Acknowledgements

Firstly, I thank my supervisors Professor Bill Brammar and Dr. Jenny Varley for supervision and many useful discussions during the course of my work at Leicester. I also thank Dr. Sue Adams for all the help and advice she has given to me, and all the members of the Joint Lab., past and present, who have made my time in Leicester an enjoyable one.

I am much indebted to Dr. Bambos Kyriacou (Dept. of Genetics, Leicester) for helpful discussions and supplying D. melanogaster adults and larvae as well as allowing me to use his Drosophila culture facilities. I thank Professor M. Ashburner (Dept. of Genetics, Cambridge) for independently assigning the cytological position of the D. melanogaster bbot gene and Professor M. Akam (Dept. of Genetics, Cambridge) for discussions on in situ hybridisation results and providing the method for non-isotopic in situ hybridisation.

I am very grateful to Professor D. Glover and Dr. B. Fenton (Dept. of Biochemistry, Dundee) for supplying D. melanogaster cell lines and anti-cyclin B antibody and for helpful discussions on immunoprecipitation.

I thank Dr. A. Moir (Kreb's Institute, University of Sheffield) for production and coupling to carrier proteins of the synthetic peptides, Dr. V. Ohanian (Dept. of Biochemistry, Leicester) for supplying the pGEX-3X expression plasmid, Miss Annemarie Simpson (Dept. of Anatomy, Leicester) for Northern blots of RNA from rat tissues and Mr. S. Byrne (Dept. of Anatomy, Leicester) for helpful discussions on many topics.

Finally, I gratefully acknowledge the financial support of the Medical Research Council and Imperial Chemical Industries plc.
For My Parents
# TABLE OF CONTENTS

Chapter 1  Introduction....................................................................................1

1.1 Introductory Overview ......................................................................................1

1.2 The Metastatic Process.....................................................................................2
  1.2.1 Tumour Cell Attachment to the Extracellular Matrix .........................2
  1.2.2 Degradation of the Extracellular Matrix............................................2
  1.2.3 Tumour Cell Movement.......................................................................4
  1.2.4 Tumour Cell Dissemination.................................................................6
  1.2.5 Arrest of Circulating Tumour Cells.....................................................6

1.3 Metastasis as a Genetically Controlled Process ................................................9

1.4 Identification and Preliminary Characterisation of the C328-10 cDNA Clone ........................................................................................................1  4

1.5 Neoplasia In D. melanogaster............................................................................1  9

Chapter 2 Materials and Methods ..............................................................2 4

2.1 Materials.............................................................................................................2 4
  2.1.1 Chemicals .........................................................................................2 4
  2.1.2 Enzymes and Proteins ......................................................................2 4
  2.1.3 Bacterial Strains and Culture Conditions ........................................2 5
  2.1.4 Cloning Vectors Used .........................................................................25
  2.1.5 Oligonucleotides Used ........................................................................26

2.2 Methods of Sterilisation.....................................................................................2 6

2.3 Nucleic Acid Methods.........................................................................................2 7
  2.3.1 Solutions Used During the Handling of Nucleic Acids......................2 7
  2.3.2 Purification of Nucleic Acids...............................................................3 3
    Purification of Genomic DNA from Tissues ...................................3 3
    Purification of Total RNA from Tissues........................................3 4
    Purification of Bacteriophage λ DNA ..............................................3 5
    Large scale Purification of Plasmid DNA ........................................3 6
    Small scale Purification of Plasmid DNA ........................................3 6
    Small scale Purification of Bacteriophage M13 Single-Stranded DNA ...........................................................................3 6

  2.3.3 Spectrophotometric Determination of Nucleic Acid Concentrations .................................................................................3 7

  2.3.4 Gel Electrophoresis..............................................................................3 7
    Agarose Gel Electrophoresis ..............................................................3 7
    Polyacrylamide Gel Electrophoresis .................................................3 8

  2.3.5 Restriction Enzyme Digestions and Southern (DNA) Blotting............3 8

  2.3.6 Northern (RNA) Blotting ....................................................................3 9

  2.3.7 Hybridisation of DNA Probes to Membrane-Im mobilised Nucleic Acids ...........................................................................3 9
    DNA..................................................................................................3 9
    RNA...............................................................................................4 0

  2.3.8 Removal of Probes and Re-use of Blots..............................................4 0
    DNA..................................................................................................4 0
    RNA...............................................................................................4 0

  2.3.9 Preparation of 32P-Radiolabelled Probes ........................................4 1
    Random Priming Method .................................................................4 1
2.7.1 Media .................................................................67
2.7.2 Culture of D. melanogaster ............................................68
2.7.3 Collection of Eggs .....................................................68

Chapter 3 Characterisation of the D. melanogaster bbc1 cDNA and Predicted Protein Product ...............69

3.1 Introduction ................................................................69
3.2 Isolation and Characterisation of D. melanogaster bbc1 cDNAs .................................................................69
3.3 Analysis of the D. melanogaster bbc1 mRNA .................................................................71
  3.3.1 Northern Blotting Analysis .........................................................71
  3.3.2 Identification of the Polyadenylation Site Within the D. melanogaster bbc1 Transcript .......................74
  3.3.3 Attempted Cloning of the D. melanogaster bbc1 mRNA 5' End ..................................................75
  3.3.4 Primer-Extension Analysis of the D. melanogaster bbc1 mRNA ..........................................................76
3.4 Analysis of the D. melanogaster bbc1 cDNA .................................................................77
  3.4.1 Comparison of the Human and D. melanogaster bbc1 cDNAs ..........................................................77
  3.4.2 Comparison of the Human and D. melanogaster bbc1 cDNAs with Other Nucleic Acid Sequences ........79
3.5 The Predicted D. melanogaster bbc1 Protein .................................................................79
  3.5.1 Comparison with the Predicted Human bbc1 Protein .................................................................79
  3.5.2 Comparison with Other Proteins .................................................................................................81
3.6 Conclusions ......................................................................81

Chapter 4 Analysis of D. melanogaster and Human bbc1 Genomic DNA Sequences ................................83

4.1 Introduction ................................................................83
4.2 PCR Amplification of the D. melanogaster bbc1 Genomic ORF Region .........................................................83
4.3 PCR Amplification of Sequences Within the Human bbc1 Gene .................................................................84
4.4 Southern Blotting Analysis of Human Genomic DNA Using an Intron-Specific Probe ........................................84
4.5 Localisation of the Human bbc1 Gene to Chromosome 16 ........................................................................86
4.6 Analysis of the 3' Extension in Human cDNA Clone 10-12 ........................................................................88
4.7 Conclusions ......................................................................90

Chapter 5 In Situ Hybridisation Studies Performed in D. melanogaster .........................................................92

5.1 Introduction ................................................................92
5.2 In Situ Hybridisation to D. melanogaster Polytene Chromosomes .................................................................92
5.3 In Situ Hybridisation to D. melanogaster Embryo Sections ........................................................................94
5.4 In Situ Hybridisation to Whole Mount D. melanogaster Embryos .................................................................95
5.5 Conclusions .........................................................................................................................98

Chapter 6 Production of Antibodies to the O. mefanogaster bbcl Protein ..................................................100

6.1 Introduction .......................................................................................................................100

6.2 Construction of the pDMI Expression Plasmid .......................................................................100

6.3 Expression of the GST-bbcl Fusion Protein in E. coli ............................................................101

6.4 Purification of the GST-bbcl Fusion Protein ...........................................................................102

6.5 Raising and Characterisation of Polyclonal Rabbit Antibodies Against the Fusion Protein ...103

6.6 Production of Synthetic Peptides and Generation of Anti-peptide Antibodies .........................105

6.7 Conclusions .......................................................................................................................106

Chapter 7 Characterisation of the O. mefanogaster bbcl Protein .........................................................107

7.1 Introduction ........................................................................................................................107

7.2 Western Blotting Analysis of D. melanogaster Proteins .........................................................107

7.3 Temporal Expression of the bbcl Protein During Embryogenesis and Later Development ....108

7.4 Immunohistochemical Analysis of D. melanogaster Embryos Using the 317 Antiserum ..........110

7.5 Immunoprecipitation Analysis of D. melanogaster Proteins Using the 317 Antiserum ........110

7.6 Western Blotting Analysis of Human Proteins Using the 317 Antiserum ... 113

7.7 Conclusions .......................................................................................................................113

Chapter 8 Discussion and Future Work ............................................................................................115

8.1 Discussion ..........................................................................................................................115

8.2 Future Work .......................................................................................................................126

Literature Cited ............................................................................................................................129
Figures and Tables

Chapter 1 Introduction
Figures:
1.3.1 C328-10 expression levels in carcinomas and fibroadenomas
1.3.2 Southern blotting analysis of human genomic DNA
1.3.3 Nature blotting analysis
1.3.4 Northern blotting analysis of cell line RNAs
1.3.5 cDNA clones isolated by Dr. J. M. Varley
1.3.6 Overlapping bbd cDNA clones

Chapter 2 Materials and Methods
Figures:
2.3.1 Oligonucleotide positions

Tables:
2.3.1 Bacterial strains
2.2.1 Oligonucleotide sequences
2.6.1 Cell lines used

Chapter 3 Characterisation of the D. melanogaster bbd cDNA and Predicted Protein Product
Figures:
3.1.1 Southern blotting of D. melanogaster genomic DNA
3.2.1 D. melanogaster bbd clones isolated by Dr. J. M. Varley
3.3.1 Northern blotting of D. melanogaster and human RNA
3.3.2 D. melanogaster developmental Northern bbd
3.3.3 Analysis of the bbd mRNA 3' end using RACE
3.3.4 Attempted cloning of the bbd mRNA 5' end
3.3.5 Primer-extension analysis of D. melanogaster bbd mRNA
3.4.1 Comparison of D. melanogaster and human bbd cDNAs
3.4.2 Northern blotting of rat tissue RNAs
3.5.1 Comparison of D. melanogaster and human bbd proteins
3.5.2 Peptide analysis of D. melanogaster and human bbd proteins
3.5.3 Comparison of bbd proteins with other proteins

Tables:
3.2.1 D. melanogaster bbd cDNA sizes
3.3.1 Densitometric Analysis of the Autoradiographs Shown in Figure 3.3.2
3.5.1 Properties of predicted D. melanogaster and human proteins

Chapter 4 Analysis of D. melanogaster and Human bbd Genomic DNA Sequences
Figures:
4.2.1 D. melanogaster bbd introns - positions and sequences
4.3.1 Human bbd intron - identification and sequence
4.3.2 Relationship between intron positions and protein sequences
4.4.1 Southern blotting of human genomic DNA - intron probe
4.5.1 PCR from somatic cell hybrid DNAs (intron primers)
4.6.1 Northern blotting of human RNA - 10-12 3' extension probe
4.6.2 PCR amplification of genomic 10-12 3' extension sequences
4.6.3 PCR from somatic cell hybrid DNAs (10-12 extension primers)
Tables:
4.5.1 Assignment of human bbcl gene to chromosome 16
4.6.1 Sizes of PCR products from amplification of genomic 10-12
3' extension sequences

Chapter 5 In Situ Hybridisation Studies Performed in D. melanogaster

Figures:
5.2.1 In situ hybridisation to polytene chromosomes
5.3.1 In situ hybridisation to D. melanogaster embryo sections
5.4.1 In situ hybridisation to whole D. melanogaster embryos

Chapter 6 Production of Antibodies to the D. melanogaster bbcl Protein

Figures:
6.2.1 Cloning the D. melanogaster bbcl ORF into pGEX-3X
6.3.1 Generation of the pDM1* expression construct
6.3.2 Expression and purification of the GST-bbcl fusion protein
6.4.1 SDS-PAGE analysis of the GST and GST-bbcl fusion proteins
6.5.1 Transfer of protein to nitrocellulose and detection with anti-bbcl antiserum
6.5.2 Preabsorption of antisera
6.6.1 Positions of synthetic peptides
6.6.2 Dot blot probed with anti-peptide antibodies

Chapter 7 Characterisation of the D. melanogaster bbcl Protein

Figures:
7.2.1 Western blotting of D. melanogaster proteins
7.3.1 Expression of bbcl protein during D. melanogaster development
7.5.1 Western blotting of proteins from D. melanogaster SL2 cells
7.6.1 Western blotting analysis of human proteins

Chapter 8 Discussion and Future Work

Figures:
8.1.1 Comparison of the D. melanogaster bbcl and Sea Urchin Histone H1 Protein Sequences
### Abbreviations Used

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>αα</td>
<td>Amino acids*</td>
</tr>
<tr>
<td>AMV RVT</td>
<td>Avian myeloblastosis virus reverse transcriptase</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>bp</td>
<td>Base pair</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>°C</td>
<td>Degrees Centigrade</td>
</tr>
<tr>
<td>CAPSO</td>
<td>3-[(Cyclohexylamino)-2-hydroxy-1-propane sulphonatic acid</td>
</tr>
<tr>
<td>CEPH</td>
<td>Centre d'Etude du Polymorphisme Humain, 27, rue Juliette Dodu, 75010 Paris, France</td>
</tr>
<tr>
<td>Da</td>
<td>Dalton</td>
</tr>
<tr>
<td>DAB</td>
<td>3,4,3',4'-Tetra-aminobiphenyl hydrochloride</td>
</tr>
<tr>
<td>dATP</td>
<td>Deoxyadenosine triphosphate</td>
</tr>
<tr>
<td>dCTP</td>
<td>Deoxycytidine triphosphate</td>
</tr>
<tr>
<td>ddATP</td>
<td>Deoxyadenosine triphosphate</td>
</tr>
<tr>
<td>ddCTP</td>
<td>Deoxycytidine triphosphate</td>
</tr>
<tr>
<td>ddGTP</td>
<td>Deoxycytidine triphosphate</td>
</tr>
<tr>
<td>ddTTP</td>
<td>Deoxythymidine triphosphate</td>
</tr>
<tr>
<td>DEPC</td>
<td>Diethyl pyrocarbonate</td>
</tr>
<tr>
<td>dGTP</td>
<td>Deoxyguanosine triphosphate</td>
</tr>
<tr>
<td>DIG</td>
<td>Digoxigenin</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco's modified Eagle medium</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethylsulfoxide</td>
</tr>
<tr>
<td>DNase</td>
<td>Deoxyribonuclease</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxynucleoside triphosphate</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>dTTP</td>
<td>Deoxythymidine triphosphate</td>
</tr>
<tr>
<td>EDTA</td>
<td>Diaminoethanetetra-acetic acid, disodium salt</td>
</tr>
<tr>
<td>EtBr</td>
<td>Ethidium bromide</td>
</tr>
<tr>
<td>FCS</td>
<td>Foetal calf serum</td>
</tr>
<tr>
<td>g</td>
<td>Grammes</td>
</tr>
<tr>
<td>GST</td>
<td>Glutathione S-transferase (from Schistosoma japonicum)</td>
</tr>
<tr>
<td>HEPES</td>
<td>N-2-Hydroxyethylpiperazine-N'-2-ethane sulphonatic acid</td>
</tr>
<tr>
<td>h</td>
<td>Hours</td>
</tr>
<tr>
<td>HRP</td>
<td>Horseradish peroxidase</td>
</tr>
<tr>
<td>HPLC</td>
<td>High performance liquid chromatography</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl-β-D-thiogalactopyranoside</td>
</tr>
<tr>
<td>Kg</td>
<td>Kilogrammes</td>
</tr>
<tr>
<td>KDa</td>
<td>Kilodalton</td>
</tr>
<tr>
<td>L</td>
<td>Litre</td>
</tr>
<tr>
<td>M</td>
<td>Molar</td>
</tr>
<tr>
<td>μg</td>
<td>Microgrammes</td>
</tr>
<tr>
<td>ng</td>
<td>Milligrammes</td>
</tr>
<tr>
<td>μl</td>
<td>Microlitres</td>
</tr>
<tr>
<td>m</td>
<td>Millilitres</td>
</tr>
<tr>
<td>mM</td>
<td>Millimolar</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger ribonucleic acid</td>
</tr>
<tr>
<td>MOPS</td>
<td>3-(N-Morpholino)propane sulphonatic acid</td>
</tr>
<tr>
<td>NBT</td>
<td>Nitroblue tetrazolium salt</td>
</tr>
<tr>
<td>ng</td>
<td>Nanogrammes</td>
</tr>
<tr>
<td>nt</td>
<td>Nucleotides</td>
</tr>
<tr>
<td>OD</td>
<td>Optical density</td>
</tr>
<tr>
<td>ORF</td>
<td>Open reading frame</td>
</tr>
</tbody>
</table>
PAGE  Polyacrylamide gel electrophoresis
PBS  Phosphate buffered saline
PCR  Polymerase chain reaction
PEG  Polyethylene glycol
pI  Isoelectric point
PIPES  Piperazine-N,N’-bis[2-ethane sulphonic acid]
PMSF  Phenylmethylsulphonyl fluoride
Q  Millipore Super-Q® system treated water
RNA  Ribonucleic acid
RNase  Ribonuclease
rpm  Revolutions per minute
SDS  Sodium dodecyl sulphate
SDS/PAGE  Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SSC  Standard saline citrate
TAE  Tris/acetate/EDTA buffer
TBE  Tris/borate/EDTA buffer
TdT  Terminal deoxynucleotidyl transferase
tRNA  Transfer RNA
U  Unit (of enzymic activity)
UTP  Uridine triphosphate
V  Volt
(v/v)  (volume/volume)
(w/v)  (weight/volume)
X-gal  5-bromo-4-chloro-3-indolyl-β-D-galactoside
X-phosphate  5-bromo-4-chloro-3-indolyl phosphate

*See Stryer (1981), table 2-1 for a list of single letter amino acid codes.
CHAPTER 1

Introduction

1.1 Introductory Overview

The study of cancer metastasis is of profound clinical importance, because the majority of deaths from cancer result not from the primary neoplasm, which can often be resected, but from secondary tumour burden in vital organs (Fidler, 1991). The high frequency of breast cancer amongst the female population of developed countries (CRC, 1990) and the poor survival rate (Brinkley and Haybittle, 1975; Zajicek, 1987) which is primarily due to metastatic spread, means that the study of breast cancer metastasis has major implications for female health.

Due to the severe impact metastatic spread of tumours has on patient survival and because of a general lack of knowledge regarding the molecular aspects of the process, a project was initiated in the ICI/JOINT laboratory to investigate alterations in gene expression between benign (non-metastatic) and malignant (metastatic) breast tumours. The results of this analysis have been detailed elsewhere (Sharp et al., 1990; Adams et al., 1992b; Sharp et al., 1992) and the reader is referred to these publications. The first section of this chapter provides a brief account of the steps involved in metastasis, with the emphasis on genes which are likely to have important roles in the dissemination of tumour cells and which one would expect to identify in a screen designed to isolate genes showing differences in expression between benign and malignant tumours. For a rigorous explanation of the metastatic process, the reader is referred to several recent reviews on this subject (Hart et al., 1989; Miller and Heppner, 1990; Price, 1990; Fidler, 1991; Liotta and Stetler-Stevenson, 1991). A description of the initial characterisation of one particular clone isolated during the screen forms the second part of the introduction.
1.2 The Metastatic Process

For a tumour cell to complete successfully the metastatic process, it must overcome many hurdles. These so called "decathlon winners" (Nicolson, 1982) are rare in the tumour cell population, thus accounting for the inefficiency of the metastatic process; less than 0.01% of circulating tumour cells form metastatic colonies (Liotta and Stetler-Stevenson, 1991). The first event in metastatic spread is migration of the tumour cell out of the normal tissue compartment in which it resides. A three-step theory of invasion involving extracellular matrix attachment and dissolution followed by tumour cell migration was first proposed by Liotta (Liotta, 1986), and is now widely accepted.

1.2.1 Tumour Cell Attachment to the Extracellular Matrix

Tumour cell attachment is mediated by cell-surface receptors of the integrin and non-integrin families which bind to matrix glycoproteins such as fibronectin and vitronectin as well as to proteins of the basement membrane (Ruoslahti and Giancotti, 1989; Albelda and Buck, 1990; Kramer et al., 1991; Rousslahti, 1991). These receptors not only bind extracellular matrix components, they also transmit signals to the cell interior (Pignatelli and Bodmer, 1988; Aznavoorian et al., 1990; Yurochko et al., 1992) and may constitute an important regulatory system in tumour spread. Indeed, addition to tumour cells of synthetic peptides that disrupt integrin-ligand interactions has been shown to inhibit not only cell binding to matrix components in vitro (Gehlsen et al., 1988; Saiki et al., 1989), but also tumour formation in several in vivo experimental systems (Humphries et al., 1986; Iwamoto et al., 1987; Saiki et al., 1989; Saiki et al., 1990).

1.2.2 Degradation of the extracellular matrix

The extracellular matrix contains many different types of molecules, such as collagen types I, II and III, elastin, proteoglycans and glycosaminoglycans (for review see Zucker, 1988). For this reason, it is not surprising that degradation of many extracellular matrix
components by a wide range of tumour- and host cell-elaborated enzymes has been reported (Matrisian et al., 1986, Liotta et al., 1981; Tryggvason et al., 1987; Biziozzi et al., 1988; Zucker, 1988; Basset et al., 1990; Liotta and Stetler-Stevenson, 1990; Mackay et al., 1990; McDonnell and Matrisian, 1990; Rochefort et al., 1990; Sloane et al., 1990). However, because collagens are by far the most common constituents of the extracellular matrix (Burgeson, 1988), enzymes capable of degrading these proteins have attracted a great deal of attention. Interstitial collagens can be hydrolysed by many proteinases including plasmin (Liotta et al., 1981), stromelysin (Chin et al., 1985) and elastase (Bartholomew et al., 1984) after initial cleavage of the molecule by a specific collagenase (Smyth Templeton et al., 1990).

Type IV collagen forms a substantial part of the basement membrane and as such represents a major barrier to tumour cell movement. An enzyme (type IV collagenase) secreted by tumour cells has been identified which specifically cleaves this molecule (Salo et al., 1983). Differences in expression level of this protease between tumour cell populations have been correlated with metastatic potential in a number of systems (Liotta et al., 1980; Salo et al., 1982; Turpeenniemi-Hujanen et al., 1985; Nakajima et al., 1987; Höyhtya et al., 1990).

The matrix metalloproteinases, collagenases, stromelysin/transin and PUMP-1 (Liotta and Stetler-Stevenson, 1990), are secreted from cells in inactive proenzyme forms. Conversion from an inactive to an active state is believed to involve a "cysteine switch" which can be activated in vitro by a number of mechanisms (Van Wart and Birkedal-Hansen, 1990). The in vivo method of activation is not known, but may involve proteolysis by a range of proteases including plasmin (Mignatti et al., 1986). Activation by plasmin could result in a proteolytic cascade mechanism whereby plasminogen activator, which is known to be expressed by metastatic tumour cells (Salo et al., 1982; Hearing et al., 1988; Duffy et al., 1990), activates latent metalloproteinases via plasmin (Mignatti et al., 1986). Plasmin may also participate directly in collagen hydrolysis (Mackay et al., 1990).
Apart from tumour cell-derived proteases, enzymes originating from infiltrating host cells may contribute to the overall proteolytic content of a tumour. Mast cells, fibroblasts, endothelial cells and macrophages have all been reported to possess collagenolytic activities (see Pauli et al., 1983). Indeed, the recently isolated protease stromelysin 3 was shown to be produced by host stromal cells surrounding breast tumours (Basset et al., 1990).

Organisms have developed highly specific mechanisms of preventing unwanted and inappropriate proteolysis (for review see Laskowski and Kato, 1980). From the point of view of the present discussion, the tissue inhibitors of metalloproteinases (TIMPs) and plasminogen-activator inhibitor type 1 (PAI-1) are of particular interest. TIMPs are ubiquitously expressed and have been shown to inhibit tumour cell invasion in vitro and tumour formation in vivo (Schultz et al., 1988; Albini et al., 1991). Down-modulation of TIMP mRNA in mouse 3T3 cells resulted in these cells becoming invasive in the human amnion assay and tumourigenic and fully metastatic in athymic nude mice (Khokha et al., 1989). Conversely, upregulation of TIMP-1 protein in B16-F10 cells reduced the in vitro invasive capacity of these cells (Khokha et al., 1992). The TIMP gene is therefore a good candidate for a metastasis suppressor gene. PAI-1 has also been shown to inhibit strongly extracellular matrix hydrolysis by certain tumour cells (Cajot et al., 1990). The presence of these inhibitors probably confines proteolytic activity to sites in close proximity to the tumour cell membrane where protease concentrations are high enough to facilitate proteolysis (Pöllänen et al., 1988; Moll et al., 1990).

1.2.3 Tumour Cell Movement

The third aspect of invasion, namely movement of tumour cells, is postulated to occur by one of two means: expansive movement associated with tumour growth (Eaves, 1973) and active cell migration. Tumour cell migration has been observed both in vivo (Haemmerli and Sträuli, 1978) and in vitro (Schor et al., 1980). Cell migration over a solid substratum requires cell attachment (see above), force generation and detachment. Force-generation is mediated by interactions between cytoskeletal components (Singer...
and Kupfer, 1986) and is transmitted to the cell surface by actin filaments, microtubules and intermediate filaments attached to integrins at adhesion plaques within the membrane (Burridge et al., 1988). The mechanisms governing cell detachment are not well understood. Proposed mechanisms include localised proteolysis (Chen and Chen, 1987; Pöllänen et al., 1988), accumulation of proteoglycans which inhibit integrin-ligand interactions (Ruoslahti and Pierschbacher, 1987) and modulation of integrin-ligand interactions by such mechanisms as phosphorylation (Tapley et al., 1989).

When discussing tumour cell movement, it is pertinent to mention chemotaxis (directed cell movement in response to a chemical gradient) and haptotaxis (undirected cell movement in response to a chemical stimulus). Factors have been purified from a number of tumour cells which stimulate in vitro and in vivo movement of other tumour cells. For example, Liotta et al. (1986) purified an autocrine motility factor (AMF) from A2058 melanoma cells that stimulated random or directed movement of the producer cell. Recently a cell-surface protein has been identified on B16 melanoma cells which may be a receptor for AMF (Nabi et al., 1990). Apart from tumour-derived motility factors, several agents have been isolated from host cells which are capable of stimulating cell movement. A "scatter" factor (Stoker et al., 1987) produced by embryonic fibroblasts stimulates movement of epithelial cells, whilst thrombospondin, which is released by activated platelets, stimulates haptotactic migration of melanoma cells (Taraboletti et al., 1987). In addition to promoting tumour cell migration, these factors may cause infiltration of the tumour by host cells such as fibroblasts, endothelial cells and macrophages. Such cells could affect tumour cell movement by production of lytic enzymes and motility factors of their own (see above).

1.2.4 Tumour Cell Dissemination

Several mechanisms exist whereby tumour cells can migrate away from the primary tumour mass, including movement through body cavities and along nerve fibres. However, the predominant means of dissemination for cells from most solid tumours is via the lymphatic
and circulatory systems (Hart and Fidler, 1980b; Weidner et al., 1991). In the lymphatic system, tumour cells may become trapped in regional lymph nodes or may "skip" these nodes and spread to distant sites (Rosen et al., 1983). Regional lymph nodes may also act as "generalising sites" in which tumour cells multiply before dispersing to distant sites through the lymphatics and blood system (Price, 1990).

Tumour cells can enter the circulatory system directly (i.e. not via the lymphatics), often through capillaries or thin-walled venules within the tumour (Miller and Heppner, 1990). Since angiogenesis is a prerequisite for the growth of any solid tumour to a size of more than a few cubic millimetres (Folkman, 1990; Fidler, 1991) many sites for tumour intravasation exist within malignant neoplasms. Indeed, a correlation has been found between the extent of angiogenesis and the number of distant metastases in breast cancer (Weidner et al., 1991). However, most tumour cells that enter the circulation rapidly die (Fidler, 1970). This is probably a result of mechanical stresses (Weiss et al., 1992), the effects of which can be decreased by aggregation of tumour cells with other tumour cells (homotypic) or with various host cells such as platelets and lymphocytes (heterotypic). Such aggregation has been shown to occur and to increase tumour cell survival in the circulation (Fidler and Bucana, 1977; Updyke and Nicolson, 1986). Formation of aggregates may also assist arrest of tumour cells in the microvasculature by physical entrapment.

1.2.5 Arrest of Circulating Tumour Cells

Tumour cells entering the vasculature typically arrest temporarily at the first capillary bed encountered (Weiss, 1980; Murphy et al., 1986). The cells may remain at this location, or, after a short period of time, may detach and recirculate. Arrest is believed to be mediated partly by physical entrapment (Fidler, 1973) and partly by specific attachment (Alby and Auerbach, 1984; Auerbach et al., 1987; Kahan, 1987). As far as the latter is concerned, it has been known for over a hundred years that cancer metastasis is not a random process (Paget, 1889). For example, breast cancer commonly metastasises to
the bone marrow (Lee, 1985), despite the fact that this is not the first organ encountered by tumour cells disseminating by a haematogenous route from a breast tumour. Studies using endothelial cell monolayers derived from specific organs show that some tumour cell types bind preferentially to endothelial cells from organs which are natural in vivo targets for these cells (Auerbach et al., 1987).

Similar results have been obtained using aortic endothelial cells cultured on organ-specific matrix components (Pauli and Lee, 1988). Likewise, experiments using explanted tissues show that tumours only develop in organs which are natural targets for tumour spread (Hart and Fidler, 1980a; Juacaba et al., 1989).

These results indicate that specific tumour cell-endothelial cell interactions are likely to be of considerable importance in determining tumour cell arrest patterns. However, little is known about the precise nature of tumour cell or endothelial cell membrane components involved in specific tumour cell arrest. An insight into the molecular processes governing tumour cell arrest may well be gained from a related phenomenon, lymphocyte homing. Exit of lymphocytes from venules and entrance into peripheral lymph nodes is an ordered event (Taylor Sher et al., 1988). The lymphocyte receptor involved in this process is a calcium-dependent lectin, L-selectin (Siegelman et al., 1986; St. John et al., 1986). The ligand bound by the lymphocyte receptor has also been identified and the gene encoding it cloned (Streeter et al., 1988). This molecule, called GlyCAM-1, is selectively expressed on nodal endothelial cells and is a member of a family of proteins called "vascular addressins" (Streeter et al., 1988).

Other molecules implicated in similar types of lymphocyte homing have been identified (for review see Butcher, 1991). For example, endothelial cell leukocyte adhesion molecule-1 (ELAM-1) has been shown to play a role in skin-homing of T cells (Picker et al., 1991). This receptor also specifically binds human colon carcinoma cells and may be involved in arrest of these cells in specific organs expressing the ELAM-1 receptor (Lauri et al., 1991). The general applicability of an addressin-type system to organ-specific tumour cell arrest has yet to be shown.
So-called "organ-specific" tumour cells have been known for many years (Brunson et al., 1978; Nicolson et al., 1978). The brain-colonising B16 variant isolated by Brunsen et al., (1978) has been shown to possess a specific 90-KDa cell-surface glycoprotein that binds transferrin (Nicolson et al., 1990). The authors propose that the protein may allow the cell to respond to low levels of growth factors present within brain and other tissues. B16 melanoma cell lines with enhanced lung-colonising ability have also been isolated and shown to express elevated (relative to parental cells) levels of the integrin α1β3 (Chang et al., 1992). These cells also displayed increased adhesion to fibronectin and subendothelial matrix, which may partly explain the elevated lung-colonising ability of these cells.

Other mechanisms probably govern tissue-specific metastasis, such as the ability of the tumour cell to penetrate the organ parenchyma and elicit an angiogenic response. Connective tissues are known to produce protease inhibitors which can block enzyme-mediated invasion (Thorgerisson et al., 1982). The presence of high concentrations of such inhibitors in other organs could also inhibit tumour cell extravasation and consequent growth. Moses et al. (1990) have identified an inhibitor of neovascularisation from cartilage which may explain why this tissue is almost never a site for tumour metastasis.

Once arrested in capillaries, tumour cells rapidly extravasate by causing endothelial cell retraction (Kramer and Nicolson, 1979) and dissolution of the underlying basement membrane by a process analogous to that described above for tumour cell invasion. Having gained access to the organ parenchyma, tumour cells are then able to grow by subverting host processes such as neovascularisation in a manner identical to that which occurred in the original tumour site. Once a secondary tumour is established, cells may disseminate from this tumour to form tertiary tumours; the so-called metastasis of metastases (Fidler, 1991).

The spread of tumour cells throughout the body is probably the most feared aspect of neoplasia and certainly the one that causes the most...
therapeutic failures. The mechanisms by which these "decathlon winners", as Nicolson (1982) has dubbed them, complete all the stages of the metastatic process is of major interest clinically and pharmaceutically. The study of this process and also the genetic changes leading to malignancy may not only lead to a better understanding of the events of metastasis, but also to better treatment and possibly a cure.

1.3 Metastasis as a Genetically Controlled Process

From the above brief account of the major steps in metastasis, it is clear that many gene-products are involved or implicated in the process. However, at present little is known about the genetic control of tumour cell dissemination. Tumours are composed of heterogeneous populations of cells which differ with respect to each other in many ways, including growth rate, antigenic and immunogenic status, response to cytotoxic agents, invasiveness and metastatic potential (Fidler and Radinsky, 1990). Since each of these attributes is itself controlled by many genes, one can see that tumours are heterogeneous both at the phenotypic and genotypic levels. This heterogeneity is generally believed to be the result of genomic instability within the tumour cells which, together with host selection pressures, leads to an accumulation of genetic changes resulting in tumour progression and, in some cases, acquisition of the metastatic phenotype. This process of acquired genetic instability was first proposed by Nowell (1976) and predicts that progression towards malignancy is accompanied by increasing genetic instability within the progressing cells. Several groups have investigated this theory by examining metastatic stability and mutation rates at selected loci of paired metastatic and low/non-metastatic cloned lines isolated from rodent tumours (Cifone and Fidler 1982; Harris et al., 1982; Kaden et al., 1989). The results of these studies show that metastatic cells do acquire mutations more rapidly than their non-metastatic counterparts, as judged by, for example, acquired drug resistance. Also, the metastatic cells show high frequencies of spontaneous deletions of DNA.
The rate of occurrence of classical spontaneous mutations in tumour cells is too low to account for the rapid changes that can take place during tumour progression (Hill, 1990). Instead, events such as gene amplification and deletion have been implicated in tumour progression. Certain proto-oncogenes have been shown to become amplified in human tumours. These include, amongst others, HER2 (c-erbB-2) (King et al., 1985; Varley et al., 1987; Liu et al., 1992), c-myc (Alitalo et al., 1983; Little et al., 1983; Varley et al., 1987) and MDM2 (Oliner et al., 1992). Amplification of the c-myc gene has been correlated with therapeutic failure resulting from metastatic spread of the primary tumour (Varley et al., 1987).

As mentioned above, deletion of DNA segments is also commonly found in tumours. According to the theory proposed by Knudson (1971), these loci contain genes which inhibit tumourigenesis/ malignancy. Early studies using cell fusions between tumourigenic and non-tumourigenic cell lines showed that most resulting fusions were non-tumourigenic and that tumourigenic hybrids consistently lost all or part of certain chromosomes (for review see Levine and Momand, 1990). Similar results were found for hybrids between metastatic and non-metastatic cells; non-metastatic hybrids resulted, except when certain chromosomes were lost (Stanbridge, 1976).

Subsequent work has identified several genes which inhibit tumourigenesis and, in some cases, malignant progression (see below). These tumour suppressor genes include the human retinoblastoma susceptibility locus (RB1) at 13q14 (Lee et al., 1987), Wilms' tumour gene (WT1) at 11p13 (Rose et al., 1990), DCC gene at 18q21 (Fearon et al., 1990), Neurofibromatosis type 1 (NF1) gene at 17q11.2 (Cawthon et al., 1990) and the p53 gene at 17p13 (Baker et al., 1989). Both RB1 and p53 proteins are implicated in control of the cell cycle and their loss can result in uninhibited cell growth and tumourigenesis (Hollingsworth et al., 1993; Perry and Levine, 1993 and references therein). However, loss of these proteins is also implicated in tumour progression, since they are believed to act as check-points in the cell cycle, preventing the cell from replicating its DNA until any damaged DNA has been repaired (DeCaprio et al.,
1989; Lane, 1992). Loss of these check-point controls could lead to replication of damaged DNA, resulting in aneuploidy, mutation and genomic instability, leading to tumour progression and malignancy. Interestingly, the product of the MDM2 gene, which is amplified in some tumours (see above), binds to the p53 protein and appears to prevent it from performing its check-point duties (Oliner et al., 1992). Hence, inactivation of a tumour suppressor gene or over-expression (as a result of gene amplification) of a dominantly acting oncogene can result in an identical result, namely tumourigenesis and malignant progression. The tumour suppressor gene WT1 encodes a zinc finger-containing transcription factor (Little et al., 1992) which may either be a suppressor of mesodermal stem cell proliferation and/or a differentiation signal in cells expressing the WT1 protein (Hastie, 1993). The NF1 protein has significant similarity to GTPase activating proteins (GAPs), which stimulate hydrolysis of GTP bound to G-proteins resulting in production of the inactive GDP-bound form of G-protein (Buchberg et al., 1990; Xu et al., 1990). Hence, the NF1 protein may be able to exert negative control over ras or ras-related protein(s), resulting in decreased response to stimuli acting through such G-proteins. The DCC protein is also interesting, in that it has properties indicating that it may function as a cell adhesion molecule. The importance of cell adhesion in controlling tumour growth and metastasis (see above), means that loss of this protein could be vital for colorectal cancer progression.

Given the complexities of metastasis, the identification and study of genes involved in this process has not been a simple task. Several methods have been used, including transfection studies and differential/subtractive screening of cDNA libraries. The former method relies on the existence of cloned genes, the products of which are thought to be involved in metastasis. Transfection of these genes, either transiently or stably, into suitable recipient cells followed by comparison of the metastatic capabilities of the transfected and parent cells allows an assessment of the affect on metastatic potential of a certain gene. This method is not without its drawbacks. Most importantly, a cloned gene must be available, thus precluding identification of novel genes. As already mentioned, a suitable cell line must be available for the transfection: one in
which addition of only one gene-product will sufficiently alter the metastatic potential such that this change can be observed. Also, an assay of metastatic potential must be used. This may either be an in *vitro* assay, such as assessing the mobility and invasive potential of cells, or secretion of enzymes such as proteases, or an in *vivo* assay of metastatic potential such as experimental of spontaneous metastasis.

Despite these problems, several genes have been studied by the method of transfection. One of the first genes to be studied by this method was the activated c-Ha-ras I gene. Muschel and colleagues (Muschel et al., 1985) found that this gene resulted in full metastatic potential when transfected into NIH-3T3 cells which were subsequently injected subcutaneously into nude mice. However, in the same study, transfection of the activated ras gene into C127 cells resulted in tumourigenic, but not metastatic, cells. Hence, as mentioned above, the recipient cell can have profound affects on the outcome of the experiment. Slightly later studies by the same group (Pozzatti et al., 1986) showed that primary rat embryo cells transformed by c-Ha-ras I alone were metastatic in nude mice, whereas those transformed by c-Ha-ras I and the E1a gene from type 2 adenovirus were non-metastatic. Hence the genetic composition of cells expressing the gene under investigation is a major component in determining the outcome of this sort of experiment.

As stated above, cells transformed with the ras and E1a gene were non-metastatic. Could the E1a gene therefore be a metastasis suppressor gene? Cook et al. (1986) reported that expression of E1a in hamster cells resulted in increased cytolytic susceptibility and Garbisa et al. (1987) showed that suppression of malignancy in the ras/E1a co-transfected rat embryo fibroblast system described above was associated with loss of type IV collagenolytic activity. Thus, in certain cell types, expression of this gene does seem to confer phenotypic attributes which would be predicted to result in a lower metastatic potential. Other genes shown to reduce metastatic potential when overexpressed in experimental systems include the nm23 gene (see below; Leone et al., 1991a) and the TIMP-1 gene (see
The second approach to identifying genes involved in metastasis is differential/subtractive cDNA hybridisation (Hart and Easty, 1991). These techniques rely on having two cell populations, either cell lines or tissue samples, which differ in their metastatic potentials. cDNAs are then identified by sole virtue of their differing abundance between the two cell populations; no assumptions being made about the nature of the gene from which they were derived. These techniques also lend themselves to the identification of novel genes; something not possible using transfection studies. A number of genes showing differential expression between metastatic and non-metastatic cell populations have been identified using these techniques, including nm23 (Steeg et al., 1988), WDNM1 (Dear et al., 1988) WDNM2 (Dear et al., 1989), stromelysin 3 (Basset et al., 1990), matrix Gla protein (Chen et al., 1990), and fibronectin (Schalken et al., 1988).

These genes were isolated during screens using cell lines or fresh tissue samples and do seem to show genuine differences in expression levels between metastatic and non-metastatic cells/tissues. However, the techniques of differential and subtractive hybridisation do have their disadvantages. Firstly, substantial amounts of undegraded mRNA must be available. Whilst this is not much of a problem when cell lines are used, it does present an obstacle when fresh tissue samples are used, both in terms of absolute amount and quality of the RNA available. A limitation facing both techniques is sensitivity. Differential screening can detect transcripts present at >0.01% of the mRNA population, whereas subtractive hybridisation is not much better, being able to detect transcripts constituting >0.001% of the mRNA pool. The use of cell lines circumvents another problem encountered when fresh tissue is used, namely stromal contamination. This can be kept to a minimum by using tissue samples with high epithelial content, but cannot overcome another problem that may lead to false positive clones: tumour cell heterogeneity. As detailed above, tumours are composed of heterogeneous populations of cells, each
with a unique gene expression profile, only one of which may have the capacity to metastasise. This decreases further the sensitivity of the screening method, since only a proportion of the cells within a tumour may be expressing a particular gene required for metastasis. Another potential problem with the use of fresh tissue samples is patient-specific gene expression patterns. To address this problem, multiple rounds of screening using total cDNA probes prepared from different patients must be performed. This approach also removes false positives arising from differences in growth rates of the tissue samples analysed. Despite all of the problems mentioned above, the use of fresh tissue samples does have one major advantage over using cell lines. Cell lines passaged in vitro for extensive time periods have the ability to acquire many changes to their expression profiles which are not related to metastasis. The above discussion points out that no approach to the isolation and/or study of genes potentially involved in the metastatic process is without its drawbacks and that careful and extensive analysis is required to ascertain whether or not genes are actually involved in metastasis.

1.4 Identification and Preliminary Characterisation of the C328-10 cDNA Clone

The experimental approach adopted in the ICI/Joint laboratory to identify cDNA sequences showing differential expression in benign (non-metastatic) and malignant (metastatic) breast tumours has been detailed elsewhere (Sharp et al., 1990; Adams et al., 1992; Sharp et al., 1992), but is also outlined below. To study gene expression patterns associated with the metastatic state, either tissue culture or biopsy-derived RNA samples can be used. Due to the possibility of gene expression being altered during in vitro cell growth (see above), fresh biopsy material was used for the study.

A shortage of normal breast tissue prevented comparison of gene expression in normal cells with that in malignant breast cells. Since normal tissue could not be used, it was decided that a comparison of gene expression in abnormal tissues would be made. The use of carcinoma-lymph node pairs in such a study has the potential to
identify genes involved in the final stages of the metastatic process i.e. growth of the tumour in the target tissue. However, tissue-specific gene expression patterns between breast and lymphoid tissues could lead to a high level of background positives, thus complicating the study.

Due to the problems associated with the above comparisons, a study of gene expression between benign and malignant breast tumours was undertaken. Such a study would be expected to identify genes involved in the early stages of malignancy, such as those encoding proteolytic enzymes and proteins involved in motility. There are several reasons favouring a study of this sort: 1. Since both lesions occur within the breast, tissue-specific expression patterns should not cause problems. 2. The comparison of two tumours (which can be matched for their proliferation rates) means that genes involved in proliferation are less likely to be identified, thus decreasing the number of positives identified. 3. Samples of benign and malignant tumours are fairly readily available and usable amounts of RNA can be purified from them.

The choice of malignant tissue was easily made, since several carcinomas had been obtained which had metastasised to lymph nodes. Reasonable quantities of RNA had also been prepared from several of these samples. A number of fibroadenoma samples had also been collected. These proliferative breast lesions are benign and their presence has not been correlated with an elevated risk of developing breast cancer (Cole et al., 1978; Hutchinson et al., 1980). Furthermore, extraction of RNA from several of these tumours yielded sufficient quantities of RNA for analysis of gene expression. The analysis of RNA from several samples of benign and malignant tumours helps to reduce two of the problems cited above, patient-specific expression patterns and stromal infiltration. A similar analysis using breast carcinomas and fibroadenomas has recently isolated the gene stromelysin 3 on the basis of its higher expression in carcinomas than fibroadenomas (Basset et al., 1990). Interestingly, this gene is expressed in the stromal cells surrounding carcinoma cells and is probably expressed in response to a signal from the tumour cells.
To identify genes expressed to a greater extent in one tumour type relative to the other, the differential hybridisation technique was employed. This method requires the construction of a poly(A)+ cDNA library from at least one of the tissues (and preferably both) which is then screened sequentially with radio-labelled total cDNA from both tissue types. Analysis of the autoradiographs for specific differences in signal strength produced by both probes will hopefully identify mRNAs which are differentially expressed in the two tissues. This technique has been used successfully by a number of other groups to identify genes which are expressed at different levels in malignant and non-malignant cells and tissues (for example, Elvin et al., 1988; Schalken et al., 1988; Chen et al., 1990 and Pogue-Geile et al., 1991).

In the ICI/Joint laboratory, two cDNA libraries were constructed and screened using the above method: one from a carcinoma (C328) and one from a fibroadenoma (F455). Differential screening analysis of the carcinoma C328 cDNA library using cDNA probes derived from poly(A)+ RNA purified from a panel of carcinomas and fibroadenomas identified a number of clones which were consistently (consistent was taken to mean no more than two contrary results in a total of ten comparisons) expressed at higher levels in one tumour type (benign or malignant) relative to the other (Sharp et al., 1990; Adams et al., 1992b; Sharp, 1992 and Sharp et al., 1992). Of the sequences not previously identified, clone C328-10 appeared to be one of the most interesting for a number of reasons. These are given below.

The fact that the sequence contained within clone C328-10 had not been previously isolated, as judged by searching the EMBL and Genbank nucleic acid sequence databases, was in itself of interest. This indicated that the cDNA may be the product of a novel gene.

Northern blotting analysis performed by Dr. S. M. Adams showed that the gene exhibited a consistent and significantly higher level of expression in fibroadenomas compared with carcinomas (figure 1.3.1), indicating that expression of this mRNA may be associated
Expression of clone C328-10 in fibroadenomas and carcinomas as determined by Northern blotting analysis. The level of expression has been controlled for differences in loading between tracks by reprobing the filter with radiolabelled cDNA synthesised by oligo dT-primed reverse-transcription of 4.65μg of total RNA from cell lines MDA MB 231, Hs578T, HeLa, fibroadenomas F524, F672 and carcinoma C690 in a ratio of 15: 8: 1: 2.5: 10: 10. Expression is displayed relative to tumour F524 (taken as 100%). Figure reproduced from Adams et al., (1992b).
with the benign phenotype. This result does not, however, determine whether the difference is causative or resultant of the benign state.

Analysis of C328-10 genomic sequences also provided interesting results. Southern blotting analysis performed on human genomic DNA cleaved separately with a number of restriction endonucleases resulted in identification of multiple DNA fragments capable of hybridising with a probe derived from the C328-10 cDNA (figure 1.3.2a). These fragments ranged in size from greater than 20 to less than 2 kilobases (Kb). Many of the bands remained when the filters were washed at high stringency (figure 1.3.2b), thus demonstrating that they were closely similar to the C328-10 cDNA sequence. A similar result was obtained when Dr. S. Adams probed a Southern blot of human genomic DNA cleaved with either SalI or SstI and separated by pulse-field gel electrophoresis (blot obtained from Dr. Z. Wong, Dept. of Genetics, Leicester). This experiment revealed that multiple hybridising bands in the size range <50 Kb to 455 Kb were present within the human genome (figure 1.3.2c). The nature of these sequences was, however, not known. It was postulated that they may represent pseudogenes, closely related functional genes (possibly isoforms), or other cross-hybridising genomic sequences.

Southern blotting analysis of genomic DNAs from many different species (so called "nature blots") was performed by Dr. S. M. Adams. Probing of these blots with sequences from the C328-10 cDNA clone identified hybridising sequences in many eukaryotes (figure 1.3.3). Some species, for example lion, mouse and roe deer, displayed multiple bands, a situation similar to that seen in humans. Other species, for example Xenopus and Drosophila melanogaster, contained a single hybridising band. This result indicated that the gene(s) encoding the C328-10 cDNA is highly conserved throughout Eukaryota.

The above preliminary results indicated that the C328-10 cDNA warranted further analysis. The 551bp C328-10 cDNA clone was judged not to be full-length by Northern blotting analysis which showed that a major ~950nt mRNA was expressed in breast tumour-derived tissue culture cells (figure 1.3.4). Also, no consensus translational start sites had been identified within the sequence.
Southern blotting analysis of human genomic DNA. 
a. 4μg of human genomic DNA (placental) was digested with 
EcoRI (lane 1), HindIII (lane 2), BamHI (lane 3) and BgIII (lane 4), separated by electrophoresis through a 0.8%
agarose gel and transferred to a nylon membrane. After hybridisation with a 32P-radiolabelled probe generated
from the 480bp Styl fragment of clone C328-10 (see figure 1.3.5), the filter was washed at a final stringency
of 2x SSC at 65°C and autoradiographed using Amersham Hyperfilm™-MP for 7 days at -70°C. 
b. 4μg of human
genomic DNA (placental) was digested with HindIII, separated by electrophoresis through a 1% agarose gel and
transferred to a nylon membrane. After hybridisation with a 32P-radiolabelled probe generated from the 307bp
Kpnl-Xhol fragment of clone C328-10 (see figure 1.3.5) the filter was washed at a final stringency of 2x SSC
at 65°C and autoradiographed using Amersham Hyperfilm™-MP for 10 days at -70°C (lane 1). After stripping
the filter and reprobing with a probe generated from the same Kpnl-Xhol fragment of clone C328-10, the filter
was washed at a final stringency of 0.05x SSC at 65°C and autoradiographed using Amersham Hyperfilm™-MP
for 14 days at -70°C (lane 2). Result obtained by Dr. S. M. Adams. 
c. A Southern blot of human genomic DNA
(CEPH panel No. 2860H7) digested with SalI (lane 1) or SstII (lane 2) and separated by pulse field gel
electrophoresis was obtained from Dr. Z. Wong (Dept. of Genetics, Leicester). After hybridisation with a 32P-
radiolabelled probe generated from the 480bp Styl fragment of clone C328-10, the filter was washed at a final
stringency of 0.5x SSC at 65°C and autoradiographed using Amersham Hyperfilm™-MP for 9 days at -70°C.
Result obtained by Dr. S. M. Adams.
Figure 1.3.3

Nature blotting analysis of genomic DNAs from human (lane 1), mouse (lane 2), lion (lane 3), roedeer (lane 4), bat (lane 5), perch (lane 6), chicken (lane 7), *Xenopus* (lane 8) and *D. melanogaster* (lane 9). 4μg of genomic DNA from each species was cleaved with *HindIII* restriction endonuclease, separated on a 1% agarose gel and hybridised with a 32P-radiolabelled probe generated from the 480bp *Styl* fragment of clone C328-10. Final washing stringency was 2x SSC at 65°C. Autoradiography was for 8 days at -70°C with Amersham Hyperfilm™-MP.
Northern blotting analysis of human breast tumour cell lines performed by Dr. S. Adams. 2μg of poly(A)^+ RNA from cell lines MDA-MB-231 (lane 1), MCF-7 (lane 2) and Hs 578 T (lane 3) was separated by gel electrophoresis and transferred to a nylon membrane. Hybridisation was in 50% formamide at 45°C with a ^32P-radiolabelled 307bp KpnI-XhoI fragment from clone C328-10. Final washing stringency was 0.5x SSC at 65°C. Autoradiography with Amersham Hyperfilm^TM-MP was for 4hr at room temperature.
In an attempt to identify a full-length cDNA sequence, Dr. J. M. Varley screened a normal human breast epithelium λgt10 cDNA library (Clontech) with the 480bp Styl fragment from clone C328-10. A number of clones were isolated and sequence information obtained. A schematic comparison of these sequences with that of clone C328-10 is shown in figure 1.3.5.

From this figure it can be seen that a set of cDNAs were isolated which contained sequences extending 3' and/or 5' with respect to the original C328-10 clone. However, the majority of these sequences were different with respect to each other, not only outside of the original cDNA sequence, but also within this sequence. Two of the clones (10-2 and 10-8) diverged with respect to each other and the original C328-10 clone at an identical nucleotide, indicating that alternative splicing may occur. Since none of the sequences around divergence sites corresponded to consensus splice acceptor sites (Jackson, 1991), it did not appear that any of the sequences represented unspliced introns. The possibility also existed that some of the divergent sequences represented cloning artefacts. Clones 10-10 and 10-12 are of interest because they contain identical sequence extending 5' of that found in clone C328-10. This additional sequence bears an initiation codon within a consensus translation start site (see below).

Re-screening of the C328 cDNA library by Dr. S. M. Adams failed to identify any clones longer than the original C328-10 cDNA. However, re-screening the F455 fibroadenoma cDNA library (also by Dr. S. M. Adams) resulted in isolation of clone Q19-5.1a. This clone is 521bp long and extends 159bp beyond the 5' end of clone C328-10. Importantly, this extra sequence agrees with the 138bp extension found in cDNA clone 10-12 isolated from the normal breast epithelium library (figure 1.3.6). The region of 5' extension within clone 10-10 that has been sequenced also agrees with Q19-5.1a (figure 1.3.5). Primer-extension analysis performed by Dr. S. M. Adams on human placental RNA indicates that clone Q19-5.1a probably contains the entire 5' untranslated region within the human bbcl mRNA (not shown).
Figure 1.3.5

Comparison of clones isolated from the human mammary cDNA library by Dr. J. M. Varley by hybridisation with a $^{32}$P-radiolabelled probe generated from the 480bp StyI fragment of clone C328-10. Restriction enzyme sites used for fragment preparation are shown. The longest open reading frame (ORF) within clone 10-12 is indicated by arrows. Numbering is with respect to the adenosine residue of the initiation codon within the ORF shown.
Figure 1.3.6

Alignment of clones 10-12 and Q19-5.1a with clone C328-10. The longest open reading frame (ORF) within clone 10-12 is shown. Numbering is with respect to the adenosine residue of the ORF initiation codon. The 3' extension within clone 10-12 is represented by an open box.
The combined lengths of clones Q19-5.1a and C328-10 is 710bp (figure 1.3.7). If a poly(A) tail of about 200nt is assumed to be present on the mRNA, which is not an unrealistic length (Baralle, 1983), then the 710bp of combined cDNA sequence within clones Q19-5.1a and C328-10 is approaching the size of the major C328-10 transcript. As mentioned above, a consensus translation initiation site (Cavener, 1987) is present within the 5' extension sequences of clone 10-10 and 10-12. This initiation site marks the beginning of the longest open reading frame (ORF) within the cDNA sequence and the only one containing a consensus translation initiation site (figure 1.3.7). The predicted protein product of the ORF is 211 amino acids long and contains a high proportion of basic residues (figure 1.3.7). The gene from which the C328-10 cDNA was transcribed has been named \( bbcl \) (breast basic conserved 1) on the basis of the above characteristics.

1.5 Neoplasia in \( D. \) melanogaster

The presence of multiple hybridising bands in human genomic DNA probed with C328-10 cDNA sequence and the isolation of a complex set of human cDNAs indicated that further analysis of human \( bbcl \) sequences, especially genomic sequences, could yield very complicated results which would be hard to interpret. The identification of hybridising sequences in many eukaryotic organisms indicated that the gene may fulfil an important and evolutionarily conserved cellular role. What is more, several species, including \( Drosophila \) melanogaster, contain a single hybridising genomic restriction fragment, indicating a single gene. The presence of a single gene could greatly simplify studies of this locus in \( Drosophila \). In addition, \( Drosophila \) genetics provides a powerful tool to analyse gene function, possibly allowing a role for the gene to be determined. \( Drosophila \) tumourigenesis has been studied for many years and more recently, a number of genes have been identified which play roles in tumourigenesis and malignancy in this organism. A group of at least 20 recessive lethal tumour
Figure 1.3.7

Sequence of the human \textit{bbcl} cDNA, compiled from cDNA clones C328-10 and Q19-5.1a, together with the translation product of the longest open reading frame. The 3' extension sequence found in cDNA clone 10-12 is also shown. Numbering is as shown in figure 1.3.6. Polyadenylation sites present in individual cDNA clones are shown by crosses above the sequence.
suppressor genes has been identified in *Drosophila* on the basis of imaginal disc overgrowth in homozygous mutants (reviewed in Bryant and Schmidt, 1990). Loss of some of these genes, such as *giant discs* (*igd*, Bryant and Schubiger, 1971), *fat* (*fat*, Bryant *et al.*, 1988) and *discs overgrown* (*dco*, Jursnich *et al.*, 1990) result in hyperplastic growths which retain a single-layered epithelial structure and the ability to differentiate when transplanted into normal hosts. Such lesions are similar to benign tumours of higher eukaryotes. Other genes such as *discs large* (*dlg*, Stewart *et al.*, 1972) and *giant larvae* (*igl*, Gateff and Schneiderman, 1974), result in neoplastic cell masses in which normal epithelial cell organisation breaks down. These malignant tumours also fail to differentiate when transplanted into normal hosts and instead continue to grow and invade host tissues, eventually killing the host. Such tumours behave very similarly to fully malignant tumours of higher eukaryotes.

A common theme in all the tumours involving mutation of these genes appears to be a breakdown in intercellular communication. This results in disc overgrowth when cells either fail to receive or do not respond to a growth inhibiting stimulus normally generated by neighbouring cells when the correct disc size has been achieved. For example, mutation of the tumour suppressor gene *giant larvae* results in malignant neuroblastomas of the larval central nervous system and imaginal disc neoplasia (Gateff and Schneiderman, 1974). This gene encodes a protein with significant similarity to the cadherin family of cell adhesion molecules (Klambt *et al.*, 1989) and the loss of this protein is postulated to prevent differentiation, and associated cessation of division, of cells normally expressing it (Klambt *et al.*, 1989). Given the important role of cell adhesion molecules in malignancy of higher eukaryotic tumours (see above), the finding that such a protein is also implicated in tumourigenesis/malignancy in *Drosophila* is particularly interesting. Other cell adhesion molecules will also probably be found to have roles in *Drosophila* tumourigenesis/malignancy as their homologues have in vertebrate malignancy (see above).

A second cloned *Drosophila* tumour suppressor gene, *discs large* is predicted to encode a guanylate kinase which is localised at septate
juncti ons (Woods and Bryant, 1991). The phenotype of \textit{dlg} is very similar to that of \textit{Igl}, and is thought to be involved in signal transduction pathways controlling epithelial cell growth and maintenance of apical-basal cell polarity in the disc epithelium (Woods and Bryant, 1991). Since homozygous mutants for \textit{dlg} contain a disorganised actin cytoskeleton, it has been proposed that the \textit{dlg} protein is involved in organising the membrane cytoskeleton (Woods and Bryant, 1991). Interestingly, a rat protein (SAP90) localised at presynaptic junctions in the cerebellum has recently been identified and shown to bear a high (54\%) degree of amino acid identity with the \textit{Drosophila} \textit{dlg} protein (Kistner \textit{et al.}, 1993). Synapses, like tight and septate junctions are areas of highly organised membrane cytoskeleton, thus indicating that SAP90 may be involved in cytoskeletal organisation.

In contrast to higher eukaryotes, very few dominantly acting oncogenes have been identified in \textit{Drosophila}. Whilst \textit{Drosophila} contains many homologues of higher eukaryotic proto-oncogenes, mutation of these genes tends to result in cell-fate changes during embryogenesis (Bryant and Schmidt, 1990) as a result of loss of function mutations. Dominant or semi-dominant gain-of-function mutations have been reported for a few proto-oncogenes in \textit{Drosophila} (reviewed in Bryant and Schmidt, 1990). However, these all result in cell-fate changes. The reason for the lack of dominant gain-of-function mutations resulting in excess cell proliferation is purely a practical one. Such mutations are likely to act as dominant lethals; a class of mutation not easily identified or studied by \textit{Drosophila} genetics.

One \textit{Drosophila} gene in which a dominant mutation has been identified is the oncogene \textit{Tumourous-lethal} (\textit{Tum-I}). The mutation causes abnormal proliferation and differentiation of the larval haematopoietic system and is lethal at the larval/pupal transition (Hanratty and Ryerse, 1981). It has recently been determined that \textit{Tum-I} is a neomorphic mutation of the recessive \textit{hopscotch} (\textit{hop}) gene (Hanratty and Dearolf, 1993), the normal product of which is required for the maintenance and continued division of imaginal disc precursor cells.
Another gene identified in Drosophila as a suppressor of abnormal development, abnormal wing discs (Awd), could be very important during malignant progression of vertebrate tumours. The Awd gene was identified in Allen Shearn's laboratory as a cell autonomous lethal mutation in Drosophila resulting in failed development of the wing and other imaginal discs (Dearolf et al., 1988). The wing discs develop a morphology indicative of disorganised cell growth and do not develop properly if transplanted into normal hosts. In this phenotype, the Awd mutants behave similarly to mutants in the second class of tumour suppressor genes described above (e.g. dig and lgl). However, the discs of Awd mutant larvae do not continue to grow and are, in fact, smaller than normal discs. Hence, this gene is not a tumour suppressor gene. The Awd gene has subsequently been shown to encode a nucleoside diphosphate (NDP) kinase (Biggs et al., 1990) which may be involved in signal transduction (Bominaar et al., 1993). The rat homologue of Awd, nm23, was cloned independently by Patricia Steeg in a study of genes differentially expressed between malignant and non-malignant rat melanoma cell lines (Steeg et al., 1988), the gene being more highly expressed in the non-malignant than malignant cell lines. The human homologue has also been shown to be more highly expressed in certain tumour types displaying low metastatic potential (Bevilacqua et al., 1989; Hirayama et al., 1991) although this has been challenged by others (Sastre-Garau et al., 1992). Interestingly, the human nm23 gene-product has recently been identified as the transcription factor PuF which is necessary for efficient transcription of the c-myc proto-oncogene (Postel et al., 1993). Since the c-myc protein functions in cellular proliferation, differentiation and tumourigenesis by influencing transcription of many genes, this provides another way by which the nm23 protein may affect cell growth and differentiation.

Given that loss of attachment, altered signalling, disorganised morphology and altered gene expression are common themes in malignancy of higher and lower eukaryotic tumours, the above findings are of considerable importance for our understanding of human tumourigenesis and malignancy and underline the usefulness of Drosophila as a model system for studying tumourigenesis. In order
to make use of the benefits that an analysis of the *Drosophila* *bbcl* gene could provide, *Drosophila* cDNA clones capable of hybridising with the human C328-10 cDNA were isolated by Dr. J. M. Varley. The results of this and subsequent experiments aimed at characterising the *D. melanogaster* and human *bbcl* genes and gene-products form the basis of this thesis.
CHAPTER 2

Materials and Methods

2.1 Materials

2.1.1 Chemicals

General laboratory chemicals were supplied by Fisons (Loughborough), BDH (Poole) or Sigma (Poole) unless otherwise stated and were of analytical grade or equivalent. Bacterial culture media were obtained from Difco (East Molesley) and animal cell culture media from Gibco BRL (Paisley). Radiolabelled nucleotides were supplied by Amersham International (Amersham) and radiolabelled amino acids by ICN FLOW (High Wycombe). Millipore Super-Q® system (Millipore Corporation, Bedford) treated water was used for all solutions.

2.1.2 Enzymes and Proteins

Restriction endonucleases were purchased from Gibco BRL, Pharmacia P-L Biochemicals (Milwaukee) or New England Biolabs (via CP Laboratories, Bishop's Stortford). T4 polynucleotide kinase was supplied by Gibco BRL. DNA polymerase I (Klenow fragment) was supplied by either Amersham International or Gibco BRL. T4 DNA polymerase was obtained from New England Biolabs. T7 DNA polymerase was obtained from Pharmacia P-L Biochemicals and Sequenase™ version 2 (modified T7 DNA polymerase) was obtained from United States Biologicals (via Cambridge Bioscience, Cambridge). Taq DNA polymerase was obtained from Amersham or Northumbria Biologicals Limited (Northumbria). T4 DNA ligase was purchased from Gibco BRL and New England Biolabs. Avian Myeloblastosis Virus reverse transcriptase was obtained from Life Sciences Inc. (St. Petersburg). Proteinase K was purchased from

TM, ® All registered trade marks are recognised as belonging to the company named.
Boehringer Mannheim (Lewes). Pronase E, RNase A (pancreatic RNase) and DNase I were obtained from Sigma. BSA (enzyme grade) was obtained from Gibco BRL. DNase-free RNase A and RNase-free DNase I were prepared by the methods described (Sambrook et al., 1989).

2.1.3 Bacterial Strains and Culture Conditions

Strains of *Escherichia coli* K12 used are given in table 2.1.1. Bacteria were grown using the solid and liquid media given below at 37°C.

Solid media:

a. Minimal agar (with supplements): K$_2$HPO$_4$, 60mM; KH$_2$PO$_4$, 33mM; (NH$_4$)$_2$SO$_4$, 7.6 mM; tri-sodium citrate, 850μM; thiamine, 1mM; glucose, 0.2% (w/v); agar, 1.5% (w/v).

b. Luria agar: tryptone, 1% (w/v); yeast extract, 0.5% (w/v); NaCl, 0.5% (w/v); agar, 1.5% (w/v).

c. H agar: tryptone, 1% (w/v); NaCl, 0.8% (w/v); agar, 1.2% (w/v).

d. BBL top agar: BBL trypticase, 1% (w/v); NaCl, 0.5% (w/v); agar, 0.8% (w/v).

Liquid media:

a. BBL broth (with supplement): BBL trypticase, 1% (w/v); NaCl, 0.5% (w/v); MgSO$_4$.7H$_2$O, 0.24% (w/v).

b. Luria broth: tryptone, 1% (w/v); yeast extract, 0.5% (w/v); NaCl, 0.5% (w/v).

c. 2xYT broth: tryptone, 1.6% (w/v); yeast extract, 1% (w/v); NaCl, 0.5% (w/v).

2.1.4 Cloning Vectors Used

The *D. melanogaster* embryo cDNA library screened by Dr. J. M. Varley was made by Dr. B. Hovemann (University of Heidelberg). An aliquot was obtained from Dr. D. J. Finnegan (Dept. of Molecular Biology, University of Edinburgh). The library was constructed in the bacteriophage vector λgt11 (Young and Davis, 1983) using random
<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>JM101</td>
<td>supE thi-1 Δ(lac-proAB) F'[tra D 36 proAB laclacZAM15]</td>
<td>(Messing et al., 1981)</td>
</tr>
<tr>
<td>DH5α</td>
<td>supE44 ΔlacU169(q80 lacZAM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1</td>
<td>(Hanahan, 1983)</td>
</tr>
<tr>
<td>Y1088</td>
<td>Δ(lac)U169 supE supF hsdR metB trpR lonA21 proC::Tn5 (pMC9)</td>
<td>(Young and Davis, 1983)</td>
</tr>
<tr>
<td>Y1090</td>
<td>araD139 Δ(lac)U169 Δ(lon) rpsL supF trpC22::Tn10(tet) (pMC9)</td>
<td>(Young and Davis, 1983)</td>
</tr>
</tbody>
</table>

Table 2.1.1 Bacterial strains utilised.
primed cDNA prepared from 1-16hr embryo poly(A)+ RNA of *D. melanogaster*. The cDNA was cloned into the unique EcoRI site of λgt11.

Routine subcloning was performed in the vectors pUC18 (Yannisch-Peron *et al.*, 1985) and M13mp18 (Norrander *et al.*, 1983), both purchased from Pharmacia. The Bluescript SK+ plasmid used for *in vitro* transcription experiments was purchased from Stratagene (San Diego). The pGEX-3X bacterial expression plasmid (Smith and Johnson, 1988) was obtained from Dr. V. Ohanian (Dept. of Biochemistry).

2.1.5 Oligonucleotides Used

Synthetic oligonucleotides used during the course of the work are shown in table 2.1.2. The fragment used for primer-extension of *D. melanogaster bbcl* mRNA was generated as follows: PCR amplification of a region of cDNA clone 4 using primers Y5A and 3605' produced a 168bp product. This fragment was cleaved using *HinfI* restriction endonuclease to generate 58bp and 110bp fragments. After separation by polyacrylamide gel electrophoresis, the 58bp fragment was purified and used as a primer. During polynucleotide kinase-catalysed radiolabelling of the primer, the strand complementary to the RNA was preferentially labelled since no 5' phosphate group was present; this sequence being derived from the 3605' primer.

2.2 Methods of Sterilisation

**Autoclaving:** Autoclaves used to sterilise media and disposable plastic-ware were as follows: a Cabburn 8cu.ft. capacity autoclave (Cabburn Sterilisers, Shoeburyness) set to attain a temperature of 121°C for 30 minutes; a model ST19 portable electric autoclave (Dixon's surgical instruments, Wickford) set to attain a temperature of 121°C for 20 minutes.

**Dry sterilisation:** A B&T "Unitemp" sterilising cabinet (Laboratory
### Table 2.1.2

Synthetic oligonucleotides used. See figure 2.1.1 for a diagram showing the positions of oligonucleotides within the human and *D. melanogaster* bbcl sequences. All oligonucleotides are shown in the orientation 5'-3'.
Figure 2.1.1

Positions of synthetic oligonucleotides and selected restriction enzyme recognition sites within the human and *D. melanogaster* *bbd* nucleic acid sequences. The open reading frame within both sequences is represented by a thick arrow and the positions of oligonucleotides by small arrows. Positions of oligonucleotides within sequence towards the 5' end of *D. melanogaster* clone #4 (see chapter 3, section 3.2) are also indicated (boxed region). The position of the 3' extension sequence present in human cDNA clone 10-12 is indicated.
The thermal equipment, Oldham) was used to sterilise glass-ware. The cabinet was set to attain a temperature of 160°C for 6 hours.

**Filter sterilisation:** Small volumes (up to 50ml) were filter sterilised by passing through Acrodiscs (Gelman Sciences, Ann Arbor) with a pore size of 0.2μm. Larger volumes (50-500ml) were filter sterilised using Nalgene 0.2μm vacuum filter sterilising units (Nalgene, Rochester, New York).

2.3 Nucleic Acid Methods

2.3.1 Solutions Used During the Handling of Nucleic Acids

**Acrylamide solution (40%):** 38g of acrylamide (Serva, Heidelberg or National Diagnostics, Aylesbury) and 2g of N,N’-methylene bisacrylamide (Serva) were dissolved in 100ml of Q water. The solution was deionised using Amberlite® MB-1 ion exchange resin (Sigma), filtered through Whatman™ No. 1 filter paper (Whatman International, Maidstone) and stored in the dark at 4°C.

**Ammonium acetate (5M):** 38.5g of ammonium acetate was dissolved in 100ml of Q water; following filter sterilisation, the solution was stored at room temperature.

**Caesium Chloride (5.7M):** 96g of caesium chloride (BRL, optical grade) was dissolved in 100ml of 0.1M EDTA (pH 8), sterilised by autoclaving and stored at room temperature.

**Caesium Chloride (1.3, 1.5 and 1.7g/ml):** Caesium chloride (BRL, optical grade) was dissolved at the specified densities in λ buffer. After checking the densities using a refractometer (Bellingham+Stanley, Tunbridge Wells), solutions were sterilised by autoclaving and stored at room temperature.

**Calcium Chloride (1M):** 21.9g of calcium chloride.6H₂O (BDH) was dissolved in 100ml of Q water, sterilised by autoclaving and frozen in 10ml aliquots at -20°C.
Chloroform/iso-Amyl alcohol: 96% (v/v) chloroform (Fisons), 4% (v/v) iso-Amyl alcohol (Fisons) was prepared and stored at room temperature.

Citric Acid (1M): 4.2g of citric acid (BDH) was dissolved in 20ml of Q water, sterilised by filtration and stored at room temperature.

Colony neutralising solution: 0.5M Trizma base, 1.5M NaCl; pH adjusted to 7.4 with concentrated HCl. Stored at room temperature.

ddNTP solutions (10mM): Powdered ddNTPs (Sigma) were dissolved in the specified volume of TE and stored at -20°C: 10mg ddTTP, 1639μl; 5mg ddCTP, 862μl; 5mg ddGTP, 806μl; 2mg ddATP, 323μl.

Denhardt’s (50x): 1% (w/v) each of Ficoll 400 (Pharmacia), Polyvinylpyrrolidone (Sigma) and BSA (fraction V, Sigma) were dissolved in Q water. The solution was divided into 2.5ml aliquots and stored at -20°C.

DEPC-treated Q water: 0.1% (v/v) DEPC was added to Q water. After shaking vigorously, the bottles of Q water were left overnight in a fume cupboard with loosened caps. The DEPC-treated Q water was then autoclaved and stored at room temperature.

dNTP solutions (0.1M): Powdered dNTPs (Sigma) were dissolved in the specified volume of TE¹ and stored at -20°C: 10mg of dATP, 186.9μl; 5mg of dTTP, 103.7μl; 5mg of dGTP, 98.6μl.

dNTP solutions (50mM): Powdered dNTPs (Sigma) were dissolved in the specified volume of TE and stored at -20°C: 5mg of dTTP, 155.8μl; 10mg of dCTP, 338μl; 10mg of dGTP, 316μl; 10mg of dATP, 339μl.

¹ This TE is specifically for random-primed oligonucleotide synthesis of DNA probes. It has the following composition: 3mM Tris-HCl (pH 7), 0.2mM EDTA (pH 8).
Direct Sequencing from PCR: All dNTP/ddNTP mixes contained 80µl of 0.5mM solutions of dCTP, dGTP and dTTP. ddNTPs are as follows: A mix (3.2µl of 5µM ddATP), C mix (80µl of 50µM ddCTP), G mix (80µl of 50µM ddGTP), T mix (80µl of 50µM ddTTP). TE was added to a final volume of 500µl.

DTT (1M): 1.55g of DTT (Sigma) was dissolved in a final volume of 10ml of 10mM sodium acetate (pH 5.6) and sterilised by filtration. 1ml aliquots were stored at -20°C.

EDTA (0.5M, pH 8): 186.1g of EDTA (Fisons) was dissolved in 1L of Q water adjusted to pH to 8.0 with NaOH pellets and 10M NaOH. Following sterilisation by autoclaving, the solution was stored at room temperature.

Ethidium Bromide (10 and 5mg/ml): Ethidium Bromide (Serva) was dissolved at the specified concentration in Q water and stored in opaque plastic bottles at room temperature.

Ethanol (100%, 80%, 70%): Ethanol (Fisons) was diluted as required with Q water and stored at -20°C.

Formamide, deionised (100%): Formamide (Fisons) was deionised using Amberlite® MB-3 ion exchange resin (Sigma), filtered through Whatman™ No.1 filter paper and stored in aliquots at -20°C.

GFM buffer: 804µl of 6.9M glyoxal (deionised); 3.89ml of 100% formamide (deionised) and 306µl of MOPS (10x) were mixed together and frozen at -70°C in 250µl aliquots.

Glycerol (45%): Glycerol (Fisons) was dissolved at a concentration of 45% (w/v) in Q water and filter sterilised. The solution was stored at room temperature.

Glyoxal, deionised (6.9M): Glyoxal was supplied as a 40% (w/v) solution (equivalent to 6.9M) by BDH. After deionising with AG® 501-X8(D) 20-50 mesh ion exchange resin (Bio-Rad, Richmond) and filtering through Whatman No.1 filter paper, the solution was stored...
in 1ml aliquots at -70°C (Sambrook et al., 1989).

Heparin (porcine grade II; 100mg/ml): 50mg of heparin was dissolved in 500µl of DEPC treated Q water and stored at -20°C.

HEPES (2M, pH 6.6): 9.53g of HEPES (Sigma) was dissolved in a final volume of 20ml Q water after adjusting the pH to 6.6 with 5M NaOH. After filter sterilisation, the solution was stored at 4°C.

HCl (0.25M): 21.55ml of concentrated HCl (Fisons) was added to 978.45ml of Q water. The solution was stored at room temperature.

IPTG (100mM): 238mg of IPTG (Sigma) was dissolved in 10ml of Q water and stored at -20°C.

λ Buffer: 6mM Tris-HCl (pH 8); 10mM MgCl$_2$ (Fisons), 100mM NaCl (Fisons); 0.05% (w/v) gelatin (Fisons). Sterilised by autoclaving and stored at room temperature.

Lithium Chloride (10M): 4.24g of LiCl (Sigma) was dissolved in 10ml of DEPC treated Q water, sterilised by filtration and stored at room temperature.

M13 Sequencing Mixes: These were prepared as given in the table below using 0.5mM dNTP and 10mM ddNTP stocks.

<table>
<thead>
<tr>
<th>Components</th>
<th>Mix</th>
<th>dTTP</th>
<th>dCTP</th>
<th>dGTP</th>
<th>ddTTP</th>
<th>ddCTP</th>
<th>ddGTP</th>
<th>ddATP</th>
<th>TE</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>500µl</td>
<td>500µl</td>
<td>25µl</td>
<td></td>
<td>16µl</td>
<td>1µl</td>
<td>1µl</td>
<td>1ml</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>500µl</td>
<td>500µl</td>
<td>500µl</td>
<td>500µl</td>
<td></td>
<td>1µl</td>
<td>500µl</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>25µl</td>
<td>500µl</td>
<td>500µl</td>
<td>50µl</td>
<td></td>
<td></td>
<td>1µl</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>500µl</td>
<td>25µl</td>
<td>500µl</td>
<td>8µl</td>
<td></td>
<td></td>
<td>1µl</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Magnesium Chloride (1M): 20.3g of MgCl$_2$·6H$_2$O was dissolved in 100ml of Q water, sterilised by autoclaving and stored at room temperature.
Magnesium Sulphate (1M): 24.6g of MgSO$_4$.7H$_2$O (Fisons) was dissolved in 100ml of Q water, sterilised by autoclaving and stored at room temperature.

Maltose (20%): 4g of maltose (Sigma) was dissolved in 20ml of Q water and filter sterilised. The solution was stored at 4°C.

MOPS buffer (10x): 0.2M MOPS, 50mM sodium acetate.H$_2$O, 1mM EDTA (pH 8). The pH was adjusted to 8 with NaOH prior to autoclaving. The solution was stored at room temperature.

PCI: 50% (v/v) phenol (containing 8-hydroxy-quinoline, Fisons), 48% (v/v) chloroform, 2% (v/v) iso-Amyl alcohol; equilibrated against 10mM Tris-HCl (pH 8) and kept in the dark at 4°C.

Phenol (liquified): Liquified phenol containing 0.1% (w/v) 8-hydroxy-quinoline and equilibrated against 100mM Tris (pH 8) was obtained from Fisons and stored at -20°C.

Sodium Acetate, (3M, pH 5.6): 40.8g of sodium acetate.H$_2$O (Fisons) was dissolved in 80ml of Q water. After adjusting the pH to 5.6 with glacial acetic acid, the volume was increased to 100ml. The solution was sterilised by autoclaving and stored at room temperature.

Sodium Acetate, (1.1M, pH 7): 14.96g of sodium acetate.H$_2$O was dissolved in a final volume of 100ml of Q water after the pH had been adjusted to 7.0 with glacial acetic acid. Following sterilisation by autoclaving, the solution was stored at room temperature.

Sodium Chloride (4M): 23.4g of sodium chloride (Fisons) was dissolved in 100ml of Q water and sterilised by autoclaving. The solution was then stored at room temperature.

SDS (10%): 50g of SDS (Fisons) was dissolved in 500ml of Q water and stored at room temperature.
Di-Sodium Hydrogen Phosphate (0.5M): 89g of Na$_2$HPO$_4$.2H$_2$O (Fisons) was dissolved in 1L of Q water. The solution was stored at room temperature.

Sodium Hydroxide (10M): 200g of sodium hydroxide (Fisons) was added slowly to 400ml of Q water. After adjusting the volume to 500ml, the solution was stored at room temperature.

Southern denaturation solution: 0.5M NaOH, 1.5M NaCl; stored at room temperature.

Southern neutralisation solution: 0.5M Tris-HCl (pH 7.4), 3M NaCl; stored at room temperature.

SSC (20x): 3M NaCl, 0.3M tri-sodium citrate (Fisons); pH adjusted to 7.0 with NaOH, autoclaved and stored at room temperature.

TAE (10x): 48.4g of Trizma base (Sigma) and 20ml of 0.5M EDTA (pH 8) were dissolved in 1L of Q water after the pH had been adjusted to 7.5 with glacial acetic acid (Fisons).

TBE (10x): 108g of Trizma base, 55g of boric acid (Fisons) and 9.3g of EDTA were dissolved in 1L of Q water.

TBE (0.5x) Acrylamide (6%) Urea: 50ml of 10 x TBE, 430g of urea (Serva) and 150ml of 40% acrylamide were dissolved in 1L of Q water and stored in the dark at 4°C.

TBE (2.5x) Acrylamide (6%) Urea: 250ml of 10 x TBE, 430g of urea, 150ml of 40% acrylamide, 50g of sucrose (BRL) and 50mg of bromophenol blue (Sigma) were dissolved in 1L of Q water and stored in the dark at 4°C.

TE: 10mM Tris-HCl (pH 7.2), 1mM EDTA (pH 8); autoclaved and stored at room temperature.

TLE: 20ml of 1M Tris-HCl (pH 9), 0.42g of LiCl (Sigma), 5ml of 500 mM EDTA (pH 8) and 1ml of 10% SDS were dissolved in 100ml of
DEPC-treated Q water, sterilised by autoclaving and stored at room temperature.

TM: 100mM Tris-HCl (pH7.5), 50mM MgCl₂. Sterilised by autoclaving and stored at room temperature.

tRNA (50mg/ml): 5mg of tRNA (E. coli type XX, Sigma) was dissolved in 100μl of DEPC treated Q water and stored at -20°C.

Tris-HCl (1M): 121.1g of Trizma base was dissolved in 800ml of Q water after which the pH was adjusted to the required value with concentrated HCl. The volume was increased to 1L and, following sterilisation by autoclaving, the solution was stored at room temperature.

X-gal (20mg/ml): 200mg of X-gal obtained from Novabiochem (Nottingham) was dissolved in 10ml of dimethylformamide (Fisons) and stored in the dark at -20°C.

2.3.2 Purification of Nucleic Acids

Purification of Genomic DNA from Tissues

Human tissue was powdered in a Mikro Dismembrator II (FT Scientific Instruments, Tewksbury) within a class II microbiological safety cabinet (Medical Air Technology Ltd., Manchester) at liquid nitrogen temperature; other tissues were powdered under liquid nitrogen in a mortar and pestle. The powder was then added to 5ml of SE buffer (150mM NaCl; 100mM EDTA, pH 8) plus 1/10 volume of 10% SDS and 1/40 volume of proteinase K (20mg/ml) in Corex® (DuPont Scientific Instruments, Delaware) tubes, mixed gently and incubated overnight at 37°C. An equal volume of PCI was added and after mixing, the emulsion was spun at 10,000 rpm in a HB-4 rotor (DuPont) within a Sorvall® RC-5B centrifuge (DuPont) for 2 minutes. The aqueous phase was transferred to a fresh tube and the phenol was re-extracted with a quarter the volume of SE buffer. Both aqueous phases were pooled and 2 volumes of absolute ethanol (Fisons) was added. Upon mixing, a precipitate appeared which was
pelleted by spinning for 10 minutes at 10,000 rpm in a HB-4 rotor. After drying, the pellet was resuspended in 500μl of TE containing 100μg/ml pancreatic RNase (DNase-free), transferred to a 1.5ml microfuge tube and incubated at 37°C for 30 minutes. 250μl of PCI was added, the phases mixed thoroughly and then separated by spinning at 13,000 rpm for 1 minute in an MSE microfuge (MSE Scientific Instruments, Crawley). The aqueous phase was removed and 50μl of 3M sodium acetate pH 5.6 was added, then 1ml of absolute ethanol. The precipitate which formed on mixing was pelleted by spinning at 13,000 rpm for 20 minutes in a microfuge, washed with 70% ethanol and air dried. The pellet was resuspended in 50-100μl of TE or Q water and the nucleic acid concentration determined by spectrophotometric analysis.

**Purification of Total RNA from Tissues**

Total RNA was purified from whole tissues by either caesium chloride centrifugation (Glisin et al., 1974) or by acid guanidinium thiocyanate-phenol-chloroform extraction (Chomczynski and Sacchi, 1987). All solutions were prepared using DEPC-treated Q water and all vessels containing RNA were soaked in DEPC-treated Q water prior to use. The two methods are described briefly below:

**Caesium chloride centrifugation:** Powdered tissues were prepared as described in section 2.3.2, above. Without thawing, the powder was transferred to 5ml of TLE and phenol extracted twice. Both aqueous phases were pooled and subjected to caesium chloride pad (5.7M CsCl) centrifugation at 25,000 rpm in a TST41:14 swing-out rotor (Sorvall®) in a Sorvall® OTD 50B ultracentrifuge (DuPont) for 18hr at 25°C. After removal of the supernatant, the RNA pellet at the bottom of the tube was resuspended in 500μl of DEPC-treated Q water and ethanol precipitated. The RNA was finally resuspended in up to 200μl of DEPC-treated Q water and stored in aliquots at -70°C. The quantity and quality of the RNA was checked by spectrophotometric analysis and agarose gel electrophoresis respectively.

**Guanidinium thiocyanate extraction:** Tissue was homogenised on ice...
in 1 ml of extraction buffer then extracted with phenol. The aqueous phase was precipitated twice with isopropanol (Fisons), washed with 70% ethanol, dried and resuspended in 0.5% SDS or DEPC-treated Q water. The quantity and quality of the RNA were checked by spectrophotometric analysis and agarose gel electrophoresis respectively.

**Purification of Bacteriophage λ DNA**

The method used for large-scale preparation of bacteriophage λ DNA was modified from a previously described method (Yamamoto *et al.*, 1970). Four well spaced plaques and two areas of bacterial lawn were picked into 250 ml of BBL broth containing 10 mM MgSO₄ and incubated at 37°C overnight with vigorous shaking. If lysis had occurred, chloroform was added to 1% and, after shaking briefly, left for 15 minutes at room temperature. After centrifugation at 8,000 rpm for 10 minutes in a GSA rotor (Dupont Scientific Instruments) the cleared supernatant was transferred to a fresh 1 L conical flask. DNase and RNase were each added to a final concentration of 1 μg/ml and incubated at room temperature for 1 hr. 10 g of NaCl and 25 g of PEG 6000 (Serva) were added and allowed to dissolve prior to incubation overnight at 4°C. Centrifugation at 5,000 rpm for 5 minutes in a GSA rotor pelleted the bacteriophage particles which were then gently resuspended in 1 ml of λ buffer and layered on top of a CsCl block gradient (3.5 ml each of 1.3, 1.5 and 1.7 g/ml CsCl in λ buffer) in a 14 ml polyallomer tube. The gradient was centrifuged in a TST41:14 rotor at 37,000 rpm for 1 hr at 20°C. The bacteriophage particles, visible as a white band, were removed from the 1.3-1.5 g/ml CsCl interface and dialysed against TE for 1 hr at 4°C to remove CsCl. Pronase was added to the dialysis tube contents at a final concentration of 1 mg/ml and the suspension was dialysed against 2 l of TNET [10 mM Tris-HCl (pH 8), 10 mM NaCl, 1 mM EDTA (pH 8), 0.01% Triton-X100 (Sigma)] at 37°C for 2 hr. Phenol extraction and ethanol precipitation yielded a stringy DNA precipitate that was air-dried and resuspended in 500 μl of TE.
Large Scale Purification of Plasmid DNA

Two methods were used for large scale plasmid DNA purification. Subsequent to purification by both methods, the quantity and quality of the DNA was checked by spectrophotometric analysis and agarose gel electrophoresis respectively.

1. Detergent lysis followed by CsCl/EtBr equilibrium density centrifugation was essentially as described in Sambrook *et al.* (1989). Briefly, bacterial cultures were lysed with Triton-X100 and supercoiled plasmid DNA separated from other nucleic acids by centrifugation in a solution of CsCl/EtBr at 39,000rpm for 48hr in a T1270 rotor (DuPont Scientific Instruments). After isolation of the supercoiled plasmid DNA, isopropanol extraction followed by dialysis removed EtBr and CsCl from the sample respectively. After ethanol precipitation, the DNA was resuspended in Q water.

2. Alkaline lysis and purification of DNA by the use of Qiagen columns (Hybaid, Teddington) was performed according to the manufacturer's instructions.

Small Scale Purification of Plasmid DNA

Isolation of plasmid DNA from 1-2ml of bacterial culture was performed essentially according to the method of Birnboim and Doly (1979), with the exclusion of lysozyme which was found to be unnecessary.

Small Scale Purification of Bacteriophage M13 Single-Stranded DNA

50μl of an overnight-grown culture of *E. coli* strain JM101 was added to 25ml of Luria broth and aliquoted into 2ml amounts in 30ml Nunc® universals (Gibco BRL). Well separated "white" plaques and a "blue" positive control plaque were transferred to individual screw-cap tubes which were incubated horizontally with vigorous shaking for 6hr at 37°C. 1.5ml of suspension was pipetted into a 1.5ml microfuge tube and centrifuged to pellet the bacteria. 1ml of
supernatant was removed to a fresh tube and 250μl of 2.5M NaCl, 20% (w/v) PEG 6000 added. After mixing, the tube was incubated at room temperature for 15-30 minutes, centrifuged for 10 minutes and the supernatant removed. The pellet of bacteriophage was resuspended in 100μl of 1.1M sodium acetate (pH 7) and extracted with 100μl of PCI. The aqueous phase was removed to a fresh tube and extracted with 60μl of chloroform:iso-amyl alcohol (24:1). The aqueous phase was removed to a fresh tube and the M13 DNA precipitated with 250μl of ethanol at -70°C for 2hr. After centrifuging for 15 minutes, the pelleted DNA was washed with 70% ethanol, dried and resuspended in 25μl of Q water.

2.3.3 Spectrophotometric Determination of Nucleic Acid Concentrations

Nucleic acid concentrations were determined by optical density (OD) measurement at a wavelength of 260nm against a blank of Q water. Optical density measurements were converted into nucleic acid concentrations using the following relationships: an OD$_{260}$ of 1.0 represents a concentration of 50μg/ml for double-stranded DNA, 40μg/ml for single-stranded DNA or RNA and 20μg/ml for oligonucleotides (Sambrook et al., 1989).

2.3.4 Gel Electrophoresis

Agarose Gel Electrophoresis

Nucleic acids were separated by electrophoresis through horizontal agarose slab gels (Seakem HGT or NuSieve GTG; both purchased from Flowgen, Sittingbourne), which varied between 0.8 and 5% (w/v). Gels were made and run in either 0.5x TAE (DNA) or 1x MOPS (RNA) buffers. Ethidium bromide was added to the gel and buffer at 0.5μg/ml for DNA gels to allow visualisation of the DNA when illuminated by ultraviolet light (254nm wavelength). When a permanent record of a gel was required, the gel was photographed using a Polaroid MP4 land camera loaded with Polaroid type 667 black and white positive film (Polaroid, St. Albans). Markers used on agarose gels were as follows: λ DNA cleaved with EcoRI and HindIII;
plasmid pAT153 cleaved with \( \text{HaeIII} \); plasmid pUC18 cleaved with \( \text{HpaII} \).

Polyacrylamide Gel Electrophoresis

Nucleic acids were separated by electrophoresis through vertical polyacrylamide gels of various percentages and types. 6% non-denaturing gels were used for purification of small DNA fragments, whilst 6% denaturing gels (containing 42% w/v urea) were used for separating the products of DNA sequencing reactions. Denaturing gels were of two types: non-gradient (entirely 0.5x TBE) or gradient (0.5x TBE at the top and 2.5x TBE at the bottom of the gel). The products of primer-extension analysis were separated on a 7.5% denaturing gel containing 42% (w/v) urea and 0.5x TBE. The RNA products of \textit{in vitro} transcription experiments were separated on a 3.2% denaturing gel containing 42% (w/v) urea, 20% (v/v) formamide and 1x TBE. All gels were run in TBE buffer of the same strength as that in the gel except gradient gels which contained 0.5x TBE in the cathode chamber and 1x TBE in the anode chamber.

2.3.5 Restriction Enzyme Digestions and Southern (DNA) Blotting

Restriction endonucleases were used according to the manufacturers' instructions and incubated with the substrate DNA until complete digestion occurred (as determined by agarose gel electrophoresis of an aliquot of the reaction against undigested substrate DNA). 1/10 the volume of loading buffer (20mM EDTA, 50% (v/v) glycerol, 0.05% (w/v) bromophenol blue) was added to the digested DNA and the sample electrophoresed through an agarose gel against a suitable size marker. After electrophoresis and photography, capillary transfer of the DNA to a nylon membrane (Hybond-N, Amersham) was performed essentially as described by Southern (1975) using 20x SSC as a transfer buffer. The filter was dried and the DNA cross-linked to the membrane by irradiation with ultraviolet light according to the manufacturer's instructions.
2.3.6 Northern (RNA) Blotting

RNA was glyoxalated (Thomas, 1980) by incubation in 33-50% GFM buffer (depending on sample volume) at 55°C for 15 minutes. 1/10 volume of loading buffer [40% (v/v) formamide, 50% (v/v) glycerol, 1x MOPS, 0.05% (w/v) bromophenol blue, 0.05% (w/v) xylene cyanol FF (BDH)] was added and the samples loaded on a 0.8 or 1% neutral agarose gel in 1x MOPS buffer. RNA size markers (BRL) were treated in an identical manner and loaded onto the gel. After electrophoresis, the RNA was transferred (without any pretreatment) to a Nylon membrane as described for Southern blotting and cross-linked by ultraviolet irradiation. To reverse the glyoxalation, the filter was baked at 80°C for 1 hr. After transfer, the marker RNA tracks were removed from the filter and stained in 0.04% (w/v) methylene blue, 500mM sodium acetate, pH5.6 for 1 minute, then destained in Q water until the RNA samples were clearly visible.

2.3.7 Hybridisation of DNA Probes to Membrane-immobilised Nucleic Acids

DNA

Prehybridisation was performed for a minimum of 2hr at 65°C in 3x SSC, 5x Denhardt's solution, 0.1% (w/v) SDS, 6% (w/v) PEG 6000 and 200µg/ml sheared, denatured salmon testis DNA (Sigma) as described (Varley et al., 1987). Hybridisation was performed in an identical solution to prehybridisation (with the exception of Denhardt's solution which was at 2x concentration) with 0.5ng/ml of 32P radiolabelled probe for 16hr at 65°C as described (Varley et al., 1987). Prehybridisation and hybridisation were performed in volumes of 20ml in sealed perspex chambers.

After hybridisation the filters were washed initially in repeated changes of 2x SSC, 0.1% (w/v) SDS at 65°C until no more label eluted from the filters (as detected using a hand-held mini monitor), then at the stringencies stated in individual figure legends (in varying concentrations of SSC containing 0.1% (w/v) SDS). Filters were blotted dry and autoradiographed at -70°C with intensifying screens.
or at room temperature, depending on the strength of signal. Three types of X-ray film were routinely used: Kodak X-Omat™ (Eastman Kodak Company, New York), Fuji RX™ (Fuji Photo Film Company, Japan) and Hyperfilm™-MP (Amersham).

**RNA**

Filters were prehybridised for 2hr at 45°C in 6x SSC, 50% deionised formamide, 5x Denhardt's solution, 0.1% (w/v) SDS and 250μg/ml denatured salmon testis DNA. Hybridisation was for 16hr at 45°C in 6x SSC, 50% deionised formamide, 1x Denhardt's solution, 0.1% (w/v) SDS, 250μg/ml denatured salmon testis DNA and 10% (w/v) dextran sulphate (Sigma) containing 0.5ng/ml of ^32P-radiolabelled probe. Prehybridisation and hybridisation were performed in volumes of 20ml in sealed perspex chambers. Filters were washed initially in 2x SSC, 0.1% (w/v) SDS at 45°C until no more label eluted from the filters, then at the stringencies stated in individual figure legends (in varying concentrations of SSC containing 0.1% (w/v) SDS). Filters were blotted dry and autoradiographed in an identical manner to Southern blot filters.

**2.3.8 Removal of Probes and Re-use of Blots**

**DNA**

Probe sequences were removed by incubating the nylon filter in 0.4M NaOH at 45°C for 10 minutes followed by incubation in 0.1x SSC, 0.1% (w/v) SDS, 0.2M Tris-HCl (pH 7.5) at 45°C for 30 minutes according to the manufacturer's (Amersham) protocol. After probe removal the filter was prehybridised and hybridised as normal, without being allowed to dry out.

**RNA**

Probe sequences were removed by incubating for 1-2hr at 65°C in 5mM Tris-HCl (pH 8), 2mM EDTA, 0.1x Denhardt's solution. After probe removal the filter was prehybridised and hybridised as normal, without being allowed to dry out.
2.3.9 Preparation of \(^{32}\text{P}\)-Radiolabelled Probes

\(^{32}\text{P}\)-radiolabelled probes were generated by three different methods:

**Random Priming Method**

The method used was that of Feinberg and Vogelstein (1983). Labelling reactions were performed at either 37°C for 30 minutes or 20°C for 4-5hrs with 2u of Klenow polymerase (Amersham). New preparations of DNA were checked for efficiency of radionucleotide incorporation as follows: 1µl (approximately 1/100) of the probe mix was pipetted onto a 2.3cm circle of DE81 paper (Whatman), dried and Cerenkov counted in a Tri-Carb\textsuperscript® Minaxi-β 4000 series liquid scintillation counter (Packard Instruments, Downers Grove, Illinois). After washing off the unincorporated nucleotide with 0.5M Na\textsubscript{2}H\textsubscript{4}P\textsubscript{2}O\textsubscript{7}, the filter was dried and counted again. Typically 70-80% incorporation was seen, and probes with over 60% incorporation were used without further purification.

**End-labelling**

Oligonucleotides and short DNA fragments were labelled using T4 polynucleotide kinase-catalysed transfer of \(^{32}\text{P}\) from \(\gamma^{32}\text{P}\)ATP to the free 5' hydroxyl group of the nucleic acid (Maxam and Gilbert, 1980). The efficiency of transfer was estimated by chromatographic analysis of a small aliquot (approximately 1%) of the reaction mixture on DE81 paper. The \(\gamma^{32}\text{P}\)ATP migrates at the solvent (0.3M ammonium formate) front, whereas the radiolabelled DNA remains at the origin. Relative proportions of radioactivity in each fraction were estimated by autoradiography of the DE81 paper. Well over 90% incorporation of \(^{32}\text{P}\) into the DNA was regularly seen.

**Reverse Transcription of RNA**

Reverse transcription of total RNA using oligo dT as a primer was used to generate a probe for standardising the Northern blot containing RNA from different stages of *D. melanogaster* development.
(figure 3.3.2b). 24μg of total RNA (6μg from each stage of development represented) was recovered by ethanol precipitation and resuspended in 32μl of Q water. To this was added 4μl (1mg/ml) of oligo dT. After denaturing at 70°C for 1 minute and annealing at 42°C for 2 minutes, 12μl of 5x reverse transcriptase buffer 1E (250mM Tris/HCl (pH 8.3); 50mM MgCl₂; 50mM DTT; 0.5mg/ml BSA; 1mM dNTPs (except dGTP); 250mM KCl), 4.6μl of cold dCTP (33pmol/μl), 4 μl of α[^32P]-dCTP and 4μl of AMV RVT were added. After incubation at 42°C for 45 minutes, the incorporation was checked in an identical manner to that used for random-primed DNA probes. 14.5μl of 1M NaOH was added and the sample incubated at 70°C for 10 minutes. 72.5μl of 1M Tris/HCl (pH 6.8) was then added followed by 95.7μl of 5M ammonium acetate and 241μl of isopropanol. After thorough mixing, the sample was centrifuged at 13,000 rpm for 20 minutes and the supernatant aspirated. The pellet was then resuspended in 100μl of Q water.

2.3.10 Comparison of Band Intensities by Scanning Laser Densitometry

To compare expression levels of the bbcl gene in different stages of D. melanogaster development, the same blot was probed individually with the D. melanogaster bbcl sequence (figure 3.3.2a) and a total reverse transcribed cDNA probe (figure 3.3.2b). The autoradiographs obtained from these probings were analysed by laser densitometry to determine relative intensities for each track. Four separate readings of different parts of each track were taken using an LKB Bromma 2202 Ultrascan laser densitometer (Pharmacia) connected to an Apple II computer running the LKB 2190-0D1 Gelscan (v4.3) programme (Pharmacia). These individual results were merged to generate a single file representative of the entire track. After subtraction of the background signal, a curve was fitted to the data and the area under the curve calculated. This value represented the total optical density of the track and was used for comparisons between tracks.
2.3.11 Purification of DNA Fragments from Gels

The method used is essentially that of McDonell et al. (1977). Briefly, the desired fragment was excised from an agarose or polyacrylamide gel, placed inside dialysis tubing containing a small amount of 0.5x TAE (agarose gels) or 0.5x TBE (polyacrylamide gels) and electrophoresed at 150V for 15 minutes in a tank containing either 0.5x TAE or 0.5x TBE depending on gel type. The eluted DNA was then ethanol precipitated, dried and resuspended in a suitable volume of Q water.

2.3.12 Generation of DNA Molecules with Blunt Ends

Blunt ends were generated by removing protruding 3' termini and filling in protruding 5' termini. Protruding 5' termini were filled essentially according to the method of Wartell and Reznikoff (1980). Briefly, DNA was included in a fill-in reaction catalysed by Klenow polymerase (Amersham). After heat inactivation (70°C for 5 minutes) of the enzyme the DNA could be included directly in ligation reactions. Protruding 3' termini were removed using the 3'5' exonuclease activity of T4 DNA polymerase (Huang and Lehman, 1972). Briefly, following restriction endonuclease digestion, dNTPs were added to a concentration of 0.1mM. 1-2 units of T4 polymerase was then added and incubated at 12°C for 15 minutes followed by heat inactivating the enzyme (75°C for 10 minutes). The DNA could then be included directly in ligation reactions.

2.3.13 Ligation of DNA Molecules and Transformation of Competent E. coli Cells

Ligations were performed using T4 DNA ligase at 15°C. Standard reactions were performed in a 10μl volume with 10-100ng of vector and an appropriate amount of insert to give a 1:1 molar ratio of vector:insert. Occasionally, greater amounts of insert were used (up to a 10-fold excess) without problem. For ligation of molecules containing cohesive ends, T4 DNA ligase from Gibco BRL at 0.1u/μl final concentration was used; incubations being for 4-24hr. For ligation of molecules containing blunt ends, T4 DNA ligase from New
England Biolabs at 200u/μl (equivalent to 3 BRL u/μl) final concentration was used and incubations were for a minimum of 24hr.

*E. coli* cells were made competent for the uptake of DNA and transformed by a variation of the method used by Mandel and Higa (1970). Briefly, an overnight culture of bacteria was diluted 1:100 into 100ml of fresh medium and grown until mid log phase (A550 = 0.3-0.6), chilled on ice for 10 minutes and centrifuged to pellet the cells. The cells were resuspended in 50ml of 0.1M MgCl₂ and pelleted again. The cells were resuspended in 25ml of freshly diluted 0.1M CaCl₂ and incubated on ice for 20 minutes, followed by pelleting. Finally, the cells were resuspended in 5ml of 0.1M CaCl₂ and kept on ice until required. For freezing competent cells, the final pellet was resuspended in 5ml of 0.1M CaCl₂ containing 12.5% (v/v) glycerol, snap frozen in a dry ice/ethanol bath and stored at -70°C. These cells were thawed on ice when required.

Transformation of competent cells was performed as follows: 95μl of 1x SSC and 5μl of ligation reaction were added to 200μl of competent cells. After mixing, the cells were incubated on ice for 45 minutes, subjected to heat shock at 45°C for 2 minutes then incubated on ice for 30 minutes. Subsequent steps depended on the type of DNA being transformed. For plasmid DNA, 1ml of Luria broth was added to the cells which were then incubated at 37°C for 30 minutes prior to plating on Luria-plates containing 50μg/ml ampicillin. The plates were allowed to dry, inverted and incubated overnight at 37°C. If colour selection by use of the chromogenic substrate X-gal was being used, the plates also had 80μl of stock X-gal and 50μl of stock IPTG spread on their surface before addition of bacteria. For M13 DNA, the transformation mix was added to 2.5ml of BBL top agar (45°C), 200μl of lawn cells (which form a confluent layer of cells in which bacteriophage plaques form), 10μl of IPTG and 50μl of X-gal. This was then poured evenly onto a H plate, allowed to set, inverted and incubated overnight at 37°C.
2.3.14 Screening Bacteriophage Plaques and Bacterial Colonies using DNA Probes

Bacteriophage Plaques

The method used is based on that of Benton and Davis (1977). Briefly, bacteriophage were plated (Luria agar) at a suitable density in 0.8% top agar and grown overnight. After chilling the plate to help prevent the top agar from becoming detached from the H agar beneath, a nylon filter (Amersham) was placed on the agar surface, orientated with 3 asymmetric marks and left for 2 minutes. The filter was peeled off and transferred (plaque-side up) to 3MM paper (Whatman) soaked in Southern denaturing solution for 5 minutes, then transferred to 3MM paper soaked in 1M Tris-HCl (pH 7.2) for 1 minute. Finally the filter was transferred to colony neutralising solution for 5 minutes, blotted and dried at 65°C for 5-10 minutes. After irradiation with ultraviolet light to cross-link DNA to the membrane, bacterial debris was removed by gently rubbing the filter in 1x SSC. The filter was then probed as described for Southern blots.

Bacterial Colonies

The method used is based on that of Grunstein and Hogness (1975). Briefly, a gridded nylon filter (orientated) was placed on an agar plate and individual colonies picked onto the membrane using sterile tooth-picks (the same colonies were also replicated onto a master plate). The plate was then inverted and grown at 37°C until the colonies became visible. At this time the filter was peeled off the plate and transferred to 3MM paper soaked in Southern denaturing solution. After 5 minutes, the filter was transferred to 3MM paper soaked in 1M Tris-HCl (pH 7.2) for 1 minute prior to transfer to 3MM paper soaked in colony neutralising solution for 5 minutes. The filter was then blotted and dried at 65°C for 5-10 minutes. After irradiation with ultraviolet light to cross-link DNA to the membrane, bacterial debris was removed by gently rubbing the filter in 1x SSC. The filter was then probed as described for Southern blots.
2.3.15 Nucleic Acid Sequencing

Methods used for the sequencing of single-stranded and double-stranded DNA templates and RNA templates are given below:

Single-Stranded Bacteriophage M13 DNA Templates

The method used was based on that of Sanger et al. (1977). Briefly, 5μl of template DNA was added to 5μl of primer mix (1μl of primer (2.5ng/μl), 1μl of TM, 3μl of Q water), heated to 65°C for 10 minutes and slowly cooled to room temperature for 15 minutes. 2μl of the annealed primer/template was pipetted into each of 4 tubes containing 2μl of G, A, T or C reaction mixes, followed by 2μl of the enzyme/label mix (0.8μl of [α-35S] dATP (410 Ci/mmol), 0.73μl of klenow fragment (Amersham) and 7.6μl of Q water). The reaction was incubated at 37°C for 20 minutes then 2μl of chase (0.25mM of each dNTP) was added to each tube. After a further 20 minutes at 37°C, the reactions were terminated by addition of 2μl of loading dye (98% (v/v) formamide, 10mM EDTA, 0.1% (w/v) bromophenol blue, 0.1% (w/v) xylene cyanol FF). After boiling and chilling on ice, the samples were loaded on a denaturing 6% polyacrylamide gel. For reading sequence less than 50nt away from the position of the primer, non-gradient gels were used, whereas gradient gels were used to obtain sequence up to 300nt away from the primer. After electrophoresis, the gel was vacuum dried at 80°C (Bio-Rad model 1125B dual temperature slab gel dryer) without fixing and autoradiographed at room temperature overnight.

Sequenase® modified T7 polymerase (United States Biochemical, Cleviland) was also used for sequencing M13 single stranded DNA templates. The kit supplied was used according to the manufacturer's instructions.

Double-Stranded Plasmid DNA Templates

Again the protocol is based on Sanger's chain termination method and is modified from that of Tabor and Richardson (1987). A T7 sequencing kit (Pharmacia) was used according to the manufacturer's
instructions. Plasmid DNA from large-scale and small-scale preparations was used with equal success. Sequenase® was used on several occasions to sequence plasmid DNA; the manufacturer's protocol was followed.

Direct Sequencing of PCR Products

The method used was that of Green et al. (1989) as modified by Dr. D. Ogilvie (ICI Pharmaceuticals, Alderley Edge). PCR products were electro-eluted from preparative agarose gels and ethanol-precipitated. 50-100ng of template DNA in 1µl of Q water was added to 6µl of primer premix [5µl of primer (100ng/µl), 18µl of 5x Sequenase® buffer, 25µl of TE, 6µl of DMSO (Fisons)], boiled for 3 minutes and snap frozen at -70°C. 4µl of enzyme/label mix (4µl of [α³⁵S] dATP (400 Ci/mmol), 8µl of 0.1M DTT, 3µl of DMSO, 19µl of TE, 1µl of Sequenase®) was added to each annealed template, then 2µl of this added to each of 4 tubes containing 2µl of G, A, T or C mix (these have the same composition as the termination mixes supplied with the Sequenase® kit, except the A mix which was 80µM each of dCTP, dGTP and dTTP and 0.032µM ddATP). After incubating at 37°C for 5 minutes, 2µl of chase (0.25mM each dNTP, 10% (v/v) DMSO) was added and incubation continued for a further 5 minutes. 2µl of loading dye was added to each tube and the samples loaded on a non-gradient denaturing 6% polyacrylamide gel after boiling for 3 minutes and chilling on ice.

Sequencing of RNA Templates

Total RNA was used as a template for RNA sequencing by the method of Smith et al. (1988). Briefly, 8ng of ³²P-radiolabelled primer (see section 2.3.17) was added to 5µg of D. melanogaster total RNA. After ethanol precipitation, the nucleic acid was resuspended in 10µl of hybridisation buffer (50mM Tris-HCl (pH 8.3), 100µM EDTA, 10mM DTT, 7.5mM MgCl₂ and 10mM NaCl) in a capped tube. The sample was incubated at 100°C for 1 minute, then 60°C for 20 minutes. Four aliquots (2.5µl each) were incubated at 42°C for 15 minutes in reactions containing 5 units of RNAsin (BRL), 7 units of AMV RVT (Life Sciences Inc.), 0.1mM each dNTP and either 75µM ddATP, 60µM
ddCTP, 30µM ddGTP or 150µM ddTTP. 2µl of formamide loading dye (see above) was added to each sample which was then heated to 100°C for 3 minutes prior to loading on a non-gradient, denaturing, 7% acrylamide gel. After electrophoresis, the gel was dried and autoradiographed at -70°C with intensifying screens.

2.3.16 PCR Methods

PCR was used for amplifying sequences from DNA/cDNA in several protocols. Amplifications were performed using a Perkin Elmer Cetus DNA thermal cycler (model No. N8010177, Perkin Elmer Corporation, Norwalk). Buffers used were either those supplied with the Taq DNA polymerase or the 7.4x buffer recommended by Anglian Biotec Ltd. (500mM Tris-HCl (pH 8.8), 120mM ammonium sulphate, 50mM MgCl₂, 75mM 2-mercaptoethanol, 50µM EDTA, 11mM each dNTP, 1.3mg/ml BSA). Reaction volumes were 50 or 100µl containing 2.5µl of Jag polymerase and overlayed with 50µl of light mineral oil (Fisons). Primers were present at 1 or 0.5µM and template DNA at 10-300ng.

Amplification of DNA Sequences using two Specific Primers

The reactions were set up according to the protocol supplied with the DNA thermal cycler and amplification performed using 30 cycles of: 94°C melt for 1 minute, anneal for 1 minute and extension at 72°C for 3 minutes. The final extension was for 10 minutes, followed by a 4°C soak. In addition, an initial 94°C melt for 4 minutes was used to ensure all the template DNA was in a single-stranded form prior to the first round of PCR. Annealing temperatures were usually 5°C below the calculated melting temperature of the oligonucleotide primers. Melting temperatures for oligonucleotides up to 20nt in length were calculated according to the following formula:

\[ Tm(°C)=4(G+C)+2(A+T) \]

where G, C, A and T are the number of respective bases in the oligonucleotide (Wallace et al., 1979). PCR experiments which generated background amplification products under the above
conditions were repeated using a modified "touchdown" PCR programme (Don et al., 1991). In this programme the annealing temperature dropped 2°C every second cycle from 70°C to 52°C, with a final 20 cycles at 50°C annealing temperature. All other parameters were identical to non-touchdown programmes.

For PCR from somatic cell hybrid DNAs, BIOSMAP™ cell hybrid DNAs (BIOS Corporation, New Haven) were used according to the supplier's instructions. Amplification was performed using the touchdown programme. Concordancy was calculated as follows: for each chromosome, the PCR result was scored according to whether it agreed (concordant) or disagreed (discordant) with the expected result. For example (see table 4.5.1 for PCR results), chromosome one is present in two somatic cell hybrids (867 and 937). However, neither of these hybrids generated a PCR product (two discordant results). Also, the two somatic cell hybrids that did generate PCR products (967 and 1079) did not contain chromosome one (again two discordant results). All other results were concordant (the hybrids did not contain chromosome one and did not generate PCR products). Hence, 21 out of 25 results (84%) were concordant, giving the concordancy value for chromosome one.

Amplification of cDNA Sequences using RACE Primers

The amplification of cDNA sequences using only one specific primer was performed according to the method described by Frohman, (1990). Different protocols were used to amplify cDNA 3' and 5' ends.

Amplification of cDNA 3' Ends

This protocol is summarised in figure 3.3.3 (see chapter 3). Briefly, 10μg of Drosophila total RNA and 10pmoles of RACE1 primer were mixed with 4μl of 5x reverse transcriptase (RVT) buffer (see below) in a final volume of 19μl and heated in a capped tube to 85°C for 5 minutes. The tube was then allowed to cool to room temperature for 20 minutes, followed by reheating to 56°C for 5 minutes. After cooling on ice for 5 minutes, 14 units of avian myeloblastosis virus reverse transcriptase (AMV RVT, Life Sciences) were added and,
after mixing, the whole incubated at 42°C for 45 minutes. 20μl of 5M ammonium acetate was then added, followed by 40μl of isopropanol. After thorough mixing, the tube was centrifuged in an MSE microfuge at 13,000 rpm for 15 minutes. After removal of the supernatant, the pellet was washed with 100μl of 70% ethanol, air-dried and resuspended in 20μl of Q water. 5μl of this was used in a 50μl PCR reaction containing 0.5μM concentrations of primers RACE2 and 5003' (see figure 2.1.1 for the position of primer 5003' in the Drosophila cDNA sequence), 2.5U of Taq polymerase and 1x Taq polymerase buffer. This reaction was amplified for 30 cycles using an annealing temperature of 52°C. The PCR product was digested with SalI, which recognises a site within the RACE2 primer and a site within the Drosophila sequence downstream of the 5003' primer (see figure 2.1.1), and cloned into the SalI site of M13mp18 prior to sequencing.

Amplification of cDNA 5' Ends

This protocol is summarised in figure 3.3.4 (see chapter 3). Briefly, 10μg of Drosophila total RNA was mixed with 10pmoles of either 5005' or 3605' primers and 4μl of 5x RVT buffer (see below) in a final volume of 19μl and heated in a capped tube to 85°C for 5 minutes. After cooling to room temperature for 20 minutes the tube was re-heated to 56°C for 5 minutes. After cooling on ice for 5 minutes, 14U of AMV RVT were added and, after mixing, the whole heated to 42°C for 1hr, followed by 52°C for 30 minutes. 7.5μl of 2M NaOH and 2.5μl of Q water were added and the whole incubated at 70°C for 30 minutes. 60μl of ethanol was then added and, after mixing, the tube was centrifuged in an MSE microfuge at 13,000 rpm for 15 minutes. After removal of the supernatant, the pellet was washed with 100μl of 70% ethanol and air-dried.

The pellet was resuspended in 19μl of Q water, to which was added 6μl of 5x tailing buffer (BRL), 1μl of 1mg/ml BSA, 2μl of 100mM MgCl2 and 4μl of 1mM dATP. After incubating at 56°C for 5 minutes and cooling on ice for 2 minutes, 20U of terminal deoxynucleotidyl transferase (Pharmacia) was added and incubated at 37°C for 30 minutes. 20μl of 5M ammonium acetate was added, followed by 50μl of isopropanol. After mixing, the tube was centrifuged in an MSE
microfuge at 13,000 rpm for 15 minutes. After removal of the supernatant, the pellet was washed with 100μl of 70% ethanol and air-dried. The pellet was resuspended in 20μl of Q water and 5μl of this used in a 50μl PCR reaction containing 0.5μM concentrations of primers 3605' and RACE1, 2.5U of Taq polymerase and 1x Taq polymerase buffer. This reaction was amplified for 30 cycles with an annealing temperature of 48°C. After amplification, the products were analysed by agarose gel electrophoresis.

2.3.17 Primer Extension

Primer extension analysis of total D. melanogaster RNA was performed essentially as described by Sambrook et al. (1989). The 58bp primer used in the experiment was obtained by the following method: A region of D. melanogaster cDNA clone 4 encompassing the translation start site of the longest ORF was amplified by PCR using primers 3605' and Y5A (see figure 2.1.1). This product was digested with restriction enzyme Hinfl to generate three DNA fragments of 103, 6 and 58bp. These DNA fragments were separated by electrophoresis through a 6% non-denaturing polyacrylamide gel (see section 2.3.4), the 58bp fragment excised and purified by electrophoresis (see section 2.3.11). 8ng of this DNA fragment was radiolabelled using [γ32P] ATP according to the method given in section 2.3.9. The radiolabelled double-stranded primer sequence was mixed with 10μg of D. melanogaster total RNA and ethanol-precipitated. The pellet was resuspended in 30μl of hybridisation buffer (40mM PIPES (pH 6.4), 80% v/v deionised formamide, 400mM NaCl and 1mM EDTA), heated to 85°C for 10 minutes, then incubated at 45°C overnight. After recovering the nucleic acid by ethanol-precipitation, the primer was extended with 14 units of AMV RVT (Life Sciences Inc.) at 42°C for 40 minutes in 20μl of the following buffer: 50mM Tris-HCl (pH 8.1), 148mM KCl, 8mM MgCl2, 150μM each dNTP, 10mM DTT, 100μg/ml BSA containing 10 units of RNAsin (BRL). 1μl each of 500mM EDTA (pH 8) and RNase (5μg/ml) were added followed by incubation at room temperature for 90 minutes. 150μl of TE containing 100mM NaCl and 200μl of PCI were added and the phases mixed by vortexing. After spinning at 13,000 rpm for 5 minutes in an MSE microfuge, the aqueous phase was removed and
nucleic acid precipitated by addition of 500 µl of ethanol. After washing with 70% ethanol and air-drying, the pellet was resuspended in 5 µl of 10 mM Tris-HCl (pH 7.5) and 5 µl of formamide loading dye added. Following heating to 100°C for 3 minutes, 2 µl of the sample was loaded onto a denaturing non-gradient 7.5% polyacrylamide gel using M13 sequencing reactions as size markers. After electrophoresis, the gel was dried down and autoradiographed overnight at room temperature.

2.3.18 In Situ Hybridisation Studies

Hybridisation to D. melanogaster Polytene Chromosomes

Polytene chromosomes from D. melanogaster (wild type strain Brighton) were prepared and hybridised with a 3H-labelled DNA probe as described by Pardue (1986). The probe was generated by random-primed synthesis (Feinberg and Vogelstein, 1983) using a 620 bp EcoRI-SphI fragment from D. melanogaster cDNA clone 6 (see figures 2.2.1 and 3.2.1) as a template. [α²H]dCTP and [α²H]dATP radiolabelled nucleotides were incorporated in the probe. Staining was with Giemsa stain according to Pardue (1986).

Hybridisation to Tissue Sections

7µM sections of wax embedded D. melanogaster embryos (wild-type strain Sevelin) were purchased from Novagen (via AMS biotechnology, Burford) and hybridised with 35S-radiolabelled RNA probes as previously described (Senior et al., 1988). The RNA probes were generated by in vitro transcription of the 620 bp EcoRI-SphI fragment of D. melanogaster cDNA clone 6 (see figure 3.2.1) that had been inserted into the polylinker region of pBluescript SK+. Sense and anti-sense probes were generated from the T7 and T3 promoters respectively.

Hybridisation to Whole D. melanogaster Embryos

DIG-labelled transcripts were generated using a RNA-labelling kit.
The template DNAs were as used for 35S-labelled transcripts. Products were separated by agarose gel electrophoresis against RNAs supplied in the kit, stained with ethidium bromide and visualised by ultraviolet light. This allowed the quality and an estimate of the quantity of transcripts produced to be determined. Incorporation of DIG into the transcripts was demonstrated by the following method. 1 μl of transcription product was mixed with 3 μl of denaturing solution (20 parts formamide, 7 parts formaldehyde, 2 parts 20x SSC, 1 part Q water) and denatured at 70°C for 15 minutes. After cooling, 8 μl of 20x SSC was added and 2 μl of this was dotted onto a piece of nitrocellulose pre-wetted with 20x SSC. Control unlabelled and DIG-labelled RNAs supplied with the kit were treated in the same manner. The nitrocellulose was washed for 5 minutes in buffer 1 (100mM Tris/HCl pH 7.5; 150mM NaCl), then blocked for 5 minutes in buffer 1 containing 1mg/ml BSA. After washing 2x 2 minutes in buffer 1, the nitrocellulose was incubated with anti-DIG antibody conjugate (1:5000) for 15 minutes in buffer 1. The filter was washed 2x 5 minutes in buffer 1, rinsed in stain buffer (100mM NaCl; 50mM MgCl2; 100mM Tris/HCl pH 9.5; 0.1% Tween 20) and developed in stain buffer containing 34μg/ml NBT and 175μg/ml X-phosphate. After colour development, the filter was washed in water, then in ethanol before drying.

Hybridisation was performed essentially according to the method of Tautz and Pfeifle (1989) with some modifications. Embryos were collected on grape plates (see section 2.7.3) washed well with PBS and dechorionated in 5% sodium hypochlorite (Aldrich) for 3-5 minutes. After washing well with Q water, the embryos were fixed for 20 minutes in equal volumes of n-heptane (Fisons) and 4% formaldehyde (freshly diluted in PBS). The lower phase was removed and replaced with an equal volume of methanol. After gentle mixing for 2-3 minutes, devitellinised embryos were removed from the bottom of the tube (any embryos still at the interface were discarded). After washing the embryos in methanol, they were either stored at -20°C or rehydrated through a graded methanol:PBS series (9:1, 7:3, 5:5, 3:7 methanol:PBS; 2 minutes in each). After a further 20 minute fixation in 4% formaldehyde/PBS and wash in PBS, the
embryos were either dehydrated through a graded ethanol series (30%, 50%, 70% ethanol; 10 minutes each) and stored at -20°C, or taken through the steps detailed below (i.e. no dehydration). Embryos that were stored in 70% ethanol at -20°C were first rehydrated through the reverse ethanol series before being taken through the following steps.

Prior to prehybridisation, the embryos were washed 3 times in PTW (PBS, 0.01% Tween 20) for 2 minutes each, followed by a rinse in 1:1 PTW:hybridisation buffer (hybridisation buffer is: 5ml deionised formamide; 2.5ml 20xSSC; 20μl tRNA (50mg/ml); 5μl heparin (100mg/ml); 10μl Tween 20; pH adjusted to 4.5 with 1M citric acid and volume adjusted to 10ml with DEPC treated Q water). Prehybridisation was for 1-2hr in hybridisation buffer at 67°C.

Hybridisation was performed overnight in 150-200μl of hybridisation buffer at 67°C in 0.5ml microfuge tubes with heat denatured probe (95°C for 3 minutes). DIG-labelled RNA probes were used at dilutions in the range 1/1000-1/5000. Two tubes of embryos contained only hybridisation buffer (see below). After hybridisation, the embryos were washed at 67°C in hybridisation buffer (20 minutes) then in 1:1 PTW:hybridisation buffer (20 minutes). Finally, the embryos were washed for 5x 20 minutes at room temperature in PTW and blocked for 1hr in 1mg/ml BSA/PTW at room temperature.

The signal was detected using a Boehringer DIG detection kit. The anti-DIG antibody conjugate was diluted 1/400 in PTW (1ml final volume) and preabsorbed for 1hr with 50μl of fixed embryos. The antibody was then diluted 1:5 in PTW (final dilution 1:20000) and added to the embryos (one of the two tubes of embryos not hybridised with RNA probes was incubated with antibody. The embryos in these two tubes served as controls for endogenous phosphatase activity within the embryos and the presence of non-specific antibody binding). After 1hr at room temperature, the embryos were washed 4x 20 minutes in PTW and rinsed twice in staining buffer. Embryos were resuspended in staining buffer containing 34μg/ml NBT and 175μg/ml X-phosphate and the reaction allowed to proceed for 15 minutes. The reaction was stopped by rinsing the embryos several
times in PTW/10mM EDTA.

Embryos were dehydrated in a graded ethanol series (30%, 60%, 80%) to 100% ethanol (5 minutes each). The 100% ethanol was changed twice and then removed as completely as possible. Embryos were resuspended in 200μl of xylene and the tube immediately layed on its side for 5 minutes (this avoided clumping of the embryos). The xylene was then changed and the embryos transferred to a clean microscope slide. After removing as much xylene from the embryos as possible (but without allowing them to dry out), a few drops of DPX mountant were added and mixed into the embryos. A coverslip was then carefully placed onto the preparation and allowed to settle by gravity (i.e. no external pressure was applied to the coverslip).

2.4 Protein Methods

2.4.1 Solutions used During the Handling of Proteins

Acrylamide (30%): 87.6g of acrylamide (Serva or National Diagnostics) and 2.4g of N,N'-methylene bisacrylamide (Serva) were dissolved in 300ml of Q water and stored in the dark at 4°C.

Bradford's Reagent: 100mg of Coomassie brilliant blue G-250 (Sigma) was dissolved in 50ml of 95% ethanol and added to 100ml of 85% orthophosphoric acid (Fisons). The volume was adjusted to 1L and the solution filtered through Whatman® No.1 filter paper. The solution was kept in the dark at room temperature.

Coomassie Blue Stain: 1g of Coomassie brilliant blue R-250 (Sigma) was dissolved in 1L of 40% methanol (Fisons), 10% acetic acid (Fisons) to form a 0.1% solution.

Electrode Buffer (5x): 15g of Trizma base, 72g of glycine (Sigma) and 5g of SDS were dissolved in 1L of Q water and stored at 4°C.

METPBS: 16mM Na₂HPO₄, 4mM NaH₂PO₄, 150mM NaCl, 5mM EDTA, 1% (v/v) Triton X-100; filter sterilised and stored at 4°C.
PBS: Dulbecco’s modified PBS (without Mg²⁺ or Ca²⁺) was prepared by dissolving one tablet (ICN Flow) in 100ml of Q water. Following sterilisation by autoclaving, the solution was stored at room temperature.

PBS/Milk: Nonfat dried skimmed milk (Gateway) was dissolved at 5% (w/v) in PBS.

PBS/Milk/Tween (PMT): Nonfat dried skimmed milk (Gateway) was dissolved at 5% (w/v) in PBS/tween.

PBS/Tween: 100μl of Tween 20 (Polyoxyethylene sorbitan monolaurate, Sigma) was dissolved in 1L of sterile PBS and stored at room temperature.

PMSF (100mM): 17.4mg of PMSF was dissolved in 1ml of isopropanol and stored at -20°C.

PNE: 16mM Na₂HPO₄, 4mM NaH₂PO₄, 150mM NaCl, 5mM EDTA; filter sterilised and stored at 4°C.

Ponceau S Stain: 0.5M Ponceau S stain in 5% TCA was obtained from BDH and used as supplied.

Sample Buffer: A stock solution was kept at room temperature (4.2ml of Q water, 1ml of 0.5M Tris-HCl (pH 6.8), 0.8ml of glycerol, 1.6ml of 10% (w/v) SDS and 15mg of bromophenol blue). Immediately prior to use, 2-mercaptoethanol was added to an aliquot of the stock at a concentration of 5% (v/v).

Sodium Hydrogen Carbonate (0.9M, pH 10.8): 7.58g of NaHCO₃ (BDH) was dissolved in about 80ml of Q water. The pH was adjusted to 10.8 with 1M NaOH and the volume increased to 100ml. After filter sterilisation, the solution was stored at room temperature.

Tris-HCl (1.5M, pH 8.8): 27.23g of Trizma base was dissolved in about 100ml of Q water. After adjusting the pH to 8.8 with 1M HCl,
the volume was increased to 150ml with Q water. Following sterilisation by autoclaving, the solution was stored at room temperature.

Tris-HCl (0.5M, pH 6.8): 6.05g of Trizma base was dissolved in about 60ml of Q water. The pH was adjusted to 6.8 with 1M HCl, then the volume increased to 100ml. Following sterilisation by autoclaving, the solution was stored at room temperature.

2.4.2 Preparation of Soluble Proteins from Tissues

Tissues were disrupted by one of three methods depending on the type of tissue.

Homogenisation using a Polytron Tissue Homogeniser

A Polytron tissue homogeniser (Northern media, Nottingham) was utilised for D. melanogaster larvae, pupae and adults. Tissues were disrupted on ice in 1-2ml of PNE containing 0.6mM PMSF for 1-5 minutes. SDS (0.05% (w/v) final concentration) was added and gentle mixing continued for 1-2 minutes. The homogenate was then centrifuged at 13,000g for 20 minutes at 4°C to remove insoluble material. After determining protein concentration by the method of Bradford (1976), the protein preparation was frozen in aliquots at -20°C.

Sonication

Sonication using a soniprep ultrasonic disintegrater (MSE) was used to disrupt D. melanogaster embryos. About 50-100 embryos were resuspended in 100-200μl of PNE containing 0.6mM PMSF and sonicated on ice for 30 seconds. 0.05% (w/v) SDS was then added and incubated on ice for 10 minutes prior to centrifugation at 13,000g for 20 minutes at 4°C to remove insoluble material. After determining protein concentration by the method of Bradford, the protein preparation was frozen in aliquots at -20°C.
Homogenisation using a Glass Homogeniser

Homogenisation of human placental tissue and tissue culture cells was performed using a 10ml glass homogeniser. The placental tissue was also powdered under liquid nitrogen in a mortar and pestle prior to homogenisation. Extraction was for 2-3 minutes on ice in 1-2ml of PNE containing 0.6mM PMSF. 0.05% (w/v) SDS was added and incubation on ice continued for 5 minutes, followed by centrifugation at 13,000g for 20 minutes at 4°C to remove insoluble material. After determining protein concentration by the method of Bradford, the protein preparation was frozen in aliquots at -20°C. The homogenisation of human tissue was performed in a class II microbiological safety cabinet (Medical Air Technology Ltd., Manchester).

2.4.3 Purification of GST-Fusion Protein Produced in E. coli

The method used was modified from that of Smith and Johnson (1998). An overnight culture of E. coli strain Y1090 containing either the parental or recombinant pGEX plasmid was diluted 1/100 into Luria broth containing 50μg/ml ampicillin and grown until an A600 of 0.5 was obtained. IPTG was added to a final concentration of 0.1mM and incubation continued for a further 1hr. The cells were pelleted and resuspended in 1ml of METPBS. Sonication on ice in an MSE Soniprep 150 probe sonicator (amplitude 20 microns) was for a total time of 30-60 seconds in 10 second pulses with 1 minute intervals. The lysate was centrifuged (5 minutes at 13,000g) and the supernatant diluted 1:1 with METPBS prior to addition to 500μl of glutathione-agarose beads (Sigma). After mixing gently at room temperature for 5 minutes, the beads were collected by brief centrifugation and washed 3 times with 20 bead volumes of PBS. The fusion protein was eluted from the beads by incubation for 5 minutes at room temperature in 500μl of reduced glutathione (freshly prepared from powder, Sigma) at 1.5mg/ml in 50mM Tris-HCl (pH 8). After brief centrifugation the supernatant was removed and the beads incubated with another 500μl of 1.5mg/ml reduced glutathione. The supernatants were assayed for protein using the method of
Bradford and stored in aliquots at -20°C. Re-use of the glutathione beads was possible after removing remaining protein by incubation in 0.9M NaHCO₃ (pH 10.8) for 15 minutes followed by extensive washing in PBS. Beads were stored in PBS containing 0.2% (w/v) sodium azide (Sigma).

2.4.4 Estimation of Protein Concentration by the Method of Bradford

This method was first described by Bradford (1976) and was used with minor alterations for the quantification of protein preparations. A standardisation curve was created for each batch of reagent using known quantities of BSA in the range 0.5-200μg. Assays were performed as follows: the protein solution to be assayed was adjusted to a volume of 100μl in a 0.5ml microfuge tube. This was then added to 2ml of Bradford’s reagent in a 3ml plastic cuvette (Sarstedt Ltd., Leicester). After mixing and allowing the colour to develop for 5 minutes, the A₅₉₅ of the sample was measured relative to a blank of 100μl of Q water plus 2ml of Bradford’s reagent. Comparison of the absorbance obtained with the standard curve allowed an estimation of the protein concentration to be made.

2.4.5 Gel Electrophoresis of Proteins

Proteins were separated by electrophoresis through SDS-containing discontinuous polyacrylamide gels (SDS/PAGE) according to the method of Laemmli (1970). Two gel systems were used, the Mini-Protean® II and Protean® II slab cells manufactured by Bio-Rad. Both systems were used according to the manufacturer’s instructions employing separating gels of the percentages shown in individual figure legends. The Pharmacia low molecular weight calibration kit was used to estimate protein sizes.

2.4.6 Western (protein) Blotting

Proteins were transferred from SDS/PAGE gels to nitrocellulose membranes (Hybond-C, Amersham) by the method of Szewczyk and Kozloff (1985) using 25mM CAPSO (Sigma) pH 10 buffer containing
20% methanol (Fisons). Transfer was performed using either the Mini Trans-blot® cell or the Trans-blot® cell (both from Bio-Rad) at 0-4°C. The Mini Trans-blot® cell was operated at 400mA for 90 minutes and the Trans-blot® cell at 80V for 2-3hr. After transfer, gels were stained with Coomassie blue stain to assess transfer. The nitrocellulose membrane was stained with Ponceau S stain to visualise marker tracks and to determine the extent of transfer.

2.4.7 Detection of Antigens on Nitrocellulose Membranes

The method used is based on those of Towbin et al. (1979) and Burnette (1981). The nitrocellulose membrane containing the proteins was incubated in PMT for greater than 1hr at room temperature to block protein-binding sites on the membrane. The membrane was then transferred to PMT containing the primary antibody at a suitable dilution (see individual figure legends) and incubated with gentle agitation for 1hr at room temperature. After 3x 5 minute washes in PBS/Tween, the filter was transferred to PMT containing the secondary antibody (HRP-conjugated goat anti-rabbit IgG, Sigma) at 1/1000 dilution and incubated with gentle agitation for 1hr at room temperature. The membrane was then washed in PBS for 3x 5 minutes prior to development. The blot was developed using the chromogenic compound DAB (BDH) with cobalt enhancement: to 98ml of PBS was added to 1ml of 3% CoCl₂ (BDH) and 1ml of a freshly made 20mg/ml solution of DAB. After mixing, 40ml of this solution was transferred to a separate container and 20µl of 30% (v/v) H₂O₂ (Sigma) added. The nitrocellulose membrane was placed in this solution and incubated with gentle agitation until protein bands were clearly visible. The reaction was terminated by rinsing the membrane in water and drying.

2.4.8 Preparation of Frozen Embryo Sections

D. melanogaster embryos were collected as described in section 2.7.3 (below) and washed extensively in PBS. The embryos were then transferred to mounting medium (O.C.T. compound, Miles laboratories, Illinois) and placed at the bottom of a plastic mould. The specimen
was frozen in a dry ice/ethanol bath and stored at -20°C until sectioned. 10μM sections were cut on a Leitz model 1720 digital cryostat and mounted on subbed slides (prepared as described in Macgregor and Varley, (1988) by dipping clean, dry microscope slides (BDH) in a solution of 0.1% gelatin (Fisons) and 0.1% chrome alum and allowing to dry overnight before use). Sections were processed immediately or stored frozen at -20°C.

2.4.9 Detection of Antigens in Fresh-Frozen Tissue Sections

The method employed was essentially that used in the Breast Cancer Research Unit at Glenfield General Hospital, Leicester. Sections were fixed in 3.7% paraformaldehyde (in PBS) for 10 minutes at room temperature, then rinsed in PBS for 10 minutes. The sections were transferred to methanol (pre-cooled to -20°C) for 4 minutes, then to acetone (-20°C) for 2 minutes and finally to PBS at room temperature. The PBS was changed twice at 5 minute intervals. All subsequent incubations were performed in a chamber humidified with PBS at room temperature without allowing the sections to dry out. All sections were blocked with 3% BSA in PBS for 10 minutes. Excess solution was removed and replaced with primary antibody at dilutions of 1/10 to 1/200 in 3% BSA (for control sections, incubation in 3% BSA was continued). Incubation for 60 minutes with the primary antibody was followed by two 5 minute washes in PBS. The secondary antibody (horseradish peroxidase conjugated goat anti-rabbit IgG) at dilutions of 1/200 to 1/500 (in 3% BSA) was then added and incubated for 30 minutes. Sections were then washed twice in PBS prior to addition of DAB solution (6mg of DAB dissolved in 10ml of 0.05M Tris/HCl (pH 7.6) plus 10μl of 30% H₂O₂, filtered if a precipitate formed) and incubation for 15 to 30 minutes. After washing in tap water for 5 minutes, the sections were stained with haematoxylin (10 seconds) washed in tap water for 5 minutes and dehydrated through graded alcohols to xylene (95% alcohol for 2 minutes, 100% alcohol for 2x 2 minutes and xylene for 2 minutes). After transfer to fresh xylene, the sections were mounted in DPX mountant (BDH).
2.4.10 Immunoprecipitation of Radiolabelled *D. melanogaster* Proteins

*D. melanogaster* SL2 tissue culture cells (obtained from Dr. B. Fenton, Dept. of Biochemistry, University of Dundee) were metabolically labelled using TRAN$^{35}$S-LABEL$^\text{TM}$ (ICN Biomedicals) $^{35}$S-labelled methionine and cysteine, according to the method given in Harlow and Lane (1988). The labelled amino acids were used at 50μCi/ml final concentration in Schneider's *D. melanogaster* medium (deficient in methionine and cysteine; Gibco BRL) containing 10μM colchicine. Labelling was for 4hr at 25°C. Immunoprecipitation reactions were performed either according to the method given in Harlow and Lane (1988) or according to a method obtained from Dr. B. Fenton. The method of Dr. B. Fenton is as follows: Labelled cells were collected by centrifugation (400rpm for 5 minutes in a Heraeus digifuge) and washed in PBS. After pelleting again, the cells were resuspended in 500μl of TET (50mM Tris/HCl (pH 8); 1mM EDTA; 0.5% Triton X-100) containing 0.5mM PMSF. The cells were disrupted by syringing through a 21 gauge needle. The lysate was then incubated with an equal volume of 10% immobilised protein G (Sigma) for 1hr on ice. After centrifugation to remove the protein G, the supernatant was divided into 5 aliquots and 10μl of rabbit antiserum added to each aliquot. Incubation on ice was resumed for 1hr, then half the volume of 20% immobilised protein G added to each sample. After 1hr on ice, the protein G was collected by centrifugation and washed 3 times in TET (the first wash also contained 1M NaCl). The pellet was finally resuspended in 50μl of SDS sample buffer and boiled for 5 minutes. 10 and 20μl aliquots were then subjected to SDS/PAGE analysis as described above. After electrophoresis, the gels were stained to reveal marker tracks and immunoglobulin bands, then soaked in 1M sodium salicylate for 15 minutes prior to drying down onto 3MM paper (Whatman). Autoradiography was at -70°C.

2.5 Preparation of Polyclonal Antibodies

2.5.1 Preparation of Polyclonal Antisera in Rabbits

About 1mg of fusion protein was produced and purified as stated...
above. The sample volume was decreased by vacuum dessication prior to dialysis against PBS at 4°C for 8hr. The protein concentration was estimated by the method of Bradford and adjusted to 1mg/ml. The protein solution was then filter sterilised and dispensed into 100μl aliquots which were stored at -20°C. To check that the protein preparation was sterile, a 20μl aliquot was spread onto a plate of Luria agar and incubated overnight at 37°C. No bacterial colonies were seen.

The GST-fusion protein was used to immunise two New Zealand White rabbits. All injections were of 100μg of fusion protein in sterile PBS (with or without adjuvant). One rabbit (317) was given an initial subcutaneous injection of fusion protein with Freund's complete adjuvant (Freund and McDermott, 1942). The second rabbit (316) was given three intravenous (i.v.) injections of fusion protein (no adjuvant) at weekly intervals. Two weeks after the final i.v. injection, both rabbits received their first boost injection (i.v.), followed six weeks later by a second boost injection (i.v.). Twelve days after the second boost injection both rabbits were exsanguinated and serum prepared (Harlow and Lane, 1988). Serum was stored in small aliquots at -70°C and kept at 4°C once thawed.

2.5.2 Pre-absorption of Rabbit Antisera

To remove immunoglobulin molecules with affinities for proteins other than the bbcl protein, antisera were incubated with a nitrocellulose-immobilised extract of E. coli expressing the GST protein. The method was developed by Mr. R. Owen (Dept. of Microbiology, University of Leicester). A 500ml culture of E. coli strain DH5α containing the pGEX-3X plasmid was grown and induced to express GST protein as described above. The cells were pelleted, resuspended in 100ml of PBS and disrupted by sonication. Strips of nitrocellulose (2cm by 7cm) were incubated in the lysate for 2hr at room temperature, washed in PBS then fixed in Ponceau S stain for 1hr at room temperature. After destaining in PBS, the filters were transferred to PBS/milk for 1hr at room temperature, then washed in PBS. Three strips were added to 4ml of antiserum in a Nunc® universal and incubated for 2hr at 4°C. The strips were removed and the incubation repeated twice for 2hr and once overnight with new
strips of nitrocellulose. The antiserum was then dispensed into small aliquots and stored at -70°C until required. This method was also used to remove immunoglobulins capable of binding to the *D. melanogaster* bbcl protein by incubating antiserum with nitrocellulose strips containing an extract of *E. coli* DH5α expressing the GST-*D. melanogaster* bbcl fusion protein.

2.5.3 Preparation of Polyclonal Antibodies in Chickens

Synthetic peptides (see figure 6.6.1, chapter 6) were generated by Dr. A. Moir (Sheffield University). All the peptides contained an additional C-terminal cysteine residue to allow coupling to a carrier protein (see below). The amino acid composition of the peptides was determined by acid hydrolysis followed by HPLC and the complexity of the synthetic product analysed by HPLC.

Coupling of peptides to carrier proteins was performed essentially as described by Sambrook *et al.*, (1989) using the sulphydryl-reactive compound MBS (*m*-maleimidobenzoic acid *N*-hydroxysuccinimide ester). The proteins used as carriers were porcine thyroglobulin and BSA (both from Sigma).

The chickens were initially given 20μg of thyroglobulin/peptide conjugate intramuscularly with adjuvant essentially as described by Gassman *et al.*, (1990). Two chickens received peptide 174 and two peptide 476. The chickens were boosted with 10μg of conjugate (plus adjuvant) 13 and 20 days after the first injection and eggs were collected from day 25. IgY antibodies were isolated from the eggs by the method of Gassman *et al.*, (1990) and stored in small aliquots at -70°C.

2.6 In Vitro Cell Culture Methods

2.6.1 Solutions used During the Culture of Eukaryotic Cells

DMEM: DMEM (without sodium pyruvate, with 4500mg/L glucose) was purchased from Gibco BRL and stored at 4°C. The medium was
supplemented with 10% (v/v) FCS (Gibco BRL), 2mM glutamine (Gibco BRL) and antibiotic (100u/ml penicillin and 100µg/ml streptomycin, Gibco BRL).

Insulin (2mg/ml): 20mg of bovine pancreatic insulin (Sigma) was dissolved in 10ml of Q water (pH adjusted to 2 with HCl), filter sterilised and dispensed into 500µl aliquots prior to storage at -20°C.

PBS (tissue culture): Dulbecco's modified PBS (without Mg²⁺ or Ca²⁺) was prepared by dissolving ten tablets (ICN Flow) in 1000ml of Q water. The PBS was dispensed (250ml aliquots) into 500ml tissue culture bottles (Gibco BRL). Following sterilisation by autoclaving, the solution was stored at room temperature.

Schneider's Drosophila Medium: This medium was bought from Gibco BRL in 100ml aliquots and stored at 4°C. The medium was supplemented with 10% (v/v) FCS and antibiotics (100u/ml penicillin and 100µg/ml streptomycin, Gibco BRL).

Trypsin/EDTA: Trypsin/EDTA (1x) was purchased from Gibco BRL, dispensed into 20ml aliquots and stored at -20°C. Once thawed, the solution was kept at 4°C.

2.6.2 Cell Lines Cultured In Vitro

The cell lines used are listed in table 2.6.1. All the human cell lines are from breast carcinomas or metastases of breast carcinomas and were obtained from Dr. P. Rye (Breast Cancer Research Unit, Glenfield General Hospital, Leicester). The Drosophila cell lines were obtained from Dr. B. Fenton (Department of Biochemistry, University of Dundee).

2.6.3 General Cell Culture

Cell manipulations were performed in a class II microbiological safety cabinet (Medical Air Technology Ltd., Manchester). Cells were routinely cultured in 10ml of medium in 9cm Nunclon® (Gibco BRL).
<table>
<thead>
<tr>
<th>Name</th>
<th>Details</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCF7</td>
<td>Human epithelial-like adherent cell from breast adenocarcinoma, obtained from a pleural effusion</td>
<td>(Soule <em>et al.</em>, 1973)</td>
</tr>
<tr>
<td>He578T</td>
<td>Human epithelial-like adherent cell from a primary ductal breast carcinoma</td>
<td>(Hackett <em>et al.</em>, 1977)</td>
</tr>
<tr>
<td>T47D</td>
<td>Human epithelial-like adherent cell from an infiltrating ductal breast carcinoma, obtained from a pleural effusion</td>
<td>(Keydar <em>et al.</em>, 1979)</td>
</tr>
<tr>
<td>SL2</td>
<td><em>D. melanogaster</em> embryonic semi-adherent cell</td>
<td>(Schneider, 1972)</td>
</tr>
<tr>
<td>Kc</td>
<td><em>D. melanogaster</em> embryonic semi-adherent cell</td>
<td>(Echalier and Ohanessian, 1969)</td>
</tr>
</tbody>
</table>

Table 2.6.1 Cell lines cultured *in vitro*. See text for details of culture methods.
D. melanogaster cells were incubated at 25°C in a humidified incubator; human cells were incubated at 37°C (5% CO₂ atmosphere) in a humidified incubator.

2.6.4 Subculturering of Cells

Suspension Cells

D. melanogaster cell lines (suspension cells). The medium containing cells was centrifuged in a sterile 10ml cone base plastic tube (Northern Media Supply Ltd., North Humberside) for 5 minutes at 600rpm in a Heraeus Digifuge® (Heraeus Equipment Ltd., Brentwood). The medium was aspirated and the cells resuspended in 10ml of fresh medium. This was diluted as required (up to four-fold) into fresh medium in 9cm dishes (10ml per dish).

Adherent Cells

Human cells (adherent cells). The medium was aspirated and the cells washed twice with 10ml of PBS. 2ml of trypsin/EDTA was then added and incubated at room temperature until the cells could be dislodged from the plate by gentle tapping (about 5 minutes). 8ml of culture medium was added and the cells diluted as required (up to four-fold) into fresh medium in 9cm dishes (10ml per dish).

2.6.5 Cryogenic Storage of Cells

D. melanogaster cells from one 9cm dish were pelleted as above and resuspended in 2ml of medium containing 10% DMSO (Sigma, cell culture grade). 1ml aliquots were dispensed into 1.5ml ampoules and cooled slowly to -196°C before storing in liquid nitrogen in a model LR40 cryostorage refrigerator (Jencons (Scientific) Ltd., Leighton Buzzard). Human cells were treated with trypsin/EDTA as above, then 2ml of culture medium containing 20% DMSO was added. 1ml aliquots were dispensed into 1.5ml ampoules and cooled slowly to -196°C before storing in liquid nitrogen.

Growth of cells from frozen samples was performed as follows: a
vial of frozen cells was placed in a water bath at the normal culture temperature of the cells until thawed. The cells were then transferred to 9ml of culture medium in a 9cm dish and incubated overnight. The following day, the medium was changed for fresh medium and incubation continued.

2.7 Laboratory Culture of D. melanogaster

2.7.1 Media

Grape Plates: 10.75g of agar (BBL) was dissolved in 237ml of Q water by heating slowly (not allowed to boil). 245ml of red grape juice (Sainsbury’s) and 5ml of acid mix [52.5ml of Q water, 3ml of green food colouring (Pearce Duff and Co., Cambridge), 4.15ml of phosphoric acid (Fisons) and 41.8ml of propionic acid (BDH)] was added and, after mixing well, dispensed into round 5ml plastic culture plates (Nunclon). After cooling, the plates were kept at 4°C until used. Prior to use, 50μl of 10% acetic acid (Fisons) was spread over the surface of the plate and 50μl of a yeast suspension (Sainsbury’s) was dispensed onto the centre of the plate.

Oatmeal Medium: 130g of ground oatmeal, 6.5g of agar (Difco) and 62g of black treacle were added to water to a final volume of 1L and brought to the boil. After simmering for 20 minutes, 11ml of 10% Nipagin (bought as p-hydroxy-benzoic acid methyl ester from Sigma and dissolved in ethanol) was added and the medium dispensed into 1/2 pint milk bottles. After cooling, the medium was seeded with a few drops of yeast suspension, foam plugs were inserted in the top of the bottles and the medium was kept at 4°C until used.

Sugar Medium: 115g of autolysed dried yeast powder (Quest International Ltd., Wirral), 65g of D-glucose (Fisons) and 10g of agar (BBL) were added to water to a final volume of 1L and brought to the boil. After simmering for 10 minutes, 15ml of 20% Nipagin was added and the medium dispensed into 1/2 pint milk bottles. After cooling, the medium was seeded with a few drops of yeast suspension, foam plugs were inserted in the top of the bottles and
the medium was kept at 4°C until used.

2.7.2 Culture of *D. melanogaster*

*D. melanogaster* were cultured on either oatmeal or sugar medium at either 18°C or 25°C using a 12hr dark/light cycle. Stocks were transferred to new medium every 3-4 weeks.

2.7.3 Collection of Eggs

About 50-100 adult *D. melanogaster* were placed in a glass tube (5cm external diameter and 15cm long) sealed at one end by gauze and at the other by a foam plug. The glass tube was then placed onto a grape plate such that the gauze was in contact with the surface of the agar and left for varying lengths of time at 25°C. Embryos were collected from the plate by washing with PBS and then processed as required.
CHAPTER 3

Characterisation of the D. melanogaster bbcl cDNA and Predicted Protein Product

3.1 Introduction

As discussed in chapter one, the genomic organisation of mammalian bbcl sequences is rather complicated (see figure 1.3.3). However, sequences capable of hybridising with the human bbcl cDNA sequence were also found in DNA from lower eukaryotes. Generally these sequences produced simpler patterns of hybridisation on Southern blots than sequences within the DNA of higher eukaryotes (see figures 1.3.3 and 3.1.1a).

Southern blotting analysis of D. melanogaster genomic DNA cleaved separately with EcoRI and HindIII restriction endonucleases demonstrated a single strongly hybridising band when probed with the human bbcl cDNA sequence (Figure 3.1.1a). The faint band seen in track 2 may represent a cross-hybridising sequence, since this band is not observed when a similar blot is probed with a fragment from the D. melanogaster bbcl sequence (Figure 3.1.1b). However, it could also be due to slight leakage of sample from lane 1 into lane 2. This result contrasted markedly with that obtained when a similar experiment was performed with human genomic DNA (see chapter 1) and indicated that analysis of the D. melanogaster bbcl gene might prove simpler than further analysis of the human gene(s).

3.2 Isolation and Characterisation of D. melanogaster bbcl cDNAs

Ten hybridising clones were isolated from a λgt11 D. melanogaster embryo cDNA library by Dr. J. M. Varley using a 480bp Styl fragment from the human bbcl cDNA as a probe at low stringency. Cleavage of these cloned DNAs with EcoRI to release the cDNA insert(s) followed
Southern blotting analysis of *D. melanogaster* genomic DNA. a. 5μg of DNA was cleaved with EcoRI (lane 1) or HindIII (lane 2) restriction endonucleases and separated on a 0.8% agarose gel. After transfer to a nylon filter, the DNA was hybridised with a $^{32}$P-radiolabelled probe generated from the 466bp fragment obtained by PCR amplification of a region of the human *bbcl* cDNA using primers 10-4 and 10-9 (see figure 2.1.1 for the positions of the primers used for amplification). Final washing stringency was 2x SSC at 65°C and autoradiography was for 14 days at -70°C with Amersham Hyperfilm™-MP.

b. Separate blot containing 5μg quantities of *D. melanogaster* DNA treated exactly as described in a., except that the DNA was hybridised with a $^{32}$P-radiolabelled probe generated from the 620bp EcoRI-SphI fragment of *D. melanogaster* cDNA clone 6, washed at a final stringency of 2x SSC at 65°C and autoradiographed overnight at -70°C with Fuji RX™ film.

<table>
<thead>
<tr>
<th>Kb</th>
<th>a.</th>
<th>b.</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td><img src="image" alt="Southern blot a." /></td>
<td><img src="image" alt="Southern blot b." /></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
by agarose gel electrophoresis allowed the number and sizes of the
cDNA inserts to be determined (table 3.2.1). Southern blotting
analysis demonstrated that each clone contained a single hybridising
EcoRI fragment ranging in size from 530 to 960bp (table 3.2.1).

To determine whether the clones contained *D. melanogaster* bbcl
cDNA sequences, the hybridising EcoRI fragment from each cDNA
cloned was subcloned into the vector M13mp18 and single-stranded
DNA prepared. Sequencing from both ends of each clone revealed that
the cDNAs were homologous to the human *bbcl* cDNA within an open
reading frame (ORF) region which comprises most of the cDNA
sequence (except clones 1 & 11 which do not contain full-length
ORFs). Figure 3.2.1 shows a schematic representation of the cDNA
clones. It can be seen that the cDNAs fall into five groups with
respect to the sequences at either end of the clones:

Group 1: clones 2, 3, 9, 10 and 11.

The 5' and 3' ends of these clones fall within sequence which has
been confirmed in at least two independent clones.

Group 2: clones 7 and 8.

The 3' ends of both these clones fall within sequence which has been
confirmed in at least two independent clones. Sequences at the 5'
ends of clones 7 and 8 extend 95 and 115bp respectively beyond
sequence which is common to both clones. These two sequences
differ with respect to each other. These clones could represent
alternatively spliced variants of the *bbcl* message such as occurs in
the mRNA for mouse α-amylase-1a (Young et al., 1981). However,
further experiments indicate that this is unlikely to be the case and
that part or all of the sequences towards the 5' ends of these two
clones are cloning artefacts (section 3.3.4).


This clone contains sequence at the 3' end which extends 17bp beyond
the 3' end of clone 3. This sequence was shown to represent part of
Table 3.2.1

Confirmation by agarose gel electrophoresis of insert sizes (previously determined by Dr. J. M. Varley) within the λgt11 D. melanogaster cDNA clones. Inserts were released by cleavage of the cloned DNAs with EcoRI restriction endonuclease, which recognises the site used for cDNA cloning (see chapter 2 for details of the λgt11 cDNA library), and separated by electrophoresis on a 0.8% agarose gel. Sizes were determined by comparison with a size marker of λ DNA (BRL) cleaved with EcoRV restriction endonuclease. Inserts which hybridised with the 480bp Styl fragment from human bbcl cDNA clone C328-10 during Southern blotting analysis are shown in bold type.

<table>
<thead>
<tr>
<th>clone</th>
<th>sizes of insert bands (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><strong>530</strong></td>
</tr>
<tr>
<td>2</td>
<td>350, 750, 1200, 1400</td>
</tr>
<tr>
<td>3</td>
<td><strong>750</strong></td>
</tr>
<tr>
<td>4</td>
<td>720, 960</td>
</tr>
<tr>
<td>6</td>
<td><strong>780</strong></td>
</tr>
<tr>
<td>7</td>
<td><strong>820</strong></td>
</tr>
<tr>
<td>8</td>
<td>860, 2300</td>
</tr>
<tr>
<td>9</td>
<td><strong>700</strong></td>
</tr>
<tr>
<td>10</td>
<td>700, 1100, 1300</td>
</tr>
<tr>
<td>11</td>
<td><strong>880</strong></td>
</tr>
</tbody>
</table>
Figure 3.2.1

Comparison of the hybridising EcoRI fragments obtained from the cDNA clones isolated by Dr. J. M. Varley from the D. melanogaster embryo cDNA library in λgt11 (see text for details). Numbering is with respect to the adenylate residue of the initiation methionine codon of the longest open reading frame identified within the D. melanogaster cDNA sequence (see figure 3.3.6). All clones contain EcoRI sites at their 5' and 3' termini as a result of the cloning method (see chapter 2).
the $bbcl$ mRNA (see section 3.3.2, below).

Group 4: clone 4.

This clone contains 250bp of sequence at its 5' end which differs from that present in any other clone. The site of divergence is 17bp downstream of the position at which clones 7 and 8 diverge. As with clones 7 and 8, subsequent experiments (section 3.3.4) indicate that this sequence is unlikely to be an alternatively spliced exon and instead probably represents a cloning artefact.

Group 5: clone 1

This clone is the smallest cDNA isolated and contains an ORF truncated at the 5' end. 81bp of sequence at the 5' end of this cDNA differ from that found in all of the other cDNAs. Since the nucleotides around the divergence point do not resemble a consensus splice acceptor site (Shapiro and Senapathy, 1987) and the positions of introns within this part of the gene are known (see chapter 4, section 4.2), the variant sequence cannot represent part of an unspliced intron. The sequence is also unlikely to represent part or all of an alternatively spliced exon, since an in-frame translation termination codon exists within the sequence but no consensus translation initiation site (Cavener, 1987) is present. Instead, it appears that the sequence is another cloning artefact.

3.3 Analysis of the $D.~melanogaster~bbcl$ mRNA

3.3.1 Northern Blotting Analysis

In order to determine the size of the mRNA produced from the $D.~melanogaster~bbcl$ gene, Northern blotting analysis was performed on total RNA isolated from adult $D.~melanogaster$. After separation of the RNA by agarose gel electrophoresis and transfer to a nylon membrane, the RNA was hybridised with a $^{32}$P-radiolabelled random primed probe generated from the $Styl$ fragment of human cDNA clone C328-10. A single hybridising band was observed corresponding to a
RNA species of about 950nt (figure 3.3.1a). This RNA species is of very similar size to the major human placental \textit{bbcl} transcript (figure 3.3.1a). Subsequent hybridisation of this blot with a $^{32}$P-radiolabelled random primed probe generated from the 620bp \textit{EcoRI-SphI} fragment of \textit{D. melanogaster} cDNA clone 6 (figure 3.2.1) also identified a 950nt RNA species in the RNA isolated from adult \textit{D. melanogaster} (figure 3.3.1b).

It is somewhat surprising that the human \textit{bbcl} probe identified both human and \textit{D. melanogaster} RNA species (figure 3.3.1a), whereas the \textit{D. melanogaster} \textit{bbcl} probe only identified a \textit{D. melanogaster} transcript (figure 3.3.1b). The reason for this is unclear, given that the final washing stringency in both cases was the same, but possible reasons are given below. Loss of RNA from the filter during removal of the human \textit{bbcl} probe prior to re-hybridisation with a probe generated form the \textit{D. melanogaster} cDNA could make detection of the human \textit{bbcl} mRNA difficult. Alternatively, if the washing temperature was a little higher than stated, or if the washing buffer was a slightly different composition, this could result in removal of the \textit{Drosophila} \textit{bbcl} cDNA probe from the heterologous (human) mRNA, whilst leaving the probe hybridised with the homologous (\textit{Drosophila}) mRNA. It is also possible that a combination of the above factors may have resulted in the observed result.

The identification of a single hybridising band indicated that either a single RNA species or several RNAs of very similar sizes were hybridising with the probe. Due to the similarity in sizes of most of the \textit{D. melanogaster} cDNAs isolated it was not possible to determine from the Northern blot which, if any, represented the \textit{in vivo} transcript.

The Northern blotting analysis described above utilised RNA extracted from adult \textit{D. melanogaster}. Since the \textit{D. melanogaster} \textit{bbcl} cDNAs were isolated from an embryo cDNA library, it was of interest to determine whether the \textit{D. melanogaster} \textit{bbcl} gene was expressed at other stages of \textit{D. melanogaster} development. RNA was therefore isolated from 0-24hr embryos, third instar larvae and pupae. These RNA samples and the previously isolated RNA from adult \textit{D.}
Figure 3.3.1

Northern blotting analysis of total RNA from adult *D. melanogaster* (lane 1) and human placenta (lane 2). 10µg of each RNA was separated on a 1% agarose gel, then transferred to a nylon filter. Hybridisation was in 50% formamide at 45°C with either a $^{32}\text{P}$-radiolabelled 480bp *Styl* fragment from human cDNA clone C328-10 (panel a) or, after removal of the human probe, a $^{32}\text{P}$-radiolabelled 620bp *EcoRI-SphI* fragment from *D. melanogaster* cDNA clone 6 (panel b). Final washing stringency was 2x SSC at 65°C in both cases and exposures were 12 days at -70°C (panel a) and 7 days at -70°C (panel b), both using Amersham Hyperfilm™-MP.
melanogaster were separated by agarose gel electrophoresis and transferred to a nylon membrane by Northern blotting. Probing this blot with a D. melanogaster bbcl cDNA sequence gave the result shown in figure 3.3.2a. To control for the amount of RNA loaded in each track, the filter was stripped and rehybridised with a probe generated by reverse transcription of D. melanogaster total RNA (hereafter referred to as "total cDNA") as described in chapter 2 (figure 3.3.2b). Scanning laser densitometry then allowed quantification of the bbcl transcript level at each developmental stage (Table 3.3.1 and figure 3.3.2c).

From this result it can be seen that maximal expression of the bbcl transcript occurs during embryogenesis, with levels of the transcript being four times as high in embryonic tissue as in adult tissue. During larval development and pupation, the amount of expression gradually decreases so that pupae contain only 4% of the transcript levels seen in embryos. Transcription rises in adult D. melanogaster, attaining a level 23% of that occurring in embryos. The high levels of transcription seen in embryos may indicate an important role for the bbcl gene during embryogenesis and parallels results obtained by in situ hybridisation studies (chapter 5) and by Western blotting analysis of D. melanogaster proteins (chapter 7).

The bbcl mRNA seen during larval and pupal stages of development may either represent remnants of the embryonic expression or de novo expression of the gene. At present, no distinction can be made between these possibilities due to a lack of knowledge regarding the turn-over rate of the bbcl transcript during different stages of D. melanogaster development. However, the low level of bbcl transcript seen in pupal RNA preparations does indicate a much reduced need for the gene product during this stage of development (see chapter 7). The possibility also exists that high levels of expression (possibly comparable to the levels seen during embryogenesis) occur in a very restricted number of cells during pupal development. Resolution of this ambiguity will require a knowledge of spatial expression patterns of the bbcl gene during pupation. The moderate levels of bbcl gene expression seen in adult D. melanogaster indicates a role for the gene product in adult tissues, but at present, the tissues
c.

% Expression

Developmental Stage

- EMBRYONIC
- LARVAL
- PUPAL
- ADULT

Graph showing expression levels across different developmental stages.
Northern blotting analysis of *D. melanogaster* RNA samples from 0-24hr embryonic (E), third instar larval (L), pupal (P) and adult (A) stages of development. 10μg of total RNA from each stage was separated on a 1% agarose gel and blotted onto nylon membrane. The membrane was hybridised in 50% formamide with either a $^{32}$P-radiolabelled probe generated from the 620bp EcoRI-SphI fragment of *D. melanogaster* cDNA clone 6 (panel a) or a $^{32}$P-radiolabelled reverse transcribed total cDNA probe (panel b) generated as described in chapter 2. Washing stringency was 0.5x SSC at 65°C in both cases and autoradiography was for 14hr at -70°C with Fuji RX™ film (panel a) or 2 days at -70°C with Amersham Hyperfilm™-MP (panel b). Marker sizes are in kilobases. Panel c shows relative expression levels of the *bbc* gene during each developmental stage analysed (see table 3.3.1; embryonic expression taken as 100%). Adjustment for differences in loading between tracks was performed as described in table 3.3.1.
<table>
<thead>
<tr>
<th>Stage</th>
<th>Measured Intensities</th>
<th>Adjusted* bbcl</th>
<th>bbcl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>bbcl cDNA probe</td>
<td>Total cDNA probe</td>
<td>Intensities</td>
</tr>
<tr>
<td>Adult</td>
<td>1089.5</td>
<td>7111.1</td>
<td>1135</td>
</tr>
<tr>
<td>Pupal</td>
<td>130.5</td>
<td>5588.8</td>
<td>178</td>
</tr>
<tr>
<td>Larval</td>
<td>643.2</td>
<td>7548.1</td>
<td>643</td>
</tr>
<tr>
<td>Embryonic</td>
<td>3006.1</td>
<td>4649.3</td>
<td>4881</td>
</tr>
</tbody>
</table>

Table 3.3.1

Expression levels of the bbcl mRNA during four stages of D. melanogaster development. Intensity values obtained when the autoradiographs shown in figure 3.3.2 were subjected to scanning laser densitometry are shown in the first two columns of figures. *The values obtained using the "total cDNA" probe were used to control for differences in RNA loading between tracks according to the following equation:

Adjusted bbcl intensity = \( b_i \times \frac{T_L}{T_i} \)

Where:
- \( b_i \) is the intensity value with the bbcl probe,
- \( T_L \) is the intensity value for larval RNA with the "total cDNA" probe,
- \( T_i \) is the intensity value with the "total cDNA" probe,
- \( i \) is adult, pupal, larval or embryonic intensity values with the bbcl probe.

†The values for bbcl expression levels during different stages of D. melanogaster development are given as percentages of the highest (embryonic) value.
which express bbd mRNA are not known.

### 3.3.2 Identification of the Polyadenylation Site within the D. melanogaster bbd Transcript

The λgt11 library from which the *D. melanogaster* bbd cDNAs were isolated was generated by random-primed first strand cDNA synthesis. Hence, none of the clones contain the diagnostic poly(A) tract seen at the 3' end of most coding mRNAs. Since the human bbd mRNA is polyadenylated (see chapter 1), it is likely that the *D. melanogaster* bbd transcript will also be polyadenylated. On this premise, an attempt was made to clone the *D. melanogaster* bbd mRNA 3' end using a PCR-based strategy called RACE (rapid amplification of cDNA ends). The technique has been described (Frohman, 1990) and is summarised in figure 3.3.3a. The product of the RACE-PCR reaction was subcloned into the vector M13mp18 using restriction endonuclease sites present within the RACE2 primer and the bbdl sequence (see chapter 2, section 2.3.16) prior to sequencing.

Figure 3.3.3b shows part of the PCR product sequence compared with the sequence at the 3' end of clone 6. The presence of 18-32 adenylate residues at the 3' end of individual PCR products indicates that the 3' end of the *D. melanogaster* bbd mRNA has been cloned. Variation in the number of adenylate residues present at the 3' end of individual PCR products results from the random nature of priming by the RACE1 primer which can prime anywhere on the polyadenylate tract. The sequence immediately adjacent to the polyadenylate tract corresponds exactly with that of clone number 6 and demonstrates that clone 6 contains all of the 3' untranslated region (UTR) encoded by the *D. melanogaster* bbd gene.

3' UTRs in *D. melanogaster* and other eukaryotic mRNAs vary greatly in length, with examples as long as the 2029nt UTR found in the mRNA for the *D. melanogaster* type I regulatory subunit of cAMP-dependent protein kinase (Kalderon and Rubin, 1988) and as short as the 30nt UTR found in mouse α-amylase-1* mRNA (Hagenbüchle et al., 1981) and the 19nt UTR found in human bbd1 cDNA 10-9 (see figure 1.3.5). The 70nt of UTR found in the *D. melanogaster* bbdl mRNA (see
Figure 3.3.3

Use of RACE-PCR to clone the *D. melanogaster* bbcl mRNA 3' end. Panel a summarises the technique of RACE-PCR (see figure 2.1.1 for the position of primer 5003'). Panel b shows a comparison of sequences at the 3' ends of the PCR product and *D. melanogaster* cDNA clone 6. Sequence forming a consensus polyadenylation signal is indicated by a horizontal line.
The D. melanogaster bbcl Gene

Chapter 3

Figure 3.4.1) lies towards the lower end of these two extremes, but is somewhat longer than the UTRs found in the majority of human bbcl mRNAs (Figure 1.3.5). The difference in lengths between the human and D. melanogaster bbcl 3' UTRs is not particularly surprising, given that these two organisms are separated by over 700 million years of evolution (Dayhoff, 1978) and that UTRs accumulate mutations at much higher rates than coding regions (Nishioka and Leder, 1979).

The polyadenylation signal within the D. melanogaster bbcl mRNA is shown in Figure 3.3.3b and, unlike the signal in the human mRNA, is not contained within the translation termination codon, but resides 42nt downstream of this point. 22nt separate the polyadenylation signal from the poly(A) tract in the D. melanogaster mRNA. Many genes contain 10-20nt of 3' UTR between the polyadenylation signal and the poly(A) addition site (Baralle, 1983; Birnstiel et al., 1985) and thus the D. melanogaster bbcl mRNA is not unusual in the length of this sequence. Commonly mRNAs contain a cytidine residue immediately prior to the poly(A) tract (Baralle, 1983; Birnstiel et al., 1985) as is seen for two of the four poly(A) addition sites found in the human bbcl mRNAs (Figure 1.3.7). However, this is not always the case and other nucleotides, including guanine in the case of the D. melanogaster and two of the human bbcl polyadenylation sites (Figure 1.3.7), occur in a substantial number of cases.

3.3.3 Attempted cloning of the D. melanogaster bbcl mRNA 5' End

The experiments described above confirmed that the sequence at the 3' end of D. melanogaster bbcl cDNA clone 6 was derived from the bbcl gene. In order to resolve the ambiguity with respect to the sequence at the 5' end of the bbcl mRNA, an attempt was made to clone this sequence using the RACE PCR method shown in Figure 3.3.4. Despite several attempts, no products were observed after the PCR step. Possible reasons for this include poor priming by the RACE1 primer, lack of cDNA synthesis by the reverse transcriptase or inefficient addition of adenylate residues to the first strand cDNA product. It is known that the terminal transferase enzyme adds
Use of RACE-PCR to amplify mRNA 5' ends. Primers 1 and 2 are nested primers within the cDNA being amplified; RACE 1 primer is as in figure 3.3.3. Taken from Frohman (1990).
nucleotides to protruding 3' termini in preference to blunt or recessed 3' termini (Roychoudhury and Wu, 1980; Michelson and Orkin, 1982; Deng and Wu, 1983). Since the cDNA product acting as a substrate for terminal transferase would probably have contained either a blunt or recessed 3' terminus, this may account for the lack of products obtained on PCR amplification of the cDNA.

3.3.4 Primer-Extension Analysis of the D. melanogaster bbcl mRNA

Primer-extension analysis (transcript mapping) was performed in an attempt to determine the position of the transcription-initiation site relative to known cDNA sequences (Sambrook et al., 1989). A 58bp primer comprising sequence just downstream of the translation-initiation site (figure 3.3.5b) was radiolabelled using $\gamma^{32}$P]ATP and polynucleotide kinase (see chapter 2), then used to prime cDNA synthesis. The products of primer-extension were separated by denaturing polyacrylamide gel electrophoresis with M13 sequencing reactions as size markers (figure 3.3.5a). A single major product was observed of size 134nt. Some minor products of smaller size can be seen which probably result from premature termination of the reverse transcriptase. Also, a couple of slightly larger products are seen, which may be due to slight variation in the site at which transcription of the D. melanogaster bbcl gene initiates. As can be seen from figure 3.3.5b, this result places the major site of transcription initiation 9nt upstream from sequence which has been confirmed in the two independent cDNA clones 7 and 8 (see section 3.2). The 46nt of 5' UTR falls within the size range found for most eukaryotic mRNAs of between 20 and 60nt (Baralle, 1983; Kozak, 1984) and is of similar length to the 51nt of 5' UTR found in the human bbcl mRNA (result obtained by Dr. S. M. Adams using the technique of primer-extension).

The above result does not unambiguously identify the D. melanogaster bbcl mRNA 5' terminus since it is possible that secondary structures present in the mRNA could result in premature termination of the reverse transcriptase prior to the mRNA 5' end. However, premature termination of the reverse transcriptase rarely occurs when primers
Figure 3.3.5

a. Primer-extension analysis of *D. melanogaster* RNA. The products of primer-extension were separated on a denaturing 7.5% polyacrylamide gel using M13 sequencing reactions as size markers. After vacuum drying, the gel was autoradiographed with Kodak X-OMAT™ film overnight at room temperature. b. Partial sequence of *D. melanogaster* bbcl cDNA showing the position of the 58bp primer (---), the predicted 5' terminus of the mRNA (▲) and the initiation methionine codon of the open reading frame (■).
The above results indicate that the mature *D. melanogaster* *bbcl* mRNA contains 773nt of transcribed sequence (figure 3.3.6). According to Northern blotting experiments (section 3.3.1) the *D. melanogaster* *bbcl* mRNA is about 950nt in length. Hence it can be estimated that a poly(A) tail of approximately 180nt is present on the mRNA. Poly(A) tracts of this length are commonly found on mRNA molecules from many eukaryotic organisms (Baralle, 1983). A similar length of poly(A) tail is predicted for the human *bbcl* mRNAs (see chapter 1).

3.4 Analysis of the *D. melanogaster* *bbcl* cDNA

3.4.1 Comparison of the Human and *D. melanogaster* *bbcl* cDNAs

Elucidation of virtually all of the *D. melanogaster* *bbcl* cDNA sequence allows this to be compared with that of the human cDNAs. Figure 3.4.1 shows an alignment of the two sequences. There is a high degree of similarity between the two sequences which is limited to the region encompassed by the longest open reading frame (ORF) in both cDNAs. The level of nucleotide identity between the sequences is 62% within the open reading frame, but drops to 32%
Figure 3.3.6

Sequence of the *D. melanogaster* bbd cDNA showing the predicted translation product of the longest ORF. The 9 unidentified nucleotides at the 5' end (N) and the poly(A) tract (\((A)_n\)) are shown. The consensus polyadenylation signal sequence is indicated by underlining.
Figure 3.4.1

Comparison of the *D. melanogaster* (D) and human (H) *bbcl* cDNA sequences (note that the polyadenylation site present in human cDNA clone C328-10 is shown). The initiation methionine codon of the ORF is indicated (=>) as is the termination codon (*). Deletions in one sequence relative to the other are denoted by dots and sequence identity by vertical lines. The consensus polyadenylation signal sequence in both cDNAs is underlined. Numbering in both cDNAs is with respect to the adenosine residue of the initiation codon of the open reading frame shown.
The coding regions are quite GC-rich (62% in both cDNAs) and contain a large number of CpG dinucleotides (50 in the *D. melanogaster* coding region and 55 in the human coding region). CpG dinucleotides are generally rare in vertebrate DNA sequences due to N⁵-methylation of the cytidine residue. This results in mutation of the cytidine to a thymidine when spontaneous deamination occurs (Bird, 1987). The presence of so many CpG dinucleotides in the human *bbcl* gene therefore indicates that this gene is in a region of undermethylated DNA, which has allowed CpGs to persist as they have in the *D. melanogaster* sequence. Such undermethylated regions of DNA usually harbour genes which are expressed ubiquitously within vertebrate organisms; the so called housekeeping genes (Bird, 1987). An analysis of tissue-specific *bbcl* expression in the rat using filters obtained from Miss A. Simpson (Dept. of Anatomy, Leicester) shows that an mRNA capable of hybridising with the human *bbcl* cDNA is expressed in all tissues examined (figure 3.4.2). Although equal amounts (30µg) of each RNA were loaded onto the gel, a control hybridisation with reverse-transcribed rat total RNA was not performed. Hence, a comparison of expression levels has not been carried out. If the human *bbcl* gene also shows ubiquitous expression, this would strongly indicate a general, "housekeeping", role for the gene in all cells. Due to the conservation of sequence between the human and *D. melanogaster* genes, it is possible that the *D. melanogaster* gene may also function as a housekeeping gene.

The presence of large numbers of CpG dinucleotides within the coding region of the human *bbcl* gene may be due to the lethality of mutations within this gene. The human and *Drosophila* coding sequences are highly conserved (figure 3.4.1), indicating that mutations have been selected against. However, there is a bias towards CpG dinucleotides not being conserved, since there are 27 and 32 unconserved CpG dinucleotides within the *Drosophila* and human *bbcl* coding regions respectively, whereas only 23 CpG dinucleotides are conserved between the two sequences. These conserved CpG dinucleotides tend to reside in highly conserved stretches of nucleic acid and encoded protein sequence. Hence, if the
Figure 3.4.2

Northern blotting analysis of total RNA prepared from adult female Wistar rat tissues. Filters (obtained from Miss A. Simpson, Dept. of Anatomy, Leicester) containing 30μg of total RNA from the tissues stated were hybridised overnight at 45°C in 50% formamide with a 32P-radiolabelled probe generated from the 480bp Styl fragment of human cDNA clone C328-10 (see figure 2.1.1). After washing in 2x SSC at 65°C, the filters were autoradiographed at -70°C for 9 days with Amersham Hyperfilm™-MP. Key: Lg, lung; H, heart; A, aorta; Li, large intestine; Si, small intestine; M, skeletal muscle; O, ovary. Marker sizes are in kilobases (Kb).
unconserved CpG dinucleotides are not required for maintenance of the protein sequence, their continued presence in the human coding sequence indicates that their rate of mutation has been lower than that of CpGs in many vertebrate genes; possibly as a result of undermethylation.

The predicted initiation methionine codons of both proteins occur in sequences which conform well to consensus translation initiation sites (Cavener, 1987) and are preceded by in-frame termination codons. In addition, both sequences contain termination codons in the predicted non-coding frames downstream of the initiation codon. Thus both cDNAs contain a single long ORF which is highly conserved at the nucleotide level and probably encodes the protein product of the bbcl gene.

3.4.2 Comparison of the Human and D. melanogaster bbcl cDNAs with Other Nucleic Acid Sequences

To identify nucleic acid sequences similar to the bbcl cDNA the EMBL nucleic acid database (release 26) was searched with both the D. melanogaster and human cDNA sequences. Sequences demonstrating similarity to both bbcl cDNAs may represent conserved regions of common function. Whilst the searches identified many sequences with varying similarities to the individual cDNAs, none of these were common to both cDNAs. This analysis, therefore, failed to identify any genes likely to encode proteins with similar functions to the bbcl protein.

3.5 The Predicted D. melanogaster bbcl Protein

3.5.1 Comparison with the Predicted Human bbcl Protein

The longest ORF in the D. melanogaster bbcl cDNA encodes a protein of 218 amino acids in length. This predicted protein is closely similar in length and sequence to the predicted product of the human bbcl cDNA (figure 3.5.1). There are 136 identical residues, resulting in 62% sequence identity. If 25 conservative substitutions are included, the level of sequence conservation rises to 74%. Proteins
Figure 3.5.1

Comparison of the predicted *D. melanogaster* (D) and human (H) *bbcl* proteins. Identical amino acids are indicated by vertical lines and conservative substitutions by colons. Spaces, denoted by dots, have been introduced into the sequences where necessary to maintain sequence alignment. Underlined sequences are consensus SV40-like nuclear localisation signals and sequences marked by an asterisk are consensus N-linked glycosylation sites.
which are highly conserved invariably perform important cellular roles, often interacting with a large number of different proteins or other macromolecules. Examples of such proteins include the core histones, actin and nm23/Awd (Korn, 1982; Wells, 1986; Rosengard et al., 1989).

Another feature of the predicted *D. melanogaster* protein is the preponderance of basic amino acids within the protein. 28% of the residues are lysine, arginine or histidine, whereas only 9% are acidic (aspartate or glutamate). These figures agree closely with those for the predicted human protein (table 3.5.1) and further demonstrate the evolutionary conservation of this protein. The basic nature of the predicted bbcl protein and the existence of two conserved SV40-like nuclear localisation signals (Kalderon et al., 1984; Goldfarb et al., 1986) within both proteins are indicative of a nucleic acid-binding protein. The predicted bbcl protein does not, however, contain either of the two known RNA-binding motifs (Dreyfuss et al., 1988; Calnan et al., 1991) nor any of the currently known DNA-binding motifs (Churchill and Travers, 1991; Harrison, 1991). It remains possible that the protein binds nucleic acid by a novel mechanism, but it is also possible that the protein performs a function within the nucleus or cytoplasm not involving nucleic acid binding.

One potential N-linked glycosylation site (Lehle and Bause, 1984) is found in the predicted *D. melanogaster* bbcl protein. The position of this motif is identical in the human protein and overlaps one of the putative nuclear localisation signals. The *in vivo* relevance of this motif remains to be established.

To analyse further the predicted protein sequence, the GCG computer programme PEPLOT was used to display various characteristics of the *D. melanogaster* and human proteins in graphic form (figure 3.5.2). These plots show acidity and basicity of the sequences (top panel), secondary structure prediction (Chou and Fasman, 1978, middle panel) and hydropathy (Kyte and Doolittle, 1982, bottom panel). Simple comparison of the two plots shows that they are very similar, as would be expected from the information given above. Closer inspection reveals some interesting features of the protein...
Table 3.5.1

Properties of the predicted *D. melanogaster* and human bbcl proteins. Molecular weight and pl values were calculated using the IBI Macvector® programme running on an Apple Macintosh computer.
Figure 3.5.2

Pepplot analyses of the predicted human (A) and D. melanogaster (B) bbcl proteins. The top panel in both sections shows a distribution of amino acid types within the protein. Red lines show basic and acidic residues (lines pointing up and down respectively), blue lines show hydrophobic residues, green lines show hydrophilic residues and black lines show proline residues. Line lengths denote relative degrees of each parameter (i.e. acidity, basicity etc.). The middle panel shows secondary structure prediction according to the method of Chou and Fasman (1978). The green line indicates α-helix forming potential and the blue line indicates β-sheet forming potential. Values greater and less than 1 indicate tendencies to form and break respectively helices or sheets. The bottom panel shows hydrophobicity according to the method of Kyte and Doolittle (1982). Positive values show hydrophobic residues and negative values show hydrophilic residues. Scales along the top and bottom of both sections show amino acid positions from the amino-terminus (0).
The 0. melanogaster bbcl Gene  Chapter 3

sequences. The top panel clearly shows that the 80 or so amino acids at the N-terminus of the proteins are either totally devoid (*D. melanogaster*) or almost devoid (human) of acidic residues. The significance of this observation is not known at present, but could indicate that this region forms a highly basic domain within the protein. The bottom panel shows that both proteins are essentially hydrophilic in nature and do not contain a hydrophobic leader sequence characteristic of secreted proteins (Devillers-Thiery *et al.*, 1975), nor a stretch of hydrophobic amino acids which could form an α-helical transmembrane domain (Michel *et al.*, 1986).

3.5.2 Comparison with Other Proteins

A comparison of the *bbcl* cDNA sequences with entries in the EMBL nucleic acid database failed to identify any sequences displaying similarity to both cDNA sequences. However, due to codon usage differences and third-base wobble, comparisons of nucleic acid sequences are not as sensitive as amino acid comparisons. Hence a search of protein sequences within the NBRF (release 36) and SWISSPROT (version 15) protein databases was performed. A large number of proteins with varying similarities to the *bbcl* proteins were found. It is of interest that both of the sequences most similar to the *D. melanogaster* and human *bbcl* proteins are from DNA-binding proteins (figure 3.5.3). The regions of similarity between the predicted *bbcl* proteins and the *D. melanogaster* P-element transposase lie 54 amino acids C-terminally to the helix-turn-helix DNA-binding motif of the transposase in a region of unknown function (Rio *et al.*, 1986). The regions of similarity between the predicted *bbcl* proteins and the Fra-2 transcription factor are 11 amino acids N-terminal to the leucine zipper motif of the Fra-2 protein (Nishina *et al.*, 1990). However, no leucine zipper is present in the predicted *bbcl* proteins and the significance of the similarity is not known.

3.6 Conclusions

Use of the human *bbcl* cDNA to screen a *D. melanogaster* cDNA library at low stringency resulted in isolation of cDNAs encoding *D.
a.

<table>
<thead>
<tr>
<th></th>
<th>Sequence</th>
<th>Length</th>
<th>Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>D. mel bbl I</td>
<td>ILFPNEKKIRAGESVXEECLATQL</td>
<td>149</td>
<td>50% over 26α</td>
</tr>
<tr>
<td>Transposase</td>
<td>ILPKINENHYVQSLKHVKLTQL</td>
<td>407</td>
<td></td>
</tr>
<tr>
<td>Human bbl I</td>
<td>ILFPRKPSAPKRGSSAEVLATQL</td>
<td>150</td>
<td>34.5% over 26α</td>
</tr>
</tbody>
</table>

b.

<table>
<thead>
<tr>
<th></th>
<th>Sequence</th>
<th>Length</th>
<th>Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>D. mel bbl I</td>
<td>KTIGAVDRRKNKSLESQRNQRLKNTKL</td>
<td>123</td>
<td>39.4% over 33α</td>
</tr>
<tr>
<td>Human fra-2</td>
<td>KTIGTVGRRRDEQLPQEQKRRKNNKL</td>
<td>136</td>
<td></td>
</tr>
<tr>
<td>Human bbl I</td>
<td>KTIGSVDPARKINTLSTEQKLVRLKNTKL</td>
<td>124</td>
<td>33.3% over 33α</td>
</tr>
</tbody>
</table>

Figure 3.5.3

Comparison of a D. melanogaster P-element transposase (a) and the human fra-2 protein (b) with the predicted human and D. melanogaster (D. mel) bbl proteins. Only the regions of highest identity between the proteins are shown. Identical amino acids are indicated by vertical lines and conservative substitutions by colons.
The *D. melanogaster bbcl* Gene Chapter 3

*melanogaster bbcl* sequences (Dr. J. M. Varley). Northern blotting analysis revealed significant differences in the level of *bbcl* mRNA present at various stages of *D. melanogaster* development, with the highest levels in embryonic RNA extracts, possibly indicating an increased requirement for the gene product during this stage of the *D. melanogaster* life cycle in comparison to other stages. Whilst the Northern blot of rat tissues indicated that the gene product may perform a "housekeeping" role, it should be borne in mind that all these tissues were from an adult rat and that the results obtained from Northern blotting analysis of *D. melanogaster* RNA do not preclude the gene being expressed in all tissues of this organism, albeit at distinctly different levels throughout development.

Various techniques were used to confirm all but 9nt of the *D. melanogaster* cDNA sequence which was shown to contain a single long ORF. The extensive similarity of the *D. melanogaster* and human ORFs suggests that this gene has been highly conserved throughout evolution.

The putative proteins encoded by the *D. melanogaster* and human cDNAs were also found to be highly conserved and to contain a large number of basic amino acids. Consensus SV40-like nuclear localisation sites were found to be present within the sequence, as was a single potential N-linked glycosylation site. However, database searches failed to identify extensive similarities to previously isolated proteins and no other protein motifs were found. The *bbcl* gene appears to encode a novel protein with unknown function(s).
CHAPTER 4

Analysis of D. melanogaster and Human bc1 Genomic DNA Sequences

4.1 Introduction

Characterisation of the D. melanogaster and human bc1 cDNA sequences is described in preceding chapters. In an attempt to characterise D. melanogaster bc1 genomic sequences, an analysis of the bc1 gene within this organism was undertaken. The results of this analysis led to a study of the human bc1 gene(s) and provided some interesting information regarding the genomic organisation of the D. melanogaster and human bc1 genes.

4.2 PCR Amplification of the D. melanogaster bc1 Genomic ORF Region

Oligonucleotide primers designed for amplification of almost the entire coding region within the D. melanogaster bc1 cDNA (see chapter 6) were used to amplify genomic DNA from this organism. The product of PCR amplification was found to be approximately 130bp larger than the product of an identical reaction performed on a control cDNA sequence. This indicated the presence of an intron(s) within the ORF-region of the gene and so the PCR product generated from a genomic DNA template was sub-cloned and sequenced.

Two introns were seen to be present within the coding region of the gene (figure 4.2.1a), one 61bp and the other 69bp in length. Whilst introns of this length are short by vertebrate standards, they fall well within the most common size range of 50 to 75bp for 80% of invertebrate introns (Hawkins, 1988). Both introns also contain the conserved GT and AG di-nucleotides found at splice donor and acceptor sites respectively (Shapiro and Senapathy, 1987). In addition, the sequences around the splice donor sites of both introns conform precisely to the consensus invertebrate motif (figure
a.

ATGGTAAGGGTAGCAGATTCACAGATAGCTAGGGCAGCGTATGAGGAAGAGCAGGCTGCAAGCTGATGACAT

MGKGNMNNMIIPNQHYNHKKNWQRH

GTGAAGATTGGTCACACGGCCCAAGGCTGCAAGCAGACGCGTCAAG

VKTWFQNPARKVRRHANKVRK

AAGGCTAAAGGCAGCTCTTCACCAACAGGAGGCTCCTCCGCTCTTGTCCTGAC

KAKAVFPRPASGPLRPVRC

CCACACATTCCCTTACACCAAGGCTGTCACGGCCCGTGGTTCCCTACCTTTGAGGAGCGTG

PTIRYHTKLRAGRGTTEEL

AGGtastgtcagcggaaasattacctcgatcctcgtgccacctaaacaaasattctctgttt

300

b.

Intron 1

AGGTAAGT..............ATCTCTGTTTACAGG

Consensus

AGGTAAGT..............TTTTCTTTTTACAGG

+- Pyrimidine

+++

Intron 2

AGGTAAGT..............AAGTAGTTTACAGG

Consensus

AGGTAAGT..............TTTTCTTTTTACAGG

+- Pyrimidine

++++
Figure 4.2.1

a. Sequence of the open reading frame region of the *D. melanogaster bbcl* gene amplified by PCR (see figure 2.1.1 for the positions of the PCR primers ORFPCR5' and ORFPCR3'). Exons are shown in upper case and introns in lower case. The predicted translation product is also shown.

b. Comparison of the intron sequences within the *D. melanogaster bbcl* gene with the consensus invertebrate motifs. Identity between the sequences is denoted by a vertical bar. The presence or absence of pyrimidine residues in the 10nt prior to the splice acceptor site is also shown.
4.2.1b). Sequences around the splice acceptor sites differ somewhat from the consensus invertebrate sequence, but retain many of the conserved residues including the CAGG tetra-nucleotide and a reasonable poly-pyrimidine tract (figure 4.2.1b).

4.3 PCR Amplification of Sequences within the Human bbcl Gene

The identification of introns within the coding region of the D. melanogaster bbcl gene was not surprising since most genes of higher eukaryotes contain introns (Shapiro and Senapathy, 1987). Of particular interest was the identification of an intron within the human bbcl gene. Primers flanking a region of the human coding sequence equivalent to the position of the first intron in the D. melanogaster ORF were used in a PCR experiment to amplify human genomic DNA. Two major products were seen, one corresponding in size to the product of PCR on a control cDNA sequence and one of a larger size (figure 4.3.1a). A faint band can be seen corresponding to a PCR product of size 570bp. The 133bp product may result from amplification of pseudogene sequences (see below). Neither the 133bp nor the 570bp products were analysed further due to time constraints.

The 341bp product was subcloned into the vector pUC18 and two subclones sequenced on both strands. In addition, direct sequencing of the PCR product was performed to resolve a base pair mis-match between the two subclones. The sequence was shown to contain an intron of 208bp in length (figure 4.3.1b). Both the splice donor and acceptor sites conform to consensus vertebrate sequences (Shapiro and Senapathy, 1987) and a poly-pyrimidine tract can be seen near the splice acceptor site. The intron within the human bbcl gene is substantially longer than the equivalent intron within the D. melanogaster gene, but is relatively short in comparison with many vertebrate introns (Hawkins, 1988).

The most interesting attribute of the human intron is the position it occupies within the gene. Figure 4.3.1c shows a comparison of the D.
a.

```
  1 2 3 4
  ←341bp
  ←133bp
```

b.

**SstI**
..CTGGAGGAGCTCAGGgtgagtactggcaqcgctggtcaggaagggcccgaagtccc
..L E E L R
ccttggttgccctcagtcgctgtgatgacattctccggaatcgcgtagctacggccttgatgaaa
goacatttgaacctttttcatctgattgctgaggcttttcatccaggcctcggggtggtgg

**SmaI**
agaaagcccgggccgtctccatctctctctctgttgtgtttgggcagGTGGCCGCGCATTCAC...
    V A G I H..

c.

**Human**
GAGGAGCTCAGGgtgagt...............cagGTGGCCGCGCAT
E E L R     V A G I

**Drosophila**
GAGGAGCTGAAGgtagttggtgagt............cagGGTGCCGCGCAT
E E L K     G A G I
Figure 4.3.1

a. Products of touch-down PCR on human cDNA clone 10-12 and human genomic DNA (from placenta) using oligonucleotide primers 505 and 620. The samples were separated on a 2.5% agarose gel. Tracks are: 1. size marker (plasmid pAT153 digested with HaellIII endonuclease); 2. no DNA control; 3. cDNA control and 4. human genomic DNA. Amplification product sizes are indicated (arrows). b. Sequence of the intron identified using PCR. Intron sequence is shown in lower case and flanking exon sequences, with predicted translation product, in upper case. The restriction endonuclease sites used for cloning the intron fragment are shown. c. Comparison of the D. melanogaster and human bbcl exons around the site of the first intron within the ORF.
melanogaster and human genomic sequences around the site of the human intron identified above. Not only do both introns reside between codons, but they occur at identical positions in both genes. This is not novel; for example, intron positions within the α- and β-globin genes of mice, rabbits and humans are precisely conserved (Breathnach and Chambon, 1980). However, it does support the theory that the two genes evolved from a common ancestral gene present before divergence of the arthropod and chordate lineages; an event that occurred about 700 million years ago (Dayhoff, 1978).

The intron position within both the human and D. melanogaster genes with respect to the ORF is also of interest because it lies very near to the boundary between the N-terminal highly basic domain and the rest of the protein (figure 4.3.2. See also chapter 3, figure 3.5.2). It has been suggested that introns mark the boundaries between protein domains (Go, 1981; Craik et al., 1983; Brändén et al., 1984) and, further, that division of genes into exons allows molecular evolution of proteins by addition and removal of entire exons from the coding regions of genes (Südhof et al., 1985a; Südhof et al., 1985b; Maeda and Smithies, 1986; Baron et al., 1991) a process known as "exon shuffling" (Gilbert, 1985). Studies on molecular evolution of the bbcl gene will require analysis of sequences from more organisms, including lower eukaryotes, which may have different gene structures.

Given the presence of one intron at an identical site within both the human and D. melanogaster bbcl genes, it was of interest to determine whether a second intron existed in the human gene and whether this was at an identical position with respect to the second intron identified in the D. melanogaster bbcl gene. Several attempts were made to amplify this region of the human bbcl gene using different primer pairs and PCR conditions. However, either no PCR products were seen, or products were obtained that corresponded in size to control amplifications of cDNAs. Two possible explanations could account for these results: either an intron exists but could not be amplified (the PCR products obtained could then represent amplification of pseudogene sequences) or no intron is present. Since an intron within the human bbcl gene could be too long to be
Figure 4.3.2

Position of introns within the human (top) and D. melanogaster (bottom) bbcl genes with respect to the highly basic N-terminal domains of the bbcl proteins. The top part of both panels represents the first 100 amino acids of the bbcl proteins (see figure 3.5.2 for the key). An expansion of the intron-containing region of the cDNA sequence is shown below this with a translation of the sequence. The position of the intron is shown. Note the two glutamate residues (E) which form the boundary between the N-terminal basic domain and the rest of the protein.
efficiently amplified by the method of PCR, no further attempts have been made to study the human \textit{bbcl} gene by this means.

4.4 Southern Blotting Analysis of Human Genomic DNA using an Intron-specific Probe.

A \textit{SstI-Smal} restriction fragment from the human \textit{bbcl} intron identified above (figure 4.3.1b) was used to probe a Southern blot of human genomic DNA cleaved with a number of restriction endonucleases. A single hybridising band was identified in all cases, indicating that a single hybridising DNA sequence was present (figure 4.4.1). This result suggests that a single \textit{bbcl} gene is present within the human genome and that the extra bands present on Southern blots probed with the \textit{bbcl} cDNA sequence represent either intron-less pseudogenes or closely related genes with cross-hybridising exonic sequences.

4.5 Localisation of the Human \textit{bbcl} gene to Chromosome 16

The use of \textit{bbcl} cDNA sequences to map the human gene to a specific chromosome could have led to artefactual results due to the large number of cross-hybridising sequences present within the human genome. Isolation of an intron and demonstration that it probably represents a unique sequence within the human genome allowed mapping of the intron to a specific chromosome using a panel of human-hamster somatic cell hybrid DNAs (Bios Corporation).

Two primers specific to the human \textit{bbcl} gene, one (10-5) intron-specific and one (10-9) exon-specific, were used to amplify human sequences within the hamster background. Two of the hybrid DNA samples contained a human sequence that was amplified to generate a product of the same size (197bp) as a control amplification performed on human genomic DNA (figure 4.5.1). Concordancy analysis was performed on these results as described in chapter 2 (section 2.3.16). Concordant results, expressed as a percentage of
Southern blotting analysis of human genomic DNA (from placenta). 4μg of DNA was cleaved with EcoRI (lane 1), HindIII (lane 2), BamHI (lane 3) or BglII (lane 4) restriction endonucleases. After separation on a 0.8% agarose gel and transfer to a nylon membrane the DNA was hybridised with a $^{32}$P-radiolabelled SstI-Smal fragment from the first intron identified in the human bbcl gene (see figure 4.3.1b). Final washing stringency was 0.5x SSC at 65°C and autoradiography was for 4 weeks at -70°C with Amersham Hyperfilm™-MP.
Figure 4.5.1

PCR amplification of human *bbcl* gene sequences from a panel of human-hamster somatic cell hybrid DNAs. After amplification with primers 10-5 and 10-9 the products were separated by electrophoresis through a 2% agarose gel and visualised by staining with ethidium bromide. Lanes are labelled as follows: M, DNA size marker (pAT153 cleaved with *HaeIII*); N, no DNA control; C, cloned human *bbcl* DNA control; Hu, human genomic DNA control; Ha, hamster genomic DNA control; numbered lanes refer to somatic cell hybrid reference numbers (see table 4.5.1).
the number of somatic cell hybrids tested (25), demonstrated that the human \textit{bbcl} gene has a 96\% probability of being on chromosome 16 (table 4.5.1).

It will be noted from table 4.5.1 that DNA from somatic cell hybrid 904 did not generate a PCR product with this pair of primers but does contain human chromosome 16. The original PCR experiment using primers 10-5 & PCR 505 generated a PCR product with DNA from this somatic cell hybrid (see Adams et al., 1992). However, this reaction has subsequently been repeated twice and no PCR product of the expected size is produced with this hybrid DNA. Identical reactions using primer-pairs 10-5 & 10-10 and 10-6 & 10-7 (see below) also fail to amplify \textit{bbcl} sequences from this hybrid DNA. Hence, it is possible that the product seen in the original experiment resulted from amplification of contaminating human DNA sequences within the reaction utilising hybrid 904 DNA. A note to this effect will be published. Failure to amplify \textit{bbcl} sequences within DNA from hybrid 904 is interesting because the results given above indicate that the human \textit{bbcl} gene is on human chromosome 16 (assignment of the human \textit{bbcl} gene to chromosome 16 has also been independently confirmed, see chapter 8). It will be noted from table 4.5.1 that only 5\% of the hybrid 904 cells contain human chromosome 16. It is conceivable that this may prevent sufficient amplification to enable the PCR product to be visualised on an ethidium bromide-stained gel. However, the presence of human chromosome 16 in only 5\% of the cells indicates that this chromosome is unstable in this cell line. It is possible, therefore, that after the cell line was karyotyped complete loss of chromosome 16 occurred. This could be verified either by rekaryotyping the cell line, or by performing a PCR reaction using primers specific for a gene known to be present on human chromosome 16 (for example, APRT).

The positive control shown in figure 4.5.1 utilised a cloned PCR product generated using primers 10-9 and PCR 620 on human genomic DNA. The 372bp PCR product (containing the intron sequence) was cleaved with \textit{Bam}HI restriction endonuclease and the 310bp fragment cloned into the plasmid pUC18. This construct then served as a substrate for PCR using primers 10-9 and 10-5. The additional
Concordancy analysis of the results of PCR of human *bcl* intron sequences from a panel of human-hamster somatic cell hybrid DNA samples. The presence or absence of a PCR product is shown in the right-most column and % concordancy in the bottom row. The presence of a specific human chromosome within a hybrid is denoted by a plus. Numbers refer to the percentage of cells containing a chromosome. D shows that a deletion within chromosome 5 exists in the human chromosome present in the hybrid.
larger PCR product seen in the control track in figure 4.5.1 possibly results from a mis-prime on plasmid DNA to generate a species which co-amplified with the desired product.

Assignment of the human \(bbcl\) gene to chromosome 16 is interesting since a number of human malignancies, including breast carcinoma, have been associated with loss and rearrangement of this chromosome (Carter et al., 1990; Sato et al., 1990; Tsuda et al., 1990; Devilee et al., 1991). More precise mapping of the human gene will be required to determine whether the \(bbcl\) gene lies near to areas of chromosome 16 which undergo rearrangement or loss in breast cancer (see chapter 8).

4.6 Analysis of the 3' Extension in Human cDNA Clone 10-12

Two bands were observed when Dr. S. M. Adams probed a Northern blot of poly(A)+ RNA from human breast tumour cell lines with a 307bp \(KpnI\)-XhoI restriction fragment from clone C328-10 (figure 4.6.1a). Since the 250nt by which these RNA species differ is approximately the same size as the 3' extension in cDNA clone 10-12, it was of interest to investigate the origin of this sequence, thereby demonstrating whether or not the 3' extension present in cDNA clone 10-12 was an artefact.

A 219bp \(AlwNI\)-\(HinfI\) restriction fragment encompassing the entire 3' extension (and nothing else) within clone 10-12 (see figure 2.1.1) was used to probe the same Northern blot of poly(A)+ cell line RNA originally probed with the \(KpnI\)-\(XhoI\) fragment from clone C328-10. The \(AlwNI\)-\(HinfI\) fragment identified two faint bands with sizes of 1200nt and 1100nt (figure 4.6.1b); however, the 950nt band identified by the \(KpnI\)-\(XhoI\) fragment from clone C328-10 was not seen. The identification of two bands by the \(AlwNI\)-\(HinfI\) fragment could possibly be explained by the presence of a second infrequently used polyadenylation site approximately 100nt 5' of that found in cDNA clone 10-12. This hypothesis could be tested by the method of RACE PCR described in the preceding chapter.
Figure 4.6.1

Northern blotting analysis of human breast tumour cell lines MDA-MB-231 (lane 1), MCF-7 (lane 2) and Hs 578 T (lane 3). 2µg of poly(A)+ RNA from each cell line was separated by gel electrophoresis and transferred to a nylon membrane. Hybridisation in 50% formamide at 45°C was firstly with a $^{32}$P-radiolabelled 307bp KpnI-XhoI fragment from clone C328-10 (panel a), then, after stripping, with a $^{32}$P-radiolabelled 219bp AlwNI-Hinfl fragment from the 3' extension in clone 10-12 (panel b). Final washing stringencies were 0.5x SSC at 65°C for the KpnI-XhoI fragment and 2x SSC at 65°C for the AlwNI-Hinfl fragment. Autoradiography with Amersham Hyperfilm™-MP was either for 4hr at room temperature (panel a) or 3 weeks at -70°C (panel b). Note that the result shown in panel (a) was obtained by Dr. S. M. Adams (see figure 1.3.4).
The result indicated that the 3' extension in clone 10-12 may represent *bona fide* *bbcl* sequence since hybridisation was only seen with transcripts of sizes larger than the major 950nt mRNA. However, the possibility remained that the hybridising sequences represented different mRNAs of similar sizes to those identified by the C328-10 probe.

To determine whether the extension is contiguous with *bbcl* genomic sequences, PCR reactions were performed on human genomic DNA using primer pairs flanking the junction between the extension and known *bbcl* sequence (figure 4.6.2). The products obtained with all possible primer pairs were of sizes predicted from the cDNA sequence of clone 10-12 (table 4.6.1). To analyse further the authenticity of the extension, primers 10-6 & 10-7 were used to amplify human sequences from the somatic cell hybrid DNA panel. Figure 4.6.3 shows the products of amplification. It can be seen that the somatic cell hybrid DNAs which produced PCR products with primer pair 10-6 and 10-7 also produced products with primer pair 10-5 and 10-9 (see figure 4.5.1). This result shows that the 3' extension sequence in cDNA clone 10-12 has a 96% probability of being present on human chromosome 16 (by concordancy analysis).

The above results indicate that the 3' extension seen in clone 10-12 represents sequence which is likely to be contiguous with *bbcl* genomic DNA and present *in vivo* within the 1200nt transcript. Both the 1200nt and 1100nt transcripts appear to be much less abundant than the smaller, 950nt, transcript; a possible result of the presence of only one consensus polyadenylation signal which resides 18-34nt 5' of the most commonly used polyadenylation sites.

Polyadenylation sites are typically about 20nt downstream of the polyadenylation signal (Baralle, 1983), and it is possible that the long 3' UTR seen in clone 10-12 is merely the result of an infrequent polyadenylation event. Other genes have been shown to produce a range of transcripts containing 3' UTRs of varying length, for example the mouse dihydrofolate reductase (DHFR) gene (Setzer *et al.*, 1980) and the *D. melanogaster* neurogenic gene Delta (Haenlin *et al.*...
PCR analysis of the 3' extension present in human clone 10-12. PCR was performed on either 300ng of human genomic DNA (placental) or 10ng of 10-12 cDNA (in the vector M13mp18) using the primer pairs stated (see figure 2.1.1 for positions within the cDNA sequence). PCR products were separated on a 5% agarose gel and visualised by staining with ethidium bromide. Lanes 1 & 2, no DNA and positive cDNA (10-12 cDNA) controls respectively plus 10-6 & 10-7 primers; lane 3, genomic DNA plus 10-3 & 10-7 primers; lane 4, genomic DNA plus 10-3 & 10-8 primers; lane 5, genomic DNA plus 10-6 & 10-7 primers; lane 6, genomic DNA plus 10-6 & 10-8 primers. M, marker (pUC19 digested with Hpal restriction endonuclease).

<table>
<thead>
<tr>
<th>Primer-pair</th>
<th>Predicted product size (bp)</th>
<th>Actual product size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10-3 + 10-7</td>
<td>246</td>
<td>240</td>
</tr>
<tr>
<td>10-3 + 10-8</td>
<td>163</td>
<td>162</td>
</tr>
<tr>
<td>10-6 + 10-7</td>
<td>319</td>
<td>307</td>
</tr>
<tr>
<td>10-6 + 10-8</td>
<td>236</td>
<td>228</td>
</tr>
</tbody>
</table>

Table 4.6.1

Predicted and actual sizes of PCR products obtained using the primer-pairs stated to amplify sequences from human genomic DNA. Predicted sizes refer to sizes of products expected if the 3' extension within clone 10-12 is contiguous with bbcl sequences (see text for details). Actual product sizes were determined by comparison with the size marker shown in figure 4.6.2, above.
Figure 4.6.3

PCR amplification of human \textit{bbc1} gene sequences from a panel of human-hamster somatic cell hybrid DNAs. After amplification with primers 10-6 & 10-7 the products were separated by electrophoresis through a 3% agarose gel and visualised by staining with ethidium bromide. Lanes are as follows: M, DNA size marker of pUC19 digested with \textit{Hpa}II (sizes in bp); N, no DNA control; C, 10-12 cDNA control; Hu & Ha, human and hamster genomic DNAs respectively; numbered lanes refer to somatic cell hybrid reference numbers (see table 4.5.1).
The D. melanogaster bbcl Gene

* et al., 1990) produce mRNAs with 3' UTRs varying in length by up to a kilobase. However, the reason why transcripts with 3' UTRs of varying length are produced from these genes is not known and may just represent natural variation in the site of polyadenylation within these transcripts.

Functional differences in polyadenylation site usage are known. For example, use of alternative polyadenylation sites within the mouse immunoglobulin μ heavy chain gene (Early *et al.*, 1980) and the human and rat calcitonin genes (Rosenfeld *et al.*, 1984) is thought to result in alternative RNA splicing events that generate different protein products. Since alternative splicing of the human bbcl gene has not been ruled out, it remains possible that alternative polyadenylation site usage affects splicing of this gene.

4.7 Conclusions

Amplification of the ORF region of the D. melanogaster bbcl gene resulted in identification of two introns within the gene. These introns were shown to contain splice sites which conform to consensus invertebrate sites and to lie between codons within the ORF. A similar analysis of human bbcl genomic sequences identified an intron corresponding to the first intron within the D. melanogaster ORF. Further analyses have failed to identify a human equivalent of the second D. melanogaster intron.

Use of the human intron sequence to probe a Southern blot of human genomic DNA identified a single hybridising band indicating that a single functional human bbcl gene may exist. PCR analysis of somatic cell hybrid DNAs using primers that specifically amplify human bbcl intron sequence mapped this sequence to human chromosome 16.

Northern blotting analysis of human breast tumour cell line poly(A)+ RNA and PCR amplification of sequences from human genomic DNA and somatic cell hybrid DNAs suggest that the 3' extension present in cDNA clone 10-12 is authentic and that substantial variation in the
site of polyadenylation occurs within transcripts generated from the human \textit{bbcl} gene. The factors governing this variation and its significance are not known.
CHAPTER 5

In Situ Hybridisation Studies Performed in D. melanogaster

5.1 Introduction

Genetic studies in D. melanogaster have resulted in identification of a large number of mutant phenotypes, many of which have been mapped to specific regions of the genome by genetic means. In addition, the polytene chromosomes of D. melanogaster allow precise cytogenetic location of sequences on a physical map of the genome. This location can then be correlated with results of genetic analyses to identify candidate mutations affecting the sequence being studied. To this end, polytene chromosomes from D. melanogaster were hybridised in situ with a radiolabelled sequence from the D. melanogaster bbc1 cDNA.

The bbc1 gene is known to be expressed in D. melanogaster embryos (see chapter 3) and determination of both temporal and spatial expression patterns within developing embryos may indicate possible functions for the gene product. A radiolabelled sequence from the D. melanogaster bbc1 cDNA was, therefore, hybridised in situ with sections from developing D. melanogaster embryos in an attempt to identify sites of expression during development. In situ hybridisation to whole mount D. melanogaster embryos was also performed to confirm the results obtained with sectioned material.

5.2 In Situ Hybridisation to D. melanogaster Polytenes

In situ hybridisation was performed as described in chapter 2 using polytene chromosomes prepared from the salivary glands of third instar larvae (larvae kindly provided by Dr. C. P. Kyriacou, Dept. of Genetics, Leicester University). After autoradiography, the chromosomes were examined and the site of hybridisation assigned as 29E-F; a site on the left arm of chromosome 2 (Figure 5.2.1). This
Figure 5.2.1

In situ hybridisation to salivary gland polytene chromosomes from third instar larvae of *D. melanogaster*. A tritium-radiolabelled 620bp *EcoRI*-*SphI* fragment from cDNA clone 6 (see figure 2.3.1) was used as a probe. Autoradiography was for 12 days at 4°C. The site of hybridisation is indicated (arrow). Magnification, 600X.
The presence of interesting markers at the same locus as the \textit{bbcl} gene might have indicated a role for the gene and allowed further investigation using molecular and genetic approaches. The chromosome region to which the \textit{bbcl} gene has been assigned is, however, poorly endowed with genetic markers. Only one marker, a transfer RNA gene (tRNA:asp2:29E; Schmidt, \textit{et al.}, 1978), has been accurately mapped cytologically to the 29E-F region and lies at 29E.

Several protein-encoding genes have been mapped to regions of the second \textit{D. melanogaster} polytene chromosome near to the position of the \textit{bbcl} gene. Examples include the catalytic domain of protein kinase A at 30C1-6 (Kalderon and Rubin, 1988); the laminin B1 chain at 28D (Montell and Goodman, 1988); \textit{gurken} at 29C (Schüpbach, 1987); the gene for glutactin at 29D (Fessier and Fessier, 1989) and the \textit{numb} gene at 30B (Uemura \textit{et al.}, 1989). The last of these, \textit{numb}, is of interest because the numb protein displays some similarities to the predicted \textit{bbcl} protein. The predicted numb protein is highly basic, especially near the amino-terminus, and may be a nucleic acid binding protein. The \textit{numb} gene also shows a highly temporal-specific expression pattern during embryogenesis which is similar to that of the \textit{bbcl} gene (see sections 5.3 and 5.4). However, a comparison of the two protein sequences shows that they share only 15% amino acid identity over the length of the \textit{bbcl} protein, thus indicating that the proteins are unlikely to have similar functions.

A number of markers not corresponding to cloned genes have been genetically mapped to the area of chromosome 2 known to contain the \textit{bbcl} gene. \textit{tip (temperature induced paralysis)} C is one of six mutations that cause paralysis at elevated (38°C) temperature (Kulkarni and Padhye, 1982). \textit{eb (extra bristles)} was isolated as a spontaneous increase in the number of bristles on the mesonotum and scutellum (Mostashfi and Koliantz, 1969). \textit{d (dachs )} was also spontaneous in origin and results in abnormal leg, wing and head morphology (Held \textit{et al.}, 1988).
In addition to the above non-lethal mutations, several lethal mutations have been genetically mapped to the region of chromosome 2 containing the \textit{bbcl} gene. \textit{Su(var)} (Suppressor of variegation) 207, 213 and 215 are three of a large group of mutations which can suppress the position-effect variegation of \textit{In(1)w}^{m4} (Sinclair et al., 1983). \textit{rag} (\textit{rags}) is a maternal effect-lethal mutant in which embryos from homozygous mothers form fragmented cuticle with large ventral holes and head defects (Schüpbach and Wieschaus, 1989). \textit{mat(2)cell} (maternal lethal (2) cellularisation defects) RH36 and QC13 are two of a series of maternal effect-lethal mutations having similar phenotypes to \textit{rag} (Schüpbach and Wieschaus, 1989). \textit{I(2)pm} (lethal (2) polymorph) results in death of developing flies throughout larval and pupal stages. The major characteristic of this mutation is severe muscular dystrophy (Benz, 1953). Lastly, are three lethal mutations which have no reported phenotype other than lethality: \textit{I(2)K} (lethals from Kofu-Kaisunama, Japan) 202 and 204 (Watanabe and Oshima, 1970) and \textit{I(2)Sp} (lethal (2) of Spiess) 14 (Spiess et al., 1963).

The above mutations, which encompass approximately a division either side of the site occupied by the \textit{bbcl} gene have not been mapped cytologically and so cannot accurately be positioned on a physical map of the polytene chromosome. Hence, any could represent mutations of the \textit{bbcl} gene. Since mutations can be very small nucleic acid abnormalities, such as miss-sense or nonsense mutations, detailed analysis of the \textit{bbcl} gene in each of these mutants could be required in order to ascertain whether or not an altered \textit{bbcl} gene exists. This would represent a large amount of work, with no certainty of successfully identifying a mutation involving the \textit{bbcl} gene. For this reason, no further work was performed in this area.

5.3 \textit{In Situ} Hybridisation to \textit{D. melanogaster} Embryo Sections

The results of \textit{in situ} hybridisation of \textit{bbcl} sequences to \textit{D. melanogaster} embryo sections is shown in figure 5.3.1. The results
In situ hybridisation to D. melanogaster embryo sections. Hybridisation was performed as stated in chapter 2 using either an anti-sense (plates A, C, E, G, I, K, M, O, Q, S, U, W & Y) or a sense (plates B, D, F, H, J, L, N, P, R, T, V, X & Z) riboprobe generated from a 620bp EcoRl-Sphl fragment of D. melanogaster cDNA clone 6 (see figure 2.1.1) in plasmid pBluescript SK+. Plates are as follows:

A & B, lateral views of 0-2hr embryos (dorsal up, anterior right (A) and left (B));
C, lateral view of a 2-4hr embryo (dorsal up, anterior right);
D, dorso-ventral view of a 2-4hr embryo (anterior, left);
E & F, lateral views of 4-6hr embryos (dorsal up, anterior left (E) and right (F));
G & H, lateral views of 6-8hr embryos (dorsal up, anterior right);
I & J, lateral views of 8-10hr embryos (dorsal up, anterior left);
K & L, lateral views of 10-12hr embryos (dorsal up, anterior right (K) and left (L));
M & N, lateral views of 12-14hr embryos (dorsal up, anterior right (M) and left (N));
O & P, lateral views of 14-16hr embryos (dorsal upper left, anterior lower left);
Q, dorso-ventral view of a 16-18hr embryo (anterior lower left);
R, lateral view of a 16-18hr embryo (dorsal up, anterior left);
S, lateral (top, dorsal up and anterior left) and dorso-ventral (bottom, anterior right) views of 18-20hr embryos;
T, dorso-ventral view of a 18-20hr embryo (anterior right);
U, lateral view of an 18-20hr embryo (dorsal up, anterior right);
V, dorso-ventral view of an 18-20hr embryo (anterior left);
W & X, dorso-ventral views of 20-22hr embryos (anterior upper left);
Y & Z, lateral views of 22-24hr embryos (dorsal up, anterior right).

Arrows indicate probable areas of probe hybridisation: an, antennal sense organ; se, subepidermal labelling; ps, posterior spiracles; tr, trachea; mt, median tooth; ph, pharynx.

Note that the brown colouration apparent in the gut region of certain embryos (e.g. G, I, K, M, N and others) is an artefact and not a result of probe hybridisation to bcl RNA, since it is seen in sense and anti-sense hybridisations.
clearly show that the \textit{bbcl} gene is expressed in a highly temporal-
specific fashion during embryogenesis. Very little expression is
evident before 14hr, at which time expression is activated rapidly. A
high level of expression is maintained for about 6hr, then drops
quickly and is undetectable at 24hr of embryogenesis.

The spatial-specific expression pattern seen for the \textit{bbcl} gene is also
interesting since a periodicity of expression is apparent. This is
indicative of a gene which is expressed within each segment of the
developing embryo. The cells expressing \textit{bbcl} message in a periodic
fashion are located subepidermally, possibly within neurogenic or
myogenic cells (M. Akam, University of Cambridge, pers. comm.).

Apart from the periodic expression discussed above, other areas of
labelling are seen on the autoradiographs shown in figure 5.3.1. This
signal appears to coincide with the tracheal system, posterior
spiracles and mandibular structures. The labelling is very intense
and may either represent high levels of \textit{bbcl} gene expression in these
structures or artefactual signals arising from binding of the probe to
cuticular material being laid down in these areas (M. Akam, pers.
comm.). The rapid decrease in signal intensity which occurs
coincidentally with the reduction in segmentally repeated labelling
and the lack of detectable signal on control autoradiographs indicate
that this signal may not be artefactual.

\textbf{8.4 \textit{In Situ} Hybridisation to Whole Mount \textit{D. melanogaster}
Embryos}

In order to confirm and possibly extend the results obtained using
sectioned material, \textit{in situ} hybridisation with \textit{D. melanogaster}
embryos was performed according to a method kindly provided by Dr.
M. Akam (University of Cambridge). Digoxygenin-labelled riboprobes
were generated as described in chapter 2 and hybridised with whole
(i.e. not sectioned) \textit{D. melanogaster} embryos. Hybridisation was
detected with an anti-digoxygenin antibody conjugated to alkaline
phosphatase. The results of this analysis are shown in figure 5.4.1. As
with the results of \textit{in situ} hybridisation to sections, a highly
Figure 5.4.1

In situ hybridisation of DIG-labelled RNA probes to whole mount *D. melanogaster* embryos. Plates are as follows: A, lateral view of a stage 12 embryo (ventral surface uppermost, anterior left) incubated without probe or anti-DIG antibody; B, lateral view of a stage 15 embryo (anterior down, ventral right) incubated with anti-DIG antibody only; C, lateral view of stage 17 (about 16 hour, anterior lower left) embryo incubated with sense probe and anti-DIG antibody; D, same as C, but anterior upper right; E, F, G, lateral views of stage 9, 11 and 15 embryos respectively (ventral up, anterior right) incubated with anti-sense probe and anti-DIG antibody; H, I, J, three dorsal views at successive focussing planes through a stage 17 (about 16 hour) embryo (anterior bottom right) incubated with anti-sense probe and anti-DIG antibody; K, L, M, three ventral views at successive focussing planes through a stage 17 (about 16 hour) embryo (anterior left) incubated with anti-sense probe and anti-DIG antibody; N, O, P, three lateral views at successive focussing planes through a stage 17 (about 18 hour; note stage 17 encompasses the time from about 16hr to hatching, 24hr) embryo (ventral up, anterior left) incubated with anti-sense probe and anti-DIG antibody. Arrows indicate areas of probe hybridisation (not all areas of hybridisation are indicated): pm, posterior midgut; am, anterior midgut; hg, hindgut; mg, midgut; de, denticles; ps, posterior spiracles; tr, tracheae; an, antennal sense organ; ko, Keilin's sense organ; dbd, dorsal black dot sense organ; vdb, ventral black dot sense organ; ci, cirri; db, dorsal bridge; cp, cephalopharyngeal plates; csc, caudal sensory cone; mt, median tooth; gb, germ band.
temporal-specific expression pattern is observed, with peak expression levels at about 16hr of embryogenesis (plates H-P, stage 17, Campos-Ortega and Hartenstein, 1985). Lower levels of expression may occur at earlier stages of embryogenesis (stages 9 and 10), possibly within the developing posterior midgut and cells of the germ band (plates E-G). The expression seen at about 16hr is clearly present within the posterior spiracles and tracheal system as well as in denticles on the ventral epidermis. Expression is also apparent in anterior sensory organs such as the antennal and maxillary organs, the black dot sense organs, Keilin's sense organs and the cirri. Hybridisation is also apparent around structures of the cephalopharyngeal skeleton such as the median tooth, dorsal bridge and cephalopharyngeal plates. Many of the structures displaying hybridisation are also detected by the 35S-radiolabelled riboprobe (compare figures 5.3.1Q & S with figures 5.4.11 & J). This suggests that the signal is not artefactual.

A number of genes have been identified that are involved in development of the larval tracheal system, including empty spiracles (ems, Walldorf and Gehring, 1992), tailless (ll, Strecker et al., 1986), crumbs (crb, Tepaß et al., 1990) and the D. melanogaster homologue of the fibroblast growth factor receptor (DFGF-R, Glazer and Shilo, 1991). Ems is a gap gene that affects structures at the anterior and posterior ends of the developing embryos including the posterior spiracles and filzkörper and posterior parts of the tracheal tree (Walldorf and Gehring, 1992). Early expression of the gene is controlled by the anterior morphogen bicoid, expression being limited to the developing head region. Later during development (stages 11 to 14), expression is more widespread, including the eighth abdominal segment where the primordia of the posterior spiracles are located.

Tailless mutant embryos also show phenotypic abnormalities associated with the embryonic termini. The posterior defects result from absence of the three terminal segments A8, A9 and A10, with a concomitant expansion of the remaining abdominal segments. No filzkörper or posterior spiracles are present in these embryos and severe disruption of the posterior tracheal tree also occurs (Strecker
et al., 1986). The posterior tissues affected by *tailless* mutations correspond well with regions of the blastoderm that express *tailless* mRNA and give rise to these organs (Pignoni et al., 1990). The *tailless* gene is also expressed in central and peripheral nervous system cells later during development and may perform an important function in these cells. The *tailless* protein appears to be a steroid receptor-type transcription factor which is expressed in response to activation of the *torso* receptor tyrosine kinase by the *torso-like* protein ligand (Sprenger et al., 1989; Stevens et al., 1990).

Interestingly, the anterior structures affected by *tailless* mutations, the cephalopharyngeal skeleton, also appear to demonstrate *bbcl* gene expression. The phenotypic abnormalities of the cephalopharyngeal skeleton seen in *tailless* mutant embryos result from shortened pharyngeal ridges (Strecker et al., 1986), all other skeletal structures being normal. Whether this observation represents a functional link between the two genes or merely a coincidence will require further analysis.

The *crumbs* gene encodes a membrane-spanning EGF-like protein which is expressed on the apical surface of epithelial cells (TepaB et al., 1990). This protein may act as an intercellular signalling molecule that is involved in formation/maintenance of ectodermally-derived epithelial structures, including the tracheal system which shows widespread disruption in *crumbs* mutant embryos (TepaB et al., 1990).

The *D. melanogaster* homologue of the FGF-R is particularly interesting in the context of tracheal development, because it is specifically expressed in this organ (and in delaminating midline glial and neural cells). Loss of function mutations within the DFGF-R gene leads to blocking of extension of tracheal cell processes that normally result in formation of the elaborate tree structure (Glazer and Shilo, 1991). It is proposed that the DFGF-R participates in receiving spatial signals that guide tracheal cell outgrowths (Glazer and Shilo, 1991). This gene is particularly interesting, because the vertebrate homologue has been implicated in angiogenesis, a process very similar to tracheal cell outgrowth.
Since the process of angiogenesis is vital for all tumours to attain a size of greater than a few cubic millimetres (Fidler, 1991), any gene involved in angiogenesis is of potential use in the development of anti-tumour drugs. Establishment of whether the human \( bbcl \) gene-product is involved in angiogenesis will require substantial work. However, the possible role of the \( Drosophila \ bbcl \) gene in tracheal development indicates that this is, at least, a possibility.

Several genes in \( D. \ melanogaster \) have been shown to affect peripheral nervous system development. These include \( Delta \ (Dl) \), \( Cut \) and \( numb \). \( Dl \) is a neurogenic gene expressed by all neurogenic ectodermal cells and which mediates commitment of non-neuronal cells to an epidermal fate (Vässin et al., 1987). Like the crumbs protein, the Delta protein is a transmembrane protein containing EGF-like repeats that may act as an intercellular signalling molecule determining the fate of non-neuronal ectodermal cells.

The Cut protein is a homeo domain-containing protein which is expressed in the cells of external sensory organs but not chordotonal (stretch) organs (Blochlinger et al., 1990). The protein is a nuclear transcription factor which may induce genes involved in external sensory organ differentiation.

5.5 Conclusions

\textit{In situ} hybridisation to \( D. \ melanogaster \) polytene chromosomes has determined a cytologic location for the \( bbcl \) gene in this organism. Although a number of uncloned markers have been genetically mapped to this region of the genome, a lack of information concerning their physical position on the polytene chromosome prevented analysis aimed at determining whether any of these markers involved the \( bbcl \) gene. Despite this somewhat disappointing result, the elucidation of a cytologic position for the \( bbcl \) gene could be of considerable use during future mutagenesis studies by allowing crosses with strains bearing deficiencies in this region of the chromosome.
In situ hybridisation to *D. melanogaster* embryo sections and whole mount embryos has identified extremely interesting temporal- and spatial-specific expression patterns of the gene during embryogenesis. The expression profiles indicate that the gene product may play a role in tracheal and external sensory organ development at about 14-18hr of embryogenesis. The gene may also function at earlier stages of development, although the expression level at these times appears much less than the peak expression observed. However, the lack of a known *D. melanogaster* strain carrying a phenotype-inducing mutation of the *bbcl* gene precludes a correlation between expression patterns and function at the present time.
CHAPTER 6

Production of Antibodies to the *D. melanogaster* bbcl Protein

6.1 Introduction

To investigate the function(s) of the *D. melanogaster* bbcl gene product, antibodies were raised against an *E. coli*-produced fusion protein containing the *D. melanogaster* protein sequence and against synthetic peptides. The fusion protein and peptides were used to raise rabbit and chicken polyclonal antibodies respectively.

6.2 Construction of the pDM1 Expression Plasmid

The pGEX cloning vectors have been designed for inducible high-level expression of coding sequences as C-terminal fusion proteins with a 26-KDa glutathione S-transferase (GST) from *Schistosoma japonicum*. The GST protein (Sj26) is expressed from the hybrid trp-lac (tac) promoter which drives high-level expression. In the absence of an inducer (such as IPTG), expression is almost entirely prevented by the lac repressor (lacI) which is encoded on the plasmid by the over-expressing lacI° allele (part of the lacZ gene was also introduced with the lacI° allele, although it has no function in the plasmid). Three versions of the vector are available (corresponding to the three possible reading frames), two of which contain a specific proteolytic cleavage site engineered into the C-terminal region of the GST protein (Smith and Johnson, 1988). Expression of the *D. melanogaster* bbcl protein utilised the pGEX-3X version of the vector which contains a blood-clotting factor Xa cleavage site.

The cloning strategy is summarised in figure 6.2.1. Briefly, the ORF within *D. melanogaster* cDNA clone 6 was amplified by PCR utilising the ORFPCR 3' and 5' primers (see figure 2.1.1). The PCR primers ORFPCR 5' and ORFPCR 3' contained recognition sites for SstII and...
Figure 6.2.1

Strategy for construction of the pDM1 expression plasmid containing the ORF from *D. melanogaster* bbd cDNA clone 6. See text for details. For details of the pGEX vector, see the text and Smith and Johnson, (1988).
BamHI restriction endonucleases respectively. These sites were originally chosen to allow cloning of the PCR product into a vector other than the pGEX vector. However, the cloning strategy described below allowed the bbcl ORF to be cloned into the pGEX vector such that all of the bbcl protein except for the initiating methionine would be produced as a fusion with the GST protein. Cleavage of the PCR product with SstII and BamHI restriction endonucleases was followed by treatment with the DNA polymerase from bacteriophage T4. This removed the protruding 3' single-stranded DNA at the SstII site and filled-in the protruding 5' single-stranded DNA at the BamHI site to create blunt ends. The pGEX-3X vector was cleaved with BamHI endonuclease followed by fill-in of the single-stranded regions with Klenow polymerase to generate blunt ends. Ligation of the vector and insert DNAs resulted in an in-frame fusion of the bbcl and GST coding regions to generate plasmid pDM1.

6.3 Expression of the Fusion Protein in E. coli

The pDM1 construct was transformed into E. coli strain DH5α. Restriction enzyme mapping and sequencing across the vector-insert junction confirmed that an in-frame fusion of the bbcl and GST coding sequences was present within the construct. Induction of cultures containing this plasmid with IPTG resulted in production of an approximately 40-KDa protein. Whilst this is considerably larger than the size of native GST protein, it falls short of the predicted fusion protein size of approximately 50-KDa. Further sequence analysis demonstrated that the bbcl sequence within the expression construct contained a single base-pair deletion relative to the original bbcl sequence. This deletion resulted in a frame-shift that caused premature termination of translation to generate a fusion protein of approximately 40-KDa in size. The deletion was found to be present in the clone (cDNA 6) from which the ORF was amplified by PCR.

In order to remove the segment of DNA containing the deletion, use was made of unique Sphi and EcoRI sites within both the bbcl cDNA clone 4 and the pDM1 construct (see figures 2.1.1 and 6.2.1).
Cleavage of the pDM1 construct and \textit{D. melanogaster} cDNA clone 4 with both enzymes followed by ligation of the desired sequences (figure 6.3.1) resulted in production of the pDM1* construct. This procedure not only removed the deletion, but also replaced most of the PCR-generated sequence. Sequencing of the insert on both strands verified that the deletion had been corrected and induction of cultures containing the altered construct produced a fusion protein of the predicted size (see figure 6.3.2a). In an attempt to maximise fusion protein production, the expression construct was transformed into \textit{E. coli} strain Y1090. This strain of \textit{E. coli} lacks the \textit{lon} protease and is routinely used for expression of exogenous proteins (Young and Davis, 1983). A time course for induction of Y1090 cultures containing the construct showed that maximal fusion protein production occurs after 4-5hr of growth at 37°C (Figure 6.3.2a).

### 6.4 Purification of GST-bbcl Fusion Protein

Expression of proteins as fusions with GST allows one-step purification of the protein in most cases. Glutathione is a substrate for GST and, when immobilised on agarose beads, allows affinity-purification of the native GST protein and soluble fusion proteins containing GST (Smith and Johnson, 1988). Fusion proteins must be soluble or be capable of being solubilised without destroying the GST tertiary structure in order to bind the glutathione agarose (Smith and Johnson, 1988).

An initial attempt at purification of the GST-bbcl fusion protein from \textit{E. coli} (strain Y1090) cells failed (data not shown) due to insolubility of the protein in cells induced for 4hr. Hence, although maximal protein production occurs after 4-5hr of induction, it appears that all or virtually all of the protein is insoluble. As a result, induction time was shortened to 1hr and purification performed as before. Figures 6.3.2b and 6.4.1 show the results of affinity purification of the fusion protein from a culture of \textit{E. coli} strain Y1090 grown for only 1hr in the presence of IPTG. Small amounts of protein corresponding in size to the fusion protein can be seen in the pellet fraction (lane 2, figure 6.3.2b), thus indicating that
Summary of the method used to generate the pDM1* expression construct. The EcoRI insert from D. melanogaster cDNA 4 was cleaved with SphI restriction endonuclease to generate two fragments. The 576bp fragment was recovered. The pDM1 construct was cleaved with EcoRI and SphI restriction endonucleases and the plasmid sequence purified away from the 620bp fragment containing bbcl sequence. Ligation of the 576bp fragment from clone 4 into the cleaved pDM1 plasmid yielded plasmid pDM1*.
Figure 6.3.2

Expression and purification of the GST-bbcl fusion protein in *E. coli*. Samples were separated on a 12% polyacrylamide SDS-PAGE gel using the Pharmacia LMW calibration kit for size determination. Panel a shows a time course for induction of *E. coli* strain Y1090 containing the expression construct. 6μg of total cell protein from each time point was loaded onto the gel. Numbers above individual tracks show the time, in hours, of induction. The fusion protein band is indicated (arrow). Marker sizes are in KDa.

Panel b shows stages in the affinity purification of the fusion protein. Tracks are: 1 & 2, soluble and insoluble proteins after sonication respectively (5μg each); 3, proteins not bound to glutathione agarose (5μg); 4, eluate from glutathione agarose (0.6μg).
Figure 6.4.1

SDS-PAGE analysis of GST and GST-bbcl fusion proteins purified from *E. coli* Y1090 cell extracts. Protein preparations were separated by electrophoresis through a 12.5% polyacrylamide-SDS gel and visualised by staining with Coomassie blue. Tracks are: 1 & 4, total cell extracts from cultures of *E. coli* not expressing the GST and GST-bbcl fusion proteins respectively; 2 & 5, total cell extracts from cultures of *E. coli* expressing the GST and GST-bbcl fusion proteins respectively; 3 & 6, purified GST and GST-bbcl fusion proteins respectively. Tracks 1, 2, 4 & 5 contain 5μg of protein, tracks 3 & 6 contain 0.5μg of protein. Marker sizes are in KDa.
not all of the fusion protein produced after 1hr of induction is soluble. Also, a band corresponding in size to the fusion protein is present in the fraction which did not bind to the glutathione agarose (lane 3, figure 6.3.2b), indicating that the beads did not bind all of the fusion protein (see below).

It is evident from figure 6.3.2b that a considerable degree of purification of the fusion protein has been obtained. The lower molecular weight bands which can be seen clearly in figure 6.4.1 appear to correspond to *E. coli* proteins which co-purify with the fusion protein and not to fusion protein degradation products. The predicted highly basic nature of the bbcl portion of the fusion protein confers a high predicted pl (9.8) on the fusion protein. Thus at the pH used during extraction (7.3), the fusion protein is expected to have a substantial positive charge. This may cause *E. coli* proteins with net negative charges to bind to the fusion protein and co-purify. The contaminating proteins do not appear to bind to the GST protein, since purification of this protein using the glutathione agarose beads does not result in detectable contamination (Figure 6.4.1).

Several hundred micrograms of fusion protein can be obtained from a 100ml culture of Y1090. The limiting factor in purifying greater amounts of fusion protein appears to be the binding capacity of the glutathione beads. The beads are claimed to have a capacity in excess of 8mg of protein/ml of swollen beads (Smith and Johnson, 1988), but in my hands their capacity was substantially lower than this. The reason for this is unclear, but may result from the bbcl portion of the fusion protein inhibiting GST binding to the glutathione agarose.

### 6.5 Raising and Characterisation of Polyclonal Rabbit Antibodies Against the Fusion Protein

Increasing the quantity of glutathione agarose beads used during the purification procedure allowed about 1mg of fusion protein to be isolated. The protein was used to immunise two New Zealand White
rabbits as described in chapter 2 (section 2.5.1). Test bleeds were taken after each boost injection.

Preliminary characterisation of the antibodies involved probing Western blots of *E. coli* extracts from cultures expressing either the GST or fusion protein as well as cultures not expressing these proteins. SDS-PAGE analysis and Western transfer were performed according to the manufacturer's (Bio-Rad) protocol. Despite the presence of visible fusion protein in the SDS-PAGE gel, no signal could be detected on the nitrocellulose membrane when this was probed with the 317 antiserum (not shown). The GST protein was, however, easily detected. The probable explanation for this soon became apparent when the gel and filter were stained to reveal proteins. Fusion protein was present in the gel, but absent from the nitrocellulose. The likely reason for this is the predicted high pi of the fusion protein (9.8). The Tris/glycine transfer buffer routinely used for Western transfer and specified in the protocol used, has a pH of 8.3. Whilst this is above the pi of most proteins, it is over one pH unit below that of the fusion protein. Since good transfer generally requires a buffer pH of at least one unit above the pi of the protein being transferred, this may account for the lack of transfer seen.

In an attempt to resolve the above problem, a transfer buffer with a pH of 10 was used (see chapter 2, section 2.4.6). This buffer has previously been shown to result in efficient transfer of highly basic proteins, including histones, from SDS-PAGE gels to nitrocellulose (Szewczyk and Kozloff, 1985). Use of this buffer did, indeed, yield visible transfer of fusion protein from the gel to the nitrocellulose (figure 6.5.1b). The transfer was not complete and a substantial proportion of the fusion protein remained in the gel (figure 6.5.1a). Probing this blot with antiserum 317 (preabsorbed with an extract of *E. coli* induced to express the GST protein) resulted in detection of the fusion protein (figure 6.5.1c). Note that the GST protein is not recognised by the antiserum despite being present on the membrane. This is due to preabsorption of the antiserum (see below).

Having demonstrated that antibodies were being generated against the fusion protein, the rabbits were exsanguinated after the second
Transfer of proteins onto nitrocellulose membranes and detection of the GST-bbcl fusion protein using antiserum 317. a. Coomassie blue-stained 12.5\% polyacrylamide SDS-PAGE gel after transfer in pH 10 buffer as described in chapter 2. Tracks are: 1, purified GST-bbcl fusion protein; 2 and 3, total cell protein extracts from cultures of *E. coli* expressing and not expressing the fusion protein respectively; 4, purified GST protein; 5 and 6, total cell protein extracts from cultures of *E. coli* expressing and not expressing the GST protein respectively. Tracks 2, 3, 5 and 6 contain 6\(\mu\)g of total cell protein. Tracks 1 and 4 contain 0.5\(\mu\)g of protein. Marker sizes are in KDa. b. Nitrocellulose membrane stained to reveal proteins transferred from the SDS-PAGE gel in panel a. Tracks are as in panel a. c. Membrane identical to that shown in panel b, incubated with antiserum 317 (1/2000 dilution, preabsorbed to remove GST and *E. coli*-reactive antibodies) and developed as described in chapter 2.
boost injection. Tests of these antisera demonstrated that both recognised the fusion protein. However, antiserum 316 gave much weaker signals than 317 and so was not used for subsequent analyses.

To remove unwanted antibodies (both *E. coli* and GST reactive) from the antiserum, the serum was preabsorbed with nitrocellulose-immobilised extracts of *E. coli* expressing the GST protein. The presence of *E. coli*-reactive antibodies within the 317 antiserum was expected, because the fusion protein extract used for injections was not totally free of *E. coli* proteins. Also, the rabbits probably experienced *E. coli* infections prior to antiserum collection.

Antiserum was also preabsorbed with extracts of *E. coli* expressing the GST-bbcl fusion protein to demonstrate that bands detected on Western blots contained bbcl protein. Figure 6.5.2 shows that preabsorption of the antiserum effectively removes reactivities to both the GST and fusion proteins at dilutions of 1/2000 of the antiserum. Importantly, removing antibodies capable of reacting with the GST protein did not prevent detection of the fusion protein. This clearly demonstrates that the 317 antiserum contains antibodies reactive towards the *D. melanogaster* bbcl protein.

6.6 Production of Synthetic Peptides and Generation of Anti-peptide Antibodies

Several peptides were produced by Dr. A. Moir (Sheffield University) corresponding to both conserved and divergent regions of the *D. melanogaster* and human bbcl proteins (figure 6.6.1). The *D. melanogaster* peptides (174 and 476) were coupled to carrier proteins (porcine thyroglobulin and BSA) by the method given in chapter 2. These conjugates were then used to inject chickens as detailed in chapter 2.

Preliminary characterisation of the antibodies used dot blots containing various protein preparations. Figure 6.6.2 shows such a blot. It can be seen that a response to the injected thyroglobulin-
Preabsorption of the 317 rabbit polyclonal antibody. 0.5μg of purified GST (G) and GST-bbcl fusion proteins (F), and 5μg of 0-24hr D. melanogaster embryo protein (E) were loaded in quadruplicate on a 12% polyacrylamide-SDS gel. After electrophoretic separation, the proteins were transferred onto nitrocellulose by Western blotting. The filter was cut into four strips each containing the three different protein preparations and incubated with the following primary antibodies: I, rabbit 317 pre-immune antibody (1/500 dilution); II, 317 post-immune antibody (1/2000 dilution); III, 317 post-immune antibody (1/2000 dilution, preabsorbed with an extract of E. coli expressing the GST protein); IV, 317 post-immune antibody (1/2000 dilution, preabsorbed with an extract of E. coli expressing the GST-bbcl fusion protein). All strips were then incubated with the same secondary antibody (1/1000 dilution horse radish peroxidase-conjugated goat anti-rabbit IgG) and developed as described in chapter 2. The fusion (F), GST (G) and probable D. melanogaster bbcl (B) proteins are indicated by arrows.
Figure 6.6.1

Comparison of the predicted *D. melanogaster* (D) and human (H) bcl1 proteins showing positions of synthetic peptides (=). Identical amino acids are indicated by vertical lines and conservative changes by dots. Note that each peptide also contains a C-terminal cysteine residue to facilitate coupling of the peptide to the carrier protein.
Figure 6.6.2

Dot blot analysis of various protein preparations probed with chicken antipeptide antibodies. Rows are: 1, thyroglobulin/peptide 174 conjugate; 2, BSA/peptide 174 conjugate; 3, thyroglobulin/peptide 476 conjugate; 4, BSA/peptide 476 conjugate; 5, thyroglobulin; 6, BSA; 7, GST/bbcl fusion protein; 8, GST protein; 9, E. coli total cell protein. 1μg of each protein preparation was applied to the filter. Columns were probed with either pre-immune antibodies (I) from four different chickens (H-K) or post-immune antibodies (II) from the same chickens. All antibodies were used at 1/1000 dilution. Chickens H and K were immunised with thyroglobulin/peptide 476 conjugate and chickens I and J were immunised with thyroglobulin/peptide 174 conjugate. The secondary antibody was a rabbit anti-chicken horseradish peroxidase conjugate used at 1/1000 dilution. All incubations were for 1hr at room temperature. The blot was developed as described in chapter 2 for Western blots.
peptide conjugate has been achieved (rows 1 and 3). It can also be seen that a response to the BSA/peptide conjugate is present (rows 2 and 4). Hence it appears that the peptide has raised an immune response and that this response is specific to the peptide (no response is seen to a control of BSA alone (row 6)). Despite this response, it appears that the antibodies are incapable of recognising the fusion protein (row 7). Likewise, no response is seen to *D. melanogaster* protein extracts either on dot blots or Western blots (not shown). Thus, although antibodies are present which recognise the peptides, these antibodies cannot recognise the same sequence when it is present within the *D. melanogaster* bbcl protein. This problem may be due to the peptide sequence being masked within the bbcl protein or because the peptide adopts a different conformation within the bbcl protein than that adopted on the carrier protein. Due to the lack of response to *D. melanogaster* bbcl protein, the chicken anti-peptide antibodies have not been used for subsequent experiments.

6.7 Conclusions

Cloning of the bbcl ORF into the pGEX-3X expression plasmid allowed a GST-bbcl fusion protein to be generated which probably contained the entire bbcl protein sequence fused to the C-terminus of the GST protein. This protein could be purified using a single-step affinity method to produce a preparation suitable for inoculation. Immunisation of rabbits with the preparation resulted in generation of an antiserum having a titre in excess of 1/2000 when used on Western blots. Preabsorption of the antiserum demonstrated that antibodies capable of recognising the bacterially produced bbcl protein were present.

Synthetic peptides were made and coupled to carrier proteins prior to injection into chickens. The antibodies recovered from eggs laid by these chickens were capable of recognising the peptide, but unfortunately did not recognise the GST-bbcl fusion protein nor the bbcl protein within *D. melanogaster* protein extracts.
CHAPTER 7

Characterisation of the *D. melanogaster* bbcl Protein

7.1 Introduction

Antibodies provide powerful tools for the characterisation of novel molecules. Techniques such as Western blotting, immunoprecipitation and immunohistochemistry can allow many attributes of a molecule to be determined. The 317 polyclonal antiserum has, therefore, been used to study various aspects of the *D. melanogaster* bbcl gene product.

7.2 Western Blotting Analysis of *D. melanogaster* Proteins

To determine whether the 317 antiserum was able to recognise *D. melanogaster* bbcl protein, preparations from various stages of *D. melanogaster* development were separated by SDS-PAGE and Western blotted onto nitrocellulose. Probing such a Western blot with antiserum 317 (preabsorbed with an extract of *E. coli* containing the GST protein) resulted in identification of a single strong band in the track containing *D. melanogaster* embryo proteins (figure 7.2.1). Weak bands are visible in the tracks containing proteins from other stages of development (see also figure 7.3.1, below). Probing blots with either the pre-immune serum or post-immune serum preabsorbed with an *E. coli* extract containing the GST-bbcl fusion protein failed to identify this band (see figures 6.5.2 and 7.6.1a), thus indicating that this band corresponds to the *in vivo* product of the *D. melanogaster* bbcl gene.

A difference between the observed size (28 KDa) and the predicted size of the *D. melanogaster* bbcl protein (25 KDa) is apparent. This may result either from post-translational modification of the protein or from the predicted highly basic nature of the protein.
Figure 7.2.1

a. Separation of *D. melanogaster* proteins by SDS-PAGE. 5μg of adult (A), third instar larval (L) and 0-24hr embryonic (E) *D. melanogaster* proteins were separated on a 12.5% polyacrylamide gel and stained with coomassie blue to visualise the proteins. 0.5μg of GST-bbcl fusion protein preparation (F) was included as a positive control. Size markers (M) are as used in figure 6.3.2; sizes in KDa. b. Western blotting analysis of *D. melanogaster* and GST-bbcl fusion proteins separated as in panel a. and transferred to a nitrocellulose membrane. The blot was divided into two and incubated with either pre-immune 317 antiserum at a dilution of 1/500 (I) or post-immune 317 antiserum (preabsorbed with an extract of *E. coli* expressing the GST protein) at a dilution of 1/2000 (II). Development was as described in chapter 2.
The D. melanogaster bbcl Gene

causing it to migrate more slowly through the gel than its size would predict. A detailed analysis of the in vivo product of the D. melanogaster bbcl gene will be required to resolve this matter.

7.3 Temporal Expression of the bbcl Protein During Embryogenesis and Later Development

The identification of a strong band within proteins extracted from D. melanogaster embryos verified the prediction made from studies of the bbcl mRNA that this protein is expressed at higher levels during embryogenesis than during other stages of the D. melanogaster life cycle. To determine at which time(s) during embryogenesis the protein is expressed, a Western blot of proteins prepared from embryos collected at specific times of development was probed with the 317 antiserum. Figure 7.3.1 shows the result of this experiment. It is apparent from this blot that levels of the protein remain remarkably constant throughout embryogenesis. The apparently large fluctuations in mRNA levels do not appear to be translated into altered levels of protein.

There are several possible explanations for this observation. The apparent high levels of mRNA expression seen in the in situ hybridisation studies (chapter 5) may be artefactual. As discussed in that chapter, this seems unlikely, but should not be discounted as an explanation for the observed protein levels. Conversely, the band being detected on Western blots may not be the product of the bbcl gene. Again, this seems unlikely given the results of control reactions performed with the antiserum. Alternatively, differences in the strains of D. melanogaster or culture conditions used may result in the differences observed. Whilst a different strain was used to prepare the sections (wild-type Sevelin) than was used for the whole mount studies and the protein preparations (wild-type Brighton), all embryos were maintained at the same temperature (25°C). As far as the whole mount studies and protein preparations are concerned, not only was the strain the same, but the embryos were collected from the same batch of flies kept under constant conditions for the entirety of the collection period.
Figure 7.3.1

a. Separation of *D. melanogaster* proteins by SDS-PAGE. 8µg of each protein preparation was separated on a 12% polyacrylamide gel then stained with coomassie blue to visualise the proteins. Tracks are: 1, 0-4hr embryos; 2, 4-8hr embryos; 3, 8-12hr embryos; 4, 12-16hr embryos; 5, 16-20hr embryos; 6, 20-24hr embryos; 7, third instar larvae; 8, pupae; 9, adult. Marker sizes are in KDa.

b. Western blotting analysis of an identical gel to that shown in panel a. After separation, proteins were transferred to a nitrocellulose membrane. The membrane was incubated with 317 post-immune antiserum at a dilution of 1/2000, then with horseradish peroxidase-conjugated anti-rabbit IgG (1/1000 dilution) and developed as described in chapter 2. The likely *D. melanogaster* bbcl protein is indicated (arrow). Marker sizes are in KDa.
If the above explanations are excluded, then it seems likely that the disparity between mRNA and protein levels is real. Due to the fact that an analysis such as that described above cannot determine protein turn-over rates, it is entirely possible that large fluctuations in protein production do occur, but that similar fluctuations in proteolysis also occur, thus resulting in relatively minor changes in the steady-state level of observed protein. Another possibility is that translational repression occurs. This process has been well documented for a number of systems, for example ferritin production (Aziz and Munro, 1987; Walden et al., 1988) and ribosomal protein synthesis (Nomura et al., 1984). Analysis of bbcl protein synthesis/degradation rates at different stages of embryogenesis will be required to resolve this matter. A Northern blot of RNA from different stages of embryogenesis would also be of use, in order to validate the results of the in situ hybridisation data.

The protein detected at early stages of embryogenesis does not appear to correspond to a period of transcriptional activity as determined by in situ hybridisation analyses (see chapter 5). However, it should be noted that a low level of transcription is unlikely to be detected due to the relatively short autoradiography time used to optimise signal intensity at later stages of embryogenesis. The same is true of the non-radioisotopic detection method. The possibility also exists that some of the protein detected in early stages of embryogenesis is maternally produced bbcl protein. A large number of proteins which perform functions early in embryogenesis have been shown to be produced from maternally-expressed genes (for reviews see Manseau and Schüpbach, 1989 and Govind and Steward, 1991). An in situ hybridisation analysis of adult D. melanogaster would be required to determine whether ovarian expression of the bbcl gene occurs.

The protein is also detectable at much lower levels in certain later stages of the D. melanogaster life cycle. A faint band is seen in protein preparations from both third instar larvae and adults, but not in pupal proteins (figure 7.3.1). The protein detected in third instar larvae could be residual protein remaining from embryonic
development or newly synthesised protein. Without knowing the turnover rate of the protein, a distinction cannot be made between these two possibilities. The results do, however, broadly agree with the results obtained from the developmental Northern blot (see chapter 3, figure 3.3.2a) which shows high levels of *bbcl* mRNA in embryos, steadily decreasing levels in third instar larvae and pupae and increased (relative to larvae and pupae) amounts in adults. Hence, it is possible that the protein detected during each developmental stage represents *de novo* synthesis of protein. An understanding of the observed differences in expression throughout development will require a knowledge of the *bbcl* protein function.

7.4 Immunohistochemical Analysis of *D. melanogaster* Embryos Using the 317 Antiserum

Expression of the protein during embryonic development was predicted from the results of the library screen (chapter 3) and from the *in situ* hybridisation study performed on *D. melanogaster* embryos (chapter 5). The results described above (section 7.3) determined the temporal expression pattern of the *bbcl* protein during embryogenesis. In an attempt to determine the spatial expression pattern of the protein during embryonic development and to see whether this parallels the expression pattern of the mRNA, the 317 antiserum was used on fresh-frozen embryo sections in an immunohistochemical analysis. Unfortunately no specific signals were observed, although a general background signal was seen (even in controls not containing a primary antibody). Possible reasons for this include masking of the epitopes recognised by the antiserum or lack of reactivity of the antiserum with native protein. Experiments which have been performed in an attempt to determine whether the antiserum recognises the native protein are described below.

7.5 Immunoprecipitation Analysis of *D. melanogaster* Proteins Using the 317 Antiserum

In an attempt to determine whether the antiserum recognised the native protein, immunoprecipitation experiments were performed. A
SL2 *D. melanogaster* embryonic cell line obtained from Dr. B. Fenton (Dept. of Biochemistry, University of Dundee) was shown to express the bbcl protein by Western blotting analysis (figure 7.5.1b). In order to verify that the bbcl protein was retained within the cells and not released into the culture medium, proteins from the cells and conditioned medium were prepared. The conditioned medium was concentrated using an Amicon micro-concentrator prior to analysis. All protein concentrations were estimated by the method of Bradford (see chapter 2) prior to SDS-PAGE analysis and Western transfer onto membranes. Probing these membranes with the 317 antiserum only identified a band in the cellular fraction (figure 7.5.1b), thus indicating that the bbcl protein remains within the cell and is not secreted.

SL2 cells were metabolically labelled using $^{35}$S-methionine and $^{35}$S-cysteine. Cellular proteins were prepared by either Dounce homogenisation or detergent lysis to generate native and denatured proteins respectively and subjected to immunoprecipitation using the 317 antiserum. SDS-PAGE analysis failed to reveal a protein corresponding in size to that detected by Western blot analysis. This may indicate that the 317 antiserum is incapable of recognising the native *D. melanogaster* bbcl protein. Alternatively, the method used for immunoprecipitation may have been at fault.

To determine whether the immunoprecipitation method was functioning correctly, a positive control antiserum (rabbit polyclonal anti-cyclin B) was obtained from Dr. B. Fenton and used in immunoprecipitation reactions. Despite performing the immunoprecipitation reaction according to two different protocols (see chapter 2, section 2.4.10), no proteins were precipitated by this antiserum or the 317 antiserum.

Under the conditions used to grow the cells, i.e. with 1$\mu$M colchicine (section 2.4.10), levels of cyclin B protein were produced that could easily be detected by the antibody on Western blots (figure 7.5.1a). Addition of colchicine to the SL2 cells resulted in raised levels of cyclin B protein relative to control cultures (Whitfield *et al.*, 1990) and had no affect on the levels of bbcl protein (compare figures
Western blotting analysis of proteins from *D. melanogaster* SL2 tissue culture cells using anti-cyclin B and anti-bbcl antisera. 5μg protein samples were separated by SDS-PAGE on 12.5% polyacrylamide gels, then transferred to nitrocellulose membranes and incubated with antiserum at the stated dilution. Development was as described in chapter 2. a. Analysis using anti-cyclin B 271 antiserum (1/500 dilution). Tracks are: M, markers; 1, 2, 3 and 4, protein from SL2 cells (10^7/ml) cultured in Schneider's medium containing 5% dialysed foetal calf serum (FCS) and 1μM colchicine for 0, 2, 4 and 6hrs at 25°C respectively; 5, protein from SL2 cells (10^7/ml) cultured for 6hr at 25°C in medium containing 5% dialysed FCS only. b. Western blotting analysis using anti-bbcl 317 antiserum (1/2000 dilution). Tracks are as in a., with additional tracks: 6, SL2 conditioned medium (SL2 cells (10^7/ml) grown for 24hr at 25°C in Schneider's medium containing 5% dialysed FCS); 7, 0-24hr *D. melanogaster* embryo protein.
The cyclin B protein contains 12 methionine residues and one cysteine residue and the bbcl protein contains 2 methionine and 2 cysteine residues. Hence both proteins should become radio-labelled with $^{35}$S-methionine and cysteine during the labelling period. Incorporation of label into cellular proteins was confirmed by the presence of radio-labelled background proteins in two immunoprecipitation reactions (not shown).

The results of the immunoprecipitation experiments were especially disappointing due to failure of the positive control. Two explanations could account for this: failure of the cyclin B protein to become radio-labelled or proteolysis of this protein after labelling. If the cyclin B protein is produced very slowly, it is possible that insufficient protein is produced during the labelling period to allow detection. *D. melanogaster* cell lines grown at 25°C have a cell cycle time of approximately 20-24hr (Sang, 1981). Hence, under the conditions used (4hr incubation with colchicine) where cells were arrested once in m-phase, most cells would have only completed one sixth of a cell cycle. This would result in approximately one sixth of the maximum amount of cyclin B production, possibly accounting for the lack of detection obtained. Whilst the incubation time for cell labelling can be increased, more cysteine and methionine must be added to the culture medium after about 6hr (Harlow and Lane, 1988). Addition of $^{35}$S-labelled amino acids every 6hr for 24hr would be prohibitive with respect to the amount of labelled amino acids used and addition of unlabelled amino acids is nonproductive.

Immunoprecipitation reactions performed by Dr. Fenton did not rely on production of radiolabelled protein, since an alternative method of detection was used (assay of maturation promoting factor protein kinase activity). Not only is it possible that this assay is more sensitive than autoradiographic detection of radiolabelled proteins but the entire cellular complement of cyclin B would be assayed, not just the radiolabelled fraction.

Proteolytic degradation of the cyclin B protein is a possible reason for the lack of detection observed. However, the extraction and immunoprecipitation reactions were performed on ice in the presence
of protease inhibitor. What is more, proteins extracted for Western blotting using the same method do not show excessive proteolytic degradation (figures 7.5.1a and b). It should be noted, however, that proteins processed for Western blotting were not subjected to incubations on ice lasting as long as those required for the immunoprecipitation reaction.

The reason(s) for failure of the immunoprecipitation reactions are not known, but may be due either to low levels of radiolabelled protein being produced and/or proteolysis of this protein subsequent to production. Due to time constraints, further attempts at immunoprecipitating the D. melanogaster bbcl protein have not been performed.

7.6 Western Blotting Analysis of Human Proteins Using the 317 Antiserum

The 317 antiserum has been used on Western blots containing human protein preparations. Proteins from human placenta, HeLa cells and three human breast tumour cell lines (T47D, Hs 578T and MCF 7) have been separated by SDS-PAGE and transferred to membranes. The 317 antiserum failed to identify a candidate human bbcl protein in these preparations at dilutions between 1/1000 and 1/200 of the antiserum (figures 7.6.1a and b). Northern blotting analysis has previously detected bbcl mRNA within RNA isolated from several of these sources, indicating that an inability of the antibody to recognise the human bbcl protein may be responsible for the lack of signal obtained. Given the high degree of similarity between the two proteins (over 70%), it was hoped that antibodies to the D. melanogaster protein would be capable of recognising the human protein. Whilst it has not been proven that the 317 antiserum is incapable of recognising the human protein, it seems likely that this is the case.

7.7 Conclusions

The 317 polyclonal antiserum raised against the D. melanogaster bbcl
Figure 7.6.1

Western blotting analysis of human proteins. 5μg of each protein preparation (0.5μg fusion protein) was separated by SDS-PAGE using 12.5% polyacrylamide gels and transferred to nitrocellulose membranes prior to probing with antisera at the stated dilutions. a. Membranes were probed with the following antisera (all at 1/200 dilution): 317 pre-immune (I); 317 post-immune (II), 317 post-immune preabsorbed with an extract of *E. coli* expressing the GST protein (III); 317 post-immune preabsorbed with an extract of *E. coli* expressing the GST-bbcl fusion protein (IV). Tracks are: *D. melanogaster* 0-24hr embryo (E, positive control); T47D cell line (T); Hs 578 T cell line (H) and MCF 7 cell line (M). Marker (S) sizes are in KDa. b. Membranes were probed with 317 post-immune antiserum at 1/500 dilution (I) and 1/1000 dilution (II). Tracks are: GST-bbcl fusion protein (F); 0-24hr *D. melanogaster* embryo (E); human placenta (P) and HeLa cell (H).
protein has been used for a number of applications intended to characterise the *D. melanogaster* and human proteins. The existence of a protein likely to represent the *D. melanogaster* bbcl protein has been shown and the expression of this protein throughout the *D. melanogaster* life cycle examined by the technique of Western blotting. The antiserum has been used unsuccessfully for immunohistochemical analysis and immunoprecipitation, which has prevented further characterisation of the *D. melanogaster* protein. The antiserum also appears not to recognise the human bbcl protein, thus preventing characterisation of this protein.
8.1 Discussion

A novel Drosophila melanogaster gene has been identified on the basis of similarity to a previously identified human cDNA sequence (Adams et al., 1992a). The cDNA product of the D. melanogaster gene contains a single long open reading frame encoding a predicted protein displaying over 70% similarity to the predicted product of the human bbcl cDNA. This similarity suggests that the D. melanogaster gene is paralogous to the human bbcl gene and so is referred to as the D. melanogaster bbcl gene. The identification of two introns within the D. melanogaster bbcl open reading frame led to the identification of an intron within sequences likely to represent the functional human bbcl gene. Demonstration that the human intron and the first D. melanogaster intron reside at identical positions within the conserved open reading frame indicates that the two genes probably evolved from a common ancestral gene that existed before divergence of the arthropod and chordate lineages i.e., the two genes are likely to be homologous.

The high degree of sequence conservation seen between the D. melanogaster and human sequences (both cDNA and protein) indicates that the bbcl gene is refractory towards mutation and undergoes a slow rate of evolution. This hypothesis is supported by the identification of hybridising bands in many species of eukaryotes using the human bbcl cDNA sequence as a probe (see chapter 1).

The large number of bands identified in human and several other vertebrates is interesting. Use of an intron-specific sequence to probe a Southern blot of human genomic DNA identified one DNA fragment likely to contain the functional gene from which the human bbcl cDNA is transcribed. The additional bands may correspond to sequences representing closely related functional genes, pseudogenes
or other cross-hybridising genomic sequences. Northern blotting experiments of human RNA from a limited number of sources indicates the presence of one major and two minor transcripts. At present it is not known whether all these mRNA species originate from a single gene. Northern blotting analysis of RNA from a number of rat tissues demonstrates that a transcript capable of hybridising with the human bcl cDNA sequence is widely expressed in this organism. Despite the uniform size of transcripts from different tissues, their sequences have not been determined and it is possible that more than one gene gives rise to the transcripts observed.

Closely related functional genes may be so closely related as to represent isogenes which are expressed in different tissues. Southern blotting analysis of human genomic DNA probed with the human bcl cDNA sequence and washed at high stringency identified a large number of hybridising bands (see chapter 1, figure 1.3.2b), indicating that these sequences are highly similar to the probe sequence. A large number of enzymes, especially those encoded by housekeeping genes, have been shown to possess isozymes expressed in different tissues or subcellular compartments or at different stages of development (Ureta, 1978). Isolation and analysis of bcl cDNAs and/or proteins from a range of human tissues and/or from different stages of development could help to elucidate whether several forms of the protein are expressed from separate genes.

Evidence also exists that some of the hybridising sequences present in the human genome may represent pseudogenes. PCR amplification of sequences from human genomic DNA using primers flanking the only intron so far identified in the human bcl gene produces two major products (see chapter 4, figure 4.3.1a). The smaller of these is identical in size to a control PCR amplification performed on human bcl cDNA and indicates that intronless bcl pseudogenes may be present within the human genome. A single major PCR product larger than the product of PCR on the cDNA control was seen, indicating that a single intron-containing bcl gene may exist in humans. Direct sequencing of this PCR product with primers within the flanking exons revealed only one intron sequence, thus suggesting that two or more introns of very similar length do not occur. A fainter PCR
product of even larger size (approximately 570bp) was detected. This may represent a second intron-containing sequence that amplified less efficiently than the smaller sequences, possibly due to its larger size or because of sequence divergence in the primer binding sites, or an artefactual amplification product. At present no information is available to distinguish between these possibilities.

The possibility exists, however, that the smaller PCR product is generated from a functional intronless bbc* gene. Such a gene could be the product of a retrotransposition event which places a functional cDNA-type sequence in the genome. An example of a retroposed gene is the rat preproinsulin I gene (Bento Soares et al., 1985). This gene is derived from the rat preproinsulin II gene by RNA-mediated duplication-transposition of a partially spliced transcript. The preproinsulin I gene lacks the 499bp intron which interrupts the preproinsulin II gene-segment encoding the C-peptide, but retains the smaller, 119bp, intron in the 5' noncoding region of the gene. The promoter region of the gene was also retained during the transposition event because the retroposed transcript initiated upstream of the promoter. Thus two functional insulin genes exist in the rat genome. Why two genes exist is not known, but they must provide a selective advantage for the rat (and mouse, which also has two insulin genes) in order to be maintained. Whilst other functional retroposed genes appear to exist (see Bento Soares et al., 1985 for references), it seems likely that the majority of intronless genes are functionless pseudogenes (see, for example, Clevland, 1983; Leavitt et al., 1984 and Limbach et al., 1985).

Regardless of the functionality of additional bbc* sequences found in the genomes of human and several other organisms, the presence of multiple bands (as opposed to two or three, which may merely result from cleavage by the restriction enzyme within the gene) on a Southern blot of genomic DNA appears to be limited to mammalian species (e.g., human, mouse, roedeer, lion, bat). Hence, the production of these sequences may represent a recent evolutionary event, possibly occurring at or after the mammalian radiation of 75 million years ago (Dayhoff, 1978). The emergence of several mammalian retropseudogene families has been linked to this evolutionary point
The *D. melanogaster* *bbcl* Gene

(Bento Soares *et al.*, 1985). Why this point in evolution was accompanied by a sudden increase in retroposition events is unknown, but may have represented a means whereby rapid genetic divergence was possible, thus facilitating the mammalian radiation.

Both the *D. melanogaster* and human *bbcl* cDNA sequences (and introns identified to date) contain high proportions of guanine and cytidine residues and a large number of CpG dinucleotides (see chapter 3). Since the majority of vertebrate coding DNA is subject to methylation (Bird, 1987; Bird, 1992) which can result in mutation of CpG to TpG dinucleotides, these sequences have been lost from most vertebrate genes (invertebrates such as *D. melanogaster* tend only to methylate intergenic DNA). The presence of a high density of CpGs in human *bbcl* sequences indicates that this gene is not subject to methylation. Vertebrate genes containing high densities of CpGs (especially in their promoter sequences) usually belong to a group of ubiquitously expressed genes known collectively as the housekeeping genes (Bird, 1987). A gene likely to represent the rat *bbcl* gene has been shown to be expressed in a range of tissues (see chapter 3) and it is tempting to speculate that the human gene might also show such expression. An analysis of the human and *D. melanogaster* *bbcl* gene promoters will be required to determine whether they contain elements associated with housekeeping genes (Dynan and Tijan, 1983; Bird, 1986; Sehgal *et al.*, 1988), but these data point towards a ubiquitous role for the gene. In keeping with this theory is the finding that the *bbcl* gene is expressed in tissue culture cells of human and *D. melanogaster* origin. It has been suggested that cells in tissue culture generally express only those genes absolutely required for cell survival (Antequera *et al.*, 1990; Bird, 1992), thus accounting for the dedifferentiated nature of most tissue culture cells. Continued expression of the *bbcl* gene in established tissue culture cells suggests that the gene is an essential gene and as such should be expressed in all cells of an organism.

The idea of a general role for the *bbcl* protein within all cells appears to be in conflict with the results of *in situ* hybridisation to *D. melanogaster* embryos which indicates a highly temporal- and spatial-specific expression pattern for the gene during
embryogenesis. Since expression of the rat gene was studied in adult tissues, it is possible that a highly controlled expression pattern occurs in embryonic rat tissues. Likewise, it is possible that ubiquitous expression of the gene occurs in adult *D. melanogaster*; expression of the gene in adult *D. melanogaster* has been demonstrated by Northern blotting analysis, although no spatial information is available.

What might the role of the bbcl protein be? The high degree of sequence conservation seen between the *D. melanogaster* and human proteins is characteristic of proteins which perform central cellular functions requiring multiple intramolecular interactions, for example actin (Korn, 1982), the core histones (Wells, 1986) and the nm23/Awd protein (Rosengard et al., 1989). All these proteins are especially highly conserved with sequence conservation between humans and *D. melanogaster* in excess of 90%. Whilst the bbcl protein is not as highly conserved, the greater than 70% conservation seen between *D. melanogaster* and humans is high for proteins from two such evolutionarily diverged organisms and predicts a conserved function for the bbcl protein within the cells of many eukaryotic organisms.

As far as specific functions of the protein are concerned, the predicted protein sequences of the *D. melanogaster* and human bbcl proteins have provided few clues. Both predicted protein sequences are highly basic, with 28% of the residues being arginine, lysine or histidine. This proportion of basic amino acids is similar to that found in the histone proteins (Wells, 1986) and is reflected in the high pl predicted for the bbcl proteins (11.4 and 12.2 for the *D. melanogaster* and human proteins respectively). The bbcl proteins are also similar in size to the histones, but have distinct sequences that show only limited similarity to histone proteins. Figure 8.1.1 shows a comparison of the *D. melanogaster* bbcl and sea urchin histone H1 proteins. Whilst this histone is the most similar to the *D. melanogaster* bbcl protein (as determined by database searches), only 17% of the residues are identical. This analysis, therefore, shows that the bbcl protein is not histone-like in sequence. It remains
Figure 8.1.1

Comparison of the *D. melanogaster* bbcl protein (D) with sea urchin histone H1 protein (H1). Identical amino acids are indicated by vertical lines and conservative changes by colons. Spaces (-) have been inserted at two positions in the histone H1 sequence to obtain the best alignment.
possible, however, that the bbcl protein is a nucleic acid-binding protein.

Apart from the basic nature of both sequences, the D. melanogaster and human bbcl proteins share several other similarities. The lack of hydrophobic leader sequences (Devillers-Thiery et al., 1975; Davis and Tai, 1980) and potential transmembrane regions (Michel et al., 1986) indicate that the proteins are not secreted nor integral membrane proteins. Indeed, Western blotting analysis of conditioned medium from D. melanogaster SL2 tissue culture cells suggests an intracellular location for the protein. Both proteins do contain two conserved SV40-like nuclear localisation signals (Kalderon et al., 1984; Goldfarb et al., 1986), indicating a possible nuclear role for the protein. However, currently known DNA- (Churchill and Travers, 1991; Harrison, 1991) and RNA-binding (Dreyfuss et al., 1988; Calnan et al., 1991) motifs are absent. It is possible that the bbcl protein binds nucleic acid in a novel manner or forms a complex with a nucleic acid-binding protein. Alternatively, the bbcl protein may function in a manner unrelated to nucleic acid binding. Indeed, the intracellular location of the protein has not been determined and, thus, the protein could reside in the cytoplasm and/or the nucleus.

A conserved N-linked glycosylation site is found in both proteins (Lehle and Bause, 1984). Whilst the majority of intracellular proteins are not glycosylated, some are, for example nuclear pore complex proteins (Holt, 1987). However, these are O-linked, not N-linked glycosylations; no N-linked glycosylations have been detected on intracellular proteins. The observed size of the D. melanogaster bbcl protein on Western blots also indicates that major post-translational modification of the protein (such as N-linked glycosylation) does not occur. Given the above information, it seems likely that the bbcl protein is not glycosylated on the consensus site in vivo. The consensus site in both proteins overlaps one of the potential nuclear localisation sites found in the proteins. Hence, it is possible that the sequence has been conserved due to a function unrelated to glycosylation of the protein.
Comparison of the predicted human and *D. melanogaster* bbl proteins with other protein sequences has revealed little similarity. Two proteins that do share regions of reasonably high identity with the bbl proteins are a *D. melanogaster* P-element transposase (Rio et al., 1986) and the human fra-2 protein, a fos-related transcription factor (Nishina et al., 1990). Unfortunately, the regions of identity lie outside of domains in both proteins with known functions.

In an attempt to correlate the position of the *D. melanogaster* bbl gene with previously identified mutations, *in situ* hybridisation to polytene chromosomes from *D. melanogaster* larvae was performed. Whilst this analysis successfully located the chromosomal region harbouring the bbl gene, this region is characterised by a paucity of identified markers. Most of these markers have also not been accurately positioned on the cytogenetic map of chromosome 2, meaning that a large amount of work would possibly have to be performed to determine whether any of the mutations affected the bbl gene. Despite the disappointing nature of this result, the assignment of a chromosomal position for the *D. melanogaster* bbl gene could be of use during future mutational analysis of the gene (see section 8.2). Flies containing one mutated copy of the gene could be crossed with a strain deficient for this region of the second chromosome, thus producing progeny lacking a functional bbl gene in which phenotypic studies could be performed.

Analysis of the *D. melanogaster* bbl gene expression pattern during embryogenesis has shown that highly specific expression patterns occur. Expression of the gene appears to occur at high levels in the spiracles and tracheal system of the developing embryo as well as in external sensory organs. This expression coincides temporally but not spatially with condensation of the CNS (Weischaus and Nüsslein-Volhard, 1986). Expression during earlier stages of embryogenesis (stages 9 and 10) does appear to occur in primordial cells of the CNS as well as in the posterior midgut invagination. However, at present the role(s) of the gene product in the organs expressing it is unknown.
Another gene-product expressed in the developing trachea of *Drosophila* embryos is the *Drosophila* homologue of the vertebrate fibroblast growth factor receptor (DFGF-R). This receptor has been proposed to participate in receiving spatial information required for the correct migration of *Drosophila* tracheal precursor cells during tracheal extension (Glazer and Shilo, 1991). There is a degree of similarity between the processes of tracheal extension in *Drosophila* and capillary outgrowth in vertebrates. Both processes require the proliferation and directed movement of cells from an existing structure, either a tracheal pit in *Drosophila* or a pre-existing capillary in vertebrates, to form a branching mono-layered tubular structure. Fibroblast growth factor is a potent endothelial cell growth factor and angiogenic stimulus in vertebrates (Folkman and Klagsbrun, 1987), proposed to function directly through receptors on the endothelial cell surface. Identification of the ligand(s) for the DFGF-R will be of interest in determining just how similar the processes of neovascularisation and tracheal extension are. If the *Drosophila* bbcl protein is also involved in tracheal development, as indicated by the *in situ* hybridisation studies, this could (albeit very tenuously) indicate a role, either positive or negative, for the human bbcl protein in angiogenesis. If this is the case, then this could indicate a differential requirement for angiogenesis between benign and malignant tumours of the female breast. There is a positive correlation between tumour angiogenesis and metastasis in breast cancer (Weidner *et al.*, 1991), which presumably relates to tumour cell access to the vasculature. However, since expression of the human bbcl gene was compared in benign (i.e. non-metastatic) and malignant (i.e. metastatic) tumours, the differential expression observed may not necessarily correlate with metastasis even if it does correlate with angiogenesis.

Antiserum to the *D. melanogaster* bbcl protein has identified a protein likely to be the product of the bbcl gene in this organism. The protein is expressed during embryogenesis and at later stages of the *D. melanogaster* life cycle to a much reduced extent. The temporal expression pattern seen for the protein as determined by Western blotting analysis reveals a different pattern to the mRNA temporal expression pattern determined by *in situ* hybridisation. The
The high level of mRNA expression detected at 14-18 hr of embryogenesis is not repeated for the protein, possibly because of alterations in the rate of protein degradation or due to translational control of gene expression (see chapter 7). Protein expression at early stages of embryogenesis in the apparent absence of zygotic gene expression may be explained by maternal gene expression, although this has not been determined.

Failure of the antiserum to recognise the bbcl protein in D. melanogaster tissue sections was disappointing. Since the antibody recognises the protein on Western blots, the possibility exists that the antibody is incapable of recognising the native bbcl protein. This could either be due to the native bbcl protein adopting a conformation that is not recognised by the antibody, or by the epitopes being masked. Polyclonal antisera raised against whole proteins contain antibodies reactive to many different epitopes on the protein, some of which will probably be recognised regardless of the protein conformation. This suggests that failure of the antibody to bind to the bbcl protein under native conditions is unlikely to be due to non-recognition of protein epitopes per se, but may result from these epitopes being obscured. A possible way in which epitopes could be hidden is by binding of other cellular molecules (such as proteins, lipids, carbohydrates or nucleic acids) to the bbcl protein.

Immunoprecipitation reactions performed to determine whether the antibody was capable of recognising the native bbcl protein failed to identify the protein in radiolabelled cell extracts. However, the reason for this may have nothing to do with the antiserum, since the control antibody also failed to identify a radiolabelled protein under the conditions used. Since two alternative extraction methods were used to prepare native and denatured protein extracts (see chapter 2), failure of either antiserum to recognise the native protein should not have caused the result seen. Possible reasons for the failure either to immunoprecipitate the desired proteins or to detect immunoprecipitated proteins may be slow synthesis rates of the proteins being precipitated or proteolysis of these proteins during the immunoprecipitation reaction respectively.
Localisation of the human \textit{bbcl} gene to chromosome 16 is of interest given the loss of sequences on this chromosome in several types of solid tumour including breast cancer (Carter \textit{et al.}, 1990; Sato \textit{et al.}, 1990; Tsuda \textit{et al.}, 1990; Devilee \textit{et al.}, 1991). The human \textit{bbcl} gene has recently been mapped to the long arm of chromosome 16 and will shortly be fine-mapped to a smaller region of this chromosome (Dr. A-M. Cleon-Jansen, Rijksuniversiteit te Leiden, Leiden, Holland. Pers. comm.). It will be of interest to see whether the \textit{bbcl} gene maps to region 16q22 which, together with 16p11 and 16p13, have been identified as the sole abnormalities in a number of human tumours (Mitelman \textit{et al.}, 1991).

Chromosome loss is usually associated with the presence of a tumour suppressor gene(s) on that chromosome (see Ponder, 1988; Hollingsworth and Lee, 1991 and Marshall, 1991 and references therein). Recently, however, loss of a gene thought to inhibit tumour progression, the \textit{nm23} gene, has been shown to occur in a number of human malignancies (Cohn \textit{et al.}, 1991; Leone \textit{et al.}, 1991b). Several potential metastasis suppressor genes have been identified, the expression of which correlate with the non-metastatic phenotype, for example \textit{nm23} (Steeg \textit{et al.}, 1988); \textit{WDNM1} (Dear \textit{et al.}, 1988); \textit{WDNM2} (Dear \textit{et al.}, 1989), \textit{TIMP} 1 and 2 (Schultz \textit{et al.}, 1988; Albini \textit{et al.}, 1991) and the fibronectin gene (Schalken \textit{et al.}, 1988).

Removal of fibronectin during \textit{in vitro} culture of murine melanoma cells has previously been shown to result in elevated metastatic potential of these cells (Terranova \textit{et al.}, 1984). Reduced adhesion of the cells to fibronectin may promote binding to and degradation of basement membranes, thus enhancing metastatic potential (Terranova \textit{et al.}, 1984). The TIMP proteins have been shown to inhibit extracellular matrix degradation by tumour cells and to reduce experimental metastatic potential \textit{in vivo} (see chapter 1). Reasons for differential expression of the \textit{WDNM1} and 2 genes in tumours of varying metastatic potential have not so far been elucidated.

The \textit{nm23} gene is of interest because the expression of this gene has been shown to correlate with reduced metastatic potential in several...
experimental systems of rodent metastasis (Steeg et al., 1988; Rosengard et al., 1989). Expression of the nm23 gene also correlates with indicators of low metastatic potential in human breast cancer (Bevilacqua et al., 1989; Hennessy et al., 1991; Hirayama et al., 1991).

The protein product of this gene is a nucleoside diphosphate (NDP) kinase (Biggs et al., 1990). It has been postulated that at least two major cellular processes involving NDP kinases, namely microtubule assembly and disassembly and signal transduction through G-proteins, may have roles in metastasis (Liotta and Steeg, 1990). Disruption of microtubule function may lead to aberrant mitosis and aneuploidy as well as other karyotypic abnormalities associated with, and possibly contributing to, the malignant phenotype. Disruption of signal transduction via G-proteins could also have wide-ranging effects due to the large number of G-protein-transduced signalling pathways present within cells. Indeed, the nm23 protein has been shown to directly activate a G-protein in vitro, implicating this protein in G-protein-transduced signalling (Bominaar et al., 1993). As mentioned in the introduction, the nm23 protein has also recently been shown to be a transcription factor required for the efficient expression of the c-myc proto-oncogene (Postel et al., 1993). This is of considerable importance, given the wide range of cellular functions influenced by the c-myc protein. However, at present the full implications of this finding are not known.

The aim of this project was to analyse the D. melanogaster bbcl gene and gene-product with a view to determining the function of the gene in this organism. Any information gained could then be used to investigate the human gene(s) with the eventual goal being to understand why the gene shows a difference in expression between benign and malignant tumours of the breast. Whilst a function for the gene has not been ascertained, a large amount of information about the D. melanogaster and human genes has been obtained and should allow isolation and full characterisation of the human gene(s) and protein(s), thereby facilitating the elucidation of a role for the gene. Further analysis in D. melanogaster also has the potential to
identify functions for the gene in this organism. Possible future studies aimed at characterising further both the D. melanogaster and human bbcl genes and proteins are described in section 8.2 below.

8.2 Future work

The work described in previous chapters has resulted in characterisation of several aspects of the D. melanogaster and human bbcl genes and the D. melanogaster bbcl protein. Several attributes of the human and D. melanogaster genes have not been determined and much still remains to be elucidated with respect to the protein products of both genes. Work which should be performed to address these matters is described below.

According to the results of primer-extension experiments (chapter 3) the D. melanogaster bbcl mRNA sequence is 9nt longer than sequence which has been confirmed in two independent cDNA clones (figure 3.3.5). To determine the identity of these nucleotides, a D. melanogaster genomic library should be screened using the cDNA sequence as a probe. Not only would this reveal the identity of these nucleotides, it would also permit characterisation of any additional introns within the genomic sequence and an investigation of the promoter structure of the gene. The last of these is potentially the most interesting due to the possibility of identifying binding sites for transcription regulatory proteins. Such sequences may provide clues to the function of the bbcl gene in this organism. Isolation of the human bbcl gene using the intron sequence as a probe would allow a similar analysis to be performed on this gene. Again, identification of binding sites within the promoter for specific factors might help to assign a function to the gene. The possibility exists, as discussed above, that the gene is ubiquitously expressed. If this is the case, then sequences characteristic of housekeeping genes (Dynan and Tijan, 1983; Bird, 1986; Sehgal et al., 1988) should be identified within the promoter of the gene.

Isolation of further unique human bbcl genomic sequences should allow mapping of the gene to a specific region of chromosome 16.
Several methods are available to do this, for example in situ hybridisation or PCR analysis using a panel of somatic cell hybrids containing subfragments of human chromosome 16 and fluorescence in situ hybridisation (F.I.S.H.) to metaphase human chromosomes. As mentioned above, fine mapping of the human *bbcl* gene to a specific region of chromosome 16 will shortly be performed (Dr. A-M. Clenton-Jansen, Pers. comm.).

Genetic analysis of the *D. melanogaster bbcl* gene should be performed in order to complement the molecular analyses described in previous chapters. *D. melanogaster* is an ideal organism in which to study the genetics of the *bbcl* gene because it has been extensively characterised at the genetic level. Recently methods for rapid mutagenesis of cloned genes have also been established (Ballinger and Benzer, 1989; Kaiser and Goodwin, 1990; Hamilton et al., 1991), making targeted disruption of the gene possible. Whilst disruption of the gene could be lethal, valuable information could still be obtained by determining the developmental stage at which lethality occurred and any phenotypic abnormalities associated with this. Given the apparent high-level expression observed during embryonic development, it is possible that homozygous mutants would not survive embryogenesis. This would not be unfavourable, because it would indicate an important role for the gene in embryogenesis. Many vitally important *D. melanogaster* developmental genes (some of which have human homologues) have been identified through their effects on embryogenesis when mutated (see chapter 5 and Akam, 1987).

Production of polyclonal antisera to the *D. melanogaster bbcl* protein has enabled preliminary characterisation of this protein to be performed. Further antisera should be generated against the *D. melanogaster* and human proteins. These antisera could be raised against whole proteins or against subfragments. Use of fragments may result in antisera capable of recognising epitopes which are hidden or have altered structure when the entire protein is expressed. Thus, antisera raised using subfragments of the *bbcl* proteins may provide a wider range of antibody specificities than antisera raised to whole proteins alone.
These antibodies could be used for a range of immunological techniques aimed at characterising further the *D. melanogaster* and human bbcl proteins. Western blotting of proteins from various human tissues would elucidate in which organs the protein is expressed, whilst immunohistochemical analyses could possibly allow the cell types expressing the protein to be determined. Immunohistochemical analysis of *D. melanogaster* adult and embryonic sections has the potential to identify which organs and cell types within this organism possess the protein. This technique would also allow an estimation of protein production in various regions of the developing embryo which could be compared with the mRNA expression levels observed by *in situ* hybridisation analyses.

Immunaffinity purification of the bbcl protein either from embryos or from tissue culture cells would allow analysis of the protein to determine whether post-translational modifications such as glycosylation and phosphorylation are present. Immunaffinity purification also has the ability to isolate macromolecules bound to the bbcl protein *in vivo*. The presence of other proteins bound to the bbcl protein could easily be determined by SDS-PAGE analysis, whilst nucleic acids can be detected by binding of fluorescent molecules such as ethidium bromide. Such analyses are dependent on the bbcl protein being recognised by antisera when bound to other molecules. Should the protein not be recognised, several methods could be used to remove (either partially or totally) the bound molecules such as altering the pH or ionic content of buffers, inclusion of detergents and enzymatic digestion. Whilst these treatments might allow the bbcl protein to be purified, it is possible that any molecules bound *in vivo* would not be isolated.

In summary, through the use of many varied techniques of molecular and cell biology as well as genetics, the potential exists to gain much information regarding many aspects of the *D. melanogaster* and human bbcl genes and their gene products. This information may not only identify a role for the gene, but may provide an explanation of why the gene is differentially expressed in benign and malignant tumours of the human female breast.
Literature Cited


The *D. melanogaster* *bcb* Gene


The *D. melanogaster* bcd Gene

---


### The D. melanogaster bbcl Gene

<table>
<thead>
<tr>
<th>References</th>
</tr>
</thead>
</table>
The D. melanogaster bcd Gene

References


The D. melanogaster bcd Gene


Folkman, J. (1990) What is the evidence that tumours are angiogenesis dependent? J.N.C.I. 82, 4-6.


The D. melanogaster BBC Gene

|------------|-----------------------------------------------------------------------------------------------------------------------------------|
The D. melanogaster bbcl Gene


The *D. melanogaster* bbcl Gene

<table>
<thead>
<tr>
<th>References</th>
</tr>
</thead>
</table>
The *D. melanogaster* bbcl Gene


The *D. melanogaster* bbcl Gene


The D. melanogaster bcd Gene

<table>
<thead>
<tr>
<th>References</th>
</tr>
</thead>
</table>
The D. melanogaster bbcl Gene References


<table>
<thead>
<tr>
<th>The D. melanogaster bbl Gene</th>
<th>References</th>
</tr>
</thead>
</table>
The D. melanogaster bbcl Gene

References


<table>
<thead>
<tr>
<th>The <em>D. melanogaster bbcd</em> Gene</th>
<th>References</th>
</tr>
</thead>
</table>
### The D. melanogaster bcd Gene

<table>
<thead>
<tr>
<th>References</th>
</tr>
</thead>
</table>
The *D. melanogaster* bbcl Gene

<table>
<thead>
<tr>
<th>References</th>
</tr>
</thead>
</table>


The *D. melanogaster* *bbcl* Gene

<table>
<thead>
<tr>
<th>References</th>
</tr>
</thead>
</table>
The D. melanogaster bcd Gene


