FUNCTIONAL ANALYSIS OF THE GENE bli-4 IN CAENORHABDITIS ELEGANS

by

ALANA KRISTINE BERGSTROME JONES

B.Sc., Simon Fraser University, 1994

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Department of Medical Genetics)

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

March 1997

© Alana Kristine Bergstrome Jones, 1997
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Medical Genetics

The University of British Columbia
Vancouver, Canada

Date April 19, 1997
ABSTRACT

Many biologically active compounds are first formed as inactive precursors. Their activation results from proteolytic cleavage, largely accomplished by a group of enzymes known as KEX2/subtilisin-like proprotein convertases, or kexins. This family of enzymes has been described in a number of species, from prokaryotes (B. subtilis), to unicellular eukaryotes (S. cerevisiae), as well as multicellular organisms (Drosophila, mouse, human). Recently, an additional member of this convertase family was identified in the nematode Caenorhabditis elegans (Peters et al, 1991). The gene encoding the proprotein protease (bli-4) produces four distinct protein isoforms, all of which show sequence similarity to members of the kexin family.

The purpose of the series of experiments reported here was twofold. First, to determine whether there was a conservation of function between the four isoenzymes produced by the gene bli-4 and kexin members from S. cerevisiae and humans. In order to achieve this goal, I have tested two putative homologues of bli-4, KEX2 and hfur, for functional rescue of the viable blistering mutation in bli-4. The second goal of this research was to analyze the control of bli-4 expression. To address this question, deletion analysis of the 5' flanking DNA was performed, searching for critical regions that may function in the control of timing and tissue specificity of bli-4.

The approach taken to address the question of functional conservation of the four isozymes of bli-4 and other members of the kexin family took advantage of the ability to genetically transform C. elegans through germline injection. A genomic fragment containing the yeast gene KEX2 (provided by R. Fuller, Stanford, CA) and a cDNA clone of the human gene hfur (provided by G. Thomas, Portland, OR) were cloned into an ectopic expression vector carrying a heat-shock
promoter (provided by A. Fire, Carnegie). The heat shock promoter (P. Candido, University of British Columbia) expresses in many of the same tissues as bli-4. A cDNA clone of one of the BLI-4 isoforms was used as a positive control. These constructs were injected into the gonad of bli-4(e937) homozygous hermaphrodites. Progeny of injected worms were heat-shocked to induce expression of the transgenes, and then screened for phenotypic rescue.

In addition to this approach, functional rescue of S. cerevisiae KEX2 deletion mutants using one bli-4 isoform was tested. The isoform chosen has a predicted structure similar to the native yeast kex2p. However, this construct, when expressed in yeast, did not rescue KEX2 mutants, even though the control experiment using a KEX2 clone did rescue the mutants.

Analysis of the 5' flanking region of bli-4 was performed by creating serial deletions in a bli-4/lacZ fusion construct. These vectors were then injected into wild-type worms and transgenics stained with Xgal to determine location of expression. This analysis revealed that there are a possible five signals controlling both tissue specificity and timing of bli-4 expression.

The results of this series of experiments show 1) that the functional role of bli-4 is conserved with yeast KEX2 and human hfur, and 2) that the control of expression of bli-4 can be at least partly explained by sequences in the 5' flanking DNA region. These conclusions emphasize the importance of C. elegans as a model organism for the study of the kexin family of convertases.
Table of Contents

Abstract ii

Table of Contents iv

List of Tables vii

List of Figures viii

Acknowledgements x

Introduction 1

The genetics of bli-4 1
Molecular structure of bli-4 4
The kexin family of convertases 7
Regulation of expression of bli-4 10
C. elegans as an experimental system for kexin regulation 12

Materials and Methods 13

Nematode culture conditions and strains 13
Preparing male stocks 15
Agarose gel electrophoresis 15
Polymerase chain reaction amplification of DNA samples 16
Cloning PCR products 17
Restriction enzyme digestions 18
Ligations 18
Gel purification of DNA samples 19
Minipreparations of plasmid DNA 20
Preparation of DNA for injection 20
Injections 21
Confirmation of the presence of injected plasmids in transgenic worms 22
Integrating transgenic arrays 22
Staining transgenic worms harbouring lacZ constructs 25
Results

Section I.  
*S. cerevisiae* KEX2 mutant rescue analysis

A. KEX2-deficient yeast are rescued by transformation with endogenous KEX2  
B. KEX2-deficient yeast is not rescued by Blisterase D  
C. Expression of Blisterase D in transgenic yeast

Section II.  
*C. elegans* transformation rescue with KEX2 and *hfur*

A. Blisterase A cDNA rescues blistering in *e937* homozygotes  
B. KEX2 rescues blistering in *e937* homozygotes  
C. *hfur* rescues blistering in *e937* homozygotes

Section III.  
Analysis of the 5’ end of *bli-4*

A. Deletion of various regions upstream of *bli-4* differentially affect gene expression
   1. pCeh288  
   2. pCeh289  
   3. pCeh291  
   4. Control constructs

B. Sequence analysis of the subregions in the 5’ end of *bli-4*
Discussion

Evolutionary conservation of function between C. elegans, yeast and human kexin family members

Transgenic rescue of S. cerevisiae mutants

Transformation rescue of C. elegans mutants

Regulation of expression of bli-4

Conclusions

References

Appendix I

Appendix II

Appendix III

Appendix IV
# List of Tables

1. Alleles of *bli-4*  
2. Intragenic complementation patterns of *bli-4* alleles  
3. Abbreviations used in this thesis  
4. Primers used to test transgenic strains  
5. Summary of *bli-4(e937)* transgenic rescue results  
6. Summary of deletion staining patterns  
7. Summary of unique *C. elegans* cosmid homologies in pCeh289  
8. Summary of BLAST search results  
9. Promoter predictions resulting from NNPP searches
List of Figures

1. Mutant phenotypes of bli-4 3
2. Molecular structure of bli-4 6
3. Kexin family members 9
4. Transgenic integration screen 24
5. Construction of pCeh228 (full-length bli-4 D cDNA) 30
6(a). Construction of pCeh295: cloning S. cerevisiae KEX2 into the hsp 16-2 expression vector 33
6(b). Construction of pCeh296: cloning human hfur into the hsp 16-2 expression vector 34
6(c). Construction of pCeh297: cloning bli-4 A cDNA into the hsp 16-2 expression vector 35
7. Preparation of deletion constructs 37
8. Summary of deletion constructs 38
9. Experimental scheme for the rescue of KEX2-deficient S. cerevisiae 42
10. Rescue of KEX2-deficient S. cerevisiae with Blisterase D 43
11. Summary of deletion constructs used in this thesis 48
12(a). pCeh289 transgenic staining pattern in eggs 51
12(b). pCeh289 transgenic staining pattern in larvae 52
13(a). pCeh291 transgenic staining pattern in eggs 53
13(b). pCeh291 transgenic staining pattern in larvae 54
14. KR3324/KR3325 Xgal staining pattern 55
15. Schematic representation of TESS and NNPP analyses results 59
16. Alignment of carboxy-terminal sequences from KEX2 and bli-4 68
Acknowledgements

Many people were critical in my pursuit of this degree. I would like to thank my supervisor, Dr. Ann Rose; thanks to assorted lab members, particularly Colin Thacker, for instruction, guidance and for keeping me on the right track; thanks to Dr. Charlie Boone and Kelly Blundell for direction and assistance; thank you to Diana Janke for injections and transgenic strains; thanks to my family for believing that I could do this; and mostly thanks to my husband David, for everything you are and for everything you do. You know I couldn’t have done this without you.
The Genetics of bli-4

The Caenorhabditis elegans gene bli-4 was originally identified in 1974 by Sydney Brenner, the pioneer of C. elegans research. In the process of screening for mutant phenotypes, he discovered a recessive, loss-of-function mutation (e937), that when homozygous gave rise to fluid-filled separations (or blisters) of the adult cuticle (Figure 1). This blistering phenotype resulting from the e937 mutation is only 80-90% penetrant in an isogenic population.

Subsequent to the isolation of the e937 allele, mutant screens have resulted in twelve additional alleles of bli-4 (Table 1) (Rose and Baillie, 1980; Howell et al, 1987; Peters et al, 1991; Thacker et al, 1995), all of which produce lethality when homozygous, with arrest in early development (Figure 1). Complementation studies of bli-4 have revealed a complex pattern of
<table>
<thead>
<tr>
<th>Allele name</th>
<th>Homozygous phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>e937</td>
<td>85% blistered</td>
</tr>
<tr>
<td>h42</td>
<td>arrested 3-fold embryos</td>
</tr>
<tr>
<td>h199</td>
<td>arrested 3-fold embryos</td>
</tr>
<tr>
<td>h254</td>
<td>arrested 3-fold embryos</td>
</tr>
<tr>
<td>h384</td>
<td>arrested 3-fold embryos</td>
</tr>
<tr>
<td>h427</td>
<td>arrested 3-fold embryos</td>
</tr>
<tr>
<td>h520</td>
<td>arrested 3-fold embryos</td>
</tr>
<tr>
<td>h670</td>
<td>arrested 3-fold embryos</td>
</tr>
<tr>
<td>h791</td>
<td>arrested 3-fold embryos</td>
</tr>
<tr>
<td>h1010</td>
<td>arrested 3-fold embryos</td>
</tr>
<tr>
<td>h1403</td>
<td>arrested 3-fold embryos</td>
</tr>
<tr>
<td>q508</td>
<td>arrested 3-fold embryos</td>
</tr>
<tr>
<td>s90</td>
<td>embryonic/L1 lethal</td>
</tr>
</tbody>
</table>
Figure 1

**Mutant Phenotypes of bli-4**

Nomarski photomicrographs of: a) Class I blistered phenotype of e937 homozygote; b) Class II q508 homozygote arresting in late embryogenesis; c) Class III s90 homozygote arresting as an L1 larva (70% of Class III homozygotes arrest at late embryogenesis, as in b)).
interactions, and have established three complementation groups (summarized in Table 2) among the alleles of bli-4 (Peters, 1991). The original mutation, e937, defines the first group, Class I. Eleven non-complementing lethal alleles, all of which arrest during the three-fold stage of embryogenesis, define the second group, Class II. These alleles, when heteroallelic with e937, result in the blistered phenotype. Penetrance of the trait generally (with the exception of h199) increases to 100%. The twelfth lethal allele (s90), which represents Class III, has a terminal phenotype of arrest in embryogenesis, or at the L1 larval stage of development, when homozygous. When s90 is in heteroallelic combination with any of the eleven Class II lethal alleles, worms arrest in early embryogenesis. However, when s90 is heteroallelic to e937, the resulting phenotype is wild-type. This is a classic case of intragenic complementation, and defines bli-4 as a complex gene.

Molecular Structure of bli-4

Cloning of the bli-4 locus was initiated by Ken Peters (1992), thus revealing the complex nature of the gene (Figure 2). The gene, bli-4, is composed of at least twenty-one exons, spanning approximately thirteen kilobases of DNA. Isolation of cDNAs indicated that there are at least four distinct protein isoforms produced by bli-4. These four isoforms all share a common region at the 5' end of the gene, but as a result of alternative splicing at the 3' end they all differ at their carboxy termini. Molecular characterization of the alleles of bli-4 has shown that the viable blister-causing mutation (e937) is the result of a deletion spanning 3.5 kb, which removes a single exon unique to one of the four isoforms. Further studies (Thacker et al, 1995) showed that the transcript containing this single
Table 2

Intragenic complementation patterns of bli-4 alleles

<table>
<thead>
<tr>
<th></th>
<th>Class I(^a)</th>
<th>Class II(^b)</th>
<th>Class III(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class I(^a)</td>
<td>Blistered (85%)(^d)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class II(^b)</td>
<td>Blistered (100%)(^d)</td>
<td>Arrest 3-fold</td>
<td></td>
</tr>
<tr>
<td>Class III(^c)</td>
<td>Wild-type</td>
<td>Arrest 3-fold</td>
<td>Arrest 3-fold/L1 larvae</td>
</tr>
</tbody>
</table>

The complementation pattern for the twelve lethal alleles and the viable allele, \(e937\), is shown. All lethal alleles (both Class II and Class III) arrest development in late embryogenesis; however, approximately 30% of \(s90\) (Class III) homozygotes survive past hatching but arrest as L1 larvae (Peters 1992; Thacker et al 1995).

\(^a\)Class I represented by \(e937\)
\(^b\)Class II represented by \(h1010\) and \(q508\)
\(^c\)Class III represented by \(s90\)
\(^d\)Percentage of blistered animals shown in parentheses.
The molecular structure of bli-4 is shown, as well as the alignment of the four cDNA clones representing the transcripts of bli-4. In addition, the causative mutations for e937 as well as the Class II alleles h1010 and q508 are shown. e937 is a 3.5 kb deletion removing exon 13. h1010 is a Tc1 transposon insertion mutant mapped to exon 9 of the common region. q508 is a small deletion that removes the 5' end of exon 12 (modified from Thacker et al, 1995). The dotted region encodes the protease domain. Shading in the 3' exons reflects the use of each in the four distinct Blisterase products.
exon is absent in $e937$ homozygotes, suggesting that it is the loss of this one isoform that is responsible for the blistered phenotype. Mapping of the causative mutations in the lethal alleles has placed all of the Class II alleles within the first twelve exons of the common region (Thacker et al, 1995; Thacker, unpub. results). These results suggest that at least one of the $bli-4$ functions is essential for $C. elegans$ viability.

Sequence analysis of the four products of $bli-4$ shows a high degree of structural similarity to the KEX2/subtilisin-like family of proprotein convertases (Thacker et al 1995). Furthermore, the first twelve exons, shared by all transcripts, have significant sequence similarity to the catalytic domain of serine endoproteases. This family of enzymes has been shown to be of importance in the enzymatic processing of many biologically important precursor proteins. Based on the mutant phenotypes of $bli-4$, it is reasonable to expect that the products of this gene are involved in the proteolytic activation of proteins essential for the early development of the worm and for the production or maintenance of the adult cuticle.

The Kexin Family of Convertases

An important requirement in all biological systems is the regulation of gene function. One level of regulation is protein activation. In many cases, proteins are first produced within the cell as inactive precursors, and do not become active until the product has been appropriately localized (in space or time). One mechanism that has evolved to control the activity of a number of gene products involves the activation of precursor molecules through proteolytic cleavage by another molecule. Examples of this mechanism were first suggested with the
discovery that pituitary hormones (Chretien and Li, 1967) and insulin (Steiner et al, 1967) are synthesized first as inactive precursors. Activation has been shown to be accomplished by a group of enzymes collectively known as the KEX2/subtilisin-like family of proprotein convertases, now often referred to as kexins (Figure 3). This family of enzymes recognizes a minimal amino acid sequence of two basic amino acids (lysine and/or arginine), and cleaves the substrate molecule at the carboxy end of this dibasic pair. For most members of the family, the recognition sequence can be expanded around this essential pair of basic residues, but in all cases tested, each can effectively cleave any substrate molecule containing only the most rudimentary recognition sequence. Sequence analysis of the four products of bli-4 has shown that the first twelve exons, shared between all four isoforms, encode a catalytic domain with significant sequence similarity to this family of proprotein convertases. The prototypic member for this family of enzymes is kex2p in the budding yeast Saccharomyces cerevisiae. This serine endoprotease is a calcium-dependent, membrane-bound molecule, localized to the late Golgi apparatus, that processes the inactive precursors of the pheromone α-mating factor and the M1 killer toxin (Fuller et al, 1989a). Mammalian members of the kexin family have been discovered in mouse and human. These include PCSK1, also known as PC3 (Seidah et al, 1991; Smeekens et al, 1991), PCSK2 (Seidah et al, 1990; Smeekens and Steiner, 1990), PCSK4 (Nakayama et al, 1992; Seidah et al, 1992), PCSK5, also known as PC6 (Lusson et al, 1993; Nakagawa et al, 1993), PACE4 (Keifer et al, 1991), PCSK7 (Seidah et al, 1996) and furin (Roebroek et al, 1986; Fuller et al, 1989b). In Drosophila, at least two members of the kexin family have been defined, Furl1 and Fur2 (Roebroek et al, 1992), both of which show sequence similarity to the human member furin. Members of the family have been classified into two groups based on the secretory pathway in which they function. Some members
Members of the kexin family of proprotein convertases

Shown are a subset of the members of the kexin family of convertase enzymes, including the prototype member kex2p from *S. cerevisiae*, furin from human, the Drosophila furin homologues Fur1 and Fur2, mammalian PCSKs and PACE4, as well as the predicted structures of the four Blisterase products (the vertical line represents the boundary of sequence shared by all four Blisterases). Shaded regions represent the protease domain, with percentage identity to the Blisterase products. Cross-hatched regions are Cys-rich. Diagonal regions are transmembrane domains. Dotted regions are secretion signal sequences (modified from Thacker *et al*, 1995).
exhibit restricted expression patterns in vivo, and participate in the regulated secretory pathway. For example, PCSK1/PC3 and PCSK2 are expressed only in endocrine and neuroendocrine tissues (Seidah et al, 1990, 1991; Smeekens and Steiner, 1990) and PCSK4 is restricted to the testes (Nakayama et al, 1992; Seidah et al, 1992). Conversely, some members are expressed in a large number of tissues and participate in the constitutive secretory pathway (Roebroek et al, 1986; van den Ouweland et al, 1990; Van de Ven et al, 1990; Bresnahan et al, 1990; Keifer et al, 1991). In particular, furin seems to have a ubiquitous role, able to functionally activate precursor substrates in most tissue types (Molloy et al, 1992). For furin, catalytic activity may be controlled by intracellular compartmentalization. Furin is a membrane-bound enzyme, localized to the trans-Golgi network (TGN) (Molloy et al, 1994). Like all members of the kexin family, furin is first synthesized as an inactive precursor (Leduc et al, 1992; reviewed in Seidah and Chretien, 1992). It becomes activated upon release from the endoplasmic reticulum and during transportation to the TGN. Substrate targets are activated by furin while being cycled via clathrin-coated vesicles from the TGN to the cell surface. Signals that may be important in this cycling are encoded in the cytoplasmic end of furin (Molloy et al, 1994; Schafer et al, 1995).

Regulation of Expression of bli-4

An additional level of control in the proteolytic activation processes is the regulation of expression of the convertases themselves. Substrate specificity is largely dependent on co-localization of both substrate and enzyme molecules within a cell type. Little is known about the tissue specific control of any of the kexin convertases, and about bli-4 in particular. Previous analyses of regulation
of expression (Aamodt et al, 1991; Okkema et al, 1993; Egan et al, 1995) in *C. elegans* have made use of reporter vector constructs, particularly *lacZ* and, more recently, green fluorescent protein (GFP) (Chalfie et al, 1994). Fragments of DNA from the 5' end of the gene of interest are fused in frame with the reporter gene and the construct injected into the gonad of N2 hermaphrodites. Deletions within this 5' fragment may also be injected in order to delineate subregions that may contain critical sequences which drive expression in a temporal and tissue specific manner. Several factors important in transcriptional, post-transcriptional, translational and post-translational control have been identified (reviewed by Krause, 1995). Most *C. elegans* genes studied thus far have proven to be transcribed by RNA Polymerase II, one of three polymerase molecules that are active in transcriptional processes. A TATA consensus binding sequence is usually found approximately 30 bp upstream of the transcriptional start site of genes transcribed by RNA Polymerase II (Krause, 1995). Specific promoter sequences have also been identified for genes with particular tissue expression. For endodermal gene regulation, a GATA sequence is generally required (Krause, 1995; Egan et al, 1995). Control of gene expression at the post-transcriptional level is also being studied. A number of *lin* genes (cell lineage control) have been shown to produce RNA transcripts that are complementary to the mRNA of the genes they regulate, indicating RNA-RNA binding interactions that are responsible for controlling gene expression (Lee et al, 1993). In addition, 3' end processing is important for the regulation of expression. Polyadenylation in *C. elegans* usually begins about 13 nucleotides downstream of a consensus sequence (AAUAAA) (Krause, 1995). The 3' poly-A signal of the gene *unc-54* has been shown to provide stability to the mature mRNA (Mello and Fire, 1995), and is thus often used in artificial gene expression constructs.
C. elegans as an Experimental System for Kexin Regulation

Understanding the biological function of these proprotein convertases can be greatly facilitated by studying genetic mutations. Experimental manipulation of the gene and observation of its function can be done in vivo with C. elegans. In this system, regulation of activation, which may be controlled by localization and expression of the convertases, can be studied. To date, only yeast and C. elegans kexin mutants have been reported. In the case of bli-4, not only is a wide set of mutations available for analysis, but the four gene products of bli-4 each resemble convertase members from both the regulated secretory pathway and the constitutive secretory pathway (Thacker et al, 1995). Clearly, this gene represents a powerful tool for understanding the function of kexin convertases in vivo. In addition, the C. elegans system as a model organism is easily exploited with its amenable genetics and molecular biology. Thus, bli-4 should provide much information about the enzymatic processing of precursor molecules not only in C. elegans, but also in a number of other organisms, including humans.

The analysis performed in this study is aimed at revealing the signals that control not only the tissue specificity of expression of bli-4, but also the timing. This investigation can be carried further by examining the control of expression of each of the four isoforms produced by bli-4. bli-4 represents an important tool for understanding the regulation of kexin convertases, both at the level of gene expression and tissue regulation. It will also serve to provide information on the evolutionary importance of these levels of regulation within the kexin family of convertase enzymes.
MATERIALS AND METHODS

Nematode Culture Conditions and Strains

All *C. elegans* strains used during the course of this study were maintained on 5 cm petri plates containing nematode growth media (NGM) streaked with *Escherichia coli* OP50 at 20°C, unless otherwise indicated (Brenner, 1974).

The standard wild-type nematode strain (+/+)) is *C. elegans*, var. Bristol, strain N2. Mutations in *C. elegans* are assigned both a gene name describing the genetic locus, and an allele designation representing the mutational event which gave rise to the associated phenotype. One of the genetic mutants used in this study, *e937*, gives a blistered phenotype (worms with blistered cuticles), and is one allele of the gene *bli-4*; thus, this mutation is written as *bli-4(e937)*. Generally, the nomenclature guidelines (Horvitz et al, 1979) require that gene names are italicized while phenotypes are not. When referring to gene products, capital letters are used; for example, the products of *bli-4* are called BLI-4 isoforms. Nomenclature of strains and materials used in this thesis is outlined in Table 3.
Table 3.

Nomenclature abbreviations used in this thesis.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>bli</td>
<td>mutations in these genes give rise to a blistered cuticle</td>
</tr>
<tr>
<td>Bli</td>
<td>the Blistered phenotype; fluid-filled separations of the cuticle</td>
</tr>
<tr>
<td>dpy</td>
<td>mutations in these genes result in a dumpy phenotype</td>
</tr>
<tr>
<td>Dpy</td>
<td>the Dumpy phenotype; short, fat body morphology</td>
</tr>
<tr>
<td>rol</td>
<td>mutations in these loci result in a rolling phenotype</td>
</tr>
<tr>
<td>Rol</td>
<td>the Roller phenotype; helical twisting of the body around the longitudinal axis, resulting in a rolling motion as the worms moves</td>
</tr>
<tr>
<td>h</td>
<td>the Rose laboratory allele designation. All alleles, extrachromosomal arrays, chromosomal rearrangements and DNA constructs designed in this laboratory are issued an h number</td>
</tr>
<tr>
<td>KR</td>
<td>the Rose laboratory strain designation. All <em>C. elegans</em> strains isolated in the Rose laboratory are issued a KR number</td>
</tr>
<tr>
<td>KRp</td>
<td>all oligonucleotides, or primers, designed in the Rose laboratory for the purpose of Polymerase Chain Reaction are issued a KRp number</td>
</tr>
<tr>
<td>pCeh</td>
<td>DNA constructs are identified as plasmids subcloned from <em>Caenorhabditis elegans</em> in the Rose (h) laboratory</td>
</tr>
<tr>
<td>In</td>
<td>designation for extrachromosomal DNA that has been integrated into the nematode genome</td>
</tr>
<tr>
<td>hEx</td>
<td>the exogenous DNA construct present within a nematode strain transformed via microinjection in the Rose (h) laboratory</td>
</tr>
</tbody>
</table>
Preparing male stocks

Although the primary sexual form of *C. elegans* is hermaphroditic, males do arise by non-disjunction of the X chromosome in natural populations with a frequency of approximately 1 in 500 (Hodgkin *et al.* 1979). To maintain populations with a high frequency of males, mating plates are established with between ten and twenty males and one or two hermaphrodites in order to encourage mating. Since hermaphrodites give rise to populations that are essentially isogenic, male stocks allow the transfer of genetic information between strains.

Agarose gel electrophoresis

Products of PCR amplification and restriction enzyme digestion were visualized by agarose gel electrophoresis. Aliquots of 1 µl were diluted in 9 µl distilled water, with 1x loading buffer (Sambrook *et al.* 1989). Each sample was then loaded into wells of an agarose gel (between 0.5% and 1.5% w/v in 0.5x TBE buffer or 1x TAE buffer), containing roughly 0.1 µg/ml ethidium bromide. The samples were electrophoresed through the gel at approximately 5V/cm for at least one hour, or until the DNA had run far enough for the individual bands to be distinguished. When excision of bands was required for purification, the above process was performed with low-melting-point agarose.
Polymerase Chain Reaction Amplification of DNA Samples

Idaho Technologies Thermocycler 1605

Amplification of DNA samples using the Idaho Technologies Thermocycler was performed in 25 µl volumes in sealed silicon glass capillary tubes. In each case, 5 µl of the appropriate DNA sample was combined with the desired concentration of Mg++, 1x buffer (supplied by Idaho Technologies), 5 µM dNTPs, 0.2 µg of each of two primers, and 0.75 units Taq Polymerase (Promega) or 1 unit Pfu polymerase (Stratagene). Amplification of whole worm DNA required lysis as described for the Perkin-Elmer protocol below, followed by addition of the above described reaction mixture. Reactions were carried out with twenty-five second denaturing at 94°C, twenty-five second annealing at the appropriate temperature determined for each pair of primers, followed by one minute extension at 72°C (longer amplification products were given 5 minute extension times). Each reaction was carried out over thirty cycles. Samples from completed reactions were transferred from the capillary tubes to microfuge tubes and stored at -20°C until further use. All samples were checked for the correct product by running a 1 µl aliquot on an agarose gel in 0.5x TBE buffer at approximately 5V/cm. DNA bands were visualized by staining with ethidium bromide (Sigma) and illuminating with UV (300 nm) light.

Perkin-Elmer for whole worm DNA amplification

Amplification reactions on whole worms carried out in a Perkin-Elmer Thermal Cycler were performed in 30 µl volumes in polystyrene microfuge tubes. Individual worms were first lysed in 5 µl lysate mixture containing 1.5 mM Mg++,
1x PCR Thermo Buffer (Promega), and 60 μg/ml Proteinase K. The samples were frozen at -70°C for at least ten minutes (or up to one week), then thawed and overlaid with 30 μl mineral oil to prevent evaporation. The samples were then incubated for one hour at 57°C followed by fifteen minute incubation at 95°C to inactivate the Proteinase K. After lysis, a mixture of 1x PCR Thermo Buffer, appropriate concentration of Mg++, 0.2 mM dNTPs, 0.2 μg of each of two primers and 0.75 units of Taq Polymerase (Promega) was added to each sample. Amplification reactions were carried out over thirty cycles, with forty-five second denaturation at 94°C, annealing at the appropriate temperature for each primer pair with a time determined based on the size of the predicted product, followed by one minute extension at 72°C. As before, samples were tested for the expected product by running a 1 μl aliquot in an agarose gel in 0.5x TBE buffer at approximately 5V/cm. Bands were again visualized by staining with ethidium bromide (Sigma) and illuminating with UV (300 nm) light.

Cloning PCR Products

DNA amplified through the PCR process was cloned into pBluescript (Stratagene) to allow for further manipulation. To ensure the purity of the DNA product, bands were first excised from agarose gels and purified as described below. When Taq polymerase was used in the amplification reaction, PCR products were treated for thirty minutes at room temperature with T4 DNA polymerase (1 unit) in 1x manufacturer’s supplied buffer (New England Biolabs) with excess (0.5 mM) dNTPs in a total volume of 50 μl. The “polished” DNA was purified by phenol extraction (see below), and cloned into Smal-cut pBluescript as
described below. If *Pfu* polymerase (Stratagene) was used for DNA amplification, the DNA fragment was cloned directly into *Sma*I-cut pBluescript.

**Restriction Enzyme Digestions**

Restriction enzyme digests of DNA samples (plasmids, PCR amplification products) were done by incubating DNA with 0.5 μl (1-5 units) of the appropriate enzyme and the manufacturer's supplied enzyme buffer at the suggested concentration, as well as 1x bovine serum albumin (New England Biolabs) in 15 μl total volume. Double digests were either done concurrently (in 15 or 20 μl total volume) if the two enzymes are optimally active in the same buffer solution; otherwise, the first digest was done as described, followed by phenol extraction of the DNA, and the second digest subsequently performed.

Overhang ends remaining after digestion were blunt ended in one of two ways: 3' overhangs were treated with T4 DNA polymerase (New England Biolabs) while 5' overhangs were blunted by the Klenow fragment (Pharmacia). Each procedure involved thirty minute incubations at room temperature with the manufacturer's supplied buffer, if necessary.

**Ligations**

Ligation of DNA fragments was performed by aliquoting at least a 3:1 molar ratio (estimated qualitatively) of insert DNA to vector DNA with 0.5 μl (200 units) of T4 DNA Ligase (New England Biolabs; Bethesda Research Laboratories) and the manufacturer's supplied 1x buffer in a total volume of 10 μl. The ligation
mixture was incubated overnight at either 16°C or 18°C, then stored at -20°C until further use.

Gel Purification of DNA Samples

Qiagen Gel Purification kit

Purification of DNA samples from agarose gels using the Qiagen Qiaex II extraction protocol followed the manufacturer's suggested guidelines. Bands were excised from gels using clean, sterile razor blades, trimmed of most agarose and transferred to a 1.7 ml centrifuge tube. Following the supplied protocol, the DNA is purified by adhesion to the Qiaex II beads, followed by several washing steps to remove residual agarose and salts. The DNA was subsequently eluted into 20 μl Tris-EDTA (pH 8.0) buffer or dH₂O, then stored at -20°C until further use.

Phenol/CIA with NaOAc/EtOH precipitation

An alternate method for purifying DNA samples from agarose gel involved excising the desired bands, as described above, then extracting the DNA in steps with phenol, phenol in chloroform:isoamyl alcohol (CIA), and finally CIA alone. The extracted DNA was then precipitated with sodium acetate and 95% ethanol at -70°C and the pellet subsequently washed in 70% ethanol, air-dried, and resuspended in dH₂O.
Minipreparations of Plasmid DNA

Bacterial transformant colonies grown on selective Ampicillin-containing Luria-Bertani medium (LB) plates were subsequently treated, following the methods of Sambrook et al. (1989), to extract the plasmid DNA. Individual colonies (from the transformation protocol described below) were used to inoculate 1 ml LB broth containing 100 μg/ml Ampicillin in 15 ml glass centrifuge tubes, then incubated overnight at 37°C in a rotating incubator. Overnight cultures were then spun down and the supernatant removed. The pellet was resuspended in 100 μl Solution I (Sambrook et al., 1989), then 1% SDS and 0.2 M NaOH were added to lyse the bacterial walls during a five minute incubation on ice. The proteinaceous elements were then precipitated by adding Solution III (Sambrook et al., 1989) and incubating on ice for an additional five minutes. The solutions were then spun down, and the supernatant, containing both DNA and RNA, transferred to a fresh 1.7 ml centrifuge tube. The DNA was either precipitated using 95% ethanol, or was first extracted in phenol and chloroform/isoamyl alcohol, and then subsequently precipitated. In all cases, the DNA pellets were washed with 70% ethanol, then air-dried and dissolved in 20 μl Tris-EDTA containing 50 μg/ml RNAse A. All samples were stored at -20°C until further use.

Preparation of DNA for Injection

Preparation of plasmid DNA for injection was performed following the prescribed protocol of the Qiagen plasmid purification kit. Overnight cultures
grown in 2 ml LB containing 100 μg/ml Ampicillin were transferred to a microfuge tube and pelleted. The cells were then resuspended in the supplied buffer containing RNase, then lysed and neutralized. The resulting precipitate was pelleted and the supernatant passed over the Qiagen extraction membrane. After washing, DNA was eluted from the membrane with 100 μl Tris-EDTA pH 8.0. Approximate DNA concentration was determined by running an aliquot of a restriction digest on an agarose gel (as described) adjacent to DNA markers of known concentration. More accurate measurements were obtained by UV spectrophotometry.

Injections

Germline transformation of *C. elegans* is performed by injection of DNA into the distal arm of one or both gonad arms in adult hermaphrodites following the method of Mello, *et al* (1991, 1995). This procedure was graciously performed by Diana Janke (Simon Fraser University). Plasmid DNA constructs were co-injected at a total concentration of 100 ng/μl with the plasmid pCes1943 which carries the dominant allele of *rol-6* (*su1006*). Worms carrying this marker exhibit the Roller phenotype, rolling around their longitudinal axis as they move in a circular motion pattern. These worms are easily recognizable among a population of non-Rollers, whose movement pattern is sinusoidal.

The co-injected plasmids carry regions of homology (for example, the Ampicillin resistance gene) which allows recombination between the marker plasmid and the plasmid of interest. As a result, the presence of the Roller phenotype is generally indicative of the presence of the plasmid of interest. Injected worms were plated individually and their progeny scored for the Roller
phenotype. This procedure was performed in the strain CB937 using yeast KEX2, human \textit{hfur} and a \textit{C. elegans bli-4} cDNA encoding Blisterase A in an ectopic expression vector using the heat shock 16-2 promoter. Additional scoring for the presence or absence of the blistered phenotype was also required. This protocol was also used with \textit{bli-4} 5' upstream DNA deletion \textit{lacZ} constructs, injected into N2 worms.

**Confirmation of the Presence of the Plasmid of Interest in Transgenic Worms**

Transgenic lines are identified by the presence of Roller individuals in each generation. Coinjection of the \textit{rol-6 (su1006)} marker plasmid, however, does not guarantee that the marker phenotype coincides with the presence of the plasmid of interest. In order to confirm that a heritable transgenic line carries the desired plasmid, a PCR assay was used. Primers were designed (using the Amplify 1.0 program for Apple Macintosh computers) to anneal to specific regions within each transgene (Table 4). Other previously designed primers were also used. PCR products would only be amplified if the plasmid of interest was carried in the transgenic strain.

**Integrating Transgenic Arrays**

Extrachromosomal arrays carried by transgenic worms were integrated into the genome using UV irradiation (Figure 4). Approximately fifty L4 to early adult worms exhibiting the Rol-6 phenotype were exposed to UV light for 25 seconds.
Each worm was then picked to individual plates and allowed to lay eggs. These progeny were then screened for the Rol-6 phenotype. F1 rollers were transferred

Table 4

**Primers used in this study to test transgenic strains**

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>KRp138</td>
<td>5' GGT GCG GGC CTC TTC GCT ATT A 3'</td>
</tr>
<tr>
<td>KRp29</td>
<td>5' TAC TCA CCC ATA TCA GTC AC 3'</td>
</tr>
<tr>
<td>KRp25</td>
<td>5' GTG CTC CCA CCC CCT ATT 3'</td>
</tr>
<tr>
<td>KRp26</td>
<td>5' CCC AAA CCT TCT TCC GAT 3'</td>
</tr>
<tr>
<td>Universal (-21)</td>
<td>5' TGT AAA ACG ACG GCC AGT 3'</td>
</tr>
</tbody>
</table>

KRp138 anneals within *lacZ* and KRp29 anneals to the first exon of *bli-4*. These were used to test the *bli-4* 5' end/*lacZ* deletion transgenics. KRp25 and KRp26 anneal to the 3' UTR of *unc-54*, and were used to test all transgenics used in the functional conservation rescue assays. The Universal (-21) primer anneals to sequences in the plasmid backbone.
Figure 4

Transgenic Integration Screen

Legend: transgenic worms carrying free extrachromosomal arrays are designated by \( hEx \), while worms with integrated copies of the arrays are designated by \( In \) (see Table 3).
to separate plates, and their progeny screened for the marker phenotype. Since F2 rollers may be either heterozygous or homozygous for the integrated transgene, F3 progeny were evaluated for lack of wild type segregants.

**Staining Transgenic Worms Harbouring lacZ Constructs**

Integrated transgenic worms exhibiting the Rol-6 phenotype were picked to a drop of distilled water on a depression slide and allowed to dessicate at room temperature. Animals were then fixed to the slide by immersion in chilled acetone at -20°C for two minutes, then air dried. Fifty microlitres staining solution (prepared as outlined below) was added, the slides placed on a 5% agar (in 2 mM EDTA pH 7.5) plate and incubated at 37°C overnight.

**Staining Solution:**

- 0.2 M sodium phosphate pH 7.5
- 1 mM MgCl₂
- Redox solution: 0.5 mM potassium ferricyanide
- 0.5 mM potassium ferrocyanide
- 0.004% SDS
- 10 µg/ml Kanamycin
- 1 µg/ml DAPI
- distilled H₂O
- 3% Xgal (dissolved in dimethyl formamide)
DNA Sequence Analysis

Restriction mapping of DNA sequences was performed with the DNA Strider program for the Apple Macintosh computer. Analysis of DNA sequences was achieved by searches in GENBANK and other databases using the BLAST network service (e-mail address: blast@ncbi.nlm.nih.gov). Computations for these searches were performed at the National Centre for Biotechnology Information (NCBI). Analysis of the promoter region of bli-4 was done using the University of Pennsylvania's Transcription Element Search Software (TESS) and the Neural Network Promoter Prediction program (NNPP) (Reese, 1994; Reese and Eeckman, 1995; Reese et al, 1996) (access to these analysis engines can be achieved through the internet address http://dot.imgen.bcm.tmc.edu:9331/seq-search/gene-search.html).

Transformations

DH5α for cloning

Cloning and maintaining stocks of plasmid DNA makes use of E. coli strain DH5α for propagation. Transformation of competent E. coli (GIBCO-BRL) was achieved by incubating 50 µl cells, thawed on ice, with 2.5 or 3 µl of either plasmid DNA or ligation mixture (in the case of plasmid construction) in chilled 15 ml centrifuge tubes on ice for at least thirty minutes. Cells were then heat-shocked at 37°C for thirty seconds, followed by an additional two minute incubation on ice. SOC (Sambrook et al 1989) medium (0.45 ml) was then added to the centrifuge tube, and the cells were incubated at 37°C in a shaking incubator for one hour.
Transformed cells were then plated (in 50 or 200 µl volumes) on LB agar plates containing 50 µg/ml Ampicillin. In some cases, selection for the presence of DNA inserted into the pBluescript backbone vector was determined by plating on selective medium containing X-gal. Recombinant plasmids produce colourless colonies due to interruption of the β-galactosidase gene in the vector, while non-recombinants produce blue colonies (Hanahan, 1983). Plates were incubated overnight at 37°C.

*S. cerevisiae*

Competent *S. cerevisiae* cells (strain Y143: MATα, KEX2Δ::URA3) were used for transformations involving the BLI-4 isoform most similar to kex2p (Blisterase D). Five to 10 µl of plasmid DNA (not treated with RNAse) was added to 100 µl cells thawed on ice. If DNA was RNA-free, 5 µl of boiled and cooled salmon testes DNA was added to aid the transformation efficiency. An equal volume of 70% polyethylene glycol (PEG) was added and the cells then mixed gently, followed by at least thirty minutes incubation at 30°C. Cells were then heat-shocked at 42°C for 5-15 minutes. PEG was removed by spinning down cells and removing supernatant; subsequently, the pellet was resuspended in 400 µl YEPD+ADE (see Yeast Growth Media) and incubated for twenty minutes at 30°C. Following this incubation, cells were spun down and the supernatant removed, and the pellet resuspended in 1 ml Tris-EDTA+sorbitol. 300 µl volumes were plated on SD-TRP (see Yeast Growth Media) plates and incubated for three days at 30°C. As a result of using selective media (lacking tryptophan), only those cells carrying the plasmid will grow.
Yeast Growth Media

YEPD+ADE
- 10g yeast extract
- 20g peptone
- 120 mg adenine
- dH₂O to 1 litre
- 2% glucose

SD-TRP
- 6.7g bacto-yeast nitrogen
- 2g “drop-out” medium (all amino acids except tryptophan)
- 20g agar (for plates)
- dH₂O to 1 litre
- 2% glucose

SGAL-TRP
- 6.7g bacto-yeast nitrogen
- 2g “drop-out” medium (all amino acids except tryptophan)
- 20g agar (for plates)
- dH₂O to 1 litre
- 2% galactose

YEPGAL
- 10g yeast extract
- 20g peptone
- dH₂O to 1 litre
- 2% galactose

Yeast Transformant Analysis

Transformed yeast colonies grown on the selective SD-TRP media were streaked onto minimal media plates (YEPD), then streaked in patches onto doubly
selective SGAL-TRP (see Yeast Growth Media) plates. The presence of galactose in the media induces the GAL promoter to transcribe the transgene. After overnight growth, the patches were replica-plated onto YEPGAL (see Yeast Growth Media) plates with a lawn of MATa-type yeast (strains 2625 and 2488). Functional rescue is indicated by the presence of halos of non-growth surrounding the patch. This work was done in collaboration with Kelly Blundell in C. Boone’s laboratory (Simon Fraser University).

Construction of full-length Blisterase D cDNA

In order to address the question of conservation of functionality between bli-4 and KEX2, I initially attempted functional rescue of *S. cerevisiae* strains mutant in KEX2 with one of the four BLI-4 isoforms most structurally similar to kex2p. The largest of the four isoforms, BLI-4D, has elements similar to those of kex2p, including a transmembrane domain (Figure 3). A yeast vector (YCp50) that makes use of the GAL promoter (Rose *et al.*, 1987) was used to carry a full-length cDNA for blisterase D. The cDNAs previously isolated (Thacker *et al.*, 1995) included a partial D isoform transcript which lacked the 5' region, including part of the common region. In order to prepare a full-length BLI-4D cDNA, the 5' sequence missing from the existing partial transcript was transferred from another of the isoforms' (BLI-4A) cDNA (Figure 5). The partial D cDNA carried in pBluescript, (pCeh 196), was digested with *EcoRI* to excise the cDNA fragment and the ends were blunt-ended using the Klenow fragment. Meanwhile, the A cDNA, also carried in pBluescript (pCeh195), was digested with *NcoI* at a site within the region of overlap between the two cDNAs. In addition, pCeh195 was
Figure 5

Construction of pCeh 228 (full-length bli-4 D cDNA)
cut with SmaI at a site within the Bluescript vector. The effect of these two digests was to remove the region shared by the two isoforms. pCeh196 was also cut with NcoI to produce homologous sticky ends between the two cDNA fragments. The resulting 1940 bp linear fragment was isolated by gel extraction purification as described. In addition, the 1126 bp fragment from the SmaI/NcoI double digest of pCeh195 was also gel purified. These two fragments were ligated together (as described) using T4 DNA ligase (New England Biolabs) to create a full-length cDNA for the BLI-4D isoform, carried in pBluescript (pCeh228).

Cloning yeast KEX2 and human hfur into hsp 16-2 expression vector

To address the question of whether kexin family members KEX2 (S. cerevisiae) and hfur (human) show evolutionary conservation of function with bli-4 in C. elegans, these were used to genetically transform bli-4 mutants. A genomic fragment of yeast KEX2 in a YCp50-based vector, kindly provided by Bob Fuller (University of Michigan Medical Center), was digested with BamHI to release the 3.3 kb fragment and cloned into the BamHI site in pBluescript. This construct was subsequently digested with KpnI and SacI and ligated into the C. elegans heat shock 16-2 (P. Candido, University of British Columbia) promoter-containing expression vector pPD49.78 (provided by A. Fire, Carnegie Institute of Technology) which had been digested at the same sites within the multiple cloning region (Figure 6a).

A cDNA clone of the human gene hfur contained in a pUC19 vector (generously provided by Gary Thomas, Oregon Health Sciences University) was digested with XbaI and SacI and ligated into the hsp 16-2 expression vector that
had been cut with \textit{NheI} (which has compatible ends with \textit{XbaI}) and \textit{SacI} within
the multiple cloning site (Figure 6b).

A positive control construct was designed to use a cDNA clone of the
Blisterase A isoform, which has previously been shown as capable of rescuing
the blistering phenotype in transgenic \textit{bli-4(e937)} animals (Srayko, 1995). The
cDNA clone (pCeh195) was digested first with \textit{KpnI}. This initial digest was
followed by a partial digest with \textit{SacI}, which has one internal site in the cDNA.
The appropriately sized fragment (2.2 kb) representing the complete cDNA was
gel purified using the Qiagen protocol (described above). This was subsequently
ligated into \textit{KpnI/SacI}-digested pPD49.78 (Figure 6c). All three constructs were
prepared for injection as described previously.

\textbf{Heat-shock induction of KEX2 and hfur transgenics and phenotypic rescue of
blistering}

Transgenic worms carrying heat shock expression constructs of yeast
KEX2 and human \textit{hfur} were maintained on Modified Youngren’s Only Broth
(MYOB) plates. Worms were first developmentally staged by allowing single
hermaphrodites to lay eggs overnight at 20°C and then removing the adult. The
eggs were then allowed to develop such that the resulting population was
approximately synchronous. Alternatively, plates were starved to induce
formation of dauer larvae; chunks of agar were transferred to new plates for two
to three hours. The dauer larvae thus transferred developed into adults within 12
to 15 hours (Lewis and Fleming, 1995). Late L4 larval populations were heat
shocked at 30°C for two hours, then returned to 20°C. Phenotypic rescue of \textit{bli-4}
Figure 6(a)

Construction of pCeh295:

Cloning *S. cerevisiae* KEX2 into the hsp 16-2 expression vector pPD49.78
**Figure 6(b)**

**Construction of pCeh296:**

Cloning human *hfur* into the hsp 16-2 expression vector pPD49.78
Construction of pCeh297:

Cloning *bli-4* A cDNA into the hsp 16-2 expression vector pPD49.78
(e937) homozygotes by yeast KEX2 and human hfur transgenes was measured by scoring adults for presence or absence of the blistering phenotype. Non-transgenic bli-4 (e937) worms were used as controls, in addition to the control strains. Transgenic worms carrying pCeh297 were used as positive controls, while a transgenic line carrying pPD49.78 was used as a negative control. In order to establish a baseline level of blistering, transgenic worms not exposed to heat shock conditions were also scored. In addition, non-transgenic worms (strain CB937) were exposed to both experimental conditions.

**Preparation of deletion constructs for analysis of the 5' end of bli-4**

Previous analysis of the expression patterns of bli-4 (Thacker et al, 1995) made use of expression vectors (provided by A. Fire) into which the 5' end of bli-4, as well as part of the first exon (the XbaI to ClaI fragment as illustrated in Figure 7) were cloned in frame with the E. coli gene lacZ. This vector system also utilizes a nuclear localization signal for ease of identification of cells in which the promoter drives expression. This original fusion construct contained approximately 5 kb of bli-4 upstream sequence. In order to further define sub-regions critical for determination of timing and tissue specificity of bli-4 expression, a limited deletion analysis was undertaken. Three separate deletions were prepared as shown in Figure 8 (a fourth was also attempted, but was discarded for reasons outlined in Appendix I). Each of these was achieved by digesting the parent plasmid (pB45.28) with two restriction enzymes in compatible buffer (as described earlier), followed by fill-in reactions. Each sample was run out on a low-melting-point agarose gel to ensure complete digestion. The digested linear plasmid was cut out of the gel and purified by Qiagen as described.
digest parent plasmid with 2 enzymes (e.g., BplI and Xhol); fill-in end overhangs

gel-purify appropriate fragment; re-circularize with T4 DNA ligase

Figure 7
Schematic representation of deletion construct preparation
Figure 8

Summary of deletion constructs used in the promoter analysis

Legend: a) genomic region spanning the bli-4 coding region. Exons are denoted by rectangles. b) DNA from an XbaI site upstream of the bli-4 coding region to a Clal site in the first exon of bli-4, contained in a lacZ reporter plasmid vector. c) deletions of the bli-4 upstream DNA fragment contained in the lacZ vector.
Ligation reactions to recircularize the plasmids were performed overnight, and competent *E. coli* transformed with the ligation mixture. Transformants were screened for the presence of the recircularized deleted plasmid. Positive colonies were grown overnight in LB containing 100 µg/ml Ampicillin and the DNA extracted and prepared for injection as described.
RESULTS

I. *S. cerevisiae KEX2 Mutant Rescue Analysis*

A. KEX2-deficient yeast are rescued by transformation with endogenous KEX2

Transformation of KEX2-deficient *S. cerevisiae* using a genomic clone of KEX2 was performed as outlined in Figure 9. The yeast genomic fragment, under the control of the GAL promoter in a plasmid vector provided by Bob Fuller (University of Michigan Medical Center), was induced by growing the transformant strain on galactose-containing media. When this strain was patched on a lawn of opposite mating type yeast, a halo formed (Figure 10), indicating that the extrachromosomally derived kex2p was expressed and activated. This experiment demonstrated that the expression vector activation and the test substrate were functional.
B. KEX2-deficient *S. cerevisiae* is not rescued by Blisterase D

Transformation of KEX2-deficient *S. cerevisiae* with a cDNA clone encoding the Blisterase D isoform of *bli*-4 was performed (Figure 9) to determine whether this *C. elegans* gene product was able to functionally substitute for the missing endogenous yeast convertase. The *bli*-4 cDNA clone was expressed through induction of the GAL promoter by growing transformants on medium containing galactose. Activity of the Blisterase D convertase was assayed by production of the active form of the α-mating factor which is a substrate of kex2p. This pheromone requires proteolytic cleavage by kex2p before being secreted. Cleavage of pro-α-mating factor is measured experimentally by the production of a halo around a patch of transformant cells plated on a lawn of opposite mating type yeast (Julius *et al.*, 1984). These experiments were performed in C. Boone’s laboratory with the assistance of Kelly Blundell (Simon Fraser University). KEX2-deficient yeast transformed with the Blisterase D cDNA did not produce a halo in the described assay (Figure 10), indicating that cleavage of the α-mating factor precursor did not occur.

C. Expression of Blisterase D in transgenic yeast

In order to confirm that the exogenous transgene does express the Blisterase D protein, Western analysis of protease-deficient transformed yeast was done. The Blisterase D-GST fusion antibody used for detection of the *C. elegans* gene product did not indicate that Blisterase D was present in the transgenic yeast.
Transform strain Y143
(MATα, KEX2Δ::URA3)

MATα, KEX2Δ::URA; bli-4 (D)
Grow on galactose to induce expression;
plate transformants on lawn of MATa yeast
(strains 2625, 2488)

Score for the presence of a halo (as in b);
this indicates cleavage of pre-α-mating factor

Figure 9
Experimental Scheme for Rescue of KEX2-deficient S. cerevisiae
Transformation of KEX2-deficient yeast with Blisterase D

The two patches on the left half of the plate are transformants expressing the endogenous KEX2 from a GAL-driven vector. The two patches on the right half of the plate are transformants containing the *C. elegans* bli-4 cDNA encoding Blisterase D.
II. C. elegans Mutant Rescue Analysis

A. bli-4(e937) mutants are rescued by the Blisterase A cDNA (pCeh297)

Previous work aimed at determining the roles of the individual products of bli-4 (Srayko, 1995; Thacker et al, 1995) has demonstrated that one of the four isoforms, Blisterase A, is able to rescue the blistering phenotype in bli-4(e937) mutant worms. As a consequence, a cDNA clone of the Blisterase A isoform (pCeh195), under the control of the heat shock promoter hsp 16-2, was injected into CB937 (e937/e937) hermaphrodites and stable transgenic lines established. This strain allows a baseline level of rescue to be established against which the abilities of KEX2 and hfur to rescue the blistering phenotype may be evaluated. Late L4 larval animals were exposed to 30°C heat for two hours, inducing expression of the extrachromosomal cDNA construct. Upon subsequent scoring, worms exposed to the heat shock conditions blistered only 43.0% of the time, while 72.7% of those not exposed to heat shock exhibited the blistered phenotype. These results are summarized in Table 5.

B. S. cerevisiae KEX2 (pCeh295) can rescue bli-4(e937) mutants

Genetic transformation of C. elegans is often used to evaluate the potential of exogenous products to phenotypically rescue mutants. In order to examine the ability of S. cerevisiae KEX2 to phenotypically rescue bli-4(e937) mutants, a yeast genomic clone (provided by R. Fuller, University of Michigan Medical Center), under the control of the C. elegans hsp 16-2 heat shock promoter (P. Candido, University of British Columbia; constructed by A. Fire, Carnegie...
Institute of Technology), was injected into CB937 (e937/e937) hermaphrodites by D. Janke (Simon Fraser University), and stable transgenic lines were obtained. Transgenic worms in the last half of the fourth larval stage (pre-adult) were exposed to 30°C for two hours to induce expression of the transgene. After heat shock, worms were scored for the presence or absence of the blistering phenotype. Transgenic worms not exposed to this heat shock treatment exhibited the blistered phenotype 72.4% of the time, while only 8.9% of those exposed to heat shock developed blisters. These results are summarized in Table 5.

The transgenic strain carrying the yeast KEX2 clone occasionally segregated worms with a dumpyish (Dpy) phenotype. Individual Dpy worms were followed through to subsequent generations; however, the phenotype was never heritable. The appearance of these worms may be the result of stochastic environmental effects, and may not reflect any properties of the yeast KEX2 transgene.

C. Human hfur (pCeh296) can rescue bli-4(e937) mutants

A human member of the kexin protease family, furin, is known to be expressed in a wide variety of tissue types and able to process a number of different substrate molecules. Both in vivo and in vitro experiments have shown furin to be able to cleave non-specific target molecules that contain a cleavage motif of R-X-K/R-R (Molloy et al, 1992). As a result of its widespread and ubiquitous activity, it was reasonable to evaluate the ability of furin to substitute for the lack of Blisterase A in C. elegans bli-4(e937) mutants. A cDNA clone of the human gene hfur (provided by G. Thomas, Oregon Health Sciences University), under the control of the heat shock promoter hsp 16-2, was injected into CB937 hermaphrodites and stable transmitting lines established as above.
### Table 5

**Summary of bli-4(e937) transgenic rescue results**

<table>
<thead>
<tr>
<th>Construct</th>
<th>Rescue</th>
<th>% Bli (heat shock)</th>
<th>%Bli (non-heat shock)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCeh295 (KEX2)</td>
<td>+</td>
<td>8.9 (41/459)</td>
<td>72.4 (416/575)</td>
</tr>
<tr>
<td>pCeh296 (hfur)</td>
<td>+</td>
<td>16.3 (90/551)</td>
<td>91.4 (655/717)</td>
</tr>
<tr>
<td>pCeh297 (bli-4 A cDNA)</td>
<td>+</td>
<td>43.0 (276/642)</td>
<td>72.7 (611/841)</td>
</tr>
<tr>
<td>pPD49.78 (negative control)</td>
<td>-</td>
<td>34.6 (191/552)</td>
<td>41.7 (221/530)</td>
</tr>
<tr>
<td>CB937</td>
<td>-</td>
<td>79.3 (658/1488)</td>
<td>74.1 (705/1656)</td>
</tr>
</tbody>
</table>

Summary of the results of the transgenic rescue experiment using *S. cerevisiae* KEX2 and human *hfur* in the inducible *hsp 16-2* heat shock vector. pPD49.78 is the heat shock 16-2 vector without any insert. Numbers in parentheses are the total number of Rol-6 transgenic worms counted for each measurement of blistering (for CB937, numbers represent total worms counted). Chi-squared statistical analysis is presented in Appendix II. Each transgenic strain, except the negative control (carrying pPD49.78), was back-crossed to another stock of CB937 in order to remove any potential mutations accumulated during strain maintenance.
Late L4 larval transgenics were exposed to two hour heat shock (30°C) to induce expression of the transgene. The worms were subsequently scored for the presence or absence of the blistering phenotype. Control animals not exposed to heat shock blistered 91.4% of the time, while 16.3% of those exposed to the 30°C temperatures developed blisters. These results are summarized in Table 5.

III. **Analysis of the 5' end of bli-4**

A. **Deletion of various regions upstream of bli-4 differentially affect gene expression**

Previous analysis of the controlling regulatory elements of bli-4 (Thacker *et al.*, 1995) determined that an approximately five kilobase region upstream of the coding region was sufficient to drive expression of lacZ in a fusion reporter construct. This plasmid was subjected to deletions through restriction digestion as described in Materials and Methods in order to further define the regions within the five kb upstream DNA that contain important regulatory information. Three deletions were prepared (Figure 11), and the individual constructs (pCeh288, pCeh289 and pCeh291) were injected into N2 hermaphrodites. Stable transmitting lines were established for each of the three constructs. Two of the three lines were then treated with UV to induce chromosomal breakage and thereby encourage integration of the extrachromosomal arrays. Individual animals (egg to adult) from each integrated line and the one non-integrated line were then stained with Xgal to detect expression of the bli-4/lacZ fusion protein. In addition, worms were stained with DAPI to allow localization of nuclei. Transgenic worms carrying the intact upstream region (KR3324 and KR3325) were
Figure 11

Deletions constructs prepared for analysis of the 5' end of bli-4

Legend: a) genomic region spanning the bli-4 coding region. Exons are denoted by rectangles. b) DNA from an Xbal site upstream of the bli-4 coding region to a Clal site in the first exon of bli-4, contained in a lacZ reporter plasmid vector. c) deletions of the bli-4 upstream DNA fragment contained in the lacZ vector.
used as positive controls, while those with 4 kb of the upstream DNA deleted (KR3002) were used as negative controls. KR3002 has been shown previously (Srayko, 1995) to exhibit no Xgal staining, indicating that essential regulatory elements are within the four kilobases of DNA deleted in the construct carried by that strain. The staining patterns of each strain, as well as the two control strains, are summarized in Table 6.

1. pCeh288

Approximately 3 kb of DNA between XbaI and BglII (see Figure 11) is deleted in pCeh288, leaving roughly 2 kb of DNA immediately upstream of bli-4 fused to the bacterial reporter gene lacZ. The non-integrated transgenic strain carrying pCeh288 (KR3205) exhibited no Xgal staining at any developmental stage. This suggests that required promoter elements that drive bli-4 expression may be found within the 3 kb region between XbaI and BglII upstream of the bli-4 ATG start codon.

2. pCeh289

The reporter construct carrying lacZ fused to DNA upstream of bli-4 is deleted between BglII and XhoI in pCeh289, leaving approximately 3 kb between XbaI and BglII upstream of the bli-4 coding region (see Figure 11). The non-integrated transgenic line carrying pCeh289 (KR3207) exhibited Xgal staining in embryos and early larval stages (L1-L2). Some limited staining was seen in older larval worms but none was observed in adults. Embryonic staining appears at the three-fold stage in a number of cells, identified as hypodermal by DAPI nuclear
Table 6

Summary of Xgal staining patterns in bli-4 upstream deletion construct transgenics.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Staining patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCeh288 (in KR3205)</td>
<td>no staining visible at any developmental stage</td>
</tr>
<tr>
<td>pCeh289 (in KR3207, KR3275, KR3276)</td>
<td>embryos, L1 and L2, limited staining at L3</td>
</tr>
<tr>
<td>pCeh291 (in KR3209, KR3277)</td>
<td>embryos, larvae</td>
</tr>
<tr>
<td>pCeh254 (in KR3002)</td>
<td>no staining visible at any developmental stage</td>
</tr>
<tr>
<td>pCeh320 (in KR3324, KR3325)</td>
<td>staining throughout development: hypodermal and vulval cells and ventral nerve cords</td>
</tr>
</tbody>
</table>
Figure 12(a)

pCeh289 transgenic staining pattern: Nomarski photomicrograph of representative KR3276 eggs
Figure 12(b)

pCeh289 transgenic staining pattern: Nomarski photomicrograph of KR3276 larvae
Figure 13(a)

pCeh291 transgenic staining pattern: Nomarski photomicrograph of representative KR3277 eggs
Figure 13(b)

pCeh291 transgenic staining pattern: Nomarski photomicrograph of representative KR3277 larvae
Figure 14

Staining patterns of positive control transgenic adult hermaphrodite (KR3324)
positioning. In both the L1 and L2 larvae, staining appears in seam hypodermal cells in the head and tail, and along the body length. A pair of hyp10 hypodermal cells appear consistently stained in the tails of L3 and L4 worms. Two integrated lines (KR3275 and KR3276) were also stained with Xgal and DAPI. Both show much the same pattern as the non-integrated line; however, one of these integrated strains also exhibits adult staining. This difference may be attributed to transcriptional interference at the site of integration. The results nevertheless suggest that the approximately 1 kb of DNA deleted in this construct contains information important in driving later (adult) expression of bli-4. Photomicrographs of embryos and larvae are shown in Figure 12 (a) and (b).

3. pCeh291

The plasmid pCeh291 has been deleted between BlpI and Xhol (see Figure 11), thus removing 2kb of DNA just upstream of bli-4. The construct contains only the 2 kb most distal to the bli-4 coding region. Strain KR3209, which carries pCeh291 in a non-integrated extrachromosomal array, shows Xgal staining in both embryos and larvae. The staining appears in hypodermal cells as indicated by nuclear positioning by DAPI staining. An integrated strain (KR3277) shows the same Xgal staining pattern. This suggests that the 2 kb of DNA most upstream of bli-4, as retained in this construct, contains elements important in the regulation of bli-4 expression in embryogenesis and early larval development, and further suggests that the 1 kb between BlpI and BglII may carry regions important for expression in adult development. Figure 13 shows representative staining patterns for KR3209 and KR3277.
4. Control constructs

Transgenic worms with the complete bli-4/lacZ fusion construct (pCeh320; provided by C. Thacker) containing approximately 5 kb of upstream DNA show extensive Xgal staining, as shown in Figure 14. The staining appears in hypodermal cells, vulval cells and in the ventral nerve cords. Staining is first detected at the three-fold stage of embryogenesis. Deletion of most of the 4 kb upstream DNA results in a total loss of expression as detected by Xgal staining (Srayko, 1995; this thesis).

B. Sequence analysis of the sub-regions upstream of bli-4

The deletion analysis described above allowed investigation of particular regions within the DNA upstream of bli-4. Fragments of the DNA sequence corresponding to each of the deletion constructs was sent by e-mail to the National Center for Biotechnology Information (NCBI) BLAST server (blast@ncbi.nlm.nih.gov) for comparison to other known sequence databases. The BLAST server detects homologies between the submitted DNA and known gene sequences. Thus, analysis by the BLAST server may suggest possible sites for DNA-DNA or DNA-protein interactions.

The DNA upstream of bli-4 contained in pCeh288 displayed homology with snRNA when aligned with sequences in the Eukaryotic Promoter Database, possibly suggesting RNA-DNA interactions within this region. However, the BLAST score for this alignment was insignificant (p=0.97), suggesting homologies only with small motifs not indicative of any biological relevance.
The bli-4 upstream DNA in pCeh289 displayed sequence homology with a number of other sequences when aligned within the non-redundant GenBank, EMBL, DDBJ and PDB databases. These include undefined 5' promoter regions from other organisms (yeast and the slime mold D. discoideum), erythroid-specific enhancer proteins (R. norvegicus), RNA polymerases (Borrelia burgdorferi and Spiroplasma citri) and zinc finger proteins (Rattus rattus and C.elegans). These alignments suggest at least one, but possibly more, sites for DNA-protein interactions in this region upstream of bli-4. As mentioned above, however, the probability values for each of these alignments are fairly low (p>0.0029), and may not be reflecting relevant interactions.

The bli-4 upstream DNA carried in the plasmid pCeh291 has approximately 75% overlap with that in pCeh289 and therefore a similar alignment with the non-redundant databases at NCBI. However, some regions unique to the non-overlapping DNA showed homology with C. elegans cosmids (listed in Table 7) identified through the sequencing projects (Cambridge and St. Louis). Some predicted genes of interest (as identified by ACeDB) include helicases, phosphatases and protein kinases. The sequence search results are summarized for each of the three deletion constructs in Table 8.

Analysis of each of the three sub-regions defined by the deletion constructs was also done using the University of Pennsylvania Computational Biology and Informatics Laboratory’s Transcription Element Search Software (TESS). The TESS search program searches for alignments between sample sequences and transcription factor sequences contained in the Transfac database. This search engine defined many transcription factor binding motifs in each deletion construct’s sequence (see Appendix IV). Sites of relevance were
Figure 15

**Schematic representation of predicted promoters based on the TESS and NNPP search results**

*Legend:* each * represents a putative promoter site based on the NNPP analysis results; each + represents a site with multiple transcription factor binding motifs as identified by TESS.
chosen based on multiple binding factors in one region. The results of the TESS analysis are consistent with the staining patterns described above.

Additional information was provided by the Neural Network Promoter Prediction (NNPP) program, which analyzes DNA sequences both for the presence of TATA-box motifs and surrounding sequence which is compatible with known promoters. NNPP analysis predicted six putative promoter motifs upstream of bli-4, all occurring within the 3 kb most proximal to the bli-4 coding region. The NNPP results are summarized in Table 9. A representation of all search results is shown in Figure 15.
Table 7

*C. elegans* cosmids identified by BLAST analysis in the DNA unique to pCeh289

<table>
<thead>
<tr>
<th>Cosmid</th>
<th>Predicted or known genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>T19D2</td>
<td>protein-tyrosine phosphatase</td>
</tr>
<tr>
<td>C05B5</td>
<td>ATP/GTP-binding site motif A (P-loop)</td>
</tr>
<tr>
<td>C18B2</td>
<td>unidentified cDNAs</td>
</tr>
<tr>
<td>T02C5</td>
<td><em>unc-2</em></td>
</tr>
<tr>
<td>F15A2</td>
<td>collagen, trehelase precursor, serine/threonine protein kinase, worm-specific proteins</td>
</tr>
<tr>
<td>B0457</td>
<td>ATP-dependant RNA helicase, ATP-dependent helicases, G-protein coupled receptor protein</td>
</tr>
<tr>
<td>K07E12</td>
<td>no known or predicted genes</td>
</tr>
<tr>
<td>ZC302</td>
<td>guanine nucleotide binding protein</td>
</tr>
</tbody>
</table>

Summary of BLAST search results for the DNA unique to the deletion construct pCeh289. Predicted or known genes listed here were identified using the *C. elegans* database ACeDB.
Table 8
Summary of BLAST search results for sub-regions in the DNA upstream of *bli-4*.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Unique BLAST search homologies</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCeh288</td>
<td>snRNA</td>
</tr>
<tr>
<td>pCeh289</td>
<td>promoters, zinc finger, tissue specific enhancer, RNA polymerase</td>
</tr>
<tr>
<td>pCeh291</td>
<td><em>C. elegans</em> cosmids</td>
</tr>
</tbody>
</table>

Summary of BLAST search results for sub-regions defined by deletions of the *bli-4* upstream DNA. Relevant homologies unique to each region carried in the plasmids listed above are reported here.
Table 9
Promoter predictions provided by the NNPP search engine.

<table>
<thead>
<tr>
<th>pCeh #</th>
<th>Promoter Sequence</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>288</td>
<td>TACAGTACTCTTTAAAGGCGCACACCCATTCACATTTAACAGAATTTT</td>
<td>566-616</td>
</tr>
<tr>
<td></td>
<td>CCGGCTACCGTATAATATGGGGCCAAAATGTTGCA TTTAACGCTCCTCATC</td>
<td>626-676</td>
</tr>
<tr>
<td></td>
<td>TCCAATCAGATAAAATACGGGCTTTGACCCA TAGCGTGAAAATCCCATG</td>
<td>932-982</td>
</tr>
<tr>
<td>289</td>
<td>TTCGACTCGTGATACCTGCGCAATTTCTGGAAATCATACTTAAGATTACT</td>
<td>1993-2043</td>
</tr>
<tr>
<td></td>
<td>TGAGATATAAAATCTCTAAAAATCTCTCCTCTCTCCCC CTAATCAATTGTAAT</td>
<td>2195-2245</td>
</tr>
<tr>
<td></td>
<td>GTGATCTGTGATGTAATCGGCTGGGCAATATCTGTTTCTGGGATCAGTTT</td>
<td>2275-2325</td>
</tr>
<tr>
<td></td>
<td>GGAAACCTATATAAATGTGTTAATAAAAATCTG GTGTTTTTTATTCTATG</td>
<td>2405-2455</td>
</tr>
<tr>
<td></td>
<td>TATTTAATTCTATTTAAAATTCAGCGAGCGAGG TAAAGATTAGTCATGGTA</td>
<td>2823-2873</td>
</tr>
<tr>
<td></td>
<td>ATTTCCCCCCATTTAAATCTCGACGCAATCTGGTT TCTCTCTTCTTACTCTT</td>
<td>2872-2922</td>
</tr>
<tr>
<td>291</td>
<td>TATTTAATCTATTTAAAATTCAGCGAGCGAGG TAAAGATTAGTCATGGTA</td>
<td>343-393</td>
</tr>
<tr>
<td></td>
<td>ATTTCCCCCCATTTAAATCTCGACGCAATCTGGTT TCTCTCTTCTTACTCTT</td>
<td>392-442</td>
</tr>
<tr>
<td></td>
<td>TACAGTACTCTTTAAAGGCGCACACCCATTCAC ATTTAACGAAACATTTTT</td>
<td>1126-1176</td>
</tr>
<tr>
<td></td>
<td>CCGGCTACCGTATAATATGGGGCCAAAATGTTGCA TTTAACGCTCCTCATC</td>
<td>1186-1236</td>
</tr>
<tr>
<td></td>
<td>TCCAATCAGATAAAATACGGGTCTTGACCCCAT AGCGTGAAAATCCCATG</td>
<td>1492-1542</td>
</tr>
</tbody>
</table>
Evolutionary Conservation of Function Between *C. elegans*, Yeast and Human Kexin Family Members

Many diverse examples of proteolytic activation of precursor molecules can be found throughout the animal kingdom. Following initial experiments three decades ago (Chretien and Li, 1967; Steiner et al, 1967), much information has been gathered regarding regulation of cellular processes through proteolysis. The kexin family of proprotein convertases, originally described by work with *S. cerevisiae*, is known to accomplish this kind of activation. The yeast KEX2 gene product, kex2p, is a Golgi-localized, calcium-dependent transmembrane protein that activates precursor forms of the α-mating factor (an essential pheromone for mating) and of the M₁ killer toxin during transport to the cell surface for secretion.

A human member of this family of serine endoproteases (furin) has also been extensively studied (Molloy et al, 1992; Molloy et al, 1994; Schafer et al, 1995; Pei and Weiss, 1995). Like kex2p, furin is a membrane-bound molecule, localized to the trans-Golgi network (TGN) (Molloy et al, 1994). It cleaves substrate molecules during a cycling process from the endoplasmic reticulum to
the cell surface. Furin is expressed in many tissue types, functioning as a ubiquitous activator of a number of substrate targets (Molloy et al., 1992; Pei and Weiss, 1995). Members of the kexin family can be classified as one of two types: those widely expressed in different tissue types (membrane-bound) and those expressed in tissue-specific patterns (non-membrane-bound). Structurally, both furin and kex2p are similar to two of the products of bli-4, Blisterase C and Blisterase D, suggesting that these two Blisterases may also be expressed in many tissue types. Conversely, the A and B isoforms are more similar to kexin family members known to be expressed in a tissue-specific manner (PCSK1/3, PCSK2 and PCSK4) (Smeekens et al., 1992; Thacker et al., 1995).

All kexin proprotein convertase family members cleave substrate molecules on the carboxy end of a pair of basic amino acid residues (lysine or arginine), thereby releasing the active form (Steiner, 1991; Seidah and Chretien, 1992). This minimal requirement for substrate recognition appears to be sufficient to induce in vitro cleavage of a substrate by any kexin. The prototypic family member, kex2p, cleaves substrate molecules containing only this dibasic motif. However, cleavage motifs can be expanded beyond this minimal recognition sequence, suggesting that regulation of this proteolytic activation process lies, at least in part, in very specific enzyme-substrate interactions. Yet since in vitro biochemical assays have shown that any kexin can cleave any molecule that contains the minimal dibasic residue motif, this suggests that a higher level of control, regulating the interaction between kexins and their substrates, must exist. There is likely to be gene regulation controlling tissue specificity of kexin protease expression. Clearly, proteolytic activation of a substrate molecule will only happen if it co-expresses with the catalytic enzyme. In unicellular organisms like S. cerevisiae, this kind of control is not possible. It may be that control of activation in these organisms is achieved through determining subcellular co-
localization, or through timing of expression. This indicates the importance of *C. elegans* as a model system for understanding the endogenous roles of kexin family members. Much of our current knowledge of this enzyme family comes from work in yeast with KEX2. However, this can not tell us about regulation through tissue specificity. The four kexin products encoded by bli-4 can tell us a great deal about the processes involved in regulation of interactions with substrate molecules. Thus, direct confirmation that the bli-4 gene products exhibit conservation of function with the prototypic kexin kex2p and with the higher kexin member from humans (furin) is essential in proving the value of *C. elegans* as an important model system in the study of kexin convertases.

**Transgenic Rescue of *S. cerevisiae* Mutants**

Transgenic rescue of mutant phenotypes has been used in many cases to demonstrate functional homology of gene products between diverse species. Previous attempts to phenotypically rescue *S. cerevisiae* KEX2 mutants with exogenous transgenes have been unsuccessful (Steiner, 1991; C. Boone, personal communication). However, genetic transformation of other yeast mutants with non-yeast transgenes (including *C. elegans* genes) has been achieved (Chaudhuri *et al.*, 1992; Chen *et al.*, 1993; Gavin *et al.*, 1995; Legget *et al.*, 1995). In this thesis, yeast carrying a deletion of KEX2 were transformed with a plasmid containing a cDNA for one of the *C. elegans* bli-4 products. A goal of this thesis was to demonstrate evolutionary conservation of function between known kexin family members, kex2p and furin, and the *C. elegans* putative members, the blisterases. Study of the kexin family of proprotein convertases has been largely limited to cell culture analysis for multicellular systems. Proof of functional identity between the
blisterases and other members of the kexin family would provide a useful model for understanding the functional role of kexin convertases in multicellular systems. The bli-4 isoform chosen (Blisterase D) is predicted through sequence analysis to be the most structurally similar to the native kex2p (see Figure 3). A cDNA encoding this single isoform was cloned into a yeast plasmid vector, placing it under the control of the inducible GAL promoter. However, induced transformants producing Blisterase D did not exhibit rescue of the mutant phenotype. As mentioned earlier, native kex2p is localized sub-cellularly to the Golgi apparatus, where it exists as a transmembrane protein, activating substrate molecules in the secretory pathway cycle between the Golgi and the cell surface. Thus, accurate localization of the exogenous molecule is essential in order for it to function. Previous studies of KEX2 have shown that kex2p is a relatively long-lived protein, with a half life of approximately eighty minutes (Wilcox et al, 1992). Following its synthesis, kex2p is transported to the late Golgi apparatus. The carboxy-terminal cytosolic tail of kex2p contains a signal required for retention in the Golgi. Both kex2p and pro-α-mating factor are co-localized to a processing vesicle, which develops into a secretory vesicle that delivers mature mating factor to the cell surface. Mutations in the cytoplasmic tail of kex2p have detrimental effects on the half-life, total cellular activity and localization of the protease. The localization or retention signals provided by Blisterase D may be sufficiently different (see Figure 16) so that transport to and retention in the Golgi apparatus either does not occur, or takes place slowly. One explanation could be that foreign proteins may be degraded through proteolysis, suggesting that the Blisterase D molecules degrade before localization can occur. To test this hypothesis, transformation of protease-deficient KEX2 mutants may be done, allowing for Blisterase D to be appropriately transported to the Golgi apparatus without degradation. This kind of analysis may determine whether Blisterase D
Comparison of:

(A) bli-4
(B) KEX2

using matrix file BLOSUM50

32.3% identity in 31 aa overlap; init: 43, opt: 49

\textit{bli-4} GDPFLDTHYFLYHSETT RTRRHKRAIVERLD
\textit{KEX2} GATFL VLYFMFFMKSRIRR SRAETYEFĐ

Figure 16

\textbf{Alignment of carboxy-terminal sequences from KEX2 and \textit{bli-4}}

Local fasta alignment of \textit{bli-4} amino acid sequence with the Golgi localization and retention signal of KEX2. The sequence on top is from \textit{bli-4}, while the bottom sequence is of KEX2. The bold residue in the KEX2 sequence represents an essential residue for appropriate localization of KEX2 in yeast.
shows conservation of function with kex2p in yeast. It is not likely that lack of substrate-enzyme specificity is responsible for the non-rescue, since it has been shown biochemically that any substrate containing the minimal recognition sequence will be appropriately cleaved by any kexin; however, this experiment has not been done using pro-α-mating factor. *In vivo* co-expression studies using other proprotein convertase family members (Thomas *et al*, 1988; de Bie *et al*, 1995) have shown that this redundancy also exists in biological systems. Fusion of the catalytic domain from bli-4 with the intracellular localization signals of KEX2 would demonstrate the ability of Blisterase D to function *in vivo* in transformed yeast. The fusion protein would be appropriately localized intracellularly, allowing analysis of the functional equivalency of Blisterase D and kex2p.

The failure to rescue the KEX2 mutant phenotype could also be the result of a number of experimental errors; as a result, it would be necessary to confirm that Blisterase D was expressed in the transgenic yeast in order to understand why rescue did not occur. The cDNA may not be transcribed or the RNA message may not be translated. Alternatively, the transgenic protein may be degraded before localization could take place. Western analysis of transgenic protease-deficient yeast was performed in order to determine whether or not mature Blisterase D protein was present. Protein extracts were stained with a GST-fusion Blisterase D-specific antibody. The results of this analysis were inconclusive, with much non-specific binding also detected. An alternate approach for confirming expression of Blisterase D in the transgenic yeast might take advantage of the reverse transcriptase PCR technique, which would determine whether or not the transgene is being transcribed.
Transformation Rescue of *C. elegans* Mutants

Transformation rescue of *C. elegans* mutants by non-nematode genes has been highly successful. There are many examples of genetic rescue by yeast, Drosophila and human transgenes (Vaux *et al*, 1992). This suggested that there was the potential to test for the abilities of yeast KEX2 and human *hfur* to phenotypically rescue *bli-4*. Germline injection of *C. elegans* is an often used and well-described method (Fire, 1986; Mello *et al* 1991; Mello and Fire, 1995) for genetic transformation. The ideal approach for transgenic rescue is to create an expression vector containing the promoter of the gene being replaced, thereby ensuring that the transgene will be expressed in the same manner in space and time as the endogenous form. Since this is not always possible (due to limited understanding of the regulatory elements for any particular gene), pre-constructed vector systems (A. Fire, Carnegie Institute of Technology) are often used. These vectors have been designed for a number of applications, from reporter assays (as described later) to expression vectors as used in this thesis. The heat-shock *hsp 16-2* (P. Candido, University of British Columbia) inducible expression vector (A. Fire, Carnegie Institute of Technology) contains the promoter elements of a heat-shock specific gene, one of three discovered through homology with known Drosophila heat shock systems (Snutch and Baillie, 1984; Jones *et al*, 1986; Stringham *et al*, 1992). The promoter is followed by a multiple cloning site which allows the transgene of interest to be readily cloned into the plasmid backbone. The 3’ untranslated region of the *C. elegans* gene *unc-54* occurs downstream of the cloning site. This region provides post-transcriptional modification information, and has been well-studied (Okkema *et al*, 1993). Thus, the heat shock promoter, induced by 30°C temperatures, initiates transcription of
the transgene. The complete message will carry the 3' UTR of unc-54, and should maintain stability long enough to be translated. In order to analyze the evolutionary conservation of function between the Blisterases and other kexin family members, the prototypic member, *S. cerevisiae* KEX2, and the ubiquitously expressed human member, *hfur*, were cloned into the heat shock vector and the constructs injected into CB937. Since the human kexin family member, furin, is expressed widely *in vivo*, it may be more closely related in function to the *C. elegans* Blisterases. Stable transgenic lines were exposed to heat shock temperatures as described in Materials and Methods and subsequently evaluated to determine whether or not they developed blisters. The strains carrying KEX2 and *hfur* developed reduced blistering in the Rol-6 animals without being exposed to high temperatures (data not shown). This is not an expected consequence, since the hsp 16-2 promoter is only active when induced by high temperatures. There is no evidence for genetic interactions between *bli-4* and *rol-6*, although the ROL-6 protein, which is expressed in the cuticle, does contain the cleavage motif predicted to be recognized by the Blisterases. However, this aberrant response was not observed in previous studies using *rol-6* as an injection marker in CB937 worms (Srayko, 1995; Thacker *et al*, 1995). It is possible that the strain used for injection may have accumulated mutations while being maintained. Consequently, the extrachromosomal array was crossed into a new CB937 stock, thereby eliminating the possibility of carrying any additional mutations. Nevertheless, it is clear that both KEX2 and *hfur* are able to phenotypically rescue the blistering mutation in *bli-4*, reducing the frequency of blistering by 60 to 75%. However, blistering is not completely eliminated. The extrachromosomal array is not integrated into the genome and may not be present in every cell. This might allow for the formation of a blister to initiate in any cell lacking the exogenous convertase (kex2p or furin). Since scoring is simply for the presence or absence of
blisters, these animals may be counted as blistered even though the foreign protein may be able to rescue the phenotype in other cells. In addition, since the blistered phenotype is adult-specific, the worms must be exposed to high temperatures at a particular time-interval when production of the adult cuticle is in process. If heat shock is too early, the exogenous protein may be degraded by the time it is needed. Conversely, if the worms are treated too late, the cuticle will already have been formed and the transgenic protein may be useless. Despite these potential confounding factors, the reduction in blistering frequency is sufficient to demonstrate functional conservation between the gene products of bli-4 with kex2p and furin. It is possible that kex2p rescues bli-4(e937) mutants more fully than furin; however, the mosaicism of expression of non-integrated transgenes and their variable heritability make this difficult to test. The controls used for transformation rescue of bli-4(e937) mutants provided somewhat unexpected results. As shown in Table 5, e937 homozygotes (strain CB937) develop blisters approximately 80% of the time. Yet surprisingly, only roughly 40% of the negative control transgenic worms, which carry only the plasmid vector without any coding information, developed blisters. This may be an unforeseen consequence of the injection process itself. The positive control transgenic worms, which carry a cDNA encoding the Blisterase A isoform, exhibited less complete rescue of the blistering phenotype than did either the KEX2 or hfur transgenic worms. Previous analyses of bli-4 (Srayko, 1995; Thacker et al, 1996) have suggested a level of redundancy among the products of bli-4. It may be that over-expression of Blisterase A alone, as in the positive control transgenic worms, is not sufficient for full phenotypic rescue. Knowledge of specific motifs recognized by the four Blisterases is not yet known. As mentioned previously, kex2p most efficiently cleaves substrate molecules containing only a pair of basic amino acids, while the preferred cleavage motif
recognized by furin is more complex (R-X-K/R-R or, less commonly, R-X-X-R) (Molloy et al, 1994; Shafer et al, 1995). This may suggest that the Blisterases also recognize a simpler motif of dibasic residues in their substrate targets. This cleavage specificity may be examined by biochemical enzymatic assays, measuring the ability of the Blisterases to cleave substrate molecules containing variations on the rudimentary dibasic motif (as reviewed in Seidah and Chretien, 1992). Functional identity with any of the individual Blisterases is not yet known. As more is learned about the regulatory signals of bli-4, we may be able to determine if particular elements determine each isoform’s expression. These could be used in a vector system as described above to drive expression of KEX2 and hfur in the manner of each particular bli-4 gene product.

**Regulation of Expression of bli-4**

Regulation of gene expression is under hierarchical levels of control. Global regulation through chromatin structure (Dixon et al, 1990; reviewed by Krause, 1995) is important in determining access of transcriptional machinery to areas of the genome. Transcriptional regulation of gene expression is under the control of RNA polymerases. In *C. elegans*, most genes are transcribed by RNA polymerase II, which binds a TATA consensus sequence generally found approximately 30 bp upstream of the transcriptional start site. Other transcriptional factors are being identified through sequence homologies as the sequencing project covers more of the genome. Control elements, such as enhancers, specific to one gene or a class of genes are also being identified. Gut-specific (Aamodt et al, 1991; Egan et al, 1995) and myosin-specific (Okkema et al, 1993) elements are also being identified by the use of transgenic strains carrying
reporter genes like *lacZ* and GFP fused to the promoter of interest. Further control of gene expression is exercised post-transcriptionally, through message splicing and 3' end processing. The final mature message is then subject to translational regulation. RNA-RNA interactions, such as that between *lin-4* and *lin-14* (Lee et al, 1993), demonstrate the kinds of control mechanisms acting on modified RNA molecules. *lin-4* produces a pair of small RNA molecules (22 and 61 nucleotides long) that are complementary to several regions in the 3' untranslated region of the *lin-14* message, suggesting that the *lin-4* RNAs act to down-regulate expression of *lin-14* by preventing translation of its RNA message. Currently, little is known about translational machinery in *C. elegans*. Nonetheless, *C. elegans* is progressively becoming a more useful system for dissecting many kinds of regulatory elements for genes of interest. Fragments of upstream or downstream DNA can be used in vector systems as mentioned above to control expression of reporter genes.

The control elements regulating expression of *bli-4* have been studied in this thesis. It is reasonable to expect that there is more than one signal directing *bli-4* expression, since it is expressed in a variety of tissue types at different developmental stages. The collections of mutants in *bli-4* also indicate complex regulatory elements, specifying expression in embryogenesis, early larval development and at the adult moulting. In order to address these issues, approximately 5 kb of upstream DNA, as well as the first coding exon of *bli-4*, was fused in frame with the bacterial reporter gene *lacZ*. A series of three deletions were created and the reporter constructs injected into N2 in order to examine the roles of each sub-region in regulating *bli-4* expression. The extrachromosomal arrays in each of the deletion-carrying transgenic lines were integrated into the genome following brief exposure to UV. Worms at all developmental stages from the integrated strains, as well as from an integrated line carrying the entire
upstream fragment (KR3324) and a non-integrated line with most of the upstream DNA deleted (KR3002), were stained with Xgal to examine the expression patterns.

In the transgenic line carrying the full, non-deleted upstream fragment (pCeh320), staining is seen in a number of cell types. Adult and larval transgenic animals show expression in all hypodermal cells, the vulva and the ventral nerve cords (Thacker et al., 1995; this thesis). Expression is first detected in hypodermal cells at the two-fold stage of embryogenesis, just prior to the developmental arrest stage in the bli-4 lethal mutants (Thacker et al., 1995). In KR3002, in which 4 kb of upstream DNA has been deleted, no staining is detected at any developmental stage, confirming that regulatory elements essential in controlling bli-4 expression are within this region. Sub-regions as defined by the series of deletions described in Materials and Methods (see Figure 11) show varying staining patterns. KR3205 non-integrated transgenic animals (carrying pCeh288) also show no staining at any time during development. Thus, essential required elements for expression lie in the most distal 3 kb of upstream DNA.

Transgenic animals which retain this most distal fragment (pCeh289), show Xgal staining in a variety of hypodermal cells in embryos and larvae, but not in adults, suggesting the presence of an adult-specific transcriptional cue in the 1 kb region between BglII and XhoI (see Figure 11). Staining appears to be similar to that in the non-deleted transgenic strain in the embryos and early larval stages, with expression detected in hypodermal seam cells. However, older larvae (L3 and L4) show restricted tissue expression patterns, with staining limited to a pair of hypodermal cells (hyp10) in the tail as well as a pair of hypodermal cells in the pharyngeal region. Two independent integrated lines carrying pCeh289 (KR3275 and KR3276) were analyzed by Xgal staining. The staining patterns are essentially the same, although one (KR3276) also shows some adult expression. It
is unclear which staining pattern reflects the true expression pattern of bli-4 and whether the expression pattern in either one of the integrated lines is being influenced by the site of integration. However, since non-integrated transgenic worms carrying pCeh289 (KR3207) show no staining in adult animals, it is likely that the integrant showing adult staining is exhibiting an aberrant expression pattern. Alternatively, the mosaic pattern of expression of this non-integrated array may be responsible for lack of detection in adult worms. Antibody staining may be done in order to confirm the wild-type expression pattern of bli-4.

Staining in KR3209 transgenics and in the integrated strain KR3277, which retain the most distal 2 kb of upstream DNA, appears both in embryos and larvae, suggesting that there may be at least one additional cis-acting factor directing adult-specific expression, located in the 1 kb region between BgII and XhoI (see Figure 11). These results suggest that there are at least three signals important in regulating expression of bli-4. In addition, the results indicate that at least one cis-acting factor may be required to act in concert with other regulatory elements for adult-specific expression. Further detailed analysis of this region must be done in order to define these regulatory signals. Sequential deletions through these regions should allow identification of small fragments whose sequence can be analyzed in more detail.

Analysis of the sequences of each of the sub-regions upstream of bli-4 made use of the NCBI BLAST server as well as the University of Pennsylvania’s Transcription Element Search Software (TESS) and the Neural Network for Promoter Predictions (NNPP). The National Center for Biotechnology Information’s BLAST server is a heuristic search algorithm used to assign significance to search tool findings using the statistical methods of Karlin and Altschul (Altschul et al, 1990). The BLAST server is most appropriate for sequence similarity searches but is not generally useful for motif-style searching.
Nonetheless, BLAST analysis of the sub-regions upstream of bli-4 may provide some insight into the regulatory elements important in bli-4 expression. Sequence homologies defined by the BLAST search engine include polymerases, tissue-specific enhancers, zinc fingers and other DNA-interacting proteins. However, none of the alignments had particularly high scores (see Appendix III), and may not reflect alignments of any biological interest.

As mentioned above, the NCBI BLAST server provides alignments of sequence rather than searching for particular motifs; consequently, TESS and NNPP programs were used to more appropriately analyze the upstream DNA. The University of Pennsylvania's Transcription Element Search Software (TESS) is designed for locating and displaying transcription factor binding sites in DNA sequence (see Appendix IV for output). TESS utilizes the Transfac database, searching for matches between strings in the sequences provided and all transcription factor sequences contained in the database. TESS analyses of each of the sub-regions defined in the deletion analysis described above identify a number of putative transcription factor binding sites (see Appendix IV). Potential sites of interest may be regions where multiple transcription factor binding domains are identified. Such regions were detected throughout the four kilobases of DNA analyzed in this deletion approach, as shown in Figure 15. These areas showing high sequence identity with transcription factor binding motifs are possible sites for transcription initiation. However, many more potential binding sites are identified by the TESS search program than are likely to be of biological interest. More refined deletions, or possibly site-directed mutagenesis, may allow for further elucidation of the roles of these binding motifs.

The Promoter Prediction by Neural Network (NNPP) program is designed to locate both prokaryotic and eukaryotic promoters in a given DNA sequence (Reese, 1994; Reese and Eeckman, 1995; Reese et al, 1996). The basis of the
The NNPP program is a time-delay neural network that consists primarily of two layers, one that recognizes the TATA-box and one that recognizes the region spanning the transcription start site. As before, each of the sub-regions described by the deletion analysis were searched using the NNPP program in order to locate potential promoter motifs. NNPP predicted six distinct promoter sites upstream of bli-4, none of which occur within the approximately two kilobases of DNA most distal to the bli-4 coding region (Figure 15). All six predicted promoter sites fall between BlpI and BglII, the same region defined by the reporter construct deletion analysis to contain information important for driving embryonic and larval expression of bli-4. In addition, these results may be supported by the TESS search software output. Three regions that contain multiple transcription factor binding motifs correspond with the locations of potential promoters identified by the NNPP program (Figure 15). Other multiple binding sites do not correspond to any predicted promoters and may instead be indicative of cis-acting factor binding sites. However, as mentioned above, it will be necessary to perform a more detailed deletion analysis of the bli-4 upstream region in order to clarify the factors involved in tissue-specific and developmental regulation of expression.

The analysis of bli-4 regulation described here provides an initial insight into the processes involved in the control of expression. However, it is important to note that only the 5' upstream DNA has been examined. There is much evidence to suggest that regulatory factors may also be found downstream of genes, as well as internally within introns. As a result, the analysis presented in this thesis must be supplemented with detailed examinations of 3' and internal influences on expression. A possible approach that will allow insight into all elements required to control expression may be cloning a non-disruptive reporter tag construct in frame into bli-4, thereby ensuring that all regulatory information
within and around the gene remain intact. Comparison of expression patterns using such a construct with those reported here may indicate whether important control elements were missing in this analysis.

Despite these potential confounding factors, the analysis described here provides unique information about the regulation of bli-4. This is the first such analysis performed for a member of the kexin family of proprotein convertases, and may be useful for understanding regulation in other family members. Sequencing of the genome of the related nematode *C. briggsae* is currently underway, and comparison of the upstream region of the *C. briggsae* homologue of bli-4 may provide clues about evolutionary conservation not only of function between kexin family members but also of regulation. It is likely that the entire control process of bli-4, and of other kexin family members, is as complex as the enzymatic roles of the convertases themselves. Both tissue-specific and developmental timing of expression must be closely regulated, suggesting that a number of transcriptional initiation molecules and processes may be involved. The analyses presented in this thesis represent a first step towards eventual understanding of these processes, in *C. elegans* and in other organisms of interest.

**CONCLUSIONS**

The goals of this thesis have been to investigate the evolutionary conservation of function between the products of the *Caenorhabditis elegans* gene bli-4 and kexin proprotein convertases from *S. cerevisiae* (kex2p) and humans (furin). The results shown here demonstrate that the four Blisterase products of bli-4 are functionally conserved with both the yeast and human
homologues. In addition, this thesis has investigated some of the regulatory elements controlling expression of bli-4. The preliminary results reported here indicate that at least three, but possibly more, transcriptional regulators are involved in determining developmental and tissue specific expression patterns of bli-4. These results underline the importance of C. elegans and bli-4 as a model system for understanding the biological roles of kexin family members throughout the animal kingdom.
REFERENCES


sequence of a mouse candidate prohormone convertase PC1 homolgous to PC2, furin and Kex2: distinct chromosomal localization and messenger RNA distribution in the brain and pituitary compared to PC2. Mol. Endocrinol. 5:111-122.


APPENDIX I

Analysis of the 5' DNA upstream of bli-4 has made use of reporter constructs, fusing fragments of upstream DNA with a reporter gene. In this thesis, three such deletions were prepared and evaluated using the bacterial gene lacZ as an indicator for determining expression patterns. A fourth deletion was also attempted, in order to break the four kilobases between XbaI and XhoI upstream of bli-4 into approximately 1 kb sub-units. The procedure used was identical for all deletions; however, in attempting to create a deletion between BlpI and XbaI, a number of difficulties were encountered. The initial step required double restriction enzyme digestion of the parent plasmid (pCeh320) with BlpI and XbaI in compatible buffer. The sticky ends were filled in and the DNA size-separated by gel-electrophoresis. The appropriately sized linear fragment was re-circularized by ligation of the blunt ends and used to transform competent DH5α bacterial cells. Transformant colonies were treated as described in Materials and Methods to purify the plasmid DNA. Purified DNA was tested by restriction enzyme digestion to ensure that the correct plasmid had been maintained. The same DNA was subsequently used for micro-injection as described in Materials and Methods. Each transgenic line created was tested by PCR for confirmation of the presence of the extrachromosomal array; however, none of the transgenic lines produced the expected product by PCR analysis. A second micro-injection procedure was performed, but again none of the lines were positive for the presence of the digested plasmid. The stock DNA used for micro-injection was tested by restriction digest to confirm its identity. Unfortunately, the results indicated that the DNA used for injection was not the same as had been confirmed previously; consequently, this deletion was not pursued for further analysis in this thesis. It is possible that the gel fragment purified from agarose
following restriction digestion also contained un-cut parent plasmid (pCeh320), which was selectively maintained during bacterial transformation. Alternatively, contamination of the DNA (at restriction digestion, gel purification, transformation) may have occurred such that the desired plasmid was not purified and maintained for use in micro-injection.
APPENDIX II

Chi-squared statistical analysis of heat shock rescue results:

H_0 = no rescue (expected frequency of blistering is the same as that for non-heat shock worms within the same strain)

df = 4 in all cases

<table>
<thead>
<tr>
<th>Construct</th>
<th>o^a</th>
<th>e^b</th>
<th>d^c</th>
<th>d^2</th>
<th>( \chi^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCeh295 (KEX2)</td>
<td>8.9</td>
<td>72.4</td>
<td>-63.5</td>
<td>4032.25</td>
<td>55.7</td>
</tr>
<tr>
<td>pCeh296 (hfur)</td>
<td>16.3</td>
<td>91.4</td>
<td>-75.1</td>
<td>5640.01</td>
<td>61.7</td>
</tr>
<tr>
<td>pCeh297 (bli-4 A)</td>
<td>43.0</td>
<td>72.7</td>
<td>-29.7</td>
<td>882.09</td>
<td>12.1</td>
</tr>
<tr>
<td>pPD49.78</td>
<td>34.6</td>
<td>41.7</td>
<td>-7.1</td>
<td>50.41</td>
<td>1.2</td>
</tr>
<tr>
<td>CB937</td>
<td>79.3</td>
<td>74.1</td>
<td>5.2</td>
<td>27.04</td>
<td>0.4</td>
</tr>
</tbody>
</table>

^a Observed
^b Expected
^c Deviation

The value of \( \chi^2 \) for pCeh295 is sufficiently high (p<0.01) that the null hypothesis may be rejected; in other words, KEX2 does rescue blistering.

The value of \( \chi^2 \) for pCeh296 is also sufficiently high (p<0.01) that the null hypothesis may be rejected; as above, hfur is capable of rescuing the blistered phenotype.

The value of \( \chi^2 \) for pCeh297 gives a probability of approximately 0.01. This may be accepted as rejection of the null hypothesis, but higher numbers of worms may be necessary for confirmation that the cDNA encoding Blisterase A does rescue the blistered phenotype.

The value of \( \chi^2 \) for both the negative control construct (pPD49.78) and the non-transgenic strain CB937 give p values greater than 0.99 and 0.8, respectively; thus, the null hypothesis must be accepted. There is no rescue of blistering in either strain.
APPENDIX III

BLAST analysis of the 5' region upstream of bli-4

<table>
<thead>
<tr>
<th>Sequences producing High-scoring Segment Pairs:</th>
<th>High Score</th>
<th>Smallest Sum Probability P(N) N</th>
</tr>
</thead>
<tbody>
<tr>
<td>gb</td>
<td>M55332</td>
<td>DDIDISIG</td>
</tr>
<tr>
<td>dbj</td>
<td>D50338</td>
<td>DDIRIPR</td>
</tr>
<tr>
<td>gb</td>
<td>U40411</td>
<td>CELB0403</td>
</tr>
<tr>
<td>emb</td>
<td>X00325</td>
<td>IDGAMMA1</td>
</tr>
<tr>
<td>emb</td>
<td>X85120</td>
<td>ASPDXD3H</td>
</tr>
<tr>
<td>emb</td>
<td>X85121</td>
<td>ASPDXD3C</td>
</tr>
<tr>
<td>gb</td>
<td>U56960</td>
<td>CELF25H10</td>
</tr>
<tr>
<td>emb</td>
<td>X74778</td>
<td>HSREPC4</td>
</tr>
<tr>
<td>gb</td>
<td>L23077</td>
<td>RATZFP15A</td>
</tr>
<tr>
<td>emb</td>
<td>X13203</td>
<td>GHLEA29</td>
</tr>
<tr>
<td>gb</td>
<td>U38750</td>
<td>ESU38750</td>
</tr>
<tr>
<td>emb</td>
<td>Z49908</td>
<td>CEC07E3</td>
</tr>
<tr>
<td>emb</td>
<td>X04775</td>
<td>DDCPRO2U</td>
</tr>
<tr>
<td>gb</td>
<td>L04272</td>
<td>MSQNCATR</td>
</tr>
<tr>
<td>gb</td>
<td>U01010</td>
<td>U01010</td>
</tr>
<tr>
<td>gb</td>
<td>U41029</td>
<td>CELF47G3</td>
</tr>
<tr>
<td>gb</td>
<td>U02970</td>
<td>PWU02970</td>
</tr>
<tr>
<td>gb</td>
<td>U60168</td>
<td>DDU60168</td>
</tr>
<tr>
<td>gb</td>
<td>L39752</td>
<td>DRO29DC3Z</td>
</tr>
<tr>
<td>gb</td>
<td>J01283</td>
<td>DDIDISIB</td>
</tr>
<tr>
<td>gb</td>
<td>U64603</td>
<td>CELC09B7</td>
</tr>
<tr>
<td>gb</td>
<td>J01284</td>
<td>DDIDISIC1</td>
</tr>
<tr>
<td>emb</td>
<td>Z11874</td>
<td>CHEGZ</td>
</tr>
<tr>
<td>gb</td>
<td>L36887</td>
<td>YSCMTCG03</td>
</tr>
<tr>
<td>emb</td>
<td>X52177</td>
<td>FFCRPOBC</td>
</tr>
<tr>
<td>gb</td>
<td>U67599</td>
<td>MJJU67599</td>
</tr>
<tr>
<td>gb</td>
<td>U42846</td>
<td>CELT19D2</td>
</tr>
<tr>
<td>gb</td>
<td>U30821</td>
<td>CPU30821</td>
</tr>
<tr>
<td>gb</td>
<td>L36899</td>
<td>YSCMTCG15</td>
</tr>
<tr>
<td>emb</td>
<td>Z70204</td>
<td>CEC11G6</td>
</tr>
<tr>
<td>emb</td>
<td>X85791</td>
<td>RNLTPKG</td>
</tr>
<tr>
<td>gb</td>
<td>L48488</td>
<td>BORRPOB</td>
</tr>
<tr>
<td>gb</td>
<td>U43281</td>
<td>SCU43281</td>
</tr>
</tbody>
</table>
emb|Z77135|CET16A9 Caenorhabditis elegans cosmid T16A9  130  0.75  2
gb|U55376|CELF16H11 Caenorhabditis elegans cosmid FI6H1I  116  0.76  4
dbj|D31785|HASMT Hansenula wingei mitochondrial DNA, ...  130  0.77  2
emb|Z79605|CEZEK678 Caenorhabditis elegans cosmid ZK678  131  0.78  2
emb|Z70206|CEC49F8 Caenorhabditis elegans cosmid C49F8  152  0.78  3
gb|L34637|HUMPECAM07 Homo sapiens platelet/endothelial ce.  140  0.80  1
gb|M35027|VACCG Vaccinia virus, complete genome.  124  0.84  3
emb|Z71517|SCYNL241C S.cerevisiae chromosome XIV reading  138  0.90  1
lcl|EPD30010 (+) Sp snRNA U2 early; range -499 to 100.  100  0.90  1
emb|X57336|SCMET19 S. cerevisiae MET19 gene for glucose...  138  0.91  1
gb|M34709|YSCG6PD S. cerevisiae glucose-6-phosphate deh...  138  0.91  1
emb|Z71518|SCYNL242W S. cerevisiae chromosome XIV reading  138  0.91  1
gb|M35027|VACCG Vaccinia virus, complete genome.  124  0.84  3
emb|Z71517|SCYNL241C S.cerevisiae chromosome XIV reading  138  0.90  1
lcl|EPD30010 (+) Sp snRNA U2 early; range -499 to 100.  100  0.90  1
emb|X57336|SCMET19 S. cerevisiae MET19 gene for glucose...  138  0.91  1
gb|M34709|YSCG6PD S. cerevisiae glucose-6-phosphate deh...  138  0.91  1
emb|Z71518|SCYNL242W S. cerevisiae chromosome XIV reading  138  0.91  1
gb|M35027|VACCG Vaccinia virus, complete genome.  124  0.84  3
emb|Z71517|SCYNL241C S.cerevisiae chromosome XIV reading  138  0.90  1
lcl|EPD30010 (+) Sp snRNA U2 early; range -499 to 100.  100  0.90  1
emb|X57336|SCMET19 S. cerevisiae MET19 gene for glucose...  138  0.91  1
gb|M34709|YSCG6PD S. cerevisiae glucose-6-phosphate deh...  138  0.91  1
emb|Z71518|SCYNL242W S. cerevisiae chromosome XIV reading  138  0.91  1
gb|M35027|VACCG Vaccinia virus, complete genome.  124  0.84  3
<table>
<thead>
<tr>
<th>Accession</th>
<th>Description</th>
<th>Similarity</th>
<th>Confidence</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>gb</td>
<td>U40410</td>
<td>CELC54G7</td>
<td>Caenorhabditis elegans cosmid C54G7.</td>
<td>131</td>
</tr>
<tr>
<td>gb</td>
<td>K00539</td>
<td>YSCMTTG03</td>
<td>Yeast (S.cerevisiae) mitochondrial c...</td>
<td>133</td>
</tr>
<tr>
<td>emb</td>
<td>Z74038</td>
<td>CEF58B4</td>
<td>Caenorhabditis elegans cosmid F58B4</td>
<td>121</td>
</tr>
<tr>
<td>emb</td>
<td>Z50176</td>
<td>CEC09G1</td>
<td>Caenorhabditis elegans cosmid C09G1</td>
<td>130</td>
</tr>
<tr>
<td>gb</td>
<td>U31083</td>
<td>PFU31083</td>
<td>Plasmodium falciparum erythrocyte me...</td>
<td>115</td>
</tr>
<tr>
<td>emb</td>
<td>V00693</td>
<td>MISC12</td>
<td>cap-ox1 genes of yeast mitochondrion.</td>
<td>133</td>
</tr>
<tr>
<td>gb</td>
<td>L15359</td>
<td>YSJATPTRN</td>
<td>Yarrowia lipolytica mitochondrial AT...</td>
<td>133</td>
</tr>
<tr>
<td>gb</td>
<td>M29112</td>
<td>DDIACTD</td>
<td>D.discoideum actin A-3-sub-1 gene, 5...</td>
<td>107</td>
</tr>
<tr>
<td>emb</td>
<td>Z47070</td>
<td>CET09B9</td>
<td>Caenorhabditis elegans cosmid T09B9</td>
<td>133</td>
</tr>
<tr>
<td>emb</td>
<td>Z19055</td>
<td>BATRYOPEA</td>
<td>B.aphidicola tryptophan operon</td>
<td>115</td>
</tr>
<tr>
<td>gb</td>
<td>U67549</td>
<td>MJU67549</td>
<td>Methanococcus jannaschii from bases ...</td>
<td>116</td>
</tr>
<tr>
<td>emb</td>
<td>Z69883</td>
<td>CEC27C12</td>
<td>Caenorhabditis elegans cosmid C27C12</td>
<td>118</td>
</tr>
<tr>
<td>gb</td>
<td>J02451</td>
<td>PFDCG</td>
<td>Bacteriophage fd, strain 478, comple...</td>
<td>132</td>
</tr>
<tr>
<td>gb</td>
<td>U09184</td>
<td>BAU09184</td>
<td>Buchnera aphidicola anthranilate syn...</td>
<td>114</td>
</tr>
<tr>
<td>gb</td>
<td>U32857</td>
<td>SCU32857</td>
<td>Saccharomyces cerevisiae VAR1 gene, ...</td>
<td>131</td>
</tr>
<tr>
<td>gb</td>
<td>M88321</td>
<td>COTLEA14A</td>
<td>Gossypium hirsutum group 4 late embr</td>
<td>131</td>
</tr>
<tr>
<td>emb</td>
<td>Z71180</td>
<td>CEF22E12</td>
<td>Caenorhabditis elegans cosmid F22E12</td>
<td>131</td>
</tr>
<tr>
<td>gb</td>
<td>M55332</td>
<td>DDIDISIG</td>
<td>D.discoideum discoidin I gamma gene,...</td>
<td>162</td>
</tr>
<tr>
<td>dbj</td>
<td>D50338</td>
<td>DDIRIPR</td>
<td>Slime mold DNA for ribosomal protein.</td>
<td>129</td>
</tr>
<tr>
<td>gb</td>
<td>U40411</td>
<td>CELB0403</td>
<td>Caenorhabditis elegans cosmid B0403</td>
<td>167</td>
</tr>
<tr>
<td>emb</td>
<td>X00325</td>
<td>DDGAMMA1</td>
<td>Dictyostelium discoideum discoidin I-</td>
<td>162</td>
</tr>
<tr>
<td>emb</td>
<td>X85120</td>
<td>ASPDXD3H</td>
<td>Artificial sequences cloning vector</td>
<td>162</td>
</tr>
<tr>
<td>emb</td>
<td>X85121</td>
<td>ASPDXD3C</td>
<td>Artificial sequences cloning vector</td>
<td>162</td>
</tr>
<tr>
<td>gb</td>
<td>U56960</td>
<td>CELF25H10</td>
<td>Caenorhabditis elegans cosmid F25H10</td>
<td>157</td>
</tr>
<tr>
<td>gb</td>
<td>L23077</td>
<td>RATZFP15A</td>
<td>Rattus rattus zinc finger protein, c...</td>
<td>126</td>
</tr>
<tr>
<td>emb</td>
<td>X74778</td>
<td>HSREPC4</td>
<td>H.sapiens simple repeat region (clon...</td>
<td>154</td>
</tr>
<tr>
<td>emb</td>
<td>Z49098</td>
<td>CEC07E3</td>
<td>Caenorhabditis elegans cosmid C07E3</td>
<td>131</td>
</tr>
<tr>
<td>gb</td>
<td>U38750</td>
<td>ESU38750</td>
<td>Ectocarpus siliculosus ribulose bisp...</td>
<td>151</td>
</tr>
<tr>
<td>gb</td>
<td>L04272</td>
<td>MSQNCATR</td>
<td>Anopheles quadrimaculatus NADH de</td>
<td>133</td>
</tr>
<tr>
<td>gb</td>
<td>U01010</td>
<td>U01010</td>
<td>Herpetomonas megaseliae kinetoplast</td>
<td>131</td>
</tr>
<tr>
<td>gb</td>
<td>U41029</td>
<td>CELF47G3</td>
<td>Caenorhabditis elegans cosmid F47G3.</td>
<td>128</td>
</tr>
<tr>
<td>gb</td>
<td>U02970</td>
<td>PWU02970</td>
<td>Protothea wickerhamii 263-11 comple...</td>
<td>126</td>
</tr>
<tr>
<td>gb</td>
<td>U64603</td>
<td>CELC09B7</td>
<td>Caenorhabditis elegans cosmid C09B7.</td>
<td>130</td>
</tr>
<tr>
<td>gb</td>
<td>U60168</td>
<td>DDU0168</td>
<td>Dictyostelium discoideum proteasome</td>
<td>145</td>
</tr>
<tr>
<td>gb</td>
<td>L39752</td>
<td>DRO29DC3Z</td>
<td>Drosophila melanogaster (subclone D2...</td>
<td>145</td>
</tr>
<tr>
<td>emb</td>
<td>Z70204</td>
<td>CEC11G6</td>
<td>Caenorhabditis elegans cosmid C11G6</td>
<td>118</td>
</tr>
<tr>
<td>gb</td>
<td>J01283</td>
<td>DDIDISIB</td>
<td>Slime mold (D.discoideum) discoidin...</td>
<td>144</td>
</tr>
<tr>
<td>gb</td>
<td>J01284</td>
<td>DDIDISIC1</td>
<td>Slime mold (D.discoideum) discoidin...</td>
<td>144</td>
</tr>
<tr>
<td>gb</td>
<td>U55376</td>
<td>CELF16H11</td>
<td>Caenorhabditis elegans cosmid F16H11</td>
<td>116</td>
</tr>
<tr>
<td>Accession</td>
<td>Description</td>
<td>Score</td>
<td>Length</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>-------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>gb</td>
<td>U42846</td>
<td>CELT19D2 Caenorhabditis elegans cosmid T19D2.</td>
<td>143</td>
<td>0.52</td>
</tr>
<tr>
<td>emb</td>
<td>Z70206</td>
<td>CEC49F8 Caenorhabditis elegans cosmid C49F8</td>
<td>152</td>
<td>0.56</td>
</tr>
<tr>
<td>emb</td>
<td>Z77115</td>
<td>CET16A9 Caenorhabditis elegans cosmid T16A9</td>
<td>130</td>
<td>0.59</td>
</tr>
<tr>
<td>gb</td>
<td>L36899</td>
<td>YSCMTCG15 Saccharomyces cerevisiae mitochondria.</td>
<td>142</td>
<td>0.59</td>
</tr>
<tr>
<td>dbj</td>
<td>D31785</td>
<td>HASMT Hansenula wingei mitochondrial DNA.</td>
<td>130</td>
<td>0.62</td>
</tr>
<tr>
<td>gb</td>
<td>L36899</td>
<td>YSCMTCG15 Saccharomyces cerevisiae mitochondria.</td>
<td>142</td>
<td>0.59</td>
</tr>
<tr>
<td>emb</td>
<td>Z69605</td>
<td>CEZK678 Caenorhabditis elegans cosmid ZK678</td>
<td>131</td>
<td>0.63</td>
</tr>
<tr>
<td>emb</td>
<td>X85791</td>
<td>RNLPKG R.norvegicus erythroid-specific enhanc...</td>
<td>141</td>
<td>0.65</td>
</tr>
<tr>
<td>gb</td>
<td>L48488</td>
<td>BORRPOB Borrelia burgdorferi RNA polymerase ...</td>
<td>141</td>
<td>0.65</td>
</tr>
<tr>
<td>gb</td>
<td>U43281</td>
<td>SCU43281 Saccharomyces cerevisiae chromosome ...</td>
<td>141</td>
<td>0.66</td>
</tr>
<tr>
<td>gb</td>
<td>U30821</td>
<td>CPU30821 Cyanophora paradoxa cyanelle, comple...</td>
<td>118</td>
<td>0.71</td>
</tr>
<tr>
<td>emb</td>
<td>X70810</td>
<td>CLEGCGA E.gracilis chloroplast complete genome</td>
<td>111</td>
<td>0.71</td>
</tr>
<tr>
<td>gb</td>
<td>U61957</td>
<td>CELAB7 Caenorhabditis elegans cosmid AC7</td>
<td>126</td>
<td>0.72</td>
</tr>
<tr>
<td>gb</td>
<td>L34637</td>
<td>HUMPECAM07 Homo sapiens platelet/endothelial ce...</td>
<td>140</td>
<td>0.73</td>
</tr>
<tr>
<td>gb</td>
<td>U40414</td>
<td>CELF53B3 Caenorhabditis elegans cosmid F53B3.</td>
<td>121</td>
<td>0.75</td>
</tr>
<tr>
<td>gb</td>
<td>U46668</td>
<td>CELF38E9 Caenorhabditis elegans cosmid F38E9.</td>
<td>110</td>
<td>0.76</td>
</tr>
<tr>
<td>emb</td>
<td>Z68134</td>
<td>CET27A8 Caenorhabditis elegans cosmid T27A8</td>
<td>118</td>
<td>0.78</td>
</tr>
<tr>
<td>emb</td>
<td>X75544</td>
<td>XPFRNAOL P.falciparum gene for beta subunit R...</td>
<td>116</td>
<td>0.78</td>
</tr>
<tr>
<td>emb</td>
<td>Z71517</td>
<td>SCYNL241C S.cerevisiae chromosome XIV reading</td>
<td>138</td>
<td>0.85</td>
</tr>
<tr>
<td>emb</td>
<td>X57336</td>
<td>SCMET19 S. cerevisiae MET19 gene for glucose...</td>
<td>138</td>
<td>0.85</td>
</tr>
<tr>
<td>gb</td>
<td>M34709</td>
<td>YSCG6PD S.cerevisiae glucose-6-phosphate deh...</td>
<td>138</td>
<td>0.85</td>
</tr>
<tr>
<td>emb</td>
<td>Z71518</td>
<td>SCYNL242W S.cerevisiae chromosome XIV reading</td>
<td>138</td>
<td>0.85</td>
</tr>
<tr>
<td>emb</td>
<td>X52177</td>
<td>PFCRPOBC Plasmodium falciparum circular DNA r ...</td>
<td>116</td>
<td>0.85</td>
</tr>
<tr>
<td>emb</td>
<td>Z69381</td>
<td>SCCXIV39K S.cerevisiae 38,855 bp segment of ch...</td>
<td>138</td>
<td>0.85</td>
</tr>
<tr>
<td>gb</td>
<td>U41749</td>
<td>CELF09E10 Caenorhabditis elegans cosmid F09E10.</td>
<td>150</td>
<td>0.86</td>
</tr>
<tr>
<td>gb</td>
<td>U31631</td>
<td>DDU31631 Dictyostelium discoideum class II ap...</td>
<td>122</td>
<td>0.86</td>
</tr>
<tr>
<td>gb</td>
<td>U25815</td>
<td>SCU25815 Spiroplasma citri RNA polymerase bet...</td>
<td>116</td>
<td>0.87</td>
</tr>
<tr>
<td>gb</td>
<td>U34610</td>
<td>MMCOL18A05 Mus musculus alpha-1(XVIII) colla...</td>
<td>137</td>
<td>0.90</td>
</tr>
<tr>
<td>gb</td>
<td>U39690</td>
<td>Mycoplasma genitalium asnS, cfxEc, d...</td>
<td>137</td>
<td>0.90</td>
</tr>
<tr>
<td>emb</td>
<td>Z66499</td>
<td>CET01B7 Caenorhabditis elegans cosmid T01B7</td>
<td>137</td>
<td>0.90</td>
</tr>
<tr>
<td>emb</td>
<td>X66160</td>
<td>PFLSRRN Plasmodium falciparum gene for a lar...</td>
<td>105</td>
<td>0.91</td>
</tr>
<tr>
<td>gb</td>
<td>M35027</td>
<td>VACCG Vaccinia virus, complete genome.</td>
<td>124</td>
<td>0.91</td>
</tr>
<tr>
<td>gb</td>
<td>U31415</td>
<td>PCU31415 Plasmodium chabaudi repeat organella...</td>
<td>138</td>
<td>0.92</td>
</tr>
<tr>
<td>gb</td>
<td>U40410</td>
<td>CELC54G7 Caenorhabditis elegans cosmid C54G7.</td>
<td>131</td>
<td>0.93</td>
</tr>
<tr>
<td>emb</td>
<td>Z50176</td>
<td>CEC09G1 Caenorhabditis elegans cosmid C09G1</td>
<td>130</td>
<td>0.93</td>
</tr>
<tr>
<td>gb</td>
<td>M55298</td>
<td>DDIDDP2 Dictyostelium discoideum plasmid Ddp...</td>
<td>136</td>
<td>0.94</td>
</tr>
<tr>
<td>emb</td>
<td>X51478</td>
<td>DDP2PLAS Dictyostelium discoideum Ddp2 plasm...</td>
<td>136</td>
<td>0.94</td>
</tr>
<tr>
<td>emb</td>
<td>A21622</td>
<td>A21622 Dictyostelium plasmid Ddp2 REP gene</td>
<td>136</td>
<td>0.94</td>
</tr>
<tr>
<td>gb</td>
<td>U41746</td>
<td>CELT18H9 Caenorhabditis elegans cosmid T18H9.</td>
<td>136</td>
<td>0.94</td>
</tr>
<tr>
<td>gb</td>
<td>U20710</td>
<td>ONU20710 Ophraella notulata isolate 126 16S r...</td>
<td>135</td>
<td>0.94</td>
</tr>
<tr>
<td>emb</td>
<td>Z77136</td>
<td>CEZC376 Caenorhabditis elegans cosmid ZC376</td>
<td>120</td>
<td>0.95</td>
</tr>
<tr>
<td>gb</td>
<td>M28500</td>
<td>EUPMIC3 Euplotes crassus G1 micronuclear seq...</td>
<td>114</td>
<td>0.95</td>
</tr>
<tr>
<td>emb</td>
<td>X95275</td>
<td>PFCOMPIRA P.falciparum complete gene map of pl ...</td>
<td>105</td>
<td>0.95</td>
</tr>
<tr>
<td>gb</td>
<td>U01011</td>
<td>U01011 Herpetomonas muscarum muscarum kinet.</td>
<td>135</td>
<td>0.96</td>
</tr>
<tr>
<td>Accession</td>
<td>Description</td>
<td>Length</td>
<td>Identity</td>
<td>Similarity</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------------------------------------------------------------</td>
<td>--------</td>
<td>----------</td>
<td>------------</td>
</tr>
<tr>
<td>Z74038</td>
<td>Caenorhabditis elegans cosmid F58B4</td>
<td>121</td>
<td>0.96</td>
<td>3</td>
</tr>
<tr>
<td>X15917</td>
<td>Paramecium aurelia mitochondrial com</td>
<td>135</td>
<td>0.97</td>
<td>1</td>
</tr>
<tr>
<td>U67549</td>
<td>Methanococcus jannaschii from bases</td>
<td>116</td>
<td>0.98</td>
<td>3</td>
</tr>
<tr>
<td>M29112</td>
<td>D.discoideum actin A-3-sub-1 gene,</td>
<td>107</td>
<td>0.98</td>
<td>2</td>
</tr>
<tr>
<td>Z19055</td>
<td>B.aphidicola tryptophan operon</td>
<td>115</td>
<td>0.98</td>
<td>2</td>
</tr>
<tr>
<td>X06424</td>
<td>Yeast mit DNA for promoter upstream</td>
<td>133</td>
<td>0.99</td>
<td>1</td>
</tr>
<tr>
<td>K00539</td>
<td>Yeast (S.cerevisiae) mitochondrial c...</td>
<td>133</td>
<td>0.99</td>
<td>1</td>
</tr>
<tr>
<td>V00693</td>
<td>cap-ox1 genes of yeast mitochondron.</td>
<td>133</td>
<td>0.99</td>
<td>1</td>
</tr>
<tr>
<td>Z69883</td>
<td>Caenorhabditis elegans cosmid C27C12</td>
<td>118</td>
<td>0.99</td>
<td>2</td>
</tr>
<tr>
<td>L36887</td>
<td>Saccharomyces cerevisiae mitochondrial c...</td>
<td>133</td>
<td>0.99</td>
<td>1</td>
</tr>
<tr>
<td>L15359</td>
<td>Yarrowia lipolytica mitochondrial AT...</td>
<td>133</td>
<td>0.99</td>
<td>1</td>
</tr>
<tr>
<td>Z47070</td>
<td>Caenorhabditis elegans cosmid T09B9</td>
<td>133</td>
<td>0.99</td>
<td>1</td>
</tr>
<tr>
<td>J02451</td>
<td>Bacteriophage fd, strain 478, comple...</td>
<td>132</td>
<td>0.99</td>
<td>1</td>
</tr>
<tr>
<td>V00602</td>
<td>Genome of the bacteriophage fd (Inov...</td>
<td>132</td>
<td>0.99</td>
<td>1</td>
</tr>
<tr>
<td>Z33164</td>
<td>M.capricolum DNA for CONTIG MC228</td>
<td>131</td>
<td>0.99</td>
<td>1</td>
</tr>
<tr>
<td>Z32679</td>
<td>Caenorhabditis elegans cosmid C05B5</td>
<td>127</td>
<td>0.99</td>
<td>3</td>
</tr>
<tr>
<td>X03991</td>
<td>Human glucagon gene</td>
<td>129</td>
<td>0.99</td>
<td>2</td>
</tr>
<tr>
<td>U67599</td>
<td>Methanococcus jannaschii from bases</td>
<td>118</td>
<td>0.99</td>
<td>2</td>
</tr>
<tr>
<td>Z68160</td>
<td>Caenorhabditis elegans cosmid D1046</td>
<td>117</td>
<td>0.99</td>
<td>3</td>
</tr>
<tr>
<td>U32857</td>
<td>Saccharomyces cerevisiae VAR1 gene,</td>
<td>131</td>
<td>0.99</td>
<td>1</td>
</tr>
<tr>
<td>U40413</td>
<td>Caenorhabditis elegans cosmid C18B2</td>
<td>120</td>
<td>0.99</td>
<td>3</td>
</tr>
<tr>
<td>Z72514</td>
<td>Caenorhabditis elegans cosmid T10B10</td>
<td>140</td>
<td>0.99</td>
<td>3</td>
</tr>
<tr>
<td>M88321</td>
<td>Gossypium hirsutum group 4 late embra</td>
<td>131</td>
<td>0.99</td>
<td>1</td>
</tr>
<tr>
<td>X75545</td>
<td>P.falciparum gene for tRNA I,A,N,L,R...</td>
<td>105</td>
<td>0.99</td>
<td>3</td>
</tr>
<tr>
<td>Z71180</td>
<td>Caenorhabditis elegans cosmid F22E12</td>
<td>131</td>
<td>0.99</td>
<td>1</td>
</tr>
<tr>
<td>U55374</td>
<td>Caenorhabditis elegans cosmid T02C5</td>
<td>116</td>
<td>0.99</td>
<td>2</td>
</tr>
<tr>
<td>Z70207</td>
<td>Caenorhabditis elegans cosmid F15A2</td>
<td>136</td>
<td>0.99</td>
<td>2</td>
</tr>
<tr>
<td>X53420</td>
<td>A.geoffroyi gamma-globin gene and L1</td>
<td>130</td>
<td>0.99</td>
<td>1</td>
</tr>
<tr>
<td>L06178</td>
<td>Apis mellifera mitochondrial genome.</td>
<td>115</td>
<td>0.99</td>
<td>4</td>
</tr>
<tr>
<td>Z54306</td>
<td>Caenorhabditis elegans cosmid B0457</td>
<td>130</td>
<td>0.99</td>
<td>1</td>
</tr>
<tr>
<td>U00054</td>
<td>Caenorhabditis elegans cosmid K07E12</td>
<td>131</td>
<td>0.99</td>
<td>3</td>
</tr>
<tr>
<td>Z73978</td>
<td>Caenorhabditis elegans cosmid ZC302</td>
<td>129</td>
<td>0.99</td>
<td>2</td>
</tr>
</tbody>
</table>
APPENDIX IV

Sample output of TESS analysis of a 5' sub-region upstream of bli-4

pCeh288:

00001
CTACTCTCAACAAGAGATATTTCTGTTCACTATTTTTTTTTCTCTGAACTT
    -----> LVa (12.00 2.0000)
    <------- GR (16.00 2.0000)
    <------- PR (16.00 2.0000)
        <------- Hb (16.00 1.6000)
        <------- Ste11 (15.00 1.5000)
            <----- GR (12.00 2.0000)

00051
CTGTGACTGAATGCACTCTCGGCACCAAATGAGAGATGCCTACGAGG
    -----> GCN4 (12.00 2.0000)
        <-----> NF-GMa (12.00 1.2000)

00101
GCGCTACTCTCCTTTTACCTACATGTATGTCTACCGAAACAAAACAATATG
        <------> E1A-F (14.00 2.0000)
        <------> Elk-1 (13.00 1.3000)
        <------ Tfl-LF1 (16.00 1.6000)
            <------> Ste11 (15.00 1.5000)
            <-----> HNF-3 (12.00 1.7143)

00151
AGGCAAATTTTTCATGGCTTTTGCTAATTTTACACTTTCCACTCCCTGTTAC
        <------> CCBF (14.00 1.1667)
        <------> Pit-1 (14.00 2.0000)
            <-------- HiNF-A (12.00 1.0000)
            <-------- HNF-4 (14.00 1.1667)
                <-----> PRDI-BF1 (16.00 1.6000)
                -----> IRF-1 (12.00 2.0000)
                -----> IRF-2 (12.00 2.0000)

00201
AAGCATATTATTTAGGTCAGGTGCTTTGCTAATTCAATTTTCAATTTTGAG
        <----- PPAR (12.00 2.0000)
            <-----> Rad-1 (14.00 2.0000)
            -----< c-Myb (12.00 2.0000)
            <-------- HiNF-A (12.00 1.0000)
                <----- GATA-1 (12.00 2.0000)
00251
AATTTGTAAACTCGTCTAGTAATTTAAAAACTCATTGATTTCTGTATAAT
        ------> MEF-2 (12.00 1.2000)
        <------ Oct-1 (16.00 1.6000)
        <------ Oct-4 (16.00 1.6000)
            <------ E4BP4 (13.00 1.3000)
            <------ NF-IL-2A (16.00 1.6000)
            <------ Oct-1 (16.00 1.6000)
            <------ Oct-2 (16.00 1.6000)

00301
GTAACACAAAAATCGAAATTCAGGTAAGCTCGATTATCCGGTCCCGACAA
        <----- Pit-1 (14.00 2.0000)
        <----- NF-GMb (13.00 1.8571)
            <------ Hox-1.3 (14.00 1.4000)
            <------ GATA-1 (16.00 2.0000)

00351
GACAAATTTGTGTTAATGCAAAAAGAGTATGCGTCTTCAAAGAGTACTGT
        -----> Pit-1 (16.00 2.0000)
        <------ TFIID (12.00 2.0000)

00401
AGTTTCTACCTCTGGTCTTCTTTAATAGGTTTCTTTTTCTCCGAT
        -----> c-Myb (12.00 2.0000)
        -----> dl (12.00 1.2000)

00451
TCCTAAAAATTGCATGTTTTTAGTTTTTCAGTAGATTAAATTCC
        <------- MEF-2 (12.00 1.2000)
        -----> SEF4 (12.00 1.7143)
        <------- Pit-1 (16.00 1.6000)
        -----> USF (12.00 2.0000)
        <------ muEBP-C2 (12.00 2.0000)
        <------- HiNF-A (12.00 1.0000)
        -----> Pit-1 (14.00 2.0000)

00501
GTTTTCATCGAAAAACTCCGTATCAAAGATCAATGAAGTTCCCAGCTA
00551
CGAGGAGTTCAGGTACTCTTTAAAGGGCGACACCCATTCACAT
        -----> MEF-2 (12.00 1.2000)

00601
TTAACAGAACATTTTCGTGATACCTGACCCGCTACCCTATATATAGGGGGCAAA
        -----> AR (12.00 2.0000)
        -----> PR (12.00 2.0000)
        <------ GR (14.00 2.0000)
            <------- c-Ets-1 (16.00 1.6000)
00651
ATGTTGCATTTAAACGCTCCTCATCGTCAATTTTCCCTGTCTAATTATT

00701
TAGAGTGCTTTTAAGTTAACTTTCATGAAAATATTATTAGAGAGTGACTGTG

00751
ATGGGCAAAACAGTCTAACGGAACAATCTTTCATTCTCGACAAATTGCTGA

00801
GTCTCGCCACGAACAAAGTCAGAAGACATAGCATATTCTTGAGAAAAGCCT

00851
TCAGTCGATTTTCATTTCCATGTTCTCTCCTCTCACACGCCCTTCAGCA

00901
TCTTCTTCTCTCGCCACAAGCTTCTCATTTTCAAATCGAATAAATAAAT
NNPP analyses of the 5’ region upstream of bli-4

pCeh288:

<table>
<thead>
<tr>
<th>Start</th>
<th>End</th>
<th>Score</th>
<th>Promoter Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>566</td>
<td>616</td>
<td>0.98</td>
<td>TACAGTACTCTTTAAAGGCGCACACCCATTCACATTTAAACAGAACATTTT</td>
</tr>
<tr>
<td>626</td>
<td>676</td>
<td>0.98</td>
<td>CCGGCTACCGTATATATGGGGCAAAATGTTGCATTTTAACGCTCCTCATC</td>
</tr>
<tr>
<td>932</td>
<td>982</td>
<td>0.90</td>
<td>TCCAATCGAATAATATCGTGTCTGGAAATCCCATG</td>
</tr>
</tbody>
</table>

pCeh289:
<table>
<thead>
<tr>
<th>Start</th>
<th>End</th>
<th>Score</th>
<th>Promoter Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993</td>
<td>2043</td>
<td>0.84</td>
<td>TTCGACTGCCTAATCTGCAATTTCTGGAATCTACTTAGATTACT</td>
</tr>
<tr>
<td>2195</td>
<td>2245</td>
<td>0.87</td>
<td>TGAGATATAATCTCAAAAAATTCCTCTTTTCCCCTAATCAATTGTGAAT</td>
</tr>
<tr>
<td>2275</td>
<td>2325</td>
<td>0.97</td>
<td>GTGATCTGTTGTTAATCTGGTGCAATATCAGTTTCTGATGTCAATGTT</td>
</tr>
<tr>
<td>2405</td>
<td>2455</td>
<td>0.91</td>
<td>GGAAACCTATATAATAATCTGGTTAATAATCTGCTGTTTTTTTTATCTAG</td>
</tr>
<tr>
<td>2823</td>
<td>2873</td>
<td>0.97</td>
<td>TATTTAATCTATTTAAAAATTCAGCGAGCGAGGTAAAGATTAGTCATGGTA</td>
</tr>
<tr>
<td>2872</td>
<td>2922</td>
<td>0.85</td>
<td>ATTTCCCCCATTTAAAATCTCGACGCATCTCTTTCTTCTTTACCTTT</td>
</tr>
</tbody>
</table>

**pCeh291:**

<table>
<thead>
<tr>
<th>Start</th>
<th>End</th>
<th>Score</th>
<th>Promoter Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993</td>
<td>2043</td>
<td>0.84</td>
<td>TTCGACTGCCTAATCTGCAATTTCTGGAATCTACTTAGATTACT</td>
</tr>
<tr>
<td>2195</td>
<td>2245</td>
<td>0.87</td>
<td>TGAGATATAATCTCAAAAAATTCCTCTTTTCCCCTAATCAATTGTGAAT</td>
</tr>
<tr>
<td>2275</td>
<td>2325</td>
<td>0.97</td>
<td>GTGATCTGTTGTTAATCTGGTGCAATATCAGTTTCTGATGTCAATGTT</td>
</tr>
</tbody>
</table>