Temporal and spatial expression patterns of the *hsp16* and *ubq-1* genes in transgenic *C. elegans*.

by

Eve Gabrielle Stringham Durovic

B.Sc., The University of Manitoba, 1984
M.Sc., The University of Manitoba, 1987

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

Department of Biochemisty-Genetics Program

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

November, 1991

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.

(Signature)

Department of Biochemistry

The University of British Columbia
Vancouver, Canada

Date Jan. 7/92

DE-6 (2/88)

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.

Biochemisty-Genetics Program

The University of British Columbia
2075 Wesbrook Place
Vancouver, Canada
V6T 1W5

Date: 22/11/91

ABSTRACT

The expression of the small 16 kDa heat shock protein gene (*hsp16*) family and of the polyubiquitin encoding gene (*ubq-1*) in *Caenorhabditis elegans* has been examined by introducing *lacZ* fusions into the nematode by transformation.

Transcription of the *hsp16-lacZ* transgenes was totally heat shock dependent and resulted in the rapid synthesis of detectable levels of β-galactosidase in most somatic tissues. Although the two *hsp16* gene pairs of *C. elegans* are highly similar within both their coding and non-coding sequences, quantitative and qualitative differences in the spatial pattern of expression between gene pairs were observed. The *hsp16-48* promoter was shown to direct greater expression of β-galactosidase in muscle and hypodermis while the *hsp16-41* promoter was more efficient in intestine and pharyngeal tissue. Transgenes which eliminated one promoter from a gene pair were expressed at reduced levels, particularly in post-embryonic stages, suggesting that the heat shock elements (HSEs) in the intergenic region of an *hsp16* gene pair may act cooperatively to achieve high levels of expression of both genes. The *hsp16* gene pairs are never constitutively expressed. In addition, their heat inducibility is developmentally restricted; they are not heat inducible during gametogenesis or early embryogenesis. The *hsp16* genes represent the first fully inducible system in *C. elegans* to be characterized in detail at the molecular level, and the promoters of these genes should be useful for studying the action of tissue or developmentally regulated genes in this organism.

Animals carrying a translational *ubq-1* construct consisting of 938 bp of *ubq-1* upstream sequences fused to *lacZ* (*ubq938-lacZ*) expressed β -galactosidase in embryos and in a tissue general manner in 20% of staining L1 larvae. Somatic expression in later stages was usually confined to body muscle. Progressively larger deletions extending from the 5' end of *ubq938-lacZ* did not significantly alter the pattern of expression until 827 bp of sequence had been removed. Thus sequences upstream of the transcriptional start site, including a G/C rich block and a sequence resembling a TATA box (GAATAA) are not required for *ubq-1* expression.

Moreover, a basal level of expression was maintained in embryos when 903 bp had been deleted. These results suggest that some of the regulatory elements required for efficient expression of *ubq-1* may reside within the transcribed region of the gene; alternatively they must lie more than 1.7 kb upstream or 0.8 kb downstream of this region. PCR analysis indicates that RNA molecules transcribed from the *ubq938-lacZ* and *ubq 827-lacZ* transgenes are trans-spliced to SL1, as is *ubq-1* RNA.

TABLE OF CONTENTS

Abstract ii

Table of Contents iv

List of Figures ix

List of Tables xi

List of Abbreviations xii

Acknowledgements xv

Dedication xvi

Chapter I: Analysis of *hsp16* Expression in *C. elegans* 1

A. Introduction 1

1. General Introduction 1

2. Events in the Stressed Cell 2

3. Functions of the Heat Shock Proteins 3

3.1 HSP110 3

3.2 HSP90 4

3.3 HSP70 5

3.4 HSP60 7

3.5 Low MW HSPs 8

4. Heat Shock Proteins in Development, Differentiation and Growth 10

4.1 Developmental Control of the Heat Shock Response in *Xenopus* .. 10

4.2 Developmental Regulation in *Drosophila* 11

4.3 Developmental Regulation of *hsp70* 12

5. The Role of Low MW HSPs in Thermotolerance 14

6. Regulation of the Heat Shock Response 15

6.1 Transcriptional Control in Eukaryotes 16

6.2 Post-transcriptional Regulation 19

6.3 Translational Control....21

6.4 Recovery from Heat Shock....23

7. *Caenorhabditis elegans* as a model system....25

8. The Heat Shock Response of *C. elegans*....26

8.1 The Small HSPs of *C. elegans*....26

9. The Present Study....30

B. Materials and Methods....31

1. Construction of *hsp16-lacZ* Fusions....31

2. Bacterial Transformations....32

3. Double Stranded Sequencing of *lacZ* Fusions....32

4. Maintenance of Strains....33

5. Establishment of Transgenic *C. elegans* Strains....34

6. Selection of Transformed Progeny....34

7. Viable Freezing of Transgenic Strains....35

8. Heat Shock Conditions....35

9. Identification of β-galactosidase Staining Cells....36

10. Preparation of Transgenic Genomic DNA....36

11. Southern Transfer and Analysis of Transgenic DNAs....37

12. DNA Dot Blot Procedures....37

13. Labelling of Radioactive Probes....38

14. Hybridization Conditions....39

C. Results....40

1. Construction of *hsp16-lacZ* Fusions and Selection of Transformants....40

2. Expression of the *hsp16-lacZ* Transgene is Temperature Dependent....43

3. Establishing Standard Heat Shock Conditions for Experiments....46

4. The Transgenes 48.1C and 41.2C are Expressed in a Tissue General Manner

....46

5. Quantitative Differences in the Tissue Specific Expression of the *hsp16* Gene Pairs 54

6. Elimination of One Promoter from an *hsp16* Gene Pair Significantly Reduces Somatic Tissue Expression Without Affecting Embryonic Expression 60

7. Southern Analysis of Transgenic Strains 64

8. Determination of *lacZ* Transgene Copy Number 69

D. Discussion 74

1. The *lacZ* Trangenes are Correctly Expressed in Response to Heat Shock *in vivo* 74

2. Transgene Copy Number *versus* Expression Levels 74

3. Expression of the *hsp16* Gene Pairs is Tissue General 75

4. Tissue Differences in the Expression of *hsp16-lacZ* Transgenes 76

5. Cooperative Interaction of HSEs May Enhance Expression 77

6. Developmental Regulation of the *hsp16s* in *C. elegans* 79

7. Conclusions 80

8. Future Prospects 81

Chapter II: Analysis of Polyubiquitin Gene (*ubq-1*) Expression 84

A. Introduction 84

1. General Introduction 84

2. Ubiquitin Mediated Proteolysis 84

3. Ubiquitin as a Regulatory Protein 86

4. Ubiquitin and Heat Shock 87

5. Ubiquitin and Chromatin 87

6. Ubiquitin at the Cell Surface 89

7. Ubiquitin and Myofibril Assembly 89

8. Ubiquitin Gene Structure 89

8.1 Polyubiquitin Gene Structure 89

8.2 Polyubiquitin Gene Expression 90

8.3 Ubiquitin Fusion Genes 93

8.4 Ubiquitin Like Genes 94

9. Trans-splicing 94

10. The Present Study 95

B. Methods 98

1. Maintenance of Strains 98

2. Construction of *ubq-1-lacZ* Fusions 98

3. Establishment of Transgenic Strains 98

4. Identification of β-galactosidase Staining Cells 98

5. Heat Shock Conditions 99

6. Preparation of RNA 99

7. Preparation of cDNA 99

8. Polymerase Chain Reactions 100

9. Separation and Analysis of PCR Products 100

10. Southern Analysis of PCR Products 101

C. Results 102

1. Construction of *ubq-1-lacZ* Strains 102

2. Expression of the *ubq938-lacZ* Transgene is Constitutive but Shows Developmental Tissue Specificity 104

3. Expression is not Diminished until 827 bp of Sequence has been Deleted from *ubq938-lacZ* 107

4. The *ubq938-lacZ* and *ubqΔ827-lacZ* Transcripts are Trans-spliced 112

D. Discussion 116

1. Expression of *ubq-1-lacZ* Transgenes in Nematodes 116

2. *Ubq-1* Expression is not Significantly Heat Inducible 118

3. Trans-splicing of *ubq-1-lacZ* Transcripts 119

4. Analysis of the *ubq-1* Promoter 119

5. Conclusions 121

6. Future Prospects 121

References 123

Appendix 136

LIST OF FIGURES

Fig. 1 Organization of the *hsp16* loci of *C. elegans* 27

Fig. 2 Construction of the *hsp16-lacZ* transgenes 41

Fig. 3 Heat shock dependence of *lacZ* expression 44

Fig. 4 Expression of β-galactosidase is proportional to temperature in *hsp16-lacZ* transgenic animals 47

Fig. 5 *LacZ* expression of the 48.1C and 41.2C transgenes in response to heat is tissue general 50

Fig. 6 Cell types expressing *hsp16-lacZ* transgenes upon heat shock 52

Fig. 7 Expression of *hsp16-exon1* fusions 55

Fig. 8 Expression of the 48P transcriptional fusion in heat shocked transgenic animals 61

Fig. 9 Southern analysis of Sma I/Apa I digested genomic transgenic strain DNAs probed with nick translated pPD16.43 65

Fig. 10 Southern analysis of transgenic genomic DNA probed with primer extended *lacZ* products 67

Fig. 11 Dot blot analysis of transgene copy number 70

Fig. 12 The non-lysosomal ATP dependent proteolytic pathway 85

Fig. 13 Organization, transcription and processing of the polyubiquitin gene, *ubq-1* of *Caenorhabditis elegans* 96

Fig. 14 Construction of *ubq-1-lacZ* fusions 103

Fig. 15 *In situ* staining of β-galactosidase activity in PC36 105

Fig. 16 Deletion analysis of the *ubq938-lacZ* transgene 108

Fig. 17 Expression of *ubq-1-lacZ* deletion transgenes 109

Fig. 18 Amplification of trans-spliced *ubq-1-lacZ* RNA 113

Fig. 19 Testing the integrity of the PCR amplification scheme 114

Fig. 20 Analysis of PCR products amplified from *ubq-1-lacZ* RNA 115

LIST OF TABLES

Table 1. Tissue distribution of β-galactosidase staining in transgenic *C. elegans* heat shocked for 15-120' at 33°C 58

Table 2. Distribution of β-galactosidase activity in trangenic strains carrying transcriptional fusions.....63

Table 3. Determination of *lacZ* trangene copy number in transgenic strains selected for extensive analysis 71

Table 4. Comparison of expression patterns between 1.48E1 strains 72

Table 5. Comparison of polyubiquitin genes among several species 91

LIST OF ABBREVIATIONS

AMP	adenosine-5'-monophosphate
ATP	adenosine-5"-triphosphate
bp	base pair(s)
10 % Blotto	0.1 g/ml skim milk powder (Carnation); 0.2 % Na azide
BSA	bovine serum albumin
C-terminus	carboxy terminus
cDNA	DNA complementary to coding strand
C. elegans	*Caenorhabditis elegans*
cpm	counts per minute
DAPI	4',6-diamidino-2-phenyl-indole
1 % Denhardt's	0.01 g/ml of each of BSA, Ficoll, PVP, SDS
DNA	deoxyribonucleic acid
dATP	deoxyadenosine-5'-triphosphate
dCTP	deoxycytidine-5'-triphosphate
dGTP	deoxyguanosine-5'-triphosphate
dH_2O	distilled water
dTTP	deoxythymidine-5'-triphosphate
ddATP	dideoxyadenosine-5'-triphosphate
ddCTP	dideoxycytidine-5'-triphosphate
ddGTP	dideoxyguanosine-5'-triphosphate
ddTTP	dideoxythymidine-5'-triphosphate
E1	ubiquitin activating enzyme
E2	ubiquitin conjugating enzyme
E3	ubiquitin ligase
E. coli	*Escherichia coli*

EDTA	ethylenediamine tetraacetic acid
HSC(s)	heat shock cognate protein(s)
hsc	heat shock cognate protein gene
HSP(s)	heat shock protein(s)
hsp	heat shock protein gene
HSF	heat shock transcription factor
HSE	heat shock element
kb	kilobase pair (s)
kDa	kilodaltons
MDa	mega daltons
mRNA	messenger RNA
mHSP	mitochondrial HSP
MMLV	Moloney murine leukemia virus
M9	3 % KH_2PO_4; 6 % Na_2HPO_4; 5 % NaCl(w/v); 1 mM $MgSO_4$
N	any base (i.e.adenine, cytidine, guanine, thymidine)
N-terminus	amino terminus
NGM	nutrient growth media
nt	nucleotide(s)
PCR	polymerase chain reaction
PPi	pyrophosphate
PVP	polyvinyl pyrolidine
RCF	relative centrifugal force
RNA	ribonucleic acid
RNase	ribonuclease
S. cerevisiae	*Saccharomyces cerevisiae*
20 X SSC	3.6 M NaCl; 0.3 M Na Citrate, pH 7.0
SDS	sodium dodecyl sulphate

SL	splice leader
SSPE	180 mM NaCl; 1 mM EDTA; 10 mM NaH_2PO4, pH7.4
TBE	90 mM Tris-borate pH 8.3; 2 mM EDTA
Taq	*Thermus aquaticus*
TE	10 mM Tris-HCl, pH 8.0; 1 mM EDTA
Tris	tris (hydroxymethyl) aminomethane
UBQ	ubiquitin protein monomer
Xgal	5-bromo 4-chloro 3-indolylgalactosidase

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor Peter Candido for his enthusiastic support during this endeavour. In particular, I am extremely grateful that he gave me the opportunity and lattitude to pursue the project of my choice even though success was not guaranteed and the necessary equipment was not initially available. Four years and a microinjection set-up later, I can honestly say that it was worth it. I would also like to thank the other members of my supervisory committee- Don Moerman, Ross MacGillivray and Tom Grigliatti-for their well-respected advice and counsel, lively committee meetings (the beer and cookies helped I'm sure) and for reading my thesis so quickly without complaining too much.

To my collaborators, Don Jones and Dennis Dixon, thanks for sharing constructs, advice and imagery from Thomas Hardy novels! To all the members of the Candido lab who have come and gone during my internment at this institution (Don, Roger, Esther, Robert, Mike, Dave, Josie, Mei and Michel) a hearty thanks for enduring the spontaneous yelps of "kill, kill, kill" during microinjection sessions, the pictures of blue worms plastered about the lab, the late night opera singing and the "question of the day".

To all my friends from the Biochemistry department and in other programs, thanks for all the intriguing and boisterous conversations and for letting me "borrow" stuff. You've given me a new perspective on life and science.

I am grateful to Andrew Fire (Carnegie Institute in Baltimore) for donating nematode expression vectors to the cause and to Jim Preiss (Fred Hutchinson Institute in Seattle) for an excellent tutorial on cellular identification in *C. elegans*.

My warmest thanks to the members of the office staff for their patient printing of my thesis and papers, and for dealing on my behalf all too many times with irate financial personnel etc.

I would like to thank my family for their support in this venture. In particular my father who as a Professor in Animal Science first introduced me to Genetics and my five elder brothers for setting such "shining examples". Last but not least, I would like to acknowledge the loving support and numerous cups of tea my husband Peter has given me throughout this degree. Thank you for believing in me.

DEDICATION

This thesis is dedicated to the memory of my mother, Gabrielle Mary Wellington Stringham, who died of cancer on August 24, 1989 when this work was still in progress. She wasn't a scientist but she believed that every individual has an obligation to make this world a better place for others. This is the ideal scientists must strive for. Let us collaborate- not compete, let us remember who we work for, why we work and that we are but tiny players in a much greater universe. Thank you Mum for making my world definitely a better place.

CHAPTER I: ANALYSIS OF HSP16 EXPRESSION IN C. ELEGANS

A: INTRODUCTION

1. General Introduction

The heat shock or stress reponse is a universal phenomenon characterized by the induction of a unique set of polypeptides called the heat shock proteins (HSPs) coupled with the simultaneous repression of normal cellular protein synthesis. Golschmidt, in 1935, suggested that heat shock alters the pattern of gene expression: this was experimentally substantiated when Ritossa documented that elevation in temperature from 25°C to 37°C induced the formation of specific chromatin puffs in the giant chromosomes of the salivary glands in *Drosophila* (Ritossa, 1962). Later studies showed that these puffs represent sites of *de novo* RNA synthesis (Tissières et al., 1974) and the link between heat shock and gene regulation was firmly established.

Heat shock in *Drosophila* induces the formation of eight distinct polypeptides (Moran et al., 1978) which range in mass from 22 to 83 kD. Generally, the HSPs can be subdivided into families on the basis of their size as determined by migration in one or two dimensional gels: the 110 kD HSPs, the 80-90 kD HSPs, the 70kD HSPs, the low molecular weight (20-30 kD) HSPs, and ubiquitin (8kD). Ubiquitin is the subject of Chapter II and will not be discussed further here. These proteins are highly conserved among both prokaryotic and eukaryotic species. For example human HSP70 shares 50 % sequence identity with HSP70 from *E. coli*, with the similarity reaching 96 % in some domains of the protein (Bardwell and Craig, 1984). This conservation suggests that HSPs have an essential protective role in all cells. However, even HSPs appear to be limited in their capacity to rescue a cell from thermal stress, since prolonged heat shock will eventually result in cell death.

Since the discovery of the heat shock reponse, it has been shown that a variety of other agents evoke similar changes in gene expression culminating in the repression of normal

protein synthesis and the induction of HSPs. These inducers include heavy metals, arsenite, ethanol, amino acid analogues, certain ionophores, mitochondrial inhibitors and viral infection (Ashburner and Bonner, 1979; Thomas et al., 1982). Thus the phenomenon is now more accurately referred to as the stress response, even though heat shock is the most frequent method employed by investigators to elicit the response.

2. Events in the Stressed Cell

The physiology and biochemistry of the cell undergo dramatic changes when confronted by stress. Firstly, there is a cessation in cellular growth, the pH plummets, and there is an increase in the levels of cytosolic calcium (Hammond et al., 1982; Findly et al., 1983). In the nucleus, protein and nucleic acid complexes become insoluble (Littlewood et al., 1987) resulting in aggregation of granular components (Welch and Suhan, 1985). In addition rod-like actin-containing filaments form and cross the nucleus (Welch and Suhan, 1985). Within the cytoplasm the proportion of actin filaments increases while the intermediate filament network collapses around the nucleus (Falkner et al., 1981). Mitochondria swell and intramembranous proteins denature (Lepock et al., 1983; Welch and Suhan, 1985). Ribosomes and mitochondria move to the perinuclear region and associate with the intermediate filaments (Welch and Suhan, 1986). The Golgi complex fragments into separate vesicles which become dispersed throughout the cytoplasm (Welch and Suhan, 1985).

Heat shock also forces a change in energy metabolism. Stressed cells tend to shift from aerobic to glycolytic pathways for energy production as the intracellular levels of ATP drop rapidly due to mitochondrial dysfunction (Findly et al., 1983; Christiansen and Kvamme, 1969; Mondovi et al., 1969; Dickson and Oswald, 1976). Lactic acid accumulates as a consequence of this shift (Hammond et al., 1982).

It is unclear which if any of these events actually triggers the stress response. In 1980 Hightower proposed that HSPs are synthesized in response to the generation of abnormal proteins. He observed that addition of the amino acid analogue canavanine, induced the synthesis

of HSPs in chick embryo and mammalian cells (Hightower, 1980). Since then several lines of evidence have given support to this argument. For example, *Drosophila* mutants which synthesize altered actin in the indirect flight muscles constitutively express HSPs in that tissue (Hiromi et al., 1986). In addition, mammalian and yeast cells which cannot degrade abnormal and short lived proteins, due to defects in enzymes of the ubiquitin mediated proteolytic pathway, constitutively express elevated levels of heat shock proteins (Finley et al., 1984; Ciechanover et al., 1984; Seufert and Jentsch, 1990; see Ch. II, section A4 for further detail).

The effect of abnormal proteins in the stress response was directly studied by monitoring the expression of a *hsp70-lacZ* fusion in *Xenopus* oocytes which had been injected with native or denatured proteins. When bovine serum albumin (BSA) or β-lactoglobulin was injected in the native form, expression of the *hsp70-lacZ* fusion was not significantly altered. However, injection of denatured BSA or β-lactoglobulin increased β-galactosidase activity 10 fold (Ananthan et al., 1986). This suggests that the mere presence of abnormally folded proteins can induce the stress response. Alternatively, it may not be the presence of abnormal proteins alone which elicits the response but rather the way those abnormalities affect downstream metabolic and developmental pathways. Degradation of proteins for example could have enormous impact on energy metabolism by inactivating enzymes of aerobic metabolism (e.g. of the TCA cycle, oxidative phosphorylation etc.) and this may be the actual trigger of the stress response.

3. FUNCTIONS OF THE HEAT SHOCK PROTEINS

3.1 HSP 110

Very little is known about the function of this family of stress proteins. In mammals, these proteins are synthesized constitutively but are induced five fold upon heat shock. HSP110 proteins localize to the nucleolus during both non-shock and heat shock conditions where they bind either directly to RNA molecules or form complexes with RNA binding proteins (Subjeck et al., 1983; Subjeck and Shyy, 1986). The yeast homologue HSP104 is a recently identified

member of the ClpA/ClpB family of nucleotide binding proteins (Parsell et al., 1991). Site directed mutagenesis of the two putative nucleotide binding sites of HSP104 revealed that both sites are essential for thermotolerance. The thermotolerant defect in *HSP104* mutants can be suppressed by over-expression of *HSP70* suggesting that the function of these two classes of proteins is at least partially interchangeable (Parsell et al., 1991).

3.2 HSP90

This class of HSPs is present in abundance in the cytoplasm of normal cells. Upon heat shock, a small proportion of HSP90 translocates to the nucleus. In vertebrates, HSP90 seems to regulate steroid hormone receptor action by masking the DNA binding site of the receptor in the absence of hormone (Sanchez et al., 1985; DeMarzo et al., 1991). HSP90 is highly phosphorylated on serines and threonines and is transiently associated with several kinases, including the tyrosine kinases encoded by oncogenes (Oppermann et al., 1981; Ziemiecki et al., 1986), and the eukaryotic translation-initiation factor kinase which modulates the phosphorylation and hence activity of IF2. For this reason it has been hypothesized that HSP90 in the normal cell regulates both transcription and translation in response to developmental cues. During heat shock, it may be detrimental to proceed with regular developmental programs; signals may be faulty, delivered inappropriately or not at all, and execution of pathways in response to those signals may be physically impossible. By inactivating hormone receptors and kinases, HSP90 may prevent potentially deleterious mistakes in developmental and growth programs.

HSP90 also complexes non-covalently with actin and tubulin of the cytoskeleton and it has been postulated that in this role, HSP90 acts as a chaperone to stabilise the cytoskeletal network during heat shock (Schlesinger, 1990).

3.3 HSP70

The HSP70 family is the most highly conserved of the HSPs. Constitutively expressed members of this family are distributed throughout the cytoplasm and the nucleus in the unstressed cell. After heat shock however, newly synthesized HSP70 localizes to the nucleus and nucleolus (Welch and Feramisco, 1984). The current model regarding HSP70 function suggests that these proteins serve as molecular chaperones by binding to hydrophobic sites of nascent or denatured polypeptides thus preventing the formation of inappropriately folded insoluble aggregates (Beckmann et al., 1990; reviewed in Lindquist and Craig, 1988; Schlesinger, 1990; Pelham, 1988). In the unstressed cell, HSP70 acts transiently to modulate the unfolding, translocation and re-folding of newly synthesized proteins across membranes (Munro and Pelham, 1986; Chappell et al., 1986; Deshaies et al., 1988; reviewed in Pelham, 1988). Functions in the heat stressed cell are thought to be similar but the target is shifted from nascent polypeptides to thermally denatured structural proteins or inactivated enzymes (Gaitanaris et al., 1990; Skowyra et al., 1990; Beckmann et al., 1990; reviewed in Pelham, 1988). HSP70 may keep these proteins soluble until they can be degraded by proteolytic pathways or in the case of salvageable proteins it may promote their re-folding.

One HSP70, the *DnaK* gene product, is an abundant protein in bacteria and is essential for cell division and growth (Bukau et al., 1989). This gene was originally isolated as a mutation in bacteriophage λ growth. In particular, *DnaK* mutants were defective in lambda DNA replication. Since then it has been shown that the *DnaK* gene product can refold and reactivate heat-inactivated RNA polymerase in an ATP dependent manner (Skowyra et al., 1990). This *E. coli* HSP70 shares approximately 48% amino acid sequence identity with eukaryotic HSP70s (Bardwell and Craig, 1984).

The Immunoglobulin heavy chain binding protein (BiP), previously identified as glucose regulated protein GRP78 of mammals, is a constitutive, non-inducible HSC (heat shock cognate) 70 which facilitates the folding and assembly of secreted or membrane proteins which have been transported across the endoplasmic reticular membrane (Munro and Pelham, 1986). A

carboxyl terminal sequence, KDEL, distinguishes BiP from cytoplasmic HSP70s and promotes its retention in the endoplasmic reticulum (ER) (Munro and Pelham, 1987).

Another HSC70 has been identified as the uncoating ATPase which binds to and dissociates the clathrin light chain (Chappell et al., 1986).

Recent studies suggest that cytosolic HSP72 and HSP73 bind co-translationally with nascent peptides because nascent polypeptides released prematurely from polysomes *in vivo* were found complexed with HSP72 and HSP73 (Beckmann et al., 1990). This provides direct evidence that HSP70 in the unstressed cell chaperones the newly synthesized proteins. In stressed cells, association with HSP72 and HSP73 was prolonged and this effect was mimicked by the addition of L-azetidine 2-carboxylic acid , a proline analogue (Beckmann et al., 1990). These results suggest that it is not heat shock *per se* but the conformational state of the protein which determines when dissociation of HSP72 and HSP73 occurs.

In yeast, there are at least 9 members of the HSP70 family; three are localized to specific cellular compartments and still others mediate translocation to these organelles. The *Kar2* gene is a homologue of *bip/grp78* and operates in the ER (Chirico et al., 1988). Another protein, mHSP70, encoded by the *ssc1* gene, is an essential protein of the mitochondrial matrix (Deshaies et al., 1988) Temperature sensitive mutants of *ssc1* fail to transport proteins into the mitochondria. This HSP70 appears to function just within the mitochondrial membrane as it was isolated in a complex with the membrane bound precursor polypeptide and an outer membrane protein (Deshaies et al., 1988). It is hypothesized that SSC1-HSP70 binds imported precursors briefly before HSP60 interaction (See section 3.4) and actually pulls the polypeptide through the membrane. As with all other HSP70s, ATP hydrolysis drives release of SSC1 from the protein (Deshaies et al., 1988).

Thus the HSP70s chaperone newly synthesized polypeptides to their final destinations within the cell. In the cytoplasm, HSP70s unfold and bind to peptides as they are translated, and escort them to various cellular organelles. Within these compartments are other HSP70s which may drive the translocation of the unfolded polypeptides through membranes into

cellular compartments while still other HSP70s and HSP60s facilitate the appropriate re-folding and assembly of the precursors within organelles.

3.4 HSP 60

The HSP60 proteins possess "chaperonin like" functions similar to the HSP70 family except that they mediate activities on the inner side of various cellular organelles in eukaryotes. Members of this family form complexes with polypeptides, have ATPase activity, and participate in the folding and assembly of proteins. These HSPs are constitutively synthesized in growing cells but are also mildly heat inducible.

In bacteria, the *groEL* gene encodes a HSP60 which is essential for the assembly of lambda phage and for growth. The demonstration that GROEL protein participates in the folding of pre-β-lactamase also suggests that HSP60s may play a role in protein secretion (Kusukawa et al., 1989). The ATPase activity of the *groEL* gene product is modulated by interaction with the *groES* gene product (Tilly and Georgopoulos, 1982; Chandrasekhar et al., 1986). Recent experiments demonstrated that GROEL could promote the assembly of RuBisCO (ribulose-1,5-biphosphate carboxylase-oxygenase) when its protein subunits were expressed from a plasmid transformed into *E. coli* (Goloubinoff et al., 1989).

Conditional mutants in the yeast *groEL* homologue, *mif-4* are phenotypically defective in mitochondrial function; they can still import proteins to the mitochondria at restrictive temperatures but cannot assemble them properly once they are inside the organelle. Analysis of the *mif-4* gene product has revealed that it is a homo-oligomer of 14 subunits which is essential for re-folding of imported mitochondrial matrix proteins (Reading et al., 1989).

In the chloroplasts of plants, the GROEL eukaryotic homologue RuBisCO-binding protein mediates the identical process, the assembly of the active RuBisCO enzyme complex (Hemmingsen et al., 1988). Thus the functions of the HSP60 family are highly conserved between prokaryotes and eukaryotes.

3.5 Low MW HSPs

In comparison to the other HSPs, the function(s) of the small HSPs are not understood. These proteins have been isolated from plants, yeast, invertebrates, and vertebrates and range in size from 15 to 30 kD. All small HSPs contain a carboxyl terminal domain which shows extensive homology to the α-crystallins of the vertebrate eye lens, suggesting that the α-crystallins were derived from an ancestral heat shock protein which possessed properties suitable for lens function. These proteins form large stable aggregates in the lens which provide the transparency of the lens for its lifetime.

All of the small HSPs isolated to date form massive aggregates *in vivo*. In humans, there is only one small HSP, HSP28, which is constitutively synthesized in small amounts. In the unstressed cell, HSP28 forms soluble aggregates of 200-800 kDa in the cytosol. Upon heat shock, HSP28 migrates to the nucleus and forms even larger, up to 2 MDa, relatively insoluble homomeric complexes (Arrigo et al., 1988). During recovery there is a return to the pre-heat shock state, and HSP28 redistributes in the soluble cytoplasmic fraction (Arrigo et al., 1988).

The localization of small heat shock proteins upon heat shock does not appear to be universal however. In chicken embryo fibroblasts, high molecular weight insoluble cytoplasmic aggregates develop from soluble particles after a second heat shock and do not re-distribute to the nucleus (Collier et al., 1988).

Tomato small HSPs also form high molecular weight insoluble cytoplasmic granules upon heat shock, and these granules have been found to be associated with mRNAs, specifically those which are not translated during heat shock (Nover et al., 1989). This suggests that one function of the small HSPs may be to protect mRNAs during stress so that they are immediately available for translation upon recovery. Alternatively, the small HSPs may sequester mRNAs during stress so that they are not translated under adverse conditions or to free up the translation machinery for HSP production.

There are also situations in which small HSPs migrate to specific organelles upon heat shock: HSP30 of *Neurospora* for example relocates to mitochondria upon heat shock (Plesofsky-Vig and Brambl, 1990).

HSP27 of *Drosophila* forms large insoluble aggregates upon heat shock which are localized to the nucleus. HSP27 is also hormonally regulated in the absence of stress (See Section 4.2). While ecdysterone induced HSP27 also localizes in the nucleus, the aggregates are soluble (Beaulieu et al., 1989). In *Drosophila* there are four small HSPs raising the possibility that the aggregates which form after heat shock may actually be heteromeric.

In a quest to determine the function of small HSPS, McGarry and Lindquist transformed flies with an antisense *hsp26* construct. While *hsp26* expression was not eliminated by this strategy, HSP26 synthesis in heat shocked flies was dramatically reduced. HSP26 repression did not affect induction of other HSPs, nor was recovery and the resumption of normal cellular protein synthesis delayed (McGarry and Lindquist, 1986). The implication here is that HSP26 does not play a major role in the regulation of the heat shock response.

Recent structural analysis of purified mouse HSP25 showed that it is organized into high molecular weight complexes of about 730 kDa or equivalent to 32 HSP25 monomers (Behlke et al., 1991). The authors proposed an arrangement of hexagonal packing to achieve the 15-18 nm in diameter spherical structure observed using electron microscopy (Behlke et al., 1991). In addition, they demonstrated that on two dimensional gels there are at least three isoforms of HSP25, two of which are phosphorylated (Behlke et al., 1991). Another study from the same laboratory isolated and sequenced phosphopeptides from HSP25 and determined that phosphorylation occurred specifically at two serine residues present in a kinase recognition region which is conserved amongst small mammalian HSPs (Gaestel et al., 1991). Phosphorylation was not required to assemble HSP25 into large aggregates however (Behlke et al., 1991). The α-crystallins have also been shown to be phosphorylated *in vivo* (Spector et al.,1985) and small HSP phosphorylation has been correlated with the acquisition of thermotolerance (See section 5). Thus while phosphorylation does not appear to be required for

complex assembly, it presumably is required for the function of these complexes during heat shock and in the vertebrate eye lens.

4. Heat shock proteins in development, differentiation, and growth

Many heat shock proteins can be induced by other agents which presumably exert deleterious effects on the cell. However, some HSPs (e.g. HSP70) have relatives (e.g. HSC70) which are constitutively expressed in the absence of heat shock. This implies that heat shock proteins perform functions which are required by a normally growing cell as well as a stressed cell. While many of these HSC proteins show temporal and spatial expression patterns, the developmental induction of the small *hsps* which do not have family members that are constitutively expressed is of particular interest. Moreover, inducibility of the heat shock response itself and the acquisition of thermotolerance seem to be under developmental regulation. The following examination of these processes provides some examples of the ever growing documentation of HSP induction in the absence of stress.

4.1 Developmental control of the heat shock response in Xenopus

Development of the heat shock response in the frog is a complex process. Prior to ovulation, in response to heat shock, oocytes translate HSPs 83, 76, 70, and 57 from pre-existing maternal messages (Bienz and Gurdon, 1982; Browder et al., 1987). After ovulation, the small HSPs are also heat inducible (Browder et al., 1987). Fertilization dramatically terminates the heat shock response, and no heat shock proteins can be induced. However, constitutive HSP70 synthesis is derived from oogenic *hsp70* RNA which has been retained (Browder et al., 1987). Extinction of the heat shock response in early embryos coincides with a period of extreme thermolability. The mid-blastula transition re-introduces the heat shock response when transcription of the embryonic genome commences (Bienz, 1984). The first HSPs which are inducible at this stage are HSP68, HSP70, and HSP87. This period also corresponds to the acquisition of thermotolerance in embryos. By neurulation, the small HSPs

and the HSP60 family as well as other members of the HSP70 family are inducible. Thus, inducibility of HSPs is acquired asynchronously (Browder et al., 1987; Nickells and Browder, 1988).

Not only is inducibility of each HSP acquired independently during development, but there is also differential expression with regards to spatial location of synthesis, temperature of maximum induction, and duration of synthesis during stress. For example, HSP35 is initially inducible at the blastula/gastrula transition stage at temperatures above 35°C and usually only for the first 40 minutes of heat shock (Nickells et al., 1989). Inducibility of this HSP drops considerably in late gastrulation, followed by a resurgence at neurulation (Nickells et al., 1989). HSP 35 is actually the glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and the inducibility of this enzyme is inversely proportional to levels of constitutive GAPDH (Nickells and Browder, 1988) suggesting that high levels of constitutive GAPDH can adequately cope with the immediate crisis presented by heat shock. Thus the inducibility of particular HSPs at certain times or places in development may reflect the metabolic state of the tissue at that point in time.

The adult frog continues to show differential synthesis of heat shock proteins in response to stress. While undifferentiated erythroblasts can synthesize a variety of HSPs, mature erythrocytes synthesize only HSP70 (Winning and Browder, 1988).

4.2 Developmental regulation in Drosophila

Evidence in *Drosophila* also indicates that *hsp* induction is absent in early development and acquired at later stages. Oocytes constitutively express *hsps* 83, 28/27 and 26, while *hsps* 70, 23 and 22 are neither expressed nor heat inducible (Zimmerman et al., 1983).

The small *hsps* of *Drosophila* (*Hsps22*, *23*, *26*, and *27*) are clustered within a 12 kb region of the genome identified as locus *67B* (Petersen et al., 1979; Corces et al., 1980; Voellmy et al., 1981). In the absence of stress, expression of the small HSPs occurs in imaginal discs and during pupation and is controlled by the moulting hormone ecdysterone (Ireland and Berger,

1982). Deletion analysis and transient chloramphenicol acetyltransferase (CAT) assays in *Drosophila* tissue culture cells showed that ecdysterone induction was mediated through multiple elements between -579 and -455 of the *hsp27* promoter while heat induction was regulated through three HSEs present between -370 and -270 (Riddihough and Pelham, 1986). Dnase I hypersensitive sites centred at -522 and -293 of the *hsp27* gene support the positioning of these elements (Costlow and Lls, 1984). Surprisingly, similar studies of the *hsp23* gene implicated sequences dissimilar to the elements defined for *hsp27*. In addition, experiments where *hsp27* gene fusions were introduced into flies by P-element mediated transformation yielded slightly different results again (Hoffman and Corces, 1986; Hoffman et al., 1987). These discrepancies may reflect the inherent problems associated with transient assays of developmentally regulated genes in cultured cells. Since there is an established tissue specificity for some of these genes (e.g. *hsp26* is expressed in the male germ line and imaginal discs (Glaser et al. 1986)) it is quite possible that data from tissue culture cells does not accurately reflect the *in vivo* situation. Nonetheless, all of the experiments suggested that the elements regulating developmental expression of the *hsps* by ecdysterone are separate and distinct from the HSEs.

4.3 Developmental regulation of hsp 70

There are many cases in which members of the HSP70/HSC70 family are synthesized in response to developmental cues. Fetal rabbits at 25 days gestation are transcriptionally heat inducible for *hsp70* while mRNA for its cognate is expressed constitutively in unstressed fetal tissues (Brown et al., 1985). Primitive cells of the avian erythroid lineage express HSP70 constitutively but are not heat responsive at either the transcriptional or translational level while cells of the definitive lineage synthesize HSP70 after heat stress through a translational control mechanism (Banerji et al., 1987). Differential synthesis of HSPs also coincides with differentiation of the mammalian testis. In the immature mouse testis, HSP72 and HSP73 are synthesized at low levels constitutively in the absence of stress, and HSP72 is strongly heat

inducible. The mature testis synthesizes HSP72 at slightly elevated levels after heat stress *via* translational control mechanisms. In addition, a novel isoform of HSP73 was detected in unstressed meiotic and post meiotic germ cells of the adult testis. Expression of this protein was not increased upon heat shock (Zakeri et al., 1990). Thus, only cells of the somatic gonad were heat responsive.

Levels of HSP70 may be even more tightly controlled during growth. Synthesis of HSP70 incresases during the late G1 to S phase, and peaks at G2 (Van Dongen et al., 1986; Simon et al., 1987). Other studies have shown that HSC70 proteins are under cell cycle control *via* an E1A like activity (Nevins, 1982; Kao et al., 1985).

What exactly is the role of heat shock proteins in development? Some investigators suggest that expression of HSPs during development is an essential pre-requisite to the development of natural thermo- resistance or as it is more commonly referred to, intrinsic thermotolerance. Usually, the acquisition of thermotolerance correlates well with the acquisition of HSP inducibility. For example, decreased risk of lethality in mouse embryos has been shown to be associated with the ability to synthesize inducible HSP68 (Muller et al., 1985). But this does not explain the temporal and spatial expression patterns observed in the absence of heat shock. Developmental programs and growth demand rapid changes in the architecture and metabolism of the organism. Not only must cells multiply but they must become specialized, migrate, and form adhesions with other cells to create functional, communicating organs. Many cells are doomed to death and their corpses must be scavenged. Thus, the disintegration, re-distribution and ordering of multiplying cellular components during devlopment may mimic the consequences of heat shock and the hormonally induced expression of hsps might meet the challenge imposed in specific situations. Alternatively, HSP synthesis may be required to direct initial changes in gene expression during development. Studies in trypanosomes have indicated that when these parasites are transmitted from their insect vectors to their mammalian hosts they undergo heat shock by virtue of an increase from 25°C to 37°C. This temperature shift induces differentiation from the promastigote into the

amastigote form and can be reproduced *in vitro* (Van der Ploeg et al., 1985). Thus heat shock may act as a switch in determining developmental pathways.

5. The role of low molecular weight HSPS in thermotolerance

The concept of thermotolerance was conceived when it was noticed that a pre-shock treatment could render an organism more resistant to a subsequent challenging stress which would otherwise be lethal (Milkman, 1966). That HSP production might play a role in thermotolerance was suggested by Mitchell et al. (1979) who found that mild heat treatments which induced the synthesis of HSPs provided protection from heat induced lethality and developmental abnormalities in *Drosophila.* In addition, agents other than heat such as arsenite, ethanol, and re-oxygenation after hypoxia can induce thermotolerance (Li and Werb, 1982).

In one study, Nguyen et al. (1989) followed the kinetics of reporter enzymes synthesized constitutively in mouse and *Drosophila* cells under the control of the SV40 promoter when the cells were subjected to heat shock. Both β-galactosidase and luciferase rapidly became insoluble in both cell types. However, enzyme inactivation during heat shock or ethanol induction was significantly reduced in pre-heat treated cells even when the challenging stress was applied 20 hours after recovery was initiated from the priming treatment. Since the mouse does not produce endogenous β-galactosidase or luciferase, the effect of thermotolerance could not be attributed to any specifically designed mouse cell functions (Nguyen et al., 1989).

Several lines of evidence suggest that the small HSPs may be involved in the acquisition of thermotolerance. Ecdysterone induction of the small HSPs in *Drosophila* tissue culture cells has been shown to confer a thermotolerant phenotype, and *Drosophila* pupae are naturally more heat resistant than other fly stages (Mitchell et al., 1979; Berger and Woodward, 1983). A thermo-intolerant *Dictyostelium* mutant is totally deficient in small HSP synthesis but otherwise synthesizes the complete set of HSPs upon heat shock. However, the levels of HSP70 after heat shock are somewhat reduced compared to wild type levels (Loomis and Wheeler, 1982), and it may be this aspect which results in the thermosensitivity of the mutant.

Heat resistant mutants isolated from Chinese hamster lung cells have increased levels of HSP27, particularly of the phosphorylated isomers of HSP27 (Chrétien and Landry, 1988). When the protein synthesis inhibitor cycloheximide was added to normal Chinese hamster cells, HSP27 phosphorylation accompanied by thermotolerance was induced, without *de novo* protein synthesis (Landry et al., 1989). Thus phosphorylation of the small HSPs may be an important component in the acquisition of thermotolerance.

6. REGULATION OF THE HEAT SHOCK RESPONSE

The rapid initiation of heat shock protein synthesis and the repression of normal cellular protein synthesis is controlled by a host of regulatory mechanisms. The following discussion is drawn largely from data in eukaryotes, but particularly with regards to transcriptional regulation, the principles can often be transferred to prokaryotes.

Prokaryotic heat shock genes form a regulon which is transcriptionally induced upon heat shock by the activity of a positive trans-acting factor. This factor is a specialized version of RNA polymerase, σ^{32} in *E. coli*, which acts through recognition of a cis-acting element present upstream of heat shock genes. Increasing the rates of σ^{32} synthesis by transforming bacteria with a plasmid containing the σ^{32} gene, *rpoH*, under control of the *lac* promoter resulted in increased production of HSPs in the absence of stress indicating that levels of the prokaryotic transcription factor directly determines the extent of the heat shock response (Grossman et al., 1987).

In eukaryotes, heat shock factor is a separate entity from RNA polymerase II, but it is the action of this factor in concert with RNA polymerase II and other proteins (e.g. TATA factor) which triggers the enormous surge in *hsp* gene expression. In addition, eukaryotes employ post-transcriptional and translational control mechanisms to fine tune the response.

6.1 Transcriptional Control in Eukaryotes

The dramatic synthesis of heat shock proteins upon elevation in temperature is largely due to *de novo* or increased transcription of heat shock genes. In eukaryotes this is mediated through the binding of a multimeric positive trans-acting factor, heat shock factor (HSF) to cis-acting sequences, the heat shock elements (HSEs), which are usually located 80-150 bp upstream of the transcriptional start site. Originally defined as a 14 bp consensus sequence of CTNGAANNTTCNAG (Pelham, 1982; Pelham and Bienz, 1982) a heat shock element is now believed to consist of at least two or three modules of the sequence NGAAN, which can be present in either orientation (Perisic et al., 1989).

Many heat shock genes such as *hsp70* of *Drosophila*, require elements other than the HSEs such as CCAAT boxes for optimal expression (Dudler and Travers, 1984). Others are expressed at specific times during normal development in the absence of heat shock through hormone binding elements which are distinct from the HSEs. (Previously discussed in section 4.2).

Recent work has focussed on the interaction between HSF and HSEs. HSF has been purified from yeast (Sorger and Pelham, 1987), *Drosophila* (Wu et al. 1987), and human (Goldenberg et al., 1988) and somewhat surprisingly, these proteins display a wide degree of divergence. Except for sequence identity in the DNA binding and the trimerization domains they appear to be totally unrelated. This is in sharp contrast to the high degree of conservation observed amongst heat shock proteins and the HSEs between species, and implies that there can be many functional variations on the theme of heat shock factor. In yeast, the gene for HSF has been recently cloned and characterized (Wiederrecht et al., 1988; Sorger and Pelham, 1988). It is a single copy gene which is essential for normal growth (Wiederrecht et al., 1988;Sorger and Pelham, 1988). Yeast must therefore require transcription of heat shock genes constitutively or other unidentifed genes are expressed *via* HSF. Heat shock factor was isolated from non-shocked yeast by affinity chromotography in which concatenated oligomers of HSEs were linked to a Sepharose column (Sorger and Pelham, 1987). Purified heat shock factor

from unshocked yeast is relatively large (150kD); this is somewhat greater than the 93kD size predicted by the sequenced gene. Deletion analysis determined that the DNA binding domain resided within a basic region in the amino terminal portion of the protein. Originally, no obvious secondary structures commonly observed in DNA binding proteins were recognized in yeast HSF (e.g. zinc fingers, leucine zippers etc.) (Wiederrecht et al., 1988; Sorger and Pelham, 1988). However, recent evidence suggests that the DNA binding domain of HSF possesses a helix turn helix structure while the domain for oligomerization resembles a leucine zipper (Clos et al. 1990; Rabindran et al. 1991; Schuetz et al. 1991).

In yeast, HSF binds to HSEs constitutively and becomes transcriptionally activated after heat shock (Sorger et al., 1987). In fact, levels of heat shock factor bound DNA isolated from control and heat shock cells are virtually equivalent. However, Sorger et al. (1987) noticed that the mobility of HSF-HSE complexes on polyacrylamide gels shifted after heat shock treatment and that at least five modified species were present. Addition of phosphatase reduced this shift in mobility. In addition, they observed that the mobility shift increased linearly with temperature increase. However, phosphorylation did not alter the affinity of HSF for HSE DNA. They suggested a model whereby in yeast, HSF binds to DNA under normal conditions. Heat shock promotes phosphorylation of HSF at multiple sites which alters the conformation of HSF such that it can now complex or interact with other constituents of the transcriptional machinery such as TATA factor and/or RNA polymerase to stimulate transcription (Sorger et al., 1987; Sorger and Pelham, 1988).

In *Drosophila* and human cells, HSF binds to HSEs after heat shock (Wu et al., 1987; Sorger et al., 1987; Kingston et al., 1987; Zimarino et al., 1987), and this change in binding affinity is accompanied by phosphorylation. Binding and activation of HSF to HSEs in *Drosophila* does not depend on *de novo* protein synthesis suggesting that post translational modifications alone may mediate conformational changes in HSF which allow it to interact with DNA. Why *Drosophila* and human cells bind HSF after heat shock as opposed to the constitutive DNA interaction in yeast is a puzzling question.

Experiments in yeast suggest that the transcriptional activity of HSF can actually be separated into two distinct domains; one is required for sustained transcription at normal growth and moderate heat shock temperatures (i.e. 15-33°C) while another transiently activates transcription during a severe heat shock (i.e. >34.5° C) (Nieto-Sotelo et al., 1990; Sorger, 1990). Recently in *Drosophila* , using an anti-HSF antibody, Westwood et al. (1991) showed that during heat shock HSF localized specifically to heat shock puffs in polytene chromosomes as well as to sites of hormonally induced puffs which regress during heat shock. These data suggest that HSF may also act as a repressor of normal gene activity during heat stress (Westwood et al., 1991).

While yeast possesses only one HSF gene, the recent discovery that there are at least two HSF genes in humans and three in tomato raises intriguing possibilities (Schuetz et al., 1991; Rabindran et al., 1991; Scharf et al., 1990). Multiple HSFs may allow for the specialization of these proteins to perform different functions. Depending on the physiological conditions of the cell, a different HSF may be required to bind to the promoter to elicit a transient or sustained response. This might explain why HSF binds to DNA after the stress has been implemented in animal cells.

Additionally, protein-DNA cross linking experiments by Gilmour and Lis (1986) and nuclear run-on transcription assays by Rougvie and Lis (1988) determined that the *Drosophila hsp70* gene is transcriptionally engaged *in vivo* under non-shock conditions, and that RNA polymerase II has begun the synthesis of a nascent transcript of about 25 nucleotides. The authors suggested that RNA polymerase is stalled at this point and is unable to elongate the transcript until heat shock activates its release. This would allow for rapid expression of heat shock genes in response to thermal stress by circumventing a potential delay in transcription due to selection and binding of a specific HSF by already initializing the task. Other *Drosophila* genes such as *hsp26* have also been shown to be transcriptionally engaged (Rougvie and Lis, 1988;Thomas and Elgin, 1988). How widespread a phenomenon this characteristic is remains to be seen.

6.2 Post transcriptional regulation

One of the first indications that heat shock interferes with processing of transcripts was supplied by the study of Mayrand and Pederson (1983) in which they found that heat shock disrupts RNP particles in *Drosophila.* More recently, it has been demonstrated that heat shock dramatically alters the structure of the spliceosome itself. Immunological studies in both *Drosophila* and mammalian cells showed that anti-snRNP antibodies, which recognize U1 snRNPs, reacted similarly with preparations from both non-heat shocked and heat shocked cells. But while non-heat shocked cells reacted with anti-Sm antibodies, which detect components of the U1, U2, U4, and U5 snRNPs, heat shocked cell preparations were unable to do so (Welch and Mizzen, 1988). Cells which had been treated with a mild heat shock prior to the challenging heat shock maintained Sm antigenicity (Wright-Sandor et al. 1989). Thus, heat pre-treatment appears to protect components of the splicing machinery.

Accumulating evidence suggests that processing of pre-mRNA is much more sensitive to the effects of heat shock than transcription. The *Drosophila hsp83* gene, which contains a single 5' intron, is expressed at normal temperatures and is strongly induced by moderate heat shocks. However, poor expresssion was observed at high temperature heat shocks due to a block in splicing of the pre-mRNA transcript. In contrast, transcription of the gene persisted at high levels (Yost and Lindquist, 1986) resulting in the accumulation of large amounts of unprocessed full length transcripts. Processing of an *adh* gene intron put under control of the *hsp70* promoter was also inhibited by heat shock suggesting that the block in splicing is general (Yost and Lindquist, 1986). This block in splicing persisted for two hours after recovery conditions were initiated. Heat shock has also been shown to disrupt cis-splicing in chicken, *Dictyostelium*, mammalian cells (Bond and Schlesinger, 1986; Kay et al., 1987; Maniak and Nellen, 1988; Bond, 1988) and trans-splicing in trypanosomes (Muhich and Boothroyd (1988). Kay et al. (1987) demonstrated that while transcription of a transfected nematode small *hsp16* gene pair was normal in heat shocked mouse cells, splicing of introns

was defective and unprocessed transcripts accumulated in the nucleus. During recovery these transcripts were processed and transported to the cytoplasm. In contrast, processing of the wild type *hsp16* gene pair was normal in heat shocked nematodes and did not result in the accumulation of unspliced transcripts. This suggests that the splicing apparatus may be less sensitive to heat stress in *C. elegans* than in other organisms. Thus, heat shock genes in *C. elegans* may be able to afford the luxury of introns which *hsps* in other animals cannot.

Yost and Lindquist (1986) demonstrated that pre-treatment of *Drosophila* cells with a mild heat shock could rescue splicing of the *hsp83* transcript during a severe heat shock which would otherwise have inhibited processing. When the translation inhibitor cycloheximide was administered with the heat pre-treatment before HSPs could be produced, splicing was later disrupted at high temperatures. Addition of cycloheximide after HSP synthesis had no effect on splicing at subsequent high temperature heat shocks (Yost and Lindquist, 1986). Collectively, the above experiments suggest that HSPs themselves play a role in the protection and/ or repair of the splicing apparatus during heat shock, either by stabilizing the components of the spliceosome or perhaps by participating directly in the process of splicing itself. Although it is not clear which HSPs specifically might have these functions, the HSP70 family is a likely candidate since its members consistently move to the nucleus during heat shock.

In most species, the canonical heat shock genes are devoid of intervening sequences; *hsp70* for example does not contain introns while the *hsc70* genes do (Ingolia and Craig, 1982). Clearly this may be simply an adaptation to avoid the shortcomings of the splicing apparatus during heat stress. *Hsp83* is the only *Drosophila* heat shock gene which possesses an intron. It has been suggested that since this gene is already expressed at reasonably high levels during normal growth, there may be enough processed transcripts available to meet the increased demand for HSP83 protein during heat stress and that translational control might operate more in the regulation of HSP83 synthesis.

6.3 Translational Control

During heat shock, the synthesis of normal cellular proteins is actually repressed in favour of heat shock protein synthesis. Thus, induction of heat shock proteins alone is not the sole change in cellular activity integral to the stress response. While transcription and processing of normal cellular RNAs (unlike hsp RNAs) are severely hampered at elevated temperatures, levels of pre-existing RNAs are quantitatively maintained (Storti et al., 1980; Findly and Pederson, 1981; Lindquist, 1981; Petersen and Mitchell, 1981). Moreover, when normal cellular RNAs are extracted from heat shocked cells they can be translated in cell free systems *in vitro*. Thus the integrity of the RNA transcripts themselves must also be maintained.

Repression of normal protein synthesis is remarkably swift and precedes the appearance of *hsp* RNA transcripts (McKenzie et al., 1975; Lindquist, 1981). Addition of actinomycin D prior to heat shock to prevent *hsp* transcription still results in repression of normal cellular protein synthesis (Lindquist, 1981). Thus repression of normal protein synthesis and induction of *hsp* gene expression can be uncoupled.

Early studies by McKenzie et al. (1975) indicated that polysomes disappeared with the repression of normal protein synthesis. Subsequently, polysomes re-associated with newly transcribed *hsp* mRNAs (Lindquist, 1980). In another study, *Drosophila* lysates purified from heat shocked cells preferentially transiated *hsp* RNA while lysates isolated from cells maintained at 25°C showed no preference when incubated at the same temperature. Thus, once specificity was established it remained a property of the lysate. Mixing the lysates resulted in translational activity intermediate between the two (Scott and Pardue, 1981; Lindquist, 1981). When the heat shock lysate was supplemented with a fractionated crude ribosomal pellet from non-shocked cells translation of normal cellular RNAs ensued. Addition of the ribosomal pellet from heat shocked cells to the 25°C lysate did not alter the translational specificity of the control lysate (Scott and Pardue, 1981). This suggests that a factor isolated with the ribosomal pellet is inactivated in heat shocked cells.

Recent experiments by Marota and Sierra (1988) have shown that in heat shock lysates, translation of *hsp* RNAs is unaffected by the addition of cap analogues while translation of 25°C RNAs is further repressed. These authors suggest that normal cellular RNAs require capping to be translated while heat shock RNAs by virtue of some property in their secondary structure may be able to bypass this necessity. They suggest that it is the inactivation of cap binding factor which results in repression of normal cellular RNA translation (Marota and Sierra, 1988; Marota and Sierra, 1989). However, heat inducible RNAs do appear to possess normal cap structures *in vivo* (Yost et al. 1990). Thus, while they may be able to bypass the requirement for capping in certain situations, it is unlikely that this alone could explain the highly efficient preferential translation of heat shock RNAs.

Rather, mutational analysis of the *hsp70* and *hsp22* genes of *Drosophila* has revealed that sequences present in the 5' untranslated leader contribute significantly to the ability of *hsp* RNA to escape the block in translation conferred by heat stress. Firstly, when the 5' end of *hsp70* including the untranslated leader was fused to other non-hsp sequences, the hybrid transcripts were efficiently translated at elevated temperatures (DiNocera and Dawid, 1983; Bonner et al., 1984). In another study, transcripts derived from *hsp70* constructs containing internal deletions in coding sequence were translated as efficiently as transcripts derived from the wild type gene. Deletion of the 5' untranslated leader, in contrast, destroyed translatability while transcription of the fusion was unaffected (McGarry and Lindquist, 1985). These experiments suggested that the untranslated leader was important for translatability but did not indicate whether this property was confined to specific sequences in the leader or resulted from properties of its secondary structure.

Leaders of heat shock genes in *Drosophila* tend to be fairly long, usually about 250 bp. To determine if specific signals are contained within this region, Klemenz et al. (1985) constructed various fusions of the *Drosophila hsp70* gene to the alcohol dehydrogenase (*adh*) gene and transformed them into flies. A fusion of the *hsp70* promoter to the *adh* gene was transcriptionally heat inducible but mRNAs were not efficiently translated. However, insertion

of the first 95 nucleotides of the *hsp70* transcript was sufficient to confer preferential translation of fusion transcripts at heat shock temperatures (Klemenz et al., 1985). Deletion analysis of the *hsp22* leader yielded similar results; 86% of the leader could be removed since only the first 26 nucleotides were required for translation. Collectively, these data imply that high levels of transcription at elevated temperatures are not sufficient to achieve translation, and that sequences in the 5' end of the leader are important for selective translation of transcripts. Similar experiments have demonstrated that large deletions (+37 to +205 of the 249 bp hsp70 leader, and +27 to +242 of the hsp22 leader) do not affect translatability (McGarry and Lindquist, 1985; Hultmark et al., 1986). On the other hand insertion of additional sequences near the 5' end may, but not necessarily hamper translation. In one situation, insertion of 39 nucleotides of sequences upstream of the transcriptional start site, at position +2 of the *hsp70* leader eliminated translation, while insertion of a duplication of the first 37 nucleotides of the leader at the same position did not diminish translation (McGarry and Lindquist, 1985). Collectively, these data indicate that while conserved sequences at the 5' end of hsp untranslated leaders are important, the structure of the leaders themselves may be equally important in signalling preferential translation of *hsp* transcripts.

There are also indications that signals in the 3' ends of hsp messages might regulate synthesis of HSPs. Experiments with human tissue culture cells have demonstrated that the *hsp70* message is more stable after heat shock and that this stability is mediated through the 3' untranslated leader (Theodorakis and Morimoto, 1987).

6.4 Recovery from heat shock

Recovery from heat shock implies a return to the normal physiological state of the cell prior to the initiation of stress. It therefore must involve the inactivation or repression of heat shock protein synthesis and the resumption of normal cellular protein synthesis. But implicit in the return to normalcy is the assumption that "all must be well" within the cell and that the cell is adequately prepared to resume normal metabolic pathways. For this to be true, *hsps*

must have successfully fulfilled their role as protectors and repairmen from the deleterious effects of stress. So what determines when a cell has truly recovered from heat stress? Generally, at least two criteria must be met for recovery to take place. First of all there must be a return to normal physiological growth temperatures or removal of the stress agent. Secondly, there are indications that minimal quantities of HSPs are required before the response is shut down (DiDomenico et al., 1982a), although HSP synthesis may not be required at all for recovery from very mild or short heat shocks. This implies that HSPs auto-regulate the extent and duration of the response.

In *Drosophila* cells, specific levels of HSP70 are consistently required before recovery from heat shock can begin (DiDomenico et al., 1982b). Inhibition of HSP70 synthesis by artificial production of *hsp70* antisense RNA delays recovery (Melton et al., 1988) while inhibition of HSP26 synthesis by antisense RNA does not (McGarry and Lindquist, 1986). Still other evidence suggests that the cell autoregulates levels of HSP70 relative to the extent of the stress (Stone and Craig, 1990). These data imply that HSP70 might be the sensor which cues the other components of the response to shut down.

While the induction of HSP synthesis is an amazingly rapid response, maximal usually within minutes, down-regulation of the response occurs over a much longer period of time, usually several hours. HSP synthesis does not cease immediately but rather decreases gradually while translation of the pre-existing normal RNAs which were sequestered during heat shock gradually resumes. Inactivation of *hsp* gene transcription and selective degradation of existing *hsp* mRNAs results in a steady decline in HSP synthesis (DiDomenico et al., 1982a). Transcriptional activation of *hsp* genes is a highly coordinated event such that HSP synthesis occurs in unison. In contrast, *hsp* mRNAs are repressed asynchronously in *Drosophila* starting with repression of *hsp70* and ending with repression of *hsp82* (DiDomenico et al., 1982b; Lindquist and DiDomenico, 1985). The fact that *hsp70* is down-regulated first provides further support to the above argument that this HSP may be the real autoregulator of the response.

What mediates the selective degradation of heat shock messages? In the previous section it was mentioned that *hsp70* mRNA becomes more stable after heat shock and that this property was contained within the 3' untranslated region of the message. Messages transcribed from fusions of the *adh* gene and the 3' untranslated sequences of *hsp70* show degradation kinetics similar to wild type *hsp70* transcripts (Petersen and Lindquist, 1989). The 3' untranslated region of *hsp70* possesses elements with identity to 3' sequences of normally unstable messages such as that of *c-myc* and *c-fos*. Not surprisingly, heat shock stabilizes *c-myc* and *c-fos* RNAs (Sadis et al., 1988). Thus while it is not clear how heat shock actually stabilizes certain transcripts, it is this process which is partially responsible for the induction of HSP70 synthesis during heat shock. During normal cellular growth and recovery, *hsp70* messages are unstable, thus maintaining a low level of available HSP70 in non-shocked cells.

7. Caenorhabditis as a model system

Caenorhabditis elegans is a small (less than 1 mm in length) free living soil nematode possessing many qualities which are advantageous to the researcher. The life cycle of this animal is short, three to four days at room temperature, and it can be simply maintained in the laboratory on Petri plates containing nutrient growth media spread with a layer of bacteria. Alternatively, large quantities of animals for biochemical analysis can be cultured in liquid in large bottles (Sulston and Brenner, 1974).

The transparency of this animal and its relatively simple anatomy has allowed the complete delineation of every cell lineage (Sulston, 1976; Sulston and Horvitz, 1977; Sulston et al., 1983; Albertson and Thomson, 1976), including cell migrations and programmed cell deaths and of the animal's neurocircuitry (White et al., 1986).

In addition, powerful genetic tools are available in *C. elegans*. Numerous mutants have been isolated and while the animal is a self fertilizing hermaphrodite, males arise spontaneously by non-disjunction at a frequency of about 1 in 500 when stocks are maintained at 20°C, allowing for crosses. Each hermaphrodite can produce about 250 progeny. The genome

is compact, about 5 times the size of the yeast genome and two thirds the size of *Drosophila*'s. Moreover, the genome contains a low proportion of short-period repetitive sequences (Emmons et al. 1980). The average intron in *C. elegans* is 50-100 bp in length. As a result of the efforts of the MRC laboratory in Cambridge and the laboratory of R. Waterston in St. Louis, most of the genome has now been cloned and ordered as a series of cosmid and yeast artificial chromosome (YAC) clones (Coulson et al., 1986; Coulson et al., 1988). Sequencing of the genome of *C. elegans* has begun.

Finally, of particular relevance to the present study, transformation, both integrative and extrachromosomal, has become an efficient and relatively simple process by which reverse genetics can be employed to study gene expression and cell biology (Fire, 1986).

8. The Heat Shock Response of C. elegans

Elevation of the temperature from 25°C to 31°C - 34°C induces the synthesis of HSPs corresponding to the HSP90, HSP70 and small HSP families characteristic of other species (Snutch and Baillie, 1983). But while *C. elegans* synthesizes all families of HSPs, research has focussed on the HSP70 family (Snutch and Baillie, 1984; Snutch et al., 1988) and the small 16 kDa HSPs (Russnak and Candido, 1985; Jones et al., 1986).

8.1 The small hsps of C. elegans

The small HSPs in *C. elegans* consist of proteins ranging from 16-25 kD (Snutch and Baillie 1983; Russnak *et al.*, 1983). To date, four different genes encoding the 16 kD HSPs have been cloned (Russnak and Candido 1985; Jones *et al.*, 1986). These are arranged as two divergently transcribed pairs; one pair, *hsp16A*, consists of the *hsp16-1* and *hsp16-48* genes, which are separated by 348 bp of DNA containing the heat shock elements (HSEs) and TATA boxes for both genes (Fig. 1). These genes have been duplicated and inverted at a site 415 bp away (Russnak and Candido 1985) to produce a large inverted repeat with four genes in 3.8 kb. The *hsp16-2* and *hsp16-41* genes reside at a second site, *hsp16B*, and contain their

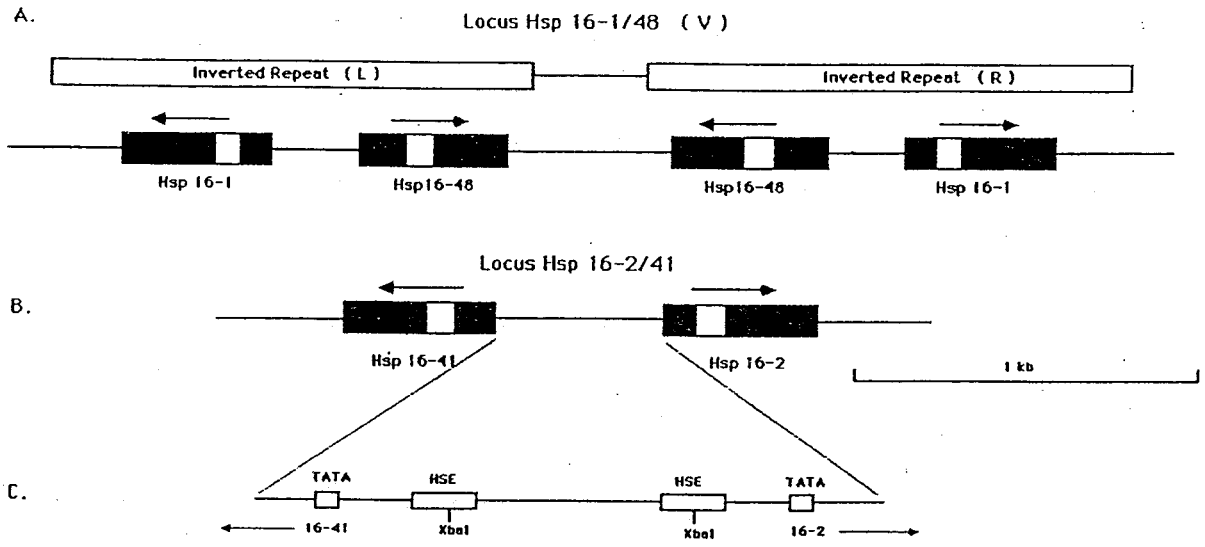


Figure 1. Organization of the *hsp16* loci of *C. elegans*. Re-printed from Candido et al. (1989). The open boxes represent introns while the filled boxes represent exons. The direction of transcription is indicated by the arrows. HSE, Heat shock element.

regulatory elements within a 346 bp intergenic region; unlike the *hsp16-48/1* genes, these are present as single copies (Jones *et al.*, 1986). Northern analysis indicates that expression of the *hsp16* genes is absolutely dependent upon heat shock (Russnak *et al.*, 1983). The polypeptides encoded by the *hsp16-1* and *hsp16-2* genes are 145 amino acids in length and share 92% identity; similarly, those of *hsp16-48* and *hsp16-41* are 143 residues in length and are 94% identical to each other. In contrast, the members of a given gene pair are less than 70% identical to each other (Jones *et al.*, 1986), and the identical residues are all encoded within the second exon; thus each gene pair encodes one of each class of HSP16s, which is defined by the sequence of exon 1. The HSP16s show a high degree of sequence similarity to both the small HSPs of *Drosophila* and to the α-crystallins of the vertebrate eye lens (Ingolia and Craig 1982; Southgate *et al.*, 1983).

Recently antibodies raised against a peptide corresponding to the carboxyl terminus of the HSP16-41 protein (anti-41/125-143) were found to react strongly with a 16 kD band and weakly with an 18 kD band on Westerns from heat shocked nematodes (Hockertz et al., 1991). Furthermore, antibodies against peptides corresponding to the carboxyl end of the HSP16-2 protein (anti-2/110-145) or to the HSP16-1 protein (anti-1/33-50) detected an 18 kD band on Westerns. The anti HSP16-2 antibody also reacted weakly with a 16 kD band. These results suggest that the 16 and 18 kD heat shock proteins of *C. elegans* are structurally related and raises the possibility that the *hsp16-1* and *hsp16-2* genes may actually encode the 18 kD heat shock proteins. Western blot analysis of a two dimensional gel using an anti-HSP16-2 antibody as a probe detected 13 heat induced polypeptides, all within the 16-18 kD region. Thus, the HSP16 and HSP18 polypeptides appear to exist *in vivo* in multiple isoforms which presumably arise from post-translational modifications. Purified native HSP16/18 proteins formed large aggregates of approximate mass of 400-500 kDa (Hockertz et al., 1991). This suggests that roughly 25-31 monomers of small HSPs potentially form these complexes, comparable to the stoichiometry of 32 monomers inferred for the murine HSP25 protein (Behlke et al., 1991).

The intergenic regions of the *hsp16-48/1* and *hsp16-41/2* gene pairs are highly conserved (86% identity) suggesting that elements other than the HSEs and TATA boxes may be important for their regulation. In particular, a stretch of alternating pyrimidines and purines and an inverted repeat may be candidates for control elements (Jones *et al.*, 1986). Kay *et al.* (1986) attempted to define the roles of these sequences in *hsp16* gene expression by transfecting mouse fibroblasts with various constructs of the *hsp16A* locus. These experiments demonstrated that a single HSE could induce bidirectional expression of the *hsp16-48/1* gene pair upon heat shock and that the number of HSEs present greatly affects the strength of the promoter. Disruption of the purine-pyrimidine stretch had no effect on expression (Kay *et al.*, 1986). However, these experiments were performed in a heterologous system and it is possible that the mouse transcriptional apparatus may not have recognized certain control elements in the nematode genes, other than the highly conserved HSEs.

Although the *hsp16A* and *hsp16B* intergenic sequences are 86% identical, the gene loci are differentially regulated (Jones *et al.*, 1989) and display differences in chromatin structure during heat shock (Dixon *et al.*, 1990). Nuclear run-on experiments determined that the transcription rates of the *hsp16* genes are similar. However, while the *hsp16-41/2* locus continues to transcribe messages at a high level after two hours of heat shock, transcription from the *hsp16-48/1* locus begins to decrease within an hour. This results in approximately a 14-fold difference in mRNA levels between the two loci in favour of *hsp16-41/2*. The disparity in mRNA levels is greatest at the embryo and L1 stages, suggesting the existence of developmental differences in *hsp16* gene expression (Jones *et al.*, 1989). The two gene clusters also exhibit differences in the timing of appearance and disappearance of DNAseI hypersensitive sites upstream of the HSEs during heat induction and recovery (Dixon *et al.*, 1990). These observations suggest there is some differential regulation of the two *hsp16* gene clusters.

9. The present study

In recent years, the volume of information regarding the function and regulation of the high molecular weight HSPs has increased dramatically. In contrast, still very little is known regarding the expression and function of the small HSPs, particularly those which are required only upon heat shock. The strict heat inducibility and the differential regulation observed in the *hsp16* gene family of *C. elegans* inspires a more careful and detailed analysis of these processes in this organism itself. Unfortunately, transformation of gene constructs designed *in vitro* into the homologous system was not a technique available to the *C. elegans* researcher until 1986 (Fire , 1986). My arrival in the laboratory in 1987 was fortuitously coincident with the accessibility of this technique and I consequently set out to evaluate the patterns of *hsp16* gene expression *in vivo*.

B: MATERIALS AND METHODS

1. Construction of hsp16-lacZ fusions

The *lacZ* gene fusions described in this study were constructed by my collaborators Donald Jones and Dennis Dixon, and by myself. The vector pPCZ1, abbreviated as 48.1C in the text, was constructed by Donald Jones. He inserted a 3500 bp HindIII-AflII fragment encompassing the *lacZ* gene (nucleotides 18 to 3518 of the expression vector pPD16.43 received from A.Fire, Carnegie Institution) by blunt end ligation into the *Hpa* I site (nucleotide 3565 of the published sequence) of the *hsp16-1* gene. Plasmid pPCZ1 contains a complete *hsp16-48/1* gene pair (hence 48.1C) extending from a *Bcl* I site at nucleotide 2280 to the *Bam* HI site at nucleotide 4186 in the published sequence (Russnak and Candido 1985). Subsequently I microinjected pPCZ1 and established transgenic lines. Construct 41.2C is the abbreviation for pDX16.31, and was the gift of Dennis Dixon who had inititiated expression studies of the *hsp16-41/2* gene pair while at the Carnegie Institute of Washington. Plasmid 41.2C contains *lacZ* as a 3200 bp Xbal-StuI fragment in the *Hpa* I site of *hsp16-2* (nucleotide 1690, Jones *et al.*,1986). The latter was contained in an *Eco* RI (nucleotide 540) to *Mbo* I (nucleotide 2870) fragment encompassing the *hsp16-41/2* gene pair. For comparative analysis of *hsp 16-48/1* with *hsp16-41/2* I microinjected 41.2C and established transgenic lines. Plasmid pPC16.48-1, referred to as 1.48E1 (48E1 represents fusion to exon 1 of *hsp16-48*), was constructed by inserting a *Sau* 3A fragment extending from nucleotides 987 to 1440 (Russnak and Candido 1985) into the *Bam* HI site of the nematode expression vector pPD16.51 (Fire *et al.*, 1990), such that the *hsp16-48* promoter was proximal to *lacZ*. pPC16.1-48, abbreviated as 48.1E1, contains the *Sau* 3A fragment in the inverse orientation such that the *hsp16-1* promoter is closest to *lacZ* . Both pPC16.48-1 and pPC16.1-48 were constructed by Don Jones. pPC16.1-48XBAI, abbreviated as 48.1XBE1 in the text, was constructed by inverting the Xbal fragment (nucleotides 1134 to 1289) between the two HSEs of the *hsp16-48/1* locus in the vector pPC16.1-48. A *Sau* 3A fragment extending from nucleotides 1121 to 1561 of the *hsp16-41/2* locus was cloned into the *Bam* HI

site of pPD16.51 such that the *hsp 16-41* promoter was proximal to *lacZ*, generating *pHS16.25* (abbreviated as 2.41E1 and also the gift of Dennis Dixon).

Plasmid pPC16.48-51, abbreviated as 48P, is a transcriptional fusion consisting of an *Mnl* I fragment (nucleotides 3085 to 3262) of the *hsp16-48/1* intergenic region cloned into the *HincII* site of the pPD16.51 polylinker. Since Mnl I cleaves DNA at a different site to its recognition sequence the resulting construct was analysed by the double stranded dideoxy sequencing method of Gaterman et al. (1988) to determine the exact sequence at the *hsp16-48/*pPD16.51 junction. The sequence of the non-coding strand at the junction was *lacZ*-pPD16.51 HincII-GTC/GATTGGCTTATATACCC-*hsp16-48* . Plasmid pPC16.41-51, referred to as 41P in the text, is a transcriptional fusion consisting of a *Taq* I fragment extending from nucleotides 1169 to 1409 in the *hsp16-41/2* intergenic region, inserted at the *Acc* I site of pPD16.51; its expected and confirmed junction sequence is *lacZ*- pPD16.51 AccI-GT/CGAAGTTTTTTAGATGCACTAGAACAA-*hsp16-41*. Subsequently, all of these constructs were microinjected into *C. elegans* oocyte nuclei and transgenic lines established. A complete list of all strains generated and their genotypes is presented in Appendix.

2. Bacterial Transformations

Fusion constructs were transformed into competent DH5α cells which had been purchased from BRL or prepared locally by the method of Hanahan (1983). Typical transformation efficiencies were in the range of 10^7-10^8 per μg of plasmid. Transformation mixtures were plated on YT Ampicillin plates and incubated overnight at 37°C for colony growth.

3. Double stranded sequencing of lacZ fusions

Plasmid DNA was isolated from *E. coli* for sequencing using a modified version of a method devised by Gaterman *et al.*(1988). Bacteria (1.5 ml) were pelleted at 10,000 rpm (8160 RCF) for one minute, resuspended in 100 μl of lysis buffer (8 % sucrose; 5 % Triton

X-100; 50 mM EDTA ; 0.5 mg/ml lysozyme) and boiled for 2 minutes. Cellular debris was pelleted by centrifugation at 4°C for 15 minutes. After removal of the pellet, the DNA was precipitated from the supernatant with 100 µl of cold isopropanol followed by another centrifugation at 4°C. The DNA pellet was washed, dried and resuspended in 40 µl of dH_2O. Typically this procedure yielded 20 µg of DNA.

Prior to sequencing, the plasmid was denatured: 8 µl of plasmid DNA (approx. 4 µg) was mixed with 1 pmol of M13 universal forward or reverse primer, 11 µl of dH_2O, and 2 µl of 2M NaOH , and boiled for two minutes. Upon removal from the boiling bath the denatured DNA was precipitated with 2 µl of 3M NAOAc, pH 6.8 and 50 µl of 95 % ethanol and incubated for 10 minutes at -20°C . Subsequently the DNA was pelleted, washed with 70 % ethanol, dried and resuspended in 7 µl of sterile dH_2O.

To anneal the primer to the template 2 µl of 5 X sequencing buffer (200 mM Tris, pH 7.5; 100 mM $MgCl_2$; 250 mM NaCl) and an additional 1 pmol of primer were added, the mixture heated at 65°C for 2 minutes and then allowed to cool to room temperature gradually over the course of about 30 minutes. Sequencing utilized commercial nucleotide mixes (US Biochemical), ^{35}SdATP and modified T7 DNA polymerase (Sequenase version 2.0-US Biochemical) and followed procedures outlined by the manufacturer. Samples were boiled prior to loading on a 6 % polyacrylamide gel and electrophoresed at 20 mA (1500 volts). Since the weak ^{35}S β emissions cannot penetrate through the urea in denaturing polyacrylamide gels, gels were washed in a solution of 5 % methanol; 5 % acetic acid to remove the urea before being dried and exposed to film.

4. Maintenance of strains

C. elegans strains were maintained on NG plates seeded with *E. coli* strain OP50 at 20°C, essentially as described by Brenner (1974).

5. Establishment of transgenic C. elegans strains

DNA was prepared for injection by a modified alkaline lysis procedure (Birnboim and Doly 1979) which included a lithium chloride precipitation step to remove RNA (Sambrook *et al.*, 1989). The DNA was finally suspended in an injection solution consisting of 2 % polyethylene glycol 6000; 20 mM potassium phosphate; 3 mM potassium citrate-pH 7.5 (Fire, 1986). Injection needles were pulled from glass capillaries (Cat. No. 1B100F-6, World Precision Instruments Inc., New Haven, Conn.) using a Frederic Haer & Co. micropipette puller, and filled with equimolar amounts of the selectable and test plasmids (40-200ng/µl of each). Wild type *C. elegans* oocytes were injected at a magnification of 400X using a Zeiss IM35 microscope equipped with Nomarski optics, and a Leitz micromanipulator according to the method of Fire (1986).

6. Selection of Transformed Progeny

Progeny transformed with the *unc-22* antisense vector, pPD10.41 were identified by visual inspection for twitching (Moerman and Baillie 1979). Strong *unc-22* mutants shake considerably but to retrieve weaker phenotypes transgenics were selected in the presence of 1 % nicotine (Moerman and Baillie, 1979). Selection of progeny transformed with pRF4, a plasmid containing the *rol-6* dominant allele su1006, was simply by visual inspection (Kramer *et al.*, 1990; Mello, personal communication) . The antimorphic allele su1006 encodes a mutant collagen and these mutants possess an altered body cuticle which forces the animal to roll onto its right side.

Each selected transgenic animal was placed on a separate plate for propagation. A line was considered established after the selectable phenotype had been sucessfully propagated for three generations.

7. Viable freezing of transgenic strains

Nematodes were preserved for long term storage by viable freezing in glycerol. Generally, animals were washed off a 9 x 1.5 cm plate almost devoid of food with cold 0.14 NaCl and centrifuged briefly at low speed (about 3000 rpm). The worm slurry was suspended in approximately 1 ml of M9 buffer and an equivalent volume of sterile Basal S containing 30 % glycerol was added. Aliquots of 400- 500 µl were frozen overnight in a -70 freezer before being transferred to liquid nitrogen for long term storage. Frozen stocks were thawed at room temperature for approximately 10 minutes before they were spotted on NG ; OP50 plates. Typically, survival rates of larval stages ranged between 60 % and 80 % while surviving adults were rendered sterile by the procedure and embryos failed to survive at all.

8. Heat Shock Conditions

Initially, the temperature response of selected transgenic lines was investigated to determine a suitable standard temperature for heat shock experiments. Transgenic lines maintained at 20°C were subjected to heat shocks of 25°C, 27°C, 29°C, 31°C and 33°C for two hours and then allowed to recover at room temperature for 15 minutes before staining for β-galactosidase activity. These experiments established that 33°C consistently induced the expression of transgenes. Routinely, transgenic lines were tested for *lacZ* expression by heat shocking at 33°C for two hours on pre-warmed NG plates spread with bacteria. A double heat shock consisted of two two-hour exposures at 33°C separated by a 30 minute recovery period at 20°C. Worms were subsequently allowed to recover for 15 minutes at 20°C before being washed off in distilled water, permeabilized by lyophilization and acetone treatment, and incubated in a histochemical stain containing 0.2 M $NaPO_4$-pH7.5, 1 mM $MgCl_2$, 5 mM potassium ferricyanide, 5 mM potassium ferrocyanide, 0.004 % SDS, 0.075 µg/µl kanamycin sulphate, 0.002 µg/µl DAPI, 0.24 µg/µl X-gal (Fire *et al.*, 1990) overnight at room temperature. Staining was often visible after only 15 minutes of incubation. As a control, some worms of each strain were incubated in staining solution without prior heat shock

treatment. No staining was ever observed in strains maintained at 20°C . Only three of 71 transgenic lines did not stain after heat shock, and in one of these instances Southern analysis indicated that the *lacZ* fusion had not been co-transformed with the *rol-6 su1006* plasmid (data not shown). After initial characterization, certain strains were selected for further analysis and subjected to heat shocks of varying duration. In addition, to study developmental stages, strains were synchronised by bleaching gravid adults to obtain eggs (Emmons *et al.*, 1979).

9. Identification of β-galactosidase staining cells

Transgenic animals were washed in distilled water, permeabilized by lyophilization and acetone treatment and incubated in a stain containing 5-bromo 4-chloro 3-indolylgalactosidase (Xgal) as described by Fire et al. (1990). Animals were permanently mounted in 80 % glycerol, 20 mM Tris-pH 8.0, 200 mM sodium azide prior to microscopic examination.

Nuclei were scored for the presence or absence of blue stain. The identity of stained cell types was determined by the size and shape of the nucleus and its position in the animal relative to other nuclei, as defined by Sulston (1976), Sulston and Horvitz (1977), Albertson and Thomson (1976), and Sulston et al. (1983). Counterstaining with DAPI (4,6-diamidino-2-phenylindole) allowed visualization of all nuclei in an animal using fluorescence microscopy (Ellis and Horvitz 1986). Frequently the staining of β-galactosidase activity was so strong that complete quenching of the DAPI was observed. This proved to be an asset as it allowed identification of some nuclei by exclusion. When cell identity could not be confidently determined, the staining nucleus was not scored.

10. Preparation of Transgenic Genomic DNA

High molecular weight transgenic DNA was isolated using a modified procedure of Reymond (1987). This method allowed for the rapid purification of DNA from numerous transgenic strains simultaneously. Nematodes were washed off crowded plates with cold 0.14

NaCl and collected in a microfuge tube. Usually a worm slurry of about 100 µl was used per DNA preparation. Digestion solution (200 µl of 50 mM Tris, pH 7.5; 20 mM EDTA; 1 mg/ ml Proteinase K) was added to the nematode slurry, mixed and then incubated at 65°C for 30 to 60 minutes. Subsequently, 200 µl of formamide was added to the microfuge tube and the mixture was incubated on ice for 10 to 15 minutes. Addition of 100 µl of 8M NH_4OAc was followed by another 10 minute incubation on ice. After RNA and cellular debris were pelleted by centrifugation at 10,000 rpm for 3 minutes, the supernatant was precipitated with 1 ml of cold 95 % ethanol and incubated on dry ice for 10 minutes. The DNA was pelleted by centrifugation at 12,000 rpm (11,750 g) for 5 minutes and the pellet resuspended in 400 µl of TNE (5 mM Tris, pH 7.5; 5 mM NaCl; 0.2 mM EDTA). Addition of 50 µl of 8M NH_4OAC and 1 ml of 95 % ethanol re-precipitated the DNA, and after incubation on dry ice and centrifugation, the pellet was finally resuspended in 50-100 µl of TNE.

DNA samples were quantified spectrophotometrically by absorbance at 260 nanometers (Sambrook et al., 1989). Typically, 50 to 100 µg of DNA was isolated from one plate (9 cm x 1.5 cm) of worms. The quality of the DNA samples was analysed by electrophoresis in 1 % agarose gels.

11. Southern Transfer and Analysis of Transgenic DNAs

Digested DNA was transferred to nylon membranes, Zetaprobe (Bio-Rad) or Hybond-N (Amersham) by Southern blotting (Southern, 1975) modified according to the manufacturers' instructions. After transfer was complete, membranes were rinsed briefly in 2 X SSC and dried at room temperature. Zetaprobe membranes were used without further treatment whereas DNA was permanently fixed to Hybond-N membranes by U.V. crosslinking.

12. DNA Dot Blot Procedures

Genomic and plasmid DNA was transferred to Hybond-N membranes using a DOT blot apparatus (Bio-Rad) and applied vacuum according to the following procedure. Six hundred ng

of DNA was suspended in 50 µl of TE (10 mM Tris, pH 8.0; 1 mM EDTA) and denatured by the addition of 0.1 volumes of 3M NaOH. The mixture was incubated for one hour at approximately 65°C and then cooled to room temperature. Subsequently, one volume of 2M NH_4 OAC, pH 7.0 was added. Prior to application, the samples were diluted to 400 µl with 2M NH_4OAc. All subsequent serial dilutions used 2M NH_4OAc. The membrane was prepared for installation in the apparatus by hydration in deionized water followed by a rinse in 1M NH_4OAc, pH 7.0. Prior to sample application the wells were rinsed with 500 µl of 1M NH_4OAc, pH 7.0. After transfer was complete the membrane was rinsed briefly in 2 X SSC and dried at room temperature. Hybridization procedures followed those previously described. Signal quantitation was by densitometry on a LKB Ultroscan XL laser densitometer. Copy number in transgenic strains was determined from points which fell on the linear part of the standard curve derived from dilutions of plasmid DNA.

13. Labelling of Radioactive Probes

Probes of high specific activity were prepared for hybridizations by either nick translation or primer extension. Nick translations of pPD16.43 followed protocols described in Sambrook et al.(1989). Typically probes were labelled with ^{32}PdATP and ^{32}PdCTP to a specific activity between 0.5 and 1 x10^8 cpm/µg using this method. Primer extension labelling of M13 templates containing the entire *lacZ* gene followed procedures outlined by Russnak and Candido (1985). This method generated probes with specific activities of approximately 10^8 cpm per µg of template. DNA molecular weight markers were created by end labelling lambda DNA that had been digested with Hind III. In this procedure 1 µg of template was incubated at room temperature for 10 minutes in a solution of 10 mM Tris, pH 7.5; 10 mM $MgCl_2$; 1 mM DTT; 10 µCi of ^{32}PdATP and 5 units of *E. coli* DNA polymerase Klenow fragment. Addition of EDTA to a final concentration of 20 mM stopped the reaction.

All radiolabelled probes were purified by chromatography on G-50 Sephadex spin columns. Usually 10,000 cpm of radioactive molecular weight marker DNA was loaded directly

into slots of agarose gels prior to electrophoresis. For most experiments, 10^7 cpm of nick translated or primer extended probe was added per ml of hybridization solution.

14. Hybridization Conditions

Usually membranes were incubated in 10 to 15 ml of hybridization solution for 4 hours prior to addition of the radioactive probe. Hybridization solution for Zetaprobe membranes consisted of 1 % SDS; 0.5 % BLOTTO; 50 % formamide; 4 X SSPE. Hybridization solution for HYBOND membranes consisted of 5 X SSPE; 5 X Denhardt's; 50 % formamide; 0.5 % SDS. In some experiments 20 µl of 10 mg/ml sheared single stranded calf thymus DNA was added to the hybridization to reduce non-specific binding.

After overnight incubation at 42°C membranes were washed in SSC and SDS with the final wash consisting of 0.1 X SSC/ 0.1 % SDS at 50 °C. Subsequently membranes were exposed to X-OMAT-AR film (Kodak) for autoradiography.

C: RESULTS

1. Construction of hsp16-lacZ fusions and selection of transformants

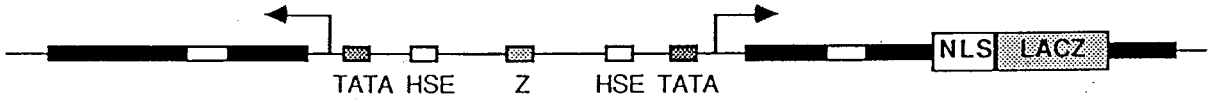
The *hsp16-lacZ* fusions used for transformation are described in Fig.2. For ease of reference, the plasmids will be referred to using an abbreviated terminology. Full designations for the constructs are given in "Methods". Constructs 48.1C and 41.2C contain the complete *hsp16-48/1* and *hsp16-41/2* gene pairs, and are in-frame translational fusions of the *lacZ* coding region with the second exon of *hsp16-1* and *hsp16-2*, respectively. Constructs 48.1E1, 1.48E1 and 2.41E1 are translational fusions to *lacZ* in which a Sau3A fragment containing the intergenic region of a gene pair was fused in-frame to *lacZ* at the Sau3A site in exon 1 of *hsp16-1*, *hsp16-48* and *hsp16-41*, respectively. Plasmid 48.1XBE1 is a translational fusion in which the Xba I fragment of the 48.1E1 promoters was inverted. Finally, 48P and 41P are transcriptional fusions containing the promoter (i.e. TATA box, HSEs and some upstream sequence) of *hsp16-48* and *hsp16-41*, respectively. All constructs contained the SV40 nuclear localization signal fused to the beginning of the *lacZ* coding region (Fire 1986).

Transgenic *C. elegans* strains were constructed by microinjecting oocytes of wild-type hermaphrodites with DNA of the desired construct, together with a selectable marker. Selection of transformants co-injected with pPD10.41, an *unc-22* antisense vector (Fire *et al.*, 1991), was by visual inspection, as *unc-22* mutants display a twitching phenotype; alternatively, animals were examined in the presence of nicotine, which enhances twitching and allows selection of weaker phenotypes (Moerman and Baillie 1979). Animals which had been transformed with *rol-6 su1006* were identified by their right rolling phenotype (Kramer *et al.*, 1990; Mello, personal communication).

The transformed lines thus obtained transmitted the marker phenotype at a frequency of 20-95%.

Figure 2. Construction of *hsp16-lacZ* transgenes. 48.1C consists of a complete *hsp16-48/1* gene pair, including the 5' and 3' non-coding sequences of both genes, with the *E. coli lacZ* gene inserted in-frame into a unique *Hpa*I site in the second exon of *hsp16-1*. 41.2C is the homologous construct using the *hsp16-41/2* gene pair. The exon 1 fusions 1.48E1, 48.1E1 and 2.41E1 were constructed by cloning a *Sau*3A fragment containing the intergenic sequence of *hsp16-48/1*, *hsp16-1/48* and *hsp16-2/41*, respectively, into the *Bam*HI site of the *lacZ* expression vector pPD16.51 (Fire et al., 1991, in press). These constructs include the first 15 (2.41E1 and 1.48E1) or 17 (48.1E1) amino acid residues of the respective HSP16. In 48.1XBE1 the XbaI fragment between the HSEs of 48.1E1 was inverted. The polyadenylation signal following the *lacZ* gene in plasmid pPD16.51 is derived from the 3' non-coding region of the myosin gene, *unc-54* (Epstein et al., 1974; Fire et al., 1990). The arrows indicate the direction of *lacZ* transcription. 48P and 41P are transcriptional fusions which remove the HSEs and TATA boxes of the *hsp16-1* and *hsp16-2* genes, respectively, but retain a single promoter (*hsp16-48* or *hsp16-41*). NLS, SV40 nuclear localization signal.

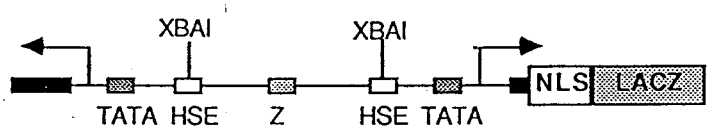
Translational Fusions (Complete):



Hsp16-48——**48.1C**——Hsp16-1

Hsp16-41——**41.2C**——Hsp16-2

Translational Fusions (Exon 1):



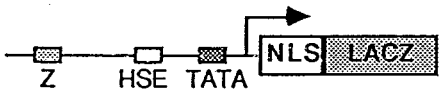
Hsp16-48——**48.1E1**——Hsp16.1

Hsp16-1 ——**1.48E1**——Hsp16-48

Hsp16-2 ——**2.41E1**——Hsp16-41

Hsp16-48/Xba-Hsp16-1—**48.1XBE1**—Hsp16-48-Xba/Hsp16-1

Transcriptional Fusions:



Hsp16-48——**48P**

Hsp16-41——**41P**

2. Expression of the hsp16-lacZ transgene is temperature dependent.

Transgenic lines were heat shocked for two hours and stained with a solution containing Xgal (see Methods) to assay for expression of the *hsp16-lacZ* fusion genes. Figure 3 shows staining of a transgenic line carrying the 48.1C transgene before and after heat shock. No staining was observed in control worms, indicating that expression of the transgene was dependent upon heat shock. Eleven lines carrying the 48.1C transgene were examined and all gave identical results. Not all heat shocked worms within a strain stained, presumably due to incomplete germ line transmission of the extrachromosomal array. Moreover, mosaicism was frequently observed among the animals which did stain. This most likely results from random loss of the array during mitosis.

To determine if the selection plasmid phenotype absolutely correlated with incidence of *lacZ* expression, rollers or twitchers and phenotypically wild type animals from the same line were picked to separate plates, heat shocked for two hours and then stained with Xgal. In all three strains tested, 100 % of the selected rollers and twitchers stained after heat shock. In two of the three strains (one roller and one twitcher), one third of the phenotypically wild type animals also stained after heat shock, even though the twitcher strain had been screened in the presence of nicotine, implying that β-galactosidase activity is a more penetrant phenotype. Since the β-galactosidase assay detects enzyme activity by a sensitive histochemical procedure, it is conceivable that a greater quantity of the *rol-6 su1006* gene product may be required to produce abnormal cuticle and that more *unc-22* antisense RNA may be needed to interfere with the wild type product to mimic loss-of-function alleles. Alternatively, since each of the wild type selection genes are expressed in specific tissues, the extrachromosomal array may have been lost in those particular tissues during development but not in others.

Figure 3. Heat shock dependence of *lacZ* expression. In situ staining of β-galactosidase activity in PC6, a strain carrying an extra-chromosomal genetic element composed of mixed arrays of 48.1C and the *rol-6* selection vector pRF4, is shown. Magnification 100 X. a) Staining of a mixed population of animals without prior heat treatment. b)Staining following a two hour heat treatment at 33°C. Staining of intestinal, body muscle, hypodermal and pharyngeal nuclei is clearly visible.


a


b

3. Establishing Standard Heat Shock Conditions for Experiments

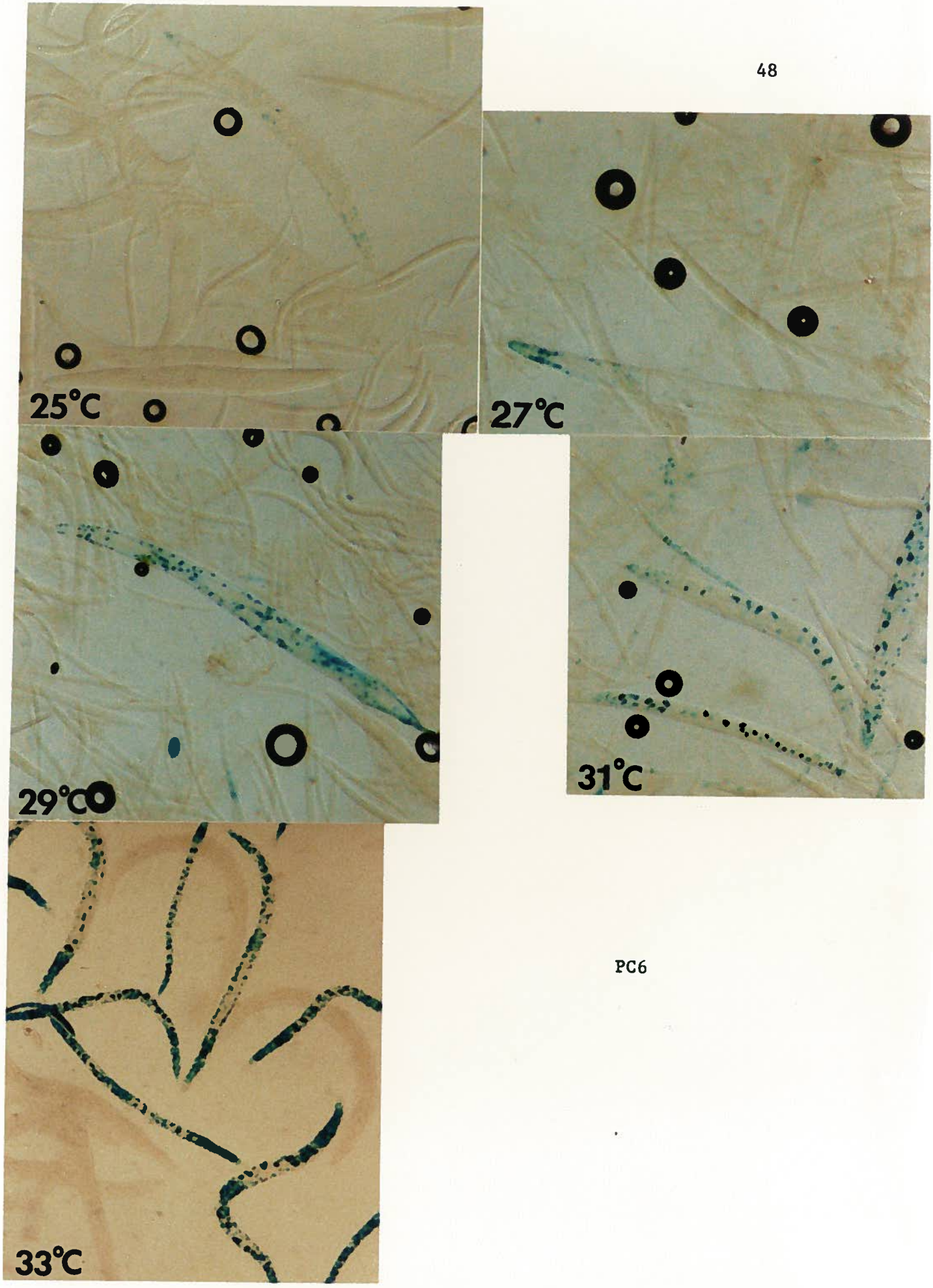
Temperature studies of transgenic lines established that the minimum temperature at which any staining was observed was typically between 29° and 31°C, but staining was sporadic and infrequent over this temperature range (Figure 4). In one strain (PC6), 1% of embryos stained at 25°C. Also, on one occasion an adult stained moderately at 25°C (Figure 4). By 31°C an extensive expression pattern was observed in PC6 animals. However, somatic expression in PC20 animals remained patchy at this temperature (Figure 4). Hence, 33°C was chosen as a suitable heat shock temperature since it provided good expression of the *lacZ* transgene in all strains.

4. The transgenes 48.1C and 41.2C are expressed in a tissue general manner

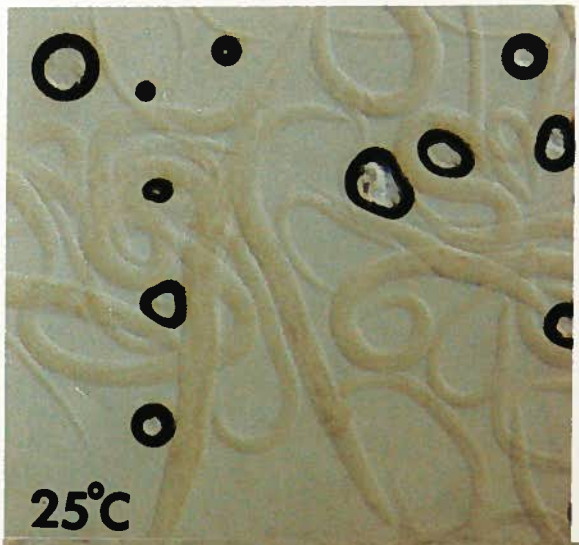
Expression of both the 48.1C and 41.2C transgenes was consistently observed in embryos from gastrulation onward, in all larval stages, and in adults. In post-embryonic stages, both transgenes were expressed in most tissues (Fig. 5). Curiously, the nuclear localization signal (NLS) in the 41.2C element did not effectively target β-galactosidase to the nucleus, even though the *lacZ* gene had been inserted into *hsp16-2* of 41.2C at the analogous position (Fig. 2) to the insertion in the *hsp16-1* gene of 48.1C. This suggests that the NLS of 41.2C is non-functional in the β-galactosidase fusion protein, while remaining active in the 48.1C fusion protein.

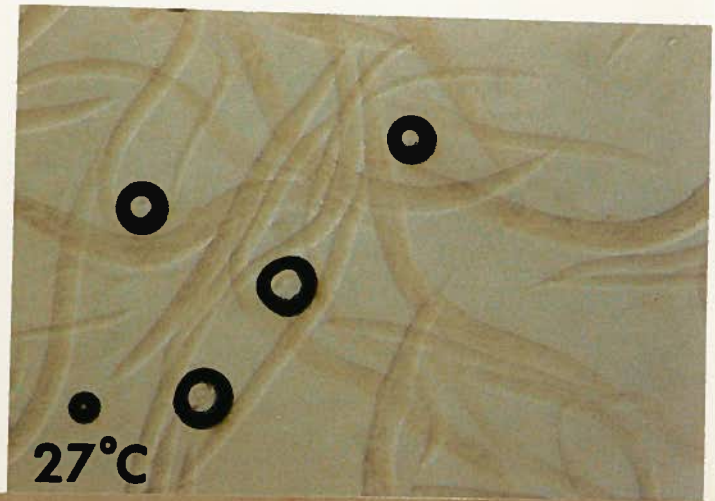
While the 41.2C transgene is expressed in numerous tissues especially the pharynx and gut, it is not expressed in the gonad. 48.1C transgenic animals also express the transgene in a tissue general fashion, and in this case the nuclear localization of the stain greatly facilitates cell identification. Examples of every cell type have been observed in strains carrying this transgene, including neurons, muscle, intestine, and hypodermis. Only the germ line failed to express this fusion gene. Figure 6 shows some examples of the tissue distribution of expression seen with these constructs. Tissues expressing the *hsp16-lacZ* fusion in these strains include nerve ganglia in the head and ventral cord; hypodermal nuclei of the lateral hypodermis, vulva,

Figure 4. Expression of β-galactosidase is proportional to temperature in *hsp16-lacZ* transgenic animals. Cultures of PC6, a 48.1C/pRF4 strain and PC20, a 2.41E1/pPD10.41 strain were heat shocked on pre-warmed plates at varying temperatures (25°C, 27°C, 29°C, 31°C, and 33°C) for two hours, recovered for 15 min. at room temperature, and then assayed for β-galactosidase activity.



PC6





PC20

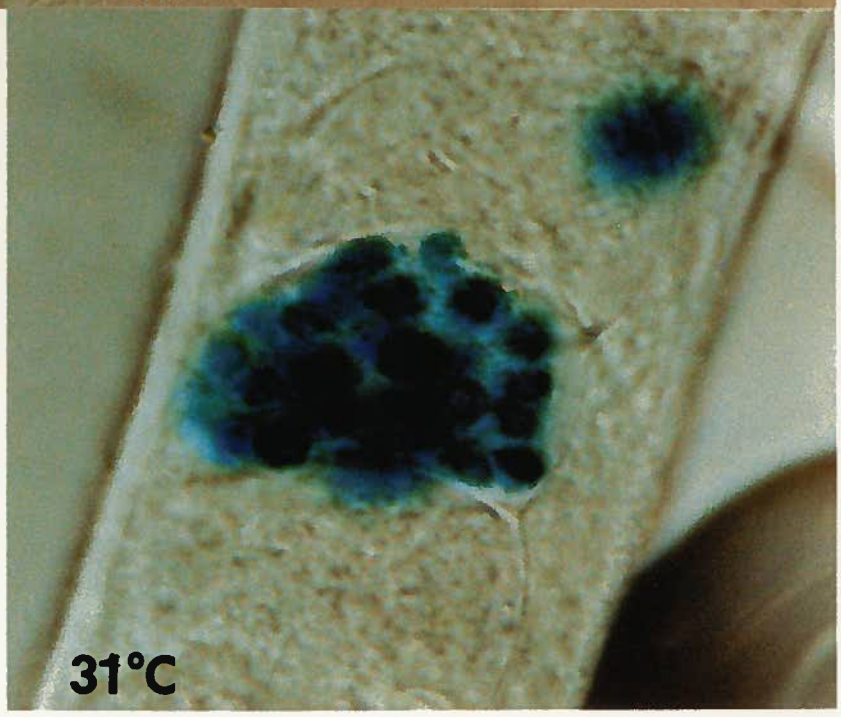


Figure 5. *LacZ* expression of the 48.1C and 41.2C transgenes in response to heat is tissue general. a) Brightfield image at 200X magnification of an adult hermaphrodite carrying the 48.1C transgene. Blue precipitate is seen in body muscle, intestinal, hypodermal, neural and pharyngeal nuclei. b) Staining in a late gastrulation/ comma stage embyo. Thirty-six staining nuclei were counted. Magnification 1250X. c) Unlocalized intracellular expression in a 41.2C adult hermaphrodite. Magnification 200X. Although the stain is diffuse, expression is evident in the intestine, pharynx, vulva, and ventral hypodermis. The germ cells of the gonad are clearly not stained.


a


c

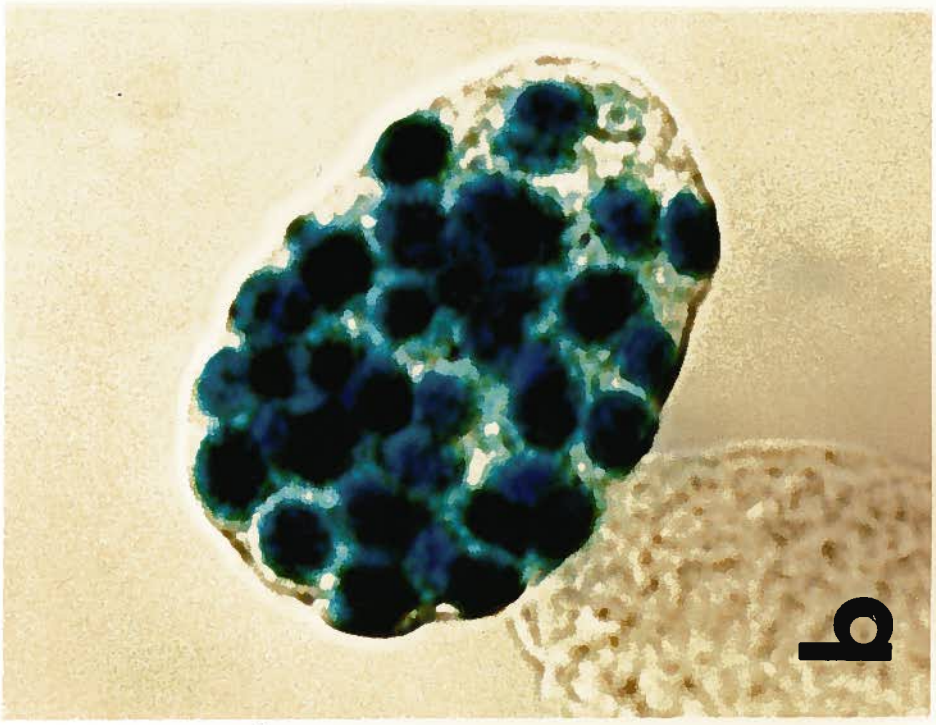
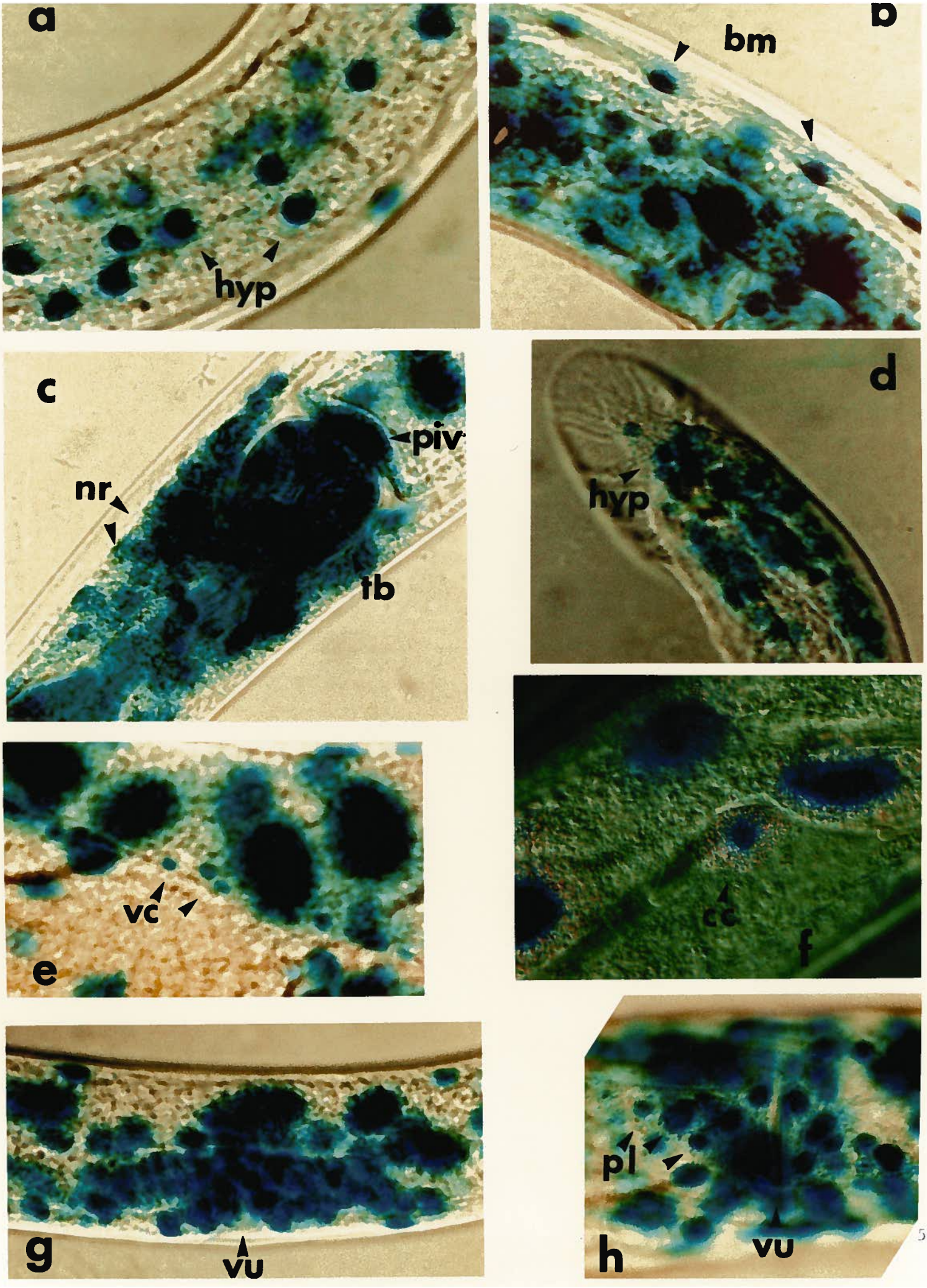

b

Figure 6. Cell types expressing *hsp16-lacZ* transgenes upon heat shock. a) Staining of nuclei of the lateral hypodermis (hyp) in a 1.48E1 animal. b) Body muscle (bm) nuclear expression in a strain carrying the 48.1C transgene. Stained intestinal and hypodermal nuclei are visible in the background, out of the plane of focus. c) Expression in the head of a 2.41E1 animal. Neurons of the nerve ring (nr) are stained, as well as those of the ventral and retro-vesicular ganglia. The pharyngeal-intestinal valve (piv) as well as muscle, epithelial and marginal nuclei in the terminal bulb (tb) are indicated. d) Expression of the 48.1C transgene in an adult male. In addition to general body muscle and hypodermis, V and T derived nuclei of the rays are stained. e) Expression of the 48.1C transgene in neuronal nuclei of the ventral cord (vc). f) A coelomocyte (cc) in the pseudocoelomic cavity of a 48.1 C adult hermaphrodite. g) Vulval expression, lateral view. P derived hypodermal nuclei of the vulva as well as mesodermal tissue above the vulva (vu) are stained. h) Vulval expression, ventral aspect. Neuronal and hypodermal nuclei of the P lineage (pl) are indicated.


a
hyp
b
bm
c
piv
nr
tb
d
hyp
e
vc
f
cc
g
vu
h
pl
vu

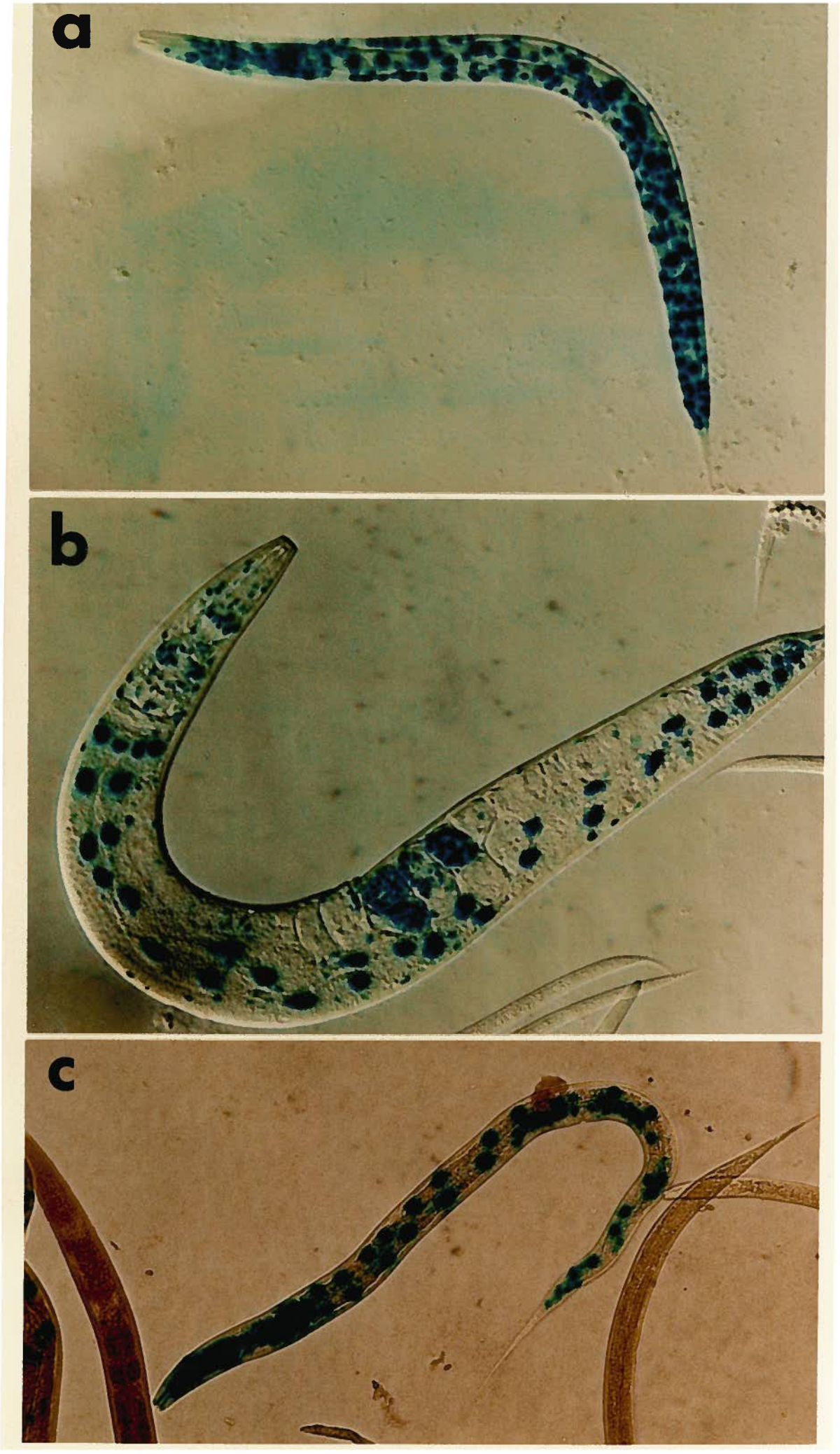
head and tail; body wall and pharyngeal muscle, somatic gonad, mesodermal tissue such as coelomocytes, and muscles involved in defecation (the latter not shown). These results indicate that the *hsp16-48/1* and *hsp16-41/2* gene pairs are both expressed in a heat inducible, tissue general manner *in vivo* and that individual cells are capable of expressing both gene pairs. On the other hand, subtle differences in the patterns of expression between the two constructs were observed; in particular, 41.2C animals tended to stain most intensely throughout the pharynx while 48.1C animals demonstrated conspicuous expression in body wall muscle. The latter characteristic was never unequivocally observed in 41.2C animals. However, these differences were difficult to document since 41.2C animals displayed diffuse staining. Thus I examined a different series of *lacZ* fusions of these gene pairs in order to assess potential differences in expression more accurately (see below).

5. Quantitative differences in the tissue specific expression of the hsp16 gene pairs.

In order to determine if differences exist in the tissue distribution of the *hsp16-48/1* and *hsp 16-41/2* gene products, I established 8 transgenic lines containing the fusion 1.48E1 and 8 containing 2.41E1 (Fig. 2). In addition, to assess whether the orientation of the intergenic region relative to *lacZ* affected expression, I also established 6 lines carrying 48.1E1, the inverse of 1.48E1. Initially, all of these strains were heat shocked for two hours at 33°C, stained and qualitatively analyzed. The typical patterns of expression observed in these lines are shown in Figure 7. Again, each transgene was expressed in a wide variety of tissues. However, certain differences were apparent. For example, all of the strains containing 2.41E1 tended to stain very well in the intestine and pharynx while 1.48E1 strains consistently showed better expression in muscle. Such differences were not exclusive between constructs but rather quantitative in nature. Some 1.48E1 animals did stain in pharyngeal muscle, but at a frequency lower than that of 2.41E1 animals.

To quantify these differences, I selected two or more strains of each construct and subjected mixed age populations of animals to varying durations of heat shock at 33°C. The

Figure 7. Expression of *hsp16-exon1* fusions. Magnification 170X. a) L4 animal carrying the 1.48E1 transgene. Staining is intense in hypodermal nuclei of the main body syncitium, head and tail, and in body muscle. Some intestinal expression is also visible. b) Adult hermaphrodite carrying the 48.1E1 transgene. Expression is prominent in embryos and in the intestine while hypodermal expression is weak. c) L4 animal carrying the 2.41E1 transgene. Staining is intense in the intestine and pharynx, and in nerve ganglia in the head.


a
b
c

results of this experiment are summarized in Table 1. Only the tissues showing the most striking differences are included for comparison. For example, all transgenic lines showed some expression in the mesodermally derived tissue of the somatic gonad, and in hypodermal nuclei of the vulva, but at frequencies which were not significantly different, so data from these tissues were not included in Table 1.

A number of general observations can be made with regard to the results in Table 1. Firstly, relatively short periods of heat shock were required to induce visible expression of the transgenes. Even a fifteen minute heat shock without a recovery period was often adequate to give visible expression of β-galactosidase in some strains. Longer periods of heat shock, however, increased the proportion of tissues staining. Mosaicism of expression was still observed, however, even after a double heat shock, since 100% staining of all tissues was never observed. It should be noted that tissue expression was simply scored for the presence or absence of any staining, i.e. all nuclei of a given tissue were not necessarily stained.

While all constructs showed a high frequency of intestinal staining at longer heat shocks, the frequency of expression dropped to 67% in 1.48E1 animals *versus* 83% and 95% for 48.1E1 and 2.41E1 animals, respectively, for 15 minute heat shocks. A striking difference was observed in pharyngeal expression. Both 48.1E1 and 1.48E1 transgenics stained infrequently in the pharyngeal intestinal valve. 2.41E1 worms, in contrast, consistently expressed the transgene in this tissue as well as in the rest of the pharynx, including pharyngeal muscle, marginal and epithelial cells. Expression in the pharynx for the other constructs was less consistent and often limited to the gland nuclei of the terminal bulb. Neural expression was most prevalent in the 2.41E1 transformants and often included ganglia of the nerve ring, the ventral and retrovesicular ganglia, and the lumbar and preanal ganglia in the tail. Neural expression in 48.1E1 animals was usually limited to head ganglia and was less prominent than in 2.41E1 transformants; it was rarely observed in 1.48E1 worms. However, expression in the latter animals was difficult to document due to the large number of intensely staining hypodermal and body muscle nuclei in the head.

Table 1. *Tissue distribution of β-galactosidase staining in transgenic C. elegans heat shocked for 15-120' at 33°C.*

TREATMENT	CONSTRUCT	INTEST	PIV	PHA	NERVE	HYP	BM
120' HS	2.41E1	97	90	90	53	70	40
30' REC	1.48E1	80	20	47	-	77	97
120' HS	48.1E1	90	0	30	30	70	75
	48.1XBE1	72	0	5	0	45	80
120' HS	2.41E1	93	77	83	33	50	57
	1.48E1	87	30	80	-	77	93
	48.1E1	95	0	25	20	60	70
	48.1XBE1	95	0	50	35	55	80
90' HS	2.41E1	90	90	87	40	43	43
	1.48E1	80	20	83	-	73	90
	48.1E1	95	0	20	10	35	55
	48.1XBE1	100	15	75	30	35	75
60' HS	2.41E1	93	70	87	30	53	53
	1.48E1	73	17	47	-	50	83
	48.1E1	95	20	55	40	45	60
	48.1XBE1	100	20	85	25	40	75
30' HS	2.41E1	100	20	53	7	27	20
	1.48E1	67	7	33	-	57	90
	48.1E1	100	0	35	25	25	70
	48.1XBE1	95	0	40	20	40	75
15' HS	2.41E1	95	13	23	3	7	0
	1.48E1	67	0	23	-	43	67
	48.1E1	83	0	0	0	20	33
	48.1XBE1	95	0	55	35	25	40

Three strains of each of 2.41E1 (PC19, 20, and 30) and 1.48E1 (PC16, 31, and 33), and two strains of each of 48.1E1 (PC52 and 55) and 48.1XBE1 (PC62 and 67) were selected for this analysis. Ten or more animals of each strain were scored for each datum. Numbers represent the percentage of staining transgenic animals which stain in the specified tissue. HS, Heat shock; REC, Recovery; INTEST, Intestine; PIV, Pharyngeal intestinal valve; PHA, Pharynx; HYP, Hypodermis; BM, Body muscle; -, Unable to determine.

All constructs seemed to be expressed at high frequency in hypodermal nuclei following extensive heat shocks (double or two hour shocks), but expression was far more consistent at shorter heat shock times for the 1.48E1 transformants. These strains were also far more likely to show expression in hypodermal nuclei of the lateral bands and of the head and tail. More significant is the prominent body muscle expression consistently observed in 1.48E1 animals as opposed to 2.41E1 and 48.1E1 animals. Thus, while the 2.41E1 construct seemed to confer greater expression in pharyngeal muscle, 1.48E1 was expressed much better in body muscle.

Collectively, these results suggest that there are differences in the priority of tissue expression between gene pairs. Surprisingly, the inverse construct 48.1E1 showed characteristics more reminiscent of those of 2.41E1, particularly with regard to intestinal and neural expression. On the other hand, the 48.1E1 construct did not express extensively in the pharynx as did the 2.41E1 fusion. Thus qualitative as well as quantitative differences in expression can be seen both between and within gene pairs.

I was curious about sequences which might confer tissue specificity for expression. The region of alternating purines and pyrimidines and the potential cruciform structure in the centre of the intergenic regions were most intriguing. To this end the XbaI fragment of the *hsp16-48/1* intergenic region was inverted in the 48.1E1 to generate 48.1XBE1 (Fig. 2) in order to see if the pattern of expression could be altered to more closely resemble that of the 1.48E1 construct (Table 1). Generally the 48.1XBE1 animals resembled most closely the 48.1E1 strains suggesting that the orientation of the sequences between the HSEs has no effect on tissue preference. However, the frequency of pharyngeal expression was slightly elevated in 48.1XBE1 animals, approaching the values observed for 1.48E1 animals, suggesting that there may be moderate influences within this region which affect pharyngeal expression.

6. Elimination of one promoter from an hsp16 gene pair significantly reduces somatic tissue expression without affecting embryonic expression.

Multiple HSEs can function cooperatively to induce high levels of heat induced transcription (Shuey and Parker 1986; Xiao *et al.*, 1991). I therefore wished to see if removing part of the intergenic region containing the HSEs and TATA motif of one gene would affect the overall expression from the remaining promoter. Using stably transformed mouse cell lines, Kay *et al.* (1986) demonstrated that the level of transcription of the *hsp16-48/1* gene pair was generally proportional to the number of heat shock elements present in the intergenic region. Since I observed differences in the priority of tissue expression which were dependent on the orientation of the gene pair relative to the reporter gene, it was of interest to determine if these differences could be attributed to a specific element in the intergenic region. Transcriptional fusions which removed the *hsp16-1* (48P) or *hsp16-2* (41P) HSEs and TATA boxes from the intergenic regions of the *hsp16-48/1* and *hsp16-41/2* loci, respectively, were constructed (Fig. 8) and tested.

All transformed lines carrying the constructs 48P (4 lines analyzed) or 41P (4 lines analyzed) still expressed β-galactosidase at high levels in embryos (gastrula stage and later), while expression in the post embryonic stages was extremely limited (Fig. 8 and Table 2). Staining in larval and adult tissues was only consistent when long heat shocks were used (two hours or double two hour heat shocks separated by a recovery period). When somatic expression was observed in 48P transformants, it was often limited to body muscle and hypodermis (64 % of staining worms treated with a double heat shock) although some intestinal expression was also observed (14 % of animals treated with a double heat shock).

Transformants carrying 41P showed more somatic expression than 48P transformants, but considerably less than 2.41E1 strains. Again, much longer periods of heat shock were required to give consistently detectable somatic expression, while embryos required only a half hour heat shock. While the intensity of expression was reduced in 41P animals, the distribution of tissues affected resembled that of 2.41E1 worms. For example, after a double

Figure 8. Expression of the 48P transcriptional fusion in heat shocked transgenic animals. Magnification 200X. a) Embryos within the uterus intensely stained even after a short heat shock (33°C, 30 min.). b) The maximum somatic tissue expression seen in 48P animals. After a double two hour heat shock with an intervening 30 minute recovery period, this animal stained in the lateral hypodermal nuclei. c) Typical 48P hermaphrodite showing minimal somatic tissue expression after a double heat shock. Staining is seen in two intestinal nuclei, in gland nuclei in the terminal bulb, and in embryos.


a


b

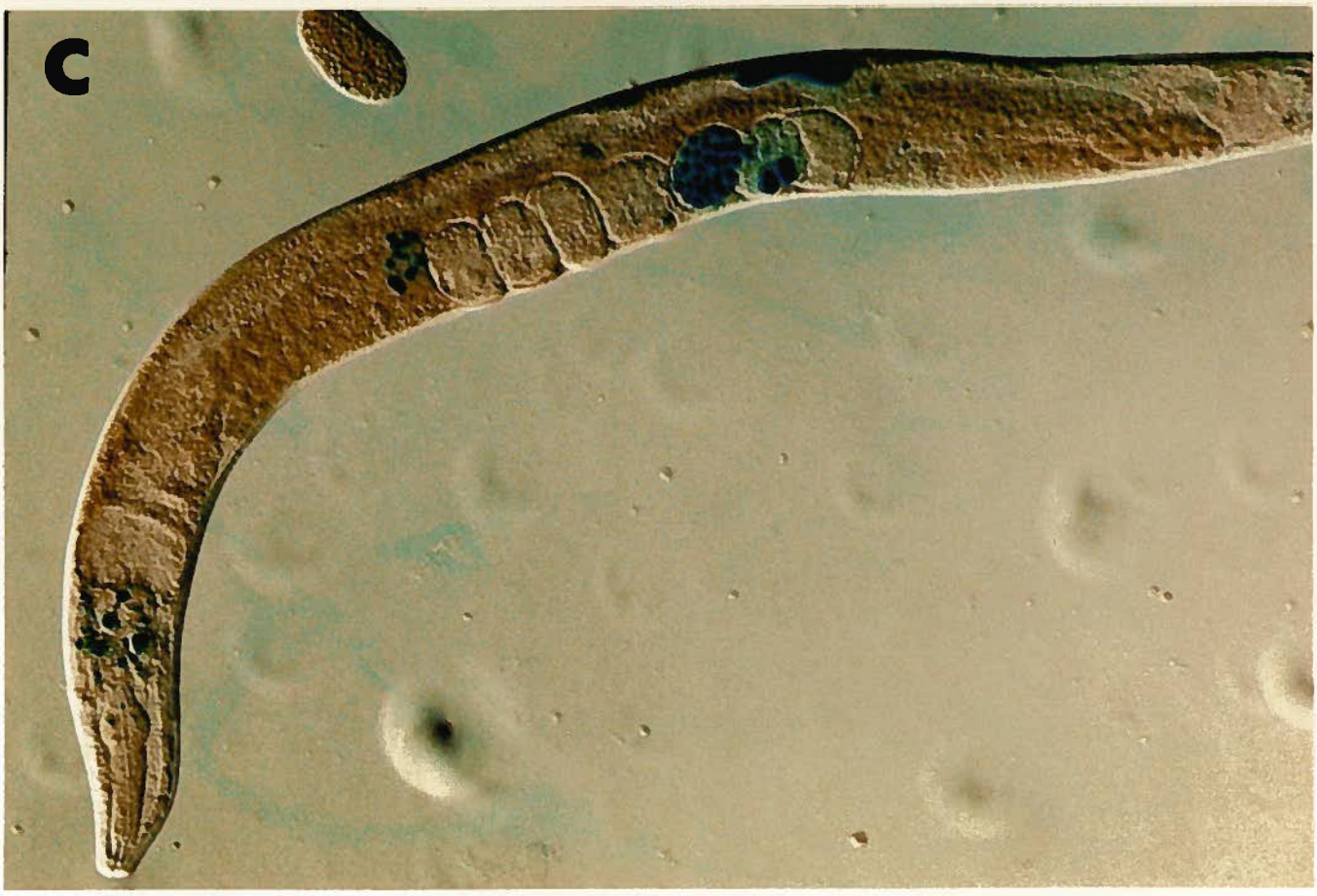

c

Table 2. *Distribution of β-galactosidase activity in transgenic strains carrying transcriptional fusions.*

For each datum, ten animals were scored for staining in the specified tissue unless otherwise indicated. TRT, Treatment; HS, Heat shock; REC, Recovery; INTEST.,Intestine; HYP., Hypodermis; PHA., Pharynx; PE, Post embryonic; -, not applicable.

TRT.	FUSION	STRAIN PC #	INTEST.	NERVE	MUSCLE	HYP.	PHA.	% PE STAGES	%EGGS
120'HS	41P	PC42	14/15	5/15	5/15	9/15	3/15	6 %	13 %
30'REC		PC45	14/20	4/20	5/20	6/20	6/20	36 %	40 %
120'HS		AVE	80 %	26 %	29 %	43 %	34 %		
	48P	PC38	4/17	0/17	12/17	2/17	0/17	14 %	20 %
		PC40	0/11	0/11	6/11	3/11	0/11	12 %	36 %
		AVE	14 %	-	64 %	18 %	-		
120'HS	41P	PC42	7	1	2	9	0	3 %	17 %
		PC45	2/11	1	7	8	2	19 %	22 %
		AVE	43 %	10 %	43 %	80 %	10 %		
	48P	PC38	6	0	8	0	0	14 %	35 %
		PC40	0	0	0	0	0	0 %	21 %
		AVE	-	-	-	-	-		
90'HS	41P	PC42	7	0	1	5	0	8 %	25 %
		PC45	5	1	7	6	5	8 %	4 %
		AVE	60 %	5 %	40 %	55 %	25 %		
	48P	PC38	0	0	1	0	0	11 %	13 %
		PC40	1	0	1	0	0	<1 %	29 %
		AVE	-	-	1 %	-	-		
60'HS	41P	PC42	9	3	4	2	1	3 %	7 %
		PC45	3	3	9	0	3	5 %	7 %
		AVE	60 %	30 %	65 %	10 %	20 %		
	48P	PC38	1	0	5	0	7	15 %	36 %
		PC40	0	0	0	0	0	0 %	12 %
		AVE	-	-	-	-	-		
30'HS	41P	PC42	10	2	4	3	0	6 %	8 %
		PC45	0	0	4	6	3	5 %	7 %
		AVE	50 %	10 %	40 %	45 %	15 %		
	48P	PC38	0	0	9	2	1	10 %	25 %
		PC40	0	0	0	0	0	0 %	10 %
		AVE	-	-	-	-	-		

heat shock, 80% of 41P animals stained in the intestine, 29% in body muscle and 43% in hypodermis, values somewhat lower but comparable in distribution to the corresponding 97%, 40%, and 70% observed after a double heat shock in 2.41E1 worms. Thus the removal of the upstream *hsp16-2* promoter sequences and/or the removal of the *hsp16-41* translational start site in the 41P construct resulted in a reduction of the overall intensity and frequency of expression but did not seem to restrict tissue distribution. It should be noted that the fragment of the intergenic region in the 41P transcriptional fusion is somewhat larger than that of 48P, and that the latter lacks the transcriptional start site of the original gene; this probably explains the much better expression of the 41P construct.

7. Southern Analysis of Transgenic Strains

Southern analysis of genomic DNAs digested with Apa I and Sma I verified the presence of the *lacZ* fusion in transgenic "twitching" strains while no significant hybridization signal was observed in control (N2) DNA (Figure 9). However, nick translation resulted in probes which did not consistently label to high specificity. For this reason, and to avoid radiolabelling non-*lacZ* pPD16.43 plasmid sequences which might cross hybridize with homologous sequences of the selection plasmid (e.g. such as the ampicillin gene sequences), M13 probes containing *lacZ* were generated by primer extension and routinely used to analyse Southern blots (Figure 10). The presence of numerous hybridizing fragments of various sizes which are smaller than the 1.9 kb *lacZ* band expected by digestion with EcoRI and EcoRV, coupled with incomplete transmission of the transformed phenotype suggests that the injected DNA formed complex extrachromosomal arrays as described by Stinchcomb et al (1985).

While equivalent amounts of DNA were loaded into each slot, the copy number of the *lacZ* transgene varied widely between strains making copy number estimates difficult. In addition many DNA samples did not appear to be completely digested, resulting in intense, but non-specific smearing of signals. Apparently, while the formamide DNA isolation procedure employed is rapid and produces high molecular weight DNA, inconsistent restriction suggests

Figure 9. Southern analysis of SmaI/ApaI digested genomic transgenic strain DNAs probed with nick translated pPD16.43. The solid arrow indicates the *lacZ* containing fragment; the clear arrow, the fragment containing other vector sequences. The latter fragment is larger in transgenic strains due to the included *hsp16* sequences. Lanes: M, Lambda DNA digested with HindIII; 1, N2 control; 2, PC1A, a 48P/pPD10.41 strain; 3, PC3A, a 48P/pPD10.41 strain; 4, PC9A, a 48P/pPD10.41 strain; 5, PC13A, a 48P/pPD10.41 strain; 6, N2 control; 7, 60 pg of pPD16.43; 8, 600 pg of pPD16.43; M, Molecular weight marker. These strains did not express the 48P transgene due to lack of a polyadenylation signal in the 3' end of the original plasmid vector. Hence this Southern was originally done to check for the presence of the *lacZ* transgene. These strains were discarded and not included in the subsequent study.

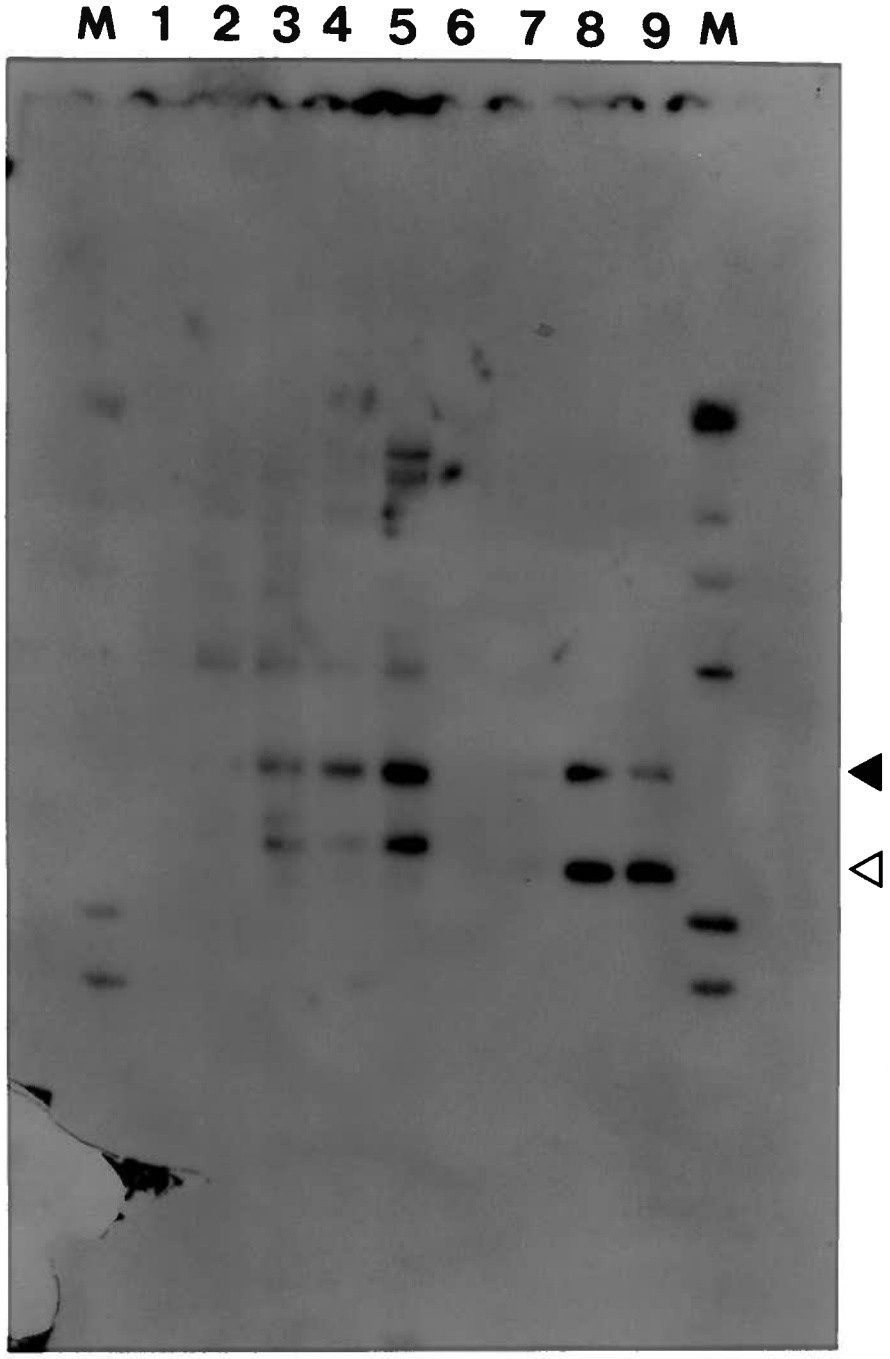
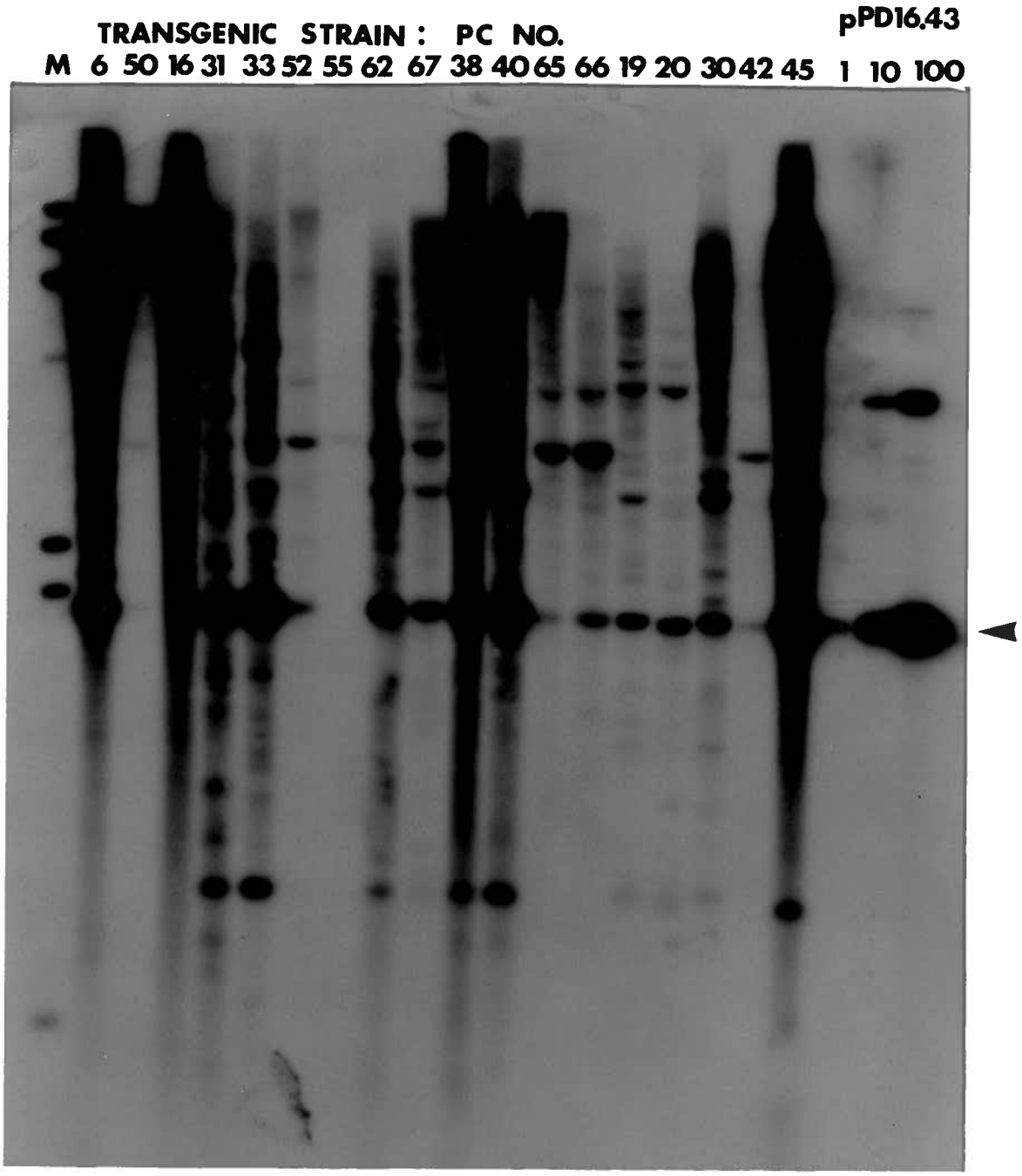

M 1 2 3 4 5 6 7 8 9 M

Figure 10. Southern analysis of transgenic genomic DNA probed with primer extended *lacZ* products. Genomic DNAs (3 μg) were digested with EcoRI and EcoRV, separated by electrophoresis on a 1 % agarose gel and tranferred to a nylon membrane by the method of Southern (1975). The arrow indicates the 1.9 kb band expected by liberation of the *lacZ* fragment by EcoRI/ EcoRV digestion. Lanes: M, Lambda/ HindIII molecular weight marker; transgenic strains as numbered; 1, 10, 100 copies of *lacZ* contained in pPD16.43.


TRANSGENIC STRAIN : PC NO.
pPD16.43
M 6 50 16 31 33 52 55 62 67 38 40 65 66 19 20 30 42 45 1 10 100

that protein contamination is an inherent problem with this procedure. To circumvent these problems, dot blots were employed to provide more accurate estimates of transgene copy number.

8. Determination of lacZ transgene copy number

By DNA dot blot analysis, I have estimated the copy number of the *lacZ* transgene in these arrays to be in the range of 5-750 copies per genome equivalent (Fig. 11; Table 3). The number of animals carrying the extrachromosomal array was inferred by selection of the marker phenotype prior to harvesting. Since in some strains one third of the phenotypically wild type animals may actually possess the array, the copy number results are probably overestimates.

In practice, because these strains transmit the selectable phenotype with less than 100 % fidelity, it is impossible to estimate exactly the percentage of cells within the animal which possess the transgene. Basically in this analysis, the likelihood of losing the array was considered to be the same in all somatic lineages as in the germ line. This assumption may not be true and certain cell lineages may preferentially retain the transgene while others may not.

The number of copies did not seem to depend on the construct injected or even on the quality of each individual DNA preparation since all lines for one construct were produced from injections with the same DNA preparation (Table 3).

To determine if copy number correlated with the intensity of expression between strains carrying the same construct, the induction of the transgenes over time was compared with the copy number (Table 4). Generally, copy number did not consistently correlate with the first appearance or extent of *lacZ* expression. Even after relatively short heat shocks (30'), PC31, with only 14 copies per genome equivalent, stained as frequently in intestine and body muscle as PC16 which contained 50 times the number of copies per genome equivalent. In addition this study indicated that the selection employed had no significant effect on the pattern of expression. Both the twitcher and the rolling strains stained equally well in body muscle.

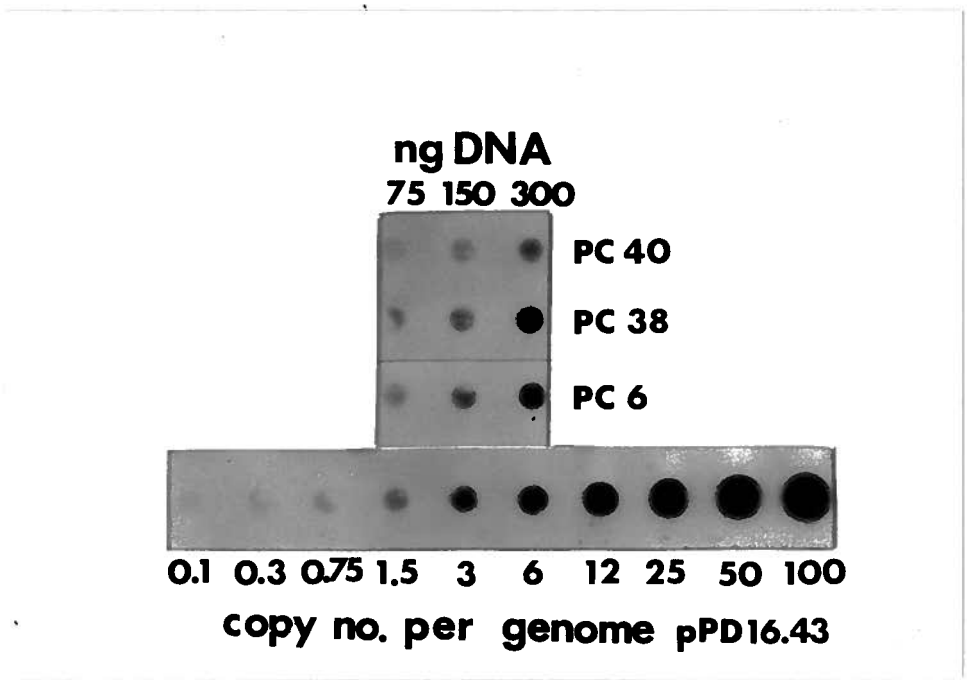


Figure 11. Dot Blot analysis of transgene copy number. Serial dilutions of genomic DNAs from transgenic strains were probed with radiolabelled primer extended *lacZ*. The copy number was inferred from the linear region of a standard curve obtained using plasmid pPD16.43 which contains one complete *lacZ* gene per genome. Using the estimated value of 100 Mb for the size of the haploid genome of *C. elegans*, I estimated that one copy of the *lacZ* gene was equivalent to approximately 1/33,000 of the haploid genome. Thus on the basis of loading 3 μg of genomic DNA the relative amount of pPD16.43 DNA required to represent a single copy gene was estimated to be 90 pg. For the DNA dot blots, 3 μg of genomic DNA could not be effectively analyzed because of the very high copy number in some strains. Samples were diluted accordingly and the results then extrapolated to represent 3 μg.

Table 3. *Determination of lacZ fusion gene copy number in transgenic strains selected for extensive analysis.* TW, pPD10.41 containing strain; ROL, pRF4 containing strain.

lacZ FUSION	STRAIN	# COPIES	% TW or ROL	COPY #/ GENOME
48.1C	PC6	41	55 %	75
	PC50	2	20 %	10
41.2C	PC65	12	30 %	36
	PC66	10	45 %	22
1.48E1	PC16	150	20 %	750
	PC31	4	30 %	14
	PC33	15	45 %	33
48.1E1	PC52	2	20 %	10
	PC55	10	20 %	50
2.41E1	PC19	6	50 %	12
	PC20	2	30 %	7
	PC30	13	35 %	38
48.1XBE1	PC62	36	30 %	120
	PC67	7	55 %	14
41P	PC42	1	20 %	5
	PC45	41	50 %	82
48P	PC38	26	10 %	260
	PC40	10	23 %	24

Table 4. *Comparison of expression patterns between 1.48E1 strains.*

TRT.	STRAIN	COPY #	SELECTION	IN	BM	HYP	PHA
120'HS	PC31	14	pRF4	8	10	5	2
30'REC	PC33	33	pRF4	9	10	8	6
120'HS	PC16	750	pPD10.41	7	9	10	6
120'HS	PC31			10	10	6	10
	PC33			9	9	8	5
	PC16			7	9	9	9
90'HS	PC31			9	8	5	7
	PC33			9	10	10	10
	PC16			6	9	7	8
60'HS	PC31			7	6	3	1
	PC33			8	10	4	5
	PC16			7	9	8	8
30'HS	PC31			9	9	4	4
	PC33			4	10	5	1
	PC16			7	8	8	5
15'HS	PC31			10	4	0	0
	PC33			9	6	5	3
	PC16			1	10	8	4

For each datum point ten staining animals were scored for expression in the specified tissue. HS, Heat shock; REC, Recovery; IN, Intestine; BM, Body muscle; HYP, Hypodermis; PHA, Pharynx.

Hypodermal expression was somewhat reduced in one of the rolling strains, PC31, but not in the other, PC33.

D: DISCUSSION

1. The lacZ transgenes are correctly expressed in response to heat shock in vivo

The in situ localization studies described here, as well as previous studies of hsp16 mRNA levels (Russnak *et al.*,1985; Jones *et al.*, 1986) demonstrate that these genes are strictly heat inducible and are not expressed constitutively at any time in development. All of the *hsp16-lacZ* transgenes in this study were expressed only in response to a heat shock. Neither the copy number (as high as 750 copies per cell in some strains) nor the complex nature of the extrachromosomal arrays seemed to interfere with the tight heat inducibility of these promoters.

2. Transgene copy number versus expression levels

Examination of multiple strains carrying each construct showed no consistent correlation between gene copy number and levels of expression. The primer extended *lacZ* products used to probe DNA should hybridize to any *lacZ* sequence of reasonable homology and length; whether it be a complete contiguous coding region or not, the probe cannot distinguish between these events. In light of the fact that different sized fragments of the selection and test plasmid sequences are presumably mixed in head to head and head to tail conformations, the actual number of complete transcribable copies of the *lacZ* transgene is probably much lower than dot blot analysis suggests. On the other hand strains which contain high copy numbers (e.g. PC16 with 750 copies) substantially increase the number of HSEs per genome. Thus, even if all copies of the transgene are not complete, a good deal of HSF could be sequestered at these HSEs during stress. It is conceivable that levels of HSF could be limiting and that strains with 750 transgene copies (or 1500 additional HSEs) may have surpassed that limit such that many HSEs remain unbound by HSF during stress.

In addition the presence of mixed arrays heightens the possibility of positional effects due to sequences derived from the selection plasmid or vector sequences. There is evidence with other genes which have been co-injected with the *unc-22* plasmid pPD10.41, that

enhancer sequences can promote spurious expression in muscle cells (Andrew Fire, personal communication). Since muscle expression was strong in strains transformed with the roller construct, pRF4, and since variability in muscle expression was between strains transformed with different fusions rather than between strains transformed with identical fusions, I believe that any such effects are minimal or non-existent in this analysis. Other studies have suggested that the roller plasmid can have negative effects on expression. This effect is particularly enhanced in late larval stages and in gene fusions which contain short promoter sequences (Andrew Fire, personal communication). Thus there is a possibility that the limited expression observed in the transcriptional fusions may in part reflect negative interactions.

3. Expression of the hsp16 gene pairs is tissue general

The staining patterns observed for the whole locus *lacZ* fusions (48.1C, 41.2C) very likely reflect the *in vivo* situation since all *hsp16* coding and flanking sequences are included in these constructs. In spite of the quantitative differences in expression between loci, as illustrated by the exon 1 fusions, most cells transcribe mRNAs from all four *hsp16* genes. Moreover, all four genes first become heat inducible in gastrulating embryos, and are inducible at all subsequent stages of *C. elegans* development. Only the germ line and the early embryo fail to express the *hsp16* genes following a heat shock. Thus the HSP16 proteins likely provide a function which is required by all cells subjected to heat shock. Like their *Drosophila* counterparts and the α-crystallins, the HSP16s of *C. elegans* form large aggregates *in vivo* (Hockertz *et al.*,1991). Speculation concerning the functions of these proteins has included roles in protecting mRNA (Nover *et al.*, 1989), the cytoskeleton (Leicht *et al.*, 1986; Vierling *et al.*, 1988) or organelles (Cooper and Ho 1987) from the effects of heat shock; however, no definitive role for the small HSPs has yet been established.

Curiously, the nuclear localization signal of 41.2C failed to direct the fusion protein to the nucleus, although this occurred efficiently in 48.1C transformants. This implies that the two fusion proteins fold differently *in vivo* such that the NLS was exposed in the 48.1C gene

product but not in the 41.2C gene product. This was surprising since the restriction site used to insert *lacZ* into *hsp16-1* and *hsp16-2* was a *Hpa* I site present in the identical position in both genes. In addition the integrity of the NLS in 41.2C was verified by restriction digestion and sequence analysis. The HSP16-1 and HSP16-2 proteins are 90% identical throughout their sequences (Jones *et al.*, 1986) and the nearest amino changes relative to the lacZ protein insertion are 9 amino acid residues away on the N-terminal side, and 26 on the C-terminal side. It is possible that these few amino acid differences can significantly alter the local folding and consequently perhaps the biological properties of these HSP16s. Alternatively, the N-terminal region of the 41.2C fusion may be unstable due to proteolysis, or some other unidentified signal within the *hsp16-2* coding region may interfere with the function of the NLS. Possibly relevant to this question is the fact that construction of the 48.1C fusion resulted in insertion of an extra 8 codons upstream relative to 41.2C. This may have served as a spacer in 48.1C, allowing better function of the NLS.

4. Tissue differences in the expression of hsp16-lacZ transgenes

The quantitative differences in expression observed between the *hsp16-48/1* and *hsp16-41/2* gene pairs, as illustrated by the exon1-lacZ gene fusions, may also reflect different cellular requirements if in fact different HSPs possess slightly different functions. The intergenic regions of these pairs are 86% identical. It is not yet clear what sequences account for the differences in tissue expression. These differences could result from either transcriptional or post-transcriptional effects; the latter might be mediated, for instance, through differences in the mRNA leaders. For example, sequences in the leaders of the *Drosophila hsp70* and *hsp22* mRNAs have been shown to be required for selective translation of these genes under heat shock conditions (McGarry and Lindquist, 1985; Hultmark *et al.*, 1986;). Previous evidence suggests that the *hsp16* pairs are differentially regulated (Jones *et al.*, 1989; Dixon *et al.*, 1990). Up to 14-fold more RNA transcripts are present per gene from the *hsp16-41/2* pair relative to the *hsp16-48/1* pair after a two hour heat shock. Moreover, it

was found that while *hsp16-2* mRNA was more abundant than that of *hsp16-1* at all points in the life cycle, the difference was most pronounced for embryo and L1 stages (Jones *et al.*, 1989). No differences in the magnitude of expression between the 48.1C and the 41.2C transgenes were observed in this study. However, since staining intensity was assessed visually and since the SV40 NLS was non-functional in the 41.2C transgenic animals, differences could easily have gone undetected.

The finding of intra-pair differences in expression of the *hsp16-48/1* transgenes implies the existence of a directionality to the intergenic region. It remains undetermined which sequences are involved in this phenomenon. Transgenic strains carrying 48.1XBE1, a construct which inverts the region of the 48.1E1 transgene between the *hsp16-1* and *hsp16-48* HSEs failed to significantly alter the pattern of expression toward that of 1.48E1 transgenic animals. The leader sequences, however, remained unchanged in this experiment and could account for this result. Interestingly, neither 48.1E1 or 1.48E1 transgenic worms stained as well in neural tissue as the whole locus fusion, 48.1C transformants. This may be the result of differences in the stability of the RNA produced from these constructs. In particular, the whole locus fusions have retained the endogenous *hsp-16* 3' sequences whereas the exon 1 fusions utilized a 3' end from a body wall myosin gene, *unc-54* (Epstein *et al.*,1974; Fire *et al.*, 1990). It is possible that the 3' ends of the *hsp16* genes may be important for mRNA stability during heat shock.

5. Cooperative interaction of HSEs may enhance expression

Elimination of the HSEs and TATA box for one gene of a pair in the 48P and 41P constructs significantly reduced the production of β-galactosidase in somatic tissues, such that double heat shocks were required to assess somatic expression, especially in 48P transgenic animals. The 48P promoter was approximately 70 bp shorter than the 41P promoter and lacked the transcriptional start site of the endogenous gene, which may account for the weaker and more restricted expression observed in strains carrying the former construct.

Nonetheless, production of β-galactosidase in somatic tissues in the fusion 41P was significantly reduced in comparison to the fusion 2.41E1. Essentially two differences exist in the construction of these fusions. The *hsp16-41* breakpoint proximal to *lacZ* in the 41P fusion is two bp upstream of the *hsp16-41* initial methionine; thus it is a transcriptional fusion. The distal breakpoint is located within the *hsp16-2* HSE such that the element is destroyed in the 41P construction. Thus while 41P contains all of the sequences between the HSEs as well as the transcriptional start site, HSE and TATA box of the *hsp16-41* promoter, the HSE and TATA box of the *hsp16-2* promoter have been eliminated. The 2.41E1 fusion on the other hand was a translational fusion of the first 15 amino acids of hsp16-41 to *lacZ* and possessed the HSEs and TATA boxes of both genes. The reduction in expression of the 41P fusion therefore may be the result of losing the translational start site of the *hsp16-41* gene; a less likely possibility is that the closer proximity of vector sequences in 41P may inhibit expression. I believe that the most likely explanation for the observed difference in expression is the existence of cooperativity between the two *hsp16* promoters in the wild type gene pairs. This is in agreement with the results obtained when the *hsp16* genes were introduced into mouse cells (Kay *et al.*, 1986). Recent evidence suggests that *in vivo* , HSTF functions as a trimer or perhaps even a hexamer (Sorger and Nelson 1989; Clos *et al.*, 1990) and binds to HSEs at repeating nGAAn motifs, which may be in either orientation (Amin *et al.*, 1988; Xiao and Lis 1988; Perisic *et al.*, 1989). The HSEs of each *hsp16* gene at each locus possess three such sequences. Thus each gene promoter is capable of binding an HSTF trimer.

It has been suggested that HSTF trimers binding at separate HSEs can interact with each other cooperatively to enhance transcription (Shuey and Parker 1986; Xiao *et al.*, 1991). The HSEs at each *hsp16* locus are only 130 bp apart, suggesting that such a cooperative interaction may be possible. It has been shown that HSEs positioned as much as 2 kb apart can activate transcription cooperatively (Riddihough and Pelham 1986), and Thomas and Elgin (1988) have shown that HSEs separated by more than 200 base pairs in the *Drosophila hsp26* gene may be brought together via folding around a nucleosome core positioned over the intervening region.

Thus a divergent arrangement of closely spaced *hsp16* genes may lead to more efficient heat inducible expression for a given number of heat shock elements.

While temperatures of 32° to 33°C are required for consistent somatic tissue expression, 29°C is sufficient over the same time period for embryos. Furthermore, transcriptional fusions were expressed extremely well in embryos while expression in the somatic tissues of later stages was minimal. These results suggest that the minimal requirements for high level expression in embryogenesis are maintained in these fusions while some element(s) necessary for somatic expression have been lost. Alternatively, it is possible that embryos have larger levels of available heat shock transcription factor (HSTF) relative to the somatic tissues of later stages. If so, then somatic gene expression may rely more heavily on the cooperative binding of HSTF to the HSEs in the intergenic region to achieve a given level of expression.

6. Developmental Regulation of the hsp16s in C elegans

Although the hsp16 genes are apparently never expressed constitutively, several lines of evidence indicate that their inducibility is regulated at certain stages in development. Firstly, no expression has been observed in the germ line of any of the transgenic strains studied to date. Secondly, embryonic expression was not consistently observed until gastrulation. It has been suggested that in *C. elegans*, zygotic genes are first transcribed at gastrulation (Hecht *et al.*, 1981) and that this stage loosely corresponds to the mid-blastula transition in amphibians (Newport and Kirschner 1982). The results presented in this study suggest that a developmentally regulated program overides the heat inducibility of the *hsp16* genes: the maternal genes are not inducible during gametogenesis or early embryogenesis, while the zygotic genes become responsive to heat shock at approximately the time of onset of general zygotic transcription. With respect to gonadal expression however, it is notable that other transgenes (such as *msp-1* which encodes a sperm specific product), fail to be expressed in the nematode gonad. This suggests that there may be a general barrier to transgene expression in this tissue (D. Dixon, personal communication).

It is interesting to compare the regulation of the *hsp16* genes to that of the related gene family in *Drosophila* . The small hsps of *Drosophila*, *hsp22, 23, 26* and *hsp27* are also developmentally regulated but expression is induced by the hormone ecdysterone during pupation and oogenesis in the absence of heat shock (Ireland and Berger 1982). Ecdysterone induction is mediated through an element which is separate and distinct from the HSEs required for heat induction (Cohen and Meselson 1985; Hoffman and Corces 1986). To date four *hsp16* genes from *C. elegans* have been cloned and seem to be purely heat inducible. It is possible, however, that there may be related members of this family which are expressed temporally in the absence of heat shock. Recently, a gene encoding an 18 kD protein which is cadmium inducible has been cloned and found to share sequence similarity to the HSP16s (C. Rubin, personal communication). Exposures to cadmium which are sufficient to activate this gene do not consistently induce the *hsp16* transgenes used in this work, but in one experiment intense staining in the pharynx, intesine and vulva of some animals (approx. 1 %) was observed.

In *Xenopus* a temporal pattern of expression is superimposed upon the heat inducibility of some of the HSPs. For instance, *hsp30* first becomes heat inducible at the tadpole stage, and the levels of mRNAS of this gene vary considerably from tissue to tissue in adult frogs, being greatest in the kidney and gut (Bienz 1984). Similar phenomena were observed in this study, but the quantitative differences in spatial expression were seen among members of the same gene family.

7. Conclusions

1) The *hsp16-lacZ* transgenes were correctly regulated with respect to heat shock in transgenic nematodes. Expression was entirely heat dependent in aggreement with previous results obtained by Russnak et al. (1985) and Jones et al. (1986). Heat shock resulted in rapid and intense induction of β-galactosidase activity whereas no staining was observed in animals maintained at normal growth temperatures. These results demonstrate the validity of employing extrachromosomally inherited transgenes to study gene expression in *C. elegans.*

2) Inducibility of *hsp16* genes appears to be under developmental regulation since expression was not inducible in the germ line or early embryogenesis for any of the transgenes studied.

3) Expression of the *hsp16* gene pairs is temporally and spatially non-specific from gastrulation onwards.

4) Sequences conferring tissue specificity are contained within the intergenic region of the *hsp16* gene pairs since orientation of this region affects the priority of tissue expression.

5) The experiments described in this study represent the first fully inducible system to be characterized molecularly in *C. elegans* and offers unique opportunities to the *C. elegans* investigator who wishes to employ a reverse genetic approach.

8. Future Prospects

This study pursued a reverse genetic approach to tackle analysis of the *hsp16s* as they are expressed in *C. elegans*. In light of the results obtained, a number of interesting avenues could be followed to elucidate further the role of the small HSPs in the stress response. An antisense strategy could be employed to approximate loss of function alleles in order to determine if the HSP16s are essential during stress or for specific components of the stress reponse. Such experiments may provide extra information regarding the identity of the HSP18s.

In addition, the differential expression revealed by the exon 1 fusions suggests that multiple enhancing elements may be present within the intergenic region. Thus, further dissection of this region would be desirable to identify these sequences. Ultimately, tissue specific enhancers may be identified which could be employed to direct expression of any gene of interest to specific tissues.

The strict heat inducibility of the *hsp16* genes constitutes a potentially powerful tool which could be used to address a variety of interesting biological questions in *C. elegans*. Using these promoters, it should be possible to achieve the heat inducible expression of almost any desired coding region. Since the *hsp16* promoters can be activated to produce a high level of

expression within minutes, it is feasible to examine the effects of a particular gene product when produced at short, specific times in development. For example, an *hsp16-mab5* fusion has been used to achieve tightly controlled, heat-dependent movement of specific migratory neuroblasts at the L1 larval stage in a *mab-5* null background (S. Salser and C. Kenyon, personal communication).

By combining tissue-specific enhancers with heat shock promoters, it should be possible to achieve heat inducible expression in specific target tissues, as suggested by Pelham (1987). Such a hybrid promoter, fused to a gene which produces a toxic product such as diphtheria toxin, might allow specific killing of selected cells under controlled conditions. In another approach, it may be possible to use a laser (White and Horvitz, 1979) to heat shock individual cells within an animal containing an *hsp16* gene fusion of interest. The transparency of *C. elegans* and its defined cell lineage (Sulston *et al.*, 1983) make it particularly well suited to such an application. Such a technique, if feasible, would provide unique opportunities to assess the functions of particular gene products in chosen cells at a given point in time, and thereby to study the roles of such products in cellular interactions during determination and development.

The induction of HSPS by a variety of biologically harmful agents including heat shock has prompted speculation that these proteins could be used to monitor environmental stress. One of the problems with this approach is that many HSPs are produced at low levels constitutively or at specific times in development in the absence of stress, thus assays would have to detect increases above background levels. The *hsp16* genes described in this study are only expressed in response to heat shock but future experiments will include determining if these genes are induced in response to chronic stress by a variety of agents. If these results are affirmative then the *hsp16-lacZ* transgenes described could be powerful tools in the assessment of environmental stress. Firstly, *C. elegans* is a natural inhabitant of soil; secondly, the histochemical staining procedure is a simple assay for expression which requires minimal technical dexterity and has a clear positive or negative result. Moreover, the assay is

sensitive, and HSP induction does not result in lethality. Finally, in conjunction with developmental time curves, these animals could provide invaluable information regarding the effects of chronic non-lethal stress on development and ageing. On the practical side, however, these experiments will necessitate the construction of homozygous integrated lines which can be easily maintained in the absence of selection.

Recently, using an X-ray mutagenesis procedure suggested by Jeff Way (personal communication), I generated stably integrated homozygous rolling lines from a 48.1C extrachromosomally transmitted strain, PC6. Initial analysis of one of these strains, PC71, substantiates the findings presented in this thesis. The *lacZ* transgene is expressed only in response to heat shock and in the tissue general manner characteristic of its ancestor, PC6 (48.1C/pRF4). Rarely, a few (< 5 %) comma to pretzel staged embryos also stain, perhaps due to positional effects. Future endeavours will include mapping experiments to determine the site of array integration, the copy number and integrity of the transgene. Since the X-ray mutagenesis procedure was technically simple and effective it should be relatively trivial to generate homozygous lines for all constructs of interest, bypassing the tediousness of injections. This would make biochemical experiments a more accessible prospect, as well as enhance the potential for using these animals as probes of environmental stress.

II:ANALYSIS OF POLYUBIQUITIN GENE EXPRESSION

A: INTRODUCTION

1. General Introduction

Ubiquitin is a highly conserved 76 residue polypeptide which is present in all eukaryotic cells (Goldstein et al., 1975). Determination of the structure of ubiquitin at 0.28 nm and later at 0.18 nm resolution revealed that it is a highly compact globular protein with an amino terminus buried within a hydrophobic core and a protruding carboxy terminal tail (Vijay-Kumar et al., 1985; Vijay-Kumar et al., 1987a; Vijay-Kumar et al., 1987b). Ubiquitin is highly resistant to thermal and chemical denaturation (Ciechanover et al., 1978; Lenkinsky et al., 1977).

Numerous cellular processes are mediated by the post translational conjugation of ubiquitin to proteins via an isopeptide bond between the carboxyl terminal glycine of ubiquitin and the epsilon amino group of an internal lysine residue in the target protein (see Hershko and Ciechanover, 1986; and Hershko, 1988 for reviews). In this fashion, it is believed that ubiquitin "marks" a protein for a particular fate: whether it be for degradation by the non-lysosomal ATP dependent pathway, or for some other protein interaction.

2. Ubiquitin mediated proteolysis

In the cytoplasm, ubiquitin is involved in the ATP dependent proteolysis of damaged or defective proteins (reviewed in: Hershko and Ciechanover, 1986; Hershko, 1988). The events involved in this procedure are summarized briefly in Figure 12 and in the text below. First, ubiquitin is activated in a two-step reaction by a specific enzyme, E1 (Step 1 of Figure 12). Ubiquitin adenylate is formed with the displacement of pyrophosphate from ATP. Subsequently, activated ubiquitin is transferred to a thiol site on the E1 enzyme and AMP is released.

In step 2, ubiquitin is transacylated from E1 to a specific cysteine residue on a carrier enzyme termed E2. Many different E2 enzymes have been identified, some of which can catalyze

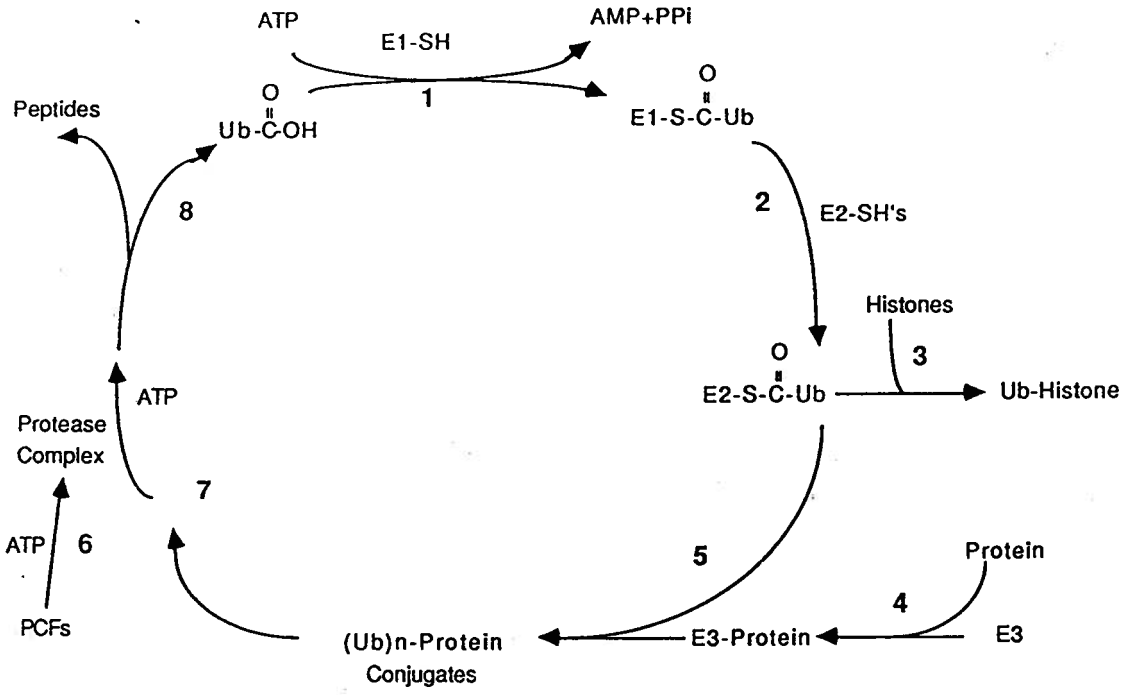


Fig. 12. The non-lysosomal ATP dependent proteolytic pathway. Adapted from Hershko (1988). See text for details. Ub, ubiquitin; E1, ubiquitin activating enzyme; E2, ubiquitin conjugating enzyme; E3, ubiquitin ligase; PCFs, protease complex factors.

the assembly of stable ubiquitin-protein conjugates which are not destined for degradation (Step 3 and Section 5).

Proteins with structures suitable for degradation are selected by and complexed to ubiquitin ligase or E3 enzymes (Step 4). One of these E3 enzymes, that encoded by the *UBR1* gene of yeast, has been shown to specifically recognize and bind only to proteins which have de-stabilizing amino terminal residues (Bartel et al., 1990) as defined by the N-end rule (Bachmair et al., 1986).

In step 5, ubiquitin is transferred from the E2 complex and linked by its COOH glycine 76 residue to ε-NH_2 groups of internal lysine residues in the E3 bound target protein. In this way, numerous monomeric ubiquitin molecules can be bound to several lysine residues in a protein, and also form branched polyubiquitin chains stemming from a single lysine. The latter formation has been implicated as a pre-requisite for degradation (Chau et al., 1989).

Protein-ubiquitin conjugates are degraded to small peptides by a large (> 100kDa) protease complex (Steps 6 and 7) the formation and action of which is dependent on ATP. Finally, in the last step of the model, free ubiquitin is recycled for use in subsequent degradations (Hershko, 1988).

3. Ubiquitin as a regulatory protein

Recently it has been demonstrated that a number of short-lived proteins are also degraded by the ubiquitin-dependent system. This includes the nuclear oncoproteins encoded by *N-myc*, *c-myc*, *c-fos*, and *E1A* (Ciechanover et al., 1991), and the yeast *MATα2* repressor (Hochstrasser et al., 1991). Degradation of cyclin by the ubiquitin pathway is the key event signalling exit from mitosis (Glotzer et al., 1991). In plants, degradation of phytochrome is accompanied by the formation of phytochrome-ubiquitin conjugates and is activated by light (Shanklin et al., 1987). Thus by determining the half-lives of regulatory proteins, the ubiquitin system may be an important regulator of cell cycle events.

4. Ubiquitin and heat shock

A common result of physiological stress such as heat shock is the accumulation of abnormal or damaged proteins. This presumably would place pressure on the ability of systems such as the ATP dependent proteolytic pathway to degrade and process the overload.

Increased ubiquitinylation of proteins has been shown to accompany heat shock (Parag et al., 1987). A temperature sensitive mouse cell line, ts85, carries a mutation in an E1 activating enzyme (Finley et al., 1984) which severely impedes the formation of ubiquitin-protein conjugates under non-permissive conditions. As a result, these cells are unable to degrade efficiently short lived or abnormal proteins at restrictive temperatures (Ciechanover et al., 1984). In addition, ts85 cells produce massive quantities of heat shock proteins at restrictive temperatures compared to wild type cells (Finley et al., 1984; Ciechanover et al., 1984). The yeast genes *UBC4* and *UBC5* encode ubiquitin conjugating enzymes (E2s) which specifically conjugate ubiquitin to short lived or abnormal proteins destined for elimination (Seufert and Jentsch, 1990). These genes are also transcriptionally heat inducible. *ubc4ubc5* double mutants accumulate free monomeric ubiquitin, fail to turn over proteins efficiently, constitutively express heat shock proteins at normal growth temperatures and are inviable at elevated temperatures or in the presence of the amino acid analogue canavanine (Seufert and Jentsch, 1990). This provides further evidence that the heat shock response is induced by the presence of excessive quantities of abnormal proteins. In retrospect then, it is not surprising that ubiquitin itself is a heat shock protein in yeast, *Drosophila*, and chicken (Bond and Schlesinger, 1985; Finley et al., 1987; Lee et al., 1988). (See section 8. for more detail.)

5. Ubiquitin and chromatin

Within the nucleus, histones are modified by ubiquitin conjugation in a reversible process which does not trigger proteolytic degradation (Busch , 1984). It has been estimated that approximately 10% of histone H2A (West and Bonner, 1980a) and 1.5% of H2B is

ubiquitinylated (West and Bonner, 1980b). Conjugation to histones H2A and H2B seems to be a characterisitic of active chromatin (Levinger and Varshavsky, 1982; Nickel et al., 1989; Davie and Murphy, 1990).

The *RAD-6* gene (*UBC2*) of yeast encodes a ubiquitin conjugating enzyme (E2) which transfers ubiquitin to histones in an E3 independent reaction *in vitro* (Jentsch et al., 1987) and this may be related to its central role in DNA repair *in vivo* (Sung et al., 1988; Sung et al., 1990). Mutants of *UBC2* are defective in DNA repair, mutagenesis and sporulation. Expression of *UBC2* is induced by ultraviolet light, and is required throughout the cell cycle but peaks in meiosis during the period of maximal genetic recombination. Unlike *UBC4* and *UBC5* , expression of *UBC2* is not induced by heat shock or starvation suggesting that this E2 is not substantially involved in the conjugation of ubiquitin to damaged proteins during stress (Madura et al., 1990). However, UBC2 has been shown to be essential for multi-ubiquitinylation and degradation of N-end rule substrates which have been recognized by the E3 ligase, UBR1 (Dohmen et al. 1991; Sung et al. 1991). Thus, this E2 catalyzes the formation of both stable ubiquitin-histone conjugates and unstable ubiquitin-protein conjugates destined for degradation.

Conjugation of E2 enzymes to ubiquitin requires formation of a thioester adduct. Site directed mutagenesis of the sole cysteine residue in *UBC2* totally abolishes ubiquitin-conjugating activity and the biological function of *UBC2* (Sung et al., 1990). This provides direct evidence that ubiquitinylation of histones is required constitutively for DNA repair as well as transcription.

Another ubiquitin conjugating enzyme, *UBC3* of yeast was originally isolated as a cell cycle mutation, *CDC34*. The protein product of *UBC3* is required for the transition from the G1 to the S phase in the cell cycle (Pringle and Hartwell, 1981) and catalyzes the conjugation of ubiquitin to H2A and H2B *in vitro* (Goebl et al. 1988). Thus ubiquitin-conjugation to histones in chromatin may be an important regulator in a large number of nuclear events.

6. Ubiquitin at the cell surface

Various membrane proteins such as the lymphocyte homing receptor (Siegelman et al., 1986), the growth hormone receptor (Leung et al., 1987), and platelet-derived growth factor receptor (Yarden et al., 1986) are conjugated to ubiquitin *in vivo*. Characterization of each of these receptors revealed two amino termini, one corresponding to the receptor polypeptide and the other to ubiquitin. While the function of this arrangement remains unresolved, it has been suggested that ubiquitin may act as a "tag" in cell guidance interactions during receptor maturation (Siegelman et al., 1986). Alternatively, ubiquitin may act in its conventional role by mediating receptor turnover since the growth hormone receptor has been shown to be short lived *in vivo* (Baxter et al., 1985; Gorin et al., 1985).

7. Ubiquitin and myofibril assembly

In *Drosophila*, the indirect flight muscles contain a myofibrillar protein called arthrin, which is a stable actin-ubiquitin conjugate (Ball et al., 1987). Ball et al. followed the incorporation of labelled [^{35}S] methionine into actin and arthrin during muscle formation and showed that arthrin synthesis lagged behind that of actin and that one arthrin molecule was produced *per* thin filament of muscle (1987). The authors suggested that arthrin may target particular actin monomers within filaments as appropriate or inappropriate binding sites for myosin heads during the cross-bridge cycle of myofibril assembly (1987).

8. UBIQUITIN GENE STRUCTURE

8.1 Polyubiquitin gene structure

The polyubiquitin gene consists of tandem repeats of a ubiquitin coding region which is transcribed as a polycistronic mRNA, translated into a polyprotein and cleaved post translationally to release ubiquitin monomers (Lund et al., 1985; Ozkaynak et al., 1984). No spacers or introns separate the ubiquitin repeats; the carboxyl-terminus of one ubiquitin

moiety is linked to the amino-terminus of the next in the polyprotein while the last ubiquitin repeat carries a one to three amino acid extension (Table 5).

The number of polyubiquitin genes per genome and the number of ubiquitin coding repeats per polyubiquitin gene varies considerably between organisms (Table 5). The yeast polyubiquitin gene, *UBI4*, contains five ubiquitin coding regions in a head to tail arrangement (Ozkaynak et al., 1984); *Drosophila* polyubiquitin contains 18 repeats (Lee et al., 1988), and *Dictyostelium* contains two genes of three and five repeats respectively (Giorda and Ennis, 1987). In spite of this variation in ubiquitin repeat number, the amino acid sequence of ubiquitin monomer is largely invariant from species to species: for example, yeast and human ubiquitin differ by only three residues (Table 5). Thus, ubiquitin structure and function are largely intolerant of sequence change.

8.2 Polyubiquitin gene expression

In yeast the single polyubiquitin gene, *UBI4*, is essential for sporulation, resistance to starvation and heat shock (Finley et al., 1987), and G_0/G_1 arrest (Tanaka et al., 1988). Deletion mutants of *UBI4* are viable under normal conditions of exponential growth, and maintain wild type levels of free ubiquitin which is presumably derived from the ubiquitin-ribosomal fusions (Finley et al., 1987; Ozkaynak et al., 1987; See section 8.3). However, under conditions of physiological stress, the fusion proteins apparently cannot maintain sufficient quantities of ubiquitin monomer in *UBI4* mutants (Finley et al., 1987) and these cells are hypersensitive to heat and are sporulation defective. These defects were complemented by the introduction of a *UBI4* minigene which contained the *UBI4* 5' upstream sequences and one ubiquitin coding repeat. Thus, *UBI4* gene function is to provide ubiquitin monomers during stress as opposed to polyprotein (Finley et al., 1987). While *UBI4* represents the sole yeast polyubiquitin gene, two mRNA transcripts of 2.6 kb and 1.5 kb were observed on Northern blots (Finley et al., 1987) and the levels of the 1.5 kb transcript were elevated upon heat shock. Thus, the *UBI4* gene is induced by heat shock at the transcriptional

Table 5. Comparison of polyubiquitin genes among several species. The conservation of amino acid sequence amongst species is shown as well as the variation in gene number, repeat number and carboxyl terminus extension.

Organism	# residues divergent from human sequence	# polyubiquitin genes	# ubiquitin repeats/ gene	carboxy terminal extension	Reference
Human	-	2	3; 9	Cys; Val	Baker and Board, 1987; Wiborg et al., 1985
Chicken	0	2	3; 4	Tyr	Bond and Schlesinger, 1985
yeast	3	1	5	Asn	Ozkaynak et al., 1984
Dictyostelium	2	>=6	3, 5, ?...	Leu; Asn	Giorda and Ennis, 1987
Drosophila	0	1	1 8	Ile Gln Ala	Lee et al., 1988
C. elegans	1	1	1 1	Asp Ile	Graham et al., 1989
Tetrahymena	2	>=4	1?, 3?,...	Ser; Gln	Neves et al., 1988

level, and sequences matching the heat shock element consensus have been identified upstream of the *UBI4* coding region (Ozkaynak et al., 1987). Multiple RNA species transcribed from a single polyubiquitin gene have also been described in *Xenopus* (Dworkin-Rastl et al., 1984).

Neves et al. (1988) showed that *Tetrahymena* possess at least four polyubiquitin genes which are differentially expressed during heat shock. Three different RNA species (5.6 kb, 1.8 kb, and 0.75 kb) were observed on Northern blots from unstressed cells whereas heat shock increased the levels of the 5.6 and 1.8 kb transcripts and induced the appearance of a 1.6 kb transcript (Neves et al., 1988). Expression of the two polyubiquitin genes of chicken, *Ubi* and *UbII*, is also differentially regulated. Levels of the two polyubiquitin RNAs were determined in three tissues- mature testis, immature testis, and liver reticulocytes. While the *Ubi* transcript was the most abundant in all tissues, the proportion of transcripts derived from the *UbII* gene increased from 8% in immature testis to 22% in mature testis. In addition, an overall increase in ubiquitin levels was observed during spermatogenesis suggesting that the increased demand for ubiquitin during spermatogenesis is met by increased transcription of the *UbII* gene (Rocamora and Agell, 1990). While the *UbII* gene seems to be important for spermatogenesis, transcription of Ubi was induced by heat shock and chemical stress (Rocamora and Agell, 1990), and in fact *Ubi* was initially identified in a screen for heat shock genes (Bond and Schlesinger, 1985).

Drosophila contains one polyubiquitin gene of 18 tandem ubiquitin repeats (Lee et al., 1988). Expression of this gene is constitutive but appears to be moderately heat inducible. Lee et al. (1988) reported a three-fold increase in polyubiquitin expression during heat shock. Using P element mediated transformation of a polyubiquitin-lacZ fusion, they determined that the polyubiquitin gene was expressed constitutively in a developmentally non-specific and tissue general manner (Lee et al. 1988).

Based on the above data, a few generalizations with regard to polyubiquitin gene expression can be made. In organisms such as *Drosophila*, where there is only one polyubiquitin gene, expression of this gene seems to be temporally and spatially non-specific

but moderately inducible by physiological stress. Heat induction can, as in yeast, result in multiple RNA species being transcribed from the same gene. Organisms such as chicken, which contain more than one polyubiquitin gene, show differential expression amongst these genes with respect to development and physiological stress.

8.3 Ubiquitin Fusion genes

Ubiquitin fusion genes consist of a ubiquitin repeat fused to one of two basic "tail" sequences which encode small ribosomal proteins (Finley et al. 1989; Ozkaynak et al. 1987; Lee et al., 1988; Redman and Rechsteiner, 1989). Thus, unlike the ubiquitin-protein conjugates which are formed post translationally by the action of ubiquitin conjugating enzymes, these gene fusions are translated into a linear bi-protein molecule which is cleaved post-translationally to release monomeric ubiquitin and the ribosomal protein in a manner analogous to the processing of polyubiquitin. It has been proposed that fusion with ubiquitin may stabilise these ribosomal proteins and/or aid in ribosome assembly (Finley et al. 1989).

Such fusion genes have been found in a wide variety of eukaryotes including yeast, *Drosophila* and mammals (Finley et al., 1989; Redman and Rechsteiner, 1989;Lee et al., 1988), and exhibit a high degree of similarity in the tail sequences. For example, the ribosomal protein tail of the *Drosophila* UB3-D fusion is 65% identical to its yeast homologue and 82% identical to the human fusion. Moreover, Southern analysis in yeast indicates that these ribosomal proteins are encoded only by fusion genes (Otaka et al., 1984). Thus it is likely that these proteins provide a basic function required by all eukaryotes.

Three of the four ubiquitin genes of yeast, *UBI1*, *UBI2*, and *UBI3* encode fusion proteins (Finley et al., 1989). Deletion of any of these genes results in a slow growth phenotype whereas deletion of *UBI4* has no effect (Finley et al., 1989). Slow growth is a consequence of loss of the fusion tail proteins since levels of free ubiquitin are almost normal in these cells and since transformation of a plasmid carrying the tail sequence alone can complement the defects (Finley et al., 1989). The tail sequences of the *UBI1* and *UBI2* genes encode the identical

52 amino acid protein which has been identified as a constituent of the large (60S) ribosomal subunit. *UBI1/UBI2* double mutants are inviable: thus these genes are essential in at least a single copy (Finley et al., 1989). The *UBI3* tail protein consists of 76 residues and is a component of the small (40S) ribosomal subunit (Finley et al., 1989).

While ubiquitin fusion genes are apparently necessary for wild type growth, they are not induced by physiological stress as is often the case with their polyubiquitin counterpart (Ozkaynak et al., 1987; Finley et al., 1987; Finley et al., 1989).

8.4 Ubiquitin like genes

Recently, several genes have been identified which show sequence similarity to ubiquitin but which do not encode products which function as conventional ubiquitin does (i.e. by ligation to other proteins). An interferon-inducible 15 kDa protein discovered in mammalian cells contains two related domains, each with approximately 30% amino acid identity to human ubiquitin (Haas et al., 1987). The human *GdX* gene contains an amino terminus of 74 amino acids bearing 43% identity to human ubiquitin (Toniolo et al., 1988).

9.0 Trans-splicing

In trypanosomes, all pre-mRNAs receive a 39 nucleotide leader RNA molecule through a trans-splicing reaction (Laird, 1989). Trypanosome genes contain no introns and this organism does not possess U1 small nuclear RNA (snRNA), the component of the spliceosome which recognizes the 5' splice site in conventional cis-splicing. In trans-splicing the spliced leader (SL) RNA provides the 5' splice site for the reaction and exists *in vivo* as an RNP particle (Van Doren et al., 1988; Bruzik et al., 1988; Thomas et al., 1988).

For years trans-splicing was considered to be a phenomenon unique to trypanosomes, being an adaptation for efficient processing of the polycistronic messages common in these protozoans, and presumably incompatible with the cis-splicing of higher eukaryotes. The discovery by Krause and Hirsh (1987) that three of the four actin genes of *C. elegans*

transcribe pre-mRNAs which acquire a 22 nucleotide leader *via* trans-splicing shattered this conception. It is now apparent that approximately 15 % of *C. elegans* transcripts are trans-spliced (Blumenthal and Thomas, 1988), including *ubq-1* hnRNA (Graham et al. 1988). Moreover, while in trypanosomes, all RNA molecules acquire the identical 39 nucleotide spliced leader, in *C. elegans* two different spliced leader molecules (SL1 and SL2) have been identified (Huang and Hirsh, 1989) and each leader is spliced to specific nematode transcripts. In addition, in *C. elegans* trans-splicing and cis-splicing can occur not only within the same nucleus but within the same transcript (Krause and Hirsh, 1987; Graham et al., 1988). Indeed *C. elegans* possesses U1 snRNA while trypanosomes do not (Van Doren and Hirsh, 1988).

Many similarities exist between cis and trans-splicing. Both cis- and trans-splicing involve the formation of branched intermediates: a lariat structure in cis-splicing (Grabowski et al., 1984; Padgett et al., 1984; Ruskin et al., 1984), versus a Y-shaped forked molecule in trans-splicing (Sutton and Boothroyd, 1988). The 5' splice site sequence of SL1 RNA resembles typical 5' splice sites of *C. elegans* introns, while the 3' acceptor site for SL1 or SL2 RNA is virtually identical to a typical 3' splice site in cis-spliced introns (Blumenthal and Thomas, 1988). Recently, Conrad et al. (1991) demonstrated that inserting a normally cis-spliced intron, devoid only of its 5' splice site, upstream of the coding region of a *vit-2/ vit-6* fusion gene converted the non trans-spliced transgene into one which was efficiently trans-spliced to SL1. These results suggested that trans-splicing by SL1 is a default mechanism which occurs whenever a 3' splice site is present unaccompanied by a 5' splice site partner. What remains undetermined is what specifies a transcript for splicing by SL2.

10. The present study

In *Caenorhabditis elegans*, the polyubiquitin gene *ubq-1* codes for eleven tandem repeats of ubiquitin (Graham et al. 1989; Figure 13). The upstream region of *ubq-1* contains several features which might be involved in gene regulation, including a cytosine-rich block, an inverted repeat possessing homology to the DNA binding site of the mammalian steroid hormone

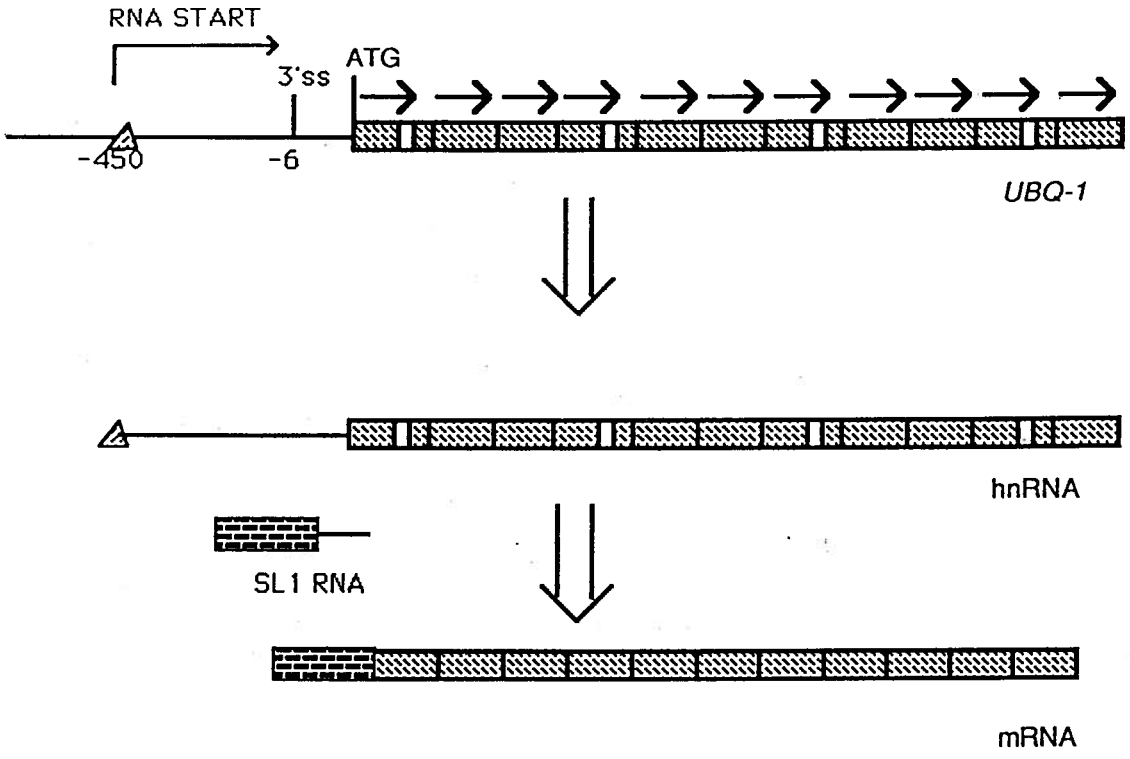


Fig. 13. Organization, transcription and processing of the polyubiquitin gene, *ubq-1*, of *Caenorhabditis elegans*. Adapted from Graham et al. (1989). *Ubq-1* encodes 11 repeats of ubiquitin. The transcriptional start site is indicated by the bent arrow. 3'ss, acceptor site for sliced leader RNA; SL1 RNA, splice leader 1 RNA.

receptor, and two sequences resembling heat shock elements. *Ubq-1* is unique among polyubiquitin genes isolated to date in that its heteronuclear RNA acquires a 22 nucleotide leader (SL1) by a trans-splicing reaction (Graham et al., 1988). Also unique is the presence of cis-spliced introns within the coding region (repeats 1, 4, 7, and 10) (Graham et al. 1989). Northern analysis shows that *ubq-1* is expressed constitutively at equivalent levels throughout development, and is not subject to regulation by heat shock or nutritional deprivation (Graham et al. 1989).

The purpose of this study was to examine the spatial expression patterns of *ubq-1* in transgenic nematodes carrying *ubq-1-lacZ* fusions. In addition, when this study was first conceived, no information regarding the specificity and mechanism of SL1 trans-splicing to pre-mRNAs in the nematode was available. Thus it was hoped that transgenic studies of *ubq-1* might provide further elucidation of this phenomenon.

B: METHODS

1. Maintenance of strains

Nematode strains were maintained as described in Chapter I methods.

2. Construction of ubq-1-lacZ fusions

Ubq938-lacZ contains 1044 bp of *ubq-1* sequences extending from a Sal I site 938 bp upstream of the initial methionine to an Eco RI site 103 bp downstream of the methionine cloned into the *lacZ* expression vector, pPD16.43 (Fire et al. 1990). Thus *Ubq938-lacZ* and all of the deletions are in-frame translational fusions containing the first 36 amino acids of the first exon of *ubq-1*. Deletions extending from the Sal I site of *ubq938-lacZ*, generated by exonuclease III digestion (Henikoff 1984) and analysed by sequencing (Sanger et al. 1977), were the gift of Don Jones.

A 1033bp Sal I to Hind III fragment of *ubq-1* 3' sequence was inserted between the Eag I and Apa I sites of pPD16.43 (Don Jones, personal communication). The resulting construct carried an additional 74 bp of polylinker derived from cloning vectors upstream of the start of the *ubq-1* 3' non-coding sequence and an additional 774 bp of *ubq-1* 3' sequence not previously published. The *ubq-1* polyadenylation signal is 110 bp downstream from the *ubq-1* stop codon and 154 bp downstream of the Sal I site.

3. Establishment of transgenic strains

Transgenic strains were established in methods described in Chapter I. For transient assays the *lacZ* fusion alone was injected at a concentration of 200 ng/µl.

4. Identification of β-galactosidase staining cells

As described in Chapter I.

5. Heat Shock Conditions

Transgenic lines were heat shocked for two hours at 33°C on NG plates seeded with OP50 and then allowed to recover at 20°C for 15'. Subsequently, worms were washed off the plates in distilled water, and stained for β-galactosidase activity as described previously.

6. Preparation of RNA

Usually four to eight plates loaded with worms was the starting material per RNA preparation. Mixed nematode populations, or embryos prepared by bleaching gravid adults (Emmons et al. 1979), were washed in cold 0.14 M NaCl, and then frozen by dripping the worm suspension into liquid nitrogen. Total RNA was isolated essentially as described by Antonucci (1985). Frozen worm pellets were powdered with a chilled mortar and pestle and then dissolved in 1-3 ml of guanidinium solution (7.5 M guanidinium chloride; 25 mM Na Citrate, pH 7.0; 0.1 M β-mercaptoethanol). The homogenate was passed 2-3 times through a syringe with a 21 gauge needle and then layered over 1 ml of DEPC-treated, sterile separation solution (5.7 M CsCl; 25 mM sodium citrate, pH 5.0). The RNA was pelleted by ultracentrifugation at 42,000 rpm (220,000 x g) for 16 hours. The pellet was then resuspended in 0.1 to 0.3 ml of sterile, DEPC-treated distilled water, precipitated by addition of 0.1 volumes of sterile, DEPC-treated 3M Na acetate (pH 5.2) and 2.5 volumes of 95 % ethanol. After centrifugation, the RNA pellet was finally resuspended in 50 µl of sterile DEPC-treated distilled water and quantitated spectrophotometrically by absorbance at 260 nm.

7. Preparation of cDNA

First strand cDNA was synthesized essentially as described by Graham (Ph.D. thesis, 1990) using a modification of procedures outlined by Maniatis et al.(1982) and Frohman et al. (1988). Ten µg of total RNA suspended in 7.5 µl of DEPC treated distilled water was combined with an equivalent volume of 40 mM methylmercuric hydroxide and incubated at 20°C for 15'

before flash freezing in a dry ice/ ethanol bath. Thirty µl of a solution containing 33 µM DTT, 1.66 X MMLV reverse transcriptase buffer (BRL), 1 unit/µl ribonuclease inhibitor (Promega or Pharmacia), 833 µM of each of dATP, dGTP, dCTP, and dTTP, and 2 µg of oligodeoxyribonucleotide RG05 or RACE (oligo-dT) was added to the frozen RNA pellet. Subsequently, 4.5 µl (900 units) of MMLV reverse transcriptase (BRL) was added to the mixture and the reaction was incubated at 37°C for 45'. Excess oligodeoxyribonucleotide was removed from the cDNA product by purification on spun filtration columns (Millipore Ultrafree-MC, 100K). Oligonucleotide RGO5 is 5' AGGGTTTTCCCAGTCACGAC 3' and is complementary to a sequence in the N-terminal coding region of *lacZ*.

8. Polymerase Chain Reactions

The cDNA product was amplified by a modification of the procedure of Frohman et al. (1988). Usually one tenth of the cDNA obtained after purification was mixed with 50 pmol of oligonucleotide OPC6, 100 pmol of oligonucleotide SL1, and 45 µl of a solution containing 77 µM of each of dATP, dGTP, dCTP, and dTTP, 10 mM Tris-HCl pH 8.4, 0.05 % Tween 20, 0.05 % Nonidet P-40, 0.5 mM $MgCl_2$, and 1 unit of Taq DNA polymerase (Promega or Pharmacia). The DNA was amplified on an ERICOMP thermocycler for 35 cycles consisting of 90 sec. denaturation at 94°C, 120 sec. annealing at 59°C, 90 sec. extension at 72°C. The amplified DNA products were separated by electrophoresis on a 2% agarose gel. Oligonucleotide SL1 is a 22mer of the sequence 5' GGTTTAATTACCCAAGTTTGAG 3', and oligonucleotide OPC6 consists of the sequence 5' GAGGATCCCGATCTCGCCATACAGCGCG3' and is upstream of RG05.

9. Separation and analysis of PCR products

The amplified DNA products were separated on a 2% agarose /1 X TBE (0.089 M Tris, 0.089 M boric acid, 0.002 M EDTA-pH 8.0) gel by electrophoresis at 100 volts in 1 X TBE buffer.

10. Southern Analysis of PCR Products

The PCR products separated on the agarose gel were transferred to a HYBOND-N nylon membrane according to the procedure of Southern (1975). DNA was immobilized on the membrane by UV crosslinking. MQIF oligonucleotide was labelled with γ^{32}PdATP in a reaction using polynucleotide kinase (Maniatis et al. 1982) and purified by centrifugation through a G-25 fine grade Sephadex spun column. The membrane was hybridized in 10 ml of a solution containing 0.5 % sodium dodecyl sulphate, 5 X Denhardt's, 5 X SSC for 10 hours at 50°C, before being washed with the final stringent wash consisting of 0.1 X SSC/0.1 % SDS at 50°C. Oligonucleotide MQIF is a degenerate 32-mer containing a BamHI site in addition to the first seven codons of *ubq-1*. Its sequence is 5' cattggatccgt ATG CAA/G ATI TTT/C GTI AAA/G AC 3'.

C: RESULTS

1. Construction of ubq-1-lacZ strains

The *ubq-1-lacZ* fusions which were analysed are shown in Figure 14. *Ubq938* contains 938 bp of sequence upstream of the initial methionine, and extends 108 bp downstream to codon 36 of *ubq-1* , which is fused to *lacZ* in the nematode *lacZ* expression vector pPD16.43 (Fire et al. 1990). Two versions of *ubq938*, one with an SV40 nuclear localization signal upstream of *lacZ* (*ubq938NLS*) and one without, were used to establish transgenic lines. A deletion series extending from the 5' end of *ubq938* was constructed by exonuclease III digestion and the resulting products were cloned into pPD16.43. None of the deletion constructs possessed the SV40 nuclear localization signal. All *lacZ* fusions contained 988 bp of *ubq-1* 3' non-coding sequence cloned into the 3' polylinker of pPD16.43, to provide the polyadenylation signal and other potentially important elements.

Transgenic lines were established by co-injecting the construct of interest, together with a selectable marker, into *C. elegans* oocytes as described (Fire, 1986). Transformants selected with pPD10.41, an *unc-22* antisense vector, were identified by their twitching phenotype (Fire et al., in press) while worms transformed with pRF4, a *rol-6* dominant marker, were identified by their right rolling phenotype (Kramer et. al. 1990; Mello, personal communication). The transgenic lines obtained transmitted the marker phenotype at a frequency of 20-80%. Southern analysis (not shown) suggested that the injected DNA formed mixed extrachromosomal arrays, as described by Stinchcomb et al. (1985) and Fire (personal communication). Some deletions (Δ183, Δ523, Δ670, Δ768) were analyzed by transient assays only; in this procedure, the test plasmid alone was micro-injected, and the progeny of injected worms were stained. *Ubq 938-lacZ*, *ubqΔ827-lacZ*, *ubqΔ903-lacZ*, and *ubq-PvuIIΔ903-lacZ* constructs were initially analysed by transient assays, and subsequently in heritable lines.

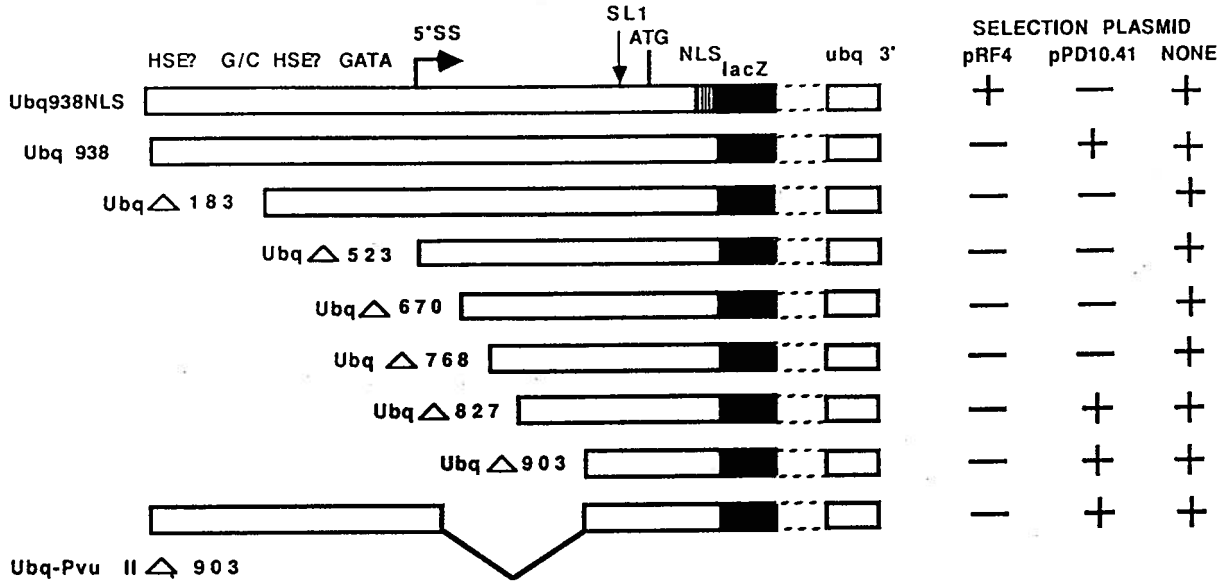


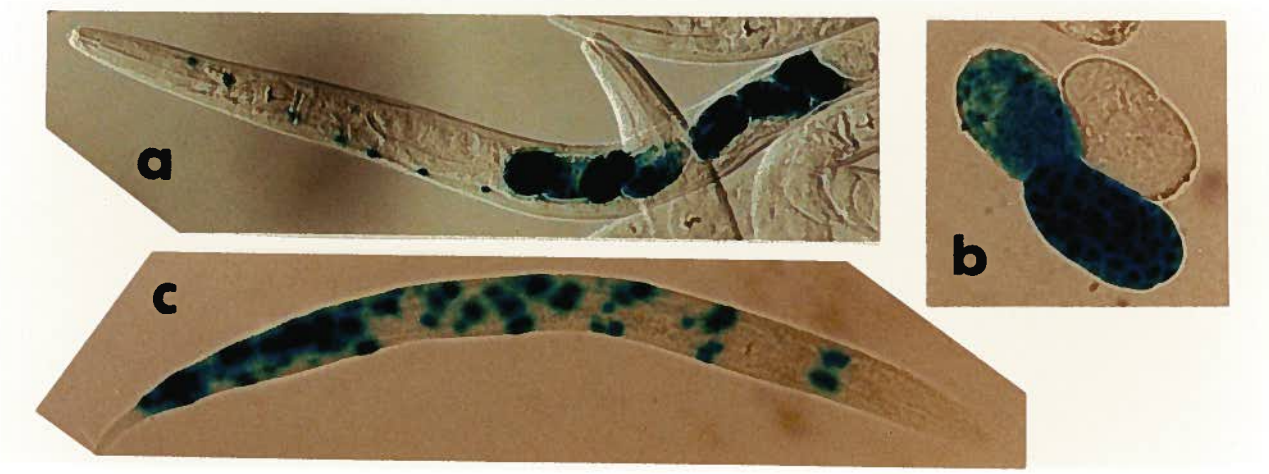
Fig. 14. Construction of *ubq-1-lacZ* fusions (not drawn to scale). Ubq938-NLSlacZ contains 1044 bp of *ubq-1* sequences 5' sequences extending from a Sal I site 938 bp upstream of the initial methionine to an EcoRI site 103 bp downstream of the methionine blunted into the Bam HI site of the *lacZ* expression vector, pPD16.43 (Fire et al., 1990). To construct ubq938-lacZ, a Hind III/Kpn I fragment containing the Sal I/Eco RI *ubq-1* fragment of ubq938NLS-lacZ was directionally cloned into pPD16.43 which had been digested with Hind III and Kpn I to remove the SV40 nuclear localization signal. Thus ubq938-lacZ and all of the deletions are in-frame translational fusions containing th first 36 amino acids of ubiquitin but are devoid of the SV40 NLS. Deletions extending from the Sal I site of ubq938-lacZ were generated by exonuclease III digestion. See Methods for details. The horizontal arrow indicates the direction of transcription from the putative start site based on S1 analysis (5'SS). The vertical arrow indicates the 3' acceptor site for trans-splicing with SL1 RNA which is located 6 bp upstream of the initial methionine of *ubq-1*. All of the deletions possessed the 3' acceptor site. HSE, putative heat shock element; G/C, sequence of cytosine residues; GATA, TATA-like sequence, GAATAA. Selection schemes tested with each lacZ fusion are designated by plus signs at the right of each fusion. Two plasmids were used to select for transformed lines, pRF4, a *rol-6* marker, and pPD10.41, an *unc-22* vector. In transient assays the test plasmid was injected alone.

2. Expression of the ubq938-lacZ transgene is constitutive but shows developmental tissue specificity

Initially transgenic lines were established which carried *ubq938* without the SV40 nuclear localization signal. The expression of the *ubq938-lacZ* transgene was constitutive at all stages of development in these strains. While some of the β-galactosidase staining was localized in the nucleus, the cytoplasm also stained extensively, making cell identification difficult. This problem was resolved by inclusion of the SV40 nuclear localization signal (Figure 15). Interestingly, there were differences in both the distribution and intensity of β-galactosidase staining between embryonic and post-embryonic stages. Embryos, from gastrulation onward, stained intensely within 15-30 minutes of the start of the incubation. In contrast, overnight incubation in the staining solution was usually required to achieve detectable somatic expression in post-embryonic stages. This suggests that embryos carrying the *ubq938-lacZ* transgene constitutively produce a greater quantity of β-galactosidase relative to somatic tissues. Moreover, the distribution of β-galactosidase activity was more limited in the post-embryonic stages. For example, in one strain, PC36 (ubq938NLS-lacZ/pRF4), one fifth of staining L1 larvae efficiently expressed the transgene in body muscle, hypodermis, pharyngeal tissue, and nervous tissue, while expression in adults was usually limited to body muscle (Figure 15). Approximately 20% of all F2 (derived from a single pre-selected rolling hermaphrodite) embryos, L1 and L2 larvae were stained, while the frequency of somatic tissue expression was a mere 5% in adults. However, 16% of PC36 adults carried stained embryos.

Subjecting *ubq938-lacZ* strains to a two hour heat shock treatment at 33°C had no detectable effect on the intensity or distribution of staining at any stage in development.

Fig. 15. *In situ* staining of β-galactosidase activity in PC36, a transgenic strain carrying an extra-chromosomal array composed of ubq938NLS and pRF4. (a) Expression in an adult after overnight incubation in stain. Gastrulation and later staged embryos are intensely stained while only a few body muscle nuclei stain in the somatic tissues. Magnification 150X. (b) Staining in comma stage (lower right) and pretzel stage (upper left) embryos. Magnification 370X. (c) Staining of primarily body muscle nuclei in an L1 larva. A few hypodermal nuclei are also stained. Magnification 370X.


a
b
c

3. *Expression is not diminished until 827 bp of sequence has been deleted from ubq938-lacZ*

Progressively larger deletions extending from the 5' end of *ubq938-lacZ* were tested by transient assays to determine if they altered the pattern of expression. Expression of the *ubqΔ183-lacZ* and *ubqΔ523-lacZ* transgenes was intense in embryos, while somatic expression in post embryonic stages was limited often to the pharynx or the body muscle in the head. Transient assays of *ubqΔ670-lacZ* and *ubqΔ768-lacZ* also failed to reveal any significant reduction in expression (Figure 16). Surprisingly, expression was not significantly altered until 827 bp of sequence had been removed to a point 330 bp downstream from the transcriptional start, or about 120 bp upstream of the initial methionine (Figure 16). In contrast to *ubq938-lacZ* embryos, which were usually stained to saturation within one hour (Figure 17a), *ubqΔ827-lacZ* embryos required four to six hours to reach comparable levels of staining (Figure 17e). Further removal of sequences to a point 30 bp upstream of the 3' acceptor site for SL1 (*ubqΔ903-lacZ*) drastically reduced expression (Figure 17g,h,i). Embryos carrying the *ubqΔ903-lacZ* transgene stained marginally after overnight incubation and only the occasional L1 larva stained (lightly) in the pharynx (Figure 17i). No somatic expression was observed in later larval stages.

These results suggest that sequences between the *ubqΔ827-lacZ* and the *ubqΔ903-lacZ* breakpoints contribute to expression while sequences upstream of the transcriptional start site (including a G/C rich block and a GATA box) are not required for embryonic expression of *ubq-1*. To test this hypothesis I assayed animals carrying *ubq-pvuIIΔ903-lacZ*, a construct with the first 538 bp of *ubq938-lacZ* fused to *ubq Δ903-lacZ*. Embryos carrying *ubq-pvuIIΔ903-lacZ* showed expression similar to that of *ubqΔ903-lacZ* (Figure 17j,k,l). To further analyze the *ubqΔ827-lacZ*, *ubqΔ903-lacZ*, and *ubq-pvuIIΔ903-lacZ* fusions, transgenic strains carrying these fusions were generated using the *unc-22* antisense vector, pPD10.41. The patterns of expression observed for these strains resembled those obtained in the transient assays.

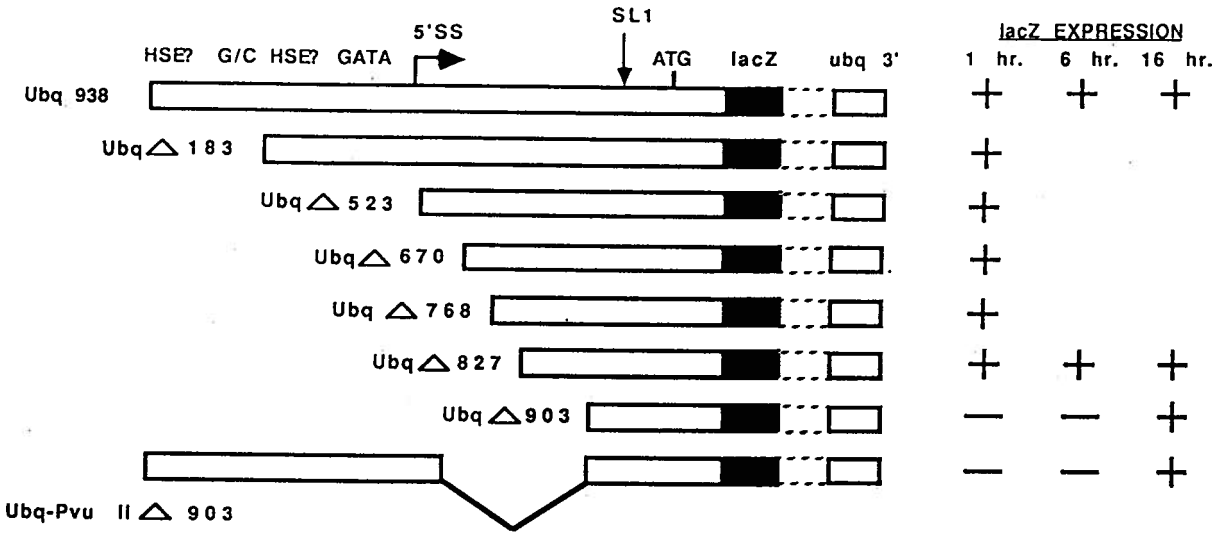
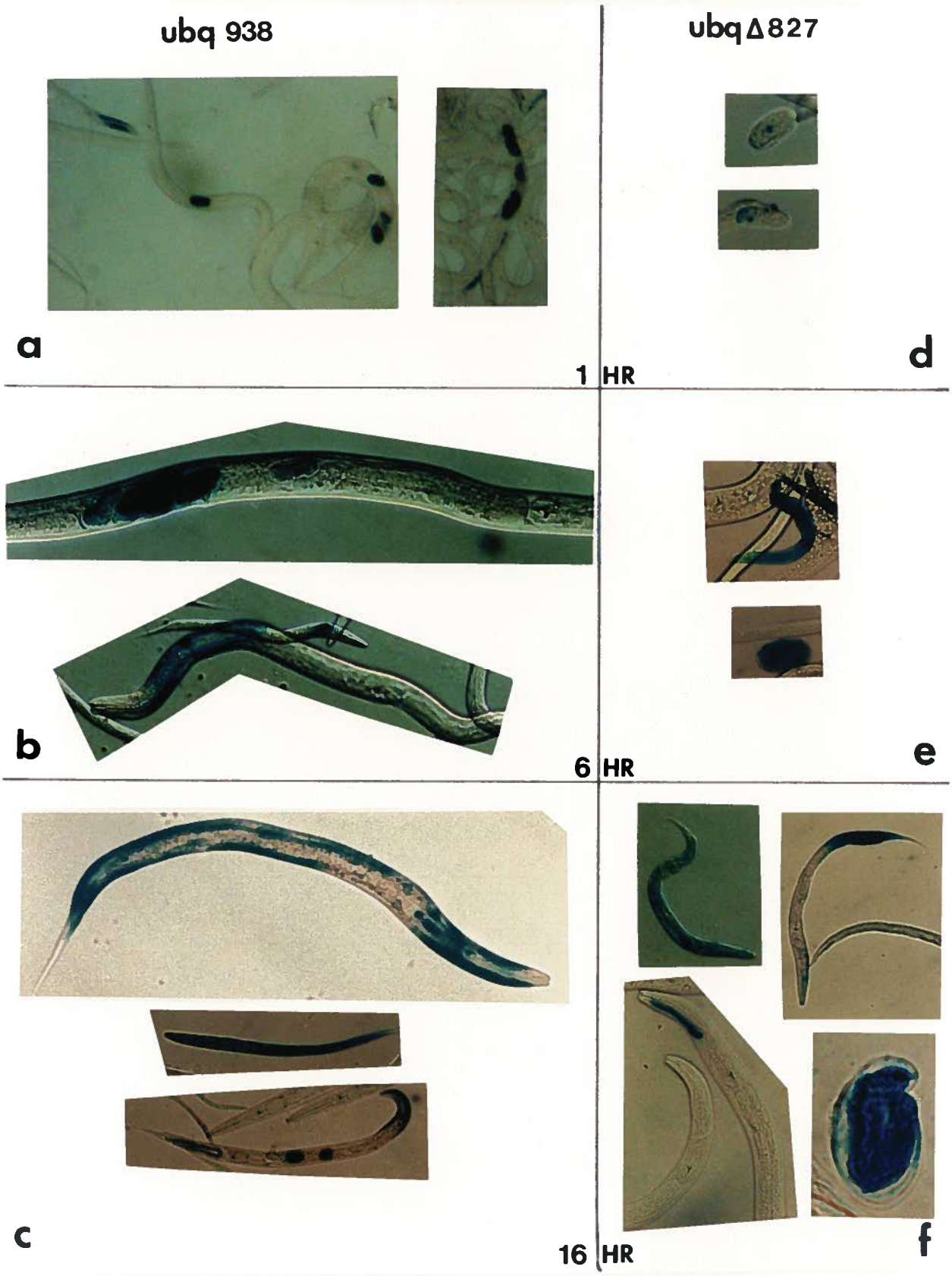
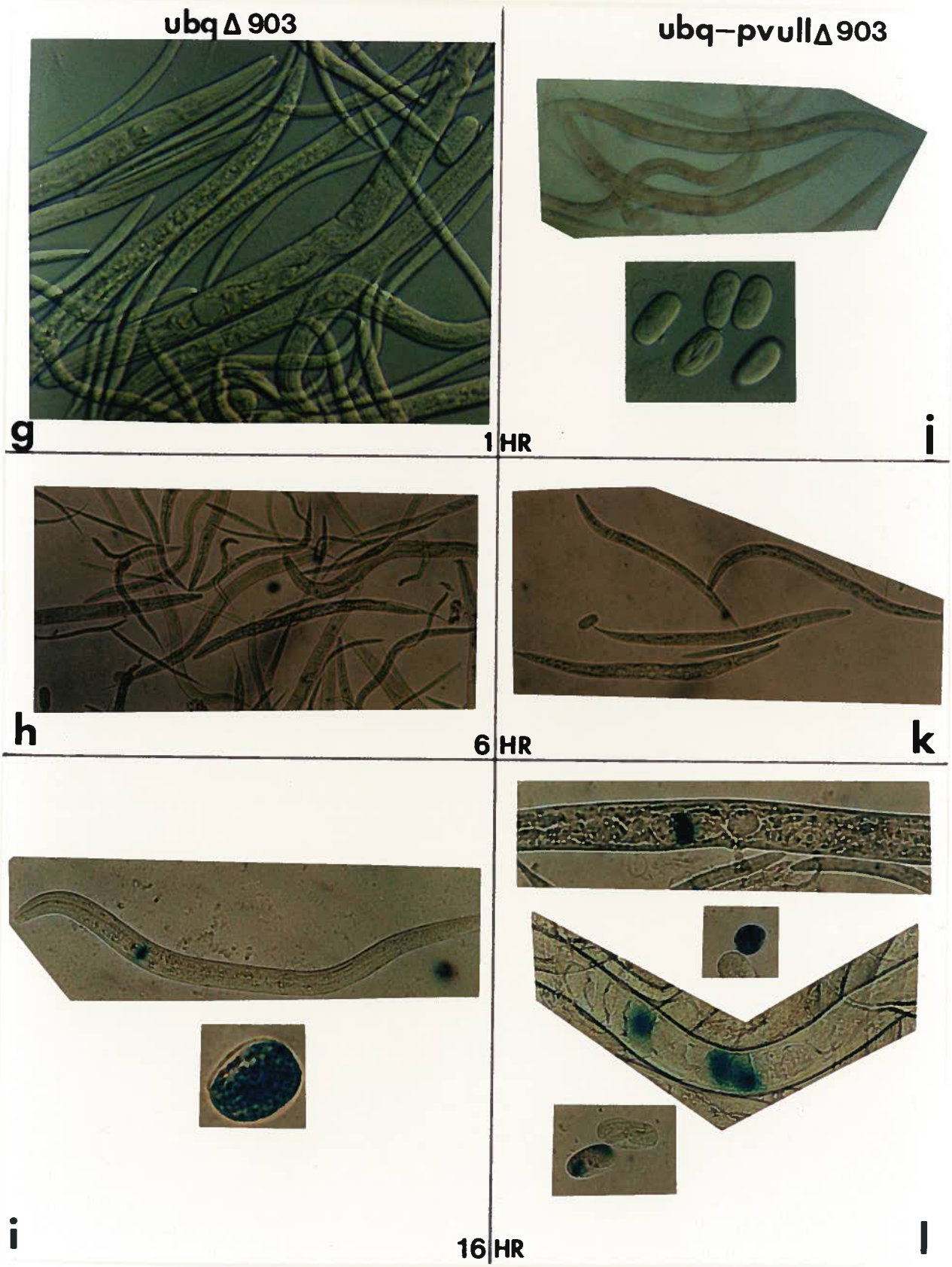


Fig. 16. Deletion analysis of the *ubq938-lacZ* transgene. Presence of β-galactosidase activity in transgenic strains and/or in transient assays after varying durations of incubation in stain is indicated by plus signs to the right of each construct.

Fig. 17. Expression of *ubq-1-lacZ* deletion transgenes. Transgenic animals were incubated for one, six, or sixteen hours in a solution containing Xgal (Fire et al. 1990) before being permanently mounted in 80% glycerol, 20 mM Tris-pH 8.0, 200 mM sodium azide for examination. (a,b,c): Staining of PC18, a ubq 938-lacZ/pPD10.41 strain. (a) After one hour, embryos within adults are saturated with stain. Staining of somatic tissues in larvae (upper left of picture) has also begun. (b) After six hours, β-galactosidase activity is increasingly apparent in the body muscle of larvae (bottom picture). (c) After 16 hours, expression in body muscle is obvious in an L3 larva, (top) and in the head of an adult (bottom). In addition, many L1 larvae are saturated with stain (centre). (d,e,f): Staining of PC4, a ubqΔ827-lacZ/ pPD10.41 strain. (d) After one hour, individual nuclei in pretzel stage embryos are staining. (e) At six hours, staining in newly hatched L1 larvae is obvious (top) and embryos are saturated (bottom). (f) 16 hours. Somatic tissue expression in the pharynx of an L4 (bottom left); throughout body of L1 (top left); and body muscle in posterior of L1 (top right). Saturated embryo (bottom right)(g,h,i): Staining of PC1, a ubqΔ903-lacZ/pPD10.41 strain. (g,h) No expression is observed in either somatic tissues or embryos even after six hours. (i) After 16 hours incubation, weak expression is visible in the terminal bulb of the pharynx of an L2 (top) and in a pretzel embryo. (j,k,l): Staining of PC17, a ubq-PvuIIΔ903-lacZ/pPD10.41 strain. (j,k) Even after six hours no expression is visible. (l) After 16 hours individual nuclei in pretzel stage embryos (bottom) are staining and occasionally moderate staining (top) is observed.


ubq 938
ubqΔ827
a
d
1 HR
b
e
6 HR
c
f
16 HR


ubq Δ 903
ubq–pvullΔ 903
g
1 HR
i
h
6 HR
k
i
16 HR
l

4. The ubq938-lacZ, and ubqΔ827-lacZ transcripts are trans-spliced

To determine if the *ubq-1-lacZ* fusion mRNAs, like those of *ubq-1*, were trans-spliced, the polymerase chain reaction (PCR) was utilized to selectively amplify only trans-spliced *lacZ* transcripts (Figure 18). Initially, the validity of this approach was verified by control experiments. First strand cDNA from both *ubq-1-lacZ* and endogenous *ubq-1* messages was generated from total RNA using the *lacZ* -specific oligonucleotide primer RGO5, or the oligo-dT primer RACE and reverse transcriptase. After the removal of excess primer, the cDNA was introduced into a polymerase chain reaction with various combinations of oligodeoxyribonucleotides and the resulting products separated by gel electrophoresis (Figure 19). When MQIF, a degenerate oligonucleotide encoding the first seven amino acids of ubiquitin, was combined with the *lacZ* oligo, OPC6, distinct bands of the expected sizes were observed in PC36 and PC4 experiments (Figure 19, Lanes 2 and 3) while the wild type negative control experiment revealed faint background bands (Lane1). Controls in which either SL1 or OPC6 alone was added to the reaction occasionally yielded very faint background bands (Figure 19, lanes 4 to 9), indicating that the majority of RGO5 or RACE had been removed.

To amplify only trans-spliced *ubq-1-lacZ* mRNAs, first strand cDNA was amplified with oligonucleotides specific for spliced leader 1 (SL1) and *lacZ* (OPC6, an oligonucleotide internal to RG05). The products of these reactions were separated on a 2% agarose gel and are shown in Figure 20 a. Bands of the appropriate sizes were consistently observed for ubq938-lacZ and ubqΔ827-lacZ; however, faint non-specific bands were sometimes observed when wild type cDNA was included as a control. To verify the identity of the bands observed in the ubq 938-lacZ and ubqΔ827-lacZ experiments, Southern analysis was carried out using probe MQIF, (Figure 20b). MQIF hybridized only to the putative ubq938-lacZ and ubqΔ827-lacZ bands, suggesting that they represent amplified trans-spliced *lacZ* transcripts. The results for ubqΔ903-lacZ and ubq pvuIIΔ903-lacZ were ambiguous: usually no amplified bands were observed (Figure 20), but occasionally faint bands of varying sizes were seen.

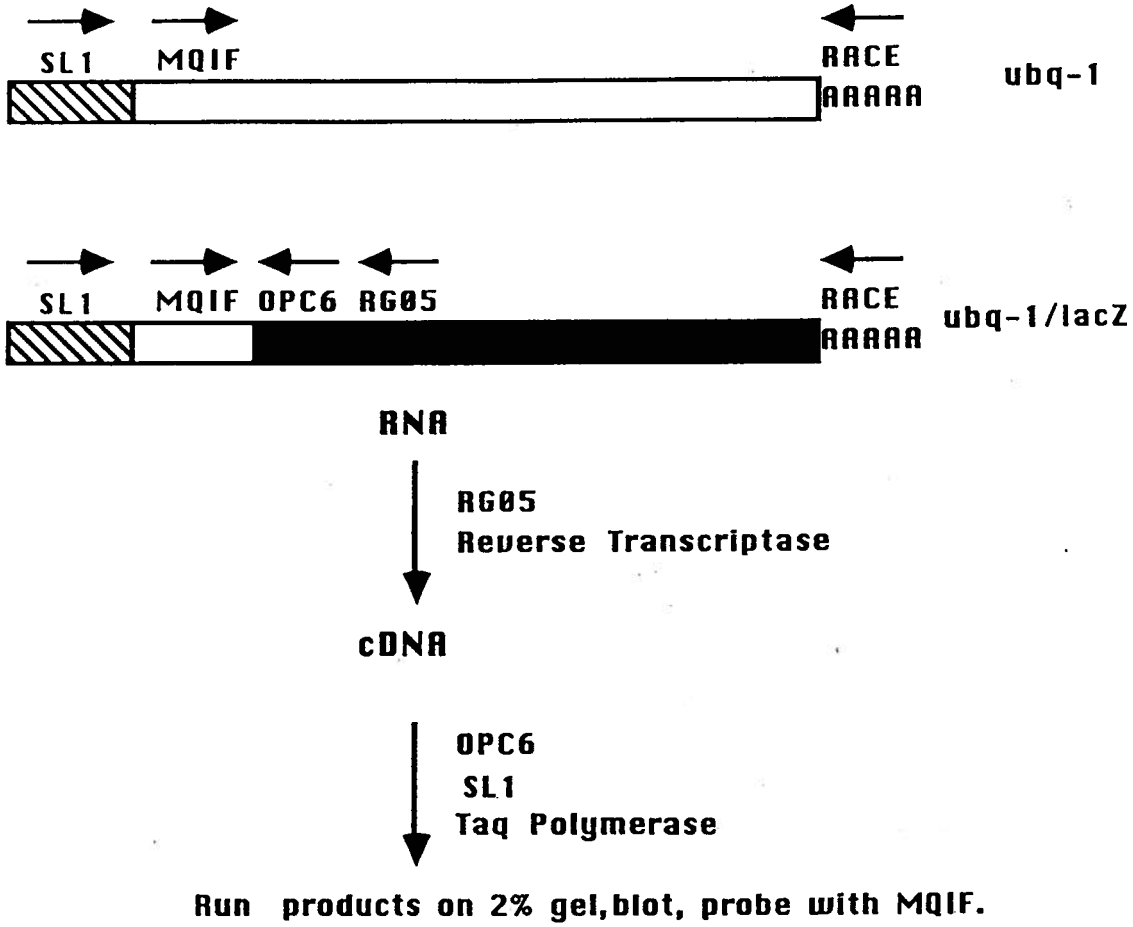


Fig. 18. Amplification of trans-spliced *ubq-1-lacZ* RNA. First strand cDNA was reverse transcribed from total nematode RNA with either oligodeoxyribonucleotide RG05 or RACE (oligo dT). The cDNA was introduced into a polymerase chain reaction with oligodeoxyriboucleotides OPC6, and SL1 and Taq DNA polymerase. The PCR products were separated by electrophoresis, transferred to a nylon membrane and probed with MQIF, an oligodeoxyribonucleotide corresponding to the first seven codons of ubiquitin.

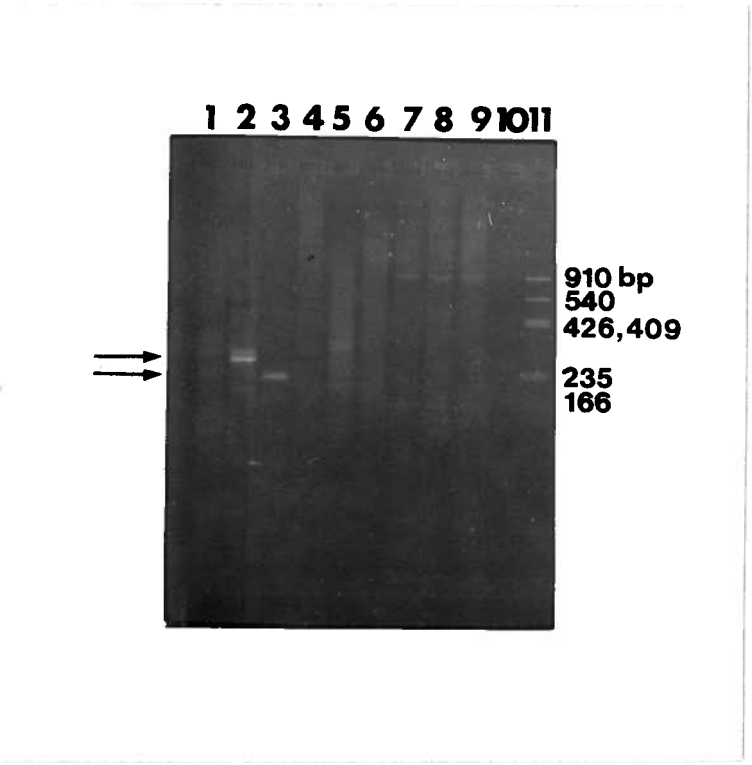


Fig. 19. Testing the integrity of the PCR amplification scheme. Various cDNAs were tested in positive and negative control experiments, and analyzed by electrophoresis on a 2 % agarose gel. See text for details. Lanes: 1, Wild type (N2) cDNA/ amplified with MQIF and OPC6; 2, PC36 (ubq938NLS-lacZ)/ MQIF and OPC6; 3, PC4 (ubqΔ827-lacZ)/ MQIF and OPC6; 4, Wild type/ SL1; 5, PC36 (ubq938NLS-lacZ)/ SL1; 6, PC4 (ubqΔ827-lacZ)/ SL1; 7, Wild type/ OPC6; 8, PC36/ OPC6; 9, PC4/ OPC6; 10, Blank; 11, pUC13 digested with DdeI molecular weight marker. The arrows indicate bands of the appropriate size present in the positive controls (lanes 2 and 3). The amplified product of PC36/ MQIF and OPC6 was expected to be 263 bp while the PC4 product was expected to be only 218 bp since it does not possess the SV40 NLS.

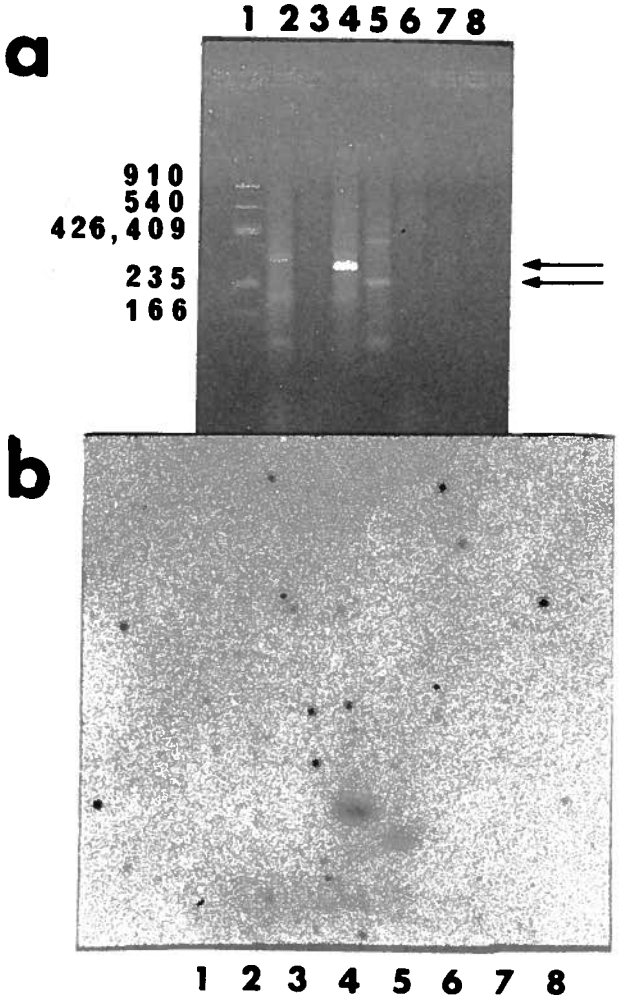


Fig.20. Analysis of PCR products amplified from *ubq-1-lacZ* RNA. RNA was reverse transcribed with MMLV reverse transcriptase and RG05. First strand cDNA was amplified with Taq Polymerase and the oligonucleotides, SL1 and OPC6. (a) The amplified products were separated by electrophoresis on a 2% gel. Lanes: 1, Molecular weight marker generated by digestion of plasmid pUC13 with restriction endonuclease DdeI; 2, Wild type control (N2); 3, Wild type (N2) control (In this experiment first strand cDNA was primed with oligo dT (RACE)); 4, PC36 (ubq938NLS-lacZ); 5, PC4 (ubqΔ827-lacZ); 6, PC1 (ubqΔ903-lacZ); 7, PC2 (ubqΔ903-lacZ); 8, PC17 (ubq-PvuIIΔ903-lacZ). The arrows indicate bands of the appropriate size present in lanes 4 and 5. The amplification product of the PC36 (ubq938NLS) experiment was expected to be 278 bp whereas the PC4 (ubqΔ827) PCR product was expected to be 233 bp since the latter construct does not possess the SV40 NLS. (b) Southern blot of the gel in (a) probed with oligo MQIF. Lanes: 1 to 8 as in (a).

D. DISCUSSION

1. Expression of ubq-1-lacZ transgenes in nematodes.

The ubq938-lacZ fusion was expressed in a tissue-general manner in embryos and in 20% of newly hatched transgenic L1 larvae. Surprisingly, the transgene was not expressed extensively in the somatic tissues of later stages. Given the known functions of ubiquitin, and the fact that *ubq-1* seems to be the only polyubiquitin gene in *C. elegans*, the expectation was that it would be expressed in most adult tissues. Indeed, using P element mediated transformation of *lacZ* fusions, Lee et al. (1988) found that the polyubiquitin gene of *Drosophila* is expressed in all tissues throughout development. In *Dictyostelium* and chicken, polyubiquitin genes are developmentally regulated, but each of these organisms possesses more than one polyubiquitin gene (Giorda and Ennis, 1987; Rocamora and Agell, 1990). Furthermore, Northern analysis of the endogenous *ubq-1* gene of *C. elegans* at various stages of development previously revealed that levels of *ubq-1* RNA remain relatively constant throughout development (Graham et al. 1989).

Thus, the tissue specificity in post-embryonic stages is surprising. The hypodermal expression observed occasionally in PC36 (ubq938-lacZ/pRF4) larvae (see Figure 15) may be an artefact due to enhancer sequences in the *rol-6* vector (pRF4). Such expression has not been confidently documented in ubq938-lacZ/pPD10.41 transgenic strains. On the other hand, the observed muscle expression seems to be real since animals carrying ubq938-lacZ constructs expressed the transgene in these nuclei regardless of which selection, if any, was employed.

It is conceivable that *ubq-1* sequences necessary for somatic expression in later stages were not included in the ubq938-lacZ construct. I therefore tested a construct which contained a further 763 bp of upstream sequence (Ubq1701-lacZ) in transient assays and found that expression of this transgene was similar to that of ubq938-lacZ. In addition to the 5' sequences, all of the *ubq-1-lacZ* fusions used in this study contained 988 bp of *ubq-1* 3' non-

coding sequence which included the endogenous polyadenylation signal. Thus it seems unlikely that critical 5' or 3' regulatory sequences were excluded from these constructs.

The endogenous *ubq-1* gene contains four typical cis-spliced introns in the first, fourth, seventh, and tenth ubiquitin coding repeats. None of these intron sequences were present in any of the *lacZ* fusions. It is possible that one or more of the *ubq-1* introns contains enhancer sequences which are necessary for maintaining expression in the post embryonic stages of *C. elegans*. Alternatively, the normal position of *ubq-1* in the genome relative to other sequences may be important for optimal expression.

Of the many functions attributed to ubiquitin, its role in targetting proteins for degradation is the best characterized. This raises the possibility that endogenous ubiquitin may have recognized the fusion protein as being abnormal ubiquitin and targetted it for proteolysis by the ATP dependent non-lysosomal proteolytic pathway. Using ubiquitin-lacZ fusions, Bachmair et al. (1986,1989) showed that ubiquitin-β-galactosidase fusion proteins were rapidly de-ubiquitinylated *in vitro* to release free ubiquitin and functional β-galactosidase, and that the half life of the released β-galactosidase depended upon the identity of the amino acid present at the mature amino terminus of the enzyme (the "N-end rule"). The ubiquitin moiety was almost always cleaved off the fusion protein after its carboxyl terminal glycine no matter which amino acid residue was at the amino end of β-galactosidase. Only when a proline residue was present at the amino terminus of β-galactosidase was the fusion protein not de-ubiquitinylated. In this case, the fusion protein was rapidly degraded.

I propose that the fusion proteins produced in this experiment could not be de-ubiquitinylated due to lack of the ubiquitin carboxyl terminus and that these proteins may have been recognized as aberrant and degraded by the ATP-dependent proteolytic pathway. Transformations of constructs which contain either one complete ubiquitin coding sequence fused to *lacZ*, or no ubiquitin coding sequences at all, should yield more extensive expression if this hypothesis is correct. Intriguingly, the *ubq-lacZ* fusion which Lee et al. (1988) used to

transform *Drosophila* contained no ubiquitin coding sequences but all of the upstream untranslated sequences.

2. Ubq-1 expression is not significantly heat inducible

Expression of the ubq938-lacZ fusion was neither more extensively distributed nor increased in level after a two hour heat shock treatment, in agreement with previous results of Graham et al. (1989) based on Northern analysis of *ubq-1* transcripts. In contrast, polyubiquitin genes of several organisms including chicken (Bond and Schlesinger 1985), yeast (*Saccharomyces cerevisiae*, Tanaka et al. 1988) and *Drosophila* (Lee et al. 1988) are induced upon heat shock. In *Drosophila*, however, the induction was weak and variable from experiment to experiment. Since modest increases in expression would be difficult to detect by either Northern analysis or *in situ lacZ* staining, we cannot rule out the possibility that *ubq-1* may be weakly heat inducible. There exist two regions upstream of *ubq-1* which possess sequence similarity to heat shock elements. The more distal of these, at -827 (Graham et al. 1989), contains two nGAAn motifs. It has been shown that while only two nGAAn motifs are required to stably bind heat shock factor *in vitro* (Xiao and Lis, 1988; Perisic et al., 1989), three such motifs are necessary to produce a strongly inducible heat shock element *in vivo* (Perisic et al., 1989; Xiao et al., 1991). By this criterion this sequence at best may represent a weak heat shock element; furthermore, it lies approximately 370 bp away from the putative transcription start site at -455. The proximal sequence at -650 contains only one nGAAn motif and thus is an unlikely candidate for a functional heat shock element. It is thus not surprising that *ubq-1* expression is not conspicuously induced by heat shock.

This does not negate the possibility that the ubiquitin mediated proteolytic pathway is heat inducible in *C. elegans*. Increased demand on this pathway after heat stress could be met by increasing the rate of ubiquitin conjugation to substrates and by increasing the quantity of available monomeric ubiquitin by post translational processing of polyubiquitin. The observations that two genes encoding ubiquitin conjugating enzymes in yeast are heat inducible

(Seufert and Jentsch, 1990) and that the *ubq-1* gene of *C. elegans* encodes a polyprotein of 11 ubiquitin molecules (Graham et al., 1989) provides credence to this idea.

3. Trans-splicing of ubq-1-lacZ transcripts

The results of the PCR analysis suggest that transcripts from the ubq938-lacZ and ubqΔ827-lacZ transgenes, like those of *ubq-1*, are correctly trans-spliced. No evidence for trans-splicing of ubqΔ903-lacZ or ubq PvuIIΔ903-lacZ was obtained, suggesting that the necessary signals had been removed. This raised the possibility that the reduction in expression observed with these constructs resulted from a failure in trans-splicing. Recent evidence however suggests that trans splicing by SL1 occurs whenever a 3' acceptor site is present unaccompanied by a 5' donor site (Conrad et al., 1991). The only apparent requirement is for an intact 3' splice acceptor site. Since both ubqΔ903-lacZ and the larger fusion ubqPvuIIΔ903-lacZ retained the 3' splice site and showed equivalent levels of expression, it is unlikely that the reduced expression seen in these transgenes is the result of a failure in trans-splicing. While PCR is a sensitive technique, presumably capable of amplifying single molecules, it is possible that there simply was not enough starting material in ubqΔ903-lacZ and ubqPvuIIΔ903-lacZ animals to allow the detection of trans-spliced products, especially given the initial reverse transcriptase step involved.

4. Analysis of the ubq-1 promoter.

A more likely explanation for the observed reduction in expression from the *ubqΔ903-lacZ* and *ubq-PvuIIΔ903-lacZ* constructs is the loss of elements of the *ubq-1* promoter. The breakpoint of *ubqΔ903-lacZ* is only 36 bp upstream of the initial methionine, and more than 400 bp downstream from a GAATAA sequence and putative transcriptional start site. While the limited expression in *ubqΔ903-lacZ* and *ubqPvuIIΔ903-lacZ* transgenics suggests that the promoter has not been completely destroyed, it seems likely that some regulatory sequences have been lost, and that these sequences are contained within the region between the *ubqΔ827-*

lacZ and *ubqΔ903-lacZ* breakpoints. This region lacks any obvious sequence similarity to known promoters or enhancers.

Additionally, it is conceivable that the *ubq-1* promoter has been completely destroyed and that *unc-22* enhancer sequences present in mixed arrays of pPD10.41 with *ubqΔ903-lacZ* or *ubq-PvuIIΔ903-lacZ* promotes the spurious expression of *lacZ* sequences observed in these strains.

Sequences upstream of the transcriptional start site, including the TATA-like sequence (GAATAA) and a long stretch of cytosine residues, appear to be unnecessary for embryonic expression of *ubq-1* . A similar situation has been documented with the *C. elegans unc-54* gene; in that case a tissue specific enhancer is sufficient for proper expression (A. Fire and S. White-Harrison, personal communication).

The constructs ubqΔ523::lacZ, ubqΔ670::lacZ, ubqΔ768::lacZ, ubqΔ827::lacZ and ubqΔ903::lacZ lack the wild type transcriptional start site and presumably utilize cryptic start sites present within the remaining 5' untranslated sequences or possibly within the vector. While it would be desirable to determine the actual start sites of these transgenes, this experiment is difficult for several reasons. Firstly, the fact that *ubq-1::lacZ* transcripts are rapidly trans-spliced means that very little of the total *ubq-1* RNA is present as unprocessed transcripts. Secondly, incomplete transmission of the transgenes as extrachromosomal arrays complicates isolation of the mass quantities of transgenic animals necessary for RNA analysis. Finally, expression of β-galactosidase is so limited in both ubqΔ903::lacZ, which does not possess the endogenous transcription start site, and in ubqPvuIIΔ903::lacZ, which does contain the *ubq-1* transcriptional start site, that the quantity of available material for RNA analysis would be reduced even further.

5. Conclusions

(1) The *ubq938-lacZ* transgene was expressed constitutively in embryos but showed developmental tissue specificity in the post embryonic stages. Somatic expression in L2, L3, L4 larvae and adults was usually confined to body or pharyngeal muscle.

(2) Heat shock did not alter the intensity or distribution of β-galactosidase staining in *ubq938-lacZ* transgenic animals. This supports previous evidence (Graham et al., 1989) that *ubq-1* is not heat inducible at the transcriptional level.

(3) Deletion analysis of the *ubq938-lacZ* transgene suggested that the region between the *ubqΔ768-lacZ* and *ubqΔ903-lacZ* breakpoints contains upstream activating sequences. A downshift in expression was first observed in *ubqΔ827-lacZ* transgenic animals and a further reduction was observed in *ubqΔ903-lacZ* animals.

(4) PCR analysis suggests that *ubq938-lacZ* and *ubqΔ827-lacZ* transcripts are efficiently trans-spliced by SL1 *in vivo*, and that there are no specific signals for trans-splicing by SL1 upstream of the *ubqΔ827* breakpoint.

6. Future Prospects

The results of this study suggested that sequences contained between the *ubqΔ827-lacZ* and *ubqΔ903-lacZ* breakpoints are important elements in the promotion of *ubq-1* expression. Future experiments should include a more detailed analysis of this region to isolate the sequences constituting the *ubq-1* promoter. In addition the potential role of introns in gene expression could be investigated by testing larger fusions which include entire ubiquitin repeats inclusive of introns.

I had anticipated that *ubq-1* would possess a strong tissue general constitutive promoter which might be confined to a small region, and this clearly was not the case. However, if future transformation experiments determine that sequences within the introns contain tissue specific enhancers, it may be possible to ligate all of these elements together to form a relatively

compact, tissue general promoter which could be used to drive constitutive expression of any coding region in *C. elegans.*

Currently, the ubiquitin-ribosomal fusion genes are being cloned and characterized by Don Jones in this laboratory. It would be interesting to determine the spatial and temporal expression pattern of these genes in *C. elegans* using the methods described in this thesis for comparison with that of *ubq-1*. In addition, deletion of the *ubq-1* gene by a recently described PCR mutagenesis procedure (Wood, *C. elegans* meeting abstracts, 1991) or over-expression of *ubq-1* under control of a heat shock promoter may assist in defining its importance for providing ubiquitin monomer in cells under various conditions.

Finally, mutagenesis of *ubq-1* and re-introduction into the nematode genome by transformation may define sites essential for the proper folding and function of ubiquitin.

III: REFERENCES

Albertson, D.G. and J.N. Thomson. 1976. Phil. Trans. R. Soc. Lond. B. 275:287-297.

Amin, J., J. Ananthan and R. Voellmy. 1988. Mol. Cell Biol. 8: 3761-3769.

Ananthan, J., A.L. Goldberg and R. Voellmy. 1986. Science 232:522-524.

Antonucci, T.K. 1985. Recombinant DNA Techniques 6:22-24. Univ. of Michigan.

Arrigo, A.-P., J.P.Suhan and W.J. Welch.1988. Mol. Cell Biol. 8:5059-5071.

Ashburner, M. and J.J. Bonner. 1979. Cell 17: 241-254.

Bachmair, A. and A. Varshavsky. 1989. Cell 56:1019-1032.

Bachmair, A., D. Finley, and A. Varshavsky. 1986. Science. 234:179-186.

Baker, R.T., and P.G. Board. 1987. Nucleic Acids Res. 15:443-463.

Ball, E., C.C. Karlik, C.J. Beall, D.L. Saville, J.C. Sparrow, B. Bullard, and E.A. Fryberg. 1987. Cell 51:221-228.

Banerji, S.S., K.Laing and R.I. Morimoto. 1987. Genes Dev. 1:946-953.

Bardwell, J.C. and E.A. Craig. 1984. Proc. Natl. Acad. Sci. USA 81:848-52.

Bartel, B., I. Wunning, and A. Varshavsky. 1990. EMBO J. 9:3179-3189.

Baxter , R. C. 1985. Endocrinology 117:650-655.

Beaulieu, J.-F., A.-P. Arrigo and R.M. Tanguay. 1989. J.Cell Sci. 92:29-36.

Beckmann, R.P., L.A. Mizzen, and W.J. Welch. 1990. Science 248:850-854.

Behlke, J., G. Lutsch, M. Gaestel and H. Bielka. 1991. FEBS Lett. 288:119-122.

Berger, E.M. and M.P. Woodward. 1983. Exp. Cell Res. 147:437-442.

Bienz, M. 1984. Proc. Natl. Acad. Sci. USA 81:3138-3142.

Bienz, M. and J.B. Gurdon. 1982. Cell 29:811-819.

Birnboim, H.C. and J. Doly. 1979. Nucleic Acids Res. 7:1513-1523.

Blumenthal, T. and J. Thomas. 1988. Trends Genet. 4:305-308.

Bond, U. 1988. EMBO J. 7:3509-3518.

Bond, U. and M.J. Schlesinger. 1985. Mol. Cell. Biol. 5:949-956.

Bond, U. and M.J. Schlesinger. 1986. Mol. Cell. Biol. 6:4602-4610.

Bond, U. and M.J. Schlesinger. 1988. Adv. Genet. 24:1-29.

Bonner, J.J. C. Parks, J. Parker-Thornberg, M.A. Mortin and H.R.B. Pelham. 1984. Cell 37:979-991.

Brenner, S. 1974. Genetics 77:71-94.

Browder, L. W., M. Pollock, J.J. Heikkila, J. Wilkes, T. Wang, P. Krone, N. Ovsenek, and M. Kloc. 1987. Dev. Biol. 124:191-199.

Brown, I.R., D.G. Low and L.A. Moran. 1985. Neurochem. Res. 10:1277-1284.

Bruzik, J.P., K. Van Doren, D. Hirsh, and J. Steitz. 1988. Nature 335:559-562.

Bukau, B. and G. Walker. 1989. J. Bacteriol. 171:6030-6038.

Busch, H. 1984. Methods Enzymol. 106:238-262.

Candido, E.P.M., D. Jones, D.K. Dixon, R.W. Graham, R.H. Russnak and R.J. Kay. 1989. Genome 31:690-697.

Chandrasekhar, G.N., K.Tilly, C.Woolford, R. Hendrix and C. Georgopoulos. 1986. J. Biol. Chem. 261:12414-12419.

Chappell, T.G., W.J. Welch, D.M. Schlossman, K.B. Palter, M.J. Schlesinger, and J.E. Rothman. 1986. Cell 45:3-13.

Chau, V., J.W. Tobias, A. Bachmair, D. Marriott, D.J. Ecker, D.K. Gonda, and A. Varshavsky. 1989. Science 243 :1576-1583.

Chirico, W. J., M.G. Waters and G. Blobel.1988. Nature 332: 805-810.

Chrétien, P. and J. Landry. 1988. J. Cell. Phys. 137:157-166.

Christiansen, E.N. and E. Kvamme. 1969. Acta. Physiol. Scand. 76:472-484.

Ciechanover, A., D. Finley, and A. Varshavsky. 1984. J. Cell Biochem. 24:27-53.

Ciechanover, A., J.A. DiGiuseppe, B. Bercovich, A. Orian, J. D. Richter, A.L. Schwartz, and G. M. Brodeur. 1991. Proc. Natl. Acad. Sci. USA 88:139-143.

Ciechanover, A., Y.Hod, and A. Hershko. 1978. Biochem. Biophys.Res.Commun. 81:1100-1104.

Clos, J., T. Westwood, P.B. Becker, S. Wilson, K. Lambert and C. Wu.1990. Cell 63:1085-1097.

Cohen, R. S. and M. Meselson. 1985. Cell 43:737-746.

Collier, N. C., J. Heuser, M.A. Levy , and M.A. Schlesinger.1988. J. Cell Biol. 106:1131-1139.

Conrad, R., J. Thomas, J. Spieth, and T. Blumenthal. 1991. Mol. Cell. Biol. 11:1921-1926.

Cooper, P. and T.-H. D. Ho. 1987. Plant Physiol. 84: 1197-1203.

Copeland, C. S., R.W. Doms, E.M. Bolzau, R.G. Webster, and A. Helenius. 1986. J. Cell Biol. 103: 1179-1191.

Corces, V., R. Holmgren, R. Freund, R. Morimoto and M. Meselson. 1980. Proc. Natl. Acad. Sci. USA 77:5390-5393.

Costlow, N. and J.T. Lis. 1984. Mol. Cell. Biol. 4:1853-1863.

Coulson, A., J.E. Sulston, S. Brenner and J. Karn. 1986. Proc. Natl. Acad. Sci. USA 83:7821-7825.

Coulson, A., R. Waterston, J. Kiff, J. Sulston and Y. Kohara. 1988. Nature 335:184-186.

Davie, J.R. and L.C. Murphy. 1990. Biochem. 29: 4752-4757.

DeMarzo, A. M., C.A. Beck, S.A. Onati, and D.P. Edwards. 1991. Proc. Natl. Acad. Sci. USA 88: 72-76.

Deshaies, R. J., B.D. Koch, M. Werner-Washburne, E.A. Craig, and R. Sheckman. 1988. Nature 332:800-805.

Dickson, J.A. and B.E. Oswald. 1976. Br. J. Cancer. 34:262-271.

DiDomenico, B.J., G.E. Bugiasky and S. Lindquist. 1982a. Cell 31:593-603.

DiDomenico, B.J., G.E. Bugiasky, and S. Lindquist. 1982b. Proc. Natl. Acad. Sci. USA 79:6181-6185.

DiNocera, P.P. and I.B. Dawid. 1983. Proc. Natl. Acad. Sci. USA 80:7095-7098.

Dixon, D. K., D. Jones, and E.P.M. Candido. 1990. DNA Cell Biol. 9:177-191.

Dohmen, R.J., K. Madura, B. Bartel and A. Varshavsky. 1991. Proc. Natl. Acad. Sci. USA 88:7351-7355.

Dudler, R. and A. A. Travers. 1984. Cell 38:391-398.

Dworkin-Rastl, E., A. Shrutkowski, and M. B. Dworkin. 1984. Cell 39: 321-325.

Ellis, H. M. and H.R. Horvitz. 1986. Cell 44: 817-829.

Emmons, S. W., M.R. Klass, and D. Hirsh. 1979. Proc. Natl. Acad. Sci. USA 76:1333-1337.

Emmons, S.W., B. Rosenzweig and D. Hirsh. 1980. J. Mol. Biol. 144:481-500.

Epstein, H. F., R.H. Waterston, and S. Brenner. 1974. J. Mol. Biol. 90:291-300.

Falkner, F.-G., H. Saumweber and H. Biessman. 1981. J. Cell Biol. 91:175-183.

Findly, R.C. and T. Pederson. 1981. J. Cell Biol. 88:323-328.

Findly, R.C., R.J. Gillies and R.G. Schulman. 1983. Science. 219:1223-1225.

Finley, D., A. Ciechanover, and A. Varshavsky. 1984. Cell 37:43-55.

Finley, D., B. Bartel, and A. Varshavsky. 1989. Nature 338:394-400.

Finley, D., E. Ozkaynak, and A. Varshavsky. 1987. Cell 48: 1035-1046.

Fire, A. 1986. EMBO J. 5:2673-2680.

Fire, A., S. White Harrison, D. Albertson, and D. Moerman. 1991. Development in press.

Fire, A., S. White Harrison, and D. Dixon. 1990. Gene 93: 189-198.

Fried, V.A., H.T. Smith, E. Hildebrandt, and K. Weiner. 1987. Proc. Natl. Acad. Sci. USA 84:3685-3689.

Frohman, M.A., M.K. Dush, and G.R. Martin. 1988. Proc. Natl. Acad. Sci. USA 85:8998-9002.

Gaestel, M., W. Schroder, R. Benndorf, C. Lippmann, K. Buchner, F. Hucho, V.A. Erdmann, and H. Bielka. 1991. J. Biol. Chem. 266:14721-14724.

Gaitanaris, G.A., A.G. Papavassiliou, P. Rubock, S.J. Silverstein and M.E. Gottesman. 1990. Cell 61:1013-1020.

Gaterman, K.B., G.H. Rosenberg and N.F. Kaufer. 1988. Biotechniques 6:951-952.

Gething, M.-J., K. McCammon, and J. Sambrook. 1986. Cell 46:939-950.

Gilmour, D.S. and J.T. Lis. 1986. Mol. Cell. Biol. 6:3984-3989.

Giorda, R. and H.L. Ennis. 1987. Mol. Cell Biol. 6:2097-2103.

Glaser, R.L., M.F. Wolfner and J.T. Lis. 1986. EMBO J.5:747-754.

Glotzer, M. A.W. Murray, and M.W. Kirschner. 1991. Nature 349:132-138.

Goebl, M.G., J. Yochem, S. Jentsch, J.P. MaGrath, A. Varshavsky and B. Byers. 1988. Science 241:1331-1335.

Goldenberg, C.J., Y. Luo, M. Fenna, R. Baler, R. Weinmann and R. Voellmy. 1988. J. Biol. Chem. 263:19734-19739.

Goldschmidt, R. 1935. Abst. u. Vererb. 69:38-69; 70-131.

Goldstein, G., M. Scheid, U. Hammerling, E.A. Boyse, D.H. Schlesinger, and H.D. Niall. 1975. Proc. Natl. Acad. Sci. USA 72:11-15.

Goloubinoff, P., A.A. Gatenby and G.H. Lorimer. 1989. Nature. 337:44-47.

Gorin, E. and H.M. Goodman. 1985. Endocrinology 116:1796-1805.

Grabowski, P.J., R.A. Padgett and P.A. Sharp. 1984. Cell 37:415-427.

Graham, R.W., D.Jones, and E.P.M. Candido. 1989. Mol. Cell. Biol. 9:268-277.

Graham, R.W., K. Van Doren, S. Bektesh, and E.P.M. Candido. 1988. J. Biol. Chem. 263:10415-10419.

Grossman, A.D., D.B. Straus, W.A. Walter and C.A. Gross. 1987. Genes Dev. 1:179-184.

Haas, A.L., P. Ahrens, P.M. Bright, and H. Ankel. 1987. J. Biol. Chem. 262: 11315-11323.

Hammond, G.L., Y.-K. Lai and C.L. Markert. 1982. Proc. Natl. Acad. Sci. USA 79:3485-3488.

Hanahan, D. 1983. J. Mol. Biol. 166: 557-580.

Hecht, R. M., L.A. Gossett, and W.R. Jeffery. 1981. Dev. Biol. 83:373-379.

Hemmingsen, S.M., C. Woolford, S.M. van der Vies, K.Tilly, D.T. Dennis and C.P. Georgopoulos, R.W. Hendris and R.J. Ellis. 1988. Nature. 333:330-334.

Henikoff, S. 1984. Gene 28:351-359.

Hershko A. and A. Ciechanover. 1986. Prog. Nuc. Acid Res. 33:19-55.

Hershko, A. 1988. J. Biol. Chem. 263:15237-15240.

Hightower, L.E. 1980. J. Cell. Physiol. 102:407-427.

Hiromi, Y., H. Okamoto, W.J. Gehring and Y. Hotta. 1986. Cell 44:293-301.

Hochstrasser, M., M.J. Ellison, V. Chau, and A. Varshavsky. 1991. Proc. Natl. Acad. Sci. USA 88:4606-4610.

Hockertz, M. K., I. Clark-Lewis, and E.P.M. Candido. 1991 FEBS Lett. 280: 375-378.

Hoffman, E. and V. Corces. 1986. Mol. Cell Biol. 6: 663-673.

Hoffman, E.P., S.L. Gerring and V.G. Corces. 1987. Mol. Cell. Biol. 7:973-981.

Huang, X.-Y., and D. Hirsh. 1989. Proc. Natl. Acad. Sci. USA 86:8640-8644.

Hultmark, D., R. Klemenz , and W.J. Gehring. 1986. Cell 44: 429-438.

Ingolia, T.D. and E.A. Craig. 1982. Proc. Natl. Acad. Sci. USA 79:2360-2364.

Ireland, R. C. and E.M. Berger. 1982. Proc. Natl. Acad. Sci. USA 79: 855-859.

Jentsch, S., J.P. McGrath and A. Varshavsky. 1987. Nature 329: 131-134.

Jones, D., D.K. Dixon, R.W. Graham, and E.P.M. Candido. 1989. DNA. 8:481-490.

Jones, D., R.H. Russnak, R.J. Kay, and E.P.M.Candido. 1986. J. Biol. Chem. 261:12006-12015.

Kao, H.-T., O. Capasso, N. Heintz, and J.R. Nevins. 1985. Mol. Cell. Biol. 5:628-633.

Kassenbrock, C. K., P.D. Garcia, P. Walter, and R.B. Kelly. 1988. Nature 333: 90-93.

Kay, R. J., R.J. Boissy, R.H. Russnak, and E.P.M. Candido. 1986. Mol. Cell Biol. 6:3134-3143.

Kay, R.J., R.H. Russnak, D.Jones, C. Mathias and E.P.M. Candido. 1987. Nucleic Acids Res. 15:3723-3741.

Kingston, R.E., T.J. Schuetz and Z. Larin. 1987. Mol. Cell. Biol. 7:1530-1534.

Klemenz, R., D. Hultmark, and W.J. Gehring. 1985. EMBO J. 4:2053-2060.

Kramer, J. M., R.P. French, E.-C. Park, and J.J. Johnson. 1990. Mol. Cell Biol. 10:2081-2089.

Krause, M., and D. Hirsh. 1987. Cell 49:753-761.

Kusukawa, N., T. Yura, C. Veguchi, Y. Akiyawa and K. Ito. 1989. EMBO J. 8:3517-3521.

Laird, P.W. 1989. Trends Genet. 5:204-208.

Laird, P.W., J.M. Kooter, N. Loosbroek, and P. Borst. 1985. Nuc. Acids Res. 13:4253-4266.

Landry, J., P. Chretien, H. Lambert, E. Hickey, and L.A. Weber. 1989. J. Cell Biol.109:7-15.

Laszlo, A., and G.C. Li. 1985. Proc. Natl. Acad. Sci. USA 82:8029-8033.

Lee, H., J.A. Simon, and J.T. Lis. 1988. Mol. Cell Biol. 8:4727-4735.

Leicht, B. G., H. Biessman, K. Palter, and J.J. Bonner. 1986. Proc. Natl. Acad. Sci. USA 83:90-94.

Lenkinsky, R.E., D.M. Chen, J.D. Glickson, and G. Goldstein. 1977. Biochem. Biophys. Acta 494: 126-130.

Lepock, J.R., K.-H. Cheng, H. Al-Qysi and J. Kruuv. 1983. Can. J. Biochem. Cell Biol. 61:428-437.

Leung, D.W., S. A. Spencer, G. Cachianes, R.G. Hammonds, C. Collins, W.J. Henzel, R. Barnard, M.J. Waters and W.I. Wood. 1987. Nature 330:537-543.

Levinger, L. and A. Varshavsky. 1982. Cell 28:375-385.

Li, G.C. and Z. Werb. 1982. Proc. Natl. Acad. Sci. USA 79:3218-3222.

Lindquist, S. 1981. Nature 293:311-314.

Lindquist, S. and B. DiDomenico. 1985. Coordinate and non-coordinate gene expression during heat shock. A model for regulation. Academic Press, New York.

Lindquist, S. and E.A. Craig. 1988. Annu.Rev. Genet. 22:631-677.

Lindquist, S.L. 1980. J. Mol. Biol. 137:151-158.

Littlewood, T.D., D.C. Hancock and G.I. Evan. 1987. J. Cell Sci. 88:65-72.

Loomis, W.F. and S.A. Wheeler. 1982. Dev. Biol. 90:412-418.

Lund, P.K., B.M. Moats-Staats, J.G. Simmons, E. Hoyt, A.J. D'Ercole, F. Martin, and J.J. Van Wyk. 1985. J. Biol. Chem. 260:7609-7613.

Madura, K., S. Prakash and L. Prakash. 1990. Nuc. Acids Res. 18:771-778.

Maniak, M. and W. Nellen. 1988. Mol. Cell. Biol. 8:153-159.

Maniatis, T., E.F. Fritsch and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press.

Marota, F.G. and J.M. Sierra. 1988. J. Biol. Chem. 263:15720-15725.

Marota, F.G. and J.M. Sierra. 1989. Mol. Cell. Biol. 9:2181-2190.

Mayrand, S. and T. Pederson. 1983. Mol. Cell. Biol. 3:161-171.

McGarry, T. J. and S. Lindquist. 1985. Cell 42:903-911.

McGarry, T.J. and S. Lindquist. 1986. Proc. Natl. Acad. Sci. USA 83:399-403.

McKenzie, S.L., S. Henikoff and M. Meselson. 1975. Proc. Natl. Acad. Sci. USA 72:1117-1121.

Melton, D.A., ed. 1988. Current communications in molecular biology. Antisense RNA and DNA, pp. 71-78. Cold Spring Harbour Laboratory, Cold Spring Harbour, New York.

Milkman, R. 1966. Biol. Bull. 131:331-345.

Miller, D. 1989. New Scientist. 1:47-50.

Mitchell, H.K., G. Moller, N.S. Petersen and L. Lipps-Sarmiento. 1979. Dev. Genet. 1:181-192.

Moerman, D. G. and D.L. Baillie. 1979. Genetics 91:95-104.

Mondovi, B., R. Stron, G. Rotilio, A.F. Argo, R. Cavaliere and A. Rossi Fanelli. 1969. Eur. J. Cancer 5:129-136.

Moran, L., M.-E. Mirault, A.P.Arrigo, M. Goldschmidt-Clermont, and A.Tissieres. 1978. Phil. Trans. R. Soc. Lond. B. 283:391-406.

Muhich, M.L. and J.C. Boothroyd. 1988. Mol. Cell. Biol. 8:3837-3846.

Muller, W.U., G.C. Li and L.S. Goldstein. 1985. Int. J. Hyperthermia. 1:97-102.

Munro, S. and H.R.B. Pelham. 1986. Cell 46:291-300.

Munro, S. and H.R.B. Pelham. 1987. Cell 48:899-907.

Neves, A. M., I. Barahona, L. Galego, and C. Rodrigues-Pousada. 1988. Gene 73: 87-96.

Nevins, J.R. 1982. Cell 29:913-919.

Newport, J. and M. Kirshner. 1982. Cell 30:675-686.

Nguyen, V.T., M. Morange and O. Bensaude. 1989. J. Biol. Chem. 264:10487-10492.

Nickel, B.E., and J.R. Davie. 1989. Biochemistry 28:964-968.

Nickells, R. W., L.W. Browder and T.I. Wang. 1989. Biochem. Cell Biol. 67: 687-695.

Nickells, R.W. and L.W. Browder. 1988. J. Cell Biol. 107:1901-1909.

Nieto-Sotelo, J., G. Wiederrecht, A. Okuda and C.S. Parker. 1990. Cell 62:807-817.

Nover, L., K.-D. Scharf, and D. Neumann. 1989. Mol. Cell Biol. 9:1298-1308.

Oppermann, H., W. Levinson and J.M. Bishop. 1981. Proc. Natl. Acad. Sci. USA 78:1067-1071.

Otaka, E., K. Higo, and T. Itoh. 1984 . Molec. Gen. Genet. 195:544-546.

Ozkaynak, E., D. Finley and A. Varshavsky. 1984. Nature 312: 663-666.

Ozkaynak, E., D. Finley, M.J. Solomon and A. Varshavsky. 1987. EMBO J. 6: 1429-1439.

Padgett, R.A., M.M. Konarska, P.J. Grabowski, S.F. Hardy and P.A. Sharp. 1984. Science 225:898-903.

Parag, H.A., B. Raboy, and R.G. Kulka. 1987. EMBO J. 6:55-61.

Parsell, P.A., Y. Sanchez, J.D. Stitzel and S. Lindquist. 1991. Nature 353:270-273.

Pelham, H. R. B., in J.K. Setlow (ed.) 1987. Genetic Engineering. 9:27-44. Plenum Press.

Pelham, H.R.B. 1982. Cell 30:517-528.

Pelham, H.R.B. 1988. Nature 332:776-777.

Pelham, H.R.B. and M. Bienz. 1982. EMBO J. 1:1473-1477.

Perisic, O., H. Xiao, and J.T. Lis. 1989. Cell 59:797-806.

Perry, G., R. Friedman, G. Shaw and V. Chau. 1987. Proc. Natl. Acad. Sci. USA 84: 3033-3036.

Petersen, N.S. and H.K. Mitchell. 1981. Proc. Natl. Acad. Sci. USA 78:1708-1711.

Petersen, N.S., G. Moller and H.K. Mitchell. 1979. Genetics 92:891-902.

Petersen, R.B. and S. Lindquist. 1989. Cell Reg. 1:135-149.

Plesofsky-Vig, N. and R. Brambl. 1990. J. Biol. Chem. 265:15432-15440.

Rabindran, S.K., G. Giorgi, J. Clos and C. Wu. 1991. Proc. Natl. Acad. Sci. USA 88:6906-6910.

Reading, D.S., R.L. Hallberg and A.M. Myers. 1989. Nature 337: 655-659.

Redman, K.L. and M. Rechsteiner. 1989. Nature 338:438-440.

Reymond, C.D. 1987. Nuc. Acids Res. 15:8118.

Riddihough, G. and H.R.B. Pelham. 1986. EMBO J. 5:1653-1658.

Ritossa, F.M. 1962. Experientia 18:571-573.

Rocamora, N. and N. Agell. 1990. Biochem J. 267: 821-829.

Rougvie, A.E. and J.T. Lis. 1988. Cell 54:795-804.

Ruskin, B., A.R. Krainer, T. Maniatis and M.R. Green. 1984. Cell 38:317-331.

Russnak, R. H. and E.P.M. Candido. 1985. Mol. Cell Biol. 5:1268-1278.

Russnak, R. H., D. Jones, and E.P.M. Candido. 1983. Nucleic Acids Res.11:3187-3205.

Sadis, S., E. Hickey and L.A. Weber. 1988. J. Cell Physiol. 135: 377-386.

Sambrook, J., E. F. Fritsch and T. Maniatis. 1989. Molecular Cloning : A Laboratory Manual, second edition. Cold Spring Harbor Laboratory Press.

Sanchez, E. R., D.O. Toft, M.J. Schlesinger, and W.B. Pratt. 1985. J. Biol. Chem. 260:12398-12401.

Sanger, F., S. Nicklen, and A.R. Coulson. 1977. Proc. Natl. Acad. Sci. U.S.A. 74:5463-5467.

Scharf, K.-D., S.Rose, W. Zott, F. Schöff and L. Nover. 1990. EMBO J. 9:4495-4501.

Schlesinger, M.J. 1990. J. Biol. Chem. 265:12111-12114.

Schuetz, T.J., G.J. Gallo, L. Sheldon, P.Tempst and R.E. Kingston. 1991. Proc. Natl. Acad. Sci. USA 88:6911-6915.

Scott, M.P. and M.L. Pardue. 1981. Proc. Natl. Acad. Sci. USA 78:3353-3357.

Seufert, W. and S. Jentsch. 1990. EMBO J. 9:543-550.

Shanklin, J., M. Jabben, and R.D. Viestra. 1987. Proc. Natl. Acad. Sci. USA 84:359-363.

Shuey, D. J. and C.S. Parker. 1986. J.Biol.Chem. 261:7934-7940.

Siegelman, M., M.W.Bond, W. M. Gallatin, T. St. John, H.T. Smith, V.A. Fried and I.L. Weissman. 1986. Science 231:823-829.

Simon, M.C., K. Kitchener, H.-T. Kao, E. Hickey, L. Weber, R. Voellmy, N. Heintz and J.R. Nevins. Mol. Cell. Biol. 7:2884-2890.

Skowyra, D., C. Georgopoulos and M. Zylicz. 1990. Cell 62:939-944.

Snutch, T. P. and D.L. Baillie. 1983. Can. J. Biochem. Cell Biol. 61:480-487.

Snutch, T.P. and D. L. Baillie. 1984. Mol. Gen. Genet. 195:329-335.

Snutch, T.P., M.F.P. Heschl and D.L. Baillie. 1988. Gene 64:241-255.

Sorger, P. K. and H.C.M. Nelson. 1989. Cell 59:807-813.

Sorger, P.K. 1990. Cell 62: 793-805.

Sorger, P.K. and H.R.B. Pelham. 1987. EMBO J. 6:3035-3041.

Sorger, P.K. and H.R.B. Pelham. 1988. Cell 54:855-864.

Sorger, P.K., M.J. Lewis and H.R.B. Pelham. 1987. Nature 329:81-84.

Southern, E.M. 1975. J. Mol. Biol. 98:503-517.

Southgate, R., A. Ayme and R. Voellmy. 1983. J. Mol. Biol. 165:35-37.

Spector, A., R. Chiesa, J. Sredy and W. Garner. 1985. Proc. Natl. Acad. Sci.USA 82:4712-4716.

Stinchcomb, D. T., J.E. Shaw, S.H. Carr, and D. Hirsh. 1985. Mol. Cell. Biol. 5:3483-3496.

Stone, D.E. and E.A. Craig. 1990. Mol. Cell. Biol. 10:1622-1632.

Storti, R.V., M.P. Scott, A.Rich and M.L. Pardue. 1980. Cell 22:825-834.

Subjeck, J.R. and T.T. Shyy. 1986. Am. J. Physiol. 250:C1-C17.

Subjeck, J.R., T.Shyy, J.Shen and R.J. Johnson. 1983. J. Cell Biol. 97:1389-1395.

Sulston, J. E. 1976. Phil. Trans. R. Soc. Lond. B. 275:287-297.

Sulston, J. E. and H.R. Horvitz. 1977. Dev. Biol. 56:110-156.

Sulston, J. E., E. Schierenberg, J.G. White and J.N. Thomson. 1983. Develop. Biol. 100: 64-119.

Sulston, J.E. and S. Brenner. 1974. Genetics 77: 95-104.

Sung, P., S. Prakash, and L. Prakash. 1988. Genes and Dev. 2:1476-1485.

Sung, P., S. Prakash, and L. Prakash. 1990. Proc. Natl. Acad. Sci. USA 87:2695-2699.

Sung, P., E. Berleth, C. Pickart, S. Prakash and L. Prakash. 1991. EMBO J. 10:2187-2193.

Sutton, R.E., and J.C. Boothroyd. 1986. Cell 47:527-535.

Sutton, R.E., and J.C. Boothroyd. 1988. EMBO J. 7:1431-1437.

Tanaka, K., K. Matsumoto, and A. Toh-e. 1988. EMBO J. 7:495-502.

Theodorakis, N.G. and R.I. Morimoto. 1987. Mol. Cell. Biol. 7:4357-4368.

Thomas, G. H. and S.C.R. Elgin. 1988. EMBO J. 7:2191-2201.

Thomas, G.P., W.J. Welch, M.B. Mathews and J.R. Feramsico. 1982. Cold Spring Harbour Symp. Quant. Biol. 46:985-996.

Thomas, J.D., R.C. Conrad, and T. Blumenthal. 1988. Cell 54:533-539.

Tilly, K. and G. Georgopoulos. 1982. J. Bacteriol. 149:1082-1088.

Tissières, A., H.K. Mitchell and U.M. Tracey. 1974. J. Mol. Biol. 84:389-398.

Todd, J.A., T.J.P. Hubbard, A.A. Travers, and D.J. Ellar. 1985. FEBS Lett. 188:209-214.

Tomasovic, S.P. 1989. Life Chem. Reports. 7:33-63.

Toniolo, D., M. Persico, and M. Alcalay. 1988. Proc. Natl. Acad. Sci. USA 85:851-855.

Van der Ploeg, L.H.T., S.H. Giannini and C.R. Cantor. 1985. Science 228:1443-1446.

Van Dongen, G., W.L.M. Geilenkreirchen, J. van Rijn and R. van Wijk. 1986. Exp. Cell Res. 166:427-441.

Van Doren, K. and D.Hirsh. 1988. Nature 335:556-559.

Vierling, E., R.T. Nagao, A.E. DeRocher and C.M. Harris. 1988. EMBO J. 7:575-581.

Vijay-Kumar, S., C.E. Bugg, and W.J. Cook. 1987a. J. Mol. Biol. 194:531-544.

Vijay-Kumar, S., C.E. Bugg, K.D. Wilkinson, and W.J. Cook. 1985. Proc. Natl. Acad. Sci. USA 82:3582-3585.

Vijay-Kumar, S., C.E. Bugg, K.D. Wilkinson, R.D. Viestra, P.M. Hatfield, and W.J. Cook. 1987b. J. Biol. Chem. 262:6396-6399.

Vitek, M. P. and E.M. Berger. 1984. J.Mol. Biol.178:173-189.

Voellmy, R., M. Goldschmidt-Clermont, R. Southgate, A. Tissières, R. Levis and W. Gehring. 1981. Cell 23:261-270.

Welch, W.J. and J.P. Suhan. 1985. J. Cell Biol. 101:1198-1211.

Welch, W.J. and J.P. Suhan. 1986. J. Cell Biol. 103:2035-2053.

Welch, W.J. and J.R. Feramisco. 1984. J. Biol. Chem. 259:4501-4511.

Welch, W.J. and J.R. Feramisco. 1985. Mol. Cell. Biol. 5:1229-1237.

Welch, W.J. and L.A. Mizzen. 1988. J. Cell Biol. 106:1117-1130.

West, M.H.P. and W.M. Bonner. 1980a. Biochemistry 19:3238-3245.

West, M.H.P. and W.M. Bonner. 1980b. Nuc. Acids Res. 8:4671-4680.

Westwood, J.T., J. Clos and C.Wu. 1991. Nature 353:822-827.

White, J. G. and H.R. Horvitz. 1979. Laser microbeam techniques in biological research. Electro-Optical Systems Design AUG.

White, J.G., E. Southgate, J.N. Thomson and S. Brenner. 1986. Philos. Trans. R. Soc. Lond. B Biol. Sci. 314:1-340.

Wiborg, O., M.S. Pederson, A. Wind, L.E. Berglund, K.A. Marcker, and J. Vuust. 1985. EMBO J. 4:755-759.

Wickner, S., J. Hoskins and K. McKenney. 1991. Nature 350:165-167.

Wiederrecht, G., D. Seto and C.S. Parker. 1988. Cell 54:841-853.

Winning, R.S. and L.W. Browder. 1988. Dev. Biol. 128:111-120.

Wright-Sandor, L.G., M. Reichlin and S.L. Tobin. 1989. J. Cell Biol. 108:2007-2016.

Wu, B., and R.I. Morimoto. 1985. Proc. Natl. Acad. Sci. USA 82:6070-6074.

Wu, B., C. Hunt and R. Morimoto. 1985. Mol. Cell. Biol. 5:330-341.

Wu, C., S. Wilson, B. Walker, I. Dawid, T. Paisley, V. Zimarino and H. Ueda. 1987. Science 238:1247-1253.

Xiao, H. and J.T. Lis. 1988. Science 239:1139-1142.

Xiao, H., O. Perisic and J.T. Lis. 1991. Cell 64:585-593.

Yarden, Y., J.A. Escobedo, W-J. Kuang, T.L. Yang-Feng, T.O. Daniel, P.M. Tremble, E.Y. Chen, M.E. Ando, R.N. Harkins, U. Francke, V.A. Fried, A. Ullrich and L.T. Williams. 1986. Nature 323:226-232.

Yost, H.J. and S. Lindquist. 1986. Cell 45:185-193.

Yost, H.J., R.B. Petersen and S. Lindquist. 1990. Posttranscriptional regulation of hsps in "Stress Proteins in Biology and Medicine". Cold Spring Harbour Laboratory Press.

Zakeri, Z.F., W.J. Welch and D.J. Wolgemuth. 1990. J. Cell Biol. 111:1785-1792.

Ziemiecki, A., M.-G. Catelli, I. Joab and B. Moncharmont. 1986. Bioc. Biop. R. 138:1298-1307.

Zimarino, V. and C. Wu. 1987. Nature 327:727-730.

Zimmerman, J.L., W. Petri and M. Meselson. 1983. Cell 32:1161-1170.

APPENDIX

Table A. Transgenic strains produced in the present study. PC, laboratory strain designation; va, laboratory allele designation; Ex, extrachromosomal; In, integrated; OR, orange cap tank.

STRAIN	SELECTION	*lacZ* FUSION	GENOTYPE	LIQUID N_2
PC1	pPD10.41	ubqΔ903	vaEx1	2:118
PC2	pPD10.41	ubqΔ903	vaEx2	3:13
PC3	pPD10.41	pPC16.48-43	vaEx3	3:117
PC4	pPD10.41	ubqΔ827	vaEx4	1:189
PC5	pRF4	pPCZ1	vaEx5	5:111
PC6	pRF4	pPCZ1	vaEx6	5:114
PC7	pRF4	pPCZ1	vaEx7	5:115
PC8	pRF4	pPCZ1	vaEx8	5:130
PC9	pRF4	-	vaIn1	3:18
PC10	pRF4	pPCZ1	vaEx10	3:18
PC11	pRF4	pPCZ1	vaEx11	5:160
PC12	pRF4	pPCZ1	vaEx12	-
PC13	pRF4	pPC16.48-1	vaEx13	6:110
PC14	pRF4	pPC16.48-1	vaEx14	6:110
PC15	pRF4	pPC16.48-1	vaEx15	6:131
PC16	pPD10.41	pPC16.48-1	vaEx16	3:121
PC17	pPD10.41	ubqPvuIIΔ903	vaEx17	1:123
PC18	pPD10.41	ubq938	vaEx18	3:121
PC19	pPD10.41	pHS16.25	vaEx19	3:117
PC20	pPD10.41	pHS16.25	vaEx20	3:13
PC21	pPD10.41	ubq938	vaEx21	2:64
PC22	pPD10.41	ubq938	vaEx22	2:64
PC23	pPD10.41	pHS16.25	vaEx23	3:124
PC24	pPD10.41	ubq938	vaEx24	4:149
PC25	pPD10.41	pHS16.25	vaEx25	4:149
PC26	pPD10.41	ubqPvuIIΔ903	vaEx26	2:118
PC27	pPD10.41	pHS16.25	vaEx27	3:108
PC28	pPD10.41	pHS16.25	vaEx28	3:118
PC29	pPD10.41	pHS16.25	vaEx29	3:118
PC30	pPD10.41	pHS16.25	vaEx30	3:108
PC31	pRF4	pPC16.48-1	vaEx31	3:135
PC32	pRF4	pPC16.48-1	vaEx32	6:132
PC33	pRF4	pPC16.48-1	vaEx33	6:132
PC34	pRF4	pPC16.48-1	vaEx34	6:131
PC35	pRF4	ubq938NLS	vaEx35	3:125
PC36	pRF4	ubq938NLS	vaEx36	6:143 (OR)
PC37	pRF4	pPC16.48-51	vaEx37	4:107
PC38	pRF4	pPC16.48-51	vaEx38	6:173 (OR)
PC39	pRF4	pPC16.48-51	vaEx39	4:134
PC40	pRF4	pPC16.48-51	vaEx40	4:134

Table A (con't):

STRAIN	SELECTION	*lacZ* FUSION	GENOTYPE	LIQUID N_2
PC41	pRF4	pPC16.1-48XBAI	vaEX41	6:234 (OR)
PC42	pRF4	pPC16.41-51	vaEx42	6:173 (OR)
PC43	pRF4	pPC16.1-48	vaEx43	2:155
PC44	pRF4	pPC16.41-51	vaEx44	2:106
PC45	pRF4	pPC16.41-51	vaEx45	3:104
PC46	pRF4	pPC16.41-51	vaEx46	2:153
PC47	pRF4	pPCZ1	vaEx47	4:X8
PC48	pRF4	pPCZ1	vaEx48	2:153
PC49	pRF4	pPCZ1	vaEx49	2:109
PC50	pRF4	pPCZ1	vaEx50	5:143
PC51	pRF4	pPC16.1-48	vaEx51	5:149
PC52	pRF4	pPC16.1-48	vaEx52	1:28
PC53	pRF4	pPC16.1-48	vaEx53	1:28
PC54	pRF4	pPC16.1-48	vaEx54	1:184
PC55	pRF4	pPC16.1-48	vaEx55	1:184
PC56	pRF4	pDX16.31	vaEx56	6:108 (OR)
PC57	pRF4	pDX16.31	vaEx57	6:120 (OR)
PC58	pRF4	pDX16.31	vaEx58	6:26 (OR)
PC59	pRF4	pDX16.31	vaEx59	6:120 (OR)
PC60	pRF4	pDX16.31	vaEx60	6:186 (OR)
PC61	pRF4	pDX16.31	vaEx61	6:186 (OR)
PC62	pRF4	pPC16.1-48XBAI	vaEx62	6:234 (OR)
PC63	pRF4	pDX16.31	vaEx63	-
PC64	pRF4	pPC16.1-48XBAI	vaEx64	6:135 (OR)
PC65	pRF4	pDX16.31	vaEx65	6:112 (OR)
PC66	pRF4	pDX16.31	vaEx66	6:112 (OR)
PC67	pRF4	pPC16.1-48XBAI	vaEx67	6:135 (OR)
PC68	pRF4	pDX16.31	vaEx68	6:25 (OR)
PC69	pRF4	pPCZ1	vaIn2	6:228 (OR)
PC70	pRF4	pPCZ1	vaIn3	6:139 (OR)
PC71	pRF4	pPCZ1	vaIn4	6:139 (OR)