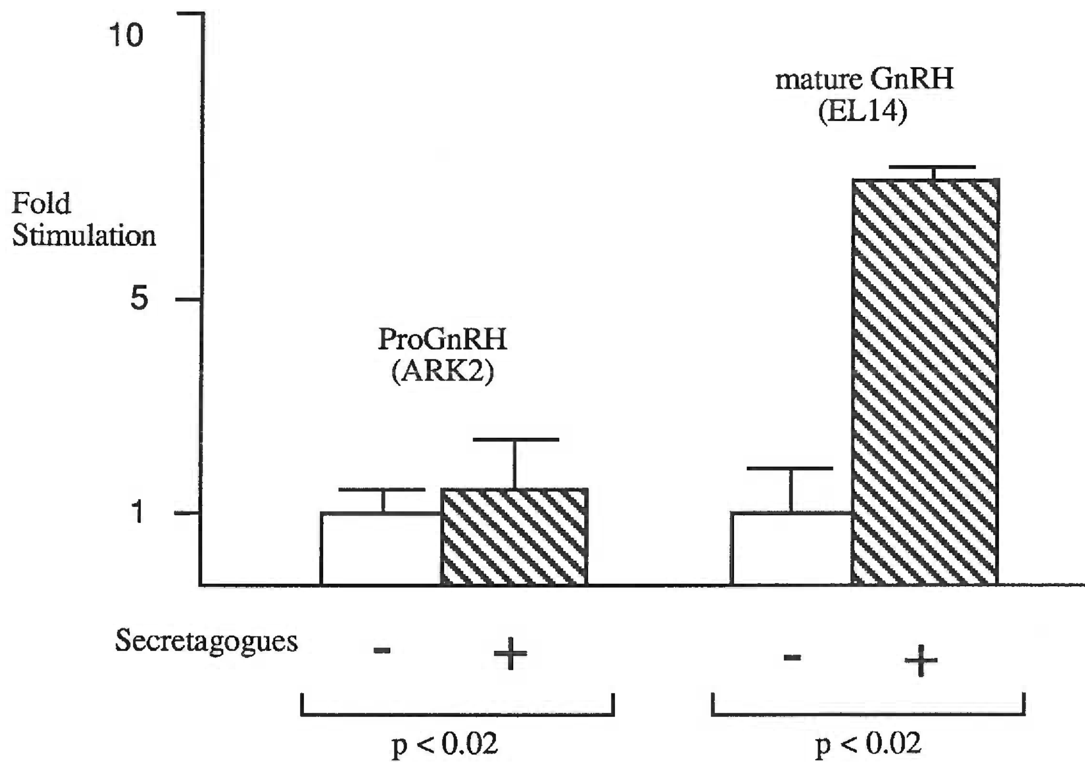


Figure 29. Regulated release study of GnRH expressed from AtT-20 cells. Parallel plates of cells were infected with VV:GnRH at 5 pfu/cell. After 16 hrs incubation, the medium was removed and replaced with fresh medium with (striped bars) or without secretagogues (0.1 mM forskolin, 0.1 mM TPA). After 1 hr incubation the media were removed and assayed for proGnRH and mature GnRH IR. The data from four experiments (n=4) was analyzed using the Student t test.

Insulin and POMC are processed in the regulated secretory pathway of AtT-20 cells and Rin m5F cells (Gumbiner and Kelly 1982) (Moore, Walker et al. 1983) (Thorne, Caton et al. 1989). To establish that hGnRH prohormone is processed in a functionally identical pathway within secretory cells, AtT-20 cells were studied (figure 29). These cells were mock infected or infected with VV:wt or VV:hGnRH for 16 hours. Media from parallel wells of infected cells incubated in the presence or absence of the secretagogues forskolin (50 mM) and 12-O-tetradecanoyl-phorbol-13-acetate (TPA, 100nM) was assayed for EL14 or ARK2 IR. VV:hGnRH infected AtT-20 cells release about 7 fold more mature decapeptide IR into the medium when secretagogues are present than when absent (figure 29). Importantly, no significant stimulated release of the precursor was detected. Medium from secretagogue stimulated mock and VV:wt infected cells had no detectable level of IR using EL14 or ARK2.

Primary bovine adrenal chromaffin (BAM) cells were also assessed for their ability to properly process and target proGnRH and the decapeptide. Twenty-four hours after plating, cells were infected with either VV:wt, VV:GnRH, or mock infected. Mature GnRH IR was not detected from mock or wt infected cell extracts after a further 16 hour incubation. Conversely, VV:GnRH infected cell extracts had 11.5 (±2) pM of GnRH IR. Next, regulated release studies were performed. After 16 hrs of VV:GnRH infection, triplicate wells were treated for 30 minutes

with either 3 mM barium chloride (in the absence of calcium which competitively inhibits the effect of barium) or TPA/forskolin or left untreated. Barium chloride elicits the regulated release of mature peptides from recombinant vaccinia infected BAM cells in culture (Thorne, Viveros et al. 1991). Medias were harvested and assayed for mature GnRH IR using EL14 antiserum. TPA/forskolin yielded a 6 fold increase in GnRH IR over untreated cells, while barium treatment elicited a 10 fold increase in GnRH IR release. Thus barium treatment of primary BAM cells is an efficient paradigm for processing and regulated release of pro-GnRH and the decapeptide.

Expression Of Pro-GnRH/globin Fusion Proteins

The experimental results that follow describe the transient expression of a series of pre-proGnRH/globin fusion proteins (figure 30). These studies were designed to test whether a peptide hormone precursor (pro-GnRH) could target a reporter protein (α-globin) into RSV. Since globin is normally synthesized in the cytoplasm of erythrocytes it probably does not contain secretory pathway sorting information.

The expression of the chimeric proteins was demonstrated using replicate wells of BSC-40 cells infected with VV:wt or viruses carrying the coding sequences for each fusion (VV:GG1, VV:GG2 or VV:GG3). After 16 hr infection, the cells were harvested. Both western blot (data not shown) and immunoprecipitation (figure 31) results show proteins of the

expected size being recognized by the antisera. Fusion GG1 (signal/globin) is 14 kD (the size of native α-globin which lacks a signal peptide) while GG2 (prepro-GnRH/globin) and GG3 (signal/globin/pro-GnRH) are 25 kD (the combined size of 9 kD pro-GnRH and 14 kD α-globin).

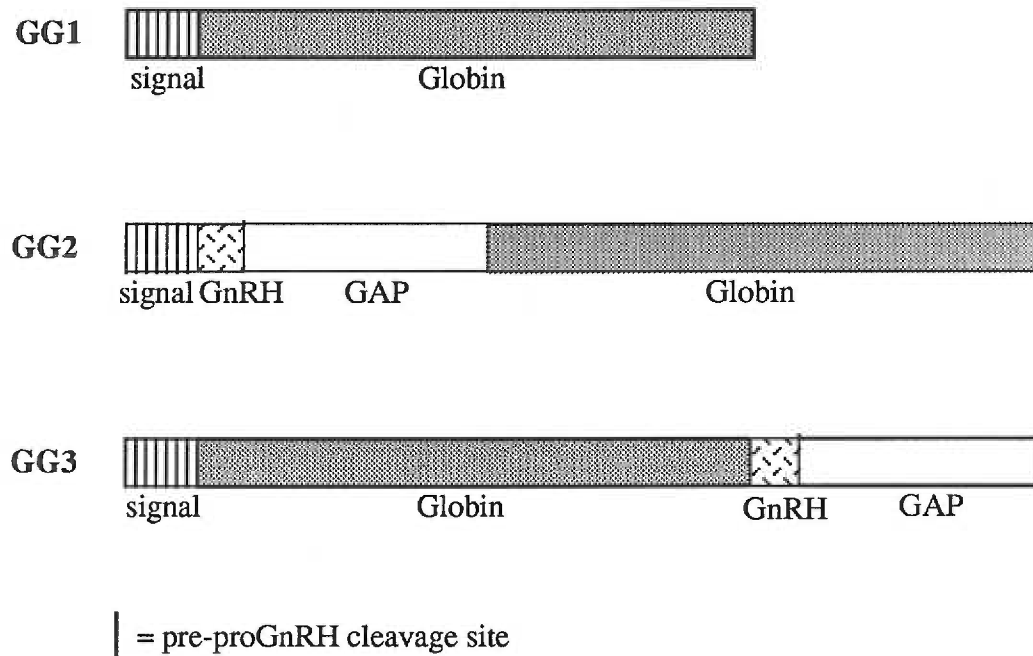


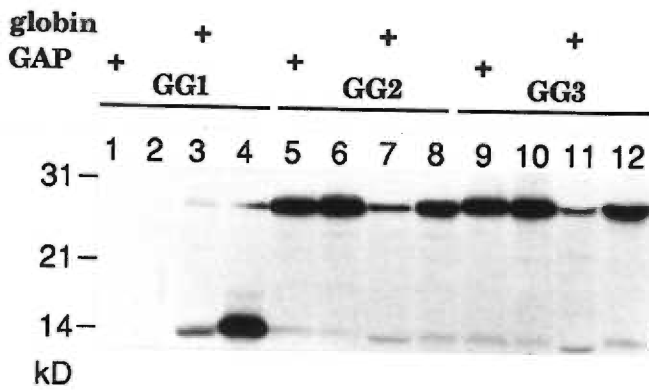
Figure 30. Schematic diagram depicting three fusion proteins that were expressed in VV infected cells. GG1 contains the signal sequence of hGnRH fused to the full length human α-globin protein. GG2 contains the full length pre-prohGnRH fused at its C-terminus to full length α-globin. GG3 contains the entire α-globin protein sequence inserted between the signal sequence and pro-hGnRH. See methods section of this appendix for details of the DNA constructions.

The crossreactivity of the antisera KN16 and anti-globin with each of the fusions was also checked by immunoprecipitation and SDS-PAGE (figure 31). As expected, GG1 was bound by anti-globin antisera and not by KN16. An excess of unlabeled α-globin significantly reduced the intensity of the labeled protein that was bound. Both GG2 and GG3 were bound by anti-globin and KN16 antisera. In addition, GG2 and GG3 were both competed away from antibody binding by an excess of unlabeled alpha-globin or synthetic GAP1-56, reducing the quantity of labeled protein recovered. Thus, each fusion protein was recognized by the appropriate antiserum and the recognition was competed by well characterized, purified synthetic peptide or natural protein.

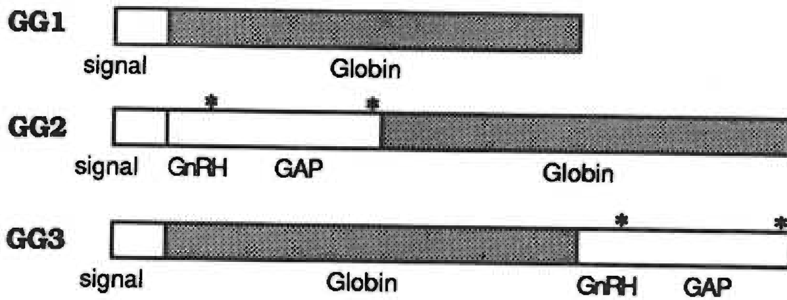
Next, the secretion of the fusion proteins was investigated. Replicate BSC-40 cells expressing the fusions were metabolically labeled for 1 hour and harvested for immunoprecipitation (data not shown). Neither GG1 nor GG2 were found in the medium of cells expressing them. Both of these proteins were found exclusively in the cell extracts. GG3 was found in cells and medium. Similar results were found when AtT20 and Rin cells transiently expressed the fusions (figures 32-34). Regulated release experiments in Rin cells show that neither GG1, GG2 nor GG3 were retained in regulated secretory vesicles (data not shown). In order to investigate the fate of these proteins, pulse-chase experiments were performed using replicate wells of Rin cells expressing each of the chimeric proteins. Fusion GG1 was found to be rapidly

turned over in the cells (t1/2=~15 min., figure 32). Fusion GG2 was predominantly turned over inside the cells at a slower rate than GG1 (t1/2=~120 min., figure 33). Fusion GG3 was predominantly secreted from cells (t1/2=~90 min., figure 34) after being modified to more than one higher molecular weight form. Similar results were obtained from pulse-labeled primary BAM cells expressing each of the fusions (data not shown). Thus a wide variety of cultured cells expressed this series of fusion proteins with very similar fates.

Figure 31. Immunoprecipitation analysis of the fusion proteins. Confluent wells of BSC-40 cells were infected with either VV:wt, VV:GG1, VV:GG2, or VV:GG3 at 5 pfu/cell. After 20 hrs incubation, the cells were methionine depleted for 30 min. and pulse-labeled with 100 μCi of [^{35}S] methionine. After 15 min. incubation the cells were harvested by scraping into ice-cold PBS, 1% Triton X-100. Cell extracts were prepared (see Methods p.38), precleared with sepharose beads and divided into 4 equal portions. Samples 1,2,5,6,9,10 received KN16 antiserum. Samples 3,4,7,8,11,12 received anti-globin serum. In addition, samples 1,5,9 received 1.5 nM/tube hGAP 1-56 and samples 3,7,11 received 1.5 nM alpha-globin. After overnight incubation, immune complexes were recovered with protein A-sepharose beads and analyzed on 12.5% SDS-PAGE. VV:wt samples incubated with either KN16 or anti-globin were devoid of signals (not shown).



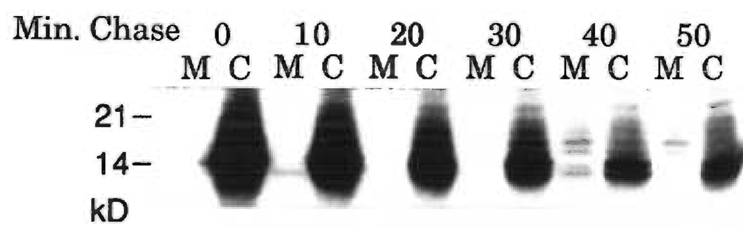
globin
GAP
GG1
GG2
GG3
1 2 3 4 5 6 7 8 9 10 11 12
31
21
14
kD



GG1
signal
Globin
GG2
signal
GnRH
GAP
Globin
GG3
signal
Globin
GnRH
GAP
* = paired basic amino acids

Figure 32. Pulse-chase study of fusion protein GG1.

Confluent replicate wells of Rin m5f cells were infected with VV:GG1 at 5 pfu/cell. After 13 hrs incubation, the cells were methionine depleted for 30 min. and pulse-labeled with 100 μCi of [^{35}S] methionine. After 15 min. incubation the cells and medium were either harvested or chased for 10, 20, 30, 40 or 50 min. in media containing a 10x concentration of L-methionine at which time cells and medium were harvested. Media (M) and cell extracts (C) were immunoprecipitated with anti-globin serum. Immune complexes were analyzed on 16% SDS-PAGE. The small amounts of higher weight globin released into the medium at 40 and 50 minutes were not reproducible and represented less than 5% of the total recovered counts. Molecular weight standards are in kilodaltons.


Min. Chase
0
10
20
30
40
50
M C M C M C M C M C M C
21–
14–
kD

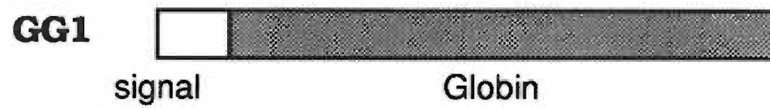
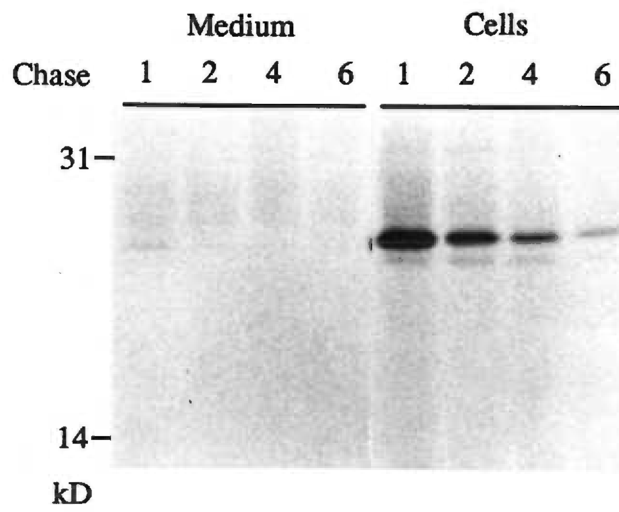

GG1
signal
Globin

Figure 33. Pulse-chase study of fusion protein GG2.

Confluent replicate wells of Rin m5F cells were infected with VV:GG2 at 5 pfu/cell. After 20 hrs incubation, the cells were methionine depleted for 30 min. and pulse-labeled with 100 μCi of [^{35}S] methionine. After 15 min. incubation the cells were chased for 1, 2, 4 or 6 hrs. in media containing a 10x concentration of L-methionine. Media was removed and cell extracts prepared for immunoprecipitation with anti-globin serum. Immune complexes were analyzed on 16% SDS-PAGE. Molecular weight standards are in kilodaltons.


Medium
Cells
Chase
1
2
4
6
1
2
4
6
31
14
kD

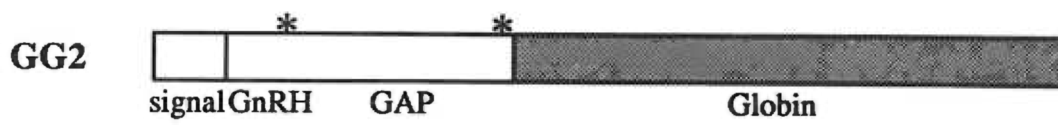
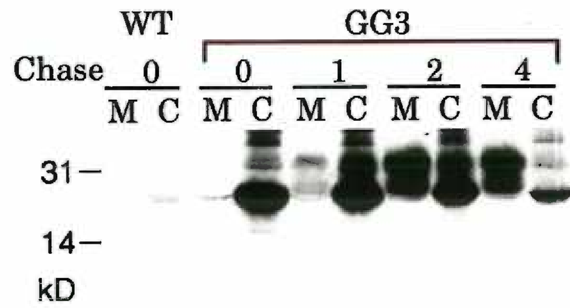

GG2
*
*
signal
GnRH
GAP
Globin

Figure 34. Pulse-chase study of fusion protein GG3.

Confluent replicate wells of Rin m5f cells were infected with VV:GG3 or VV:wt at 5 pfu/cell. After 13 hrs incubation, the cells were methionine depleted for 30 min. and pulse-labeled with 100 µCi of [^{35}S] methionine. After 15 min. incubation the cells and medium were either harvested (0 time) or chased for 1, 2 or 4 hrs. in media containing a 10x concentration of L-methionine. Media (M) was removed and cell extracts (C) prepared for immunoprecipitation with anti-globin serum. Immune complexes were analyzed on 16% SDS-PAGE. Molecular weight standards are in kilodaltons. Note GG3 is apparently modified to several higher weight forms and released into the medium.



* = paired basic amino acids

Discussion and Conclusions

This appended section describes experiments designed to test the hypothesis that a signal sorts human pro-GnRH and a fused reporter protein (human α-globin) into the regulated secretory pathway. The data from the pro-GnRH expression and secretion experiments support the following conclusions. 1) Recombinant VV derived hGnRH prohormone and decapeptide is immunologically indistinguishable from native protein. 2) GnRH prohormone is processed only by specific cell types. 3) The prohormone is processed to mature decapeptide in AtT-20 cells. 4) Processing of proGnRH occurs in the regulated secretory pathway of AtT-20 cells. In addition, they establish precedents for the use of prohGnRH as a model for the study of protein sorting into the regulated secretory pathway.

The three fusion proteins were efficiently expressed in transient expression studies, as determined by antibody reactivity using two criteria, 1) western blot (data not shown) and 2) immunoprecipitation (figure 29). Time course studies examining cells and media were used to demonstrate that two fusion proteins (GG1 and GG2) were not secreted from cells. Instead they both were degraded intracellularly over different time courses. GG1 was turned over much more rapidly (15 min.) than was GG2 (120 min.). The third fusion protein (GG3) was modified with an unidentified moiety to a higher relative molecular mass and secreted constitutively from cells. Since none of the proteins is sorted to the regulated pathway, this

experimental paradigm will be difficult to apply toward determining whether a sorting signal exists in pro-GnRH for fusion protein sorting and targeting.

Studies using analogous constructs stably expressed in GH3 cells (a rat anterior pituitary somatolactotroph line) found similar results. When chimpanzee α-globin was fused to the C-terminus of the somatostatin (SRIF) precursor (Stoller and Shields 1989) forming Pro-Glo (akin to GG2), the fusion was predominantly (70%) degraded. The remaining 30% was either secreted constitutively in an unprocessed form (18%) or processed at an internal paired basic residue site and stored for regulated release (12%). When a fusion consisting of the β-lactamase signal peptide fused to α-globin (Sig-Glo) was similarly expressed in GH3 cells it was translocated into the ER lumen, cleaved correctly by signal peptidase and rapidly degraded in a chloroquine insensitive compartment (i.e., non-lysosomal). The results from the SRIF/globin fusions expressed in GH3 cells are in good agreement with those presented in this appendix.

In disagreement, however, is the finding that Xenopus laevis oocytes are capable of processing and secreting α-globin when injected with mRNA encoding the β-lactamase/α-globin fusion (Sig-Glo) referred to above ((Simon, Perara et al. 1987), V. Lingappa, UCSF, personal communication). Thus an identical fusion protein is tolerated in an amphibian oocyte while in mammalian cells it is rapidly degraded. It is more likely that this phenomenon is cell type specific rather than

species specific. Those cells that rapidly degraded the signal/globin fusion (GG1 or Sig/Glo) are specialized, terminally differentiated secretory cells (i.e., endocrine or epithelial), while the X. laevis oocyte is an undifferentiated secretory cell.

The GnRH-globin series of chimeric proteins described above may be useful in determining what mechanism(s) is involved in secretory pathway "proofreading" and degradation and its biological significance. There is evidence for a quality control mechanism on protein synthesis in the ER or early Golgi compartment (Bole, Hendershot et al. 1986; Chen, Bonifacino et al. 1988; Kozutsumi, Segal et al. 1988; Lippincott-Schwartz, Bonifacino et al. 1988). Nascent proteins synthesized by membrane bound ribosomes are retained in the lumen of the ER until properly configured at which time they are released for transit to the Golgi stacks. The ER retention mechanism involves the immunoglobulin heavy-chain binding protein (Bip). It may be possible to co-precipitate Bip along with the fusions by using less stringent washing conditions thus demonstrating that it is a bound component. Sequential C-terminal deletions of GG2 and 3 may show what sequences are involved in the retention and degradation of the proteins. However, the involvement of higher order structure will be difficult to investigate.