A Screen for *Drosophila* Genes Revelant to the Nervous System

A thesis for submission for the degree of Doctor of Philosophy at the University of Glasgow

Simon R Tomlinson

Institute of Genetics
University of Glasgow

September 1994.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

To mum and dad

The research reported in this thesis
is my own work except where otherwise
stated and has not been submitted
for any other degree.

Simon R Tomlinson

List of Figures

All references to figure numbers in the text are prefixed with the term "figure" to distinguish from cross-referencing to chapter sections. In this list the term "figure" has been omitted.

Abbreviations

Standard abbreviations for elements of the periodic table have been used throughout, together with the amino acid abbreviations described by Dayhoff (1968). Additionally, the following abbreviations have been used:-

AA	Amino acid
ATP	Adenosine triphosphate
ATPase	ATP hydrolysing enzyme
bp	Base pairs
dCTP	Deoxyadenosine triphosphate
cDNA	Copy DNA
DEPC	Diethylpyrocarbonate
DIG	Digoxigenin
DNA	Deoxyribonucleic acid
DNase	DNA hydrolysing enzyme
DTT	1,4-Dithiothreitol
EDTA	(Ethylenedinitrilo) tetra-acetic acid
EGTA	(Ethylene-bis(oxyethylene-nitrilo))tetraacetic acid
ER	Endoplasmic reticulum
ExoIII	Exonuclease III
GCN	Guanidine thiocyanate
HAP	Hydroxylapatite
IPTG	Isopropyl-β-D-thio-galactopyranoside
kb	Kilobases
kDa	KiloDaltons
Klenow	Klenow fragment of *E.coli* polymerase I
mg	Milligrams
ml	Millilitres
min	Minutes
mRNA	Messenger RNA
ODx	Optical density (x=wavelength in nanometres)
OrR	Oregan R wildtype strain of *Drosophila*
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PERT	Phenol enhanced reassociation technique
PFU	Plaque forming unit
pH	$\log_{10}$ of the recipical of hydrogen ion concentration

PolyA+	Poly adenosine tailed RNA molecule
PPi	Pyrophosphate
RNA	Ribonucleic acid
RNase	RNA hydrolysing enzyme
RoT	A multiple of the initial RNA concentration and time
Tris	2-Amino-2-(hydroxymethyl)-1,3-propanediol acetate
UV	Ultraviolet
Xgal	5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside
X-phosphate	5-Bromo-4-chloro-2-inodolyl-phosphate
µl	Microlitres
µg	Micrograms
ng	Nanograms
3'	Three prime
cAMP	Cylic adenosine monophoshate
GSN	Guanidine thiocyanate
PKA	Protein kinase A
PKC	Protein kinase C
PVP	Poly vinyl pyrolidine
BSA	Bovine serum albumin
tRNA	Transfer RNA
RP49	Ribosomal protein 49 (*Drosophila*)
DEPC	Di-ethyl pyrocarbonate
TCA	Tri-chloro acetic acid
5'	Five prime

Acknowledgements

I wish to take this opportunity to thank all those people in the Institute of Genetics and the Robertson Institute of Biotechnology, Glasgow, who provided help and support during this thesis work. Special thanks are due to the people on the fifth floor, the media ladies, to Gareth Griffith for many helpful discussions and Gus Souza for introducing me to computer based sequence analysis.

Great thanks are also due to Avind Hundall, Leslie Glass, Alan Griffin, and Margeret Adams to whom I still owe many meals and many more beers. Lastly, I'd like to thank Gillian Scott for helping to keep me sane by driving me mad.

Abstract

Understanding the molecular and genetic control of many brain functions remains one of the most challenging problems of modern biology. Although a number of brain structural and functional mutants have been isolated, their number and diversity is small compared to the complexity of the *Drosophila* nervous system. Using various differential screening methods in conjunction with a method of reverse genetics developed in the laboratory, the aim of the project was to gain molecular and eventually genetic access to this missing diversity.

Following an assessment of differential and subtraction screening methods a differential screen was performed to isolate a set of about twenty cDNA clones representing genes expressed strongly in the head and weakly in the body. These clones were sifted further using partial sequence and approximate pattern of expression. Several clones were putatively identified as being derived from genes with homologues in other organisms including one previously known head specific gene, *nina*E. Another clone from the screen, pST123, was shown to be a cloned fragment of a message predicted to encode an unknown homologue of synaptobrevin, a family of molecules known to be involved in synaptic transmission. A third clone class, represented by the pST51 cDNA, was derived from a Na/K ATPase β-subunit gene family member, a family of molecules involved in establishing membrane potentials in the nervous system, as well as in other processes such as cell-cell recognition. These two cDNAs were selected for further study, together with two other cDNA clones, pST162 and pST170, which defined genes with interesting patterns of transcription.

A more detailed molecular characterisation of the four cDNA clones was performed to confirm the predicted gene expression pattern and to further investigate the molecular biology. For example, the β-subunit cDNA insert from pST51, was sequenced and a full length predicted amino acid sequence identified. Phylogenetic analysis revealed that the encoded polypeptide was most closely related to the only other arthropod Na/K β-subunit known, that of the brine shrimp. Together they define a new class of β-subunit, clearly distinct from the three mammalian types. Expression analysis included head/body northerns and *in situ* hybridisation to head and head/body frozen sections using cDNA derived probes. This showed an expression pattern consistant with a gene expressed exclusively in the nervous system. A set of related cDNA clones identified using the pST51 cDNA insert reveal the likely presence of multiple transcripts generated with a

variety of 5' and 3' untranslated regions. The analysis of the related cDNAs also identified the likely presence of a message encoding a second form of the β-subunit. Genomic Southerns and *in situ* localisation to polytene chromosomes indicate a single copy gene, although lower stringency hybridisation experiments identify the presence of several related loci in the *Drosophila* genome.

A similar analysis has been performed on the other selected cDNAs. Of these, the study of the pST123 cDNA derived from a novel *Drosophila* synaptobrevin gene has proved to be the most informative. Interestingly, synaptobrevins have recently been identified as the target for tetanus toxin cleavage. From the predicted amino acid sequence it appears that the gene product should be cleaved by this toxin. A *Drosophila* synaptobrevin locus had previously been identified, but it appears that this locus produces proteins unlikely to cleave with the toxin, and has recently been reported as having a predominantly non-neuronal expression profile. *In situ* hybridisation experiments presented in this thesis suggest a neuronal role for the subunit isolated in this screen.

Chapter1 Introduction

Please note
This section briefly introduces the experimental approaches that have been employed to study the *Drosophila* nervous system and introduces the methodologies discussed later in this thesis. Extensive introductions to the biology of specific gene families are provided at the beginning of each individual results chapter.

Understanding the functioning of the nervous system is one of the major challenges remaining to biology. In the past century a set of powerful tools have been developed in order that one might begin to dissect the nervous system of higher animals. *Drosophila* as an experimental system is peculiarly well suited to two of the most potent technologies; genetics and molecular biology.

The *Drosophila* nervous system is compacted into an organism 6mm in length, with a mass of 1000th of a gram. This nervous system comprises perhaps 10% of this by mass. This nervous system is constructed from about 10^5 cells, which can be seen to be ordered in complex interconnecting networks using light microscopy. The complex nature of the cellular nervous system is mirrored at the level of the whole organism phenotype, which has been demonstrated most clearly by a set of remarkable behavioural studies into *Drosophila* courtship- a set of responses that employ an immense set of interacting inputs known to consist of specific behavioural elements, some of which are performed sequentially. Often each element can be broken down further to consist of the exhibition and processing of visual, chemical, auditory and mechanical components (e.g. Burnet and Connolly 1974).

Not surprisingly, this cellular and behavioural complexity is mirrored by the transcript and protein diversity. For example, it has been estimated by solution hybridisation that 11,000 transcripts are present in the *Drosophila* head (Levy and McCarthy 1975, Izquierdo and Bishop 1979, Levy and Manning 1981) and several thousand of these are reasonably expected to be expressed solely in the nervous system.

The process of unravelling the study system employs a reductionist methodology, essentially the system is broken down into it's component parts and these parts are studied individually in order to gain understanding of the whole system. Central to this methodology is the concept of the gene which is considered as a reductionist base unit. In this thesis the term "gene" is used in a molecular genetic sense which is "a segment of DNA involved in producing a polypeptide chain..."(Lewin 1987). Such a gene can be said to be defined when it can be distinguished from other such pieces of DNA in the genome. A gene may be defined genetically if the functional expression of the polypeptide chain is heritably altered by a single change in the genome. A gene may be defined by molecular cloning if it is singularly isolated as a cloned fragment of the total coding capacity.

Only relatively few nervous system genes have been defined genetically. Often this is because elaborate screens have to be designed in order to sift out the rare individual carrying the mutation of interest. Nervous system genes that have been genetically defined range from those involved in physiological processes e.g. *sh* (Catsch 1944, Kaplan and Trout 1969) *ninaA* (Matsumoto *et al.* 1986), nervous system development e.g. *pas* (Thomas and Wyman 1984) structural brain mutants e.g. *sol* (Fischbach and Heisenberg 1981) those involved in aspects of behaviour such as mating e.g. *stuck* (Beckman 1970), circadian rhythms e.g. *per* (Konopka and Benzer 1971) and learning and memory e.g. *dunce* (Dudai *et al.* 1976). Once the gene has been genetically defined, it may also be molecularly cloned.

A powerful approach to identify new genes has been the use of the technique of enhancer trapping, pioneered in *Drosophila* by O'Kane and Gehring in 1987. This strategy uses a transposon called a P-element engineered to contain a reporter gene, such as *lacZ* (encoding β-galactosidase) under the control of a weak promoter. The P-element is mobilised in the genome, and when it comes within the proximity of a local enhancer element this enhancer can drive expression of the reporter gene. This expression can be detected using a chromatic substrate of β-galactosidase.

A few genes are defined by direct protein purification of the gene product from *Drosophila.* Notable successes including a number of genes involved in acetylcholine metabolism such as *cha* (Crawford *et al.* 1982, Slemmon *et al.* 1982, Salvaterra *et al.* 1985 Itoh *et al.* 1986)

A commonly employed method to define a gene is by gene cloning by screening with a mammalian probe. This requires the use of a variety of methods of homology screening. It makes the assumption that structural features of the functionally defined locus e.g. DNA or amino acid sequence are sufficiently conserved between *Drosophila* and mammals. It also makes the assumption that should a motif be conserved at the polypeptide level it will perform the same function in the target organism.

Even with the limitations of homology screening a wide variety of genes have been successfully cloned by homology to mammalian genes. These include transmitter related genes such as several acetylcholine receptors, *ard, als* and *sad* (Hermanns-bergmeyer *et al. 1986* Hermanns-bergmeyer *et al.* 1989, Schloss *et al.* 1988, Sawruk *et al.* 1988, Sawruk *et al.* 1990, Wadsworth 1988 and Bossy *et al.* 1988), second messenger genes such cAMP dependent protein kinase (PK-A Foster *et al.* 1988), genes involved in phototransduction such as *nina*E the major opsin in the *Drosophila* eye, and genes involved in a variety of physiological functions including a *Drosophila* synaptobrevin (Sudhoff *et al.* 1989) and Na+/K+ ATPase α-subunit (Lebovitz *et al.* 1989).

If one adds together the full set of defined *Drosophila* genes specifically involved in the functioning of the nervous system then these genes might number one hundred. Even though each gene might contribute several distinct mRNA and even protein products, it is clear that most of the complexity reported by studies of mRNA and protein diversity remains to be studied.

Since studies of mRNA complexity and indeed protein complexity report such a wealth of structural diversity, is it possible to use these systems to clone and so study these missing genes? One possible approach was pioneered by S. Benzer whose group used banks of antibodies raised against *Drosophila* brain proteins to screen a *Drosophila* expression library (Fujita *et al.* 1982). Two cDNA screens have also been performed to isolate genes specific to the *Drosophila* nervous system. It is this second approach that I have adopted to isolate a distinct set of such genes.

The first *Drosophila* nervous system-specific cDNA screen was that of Levy *et al.* 1982. This employed tissue specific cDNA probes generated from head and body polysomal RNA and used this to probe duplicate filters lifted from a

genomic library. They isolated 20 non-cross hybridising 'lambda genomic clones. These clones were initially defined by their developmental expression patterns and chromosomal map location. The loci have subsequently been analysed in a number of laboratories:-

One of the genomic clones from the Levy *et al.* 1982 was shown to be from the *nina*E gene. A number of *Drosophila* mutants have been isolated that are defective in either reception of light or the processing of the light signal in the photoreceptor cells. One class of these mutants are the *nina* (neither inactivation nor afterpotential series (A to E). One of these mutants *nina*E has been shown to encode the major opsin of the eye (O'Tousa *et al.* '85). This gene in *Drosophila* was cloned by homology to a bovine opsin by this group, and was found to map to the same chromosomal locality as *nina*E and a genomic fragment was able to rescue the *nina*E phenotype. The locus was also independently cloned by Zucker *et al.* 1985 by deficiency mapping. Most importantly, the expression of *ninaE* is very head elevated over body and is restricted in development to the late pupal stage where the eye takes up it's mature, adult characteristics.

A second genomic clone from the Levy *et al.* 1982 screen maps to the chromosome close to another genetically defined member of the *nina* group, *ninaC*. The Levy *et al.* 1982 genomic insert was used in an attempt to rescue the mutant phenotype by P-element transformation. Transformant flies were assayed by their response to blue light, which is characteristically abnormal in the mutant (Montell and Rubin 1988). The genomic fragment was able to rescue this phenotype and so encodes the *nina*C gene product. This gene was also shown independently to be head elevated by Montell and Rubin (1988).

Another of the Levy *et al.* 1982 clones was a likely candidate to encode the *trp* gene product on the basis of its pattern of expression and chromosomal map position. *Trp*, transient receptor potential, is one of several mutants which have been identified that affect intermediate stages in phototransduction. P-element rescue experiments were performed and were able to rescue the mutant phenotype (Montell *et al.* 1985).

Ray and Ganguly (1992) have recently identified another of the Levy *et al.* 1982 positive clones as encoding a γG protein subunit. G proteins are heteromeric proteins composed of α, β and γ subunits that are involved in a

wide variety of signal transduction processes. They are found in abundance in the nervous system. The expression patterns reported by Levy and Manning 1982 were used to identify a likely candidate clone. Sequencing and database searching was used to confirm the identity of the encoded subunit.

A second *Drosophila* head specific cDNA screen has been published by Palazzolo *et al.* 1989 who reported the isolation of 436 non cross-hybridising cDNA clones that were expressed in the adult head and not the early embryo. The early embryo was chosen as this tissue is devoid of nervous system, but contains many *Drosophila* house keeping genes. They employed an elaborate subtraction cloning/screening strategy to identify a vast set of rare genes. Unfortunately, considerable problems were encountered in this screen because the subtraction procedure used caused the cDNA fragments to break down to an average size of only 500bp. They report that this subtraction process enabled then to double the abundance of a given head specific gene in their library. But by their efforts they estimate that they have cDNA clones that account for 30% of the genes expressed specifically in the adult head (Palazzolo *et al.* 1989).

Interestingly, one of the Palazzolo *et al.* 1989 clones has been identified as encoding a new class of G protein called $G_{q\alpha}$ (Lee *et al.* 1990). Several screens had been performed previously by low stringency hybridisation using bovine probes (reviewed Buchner 1991) but these screens had failed to identify the expected homologue. Lee *et al.* 1990 opted to sift the Palozzolo clones by using cDNA probes derived from *eya* mutants. This allowed the definition of a subset of cDNA clones derived from mRNAs specifically located in the *Drosophila* eye. Further analysis identified the Gqα encoding gene. The protein showed only 50% homology with most known Giα subunits.

Chapter 2 Methods

Molecular biology techniques were performed as described by Sambrook *et al.* 1989 unless otherwise described.

2.1 General Information

It is not practical to describe the suppliers of every consumable, chemical or biochemical used in this thesis. Generally, chemicals were obtained from BDH and enzymes from Gibco-BRL or Promega. Where the supplier of a particular reagent was found to be critical to the success of the experiment these are referenced in the text.

2.1.1 Specified Suppliers of Consumables

Amersham
Amersham International plc
Amersham Place
Little Chalfont
Bucks HP7 9NA

BDH
Merck Ltd
Merck House
Poole Dorset
BH15 1TD

Beckman
Beckman Instruments Ltd
Glenrothes
Fife
Scotland

Boeringer Mannheim
Boeringer Mannheim plc
Bell Lane
Lewes
East Sussex BN7 1LG

Gibco-BRL
Gibco Ltd
P.O. Box 35
Trident House
Renfrew Road
Paisley Scotland

IBI
IBI Ltd
36 Clifton Road
Cambridge England
CBI 42R

Pharmacia
Pharmacia Biotech. Ltd
Day Avenue
PO Box 100
Knowlhill
Milton Keynes MK5 8PB

Promega
Promega Ltd
Southampton UK

Sigma
Sigma Chemical Co Ltd
Fancy Road
Poole, Dorset
BH17 7TG

Stratagene
Stratagene Ltd
Cambridge Innovation Centre
Cambridge Science Park
Milton Road
Cambridge CB4 4GF

USB
As Amersham

Whatman
Whatman International
St. Leonard's Road
20/20 Maidstone, Kent
ME16 OLS

Vecter Laboratories
16 Wulfie Square
Bretton
Peterborough PE3 8RF

2.1.2 Bacterial Strains Used

NM621 (Whittaker *et al.* 1988) is *hsdR, mcrA, mcrB, mcrD, recD,* 1009 *thy+* and was used to propagate bacteriophage lambda. XL1-B (Bullock *et al.* 1987) is *recA1, endA1, gyrA96, thi, hsdR17, supE44, relA1, lac,* {F' *proAB, lacIq* ZM15, Tn10(tetR)} and was used to support the growth of all plasmid vectors.

2.2 Basic Laboratory Methods

2.2.1 Ligation of Double Stranded DNA

Ligations were performed as described in the Promega Protocols and Applications Guide (Promega). Ligation reactions for plasmid DNA were established in a final volume of 10µl containing about 200ngs of DNA with an insert ration of 2:1 insert:vector and 1-2units T4 ligase. Incubation was overnight at room temperature for ligation of 3' and 5' overhangs and at 14°C

overnight for blunt end ligations. Blunt end ligations were also carried out as described in section 2.12.5.

2.2.2 PCR Amplification of DNA

PCR was performed as described by Promega Protocols and Applications Guide. All reagents used were supplied by Promega, except dNTPs which were from Pharmacia.

2.2.3 Staining and Detection of DNA in Agarose Gels using Ethidium Bromide.

Staining was carried out essentially as described in Sambrook *et al.* 1989 using an excess volume (~100-300mls) of freshly prepared 0.5μg/ml ethidium bromide solution. The gel was then transferred to a bath of distilled water and gently rocked for ten minutes. The gel was examined and photographed under medium wavelength UV irradiation.

2.2.4 Plasmid Preparation Using Alkaline Lysis

2.2.4.1 Small Scale "Minipreps"

Small scale plasmid preparations were made using a slight modification of the method of Sambrook *et al.* 1989. Plasmids were not phenol/chloroform extracted but instead filtered after addition of solutionIII using lint-free tissue in a 1ml syringe. pBluescript (Stratagene) routinely gives 3-5μg of plasmid per ml of culture.

2.2.4.2 Large Scale Preparations

These were performed using a scaled down version of the protocol described in Sambrook *et al.* 1989 and in section 2.2.4.1. Plasmid extracted from a 100ml culture was further purified using a caesium chloride/ethidium bromide equilibrium gradient at 45k in an SW50.a (Beckman) rotor at 15°C. After centrifugation, the supercoiled plasmid was removed from the column as described in Sambrook *et al.* 1989 except that the plasmid was recovered from the ethidium bromide/caesium chloride buffer by extraction using caesium chloride saturated isopropanol. The plasmid was recovered by ethanol precipitation (2.2.10) after the caesium chloride concentration had been reduced five fold by addition of distilled water.

2.2.5 Restriction Digestion of Plasmid DNA

Restriction of plasmid DNA was performed as described in the Promega Protocols and Applications Guide. Minipreped DNA was routinely restricted with up a 10 fold excess of enzyme (e.g. 2μg of plasmid restricted with 5-10 units enzyme for two hours).

2.2.6 Purification of Plasmid DNA from Agarose Gels

Plasmid was purified by electroelution from agarose gel slices into a dialysis bag. This method is described in Sambrook *et al.* 1989. Yield was typically greater than 50%.

2.2.7 Dialysis Of DNA

This method was performed as described in Sambrook *et al.* 1989. Dialysis was typically in a 1000 fold excess volume of T.E pH 8.0 at 4°C with four changes of buffer over 24-48 hours with gentle stirring.

2.2.8 Spectroscopic Determination of DNA and RNA Concentrations

This was as described in the Beckman Du-50 users manual. Samples must be diluted to the accurate range of the machine which was between an OD reading of 0.05 and 0.5. Concentration was calculated given that OD260 of 1.0 corresponds to a concentration of 40μg/ml for RNA and 50μg/ml for double stranded DNA. Extinction coefficients were used to calculate the relationship between OD_{260} and concentration for oligonucleotides as described by Sambrook *et al.* 1989.

2.2.9 DNA Spotting Method to Estimate Concentration of Nucleic Acid

This method was performed as described in Sambrook *et al.* 1989. Samples were serially diluted and 1μl of each dilution placed onto a 1% agarose plate containing 0.5μg/ml ethidium bromide. A serially diluted control of known concentration was similarly plated. The spots were examined under UV irradiation and concentration assessed using intensity of ethidium staining.

2.2.10 Precipitation of Nucleic Acids Using Ethanol and Isopropanol

Samples were precipitated as described Sambrook *et al.* 1989. Typically samples were treated with a tenth volume of 3M NaOAc pH 7.0 and two volumes of ethanol mixed and then chilled on ice. The DNA was recovered by centrifugation. For RNA precipitations 3M NaOAc pH4.5 was used instead.

The protocol was occasionally modified by the substitution of another cation for sodium, for example lithium, when this was required by a specific protocol. 0.67 volumes of isopropanol can be used to substitute for ethanol if the volumes are more convenient. Precipitations with isopropanol were not chilled.

2.2.11 Concentration of Nucleic Acids By Vacuum Drying

This method was commonly used to concentrate radioactive nucleotide solutions and to concentrate oligonucleotides. The samples were simply spun under vacuum in a Gyro-Vap desiccator (Howe) with gentle heating to 40°C. Safe handling methods for radioactive materials was employed as required.

2.2.12 Non-Denaturing Agarose Gels

This method used was exactly as described in Sambrook *et al.* 1989. The TBE buffer system (90mM Tris, 90mM boric acid, 2mM EDTA) was used in preference.

2.2.13 Sodium Hydroxide Denaturing Agarose Gels

This method was performed exactly as described in Berger and Kimmel 1987. The gels cannot be stained easily with ethidium bromide.

2.2.14 Measurement of pH

The pH of small samples was taken using papers (BDH), while that of larger samples was taken using an electronic pH meter as described in the user manual. pH of Tris-HCl containing buffers was measured using a specific electrode on the same device.

2.2.15 Using Filter Binding To Assay Radioactive Incorporation

A known volume of each sample to be assayed were spotted onto glass fibre filters (Whatman GF/C paper, Whatman), air dried and then washed in an excess volume of 5% TCA and 20mM NaPPi. The amount of incorporated activity is given by the radioactivity bound to the filter (the high molecular weight material) to that total amount spotted onto the filter. Activity was counted by Cherenkov counting or by scintillation counting in Ecoscint (Gibco-BRL).

2.2.16 Safe Handling of Radioactive Material

Methods for the safe handling of radioactivity were followed as described by Zoon 1987 and in the departmental and university safety manuals.

2.2.17 Light Microscopy

Microscopy of chromosomes and *in situ* hybridisations was performed using a Ziess inverted microscope as described in the operating instructions for that microscope. Photography was using Kodak Ektra ASA-25 print film.

2.2.18 G50 Chromatography

Sephadex G50 is a resin that retards the progress of small molecules relative to macromolecules. G50 columns were routinely poured in 1ml plastic syringe "bodies" plugged with a small amount of siliconised glass wool. The G50 slurry was added to the column and allowed to pack until the column height was about 0.8cm. The column was then washed with sterile water, and then the sample added. Sample volume was made up to 200µl and a tiny amount of orangeG dye added before addition to the column. The sample was allowed to run into the column and then 100µl aliquots of water added. Aliquots of the same volume were collected from the column.

2.2.19 Autoradiography

Autoradiography was performed as described by Amersham and Sambrook *et al.* 1989. The filter was first sandwiched in Saran wrap (Sigma) and then placed between intensifying screens. A piece of Fuji X100 (IBI) film was placed onto the film under safe-light illumination and the sample sealed in a light-tight cassette. Autoradiography was carried out at -70°C. To perform quantitative autoradiography the film was preflashed by flashing with a gun to give a background exposure level at OD650 of 0.05. The film was developed using an XOMAT automated film developing system.

2.2.20 'Lambda Maxipreparations

Phage were grown as described in Sambrook *et al.* 1989 except that NM621 host cells were used. Removal of contaminating genomic DNA was performed by banding 'phage using a CsCl gradient. This was prepared by adding 7.1g of CsCl to the phage lysate and loading into a Ti70.1 rotor (Beckman). Ultracentrifugation was at 49k, 15°C for 12 hours. The banded 'phage were removed and dialysed (methods section 2.2.7) against T.E (10mM

Tris pH 8, 1mM EDTA). The DNA was then extracted from the 'phage heads as described by Sambrook *et al.* 1989.

2.3 Microbiological Techniques

2.3.1 Maintenance of Bacterial Stocks

Bacterial stocks were maintained in LB (0.1% NaCl, 0.1% bacto-typtone, 0.05% bacto-yeast extract) made up to 20% glycerol. Storage of these stocks was at -20°C or -70°C. Short term stocks were made by plating out a sample of this onto LB plates, growing overnight at 37°C before storage at 4°C. The 4°C stocks remained contamination free for about two weeks.

2.3.2 Plating of Bacteriophage Lambda

Bacteriophage lambda was always grown upon NM621. The plating procedure was performed as described in Sambrook *et al.* 1989 except that 'phage were attached to the host cell at room temperature for 20 minutes and LB used as the culture media. No additional maltose was added to the media.

2.3.3 Plating of *E.coli*

This was performed exactly as described in Sambrook *et al.* 1989.

2.3.4 Transformation of *E.coli* with Plasmid DNA

Competent cells were prepared using the $CaCl_2$ method as described in Sambrook *et al.* 1989. Competent cells were stored overnight at 4°C prior to use. Typical transformations employed 50ng plasmid DNA in 200µl competent cells. The plasmid drug resistance marker from such a transformation was expressed in a total volume of 1ml before 300µl was plated onto each plate.

2.3.5 Selection of Recombinant Plasmids using XGAL/IPTG

This protocol was performed as described by Stratagene. This was used to select for recombinant pBluescript plasmids; presence of the insert within the polylinker blocks the production of a full length *lacZ* gene product. This product from the *lac* operon of *E.coli* acts upon XGAL to produce a blue colouration. In order to score the blue colour it was found necessary to allow the plates to sit at 4°C for 4 hours after overnight growth of the colonies.

2.3.6 Excision of pBluescriptII SK$^-$ Plasmid from 'Lambda ZAP II

This allows the mass transferral of inserts between a 'phage and plasmid vector. The protocol was performed essentially as described Short *et al.* 1988 using R408 helper phage (Stratagene). The only modification was that plates were grown at 42^{o}C to ensure that the colonies were no longer able to produce single stranded 'phage but where still able to replicate the required plasmid.

2.4 Laboratory Culture of *Drosophila*

Basic techniques fro the laboratory culture of *Drosophila* were as described by Ashburner 1989 and Roberts 1986. All cultures unless otherwise stated were at 25^{o}C.

2.4.1 Large Scale Culture

Large scale culture was used to prepare material to make head and body RNA. As such, the flies must all be of uniform size and in perfect condition at the time of harvesting of the culture.

The culture was started using a single sibling cross in a vial. The offspring of this cross were passed into larger bottles (half pint milk bottles) and grown on standard media (Ashburner 1989). This culture was passed through several generations into more and more milk bottles until the flies were sufficient in number to be grown in a fly cage. Fly cages are simply perspex boxes with gauze on one side to allow gas exchange and a round hole in the lower face just smaller than a petri dish. A petri dish containing egg laying media (Ashburner 1989), smeared in live yeast was placed plugging the round hole an cage. The dish was replaced every 24hours. Eggs were harvested off the plate by washing in water and sieving.

Eggs were placed into a perspex box containing a 2cm layer of larval culture media (Ashburner 1989). The lid of this box was tightly sealed, but with a 4cm^2 gauze panel cut in it to allow gas exchange. The larvae were grown until they began to reach late L3 ("crawler stage") and then were transferred by sieving to bottles containing standard egg lying medium. They were then allowed to eclose.

2.4.2 Small Scale Culture and Maintenance of Stocks

Stocks were usually grown in "shell vials" in the standard manner as described in Roberts 1986. Phenotypes and general condition of the flies were checked at every or every other generation.

2.4.3 Preparation of Developmentally Staged Material

The method described in section 2.4.1 was modified to generate staged material. Staged embryos can were prepared from the egg plates by taking a timed laying of eggs and then ageing the plate at 25°C. For example, a plate from a six hour laying contains embryos from 0-6 hours, age this plate for 6 hours at 25°C and they are 6-12 hours old. Approximate staging was confirmed by visual inspection under light microscopy.

Staged larvae were prepared by ageing the egg plate further for L1's or by growing on rich media for L3's. The exact proportion of a given stage in any timed collection critically depends upon local conditions such as time of day that the eggs were collected.

Pupal stages were more difficult to collect. Timing of L3 cultures does not work well because crowding of larvae has a profound effect upon the pupation time. The method used employs the fact that pupae produce an air bubble six hours into pupation (modified from Ashburner 1989). This was used as a method of timing by washing a culture removing all those pupae that float (these are greater than six hours old). The culture was then aged and floated again. Those pupae that now float have passed through the six hour mark since the last washing of the culture and so can be staged. Pupal development calibrated using this method were then aged on damp 3MM (Whatman) paper to produce staged material. Stages were confirmed using morphological markers.

2.4.4 Isolation of Large Quantities of *Drosophila* Head and Body Tissue

Large scale fly cultures were grown as described in 2.4.1. It is very important to ensure that all the flies very healthy and of maximum body size. The head and body isolation procedure is best performed in a fume hood.

Flies were first frozen in liquid nitrogen and poured into a pre-chilled falcon tube in the lid of which had been drilled several 2-3mm holes (these are important!). Liquid nitrogen was allowed to evaporate until the flies were no

longer immersed in the liquid and then the tubes were vigorously vortexed for 30seconds.

The vortexed mixture was then placed into a pre-chilled soil sieve with a 700μm mesh size. The mixture was sieved through the mesh using a nylon brush. Fly bodies were retained on the surface of this sieve but head and leg/wing debris pass through. The sieving process was repeated with a 580um diameter sieve and in this case the heads should be retained. Thus the whole fly mixture was separated into three fractions, head, body and debris. Purity of each fraction was assessed using light microscopy and the still frozen fractions used immediately for RNA extraction (section 2.7).

2.5 Preparation of High Quality Genomic DNA from *Drosophila*

The two protocols described here were both used to prepare genomic DNA. The phenol/chloroform method was used as a quick method that does not require ultracentrifugation. The CsCl method was used as the method of preference.

2.5.1 Using CsCl

1g of material produces 1mg of genomic DNA. The flies or eggs were ground up in liquid nitrogen in a mortar and pestle until a fine powder. This was allowed to warm up and almost defrost before 9mls of ice chilled homogenisation buffer (10mM Tris pH7.5, 60mM NaCl, 10mM EDTA, 0.15mM spermine, 0.15mM spermidine) was added.

Homogenisation was performed in a Wheaton glass homogeniser, filtered through a sterile gauze to remove debris and the homogenate placed on ice. The homogenate was spun at 4000 xg for ten minutes at 4°C. The pellet, which contains the cell nuclei and hence the DNA was treated with a half volume of nuclear lysis buffer (10mM Tris pH7.5, 60mM NaCl, 10mM EDTA) and made to a final concentration of 0.1% SDS. This was incubated overnight at 37°C together with 100μg/ml Proteinase K.

After incubation, 1.25g of CsCl per ml was added and the samples loaded into sealable polyallomer tubes (Beckman) which were then sealed. The samples were centrifuged at 45K in a Ti70 rotor (Beckman) for 24 hours.

The gradient was then dripped into Eppendorf tubes and the fractions containing the DNA identified by their viscosity. Selected samples were then dialysed (section 2.2.7) before use.

2.5.2 Using Phenol/Chloroform

This method is a modification of the 2.5.1 in which the samples are taken after overnight incubation in extraction buffer and extracted in an equal volume of 50% phenol and 50% chloroform. The aqueous phase was then retreated four times with phenol/chloroform. The DNA sample was then dialysed directly as described in section 2.2.7.

2.6 Southern Blotting

This method was modified from Sambrook *et al.* 1989 but with modification described by Amersham for the use of Hybond-N hybridisation membrane.

2.7 Isolation of Total RNA from *Drosophila*

This protocol was performed under RNase free conditions (section 2.7.1). The two protocols described here both employ a common extraction procedure. Method 2.7.2 was used in preference, although method 2.7.3 also gives good quality RNA and was used on occasions when a number of batches were processed together.

2.7.1 Preparation of Solutions, Plasticware and Glassware for RNA Work

RNA is only slightly less chemically stable than DNA, but the ubiquitous presence of RNases in the working environment necessitate the use of special working conditions in order to preserve the RNA in intact form.

Glassware was treated by baking at 250°C for at least 48 hours. Open containers are sealed with aluminium foil during this process. After baking the glassware was stored in a closable cupboard until use.

Plasticware was treated by total immersion in a 0.1% suspension of DEPC. The DEPC suspension was prepared by mixing a quantity of the solution using a magnetic stirrer for 15minutes before use. Plasticware that cannot be autoclaved is simply rinsed after DEPC treatment.

Solutions used for RNA work are made 0.1% DEPC, left to sit overnight and then autoclaved. Solutions such as Tris-HCl and EDTA cannot be treated in

this way and so must be made from the purest unopened chemical stock available.

2.7.2 Isolation of RNA Using Caesium Chloride and Guanidine thiocyanate

This method is a modification of the method of Chirgwin *et al.* 1979 as described by Sambrook *et al.* 1989. 10mls of homogenisation buffer (4M guanidine thiocyanate, 0.1M Tris. pH8.0 and 0.1μl/ml antifoam A) was used per gram of tissue. Before homogenisation the flies are anaethesised under carbon dioxide. If material was used that had been frozen at -70°C then it was allowed to warm up slightly before homogenisation. This is to prevent crystallisation of the guanidine thiocyanate from the homogenisation buffer.

Homogenisation was performed using a Polytron homogeniser at 5000rpm for 60 seconds. The homogeniser produces dangerous aerosols and so the homogenisation was performed in a fume hood at room temperature. After homogenisation, skeletal debris was removed by sieving the homogenate through a muslin cloth. The homogenate was then centrifuged at 4000xg for five minutes at room temperature. The supernatant was then collected and the pellet discarded.

The homogenate was layered onto a cushion of 5.7 M caesium chloride, 4mM EDTA pH 7.5 in an open ultracentrifuge polyallomer tube (Beckman). Two sizes of tube were routinely used depending upon the scale of the preparation. Small scale preparations employed a SW50.1 (Beckman) rotor. For this rotor a 1.3mls caesium chloride cushion with 3.6mls of homogenate was used. Ultracentrifugation was performed at 47k for a minimum of eight hours at 20°C. The second rotor was an SW28 with 40ml buckets and in this case each tube contained 9mls of caesium chloride and 29mls of homogenate. Ultracentrifugation was performed at 27k for at least 26 hours and at 20°C.

After centrifugation the tubes were carefully removed from the rotor. The RNA was expected to be in a glassy pellet at the bottom of the tube. DNA and protein remains at the homogenisation buffer/caesium chloride interphase and so to prevent contamination with these macromolecules the buffer was carefully aspirated off using a weak vacuum pump to leave the caesium chloride solution in the bottom of the tube. The tube was then quickly inverted to remove this solution. In order to prevent contaminants running

back into the RNA sample after the tube is tipped back upright, all but the last centimetre of the tube was cut off using a hot scalpel. This leaves the pellet containing the RNA in a small dish. The cut down tube is allowed to dry inverted over lint free tissue, although the pellet was never allowed to dry fully.

The pellet was resuspended in 5% phenol pH8.0 and 3% sodium lauryl sarcosine. Once resuspended the RNA was extracted once with 50% phenol pH8.0, 50% chloroform and the aqueous phase precipitated using sodium acetate pH4.0 and ethanol (section 2.2.10).

The precipitated RNA was stored conveniently under ethanol at -20°C before use. The quality of the RNA preparation was assessed initially by running a sample out on a 1% agarose 1xTBE buffered gel and ethidium staining. The concentration was determined by spectroscopic means (section 2.2.2.8). The quality of preparations were also assessed by northern blotting (section 2.9).

2.7.3 Isolation of RNA Using GCN and Phenol Chloroform

This method is a modification of that described by Chomczynski and Sacchi 1987. Homogenisation was performed as section 2.7.1 using the same buffer. The homogenate was made 0.3M for sodium acetate pH4.0. To the aqueous mixture was added an equal volume of phenol and one eighth volume of chloroform. The phenol was prepared by adding hydroxyquinoline equilibrating pure phenol with one sixth volume of 1M Tris pH8.0. The aqueous phase was then replaced with sterile water. The mixture of homogenate, phenol and chloroform was vortexed vigorously and allowed to sit on ice for 20minutes. After incubation on ice the mixture was centrifuged at 4000xg to separate the organic and aqueous phases. The aqueous phase was transferred to a fresh tube, to which was added an equal volume of isopropanol and the tube vigorously mixed. The preparations were then incubated on ice for 1 hour.

To recover the precipitated RNA the samples were centrifuged at 7000 xg, 4°C for twenty minutes. The supernatant was removed and the pellet resuspended to a final approximate concentration of 1mg/ml. The quality and yield of RNA was assessed using denaturing (section 2.9) and non-denaturing electrophoresis (section 2.2.12) and by spectroscopic means (section 2.2.8). Typical yields from the preparation were 1mg of RNA per gram of tissue.

2.8 Preparation Of PolyA+ RNA Using Oligo dT Cellulose

The method described here is a modification of the method of Aviv and Leder 1972. This protocol was performed under RNase free conditions (section 2.7.1).

The oligo dT cellulose used in all poly A^+ preparations was Pharmacia type VII. Prior to use the resin was swollen by immersion in 40 volumes of 1x Binding Buffer (0.5M NaCl, 10mM Tris pH7.5. 0.1% sarcosine and 1mM EDTA) for 1hour at 4°C with occasional mixing The swollen resin was washed in 40 volumes of 0.1M sodium hydroxide. This wash, and all subsequent washes was performed at room temperature. Washing was achieved by gently mixing and then spun down spinning for a few seconds in a centrifuge. The resin was then washed in 40 volumes of RNase free water. Washing in water was repeated until the pH became neutral. After neutralisation the resin was washed in 40 volumes of 1x binding buffer.

Prior to binding to the resin the RNA must be brought to a suitable buffer conditions and also denatured. The RNA solution was made 1x binding buffer such that the final RNA concentration is 1mg/ml. The RNA in 1x binding buffer was then heated to 70°C by emersion in a water bath for 5minutes, quickly chilled on ice and then added to the drained resin. A packed resin volume of 1ml can be used to process 10mg of total RNA.

Binding was performed at room temperature on a rotary mixer for 30minutes. After binding the resin was spun down to the bottom of the tube by centrifuging briefly at low speed and then the supernatant removed. The supernatant should not contain any poly A^+ RNA but was kept on ice as a precaution. The resin was then washed with three changes of 40 volumes 1x binding buffer. These were washed for 5minutes each on the rotary table before the resin was spun down.

The RNA was eluted from the resin by treatment with water that had been heated to 70°C. A volume of heated water equal to that of the resin was added to the drained resin and the mixture vortexed. The suspension was spun down by quick low speed centrifugation and the supernatant removed. The washing of the resin with water was repeated four further times. The supernatant contains the RNA enriched in poly A^+ message.

A second cycle of binding and elution of the RNA was then performed. A small volume of each eluted fraction was spotted onto an ethidium plate and this used to determine the RNA concentration as described section 2.2.9. If the poly A^+ RNA concentration was expected to be below 5ng/μl then the RNA cannot be detected directly by use of ethidium plates. In this case the RNA samples were separately precipitated with ethanol, resuspended in a smaller volume and then assayed as section 2.2.9. Ethanol precipitation was performed as described in section 2.2.10 except that the sodium acetate and ethanol precipitated samples were centrifuged at 50k in a Beckman ultracentifuge SW50.1 rotor for 30 minutes at 4°C.

PolyA^+ RNA was resuspended to an approximate concentration of 1μg/μl. The yield and quality of the preparation was assessed using denaturing (section 2.9), and non-denaturing electrophoresis (section 2.2.12) and by spectroscopic means (section 2.2.8). Typical yield for a preparation are 20μg per mg of total RNA.

2.9 Northern Blotting

Northern blotting was performed as described in Sambrook *et al.* 1989 using formaldehyde agarose gels. Samples were prepared as described except that 10μg ethidium bromide was added to each RNA sample. This enabled direct visualisation of the sample in the gel and on the surface of the membrane. Blotting was onto Hybond Ce (Amersham) which is a nylon reinforced nitrocellulose membrane.

2.10 Preparation of Radiolabelled Probes

Radioactive materials were handled as described in section 2.2.16.

2.10.1 Using Random Priming of DNA Templates

Random prime labelling of DNA was performed exactly as described by Feinberg and Vogelstein 1983 except that samples were denatured by heating at 70°C for 5minutes and chilled on ice and Klenow fragment was used. Typical reactions contained 5-100ngs DNA, 30μCi $-^{32}$ P dCTP and 2 units of Klenow in a total volume of 50μl. Incubation overnight at room temperature was preferred. The probe was then treated as described in section 2.2.18.

2.10.2 End labelling DNA Using Klenow

This is a simple method whereby the radioactive nucleotide is incorporated by filling in the 5' overhang remaining after digestion with a suitable restriction enzyme. The buffer used was simply the restriction digest buffer; after digestion of the template with restriction enzyme 30μCi of an appropriate ^{32}P-dNTP was added together with 1-2units of Klenow. The reaction was allowed to incubate at 37°C for 1 hour. The probe was then treated as described in section 2.2.18.

2.10.3 End Labelling of Oligonucleotides Using T4 Polynucleotide Kinase

This method was performed exactly as described in the Promega Protocols and Applications Guide using standard Promega buffers and up to 5μg of single stranded oligonucleotide.

2.10.4 Nick Translation

This protocol was carried out essentially as described in Sambrook *et al.* 1989 using 4-500ng DNA template and 40μCi α^{32}P-dCTP radionucleotide.

2.11 Hybridisation of Radiolabelled Probes to Hybridisation filters

These procedures were carried out using "safe" radioactivity handling conditions as described section 2.2.16.

2.11.1 Hybridisation to Filter Bound DNA at High Stringency

This was performed as described in Sambrook *et al.* 1989 and optimised by Amersham for Hybond N membrane. The Amersham protocol was modified to use a more complex hybridisation solution containing 6x SSC, 1% SDS, 0.1% PVP, 0.1% Ficol, 0.1% BSA, 100μg/ml sheared, denatured, herring sperm DNA (prepared as described Berger and Kimmel 1987), 12.5mM $NaPO_4$ pH6.8, 0.05% NaPPi and 1mM EDTA. Hybridisations were carried out in a rotary oven (Techne) or using sealed plastic bags in a shaking water bath.

After hybridisation the filters were washed in an excess volume of 2x SSC for twenty minutes followed by two washes in 2x SSC, 1% SDS and then two in 02xSSC, 0.1% SDS. All washes except the first were with prewarmed solutions at 68°C. The first wash was at room temperature. The filters were then autoradiographed as described section 2.19.

2.11.2 Hybridisation to Filter Bound DNA at Low Stringency

Low stringency hybridisation was performed as described by Cassilli *et al.* 1992, except that the final washes were in 2x SSC at room temperature, 2x SSC at 50°C, 0.2x SSC at 50°C and finally 0.2x SSC 68°C. The damp hybridised membrane was washed and autoradiographed at each successive stringency. The final wash is equivalent to a high stringency probing under more standard conditions described in section 2.11.1.

2.11.3 Hybridisation to Filter Bound RNA at High Stringency

This was exactly as described in Sambrook *et al.* 1989 for nitrocellulose membranes, but using Hybond Ce, a reinforced nitrocellulose membrane.

2.12 ExonucleaseIII Deletions

The method used for this procedure was a slight modification to that described by Promega. It is derived from the method of Henikoff 1984. The process utilises the ability of exonucleaseIII (exoIII) to progressively digest one strand of double stranded DNA starting at the molecules termini. This digestion can be performed from one DNA end only if the other is "protected" from digestion. This was either by the use of 3' overhanging termini or by the incorporation at one termini of an α-phosphorothioate derivative base. Digestion with exoIII produces a long single stranded overhang that is then removed using S1 nuclease and the deleted molecules cloned.

2.12.1 Protection by 3' Overhang

10μg of plasmid (2-3μg of insert) was digested with the enzyme (as described in section 2.2.5) used to generate the 3' overhang. A small sample of this reaction was taken (200ng) in order to assess the extent of cutting using gel electrophoresis as described in section 2.2.12. If the digestion had progressed to completion, digestion with the second enzyme was performed; digestion conditions were adjusted to suit the second enzyme. The extent of cutting by this enzyme cannot be assessed by electrophoresis. The success of this restriction was tested using a test exoIII digest.

The double digested template was extracted once with an equal volume of phenol pH8.0 and chloroform. After addition of the phenol/chloroform, DNA was then precipitated using sodium acetate and ethanol (section 2.2.10). The

sample was resuspended in TE pH8.0 (section 2.1.4) at a concentration of 200ng/μl.

2.12.2 Protection from Digestion by Incorporation of α-Phosphorothioate

An alternative means of template preparation employed a fill-in reaction using α-phosphorothioate (thio) base derivatives. ExoIII is unable to digest through termini containing these bases. To prepare this template, the protection enzyme was a 5' overhang generating enzyme. After digestion thio dNTP mix to a final concentration of 10μM was added together with 2units of Klenow enzyme. After mixing the reaction was incubated at room temperature for 15minutes. The reaction was stopped by heating to 65°C for 5minutes followed by extraction with an equal volume of 50% phenol pH8.0 and 50% chloroform. The DNA was ethanol precipitated using sodium acetate (section 2.2.10) and then resuspended in the second enzymes reaction buffer. A small sample of this single cut template was retained (200ng) for test digestion with exoIII. The remainder was digested with the second enzyme. The template was then extracted once with an equal volume of 50% phenol pH8.0 and 50% chloroform, before being again precipitated with ethanol and sodium acetate. The template DNA was resuspended in TE pH8.0 to a concentration of 200ng/μl.

2.12.3 Test Digestion with ExonucleaseIII

The templates from both preparation methods were tested by a test digestion. The 10μl reaction contained 6.6mM Tris pH8 and 66mM magnesium chloride together with the test template at a final reaction concentration of 20-50ng/μl. To this was added 100units of exoIII enzyme and the reaction incubated for 15minutes at 37°C. The reaction products were then electrophoresed through a 1% non-denaturing agarose gel (as described section 2.2.12). Digestion of the double cut template was indicated by a change in migration size of the products to the undigested template.

2.12.4 Full Scale ExonucleaseIII Digestion and S1 Nuclease Treatment

ExoIII digestion was performed using 9.5μg template at 0.2μg/μl in ExoIII reaction buffer (6.6mM Tris pH8 and 66mM magnesium chloride and 30-50 units of enzyme per microgram of template). Typical final reaction volume was 60μl. Before addition of the enzyme the reaction was prewarmed at 37°C for 5 minutes. Aliquots of this reaction were taken every 30seconds. At 37°C exoIII progresses 450bp/min. The number of aliquots taken was

adjusted so that the plasmid insert can reasonably be spanned by the deletion series. Each 2.5µl aliquot was placed into 7.5µl ice cooled S1 nuclease reaction containing 2units of S1 nuclease in 1x S1 nuclease buffer (45mM KOAc pH 4.6, 350mM NaCl, 1.3mM $ZnSO_4$ and 7% glycerol). All the aliquots were then returned to room temperature and incubated for 30minutes. The reaction was stopped by incubation at 37°C for 10 minutes in 30mM Tris base and 5mM EDTA.

2.12.5 Ligation and Transformation of Plasmids

Since the templates are linear molecules containing inserts of various sizes with an attached vector sequence it is only necessary to recircularise the plasmid in order to obtain a viable recombinant clone. Standard ligation conditions (2.1.5) were adjusted to favour intramolecular ligations of blunt ended DNA.

Each timepoint was made 1x for the exoIII Klenow reaction mix (section 2.10.1) and the reaction, which contains 0.1units of Klenow was incubated for 10minutes at 37°C. The ligation reaction was performed using 25% sample from the Klenow reaction and 75% ligation mix (1x Promega ligase buffer, 5% PEG, 1mM DTT 5U/ml T4 ligase). Samples were ligated overnight at 15°C. The ligated plasmids were then used for XL1-B transformation (section 2.3.4).

2.13 Sequencing of Double Stranded Plasmid

The protocol used was a modification of that of Sanger *et al.* 1977. Templates for double stranded plasmid sequencing were produced using the method in section 2.2.4.1 and 2.2.4.2. Sequencing reactions were performed using the Sequenase II kit (USB Biochemicals) exactly as described by USB, except for the following modifications:-

5µg of double stranded plasmid (containing 1-3µg of insert) was used per synthesis reaction. 10ng of each primer was used per template. The vector was denatured by making it 200mM sodium hydroxide at a plasmid concentration of 250µg/ml. The mixture was incubated at 37°C for 30minutes before being made 250mM for ammonium acetate pH4.6 and precipitated by addition of 2 volumes of ice cold ethanol and chilling for 30minutes at -70°C. The template was then recovered by centrifugation at 4°C and resuspended in standard Sequenase annealing buffer mix except that

DMSO had been added to 10%. The Sequenase protocol was otherwise unchanged except that the termination reactions were also performed at 42°C.

2.14 Differential Screening

The method employed was a modification of St. John and Davis 1979.

2.14.1 Preparation of Ordered Arrays of Plaques and Plaque Lifts

A device for the picking of ordered arrays was developed in the laboratory (Mackenzie *et al.* 1989). Ordered arrays were produced using this device. The 'phage library was first plated out randomly as described in section 2.3.2. 200 'phage were plated onto a single plate. Single plaques were picked off this plate using fine glass rods that slip through a hole, which is one of many drilled in an ordered array in the top of the device. Once all the holes are filled, a plate, containing a lawn of host bacteria (section 2.3.2), was raised until it touches the suspended rods. The plate was held in position for 30 seconds before the rods lowered. This plate was then cultured until the array of plaques was clearly visible. This array can be placed in the device in the same orientation as before and used to "recharge" the rods. Duplicate replicas were then taken and grown as for the master plate.

2.14.2 Plaque Lifts for Differential Screening

Filters replica lifts were produced from array plates using Hybond-N membrane exactly as described by Amersham. DNA was baked to the filters for 2-3hours at 80°C. After baking the filter was washed in a large volume (200mls) of 2x SSC (20x SSC; 3M NaCl, 0.3M sodium citrate pH 7.0) at 85-90°C for 20minutes. The filter was rubbed to remove any adhered agarose. The filter was then washed in 1x Triton-x100 for five minutes at room temperature, before being quickly washed in 2x SSC at room temperature. The filters were then blotted dry.

2.14.3 Preparation of Probes by Reverse Transcription

For the synthesis reaction one of two reverse transcriptases were used. The first AMV (Pharmacia) and the second an engineered MMLV enzyme, Superscript (Gibco-BRL). The protocol shown is for Superscript, but is essentially the same as that used for AMV except that the reaction conditions have been adjusted to suite the second enzyme [this involves changing the salt conditions and decreasing the temperature from 45°C to 42°C]. AMV buffer

was as recommended by Pharmacia and Superscript buffer was supplied by Gibco-BRL.

The template mRNA must first be denatured by heating together 1μg of mRNA in 5μl of RNase free water together with 100ng of oligo dT primer (oligo dT_{12-18} from Pharmacia) at 70°C for 5 minutes. After denaturation the template was then chilled on ice. The final reaction mix was then made up to contain 1x Superscript reaction buffer (Gibco-BRL), 10mM DTT, 0.5mM of dNTP other than dCTP, 70μCi ^{32}P dCTP (800 Ci/mmole) and 10 000u/ml of Superscript. Typical reaction volume was 20μl. The reaction was incubated at 45°C for 1 hour. After this, one fifth volume of a mix containing 0.5mM dCTP and 10 000u/ml superscript was added and incubation continued for a further 60 minutes.

The mRNA template was hydrolysed by making the reaction 0.3N for NaOH and 10mM for EDTA and incubating at 65°C for 45 minutes. The probe was separated from the unincorporated nucleotides by Sephadex G50 chromatography as described section 2.2.18. After Sephadex G50 chromatography the probe was treated by passing through hydroxylapatite as described in section 2.15.2. This has been observed in the laboratory to reduce background in cDNA probes. Since the probe is all single stranded, all the probe should elute in one fraction. It is then used for hybridisation as described in section 2.15.2.

2.14.4 Prehybridisation, Hybridisation and Washing of Filters

Prehybridisation and hybridisation was carried out essentially as section 2.11.1 except that the reaction buffer also contained 100μg/ml polyA (Pharmacia). Filters were placed singly in bags and kept at 68°C using a water bath. Prehybridisation was for 12 hours and hybridisation for 18-24 hours. The filters were washed as section 2.11.1.

2.15 Subtraction Screening

This protocol is a modification of that described in an EMBO practical course guide (EMBO, Heidelberg, West Germany). The protocol requires the use of both RNase free conditions (section 2.7.1) and safe handling of radioactivity (section 2.2.16).

2.15.1 Solution Hybridisation

Probes for use in subtractive hybridisation experiments were prepared as section 2.14.3. Probes were routinely prepared using 150μCi dCTP (800Ci/mmole). To the probe in aqueous solution was added a 20-30fold molar excess of driver mRNA. This was polyA+ RNA prepared as described in sections 2.7 and 2.8. Both the cDNA and mRNA were then precipitated in sodium acetate pH 4.5 and ethanol (2.2.10).

RNA and cDNA should be soluble at 10μg/μl (EMBL practical guide), although in practice 4-5μg/μl seems more realistically attainable. 1x subtraction buffer (120mM $NaPO_4$ pH6.8, 820mM NaCl, 10mM EDTA and 30μg/ml poly-U (Pharmacia)) was used for the hybridisation. Solubalisation of the probe was performed in water before addition of the 2.5x buffer to a final 1x reaction concentration. Solublisation was aided by heating the components to 65°C for 2 minutes, vortexing and then chilling on ice. The sample was also drawn repeatadly into an eppendorf tip, although when this method was used the sample volume must be adjusted as evaporation of the buffer. Once in solution in 1x buffer the reaction was transferred to a fresh tube and a layer of oil added to prevent evaporation during incubation. The sample was incubated, without shaking, for 24 hours at 68°C.

2.15.2 Hydroxylapatite (HAP) Chromatography

This method is modified from Britten *et al.* 1974. The HAP chromatography was performed in a water jacketed column at 60°C. This was performed by enclosing a 1ml syringe in a 2ml and arranging the outer to act as a water jacket. The columns allowed to equilibrate for 1 hour at 60°C before use. The inner syringe barrel was plugged using siliconised glass wool.

The HAP resin used was first batch tested to ensure that the single and double stranded DNAs elute with the appropriate fractions. If DNA grade HAP (Boeringer Mannheim) is used as here, the elution profile of the resin has already been calculated. To prepare the HAP resin it was hydrated in 20 volumes of PB40 buffer (40mM $NaPO_4$ pH6.8, 0.05% SDS). The resin was mixed in the buffer and the HAP allowed to settle. The milky supernatant was removed and the process repeated until the supernatant was almost colourless. The resin was then added to the column until a 0.5ml packed bed volume had formed. The column was equilibrated using 4 column volumes of PB40 before use.

To prepare the solution hybridisation for loading onto the column the oil was first pipetted off the mix. The probe was then loaded directly onto the column in an excess of PB40 buffer that had been heated to 60°C. The head of liquid was allowed to run through the column. The eluate was collected and stored. The column was then washed in six washes column volumes of PB40 that had been preheated to 60°C. Each fraction was collected and stored. The single stranded fraction was eluted using six column volume washes of PB120 (120mM $NaPO_4$, 0.05% SDS). The double stranded fraction was then eluted in six washes of 1column volume of PB400 (400mM $NaPO_4$, 0.05% SDS). The remaining radioactivity on the column was assessed by removing the resin or washing in several volumes of PB400 at 95°C. Each fraction was counted and an elution profile constructed for the column by Cherenkov counting as described in section 2.2.15. The single stranded fraction which is to be used as a probe was made up to a 6x SSC final salt concentration and added to a prehybridised filter as described in section 2.14.4.

2.15.3 A Second Round of Subtraction?

After the first subtractive hybridisation if the fraction of double stranded cDNA was lower than expected then a second round of subtraction was performed. To do this the probe from the first round needs to be removed from PB120. To do this the fraction was desalted using a G50 column (section 2.2.18) and taking the peak radioactive fraction(s). The probe volume was reduced with the addition of an equal volume of butanol and removal of the organic phase until the sample could be precipitated using ethanol and NaOAc pH4.5 (section 2.2.10). The pellet was resuspended in a suitable volume of RNase free water (~20µl) and another 20-30µgs driver mRNA added. The hybridisation and HAP chromatography were then repeated as described in sections 2.15.1 and 2.15.2 and the single stranded fraction used as a probe.

2.16 Construction of a Genomic Library

λGem11, a bacteriophage lambda replacement vector, was supplied by Promega and had already been prepared for ligation of *Sau*3A1 partially digested DNA. Construction of the library followed the method described in the Promega Current Protocols and Applications Guide for the supplied vector.

300μg of DNA was restricted with *Sau*3A1 such that the average fragment size was 14-22kb when assessed using a 0.4% agarose gel. The DNA was size fractionated by passing through a 40% to 10% sucrose gradient at 26k, 20°C for 18hours in a SW28 rotor (Beckman). After fractionation, 0.5ml samples were taken from the gradient by insertion of a needle into the base of the tube and allowing the sucrose to drip out. Each fraction was electrophoretically sized (compared to size markers treated with varying concentrations of sucrose) and fractions containing DNA in the 18-22kb size range retained. These fractions were dialysed against T.E (10mM Tris pH 8.0, 1mM EDTA) and the DNA precipitated using ethanol and sodium acetate (section 2.2.10). The samples were resuspended in 10μl of T.E and an aliquot used to again check the size of the fraction. The fraction containing fragments averaging 18kb was used for ligation to the vector.

Ligations were prepared as described by Promega and the library packaged using Gigapack Gold II packaging mix (Stratagene). The library was amplified using NM621 host cells by plating 20,000 PFUs/plate onto square 10cm plates and harvesting the phage as described by Sambrook *et al.* 1989.

2.17 *In situ* Hybridisation to Whole Mount Embryos

The method described in this section was obtained from the laboratory of D.Glover (Dundee Biochemistry Department) and employs a Boeringer Mannheim Digoxygenin (DIG) labelling and detection kit.

The probe used is generated by random priming (refer section 2.10.1) to incorporate a nucleotide to which is conjugated the DIG label. Hybridisation of this DIG labelled probe to the tissue sections is detected using an antibody specific for DIG (which does not occur naturally in *Drosophila*). This antibody carries alkaline phosphatase. The presence of this phosphatase is detected by the use of a colour reaction using X-phosphate and NBT.

2.17.1 Preparation of Embryos

Mixed stage embryos were prepared as described in section 2.4.3.

The embryos were dechorionated by immersion in a 50% solution of household bleach for 4 minutes at room temperature. They were then rinsed in copious quantities of tap water to remove all traces of the bleach.

The embryos (of about $2cm^3$ settled volume) were then placed in 5mls of Paraformaldehyde/PIPES solution. [This solution was freshly made, in a fume hood, and contained 4% paraformaldehyde, 100mM PIPES (pH 7.0), 1mM EGTA (pH 8.0) and 2mM magnesium sulphate]. To the paraformaldehyde/PIPES and embryo mix was added an 5mls of heptane. This mixture was then gently rocked on a mechanical rocking device for 90 minutes. After this, half the upper organic phase and all the lower aqueous phase was removed using a Pasteur pipette.

To the mix of embryos and organic phase was then added 5mls of a solution containing 90% methanol and 10% 0.5M EGTA pH8.0. The mixture was gently shaken for 5 minutes and then the phases allowed to separate. Those embryos that have been devitellinised sink to the bottom of the tube and were collected using a wide bore Pasteur pipette.

To the dechorionated and devitellinised embryos was added 1ml of 90% methanol and 10% 0.5M EGTA pH8.0. The tube was inverted, the embryos allowed to settle, and then the solution removed. This rinse was repeated twice. The embryos were then rehydrated by mixing in a series of methanol and paraformaldehyde/PIPES mixture. The first solution was added contained 70% methanol with the balance (30%) made up from Paraformaldehyde/PIPES solution. The embryos were mixed and then remained in this solution for 5 minutes before being successively transferred to solutions containing 50%, 30% and 0% methanol. The embryos were immersed in the last solution for 20 minutes at room temperature with occasional inversion of the tube.

2.17.2 Preparation of DIG Labelled Probe

Fragments used to generate probes were excised from the vector using a suitable restriction digest and gel purified as described in section 2.2.6. 200ngs of this gel purified fragment was then used to produce each DIG labelled probe. Briefly, the DNA was denatured at 100°C for 5 minutes and quickly chilled on ice before addition to the labelling mixture. The labelling reaction was performed using the Boeringer Mannheim kit and as described in the instructions for this kit, except that 10μCi ^{32}P dCTP was added to the mixture in order to monitor the yield of the synthesis reaction. Incorporation of dCTP was assayed as described in section 2.2.15. Probes were precipitated by adding an equal volume of 4M ammonium acetate and adding EDTA

pH8.0 to a final concentration of 10mM. This mixture was vortexed briefly and then two volumes (100µl) of ice cold ethanol added. After mixing, the precipitation was placed in a bath of dry ice and ethanol.

After 30 minutes in the ice/ethanol mix the mixture was centrifuged in a microfuge on full speed for 10 minutes. The supernatant was carefully removed and the pellet air dried by inversion of the tube over a piece of clean tissue. After drying the pellet was resuspended in 150µl in hybridisation solution (50% formamide, 5x SSC, 100µg/ml sheared herring sperm DNA, 50 µg/ml heparin and 0.1% tween-20).

2.17.3 Prehybridisation and Hybridisation

The embryos were then transferred to 1ml of PBT (PBS which is 130mM NaCl, 3mM KCl, 9mM NaH_2PO_4, 2mM KH_2PO_4 with 0.1% Tween-20 added), gently mixed by inversion and allowed to settle for 10 minutes. The process was repeated twice into fresh tubes of PBT. The supernatant was then removed and 1ml of 50µg/ml proteinase-K in PBT was added. The digestion was incubated for 3.5minutes at room temperature. The exact digestion properties of the enzyme were batch tested before use. The reaction was stopped by washing the embryos in two washes of 1ml of 1mg/ml glycine in 1x PBT. Washes were for one minute each. The embryos were then washed twice by gentle inversion in 1ml of 1x PBT. They were then fixed in 1ml of freshly made paraformaldehyde/PIPES for 20 minutes. After fixation the embryos were washed in six changes of 1x PBT by gentle inversion and settling for 5 minutes per wash.

The embryos were now prehybridised. The embryos were transferred to a 1.5ml eppendorf tube and washed in 1ml of 50% PBT and 50% hybridisation mix (hybrix; 50% formamide, 5x SSC, 100µg/ml sheared herring sperm DNA, 50µg/ml heparin and 0.1% tween-20). They were washed by gentle inversion for 10 minutes and then washed in 1ml of hybrix solution. The embryos were transferred to a fresh 1ml of hybrix and placed at 45°C for 1 hour and occasionally mixed by gentle inversion. After 1 hour the volume was reduced to 100µl and the probe was then added. The added probe mix volume was half that of the settled embryos. Hybridisation was performed for 8-12 hours at 45°C. The hybridisation may be gently mixed by rolling within a hybridisation tube.

2.17.4 Washing and Antibody Hybridisation

After hybridisation with the DIG labelled probe the excess probe is removed by first washing in 500µl of Hybrix for 20 minutes. Mixing was by occasional gentle inversion. The embryos were then washed in two changes of 500µl of 50% 1x PBT, 50% Hybrix for 20 minutes each. The embryos were then washed with two changes of 500µl 1x PBT. All these washes were at 45°C and the wash solutions were pre-warmed to 45°C before use. The embryos were then washed with four changes of 1xPBT for five minutes each change at room temperature. Each wash volume was 500µl.

The embryos were then treated with pre-absorbed Anti-DIG antibody from the Boeringer Mannheim detection kit. To preabsorb the antibody a mix containing 1µl of antibody per ml of 1xPBT was treated with 100µl/ml (settled volume) of embryos. These embryos had been treated as the those already described, but not prehybridised. The antibody/PBT mix was treated with these embryos for 1.5 hours by rolling on a rotary mixer at room temperature. The embryos used for preabsorbtion of the antibody were then discarded. The preabsorbed antibody was then added to the main batch of embryos. A total volume of 1.5mls was used per 50µl of settled embryos and the antibody hybridisation mixture was mixed using a rotary mixer at room temperature for 1.5 hours.

The embryos were then washed quickly in 1ml of 1x PBT at room temperature and then washed again in three washes of 1x PBT for 20 minutes each. All washes were using the rotary shaker. The embryos were then added to a 50% mix of 1xPBT and NMTT (100mM NaCl, 50mM $MgCl_2$, 100mM Tris pH 9.5, 0.1% tween-20 and 1mM lavamesol), mixed for five minutes using the rotary mixer before being transferred to 1ml of NMTT. The embryos were washed in this solution for five minutes using the rotary shaker at room temperature and the wash repeated twice.

2.17.5 Colour Detection

Colour detection was using the Boeringer Mannheim colour reagents as described in the Boeringer Mannheim protocol. To 1ml of NMTT was added 3.5µl X-phosphate and 4.5µl NBT (from the kit). To this mixture was added the embryos. The colour reaction was mixed by rotation in the dark at room temperature. Samples of this reaction were taken over successive time periods ranging from 30 minutes to 4 hours. To stop the colour development in these

samples they were first rinsed in 1ml of 1x PBT at room temperature before being transferred to 300µl absolute ethanol for storage in the dark at 4°C until mounting. Transfer to absolute ethanol from PBT was via a dehydration series of 30, 50 and 70% ethanol. Each intermediate wash was for 5 minutes. Progress of staining was assessed by examination under a light microscope (section 2.2.17).

2.17.6 Mounting of Embryos

The mountant used was GMM (Roberts 1986) which is Canada Balsam based. Several strips of sticky tape were added to a slide such that they could act as a rest for a coverslip. This rest was to prevent the cover slip from crushing the whole embryos. A suitable quantity of embryos (30-40 embryos initially) were drawn up into a 1ml widened eppendorf tip. The tip was held vertically for a few second until all the embryos fall down into a drop of liquid slowly exuded from the tip. This drop of liquid was then touched on the surface of the GMM on the slide, the embryos were released into the mountant in a drop of ethanol. A coverslip can then be placed over the mounted embryos.

2.18 *In situ* Hybridisation to Head and Head/Body Sections

2.18.1 Fixation and Sectioning of Tissues

To prepare head sections the flies were placed in a Heisenberg collar (as described Ashburner 1989), soaked in OCT mountant (BDH) for 30 minutes at room temperature before being frozen in a cryostat. 10µM sections were cut, placed onto gelatinised slides and allowed to dry at room temperature for 1-2 minutes. Head/body sections were prepared in a similar manner, except the flies were not mounted in the collar. After drying the sections were then fixed in freshly made PLP (2% paraformaldehyde, 0.1M $NaIO_4$, 75mM lysine and 1x PBS).

2.18.2 Treatment of Sections for Hybridisation

Fixed sections were then rinsed in 1x PBS for five minutes, before being placed in 0.2N HCl for 20minutes. The sections were then washed in 2x SSC for 30minutes at 65°C. Sections were then pronase treated by covering with a solution of 350µg/ml pronase in 1x TE for 10minutes at 37°C. To stop the proteolysis the slides were rinsed in distilled water, incubated in a large volume of 2mg/ml glycine in 1x PBS for 1 minute and then washed in 1x PBS for 5 minutes.

The sections were then treated with 4% paraformaldehyde in 1x PBS for twenty minutes to fix the sections. The sections were then briefly washed in 1x PBS. The sections were then treated to an acetylation reaction by bathing the slides in a gently stirring freshly made mixture of 0.1M triethanolamine, 0.25% acetic anhydride in 1x PBS.

2.18.3 Prehybridisation and Hybridisation of the DIG Labelled Probe

The slides were placed in 2x SSPE (300mM NaCl, 18mM NaH_2PO_4, 2mM EDTA) for 10 minutes and then the excess liquid removed from the slide using a tissue. The sections were then bathed in prehybridisation solution (5x SSPE, 50% formamide, 5% dextran sulphate, 0.05% PVP, 0.05% Ficol, 0.05% BSA, 500 μg/ml sonicated salmon sperm DNA, 250 μg/ml yeast tRNA) and covered with a siliconised cover slip. 200μl of prehybridisation solution was used for each slide. Prehybridisation was carried out for 1 hour at 42°C.

DIG labelled probes were prepared as described for the probe prepared for whole mount embryos. The probe was denatured by boiling for two minutes before being resuspended in prehybridisation solution at a final concentration of 50ng/ml. The prehybridisation solution was removed from the slides and the sections covered in about 150μl hybridisation solution. The slides were incubated overnight at 42°C.

After hybridisation the slides were washed to remove the non-hybridised probe with 5 washes in 2x SSPE at room temperature for 15 minutes, two washes in 1x SSPE at room temperature for 15 minutes and two washes in 0.5x SSPE at room temperature for 15 minutes. Finally the slides were washed in 2mM NaPPi, 1mM NaPO4 and 1mM EDTA at 42°C for 15 minutes.

2.18.4 Immunological Detection

Slides were washed in an excess volume of bufferI (100mM tris pH 7.5, 1mM EDTA) for 10 minutes at room temperature. Sections were then incubated in 200μl of 2% sheep serum in bufferII (0.3% Triton X-100 in bufferI) for 60 minutes. This solution was removed from the slides and replaced with the anti-DIG antibody supplied in the Boeringer kit. 200μl of a 1:500 dilution of

this antibody in solutionII was added to each slide. The slides were covered with a coverslip and incubated at room temperature for 3 hours.

After incubation the unbound antibody was removed by washing five times in bufferI for 15minutes each wash. The slides were then washed once in bufferIII (100mM Tris pH 9.5, 100mM NaCl, 50mM $MgCl_2$) for 3minutes. 150µl of solutionIV(10mM Tris pH 8.0, 1mM EDTA) was then used to cover the sections and the slides incubated in a light tight box to allow colour development to take place, which took between 2 and 24hours. The colour reaction was stopped by washing in an excess of TE pH8.0 for 10minutes and then by three further washes in distilled water. Finally excess water was removed from the sections and 2-500µl of Aquamount (Sigma), prewarmed to 40°C added to each slide. A coverslip was carefully placed on this mountant.

2.19 Polytene Chromosome Hybridisation

Dissection of salivary glands and squashing of chromosomes were carried out as described in Roberts 1986.

2.19.1 Pretreatment of Chromosomal Squashes

Slides containing the squashed chromosomes from three glands were kept at 4°C to partially dehydrate for up to two weeks. After this the slides were frozen in an ethanol/dry ice bath and the coverslips removed using a fine blade. The slides were then placed in a 25% ethanol, 75% glacial acetic acid bath for 10 minutes and then into 100% ethanol for 10 minutes before being air dried. Chromosomes were now examined to check the quality of the preparation. Slides containing chromosomes that appear glassy were discarded; these chromosomes were not sufficiently flat on the slide to be of use in the *in situ* procedure.

Slides were incubated in a bath of preheated 2x SSC at 65°C for 30 minutes and then rinsed in two washes of 2x SSC at room temperature. The chromosomes were then treated to an acetylation reaction by bathing the slides for 10 minutes in a gently stirring freshly made mixture of 0.1M triethanolamine, 0.25% acetic anhydride. The chromosomes were then again washed briefly in 2x SSC. The slides were then dehydrated by passing through 70% and 95% ethanol. The chromosomes were then denatured using a bath of 0.07N NaOH for 3 minutes at room temperature. The slides were once again dehydrated using 70% and 95% ethanol before being air dried.

2.19.2 Hybridisation of Chromosomal Squashes

Probes were synthesised using a standard nick translation reaction (section 2.10.4), except that reactions were performed in the presence of 1mM biotin 16dUTP (Boeringer Mannheim). A trace 32P dCTP (10μCi) label was also added to monitor the progression of the synthesis reaction. The precipitated probe from 500ng of cDNA plasmid was resuspended in 75μl of chromosome *in situ* hybridisation solution (0.6M NaCl, 50mM $NaPO_4$ pH 6.8, 5mM $MgCl_2$, 0.02% ficoll, 0.02% bovine serum albumin, 0.02% polyvinylpyrolidine).

The probe was denatured in hybridisation buffer by boiling for 3 minutes and then chilling on ice. The probe was applied to the area of the slide containing the chromosomes. About 25μl of probe was used per slide and covered with a coverslip. The slide was sealed using Cow Gum softened in diethyl ether before use. The slide was incubated for 12-18 hours at 58°C. After incubation the coverslip was removed and the slides washed in three changes of 2x SSC for 20 minutes each at 53°C.

2.19.3 Signal Detection

Detection utilised the Vectastain kit (Vector Laboratories). A detection mixture was prepared by adding one drop of grey tube A to one from tube B into 2.5mls of 1x PBS. This was allowed to stand for 30 minutes.

Slides were washed twice in 1x PBS for 5 minutes at room temperature followed by a 2 minute wash at room temperature in 1x PBS, 0.1% triton X-100. The chromosomes were rinsed in 1x PBS before being immersed in the Vectastain mix and covered with a coverslip. They were allowed to incubate for 45 minutes at 37°C. Coverslips were removed by immersion in 1x PBS and then the sections were rinsed twice in 1x PBS for 5 minutes at room temperature. The slides were then washed in 1x PBS, 0.1% Triton for 2 minutes at room temperature before being rinsed in 1x PBS.

The reaction substrate was prepared by mixing 0.5mg/ml DAB with 1% of 30% H_2O_2. Slides were placed in a tray and the chromosomes immersed in staining solution. The tray was made light tight and the slides were incubated for 1 hour at room temperature. The colour reaction was stopped by quick

immersion in distilled water and then into 1x PBS. The slides were then stained with Geimsa stain (Sigma) using a solution of 5% Geimsa stock in 10mM NaH_2P0_4. Staining was carried out for 12 minutes. The chromosomes were mounted using DPX mountant.

2.20 Site-Selected Mutagenisis

This is a procedure can be used to select for P-element insertions within or close to genes of interest (Kaiser and Goodwin 1990).

2.20.1 Preparation of Mutagenised Flies

The mutagenic cross was performed as described by Kaiser and Goodwin 1990 and flies kindly provided for screening by S. Goodwin, K. Linruth, T. Davis, P. Emery and T. Asimov (Glasgow Genetics Department). Mutagenised females were placed into cages of 100 individuals together with 20 wildtype males. Eggs were collected from the cage plates as described in section 2.4.3. These were then used to prepare DNA essentially as described in Goodwin and Kaiser 1991.

2.20.2 Screening DNA Prepared from Batches of Flies

DNA was screened by performing a PCR reaction using a P-element specific primer and a primer that is specific for the gene of interest. PCR was as described in section 2.2.2. 10-50ngs of DNA preparation was used for each reaction. The samples were then electrophoresed through a 1% agarose gel and then blotted before being probed with a gene specific probe. Cages that correspond to positive DNA mixtures were sub-divided and the screening process repeated.

2.21 Synthesis of Oligonucleotides

Oligonucleotides were synthesised using an automated DNA synthesiser as described in the machines operating manual. To free the oligonucleotide from the glass bead synthesis matrix the beads were treated with 1ml of 37% NH_4OH at room temperature. The NH_4OH was then removed from the glass beads. To the NH_4OH solution containing the oligonucleotide was added an equal volume of fresh NH_4OH, the tube sealed, and the sample incubated at 50°C overnight. The oligonucleotide was then recovered by vacuum drying. The sample was resuspended in 1x TE (10mM Tris pH 8.0, 1mM EDTA) at an approximate concentration of 0.5mg/ml.

2.22 Computer Methods

The methods described in this section are those applied to analysis of sequence data. Only the briefest description of each program/procedure is given.

2.22.1 Input of Sequence into MacVector Using IBI Gel Reader

A simple pen is used to mark a given band on the gel and it is recorded as being in the A,C,G or T track. Reading of each track was performed twice and errors checked between versions.

Sequence was input into MacVector software (IBI). Short and long runs of a particular reaction were read into the same data file by eye, or computer searching. On average, each complete sequence extension from a single primer was input and checked three times before further analysis.

2.22.2 Transferral Of Sequence Data to and from the Wisconsin GCG Package

Transfer of the Mac based file to GCG and back again was performed by saving the file in GCG from MacVector and transferring the file as text using the kermit communications protocol.

2.22.3 Motif Searching of Amino Acid and DNA Sequence

Predicted amino acid coding regions and electronic translation of DNA sequence was as described in section 2.22.9. Motif searching was performed using the program MacPattern or the GCG based program Motif. Both programs search the Prosite motifs database, although slightly different results were observed using each program. This was not due to versions of the pattern database. Transfer between GCG and Mac was performed as section 2.22.1.

2.22.4 Database Searching Using FastA and the GenEMBL Database

This search tests all possible alignments and weights the scores of each depending on the set parameters. The search employs a mathematical sifting algorithm before homing in on the set of highest matches (as described Pearson 1990). Because of this the weighting match and the extension are

critical to success and so these were varied about reasonable values as described in the GCG manual.

2.22.5 Database Searching Using TFastA and DNA Databases

Searches were as described in the GCG manual. The number of scores taken for each search was set to a high value, typically 100. This search translates the entire searched DNA database into all possible protein coding frames before matching to the search peptide, it was therefore very computationally intense.

2.22.6 Database Searching Using FastA and Protein Databases

This search uses the PAM250 residue match weighting matrix, which is used to construct a score by weighting the mismatches based upon their biochemical similarities. Scores are then sifted by the program as described in section 2.22.4.

2.22.7 Contig Assembly Using GCG Gelassemble

Contig assembly was performed by imputing sequence into this program using GelEnter, aligning using GelOverlap and assembled using GelAssemble as described in the GCG manual. The contig alignment was finally checked by eye. Any errors or discrepancies in the data were marked and reference made back to the original autoradiographs before output of the completed consensus sequence.

2.22.8 "Translation"of DNA Sequence Using "Translate"

Prediction of protein coding regions were made on the basis of homology or using Testcode (2.22.11), Codonpreference (2.22.10) or on the basis of direct details of coding region. Translation to a GCG peptide file was by the use of GCG Translate.

2.22.9 Assessing Codon Preference

GCG Codonpreference program can be used to define coding regions on the basis that a particular organism may only use a subset of the synonymous codons available. *Drosophila* codon usage tables for strong medium and weakly expressed genes were kindly provided by D. Crompton (Glasgow

Genetics Department). Codon preference was used as one basis for determining likely coding regions in unknown sequences.

2.22.10 Testcode Predictions of Coding Capacity

This program can also be used to predict coding regions based upon GCG bias at a particular triplet position. It is not always reliable; for example it fails to locate the known coding region of some *Drosophila* genes (data not shown).

2.22.11 Polypeptide Secondary Structure Predictions

Predictions of amino acid sequence secondary structure were made by the use of the GCG programs ProteinStructure, PepMap and PepPlot exactly as described in the GCG manual. Kyte and Doolittle (1982) hydropathy predictions were also made using the IBM-PC program SeqAid (D. D. Rhoads and D J Roufa 1989, Molecular Genetics Laboratory, Centre for Basic Cancer Research, Kansas State University).

2.22.12 Multiple Sequence Alignment and the Production of a Boxed Alignment

Wisconsin GCG includes a multiple alignment program called Pileup. This was used as described in the GCG manual to produce multiple alignments, together with the program ClustalV (Higgins and Sharp 1988). Output from these programs was used in the program Boxshade (Kay Hoffman, University of Cologne, Germany) to produce a shaded output.

2.22.13 Construction of Phylogenetic Trees

The output from ClustalV was used to construct phylogenetic trees based upon the neighbour joining method of Saitou and Nei 1987. These trees were tested for significance using the "boot-strapping" method using ClustalV (Higgins and Sharp 1988). Several methods of weighting the scores were used and compared to produce each phylogeny.

Chapter 3 Isolation of Differentially Expressed Genes

3.1 Introduction

Differential or +/- screening is a method that allows the isolation of genes based solely on their pattern of expression. The technique is able to detect messages that comprise as low as 0.1% of the mRNA population (Sambrook *et al.* 1989).

Differential screening involves the isolation of message populations from two cell types, the preparation of representative probes from these two populations by reverse transcription in the presence of radiolabelled nucleotide and the subsequent hybridisation of these probes to duplicate blots of a cDNA or genomic library. This process is shown as figure 3.1. Although the technique is conceptionally simple, it is limited in sensitivity. This is an inherent limitation of the technique, which is simply explained below.

Consider two probes, one constructed from 1μg of a plasmid insert and one from the same molar quantity of mRNA. If each template is used to incorporate an identical quantity of radioactive label. But for a given species within the cDNA population the incorporation of label is dependent upon its frequency within the resultant cDNA population of the probe; an mRNA present at 0.1% of total message will incorporate 0.1% of the total radioactive label used. But it is not practical to increase the amount of radioactive incorporation by one thousand fold in order to produce a cDNA probe of equivalent activity. Even if practical, the incorporated label in cDNA species other than that of interest would proportionally increase. This fraction of the probe makes no contribution to the signal, but adds to the overall background hybridisation. Hence beyond a threshold value increasing probe activity can decrease the detection sensitivity.

To maximise range of cDNAs isolated a more sensitive modification of the differential technique can be used called subtractive hybridisation. A schematic of the general protocol is shown in figure 3.2. In this experiment the aim is to remove sequences common to the two mRNA populations by hybridisation and leave a probe enriched in species of interest.

The first aim of this chapter was to evaluate the different screening options and then to select a technique that was most appropriate to the question being

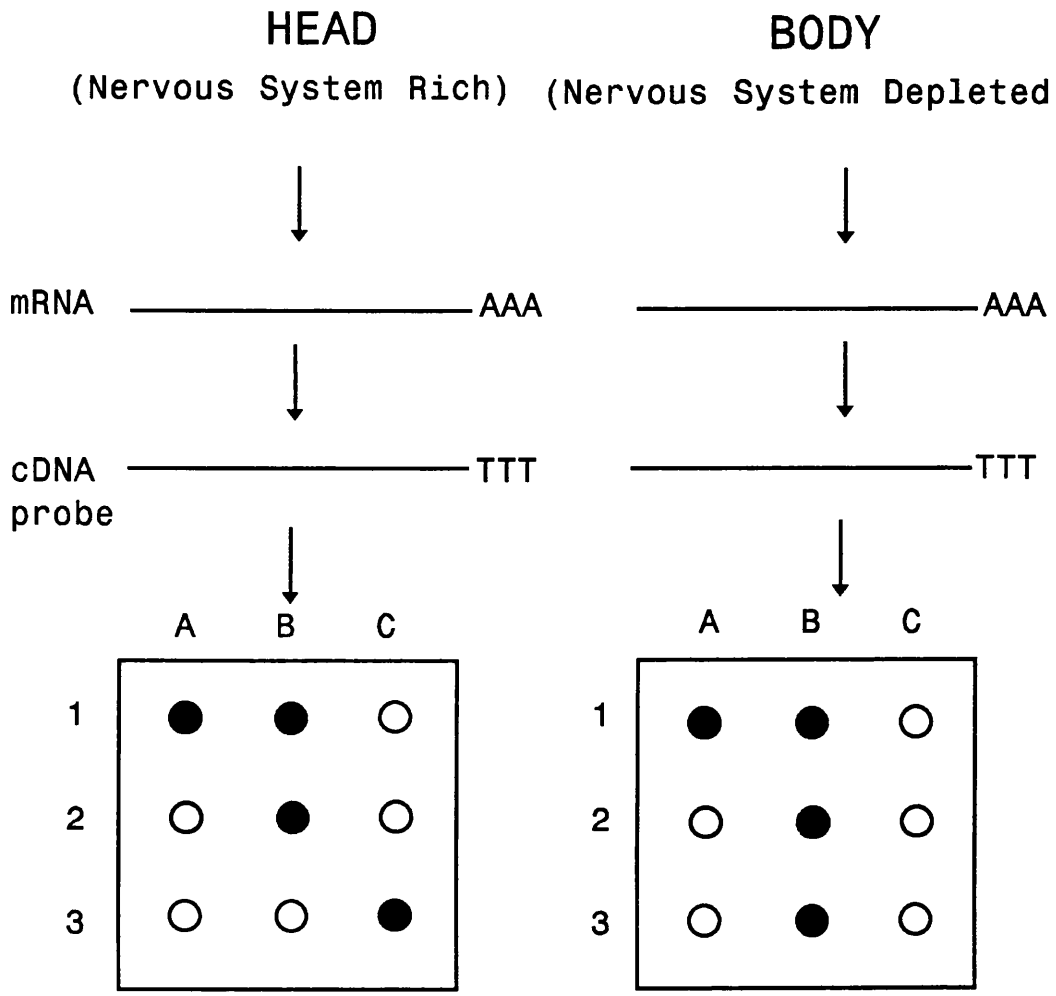


Figure 3.1 Differential Screening

The figure shows the differential screening procedure used to identify cDNAs that represent messages that were more abundently expressed in the head than the body. The grid of circles represents the grid of signals obtained after hybridisation to duplicate filters lifted from an ordered array of plaques. Black plaques are those with inserts that hybridise to the probe, clear plaques are those that do not. Some plaque DNAs hybridise to both probes e.g. **A1**, while others hybridise only to the head probe e.g. **C3**. Other plaque DNAs only hybridise to the body probe e.g. **B3**. Clones that do not hybridise with either probe may contain no insert or may represent a message that is below the detection level of differential screening.

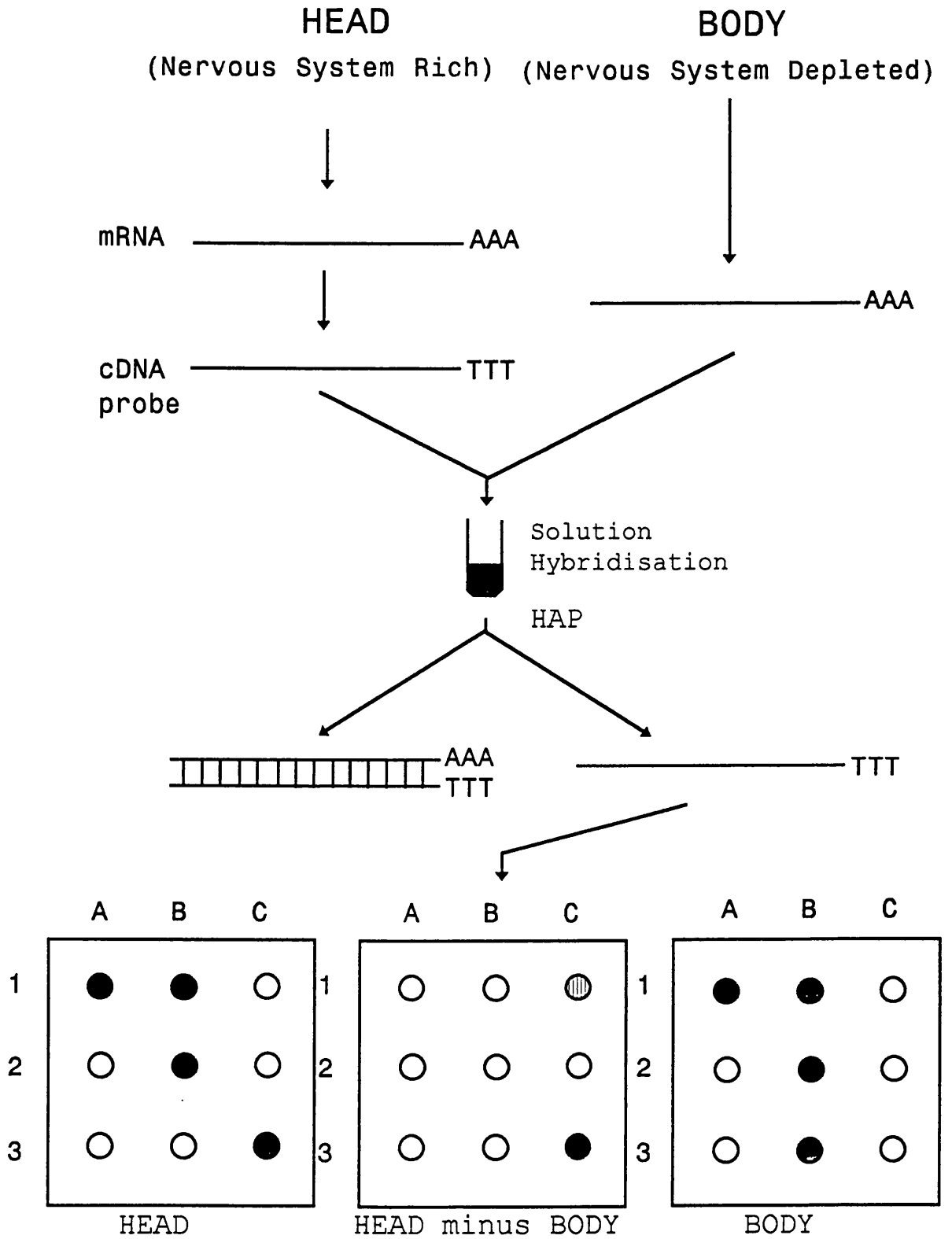


Figure 3.2 Subtracted Screening
In this procedure the head mRNA is reverse transcribed to produce a cDNA probe. This probe was then hybridised to a twenty fold excess of body driver mRNA. After hybridisation the common sequences are removed in double stranded form. The single stranded fraction, which is enriched in head specific transcripts, is then used as a probe onto an array of clone DNA. The result is that all those signals generated from plaque DNA containing non-differential inserts are eliminated. Weaker head specific signals, often difficult to identify using differential screening become visible (e.g. **C1**).

addressed. The selected technique was then used to isolate a selection of head elevated cDNAs that were likely candidates for genes with specific functions within the nervous system.

Later characterisation of these clones sought to confirm the predictions made about the transcription patterns detected using the selected cDNAs. Simple and quick methods were then used to sift the positives that had already been studied or had interesting patterns of expression or sequence homologies. At the end of this chapter, I hope it will be clear that the cloned genes have the predicted pattern of expression and why the four clones discussed in subsequent chapters were selected for further study.

3.2 An Example Subtraction and Differential Screen

In order to illustrate the technical considerations involved in performing subtractive and differential hybridisation I present here one of a set of trial experiments performed using these methods.

3.2.1 Preparation of Test Array

An array of 900 genomic clones was generated from an amplified 'lambda Gem11 (Promega) library constructed as described in methods section 2.16 and containing $5x10^6$ primary recombinants. The plating and replication are as described in methods section 2.14.1. A set of at least 12 duplicate arrays were generated using Hybond-N membrane (Amersham) and the plaque lifting protocol as described methods section 2.14.2. Test arrays were individually prehybridised in 10mls of buffer overnight (methods section 2.14.3) and then frozen at -20°C until further use. Before use the membranes were defrosted and again prehybridised. The same test array was used for all the experiments described in this section.

3.2.2 Preparation of Head and Body mRNA

Head and body RNA was prepared using the guanidine thiocyanate and caesium chloride method (as described methods section 2.7.1). Over 10mgs of body total and 1mg of head total RNA were produced for these experiments. Contamination of head material with bodies was undetectable by visual inspection and of bodies with head at less than 1 in 3000 by mass. Concentration of both RNAs was assessed by non-denaturing agarose gel electrophoresis against a concentration standard (a well characterised body

preparation). Concentration was also assessed using OD_{260} readings (methods section 2.2.8).

PolyA+ RNA was selected by two passes over oligo dT cellulose exactly as in methods section 2.8. After resuspension about 1μg $polyA^+$ message was electophoresed through a denaturing formaldehyde gel, blotted and probed with a control probe generated from α1 tubulin (Kalfayan and Wensink 1982). Both preparations gave a single sharp band suggesting that the mRNAs were undegraded (data not shown). Final assessment of concentration was made from this gel and a second non-denaturing gel. RNA was aliquoted and stored at -20°C prior to use.

3.2.3 A Test Differential Screen.

Probes were prepared by reverse transcription as described methods section 2.14.3. After synthesis an aliquot of each probe was scintillation counted (methods section 2.2.15) and the overall incorporation of radionucleotide into the probe calculated. A small sample of each probe was saved and ran out on a sodium hydroxide denaturing gel. The activity in each probe was adjusted and the probes then added to the hybridisation. Hybridisation of these experiments was in 5mls per bag overnight. The filters were washed and autoradiographed as described in methods section 2.14.4. The resulting autoradiograph from an overnight exposure is shown as figure 3.3.

3.2.4 A Test Subtractive Hybridisation

The protocol is as described in figure 3.2 and in methods section 2.15. 20μg of body mRNA driver was used as driver, the probe was synthesised using 200μCi ^{32}P dCTP which, when assayed by scintillation counting gave a probe with total incorporated activity of $1x10^9$dpm. The hybridisation reaction was performed in 10μl total volume overlaid with an excess volume of oil. The tube was then placed inside a glass vial which contained a non-covering volume of water and incubated without shaking for 24 hours at 68°C.

The incubation mixture, which should now contain both double and single stranded material was subject to HAP chromatography as described in methods section 2.15.2. Each radioactive sample from the column was scintillation counted and the activity profile is shown as figure 3.4. In the

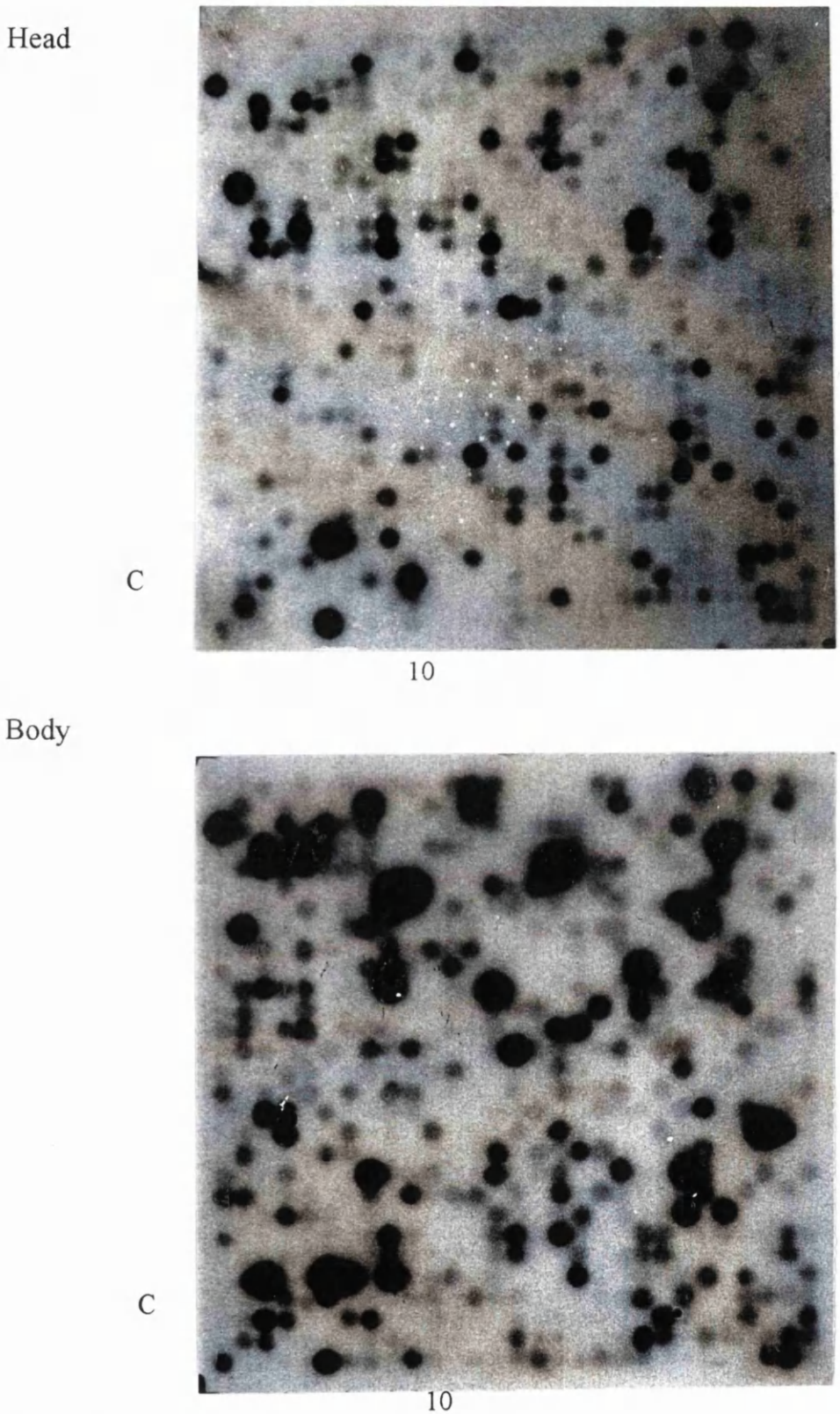


Figure 3.3 A Test Differential Screen

Differential screening was performed as described in methods section 2.14. Probes were generated with identical incorporated activity from head and body polyA+ mRNA and used to hybridise duplicate blots of an array of 1000 genomic lambda clones. The filters were identically washed to a stringency of 0.2x SSC at 68°C. The reproduced autoradiograph was generated after 24 hours exposure using intensifying screens. The clone labelled **C10** shows much stronger hybridisation to the head probe than to the body.

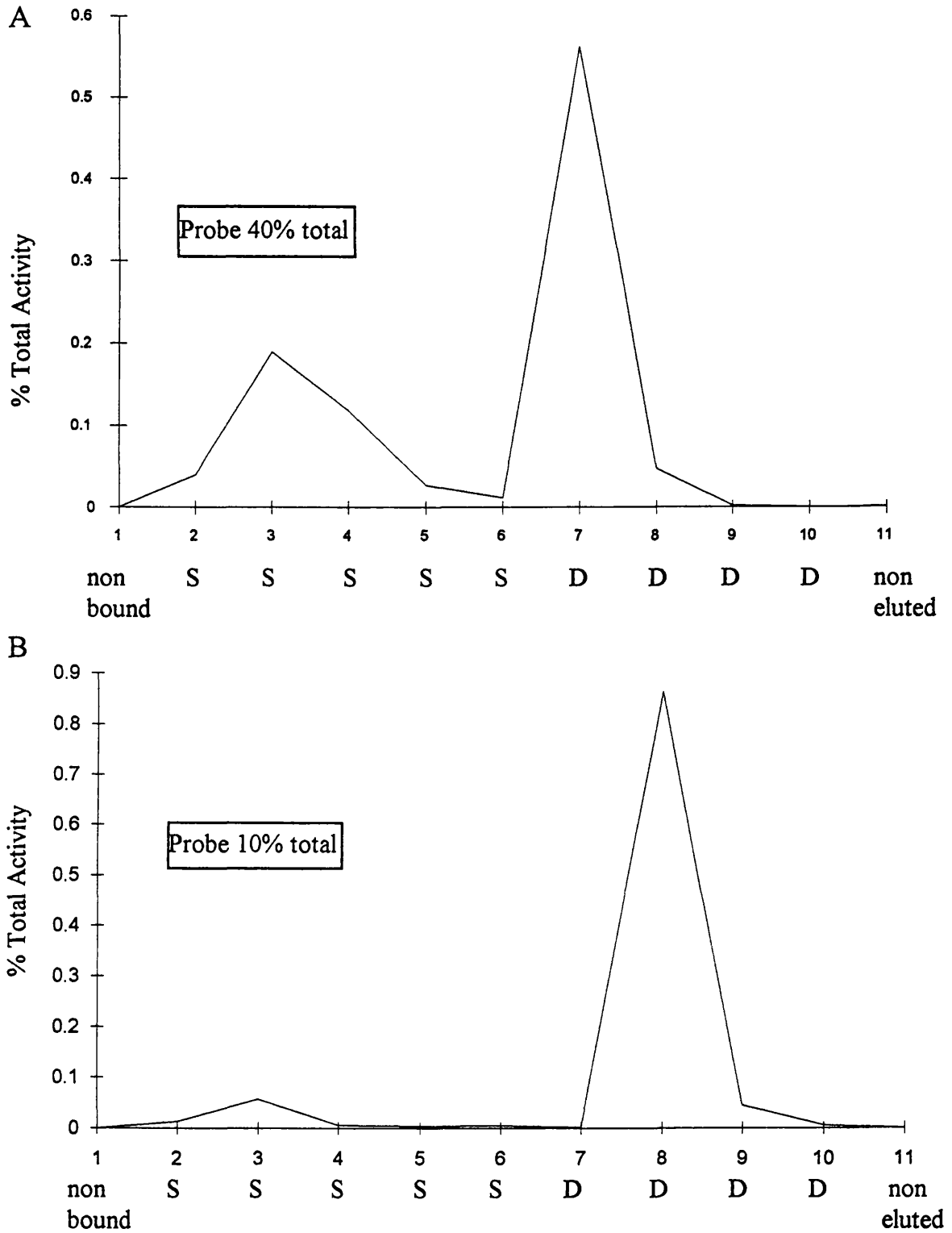


Figure 3.4 Activity Profile for Test Subtraction

These two profiles are from a head cDNA probe subtracted with an excess of body mRNA (**A**) and a body cDNA probe subtracted with a body mRNA driver (**B**) as described methods section 2.15. Activity is scored as a percentage of total incorporated activity (methods section 2.2.15). Total probe incorporated activities were $1x10^8$dpm for the body probe and $6x10^7$dpm for the head probe. 70% of this activity and 50% of this activity can be accounted for in the body and head probes respectively after hydroxylapatite chromatography. Unfortunately, after probe processing, and analysis only 10% of the activity was eventually used to hybridise to the array. Eluted fractions are numbered 1 to 11.

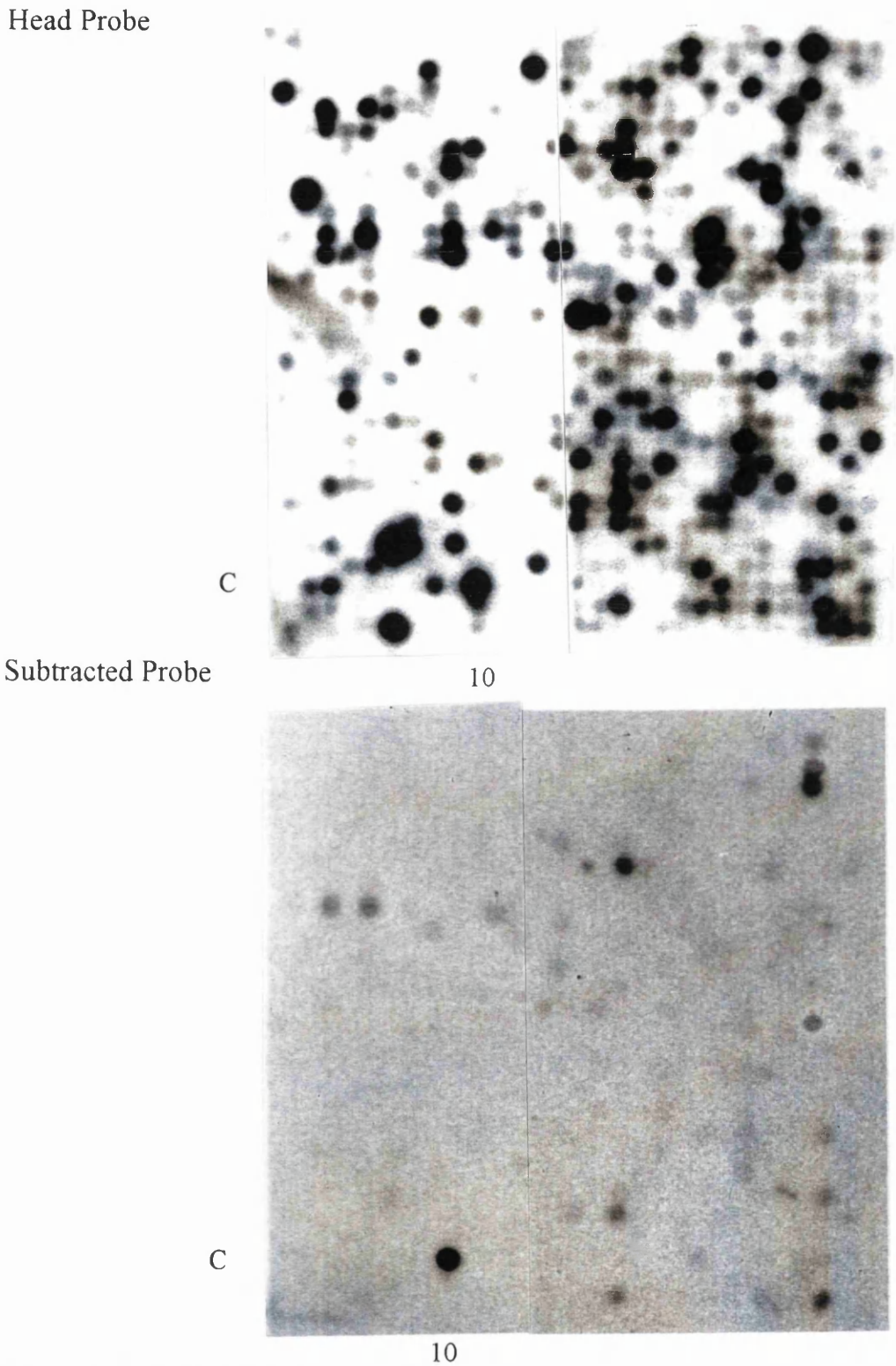


Figure 3.5 A Test Subtractive Screen

A head minus body subtracted screening was performed as described in methods section 2.15 on an array of 1000 genomic lambda clones. The filter used are duplicates of those shown in figure 3.3. The filters were washed in either 0.2x SSC at 68°C (left) or 0.1x SSC at 68°C (right). A duplicate filter, probed with head polyA+ RNA and washed at the same two stringencies, is shown for comparison. The body minus body subtracted probe, generated as described in section 3.2.3 was also used as a hybridisation probe but gave no detectable signal (data not shown). Several show strong hybridisation to the subtracted probe including a strong positive at grid position **C10** which was detectable using differential screening.

profile two large activity peaks are shown corresponding to the single and double stranded nucleic acid. 70% of the single stranded cDNA has been converted into double stranded form.

The single stranded fraction is now a probe rich in head specific cDNA species. This probe must now be treated to remove the phosphate buffer, this is performed using G50 fractionation as described methods section 2.2.18. The probe was then precipitated using ethanol, resuspended in 200µl of water, and added to a prepared hybridisation bag. Hybridisation, filter washing and autoradiography are exactly as described for the differential screen part of this experiment. Figure 3.5 shows several exposures of the resulting autoradiograph together with the differential screen results for comparison.

3.2.4 Test Screen Conclusions

From the results of this experiment it appears that both forms of screening have proved highly successful. However, there are a number of areas within the subtraction experiment that were cause for concern.

Consider the comparison between the differential and subtraction results by examining figure 3.5. Clearly both methods identify several clones as being very likely candidate head elevated or head specific clones. A longer exposure of the subtracted filter further emphasises that there may exist a whole set of rarer cDNAs not easily identifiable using differential screening. These clones would be good candidates for rare brain specific genes.

If one examines the relative intensity of the abundant differential clones one can see that a considerable loss in signal has occurred during the subtraction procedure. Although reasonable for this experiment, it is often necessary to perform a second subtraction upon the products of the first hybridisation. This is usually because the hybridisation has not reached sufficient R_0t during its first 24 hours. Losses from this second round become critical to the success of the experiment.

Experimental loss of activity is a well documented feature of the subtraction procedure (e.g. Sambrook *et al.* 1989, Berger and Kimmel 1987). Single stranded DNA tends to be very "sticky" because of the exposed charged phosphate backbone, and so elaborate procedures need to be employed in order to minimise losses. Even so, as much as 5% of activity can be lost per

pipetting or precipitation (data not shown). Large losses are observed during the resuspension of the solution hybridisation mixture. The phosphate used tends to make the pellet viscous and so is added last. Therefore 20μg of mRNA and 200μCi of probe are resuspended in 5μl of water. Loss at this stage alone is 30%. But complete resuspension is very important, small areas of poorly resuspended probe that later resuspend during hybridisation will not reach the required R_0t and contaminate the subtracted probe with non-subtracted material. Spectacular losses were also observed during the desalting and precipitation of the subtracted probe prior to hybridisation or a second round of subtraction

In this set of experiments, great problems were observed with the hybridisation membrane. It is clear that the hybidisation membrane supplied to the laboratory shows batch variability. In this experiment problems with a current batch of membrane lead to reference to an older frozen set of filters. These were used for this experiment. Although it was no longer possible to refer back to the original culture plates, this set of hybridisations was used to test as many of the different conditions as possible.

The problem areas of the subtraction method have been overcome in a number of laboratories using a range of modifications such as the use of different hybridisation protocols such as PERT (Kohne *et al.* 1977), the use of different single stranded selection protocols e.g. using biotin (Sive and John 1988) or by the use of a conventional cloning (e.g. Palazzolo *et al.* 1989) or by PCR (e.g. Sturzl and Roth 1990). However, these methods each have technical limitations.

Progress towards developing a workable subtraction procedure in the laboratory has proved very slow. As a pilot screen it was decided to use differential screening to isolate a small set of head elevated clones. This was to look in more detail at some of the clones likely to be isolated by the larger screen. The pilot study proved useful in the laboratory, and a second screen was performed by C. Millagan (Glasgow Genetics Department) using a procedure optimised from my own screen.

3.3 Choosing the Differential Screen

A number of differential screens have been performed previously. Of these the Levy *et al.* 1982 screen most closely resembles the method performed

here. The distinctions between the two screening procedures were designed to enhance the sensitivity of the screening method.

Many of Levy *et al.* 1982 head elevated clones were from the eye. Several eye specific clones have now been identified and studied including a number of opsins (as discussed in chapter 1). To avoid re-isolating these clones a strain of *Drosophila* was used called *eya* which lack the eye structures and most of the parts of the brain involved in visual processing (Sved 1986, Bonini *et al.* 1993). At first an OrR, wildtype head library was screened with OrR head and body cDNA probes, the more difficult to prepare *eya* head probe was to be used as a second screen. During this differential screen an *eya* head library became available and so this was used in preference for later cycles of the screen.

Levy *et al.* 1982 screened a genomic library. Chance of a given sequence being represented in this library depends only upon the size of genomic insert and copy number of the gene. Chance or representation in a cDNA library is dependent on the abundance of the template mRNA. Thus a few common cDNAs, perhaps non-differentially expressed, can make up a significant proportion of the signal on a cDNA library blot. These can cause significant problems by obscuring the signal from weaker hybridised clones. By using an ordered array, prepared using a device developed in the laboratory (Mackensie *et al.* 1989), problems of obscuring become less critical with the cDNA library. Since the screen is neither exhaustive nor searching for rare genes the representation advantages of the genomic library are no longer germane. With a set of cDNA clones one can quickly progress in the downstream analysis and so the cDNA libraries were chosen for screening.

3.4 The Differential Screen

The mRNA used to synthesise the probes was generated from *Drosophila* head and body OrR material as described in methods section 2.14.3. Care was taken to match the activity in each probe in order that the hybridisations be quantitative. The only other modification to the Levy *et al.* 1982 procedure was that the probes after synthesis were passed over hydroxlapatite, a treatment that seems to reduce the background hybridisation level under certain circumstances (unpublished observations, Kaiser laboratory, Glasgow Genetics Department).

In order to determine that a clone was likely to be differential several duplicate hybridisations were performed and autoradiographs taken after exposing the film over several time durations. Candidate clones of genes that seemed more abundantly expressed in the head than body were selected and picked directly from the array. These clones were then each streaked out onto plates and single plaques chosen. This is to prevent confusion caused by mixed plaques. The rescreen array of these plaques and a number of clones that didn't seem to hybridise with either probe were replated in another array. This array was interspersed with negative clones that were made from vector 'phage that lacked an insert. The array was rescreened with head and body probes and used to select candidate differential cDNAs (figure 3.6).

Selection of the particular positive was based upon the assessment of duplicate blots and multiple exposures. Positive clones that had originated from the first array were from an OrR head library constructed by S.Russell (Glasgow Genetics Department) in the vector NM1149 (please refer figure 3.6). Some of the positives from this section of the screen were very strongly differential. These were likely candidates to be opsins or one of the other clones found by Levy *et al.* 1982. These were of little long term interest to the laboratory, but one, designated λST41, was selected as representative of the class. No other clone from this library was selected for further analysis. Not all the differential candidates from this library will be from the eye. The other positive clones from this library remain to be assessed; they have not been discarded.

In total twenty, clearly differential clones were selected for further study. λ ST41 from the NM1149 library and the remainder of the clones selected were in the ZAPII 'phagemid vector from Stratagene (Short *et al.* 1988). The library used was constructed in the laboratory by R.C Mackensie from pure head material harvested from *eya Drosophila melanogaster.* [The library is known in the laboratory as Ron's head first packaging].

Bacteriophage vectors are more difficult to work with than plasmid vectors. ZAP II has an in built mechanism that allows the insert to be excised from the 'phage vector as single stranded 'phagemid. In a host *E.coli* cell this 'phagemid is converted to a double stranded plasmid. In effect the directionally cloned ZAPII inserts can be converted to directionally cloned

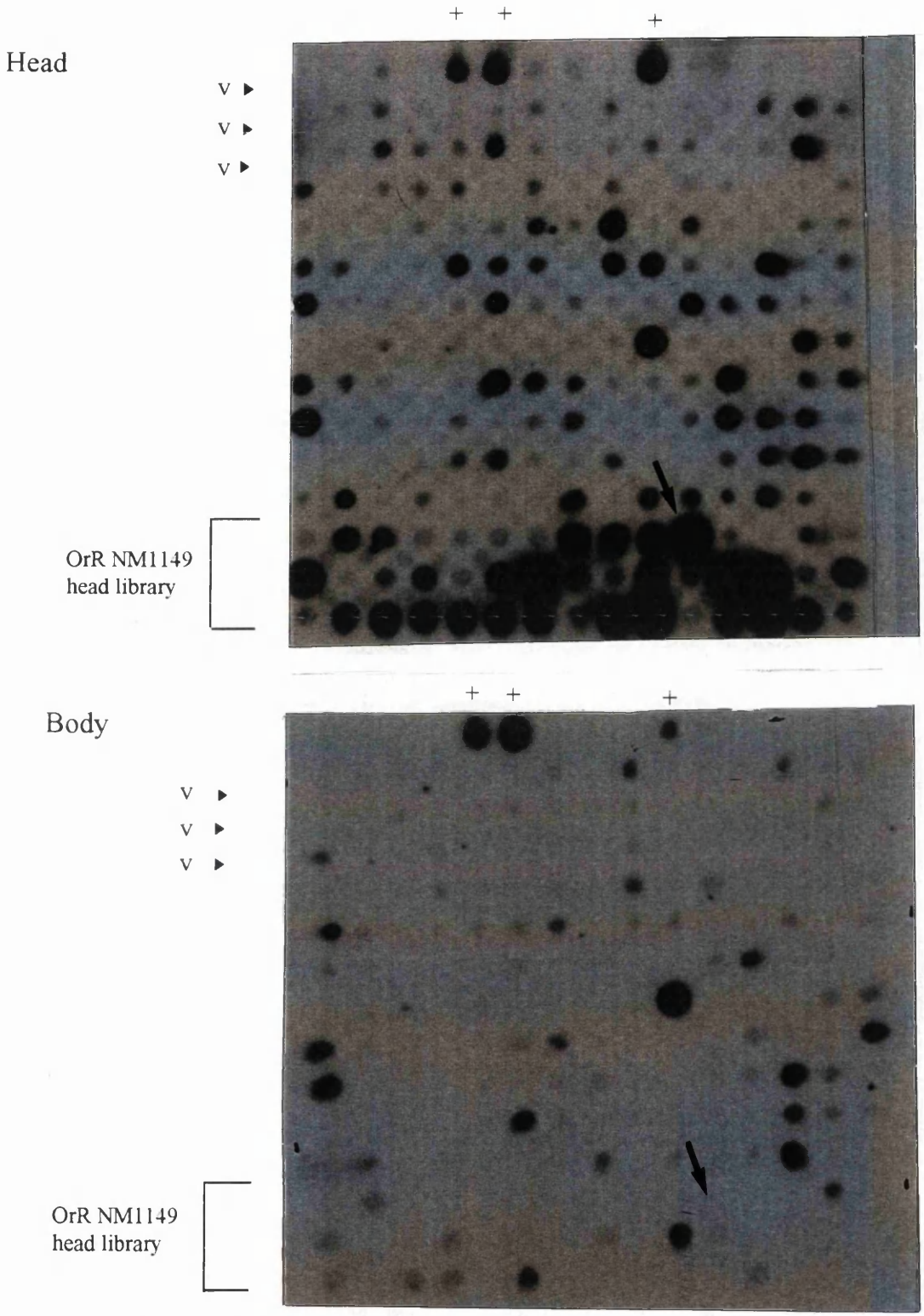


Figure 3.6 The Differential Screen

Screening was performed as described in methods section 2.14 and section 3.4. The potential head specific clones from the first screen (not shown) were replated again probed with head and body derived cDNA probes of equal activity. The two reproduced autoradiographs were generated from an array of these positives interspersed with rows of the non-recombinant 'phage vector (**v**). Positive, non-differential controls(+) from the primary screen were also plated. The arrow indicates a strongly differential clone; in this case the clone is λST41. Clones isolated from the NM1149 library are indicated, all other clones are from the ZAPII *eya* head library.

plasmid inserts *en mass*. This procedure was performed as described in methods section 2.3.6. Single excised plaques were grown through several plating cycles and then used to produce plasmid maxipreps (methods section 2.2.4.2) of the candidate clones.

λST41 was plated and a single plaque selected. This was used to seed a 'phage maxi-prep as described in methods section 2.2.20. Inserts may be released from NM1149 by double digestion with *Eco*RI and *Hind*III. Single and double digestion with these enzymes indicated that the clone possessed only one of each site that liberate upon double digest an insert of 1.6kb (data not shown). This insert was used first produced from the 'phage vector until it was subcloned into pBluescript II SK$^-$.

Subcloning of the insert was performed by digesting the vector and 'phage with both enzymes and then ligating a mixture of the two reactions with a molar equivalent ratio of vector and insert. This reaction was performed under standard ligation conditions (methods section 2.2.1). The products of the ligation were selected using blue/white selection in XL1B cells (methods section 2.3.5). Two white plasmids were selected, screened to purity, and checked to contain a cross hybridising insert of the same size as λ41. These two clones were designated pST41-1 and pST41-2. Both clones appear identical from restriction mapping (not shown). Only pST41-1 was used in later experiments.

3.5 Restriction Analysis of Positive Clones

Plasmid clones from the screen were first digested with *Eco*RI and *Xho*I as single and double digests and electrophoresed through a 1% agarose gel together with each uncut plasmid. All except one plasmid, pST157, revealed the presence of inserts within the plasmids. Several of the plasmids did not digest with *Eco*RI, although all cut at least once with *Xho*I. Several internal *Xho*I sites were found in clones. This was an important observation, as C Mackensie (Glasgow Genetics Department) who constructed the library, has evidence to suggest that the methylation used to protect against restriction of internal *Xho*I sites was not fully functional. A pair of restriction enzyme cleavage sites located further into the vector and flanking the polylinker were used in subsequent experiments to free the insert from the vector. In the reverse northern experiments presented in the next section the enzyme used was *Pvu*II, which also cuts at two vector sites flanking the polylinker. All

plasmids, except pST157, release an insert after digestion with this enzyme. The failure of some plasmids to digest with *Eco*RI is supported by junction point sequence data discussed in section 3.7.

3.6 Reverse Northerns

In this procedure, each of the plasmids are digested with *Ava*II to release the insert and the products electrophoresed through an agarose gel. The result is blotted (methods section 2.6) and then hybridised with cDNA probes produced by reverse transcription of head and body mRNA, (as described in methods section 2.14). In this experiment insert from λST41 and from a non-differential control α1 tubulin (Kalfayan and Wensink 1982) was also used. α1 tubulin was supplied as a kind gift from S. Russell (Glasgow Genetics Department) and the insert was released from the plasmid vector by digestion with *Eco*RI and *Bam*HI. Figure 3.7a and figure 3.7b shows the agarose gel and resultant autoradiography of the blotted gel.

The filters were then stripped of probe by washing in 0.01% SDS at 80°C for 30 minutes and checked to be clean of signal by autoradiography (methods section 2.2.19). They were then reprobed with cDNA probes prepared from mid-pupal and mixed embryo mRNA (methods sections 2.14.3). They were treated as for the head and body probe. The resulting autoradiograph is shown as figure 3.7c and 3.7d. All the blots have been treated with equivalent probes of matched specific activity.

3.6.1 Conclusions from Reverse Northerns

By comparison of the control to intensities of hybridisation to each other it is clear that nearly all the clones represent genes that are more abundantly transcribed in the head than in the body. The exception to this is pST78 which seems to be found in all tissues.

A developmental profile can be constructed using the reverse northern data. This is shown as figure 3.8. No clone seems to be more abundantly expressed in other tissues than it is in the head. If one considers what one expects from a nervous system specific gene it is that it is expressed more in head than in bodies and less so in other tissues. Head material contains proportionally more neuronal tissue than body tissue and than tissue from pupal and embryo. Of course this is a generalisation, but if a gene was more expressed in say embryos than in head then this suggests it has a specific role at that stage.

A Head cDNA Probe

B Body cDNA Probe

C Mixed Embryo cDNA Probe

D Mid-Pupal cDNA Probe

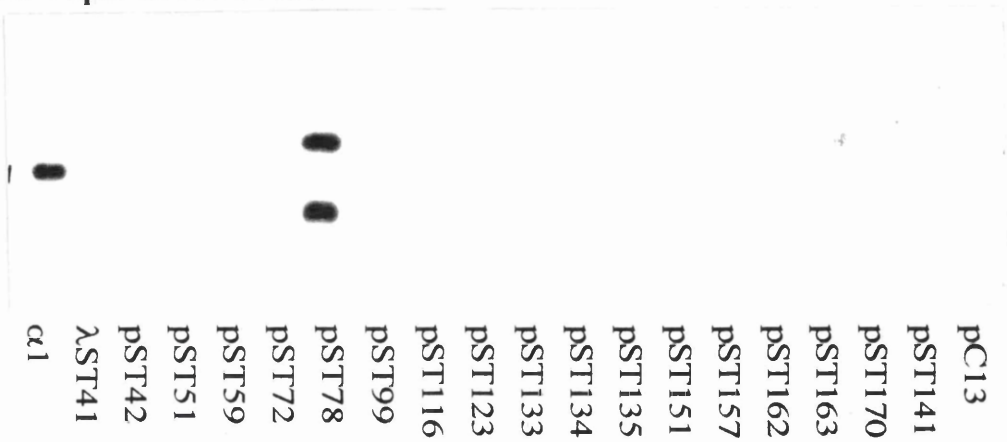


Figure 3.7 Reverse Northern Experiments

Reverse northerns were performed as described in methods section 2.14 and 2.6. Each plasmid has been restricted with *Ava*II, except λST41 which was cleaved with *Eco*RI and *Hind*III to release the insert. Southern blot filters were probed with head (**A**), body (**B**), mixed embryo (**C**) and mid-pupal (**D**) cDNA probes of equivelant activity. Several durations of autoradiography were used for probes of equal incorporated activity and representative exposure intensities are shown. α1 tubulin is also present on the blots as a positive control.

Length of insert in Kb	Clone	Head	Body	Mid Pupal	Mixed Embryo	Comment
	α1Tubulin	*100*	*100*	*100*	*100*	Control
1.6	λST41	1000	ND	ND	ND	
0.8	pST42	100	ND	20	10	
1.4	pST51	100	ND	10	1	Group A
0.5	pST59	500	20	10	ND	Group B
0.7	pST72	500	50	10	ND	
1.4	pST78	1000	1000	1000	1000	
1.4	pST99	100	ND	10	ND	Group A
1.1	pST116	100	10	10	10	
2.3	pST123	100	ND	10	10	
0.6	pST133	1000	20	10	ND	Group B
0.4	pST134	10	ND	10	ND	
2.6	pST135	100	50(LB)	10(ALL)	0.1	Group C
2.7	pST151	100	ND	10	0.1	Group C
?	pST157	ND	ND	ND	ND	
2.1	pST162	100	ND	10	1	
0.6	pST163	100	ND	10	10	
0.55	pST170	500	ND	10	ND	Group B
3.4	pST141	500	100	10	10	
1.6?	pC13	1000	50	20	ND	

Figure 3.8 Predicted Expression Patterns Derived from Reverse Northerns
Level of expression has been tabulated by examining a range of exposures of the "reverse northern" blots and normalising expression levels against that of α1 tubulin which is given a nominal expression score of 100. The loading of DNA in each track was approximately constant with the exception of the insert from λST41 which is considerably underloaded when compared to the plasmid derived cDNAs. Clones that have been found to be related are labelled as group A to C. Grouping of these clones is based upon partial sequence data and upon restriction map data (discussed in later sections). Clones that are related show similar expression patterns. Slight differences between the expression of grouped clones may be biologically significant. ND stands for non-detected, LB for lower band.

Reverse northerns do not assess the range or stage specificity of different transcripts, what is assessed here are all transcripts that hybridise to a particular DNA fragment. It is therefore possible for pST78, which gives a strong signal with both head and body probes, to be composed of two different sized, equally abundant transcripts one head specific and one body specific. The differential screen and reverse northerns would score the clone as non-differential and it should slip through the screening net.

Reverse northerns are not as sensitive a technique as northern blotting. The technique is hampered because of the nature of the cDNA probe (see section 3.1). It is reasonably expected to be more sensitive than screening a plaque array both because much more target DNA can be placed on the filter and because background hybridisation to the vector that often occurs (data not shown) is more easily accounted for. The technique, like any detection system has a sensitivity window; those molecules above and below the limit of detection are scored equally as 100% and 0% abundance respectively.

3.7 Sequencing Extensions From 5' and 3' Ends of Each cDNA

Sequencing was performed on each clone in order to determine whether the clone was related to other clones in the set, related to previously sequenced *Drosophila* genes or to other sequences in sequence database. The technique also gave valuable information about the 5' and 3' cloning termini.

Each cDNA clone was sequenced from vector based T3 and T7 primers as described in methods section 2.13. Each sequence was assessed to determine the cloning junction points and then the insert sequence input into GCG sequence analysis package as described methods section 2.22. Junction points are shown tabulated as figure 3.9. All 3' junction points for which sequence data is available, are as expected, although the 5' junction points show several alterations about the *Eco*R1 site. The absence of the *Eco*RI site is predicted from restriction data in these clones as shown.

3.8 Database Searching with Partial Sequence

DNA searches were performed through the full GenEMBL database. Putative peptides in three forward frames were also constructed and used to search peptide databases and GenEMBL using the TFastA searching method (methods section 2.22.5).

A

Clone	5' Sequence Junction Point	EcoRI sites
expected	.CCCCCGGGCTGCAGGAATTCGGCACGAG..3'	0
pST42	.CCCCCGGGCTGCAGGAATTCGGCACGAG..	1
pST51	.CCCCCGGGCTGCAGGAATTCGGCACGAG..	1
pST59	GAATTCGGCACGAG..	2
pST72	...CCCGGGCTGCAG**GAAT-C**GGCACGAG..	1
pST78	......GGGCTGCAGGAATTCGGCACGAG..	1
pST99	.CCCCCGGGCTGCAG**GA-TTC**GGCACGAG..	0
pST116	.CCCCCGGGCTGCAGGAATTCGGCACGAG..	1
pST123	..........**GCAGACGCAGGCACAGGT**..	0
pST133	.CCCCCGGGCTGCAGGAATTCGGCACGAG..	2
pST134	..........TGCAGGAATTCGGCACGAG..	2
pST135	.CCCCCGGGCTGCAGG**CTCGAG**[PolyA tail]..	0
pST151	.CCCCCGGGCTGCAGG**CTCGAG**[PolyA tail]..	0
pST157	This clone produced no readable sequence	?
pST162	..................CGGCACGAG..	1
pST163	.CCCCCGGNCTGCAGGAATTCNNCACGAG..	2
pST170	.CCCCCGGGCTGCAG**GAAAATCATCATCG**..	1
pST141	............CAGGAATTCGGCACGAG..	1

B

Clone	3' Sequence Junction Point	XhoI sites
expected	....$(A)_{XX}$CTCGAGCCCCCCGGG....	
pST42		1
pST51	....$(A)_{16}$CTCGAGCCCCCCGGG....	1
pST59	....$(A)_{15}$CTCGAGCCCCCC....	1
pST72	....$(A)_{18}$CTCGAGC....	1
pST78		1
pST99	....$(A)_{17}$CTCGAGCCCCCCG....	1
pST116	....$(A)_{16}$CTCGAGCCCCCCGGG....	1
pST123	....$(A)_{16}$CTCGAGCCCCCC....	2
pST133		
pST134	....$(A)_{17}$CTCGA....	2
pST135	....$(A)_{21}$C...	2
pST135R	....$(A)_{20}$CTCGAG....	2
pST151	....$(A)_{21}$CTCGAGCCCCCCGGG....	2
pST151R	....$(A)_{21}$CTCGAG....	2
pST157	This clone produced no readable sequence	1
pST162	....$(A)_{18}$CTCGA....	1
pST163		1
pST170	....$(A)_{23}$CTCGAGCCCCCC....	1
pST141		1

Figure 3.9 5' and 3' Cloning Junctions

In each table the expected junction points are shown for comparison.

A shows the 5' junction points, **N** indicates difficult to read residues, dashes indicate deletions in the junction point. Bold sequences are those that lack a 5' *Eco*RI site and the number of *Eco*RI sites found by restriction mapping is shown for comparison. Predicted *Eco*RI sites are shown underlined.

B shows the 3' junction points. Sequences are underlined at the predicted *Xho*I site. The number of *Xho*I sites from restriction mapping are also shown. Several clones have not been sequenced across their 3' junction points. Sequences labelled pST151R and pST135R are from clones which appear to have 3' sequences at both termini.

3.8.1 Matches to Other Clones

Sequence data was used to group the clones. This is only a limited grouping as clones that differ in their flanks but not in there central sequence would not be grouped together. C. Milligan (Glasgow Genetics Department) in a parallel screen used cross hybridisation as well as sequence data to group clones. No further homologies were added to those found using sequence data alone. After my own sequence comparison, three sets of related clones were identified. pST135 and pST151 group together, pST51 and pST99 are related as are pST170, pST133 and pST59. Interestingly, C. Millagan also has two clones that are homologous to the pST170 group called pC12 and pC13.

The pST170 and pST51 groups can be detected to differ about their 5' termini. Both these groups have been examined in more detail in later chapters (chapter 6 and 5 respectively). The grouping of pST135 and pST151 is interesting because the clone seems to be the product of a head to head double insert. [Both pST135 and pST151 lack an internal *Eco*RI site that would be expected to mark the junction between the two inserts. The cloning event may have been more complex than first appeared]. Such an event is expected to be rare in the library and so two clones with exactly the same configuration should only arise from a single event that has been duplicated during amplification of the cDNA library. The single base pair variation in the 5' cloned polyA tail length is rather at odds with this, although it could arise due to a local compression on the sequencing gel.

Reference back to the reverse northern data (figure 3.8) shows that the predicted expression patterns for each members of the group are in close correspondence to each other. The most significant departure is that pST135 gives a score in the body probe that is not detected with pST151. This may be faint background hybridisation, but it may be that pST135 and pST151 are more dissimilar than is apparent from the restriction and partial sequence data available.

3.8.2 No Detectable Similarity to GenEMBL Sequences

DNA searches are not sensitive enough to detect anything other than very highly conserved genes. Most conservation is within protein coding regions, but as there can be as many as six codons for a given AA then conservation becomes difficult to detect at the DNA level. To increase the sensitivity the

DNA sequence can be translated, but this must be performed in three forward frames and is prone to frame-shift errors within the sequence. This is an inherent problem with sequence data. Single stranded sequence submitted by the human genome sequencing project has a reported error rate of nearly one percent (Adams *et al.* 1991). As most of my own sequencing errors seem to be missing bases (as determined by those cDNAs that have been fully sequenced) it seems likely that even if all the DNA sequence corresponds to peptide sequence, this peptide will be split across a number of frames. It becomes very difficult to detect weak homologies in this way. These problems are compounded if, for example, the cDNA contains a long 5' and 3' untranslated region.

However, imperfect sequence data cannot fully account for the negative matches of many of the cDNAs. Many of the sequence runs will fall into poorly conserved leader sequences that are of little use in tagging a cDNA as being derived from a member of a particular gene family. Indeed, sequence error rates as high as 1% deletions and 5% substitutions can still allow most homologies to be identified (States and Botstein 1991). Certainly identity to known *Drosophila* genes would not be clouded by imperfect sequence.

3.8.3 pST41 is a cDNA clone derived from the *Drosophila ninaE* Gene

Sequence was obtained from pST41 from the 5' and 3' ends of the clone and put through the GenEMBL database as a DNA search using FastA (methods section 2.22.4). Figure 3.10a shows the alignment of the published sequence of the opsin and the 5' sequence from pST41. Clearly, pST41 is a cDNA clone of a *ninaE* transcript. The isolation of an opsin from the male head OrR cDNA library, together with it being the most abundant clone I have isolated, supports the use of the *eya* strain of fly to isolate genes from the fly brain not specific to the eye.

NinaE was originally cloned by homology to the bovine sequence (Zucker *et al.* 1985) and by genome walking from the genetically defined locus (O'Tousa *et al.* 1985). It has also been found as one of the head specific positives from both the Levy *et al.* 1982 and the Palazzolo *et al.* 1989 screen. The hybridising transcript is found abundantly in heads and very weakly in late pupae (reproduced as figure 3.10b). It is not expressed at other stages. From the reverse northern analysis it can be seen that this expression pattern has been correctly predicted by the analysis.

A

```
                        10        20        30        40        50
ST41              AACGACCAATCGCCGCGACTAGTCCGCCCCAGTGAAATATTCAGAATCCA
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
ninae   GCAGGTTTCCAACGACCAATCGCCGCGACTAGTCCGCCCCAGTGAAATATTCAGAATCCA
                10        20        30        40        50        60

                60        70        80        90       100       110
ST41    GGAACCCTTTATGTAAAAAGTGTTAGAAATATTGTTAGTGAATTTGCAGCTTTTTATGTA
        || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ninae   GG-ACCCTTTATGTAAAAAGTGTTAGAAATATTGTTAGTGAATTTGCAGCTTTTTATGTA
                 70        80        90       100       110

               120       130       140       150       160       170
ST41    GACAGTGTGATATAGGCGGGATATAGTGACGCAGCCAGTAACCAAAACACAATGGAGAGC
        |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
nina    GACAGTGTGATATAGGCGGGAAATAGTGACGCAGCCAGTAACCAAAACACAATGGAGAGG
                120       130       140       150       160       170
```

B

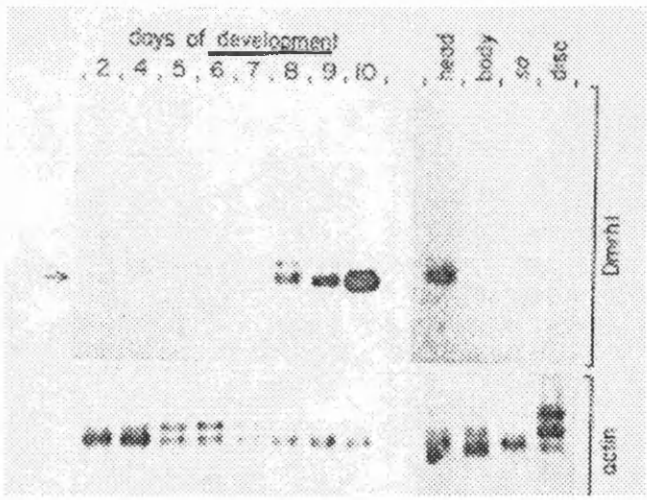


Figure 3.10 pST41 is a cDNA clone of an Opsin Transcript
A shows the pST41 sequence derived from extension from the T3 primer (**ST41**) aligned to the published sequence of Zucker *et al.* 1985 (**ninae**) using FastA as described in methods section 2.22.4. Numbering is in base pairs from the 5' end of the cDNA.
B shows the developmental expression of the opsin reproduced from Zucker *et al.* 1985. The arrow shows the migration of the 1.5kb transcript. An actin probe was used as a control for loading as shown. The gene is expressed specifically in the head, as opposed to the body. The mid-pupal RNA used in the reverse northern experiments was generated over 24hours between 7.5 and 8.5 days. The weak expression expected from the northern was not detected on the reverse northern (a less sensitive technique).

3.8.4 Identification of pST123 as Representative of a Novel Synaptobrevin Gene Family Member

Sequence was obtained from pST123 from the 5' and 3' ends of the clone and used to search through the GenEMBL database as a DNA search using FastA (methods section 2.22.4). Homology was detected to a family of molecules expressed abundantly in the nervous system as part of the synaptic transmission system. A synaptobrevin locus had been previously identified from *Drosophila* by homology to the bovine sequence (Sudhoff *et al.* 1989). However, the alignment shown to this sequence and the pST123 cDNA indicates that this is not the previously studied cDNA (figure 3.11). The predicted reading frame of the published *Drosophila* sequence can be overlaid over the 5' pST123 sequence and changes seen in the sequence characteristic of evolutionary change (data not shown). pST123 is 2.3kb long but the two mRNA species seem in head populations by Sudhoff *et al.* 1989 are 0.85 and 0.9kb. There is no identifiable homology at the 3' ends of the sequences.

3.8.5 A cDNA Clone Derived from Na+/K+ ATPase β-Subunit Gene Family Member

The sequence obtained from the 5' of pST51 and pST99 showed homology to sequence of the Na+/K+ ATPase β-subunit gene family. However, this was only identified once over 300bps of sequence was obtained from the 5' end of pST51 and database searches performed with predicted peptides (as shown in figure 3.12). pST51 was originally selected for further study upon the basis of expression pattern, but the first sample exoIII deletion clone sequenced positively identified the clone. Later studies (chapter 4) show that the predicted amino acid sequence is most closely related to that of the shrimp (49% identity) but even with full length sequence DNA comparing programs cannot detect any significant homology at the DNA level.

3.9 Selection of Clones for Further Study

Figure 3.13 shows a summary of what is known about the clones from this screen. pST51 and pST123 were selected on the basis of interesting homology. Two other clones were selected for study, pST170 and pST162. These clones were selected because they are derived from genes that are most abundantly transcribed in head tissue.

A

The best scores are:		initl	initn	opt..
gb_om:btmrnasb	X76199 B.taurus mRNA for synaptobrevin. 1...	209	348	418
em_om:btmrnasb	X76199 B.taurus mRNA for synaptobrevin. 1...	209	348	418
em_in:dmsybanb	L14270 Drosophila melanogaster synaptobre...	193	302	360
gb_in:drosybanb	L14270 Drosophila melanogaster synaptobr...	193	302	360
gb_ro:ratvampb	M24105 Rat vesicle associated membrane pr...	167	290	383
em_ro:rnvampb	M24105 Rat vesicle associated membrane pro...	167	290	383
gb_pr:humsyb2a3	M36203 Human synaptobrevin 2 (SYB2) gene...	202	258	348
em_pr:hssyb2a3	M36203 Human synaptobrevin 2 (SYB2) gene,...	202	258	348
gb_pr:humsyb1a3	M36198 Human synaptobrevin 1 (SYB1) gene...	201	257	336
em_pr:hssyb1a3	M36198 Human synaptobrevin 1 (SYB1) gene,...	201	257	336
gb_ov:fscvampl	J03777 T.californica synaptic vesicle-ass...	207	247	400
em_ov:tcvampl	J03777 T.californica synaptic vesicle-asso...	207	247	400
gb_in:aplbrevin	U00997 Aplysia californica synaptobrevin...	187	235	311
em_in:aplbrevin	U00997 Aplysia californica synaptobrevin...	187	235	311

B

```
                                                10        20        30
ST123.                                          CAGCAGACGCAGGCGCAGGTCGATGAGGTN
                                                |||||||| || ||  |||| || |||||:
dmsyba AGGAACAACAATGCGGCCCAGAAGAAGCTGCAGCAGACCCAAGCCAAGGTGGACGAGGTG
          1700      1710      1720      1730      1740      1750

               40        50        60        70        80
ST123. GTGGACATCATGCGCACGAACGTGGAANAAGGTGCTGGAGC--GACAGGAAGCTGTCGGA
       || |  || |||||  ||||||||  |:|||||  |||||||  |||  |||||| |||||
dmsyba GTCGGGATTATGCGTGTGAACGTGG-AGAAGGTCCTGGAGCGGGACCAGAAGCTATCGGA
          1760      1770      1780      1790      1800      1810

       90       100       110       120       130       140
ST123. GCTGGACGACAGTGCCGATGCCTTGCAGCAGGGTGCCTCGCAGTTTGAGCAGCAGGCGGG
        |||| |||  |||| |||    || |||||||  || || ||||  |||||||||||  ||
dmsyba ACTGGGCGAGCGTGCGGATCAGCTGGAGCAGGGAGCATCCCAGTCCGAGCAGCAGGCCGG
          1820      1830      1840      1850      1860      1870

      150       160       170       180
ST123. CAAGCTCAAGAGGAAATTCTGGCTCCAGAACTT
       |||||| ||| | ||
dmsyba CAAGCTGAAGCGCAAGCAATGGTGGGCCAACATGAAGATGATGATCATCCTGGGCGTGAT
          1880      1890      1900      1910      1920      1930
```

Figure 3.11 Homology Searching with pST123 Derived Sequence

A shows the scores obtained by searching GenEMBL with sequence obtained by extension from a 5' vector based primer (methods section 2.13) using FastA (methods section 2.22.4). **Initl** is the score of the initial match, **initn** the highest score of extented homology surrounding the initial match and **opt** is the optimised score after final alignment of the match.

B shows FastA generated alignment of the pST123 5' sequence to a known *Drosophila* synaptobrevin cDNA(from Sudhoff *et al.* 1989).

A

The best scores are:		initl	initn	opt..
gb_pl:chntxx	Z00044 Tobacco chloroplast genome DNA. 12/93	70	169	71
em_or:chntxx	Z00044 Tobacco chloroplast genome DNA. 12/93	70	169	71
gb_pl:chosxx /rev	X15901 Rice complete chloroplast genom...	56	161	57
em_or:chosxx /rev	X15901 Rice complete chloroplast genom...	56	161	57
gb_pl:mtpacg /rev	X55026 P.anserina complete mitochondri...	72	160	123
em_or:mtpacg /rev	X55026 P.anserina complete mitochondri...	72	160	123
gb_pl:panmtpacga /rev	M61734 Mitochondrion Podospora ans...	72	160	123
em_ba:batrpeg	Z21938 B.aphidicola trpE and trpG genes fo...	114	157	123
gb_ba:batrpeg	Z21938 B.aphidicola trpE and trpG genes fo...	114	157	123
em_in:cef22b7	L12018 C. elegans cosmid F22B7. 4/94	67	156	77
gb_in:celf22b7	L12018 C. elegans cosmid F22B7. 4/94	67	156	77
em_ro:rncryg /rev	M19359 Rat gamma-crystallin gene clust...	63	156	85
gb_ro:ratcryg /rev	M19359 Rat gamma-crystallin gene clus...	63	156	85
gb_pr:humretblas	L11910 Human retinoblastoma suspectibil...	77	154	87

B

The best scores are:		initl	initn	opt..
gb_in:asnakatp	X55780 A.salina mRNA fro NA,K-ATPase beta...(3)	96	96	**119**
em_in:asnakatp	X55780 A.salina mRNA fro NA,K-ATPase beta...(3)	96	96	**119**
gb_ov:tcatpbr	X03471 Torpedo californica mRNA for (Na+ a...(2)	77	77	**83**
em_ov:tcatpbr	X03471 Torpedo californica mRNA for (Na+ a...(2)	77	77	**83**
em_ov:aaspab	X76109 A.anguilla mRNA for sodium/potassium...(1)	73	73	**80**
gb_ov:aaspab	X76109 A.anguilla mRNA for sodium/potassium...(1)	73	73	**80**
em_in:cek07e12	U00054 Caenorhabditis elegans cosmid K07E...(2)	51	72	**64**
gb_ov:bmnkab1	Z11797 B.marinus mRNA for Na, K-ATPase bet...(2)	72	72	**74**
gb_in:celk07e12	U00054 Caenorhabditis elegans cosmid K07...(2)	51	72	64
em_ov:bmnkab1	Z11797 B.marinus mRNA for Na, K-ATPase bet...(2)	72	72	**74**
gb_pl:mpomtcg	M68929 Marchantia polymorpha mitochondrion...(1)	71	71	73
em_or:mimpcg	M68929 Marchantia polymorpha mitochondrion,...(1)	71	71	73
gb_ba:ecol10k	D10483 E.coli K12 genome, 0-2.4min. region...(6)	44	70	44
em_ba:ecapah02	D10483 E.coli K12 genome, 0-2.4min. regio...(6)	44	70	44
gb_om:oaatpbr	X03883 Sheep kidney mRNA for (Na+/K+) ATPa...(1)	69	69	**72**
gb_pr:humnakatpa	M25159 Human Na,K-ATPase beta subunit (...(1)	69	69	**72**
gb_ro:mmnakatp	X16646 Mouse mRNA for Na,K-ATPase beta su...(3)	69	69	**72**
gb_ro:ratatpbsa	M14137 Rat brain Na- ,K- ATPase beta sub...(2)	69	69	**72**
gb_om:cfatpbr	X05297 Dog kidney mRNA for (Na+/K+)-ATPase...(2)	69	69	**72**
gb_om:ssatpbr	X03937 Pig mRNA for (Na+,K+)-ATPase beta-s...(2)	69	69	**72**

Figure 3.12 Database Searching with a pST51 Derived Sequence
A shows the results of a DNA search through GenEMBL using the initial 5' sequence obtained from pST51 using a vector based T3 primer using FastA (performed as described in methods section 2.22.4). No significant match was found. **Init1, initn** and **opt** are as described in figure3.11.
B shows the result of searching the GenEMBL database with a peptide derived from the predicted first 30 AA of the protein coding sequence, located between ~200 and ~300bp from the 5' termini of the cDNA insert. Initial searches employed a range of predicted peptides in all forward frames (data not shown). Optimised scores to ATPase β-subunits are shown in bold. The matching translation frame is shown bracketed.

clone number	3'XhoI by seq	3'XhoI by map	internal XhoI	5'EcoRI by seq	5'EcoRI by map	orientation by seq	size by rest. map	sequence 5'	sequence 3'	detect Head	detect body	detect pupae	detect embryo	Matches	selected? because	Head elevated by Northern
ST41	ND	ND	ND	ND	ND	5'-3'	1.6	Y	Y	1000	UD	UD	UD	*ninaE*		Y
pST42	ND	Y	N	Y	Y	5'-3'	0.8	Y	Y	100	UD	20	10			
pST51	Y	Y	N	Y	Y	5'-3'	1.4	Y	Y	100	UD	10	1	beta subunit and pST99	homology	Y
pST59	Y	Y	Y	Y	Y	5'-3'	0.5	Y	Y	500	20	10	UD	pST133, and 170 pC12 and 13	related	(Y)
pST72	ND	Y	N	Y	Y	5'-3'	0.7	Y	Y	500	50	10	UD			
pST78	ND	Y	N	Y	Y	5'-3'	1.4	Y	Y	1000	1000	1000	1000			
pST99	Y	Y	N	N	N	5'-3'	1.4	Y	Y	100	UD	10	UD	beta subunit and pST51	related	(Y)
pST116	Y	Y	N	Y	Y	5'-3'	1.1	Y	Y	100	10	10	10			
pST123	Y	Y	Y	N	N	5'-3'	2.3	Y	Y	100	UD	10	10	synaptobrevin	homology	Y
pST133	ND	Y	N	Y	Y	5'-3'	0.6	Y	Y	1000	20	10	UD	pST59 and 170 pC12 and pC13	related	(Y)
pST134	Y	Y	Y	Y	Y	5'-3'	0.4	Y	Y	10	UD	10	UD			
pST135	Y 5'and3'	Y 5'and3'	Y	N	N	back to back	2.6	Y	Y	100	50(LB)	10(all)	0.1	pST151		
pST151	Y 5'and3'	Y 5'and3'	Y	N	N	back to back	2.7	Y	Y	100	UD	10	0.1	pST135		
pST162	Y	Y	N	Y	Y	5'-3'	2.1	Y	Y	100	UD	10	1		expression	Y
pST163	Y	Y	N	Y	Y	5'-3'	0.6	Y	Y	100	UD	10	10			
pST170	Y	Y	N	N	Y	5'-3'	0.45	Y	Y	500	UD	10	UD	pST59 and 133 pC12 and 13	expression	Y
pST141	ND	Y	N	Y	Y	5'-3'	3.4	Y	N	500	100	10	10			
pC13	Y	Y	N	?	Y	?5'-3'	1.6?	Y	Y	1000	50	20	UD	pST59 170 133 pC12	expression	Y
TOTAL	all tested	all cut	5/18 cut	10/16	13/17	16/18	Av=1.4	Av=240bp						14 indep^{t} cDNAs	8/18	all tested

Figure 3.13 Summary of Clone Data

The information shown in the table was used to select clones for further study. Abbreviations used in the table are N-No, Y-Yes, ND not determined, UD- undetected, LB-Lower band, Av- average. Data for ST41 is from the lambda clone, all other cDNAs were in pBluescriptII SK^{-}. Some of the data for pC13 (and pC12) was kindly provided by C.Millagan (Glasgow Genetics Department). Northern analysis, with the exception of pST123, was performed after selection of clones. Northern data for ST41 was inferred from Zucker *et al.* 1985. Clones marked related cross-hybridise to clones selected for further study. Because of this relationship, they must also identify head elevated bands if used to probe a northern blot (and hence are shown bracketed in the northern column).

3.10 A Retrospective Examination of the Differential Screen Presented in Section 3.4

Data obtained on the four groups of clones selected for futher study is very indicative of the success of this screen. For example, northern blotting experiments with the four selected clones confirm a head elevated pattern of expression (refer figures 4.15, 5.9, 6.2 and 6.4).

One interesting feature of the screen is that many of the positives from the first round did not pass through the secondary screen. There was an 80% failure rate at this stage (data not shown), although it is difficult to assess accurately because weak negative clones were used to fill up the other spaces in the secondary screen array (and so the primary screen can be considered a pre-screen to remove strong non-differentials). This is interesting because the accuracy when moving from the secondary screen to the tertiary reverse northern head body screen was 95% (data not shown). The use of the ordered array ensures that the identification of positive plaques is very accurate. Great care was taken when selecting plaques corresponding to potential positives. The selective nature of the secondary screen is argued for by the under representation of strong differentials in the secondary screen's body array (see figure 3.6). This array was exposed for twelve hours but only a single abundant non-differential clone can be seen. Many more of this class can be seen on the primary arrays.

There is another possibility to account for the relative screening efficiencies. In the primary screen, although probes were of equivalent specific activity, fine tuning of exposure level was made by reference to the non-differential plaques of the array. This is reasonable, since the most common cDNAs in the library are non-differential. In the secondary screen, fine tuning was performed by comparison to a negative control, which was the non-recombinant vector of the two libraries. This makes the assumption that the background hybridisation to the plaques is equal for both probes. Accumulated evidence in the laboratory suggests that this is often not the case; head probes very often give a higher vector background than body probes (unpublished laboratory data, Glasgow Genetics Department). The reason for this is not clear but is supported by independent data from four workers in the laboratory. In the secondary screen, final comparison was made of the twelve hour body autoradiograph to the five hour head autoradiograph. If equivalent time comparisons are made many more plaques

appear differential than in the first round of analysis. Fortunately, the 'phage used to make the secondary array have not been discarded.

It is also interesting to note why the difference between a 12 and 24 hour exposure of filters in a differential screen is more critical than it first appears. If the genuine level of head elevation is 100 fold then this plaque may give a signal in head of 100 units and with body of 1unit. But if general level of background is also 100 units then the actual signals recorded are for head (100+100)200 units and body (100+1)101 units. This is just a two fold difference in intensity and so could easily be obscured by the difference between a 12 and 24 hour exposure. It is clear to that this problem becomes more critical for weaker signals and higher background levels.

3.11 Conclusions

This chapter describes the identification of a number of cDNA clones of mRNAs expressed strongly in the head and weakly in the body of *Drosophila melanogaster*. Preliminary characterisation identified several of the clones as being derived from members of known gene families. Several of these clones, pST51 and pST123, have been selected for further study on this basis. Two other clone groups, represented by pST170 and pST162, have also been selected for further study. These clones were selected upon the basis of the expression pattern.

Chapter 4 A Novel Member of the Na+/K+ ATPase β-Subunit Gene Family

4.1 Introduction

Na+/K+ ATPases are pumps found on the surface of nearly all animal cells that transport Na+ ions out of the cell and K+ ions into the cell, with the concomitant hydrolysis of ATP. The pump acts to establish and maintain an electrochemical gradient across the plasma membrane. This gradient is of crucial importance to the cell, both because of its direct chemical effects, and because many ion and solute co-transport systems employ the gradient as a source of energy. In addition to this house-keeping role, the Na+/K+ ATPase also has an important, specialist role in excretory systems and in excitatory tissues (nervous system and muscle). The specific nervous system functions of the Na+/K+ ATPase includes both nerve impulse generation and a role in glial cells to buffer changes in extracellular K+ (reviewed Horisberger *et al.* 1991).

Na+/K+ ATPase activity has been studied for about sixty years (reviewed Dean 1987). The establishment of an electrochemical gradient was first detected as early as 1930 using directional pumping activity found in frog skins (Dean and Gatto 1937). The active process of Na+ and K+ exchange was shown experimentally in the early 1940s when it was shown that rats fed upon a low K+ diet had unusually high concentrations of Na+ in their muscle tissue, but this could be rapidly replaced by K+ when it became available (Heppel *et al.* 1940). This suggested the presence of an active exchange mechanism, but it wasn't until 1948 that a specific Na+ pump was implicated in this exchange process (Ussing 1951) and until 1957 that the link to ATPase activity was established (Skou 1957). Since these early pioneering studies a wealth of molecular, pharmacological and physiological knowledge has been accumulated. A brief summary of what has so far been learned, with special emphasis upon the β–subunit, is presented below.

In mammals the Na+/K+ ATPase functional unit has been shown to be heteromeric. The 112kDa α-subunit contains the catalytic regions of the enzyme, including the binding sites for Na+, K+, ATP and ouabain (a glycoside inhibitor of the pump). This subunit has been isolated from a wide

variety of organisms including human, *Torpedo*, *Xenopus*, *Artemia* and *Drosophila*. These subunits are very highly conserved; for example the *Drosophila* and human (α1) subunits are 90% identical at the AA level. The subunit is a transmembrane type four protein, predicted to span the membrane between six and eight times (reviewed Horisberger *et al.* 1991, Lingrel 1992, Inesi and Kirtley 1992).

By contrast the role of the small 35kDa glycosylated β–subunit is much less well understood. It is known that the pump is not active without this polypeptide, but the exact function of this subunit is still rather elusive. However, several experiments using injected mRNA into *Xenopus* oocytes and somatic cell lines have shown that the subunit is required for translocation of the newly synthesised pump through the ER and Golgi complex and onto the cell surface. It has also been shown to be required for the formation of a mature, active pump complex on the cell surface. A basic model has been proposed for the role of the β-subunit, this is shown as figure 4.1.

Study of the β-subunit is difficult *in vitro* because most of the model eukaryotic systems express versions of the Na+/K+ ATPase. An active pump can be formed from exogenous α- subunit and endogenous β-subunit. Approaches to the study of these pumps in expression systems initially used endogenous/exogenous subunit mixtures (e.g. Noguchi *et al.* 1987), but later used selective inhibition of endogenous pumps (Jaisser *et al.* 1992), or used a yeast expression system in which no closely structurally related pump exists (Horowitz *et al.* 1990) and so specific pump inhibitors could be used to assay the activity of the exogenous pump.

The early studies of Noguchi *et al.* 1987 in *Xenopus* oocytes have shown that both the α- and β-subunits are required to be co-transfected to produce exogenous pump on the surface (assayed by ATP hydrolysis, and ouabain binding). These studies are supported by studies in yeast (Horowitz *et al.* 1990) which have shown that in this system both α and β-subunit need to be co-expressed to produce ouabain binding and ATP hydrolysis on the cell surface.

The native β-subunit has been shown to be a glycoprotein by treatment with gycosidases (Fambrough 1983, Sweadner and Gilkeson 1985, Brown *et al.* 1987, Marshall and Hokin 1979, Smith *et al.* 1987). Additionally, several

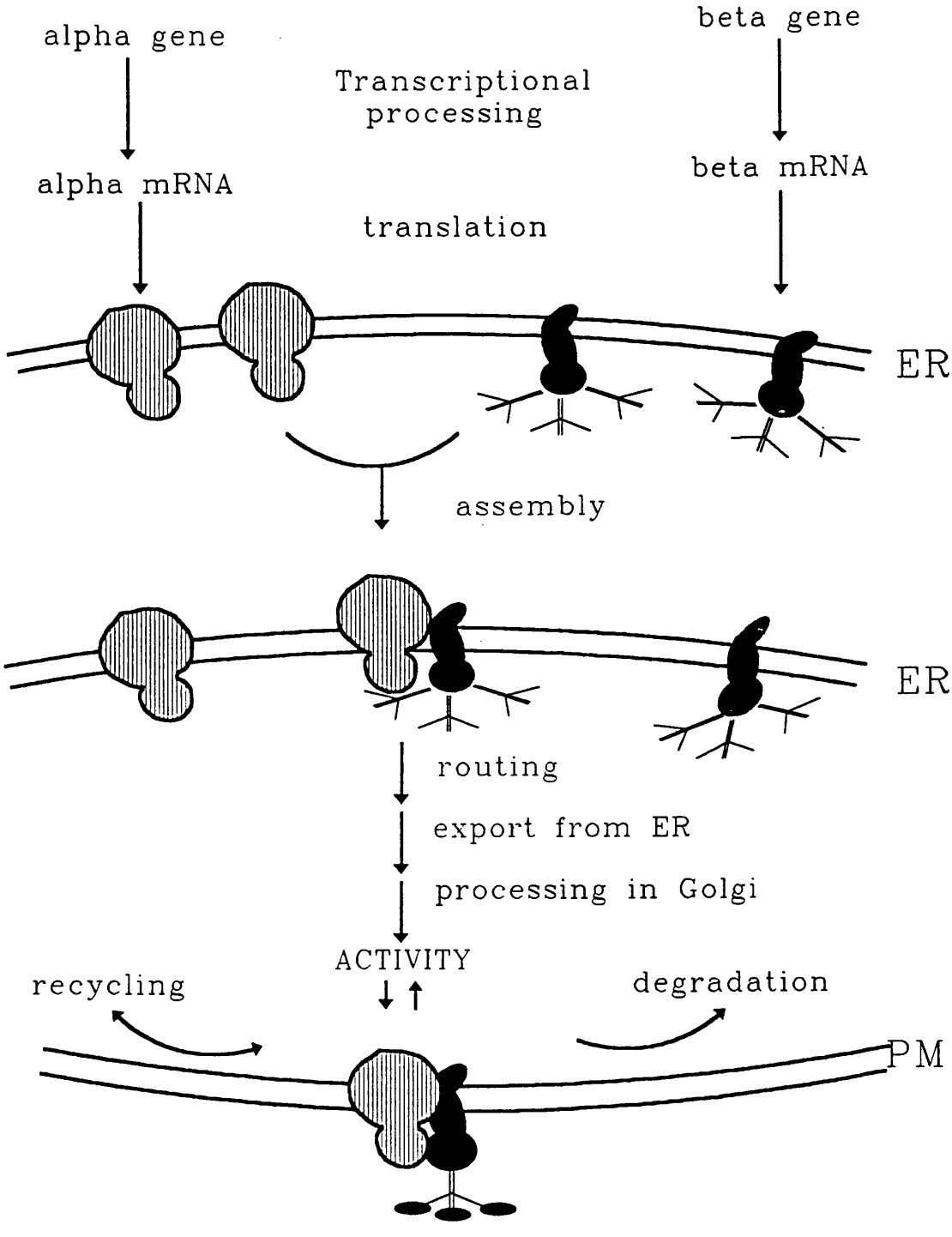


Figure 4.1 The Proposed Role of the Na^+/K^+ATPase β-Subunit.
The figure shows a model of Na^+/K^+ATPase synthesis and assembly, where ER is the endoplasmic reticulum, PM is the plasma membrane. Both subunits are cotranslationally inserted into the membrane, associate, and then the resulting structure leaves the ER and is processed in the Golgi and routed to the surface of the plasma membrane. Glycosylation of the β-subunit is shown by the branching. These sugar residues are modified during Golgi passage before the final functional unit is produced upon the cell surface. (Redrawn and slightly modified from McDonough *et al.* 1990).

studies in mammalian cell lines have shown that treatment with the glycosylation inhibitor tunicamycin does not cause a decrease in the formation of αβ complexes, their transport, or their degradation rate (e.g. Zamofing *et al.* 1988 and 1989). However it does result in a decrease of β-subunit on the cell surface and a parallel decrease in the α-subunit, which is not itself a glycoprotein. Controls were performed to show that the inhibitor, tunicamycin, was not having a non-specific effect.

Other studies have also implicated the β-subunit in altering the stability of the pump complex (Geering 1989). It has been observed that the pump undergoes a transition from trypsin sensitive to resistant about 20 minutes after synthesis (Geering 1989). This change is associated with the pump taking on its mature, functional properties, for example, ATP binding (Caplan *et al.* 1990). It has been shown in *Xenopus* oocytes that this change from inactive/sensitive to active/resistant forms can be triggered by the co-injection of β-subunit cRNA with the α-subunit cRNA that alone had previously produced the trypsin sensitive form (Geering 1989).

Using deletion studies in a *Xenopus* oocyte transfection assay several workers have began to map the regions of the β-subunit responsible for the association with the α-subunit and the progression through the ER and Golgi. For example from the study by Renaud *et al.* 1991 it appears that the N-terminal (cytoplasmic) domain is not required for these processes. The transmembrane region can be deleted by up to 14 residues (plus four flanking the region) and yet the two subunits can still associate. However, removal of more than four residues from the transmembrane region resulted in a loss of ability to transport the pump complex to the cell surface, leaving the αβ complex trapped in the Golgi.

This requirement for the β-subunit to be expressed with the α-subunit allows the possibility of additional regulation. It has been shown that in some systems subunit expression is co-ordinately regulated. This has been demonstrated for the response of kidney cell lines to thyroid hormone, which up-regulates the pump by increasing the expression of both subunits (McDonough *et al.* 1988). However, in other systems, for example mammalian LLC-PK1 cells in response to low K+ concentration in the medium only the β-subunit, and not the α-subunit, is up-regulated (McDonough *et al.* 1990). This two-fold up-regulation is paralleled by the doubling of pump activity. This is interpreted as

a stabilisation effect of the β-subunit on the α-subunit (McDonough *et al.* 1990).

Both the α-subunit and β-subunits are encoded by multiple loci. In mammals, three separate loci have been identified. The expression patterns of these three genes are complex. For example, consider the study of Emanuel *et al.* 1987 on the three rat α-subunits. A summary of the molecular properties of these three genes is shown in figure 4.2.

From figure 4.2 it is clear that the molecules are under complex regulation, which perhaps reflects the complex cellular demands placed upon the Na+/K+ ATPase. It is also interesting to note that several isoforms are abundantly or exclusively, expressed in the nervous system.

It has been shown that different α-subunits produce pumps with varying affinities for ATP, ouabain and Na+ (e.g. Sweadner 1989). These modified properties have almost exclusively been shown to be associated with the α-subunit; for example if one injects α-subunit or β-subunits from a ouabain resistant pump form into oocytes which possess a sensitive pump, gain of resistance is associated with the exogenous α-subunit and not the β-subunit (e.g. Jassier *et al.* 1992).

Three groups of β-subunit have so far been identified, called β1 to β3. All known Na+/K+ ATPases are 295-305 AA long and about 35kD. They all have a single putative transmembrane region that spans the membrane once. They all possess six conserved cysteine residues (a seventh occurs in the transmembrane region of type β1's only) and these have been shown to form three disulphide bonds in the case of the β1 form (Brown *et al.* 1987). All the subunits mRNAs possess consensus glycosylation sites, and these residues are conserved within each of the types, but not always between types. These common features are shown in figure 4.3.

Different β-isoforms are distinguished from each other principally upon the basis of amino acid sequence homology. An alignment between representative sub-family members is discussed in detail later in this chapter (see section 4.5.3). β-subunits are more highly conserved within types between species, than between types and within a species. Some individual residues are also isoform specific; the seventh conserved cysteine within the transmembrane

A. Expression in Tissues

Isoform	Transcripts	Tissue and stage that transcript is expressed in											
		Brain			Heart			Kidney			Liver		
		Fetal	2week	Adult	Fetal	2week	Adult	Fetal	2week	Adult	Fetal	2week	Adult
α1	4.5Kb	*	***	**	**	***	**	**	**	****		*	*
α2	4.5Kb		****	***					**	**		**	*
α3	6.0Kb		****	***	*		**	**				**	
	4.5Kb									**		**	
	4.0Kb		****	***			**	***		*		*	

B. Expression in Nervous System Cell Lines

Isoform	Transcript	Glial cell lines		Neuronal cell lines					
		B15	B82	B35	B50	B60	B103	Pc12 NGF+	Pc12 NGF-
α1	4.5Kb	***	**	**	**	**	*	****	****
α2	4.5Kb	**	*	*	*	*	*	***	*
α3	6.0Kb				***				
	4.5Kb	**	**	**	***	**	**	**	**
	4.0Kb				***				

Figure 4.2 Differential Expression of Na/K ATPase α-subunit mRNAs in (A) Rat tissues and (B) cell lines.
The data in these tables is a summary of the data presented in Emaunel *et al.* 1987 where *- very weakly expressed, **- weakly expressed, ***- strong expression and ****- abundant expression.
The table is a guide only, the ratings are <u>not</u> quantitative. Please refer to the text for more details.

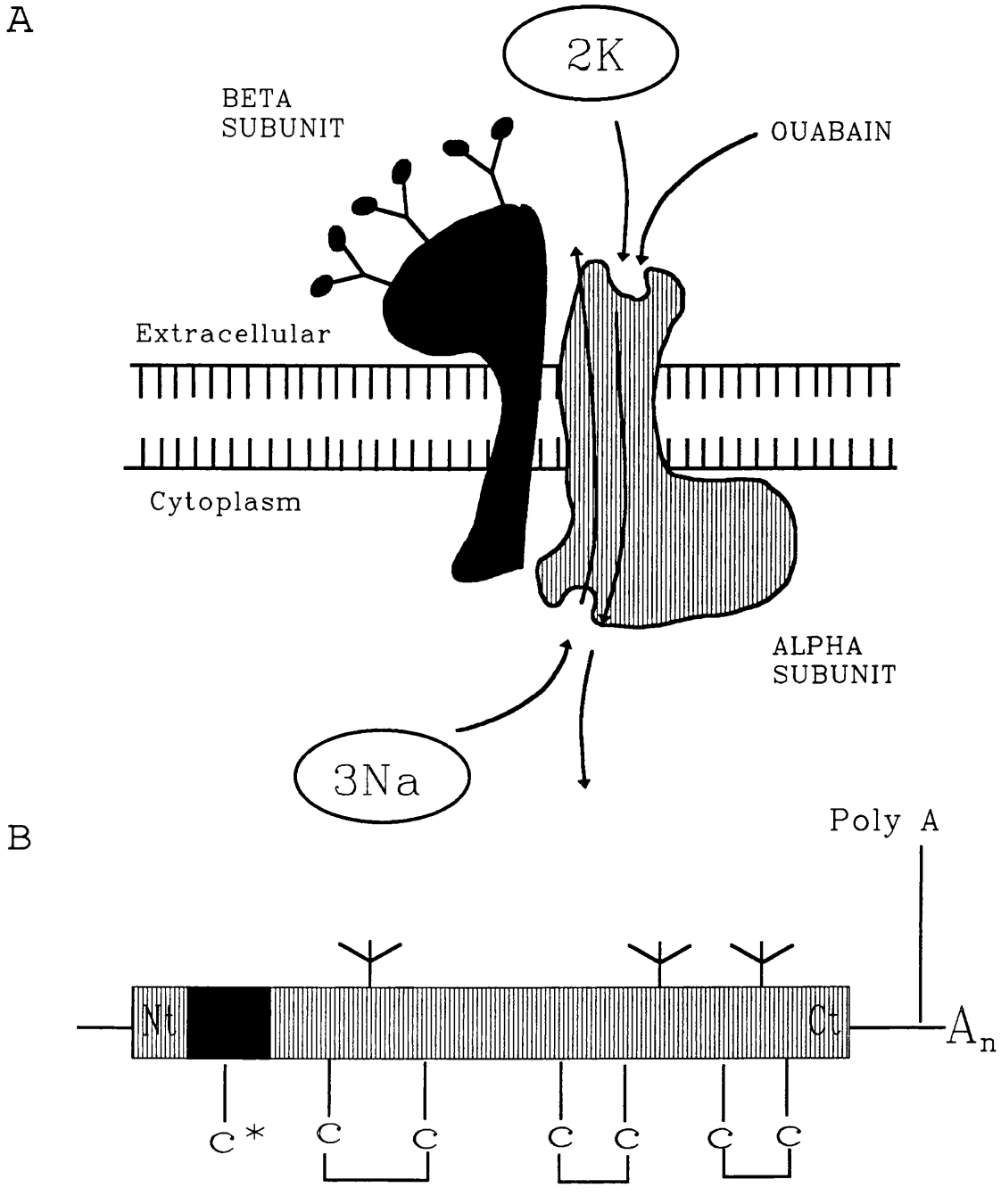


Figure 4.3 General Structural Features of Na+/K+ ATPases

A shows a simple model of the structure of the Na+/K+ ATPase, redrawn and modified from Horisberger *et al.* 1991. The figure shows the cation binding sites and the site of binding of the Na+/K+ ATPase inhibitor ouabain. Only one of each subunit is shown for clarity.

B shows the conserved structural features of the Na+/K+ ATPase β-subunit mRNA and protein. All predicted amino acid sequences possess six conserved cysteine residues (C), a seventh (C*) being present in β1 subunits only. The six cysteine residues have been shown to form disulphide bonds *in vivo* in the conformation shown (Brown *et al.* 1987). The three ASN glycosylation sites are shown as branched structures. The thirty AA transmembrane region is shown blacked. The approximate size of the mRNA molecule is 1.8kb, excluding polyA tail, and is predicted to encode a polypeptide of ~300AAs.

regions of type one subunits is an example of this. As well as being a useful means of classification, this suggests that individual β-subunit isoforms have distinct roles that have led to the conservation of functionally important differences across evolutionary time. These differences are superimposed upon a common ability to produce an active pump, even with α-subunits from a different species.

In addition to the β-subunit's intracellular transport/maturation roles it has also been proposed that the glycoproteins might directly modulate the activity of the pump on the cell surface. A number of studies sought to map biochemical properties of the pump such as substrate and ouabain binding onto the subunits; all these studies implicated sites within the α-subunit. Differences in functional properties between pump isoforms also seem to correlate with the α-subunit. However, a more recent study in *Bufo marinus* (Jassier *et al.* 1992) has tested in *Xenopus* oocytes the properties of α1β1 and α1β3 and shown a significant difference between the two pumps. This study exploited the fact that the *Bufo marinus* α-subunit produces a ouabain resistant pump, and so the "noise" in the system from the endogenous pump could be removed by pre-treatment with ouabain. The reported difference may be of physiological significance as, for example, the two β-subunits in the kidney respond differently to aldosterone. The pumps are differentially sensitive to concentrations of intracellular K+ ions at concentrations that have been experimentally recorded within the system.

One β-subunit, β2, has also been shown to have an isoform specific role in cell-cell recognition (Gloor *et al.* 1990). This subunit is expressed in glial cells and is thought to be important in the interaction of the glial cells with intracellular components. This subunit is associated with an α-subunit at the cell surface and does produce an active pump in combination when injected into *Xenopus* oocytes. This form of β-subunit is heavily glycosylated (eight consensus ASN glycosylation sites), and, by analogy to other cell-cell recognition molecules this is thought to be of significance in the molecules role as an adhesion molecule.

The variety of subunits that have already been defined is complicated by the presence of related pumps. The gene superfamily currently includes two other homologous members, the K+/H+ antiport that plays an important role in mammalian gut acidification and the Ca2+/Na+ exchanger in the sarcoplasmic

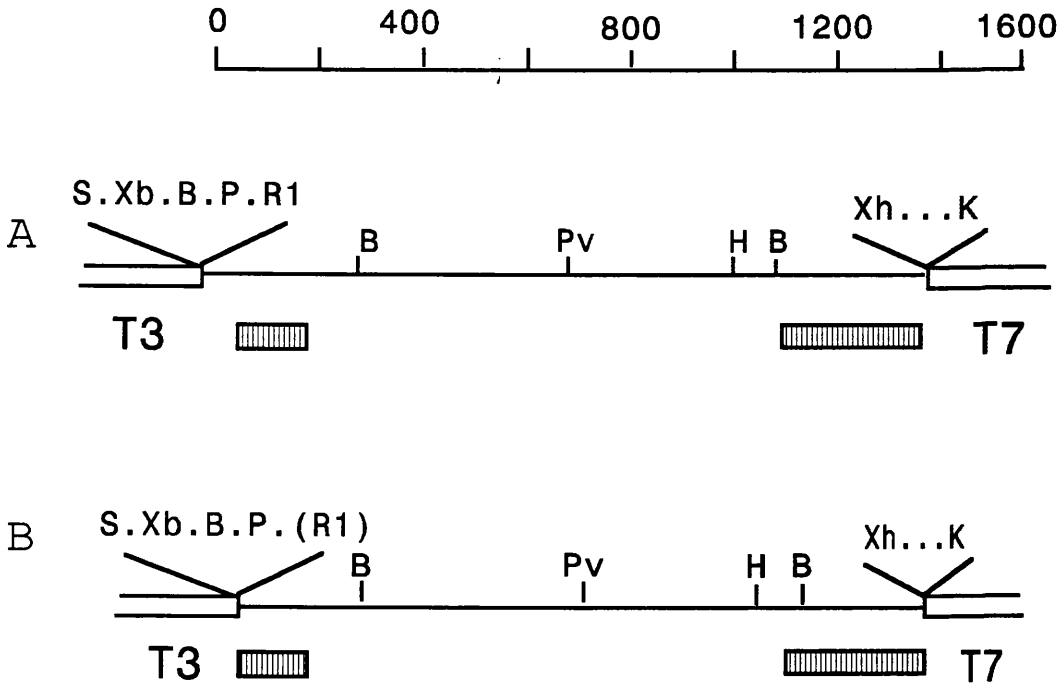
reticulum. The α-subunits are 70% related between the α1 Na+/K+ ATPase and the K+/H+ ATPase and 40% α1 between the Na+/K+ ATPase and the Ca2+/Na+ ATPase. However, only the K+/H+ ATPase has an identified, cloned β-subunit. It is about 30% identical to the β1-subunit of mammals (e.g. Shull 1990). The sarcoplasmic Ca2+ pump is thought to function *in vivo* in the presence of a β-subunit, although its presence is not required for function of this pump *in vitro*. Remarkably it has been shown that the proton pump β-subunit can associate with a Na+/K+ ATPase α-subunit and produce an active Na+/K+ ATPase on the cell surface in *Xenopus* oocytes. The features that confer function are likely to be within the evolutionary conserved regions of these subunits

In *Drosophila*, only the α-subunit has so far been cloned (Lebovitz *et al.* 1989). *Drosophila*, unlike the mammalian species, only seems to have a single, highly conserved, α-subunit. An independent attempt have been made to isolate further loci (Varadi *et al.* 1989) but these have so far failed, although they do find other related family members (a sarcoplasmic Ca2+ pump homologue). The single identified locus produces three different mRNAs that are present at all stages of development in varying amounts. Sequence data and restriction mapping of cDNAs representative of two of the forms reveals that the transcripts differ in the 3' untranslated regions of the cDNAs. An antibody specific for the α-subunit (raised against a chicken α1 antigen) identifies the gene to be expressed in large amounts in the central nervous system (head and thoracic ganglia), in the malpigian tubules and to a lesser extent in the musculature. Interestingly, the α-subunit was only able to interact weakly with the endogenous β-subunit when expressed in mouse L-cells, suggesting that this subunit at least may have a requirement for a more closely related β-subunit.

In this chapter I will describe the characterisation of a candidate cDNA with the capacity to encode the first described Na+/K+ ATPase β-subunit identified from *Drosophila*.

4.2 Restriction Mapping of pST51 and pST99

The differential screen (chapter 3) yielded two independent clones with were derived from messages from a Na+/K+ ATPase β-subunit gene family. The two clones are called pST51 and pST99. Both these cDNA clones were restriction mapped using six cutter enzymes that cut once within the



C

```
CCCCGGGCTGCAGG-ATTCGGCACGAG---------------------------------------- pST99
CCCCGGGCTGCAGGAATTCGGCACGAGTTAAATGCAAGTCACAACAATACTATGGTTACAGACGTTTAAA pST51

----------------------------------------------ATAACGACTTAAAACAAAAA pST99
                                               ::::::::::::::::::::
AATATATAAAAAAAATTTTGCACCAGAACAGTTGATAAAAGAAACTTAAAATAACGACTTAAAACAAAAA pST51

TAAATTCGAAATAATTCAACTTGAAAGACAATACGCCCATCAACACACTACAAAGTTAAAATGGCCGATA pST99
:::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
TAAATTCGAAATAATTCAACTTGAAAGACAATACGCCCATCAACACACTACAAAGTTAAAATGGCCGATA pST51
```

D

```
TAAAGGAATGGAATATAGAAAAAAAAAAACTGACAACAAAACCCACACAGAATAGATACGAGAAAAAtGTA pST99
:::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
TAAAGGAATGGAATATAGAAAAAAAAAAACTGACAACAAAACCCACACAGAATAGATACGAGAAAAATGTA pST51

GTAGCGCAGAAACGTCAGGTTAAAGCAAAAACAAAAACAAAAAAAAAAAAAAAAAACTCGAGGGGGGGCCC pST99
::::::::::::::::::::::::::::::::::::::::::::::::::: ::::::::::::::::
GTAGCGCAGAAACGTCAGGTTAAAGCAAAAACAAAAACAAAAAAAAAAAAAAAAA-CTCGAGGGGGGGCCC pST51
```

Figure 4.4 Restriction maps of pST51 and pST99

A and **B** show restriction maps for the pST51 and pST99 cDNA inserts respectively, where S-*Sac*I, Xb-*Xba*I, B-*Bam*HI, R1-*Eco*RI, RV-*Eco*RV, P-*Pvu*II, Xh-*Xho*I, H-*Hind*III, and K-*Kpn*I restriction sites. Boxed region between the two maps indicated the region of known sequence identity between the two sequences. pST99 has not been sequenced beyond the reach of vector based primers. Scale is in basepairs.

C and **D** show the 5' and 3' aligned junction point sequence. The 5' termini of the pST51 cDNA is 93bps longer than that of pST99. The 3' termini differ only in a single base pair in the cloned polyA tail. Vector sequence is shown underlined, with the vector based *Eco*RI(5') and *Xho*I(3') sites shown in bold. The 5' *Eco*RI site is defective in the pST99 cDNA clone.

pBluescript II SK- polylinker. The resulting maps, figure 4.4a and 4.4b, and the previously performed 5' and 3' sequencing runs, both suggest that pST51 is an extended version of pST99. pST51 extending 100bp further 5' and was therefore selected for further analysis. The relationship between these sequences is shown in figure 4.4c and 4.4d and will be returned to in more detail in section 4.7.

4.3 Deletion of Clone pST51 using ExonucleaseIII

ExoIII can be used to generate a set of deletions across a piece of DNA for use as sequencing templates. The method used here is a modification of Henikoff 1984 and is described in methods section 2.12. Use of the precise buffer conditions is critical to the success of the experiment. Promega supply a quality controlled buffer system and this was used in the early experiments (a kind gift from G.Griffith, Glasgow Genetics Department). Once reliable exoIII deletions could be produced (using the pST123 plasmid described in section 5.2) this working system was used to construct an exoIII deletion kit. This kit was distributed to a number of laboratories and has been used successfully in most of these (C. Milligan A. Griffin and E. Galleger, Glagow Genetics Department and Robertson Institute of Biotechnology, Glasgow).

The exoIII procedure used to generate one set of pST51 exoIII deletions is shown in figure 4.5. Digestion with *Sac*I and *Eco*RI was similarly used to generate templates for sequencing the other strand of DNA. 10µg of CsCl prepared plasmid DNA was digested with 40units of *Kpn*I for three hours to generate the 3', exoIII protected cut. A sample of this digest was then electrophoretically seperated on a 1% agarose TBE gel to assess the extent of cutting (data not shown). Failure to digest to "to completion" results in contamination of the deleted plasmids with undeleted plasmids. Buffer conditions for the remainder of the *Kpn*I restricted plasmid were adjusted to suit the second enzyme and the template digested for three hours with 40units of *Xho*I. To assess the extent of digestion with the second enzyme, (which must also digest to completion), test exoIII digestions were performed with 100ng of template and 80units of exoIII for 15mins at 37°C (data not shown). Under these conditions double cut plasmid (*KpnI/XhoI*) should be expected to show deletion, while the single cut plasmid (*Sac*I only) should be protected. However, under the conditions specified by Promega, (modified from Henikoff 1984) both the single and the double cut templates digested to completion. This experiment was repeated twice for three different combinations of

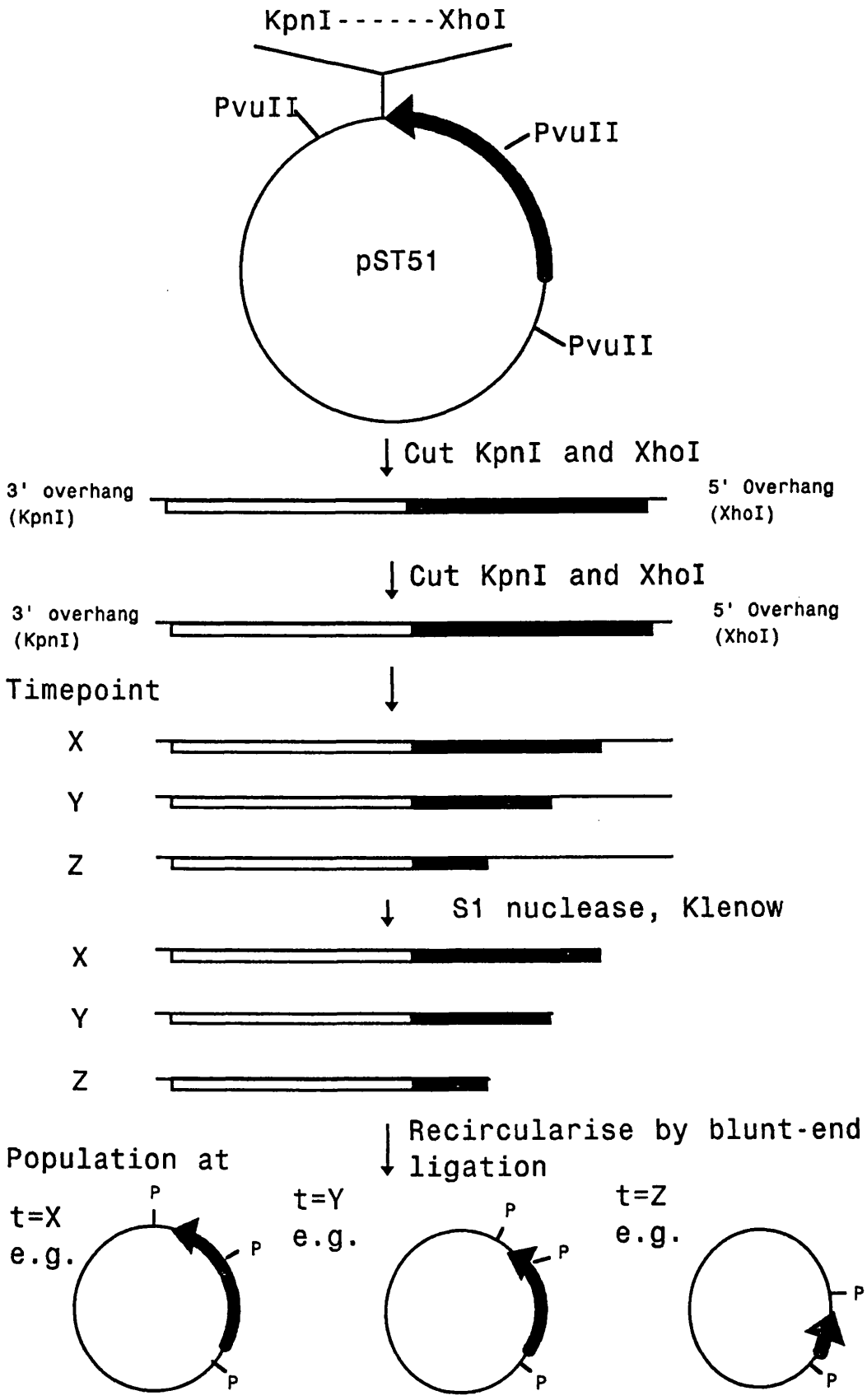


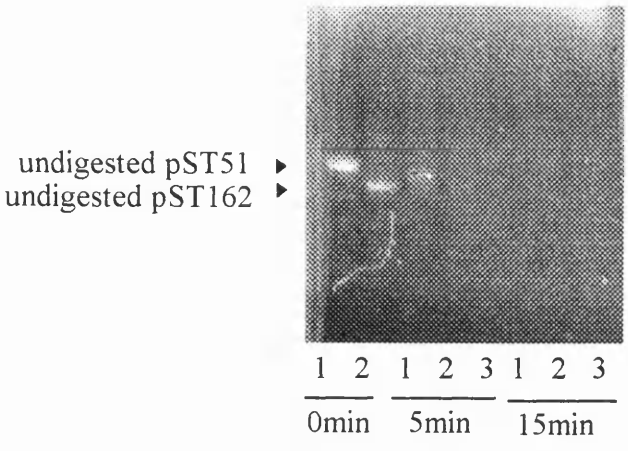
Figure 4.5 Procedure for ExoIII Deletion of pST51.
Insert sequence is shown blacked, the arrow points 3' and so indicates the direction of cloning of the cDNA insert. Abbreviations are **t** for time and **P** for *Pvu*II restriction cleavage sites. The figure is modified from the Promega Protocols and Applications Guide (1991).

restriction enzymes and plasmid and always gave the same result. It appears that the 3' overhang enzymes were cutting, but were unable to protect from the exoIII digestion.

Titration experiments were then performed using single and double cut templates that showed that under shorter time conditions the single cut (3' overhang, protecting) enzyme templates were partially resistant to exoIII after 5 minutes, but most of the template was digested upon longer incubation (figure 4.6a). Since the enzyme progresses 450bp/min at 37°C the degree of protection from non-specific degradation was sufficient to allow the deletion experiments to be performed. In addition, exoIII deletions were carried out under conditions of reduced enzyme concentration to try to limit the non-specific exonuclease digestion of the 3' protected end (300units exoIII was used per 10μg of template). Later experiments showed that this non-specific digestion was enzyme batch dependent (data not shown). As an additional control pST51 was digested with *Xho*I only and then this template exoIII digested to give an indication of the "both-sides" rate. Samples of *Kpn*I/*Xho*I and *Xho*I only cut template were digested with exoIII as described in the methods section 2.12.3 with timepoints ranging from 0 to 180 secs. The resulting samples were treated with S1 nuclease and then an aliquot electrophoretically seperated through an agarose gel. This gel is shown in figure 4.6b. From figure 4.6b it can be calculated that the rate of digestion for the *Kpn*I/*Xho*I template was 550bp/min and for the *Xho*I template 1250bp/min. Thus the rate of digestion of the double cut template is close to described in the literature (450bp/min) and also ~50% of the rate of the single cut template as expected.

Each timepoint sample was treated with Klenow to flush the DNA termini and then recircularised with ligase and the ligation products used to transform XL1B competent cells (see methods section 2.3.4). Figure 4.7a shows the numbers of transformants obtained from both the *Xho*I/*Kpn*I experiment. Five colonies from each timepoint were selected at random and minipreped (methods section 2.2.4.1). After electrophoresis of the uncut plasmid clones through a 0.7% agarose gel, selected plasmids were digested with *Pvu*II to release the inserts. The digestion of ten such plasmids from the *Xho*I/*Kpn*I experiment is shown in Figure 4.7b. The restriction pattern is consistent with the clones being progressive deletions of pST51 from the *Xho*I side of the clone.

A



B

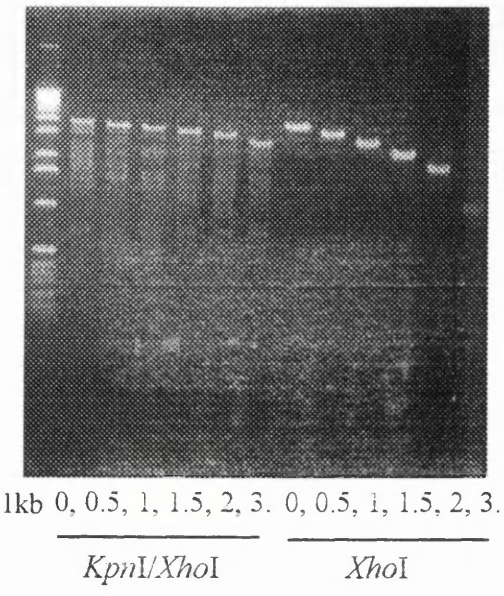


Figure 4.6 ExoIII Deletion of the pST51 Insert

A shows an ethidium bromide stained agarose gel demonstating that exoIII shows a limited ability to delete templates from "protected" termini. The three templates used for the experiment were **1**-pST162 restricted with *Kpn*I (protected termini), **2**-pST51 also cut with *Kpn*1 and **3**-pST51 cut only with *Xho*I (unprotected termini). These templates were digested with ExoIII as described in the text and samples of the reaction stopped after five and fifteeen minutes.

B shows the products of exoIII and S1 nuclease digestion of *Kpn*I/*Xho*I and *Xho*I digested pST51 plasmid after agarose gel electrophoresis. Digestions were performed under the modified conditions described in the text. Samples of the exoIII reaction were removed at the time intervals shown. The digestion rate of the *Xho*I single cut template is double that of the *Xho*I/*Kpn*I digested template. Time is in minutes.

A

	Timepoint(secs)	Total number of transformants	Comments
(kpn1/Xho1) pST51	0	1000	This was ligated
	30	500	
	60	500	
	90	1000	
	120	500	
	150	200	
	180	200	
-ve control	NA	0	Cells only
+ve control	NA	$1x10^5$ per µg	pBluescript SK minus

B

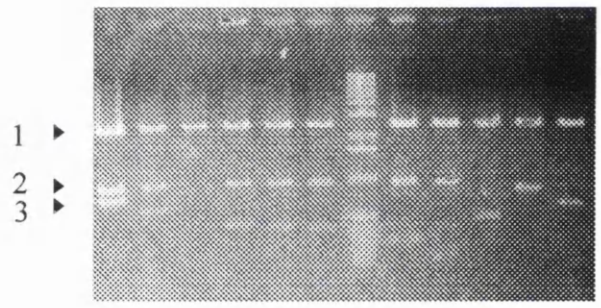


Figure 4.7 Selection of ExonucleaseIII Deletion Plasmids

A shows the numbers of transformants obtained from each timepoint of the deletion experiment. Each ligation reaction contained about 100ng of plasmid DNA. The transformation competence of the cells was determined using a plasmid control of known concentration.

B shows an ethidium bromide stained agarose gel electrophoresis of ten single plasmid minipreps from the deletion experiment after digestion with *Pvu*II. *Pvu*II cleaves pBluescriptII either side of the polylinker and so acts to release the vector from the insert. The 1.4kb pST51 cDNA insert also contains an internal *Pvu*II site located about 1kb from the 5' *Pvu*II site. Thus three bands are possible, a vector band of 2.7kb (**1**) that should not delete, a 5' band of 1kb(**2**) that should only delete after the 3' 800bp (**3**) fragment. The size marker is a 1kb ladder from Gibco-BRL.

Deletion plasmids were selected from each deletion series to provide sequencing prime sites 200-250bp apart along the cDNA insert. These clones were then used to prepare sequencing grade template (methods section 2.2.4).

4.4 Sequencing of ExoIII Deletion Clones

Sequencing template was prepared and reactions performed, exactly as described in the methods section 2.2.4 and 2.13. Sequence was digitised using a Gelreader and Macvector software (IBI) and into the Wisconsin-GCG sequence analysis package running on a DEC-VAX computer. The sequence runs were then aligned using the GCG Gel Assemble program (see methods section 2.22.7). The output alignment was used to error-check the final sequence against the original autoradiographs. One sequencing gap remained in one strand, and this was "closed" using a single 3' pointing oligonucleotide primer. This oligonucleotide was designed and constructed as described in methods section 2.21 Figure 4.8 shows the final sequencing runs used to construct the full length sequence of the pST51 insert.

4.5 Analysis of Sequence

4.5.1 General Sequence Features

The full length sequence was 1309 bp long excluding the 3' 16bp polyA tail remnant. The sequence is shown in figure 4.9. The GCG Map program was used to predict the restriction map of the cDNA; it showed very close correspondence to that obtained using restriction enzymes.

The cDNA contains a 311 AA open reading frame producing a predicted polypeptide that shows homology to the β-subunit of the Na+/K+ ATPase from a variety of organisms. This reading frame is bounded at the 5' end by a potential ATG start site at DNA residue 182-185 (see figure 4.9). Several inframe stop codons are located further upstream of the potential start codon. The consensus sequence flanking translational start sites has been reported to be (C/A) A A (A/C) **A T G** (Calvener 1987). The sequence of the predicted start site, T A A A **A T G** is in good accordance with this consensus. Indeed, the average match of a given start site to the consensus is reported by Calvener to be 3.1 out of 4 of the upstream residues. It is interesting to note that another potential start site exists at AA residue 16. It has the sequence G A A A **A T G**, again with good match to the consensus. This raises the interesting possibility that multiple peptides could be produced from this single transcript.

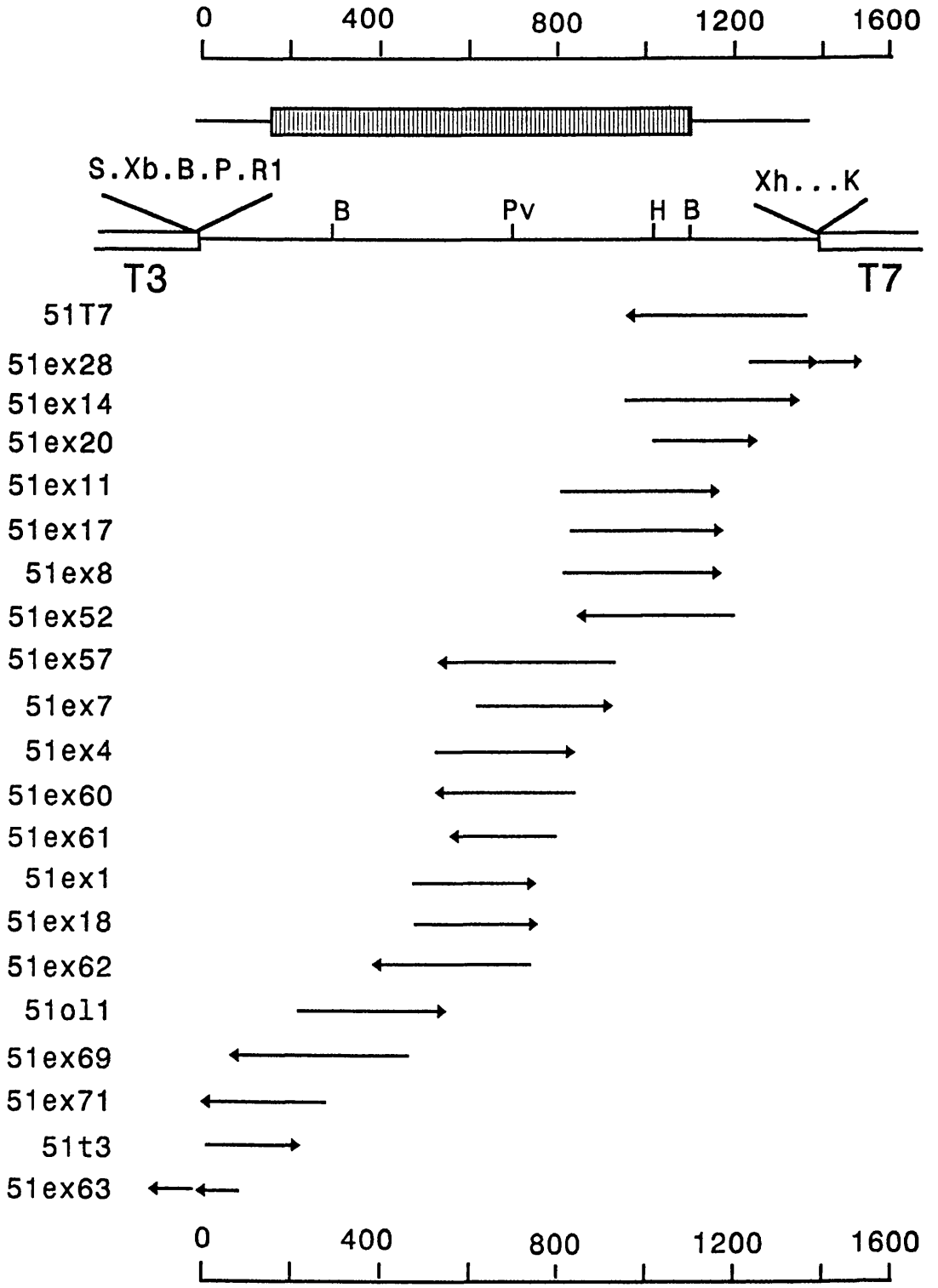


Figure 4.8 Sequencing Extensions Used to Construct the Full Length Sequence of the pST51 cDNA Insert

Arrows indicate the direction of a sequencing run. Plasmids used are labelled along the left edge of the figure, the sequencing primer used was the vector based T3(5' to 3') or T7(3' to 5') unless otherwise indicated. A single oligonucleotide primer was used to prime one sequence extension (see main text); this sequence is labelled pST51ol1. The restriction map and coding region is shown above for comparison. The coding region is shown as a hatched box. Scale is in basepairs. Broken arrows indicate those runs that cross the cloning junction.

```
                CCCCCGGGCTGCAGGAATTCGGCACGAGTTAAATGCAAGTCACAACAATACTATGGTTACA      33

GACGTTTAAAAATATATAAAAAAAATTTTGCACCAGAACAGTTGATAAAAGAAACTTAAAATAACGACTT     103

AAAACAAAAATAAATTCGAAATAATTCAACTTGAAAGACAATACGCCCATCAACACACTACAAAGTTAAA     173

ATGGCCGATAAAAAAATTGGTGAATACTATGCACCACCTGTGAAAATGGGCAAATGGGAGGGTTTCAAAA     243
M  A  D  K  K  I  G  E  Y  Y  A  P  P  V  K  M  G  K  W  E  G  F  K  K       24

AATTCCTATGGAACAGTGAAACTAGCCAATGCCTTGGACGCACCGGATCCAGTTGGGCGAAAATTCTCCT     313
   F  L  W  N  S  E  T  S  Q  C  L  G  R  T  G  S  S  W  A  K  I  L  L       47

ATTTTACATAATTTTTTATGCGGCGTTAACTGGTTTTTTTGCTGCAATTTTCACTGTATTTTATCAAACT     383
 F  Y  I  I  F  Y  A  A  L  T  G  F  F  A  A  I  F  T  V  F  Y  Q  T         70

TTGGACAATGAAAAGCCAAAATGGATGCTTGACAATGGTTTGATAGGGTCCAACCCAGGTCTAGGCTTTC     453
L  D  N  E  K  P  K  W  M  L  D  N  G  L  I  G  S  N  P  G  L  G  F  R       94

GACCAATGCCACCGGAAGCGAATGTTGAGAGCACATTAGTTTGGTACGAGTCATCAAAGAAGGATAACTA     523
   P  M  P  P  E  A  N  V  E  S  T  L  V  W  Y  E  S  S  K  K  D  N  Y       117

TAAGTACTGGGTGGACGAGACTTCACGTTTTCTGAAATCGTACCAAGATCTTGAGAAGCAAAATCAAGTG     593
 K  Y  W  V  D  E  T  S  R  F  L  K  S  Y  Q  D  L  E  K  Q  N  Q  V         140

AACTGTAGCTTTGAACATCCACCACAAGACGACAAGGTCTGCGGCATTGACTTCTCCAGCTTCTCGCCAT     663
N  C  S  F  E  H  P  P  Q  D  D  K  V  C  G  I  D  F  S  S  F  S  P  C       164

GTACAGCTGACAACAACTTTGGCTACCACGTTGCCCGGCCATGTATATTCCTTAAGTTGAATAAGATTTA     733
   T  A  D  N  N  F  G  Y  H  V  A  R  P  C  I  F  L  K  L  N  K  I  Y       187

CAATTGGATACCAGAAATTTACAACGATTCTAAGACTTTGCCAGATCATATGCCAGAGGAACTAAAACAG     803
 N  W  I  P  E  I  Y  N  D  S  K  T  L  P  D  H  M  P  E  E  L  K  Q         210

CACATCAAGGAAAAGCAAAGTCTTAGGCCCAATGAAACAAATGTAGTTTGGGTCTCGTGCGAGGGAGAAA     873
H  I  K  E  K  Q  S  L  R  P  N  E  T  N  V  V  W  V  S  C  E  G  E  N       234

ATCCCGCTGATGTCGAGAACATAAAAGCACGCGACTACTATCCCCGAAtGGGATTTCCTCGTTACTATTT     943
   P  A  D  V  E  N  I  K  A  R  D  Y  Y  P  R  M  G  F  P  R  Y  Y  F       257

TCCGTTTAAAAACATTCAGGGATATATACCGCCCATTGTTGCTGTTCAATTTACCGTTGAAACCGGCGTT    1013
 P  F  K  N  I  Q  G  Y  I  P  P  I  V  A  V  Q  F  T  V  E  T  G  V         280

TTGATCAACATTGAATGTAAAGCTTGGGCCCGCAACATTAATCACGACCGTTCAGACAGAAGAGGATCCG    1083
L  I  N  I  E  C  K  A  W  A  R  N  I  N  H  D  R  S  D  R  R  G  S  V       304

TTCACTTCGAGTTGATGGTTGATTAAGAGAGTCGTTGGGAAGATGTCATTGGAGAGAAGAGATCGGCCGT    1153
   H  F  E  L  M  V  D  <                                                    311

ATTATAAAGTGGGCATCAAAAAATAAATGAGAAAATAAAATAAAGGAATGGAATATAGAAAAAAAAAACT    1223

GACAACAAAACCCACACAGAATAGATACGAGAAAAATGTAGTAGCGCAGAAACGTCAGGTTAAAGCAAAA    1293

ACAAAAACAAAAAAAAAAAAAAAAACTCGAGGGGGGGCCC    1317
```

Figure 4.9 Full Length Sequence of pST51

The vector sequences flanking the cDNA insert are shown underlined. The 5' vector based *Eco*RI site and 3' vector based *Xho*I site are shown in bold. The DNA sequence is numbered starting from the most 5' base of the cDNA, and ending the last base of the polyA tail. The vector sequence is not numbered. The predicted amino acid sequence is shown aligned to the DNA sequence and numbered from the predicted start methionine.

The shorter polypeptide would lack the N-terminal region which is well conserved in evolution (see figure 4.12) between family members and produce a smaller intracellular domain than reported for any other ATPase.

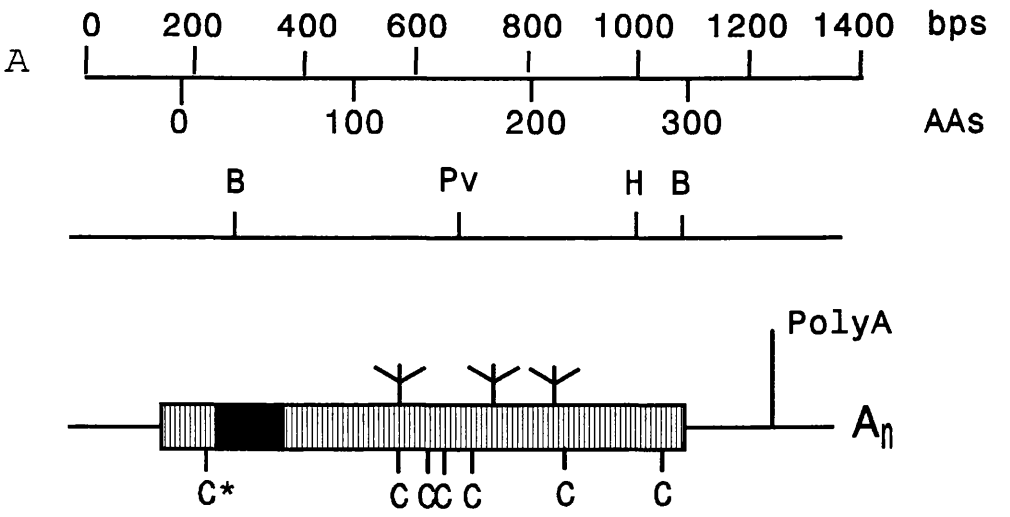
The ORF is terminated at the 3' end by a TAA stop codon encoded by DNA residues 1105-1108 (figure 4.9). The remainder of the cDNA, some 202bp, is presumed to be 3' untranslated region. The putative 3' untranslated region does not contain a consensus polyA addition signal within 100bp of the polyA tail. However, three clustered consensus sequences are located starting at residues 1182, 1194 and 1198bp (figure 4.9). These may be the signals used. Interestingly, the most closely related subunit, that of the brine shrimp, also lacks a polyA signal close to the 3' end but also has a polyA consensus site 100bp upstream (Bhattacharyya *et al.* 1990). It is also certainly possible that the β-subunit mRNA uses a non-standard polyA signal. The cDNA could also be internally primed from an A rich region of the mRNA; a possibility returned to in section 4.7.1.

4.5.2 Predicted Amino Acid Sequence Features

As described in section 4.1, all β-subunit family members possess characteristic protein motifs; they all are proposed to span the membrane once close to the C-terminal end of the protein, all possess three pairs of conserved cysteine residues in the cytoplasmic (C-terminal) domain and all possess GSN glycosylation consensus sites also in the cytoplasmic domain. A combination of multiple sequence alignment (GCG Pileup, methods section 2.22.11) a protein motif searching program (Mac-Pattern, methods section 2.22.12) and a Kyte-Doolittle Hydropathy plot program (GCG Pepmap and Pepplot, methods section 2.22.12) were used to identify sub-sequence features of predicted polypeptide. The results of the analysis are shown in figure 4.10.

From the figure, it is clear that the pST51 polypeptide, henceforth called Dmbeta, possesses characteristic features of a Na+/K+ ATPase β-subunit gene. The possession of these functionally significant motifs strongly argues that Dmbeta is not only evolutionally related to the β-subunit genes, but could act as a functional β-subunit.

In the next section I will describe the multiple alignment of β–subunit super-family members. I will then return in more detail to the functional significance of the motifs described in this section.



B

Motif	Peptide matching Residues	Probablity of occurence	Identified using
ATPase beta-1	22-43	8.4x10-11	Mac Pattern
Transmembrane	44-74	NA	GCG pepmap and plot, GCG motifs
Asn glycosylation	141-144, 195-198, 221-224.	4.5x10-3	Macpattern, GCG motifs
Cysteine residues	34, 142, 154, 164, 178, 230, 286	0.05	by eye!
*cAMP phopho site	300-303	**0.01 0.013**	Macpattern
*CK2 phopho site	112-114, 130-133, 199-202, 313-316, 350-353, 359-362.	**0.01 0.014**	Macpattern
*PKC phospho site	111-113, 112-114, 124-126, 217-219, 298-301	**0.01 0.012**	Macpattern

C

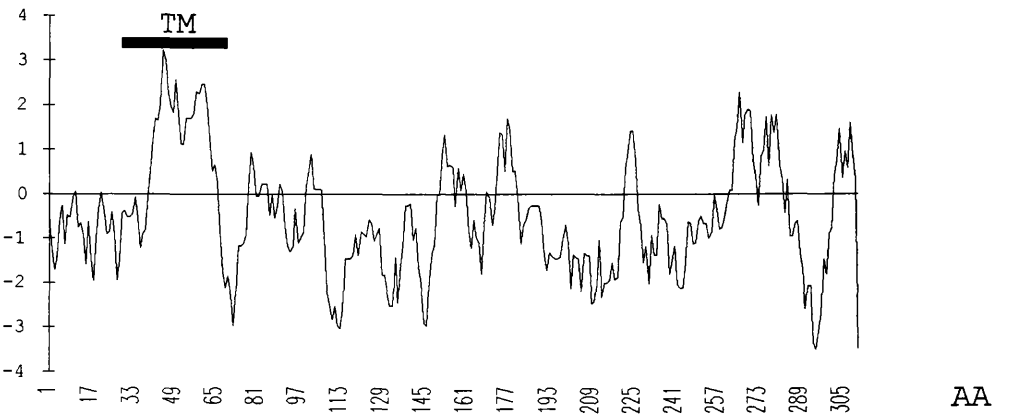


Figure 4.10 Sub-sequence features of pST51

A shows the pST51 cDNA with the predicted protein coding region in bold. The transmembrane region is blacked, branch structures mark the position of potential glycosylation sites, and **C**s indicate cysteine residues. The star indicates a residue that is not conserved between classes of β-subunit. DNA restriction map and scale are shown for comparison and labelled as figure 4.4.

B shows the petide sequence features used to construct figure 4.10a. Probability of a particular consensus occuring by chance is shown. Bold figures indicate values corrected for AA composition of this cDNA.

C show a plot of hydopathy, calculated using the Kyte and Doolittle 1982 method. The region labelled **TM** is likely to span the plasma membrane (once). Scale is in basepairs (bps) and amino acid residues (AA).

4.5.3 Comparison of Dmbeta to Other Family Members

Database searches were performed using the Dmbeta sequence through both Swissprot and NBRF protein databases using the FastA algorithm (methods section 2.22.4). A search was also performed using the amino acid sequence though the combined GenEMBL database and using the TFastA program. From these searches a set of thirty β–subunit family members was identified. Any DNA sequences were translated and the whole set of sequences aligned using the GCG Pileup program. From this set, representative proteins from each class were selected. Classes were Na+/K+ ATPase β1, β2, β3, and an unassigned β-subunit from shrimp and several proton pump β-subunits. This more manageable set was then input into Boxshade (methods section 2.22.13) and this program used to produce figure 4.11.

A strong N-terminal region of conservation can be seen co-incident with the transmembrane spanning region (shown underlined in the figure). As described in the introduction to this chapter, this region is also implicated in the translocation of the pump complex form the Golgi to the cell surface in other subunits. Several residues, for example the GRT peptide, glycine, arginine and threonine, at positions 36-38 is conserved in all known subunits including Dmbeta. Overall, the level of amino acid conservation is rather low, Dmbeta is just 29% identical to the nearest mammalian subunit. However, Dmbeta is 48% identical to the only other arthropod sequence cloned, that from *Artemia*, the brine shrimp. A number of residues show conservation between these two subunits and no others. For example, the peptide MADKK at the N-terminus is conserved between this pair of ATPases and no others. The conservation of these N-terminal residues is also strongly supportive of the independently predicted Dmbeta translation start site. A similar argument can be applied to the termination site.

Figure 4.12 shows the alignment of the brine shrimp and Dmbeta polypeptides with the identical residues, transmembrane region, and conserved cysteine residues shown for reference. Overlaid upon these are the predicted glycosylation sites and several different consensus phosphorylation sites. Since the two peptides are 49% identical it might be expected that regions important in regulation/function would be conserved between subunits. Dmbeta contains a single consensus cAMP dependant kinase site, but this is not found in the other amino acid sequences. Dmbeta also contains five PKC

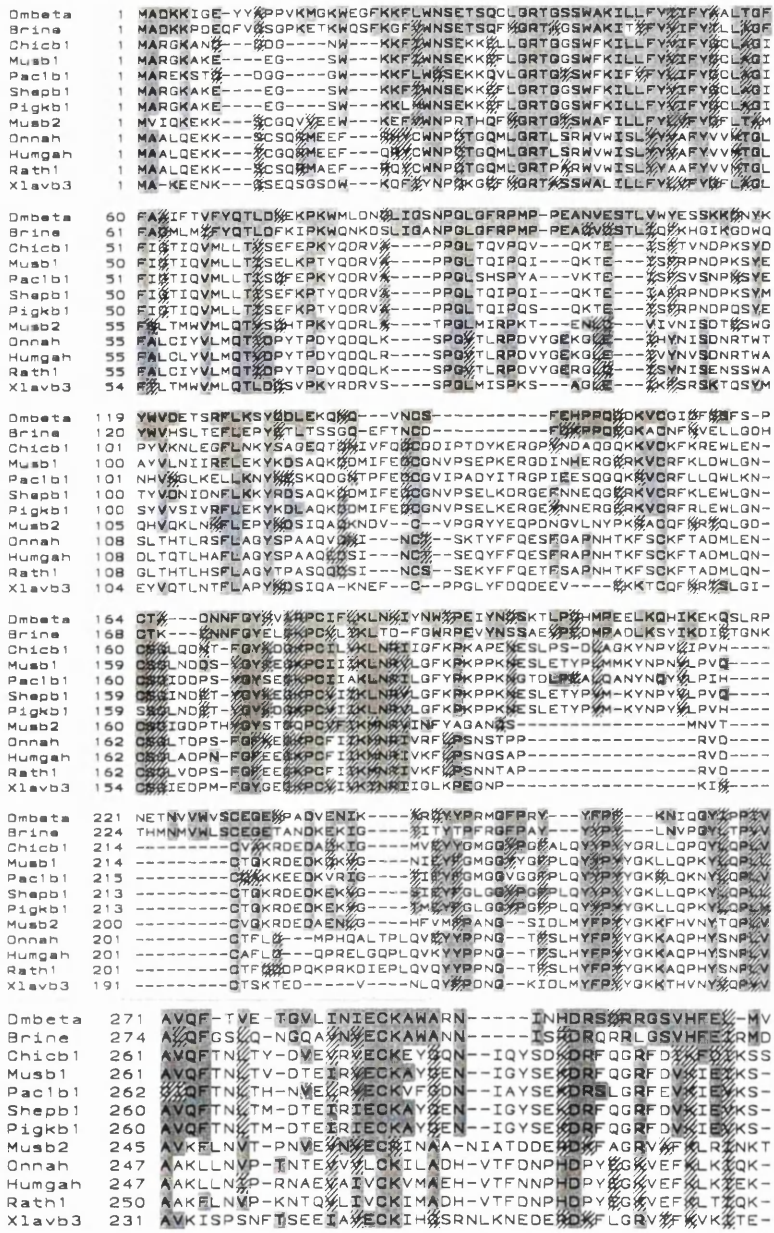


Figure 4.11 Multiple Alignment of Representative ATPase

The multiple alignment was produced as described in methods section 2.22.13. Residues identical to Dmbeta at a particular position are blacked, and those that are similar are hatched. The proteins used were from *Drosophila* (**Dmbeta**), brine shrimp (**Brine** from Bhatttacharyya *et al.* 1990), chick β1 (**Chicb1** from Takeyasu *et al.* 1987), mouse β1 and β2 (**Musb1** from Gloor 1989 and **Musb2** from Gloor *et al.* 1990), pacific ray β1 (**Pac1b1** from Noguchi *et al.* 1986), sheep β1 (**Shepb1** from Shull *et al.* 1986), pig kidney β1 (**Pigkb1** from Broude *et al.* 1987), rabbit gastric pump β-subunit (**Onnah** from Reuben *et al.* 1990), human gastric (**Humgah** from Ma J. y. *et al.* 1991), rat gastric (**Rath1** from Shull 1990) and *X.laevis* β3 (**Xlavb3** from Good *et al.* 1990).

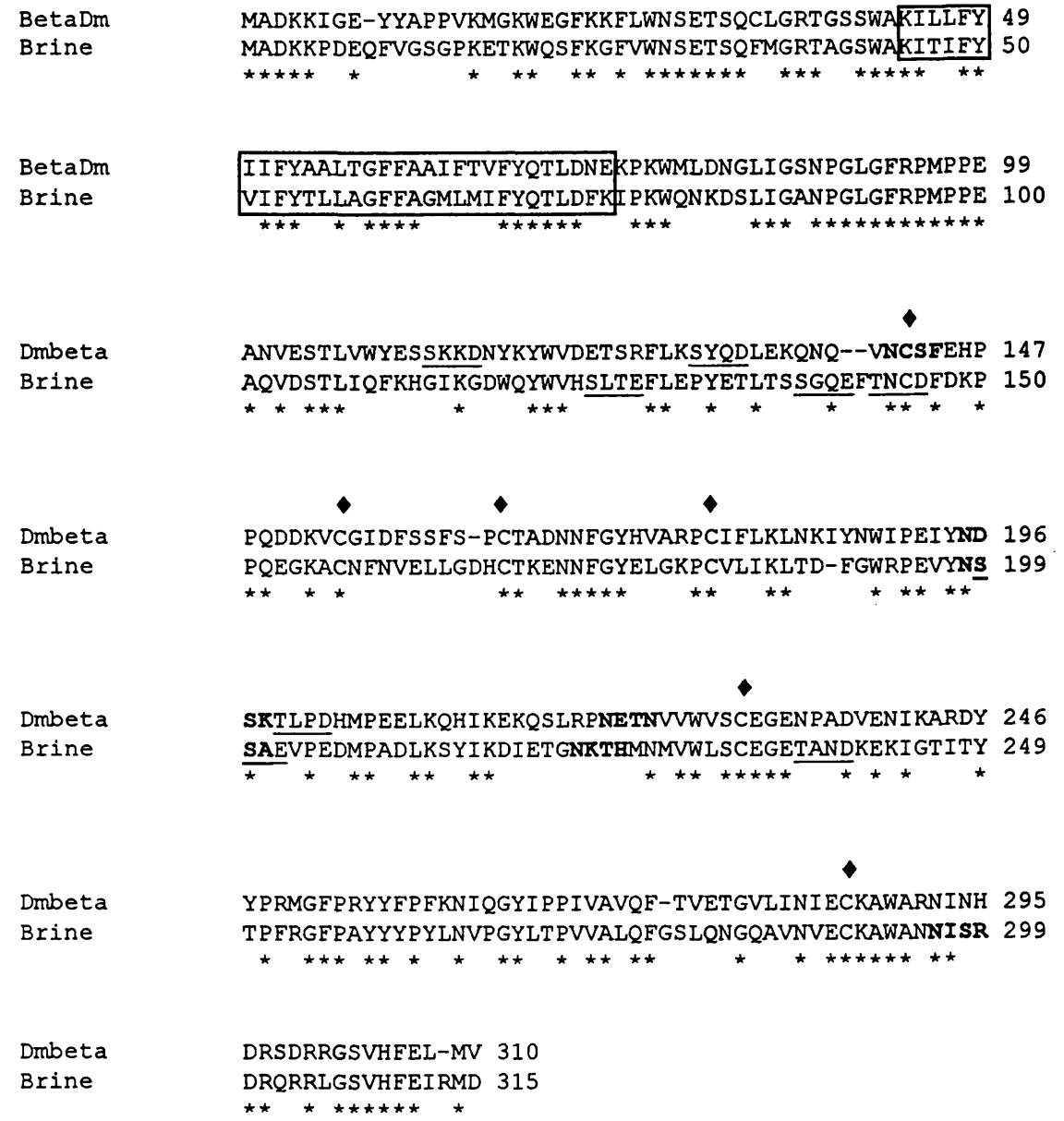


Figure 4.12 Alignment of β-Subunits from *Drosophila* and Brine Shrimp Showing Subsequence Features.

The multiple alignment was prepared using ClustalV (section 2.22.12). The *Drosophila* subunit is labelled Dmbeta and the shrimp subunit brine (peptide sequence as reported Bhattacharyya *et al.* 1990). Stars indicate identity at that particular position. Transmembrane region is shown boxed. Consensus glycosylation sites are shown in bold, while the six conserved cysteine residues are shown labelled by the symbol ♦. Underlined sequences are consensus PKC sites. Gaps in the sequence alignment are shown dashed.

consensus sites while the brine sequence contains a single site that does not match in location to that found in Dmbeta. However, there is a weak correspondence in location of the PKC sites as shown. The Na+/K+ ATPase is known to be under protein kinase regulation in mammals since alterations in levels of cAMP or PKC cause alterations in pump activity in mammals (Marver *et al.* 1986, Smart and Deth 1988). However, the α-subunit is thought to be the modified subunit, (Bertorello *et al.* 1991, Ling and Cantely 1984, Chilbalin *et al.* 1992). However, in Ehrlich ascites tomour cells it has been suggested that the β-subunit is regulated by phosphorylation (Spector *et al.* 1981). When considering potential phosphorylation sites it must always be considered that the potential target sites occur often by chance in a random sequence.

Potential predicted glycosylation site patterns are relatively constant within subunit groups. For example, the location and number of glycosylation sites are conserved between all the mammalian Na+/K+ ATPase β1-subunits, but do not match well the six consensus GSN glycosylation sites found in Na+/K+ ATPase β2-subunits. Both *Drosophila* and brine shrimp subunits possess three consensus sites and as shown in figure 4.11 two of these sites match very well in location. The third *Drosophila* site is situated over a conserved cysteine residue; this could hinder glycosylation at this consensus site. The two matching pairs of glycosylation sites are not located within well conserved regions of the molecule, either when considered as DNA or protein sequence. Thus the preservation of both site pairs is not a trivial consequence of the consensus sequence N X (S/T) occurring by chance and being preserved through evolution by being located in a region conserved for an unrelated reason. The evidence from the functional studies of other subunits, plus the preservation of site location and number between the *Drosophila* and shrimp, strongly suggests that these sites are of functional significance. Interestingly, Peterson and Hokin 1980 have shown that the native shrimp β-subunit is glycosylated and Baxter-Lowe *et al.* 1988 have shown that in a cell-free translation system β-subunit can be glycosylated.

Another notable feature of figures 4.11 and 4.12 is that the alignment shows a number of insertion, and/or deletion events have taken place since the divergence of the family members. Indeed, even the alignment between shrimp and Dmbeta requires five small deletions. (The actual number and size depends somewhat on the gap and extension of gap weighting used by the

alignment program). To assess the extent of rearrangements in a less context sensitive way than multiple alignment, matrix similarity alignments were performed, using a modification of the method of Pustell and Kafatos 1982 (methods section 2.22.12), shown in figure 4.13.

When Dmbeta is compared to itself the single diagonal trace of capital A's in figure 4.13 indicate identity. The polypeptide shows no evidence of internal duplication, as shown by the absence of points outside the diagonal line. Similarly, there is little evidence of duplication events when Dmbeta is compared to the brine shrimp sequence. Interestingly, a very weak inverted region of homology (score 30% over 30 AA) can be detected between a portion of the transmembrane region and at 260-280 AA along the molecule. This is only evident when shrimp is compared to itself or to Dmbeta with one sequence in reverse , but not when Dmbeta is compared to the reverse of itself. This suggests that at least one inversion event has occurred in the shrimp line after the divergence from *Drosophila* (data not shown).

The similarity matrix comparison between Dmbeta and shrimp shows that the regions (as opposed to individual amino acid residues) of conservation are principally between amino acid 1 to 130 and 230-300. Of these regions, the N-terminal stretch, which includes the transmembrane section, is the most highly conserved. This general trend is reiterated when Dmbeta is compared to representative members of each sub-family. However, the conservation of the C-terminal domain becomes much less apparent for the H/K+ ATPases and for members of the β3 class of Na+/K+ ATPase.

4.5.4 A Phylogenetic Tree of the β-Subunit Gene Family

The Pileup and ClustalV multiple alignment data were used to construct a dendrogram, which is shown as figure 4.14a. The algorithm used to construct the dendrogram was not a genetic one and the figure should not be considered as a phylogenetic tree. The data were used to construct a phylogenetic tree using the neighbour-joining method of Saitou and Nei (1987) as implemented in the ClustalV program (methods section 2.22.12). This is shown as figure 4.14b. The phylogenetic tree is unrooted, i.e. does not display the direction of evolution, but the tree does show evolutionary distance. The significance of the branchings has been tested by using bootstrapping (Felsenstein 1985) and all branches are very strongly supported by the data except for the relative groupings of the three gastric subunits and that of the sheep, pig and mouse β1

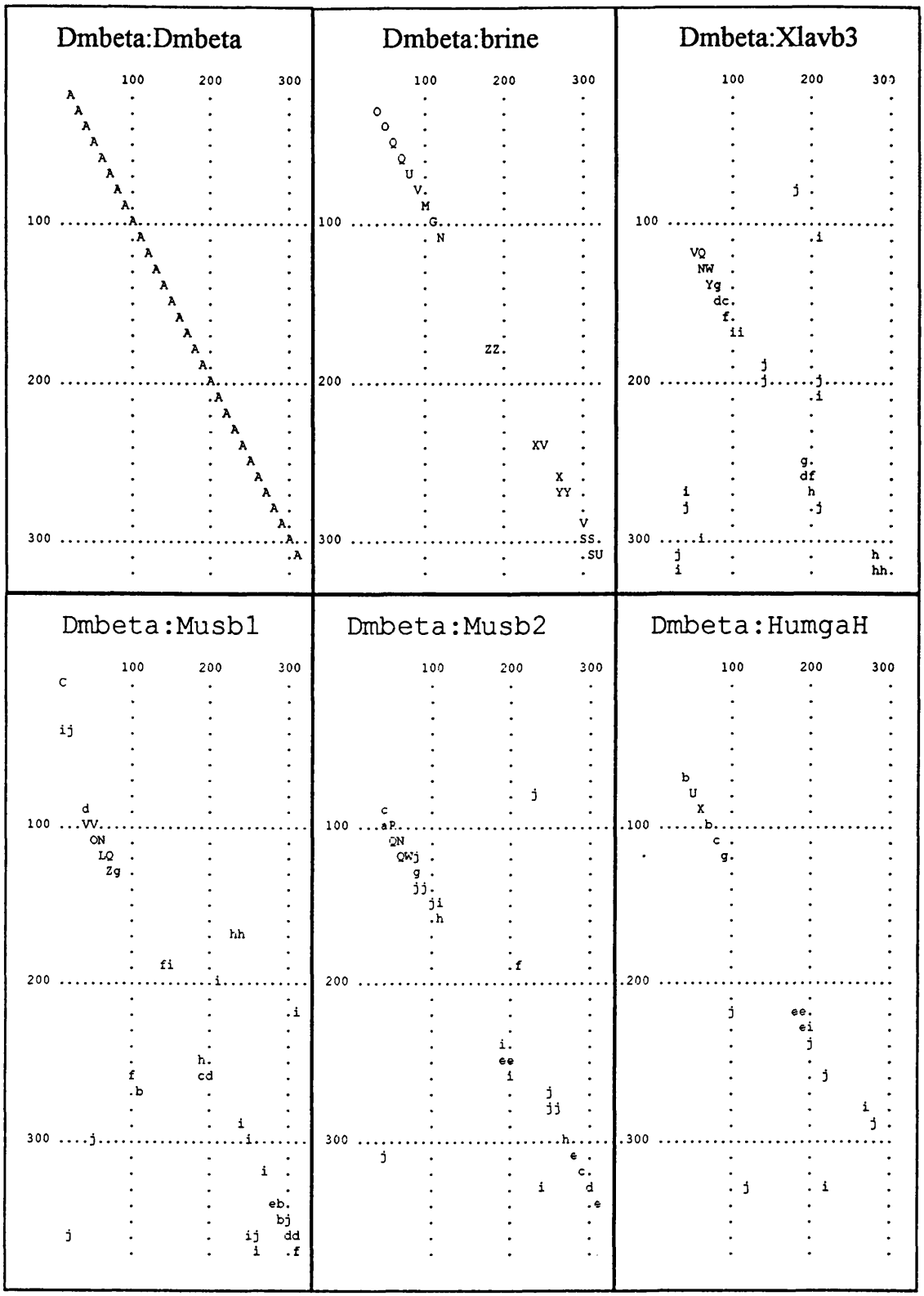


Figure 4.13 Pairwise Alignments of β-Subunits

Polypeptide sequence labels are as figure 4.11. Numeric values are amino acid positions from each start methionine. Scores are calculated using the Frinstensky *et al.* 1982 implementation of the Pustell and Kafatos 1982 method (methods section 2.22.12). Upper case letters represent scores ranging from 100% (**A**)to 50%(**Z**), lowercase range from 48-49% (a) to 30-31%(j). Scores are averaged over a range of 10 amino acids.

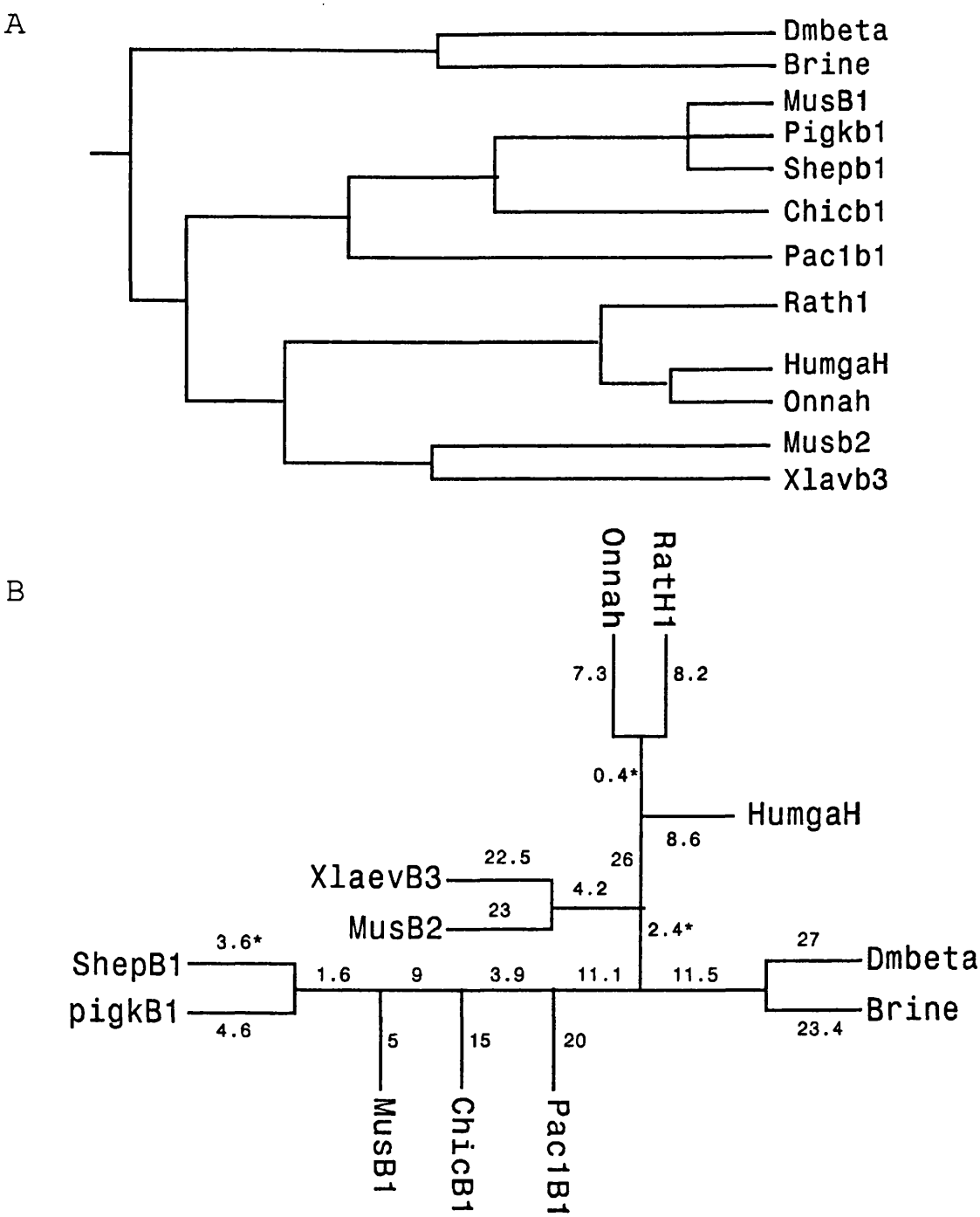


Figure 4.14 A Dendrogram and Phylogenetic Tree of ATPase β-subunits.
A shows a dendrogram for the sequence relationship between the different aligned subunits as shown in figure 4.11. It was generated using the GCG Pileup program (methods section 2.22.12).
B shows an unrooted phylogenetic tree generated using the ClustalV program (methods section 2.22.13). Numeric values are percentage divergence, stars indicate those branchpoints that are not strongly supported by the data as calculated using bootstrapping. These values are not corrected for multiple substitutions. The subunits used are *Drosophila* (**Dmbeta**), shrimp (**brine**), chick β1 (**Chicb1**), Mouse β1 and β2 (**Musb1** and **Musb2**), Pacific ray (**Pac1b1**), Sheep β1 (**Shepb1**), pig kidney β1 (**Pigkb1**), rabbit gastric pump β-subunit (**Onnah**), human gastric (**Humgah**), rat gastric (**RatH1**) and *X.laevis* β3 (**Xlavb3**). References for these subunits are as described for figure 4.11.

subunits. The phylogenetic tree can be rooted by placing the root along the longest branch or on biological grounds. I have avoided rooting the tree along it's longest branch (and so assigning the *Drosophila* sequence as ancestral) because I have constructed further trees which employ a distance correction of Kimura's (Kimura 1983) to correct for multiple evolutionary changes at the same site. In this tree (not shown) the gastric sequences could be considered as ancestral.

The alignment shown in figure 4.14 is consistent with the grouping of subunits described in previous papers. β1-subunits are all closely grouped, and within this group all the mammalian sequences are more closely related than the chick or ray subunits. Dmbeta and the brine shrimp subunit are placed as a group outside previously described classes, most closely related to the Na+/K+ ATPase β–subunit classes. It is more difficult to determine which sub-class the Dmbeta group most closely resembles.

Even though there is low level of homology with other β-subunits, studies on the brine shrimp subunit have conclusively shown that this subunit purifies with the α-subunit (Peterson and Hokin 1980). Therefore, although these two subunits look unlike any particular previously classified group, the brine shrimp by experimentation and Dmbeta by implication, are likely to encode Na+/K+ ATPase subunits. Indeed, the evolutionary distance between brine shrimp and Dmbeta is less than that between β1 and gastric subunits which have been shown to be functionally interchangeable (Horisberger *et al.* 1991b). Thus Dmbeta and the brine shrimp subunit define a new class of β–subunit. In suggesting that Dmbeta and the brine shrimp subunit form a new arthropod ATPase subunit group I imply that these two subunits are not only considerably diverged from the others but also that this divergence is associated with the selection for a related, but subtly novel function within this new group.

4.6 Head/ Body Northern Blot

Figure 4.15 shows a northern blot (methods section 2.9) produced by probing head and body polyA+ RNA with a ^{32}P labelled DNA probe generated from the whole insert sequence of pST51. The probe clearly identifies at least five transcripts at high stringency, ranging in size from about 1.6kb to about 6kb. A shorter exposure (not shown) of the blot fails to resolve the large diffuse band, and so this may contain a variety of transcripts. In addition, a longer

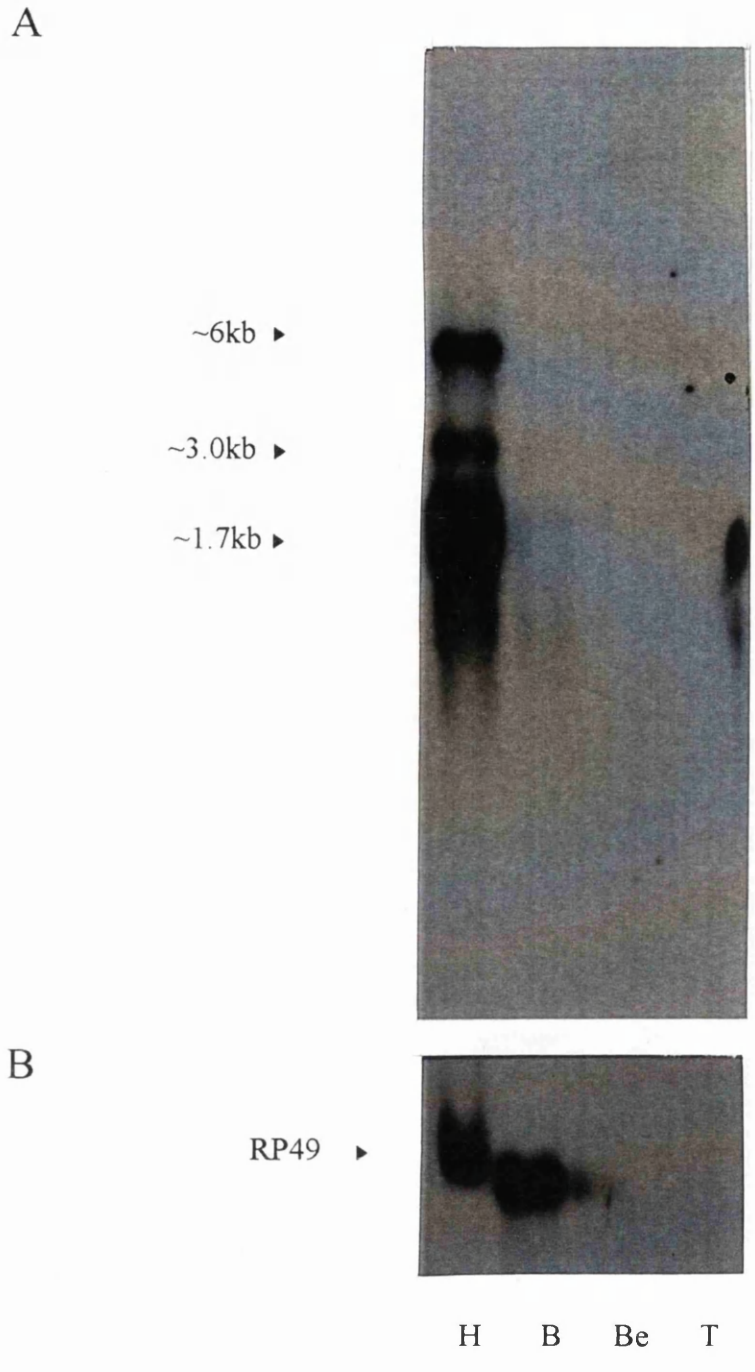


Figure 4.15 A Head/Body Northern Blot Using a Whole Insert Probe from pST51

The blot was produced as described in methods section 2.9. Each track contained an estimated 1-2μg of polyA mRNA. Lanes were loaded with wild-type head polyA+ (**H**), body polyA+ (**Bo**), *eya* polyA+ body (**Be**) and total RNA (**T**).
50ngs of the gel purified whole insert cDNA fragment from pST51 was labeled with 40μCi of incorporated activity of ^{32}P using the random priming method (methods section 2.10.1) with radiolabelled dCTP. The blot shown (**A**) is from a two week long exposure of the autoradiograph.
B shows the blot probed with an RP49 probe prepared as above as a control for lane loading, the head (**H**) and body (**Bo**) polyA lanes are approximately evenly loaded. Sizes shown are in base pairs, estimated from ribosomal bands and using the duplicate blot probed with pST123 as confirmation of the size estimates.

exposure indicates the presence of still further bands of several sizes above 6kb. These may represent larger transcripts or transcript processing intermediates. Hybridisation to the highly expressed bands will tend to obscure local rarer bands and so the estimate of transcript number is likely to be an under estimate.

The pST51 probe is also able to hybridise to body mRNA. However, by comparison to the rp49 (O'Connell and Rosbach 1981) control it can be seen that the relative abundance in head compared to body mRNA is of the order of 20 to 1 or greater. This is consistent with the selected expression pattern. It can also be seen from the blot that some of the transcript bands are not in the same proportions as seen in the head message. For example, the 3.4kb transcript is considerably under-represented in the body compared to the proportions of the other body transcripts. This is the first evidence that the multiple transcripts may be differentially regulated.

The presence of multiple transcripts is a common feature of Na+/K+ ATPase β–subunit. However, in this system it is especially interesting because the Na+/K+ ATPase α-subunit appears to be specified by a single gene which produces several transcripts in the head, but the data is consistent with these being produced by variation of the 3' untranslated region and not within the protein coding sequence. Thus in *Drosophila* much of the diversity of function furnished by the α-subunit variation is absent. Since Na+/K+ ATPases apparently serve analogous purposes between mammals and *Drosophila* it was interesting to look at the possible diversity facilitated by the various β–subunit transcripts.

4.7 Isolation of cDNA Clones Related to pST51

To isolate cDNA clones related to pST51 a conventional cDNA library screening approach was chosen. Several head cDNA libraries have been constructed in the laboratory or were available for use. The *eya* head library was again selected for this experiment, not only because of the ease of use of the vector system but also because it has proven to be the only library available that has given consistent, repeatable results. Even the NM1149 male head cDNA library from which the pST41 clone was isolated has given a number of unusual clones in other workers hands (C. Millagan and D. Harbison, Glasgow Genetics Department pers. comms.). The use of the *eya* library precludes the

isolation of any eye specific pST51 related cDNAs, but there is little long term interest in the laboratory to isolating such clones.

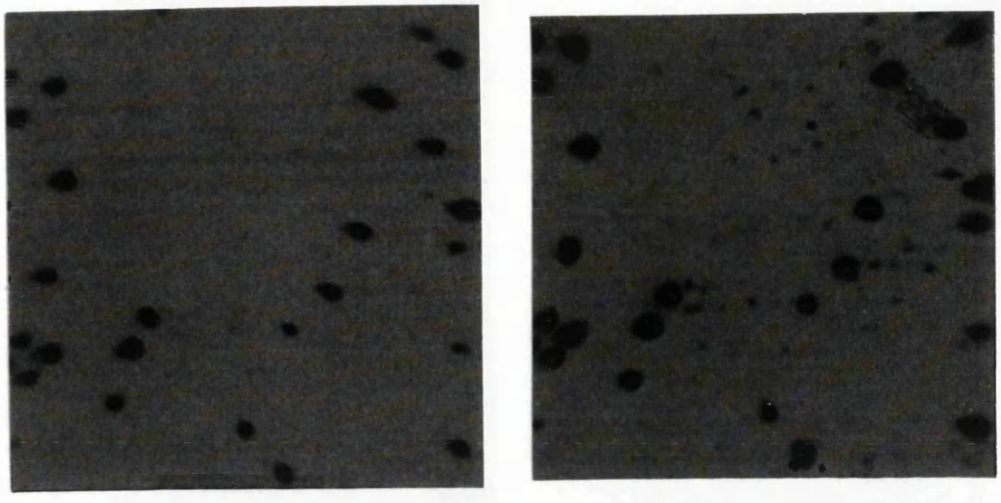
The cDNA screen was performed using a whole pST51 gel isolated insert (methods section 2.2.6) probe generated by random priming (methods section 2.10.1) in the presence of ^{32}P-dCTP. This probe was used to screen filters lifted from randomly plated ZAPII *eya* head library (methods sections 2.3.2, 2.14.2 and 2.14.4) and as shown in figure 4.16. Agar plugs containing positive clones from the secondary screen were used to produce plasmid borne cDNA clones (methods section 2.3.6). DNA from these colonies was digested to release the inserts and reprobed with the pST51 whole insert probe. Positive plasmids from this cycle of screening were again used to transform XL1-B *E.coli* cells. Single colonies were picked from this transformation were used for further analysis. Twelve independent positive clones were purified in this way; 50% of those selected at the primary screen.

The cDNA clones were restriction mapped using standard six-cutter restriction enzymes. Figure 4.17 shows the resultant restriction maps aligned about the common internal *Bam*HI cleavage site.

The inserts range in length from 0.8kb to 2.9kb. A 0.9kb *Bam*HI fragment is common to all but two cDNAs, pST51r46a and pST51r14. It appears that these two cDNAs are too short to encode a complete *Bam*HI fragment. These are likely to be the products of premature termination of the reverse transcription event rather than genuine full length cDNAs since sequencing runs from a vector based T3 primer fail to identify a suitable termination stop codon, see later in this section.

Of the remaining cDNAs, all appear of a size consistent with size range expected from the northern blot, once the polyA tail length is taken into account. The cDNAs all have very closely related restriction maps. I estimate (from the sequenced and restriction mapped pST51) that the maps are accurate to about 50bp. Hence small differences between cDNA restriction maps, e.g. between the placement of the paired *Hind*III and *Bam*HI sites may be artefacts of the mapping process. Several of the map differences were investigated further and the results of these are discussed in the next three subsections.

A



B

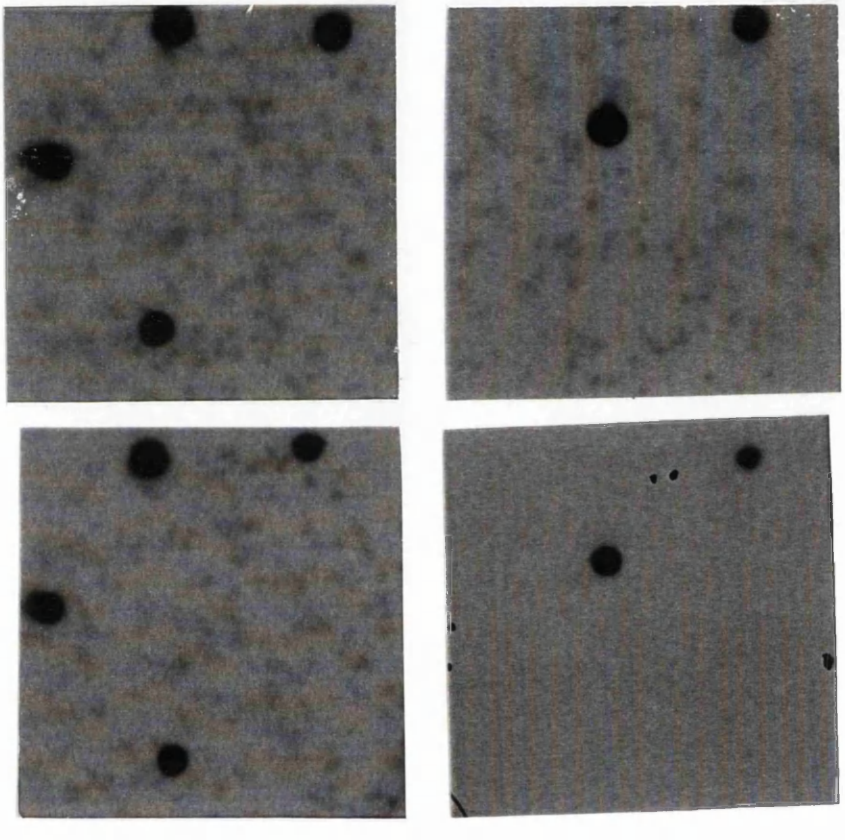


Figure 4.16 Screening a cDNA Library for Related Clones to pST51
Screening was performed as described in methods sections 2.3.2, 2.14.2, and 2.14.4.

A shows autoradiographs of duplicate filters probed with the pST51 whole insert probe. The plate contained about $1x10^4$ plaques. About 40 positive clones can be identified from this particular culture plate.

B shows autoradiographs of positives from the second round of screening. Only a proportion of the plaques are clearly positive at this stage because the large plug of agarose taken from the region around primary positive contains a number of different phage (about 100).

Secondary positives were rescreened by minipreping (methods section 2.2.4.1) after excision of the pBluescript II phagemid (methods section 2.3.6, data not shown).

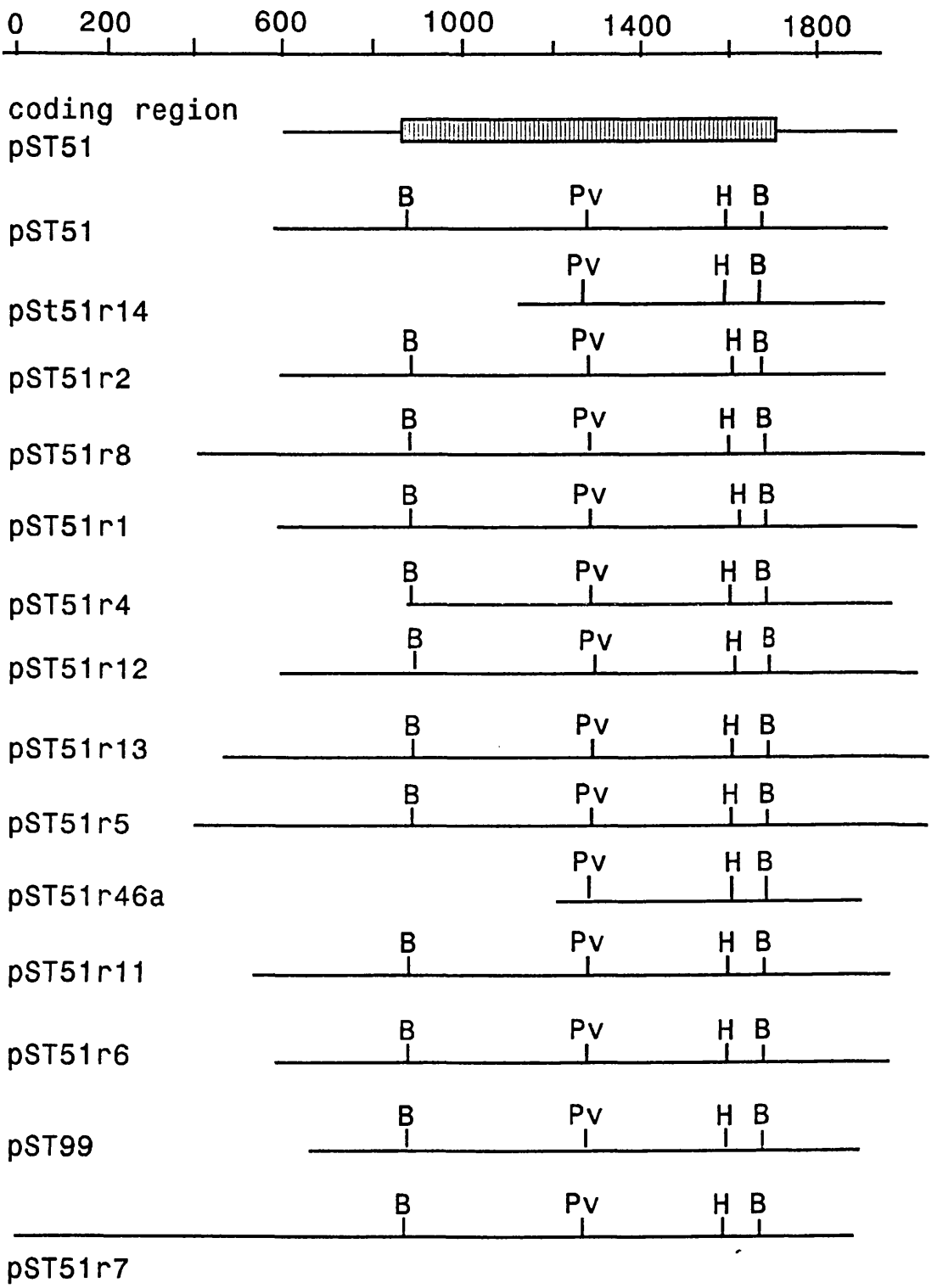


Figure 4.17 Restriction Maps of pST51 Related Clones
Restriction sites shown are **B**-*Bam*HI, **P**-*Pvu*II, **H**-*Hind*III and **B**-*Bam*HI. All plasmids were also digested with *Sac*I, *Xba*I, *Pvu*II, *Eco*RI, *Xho*I and *Kpn*I. The hatched region is the predicted polypeptide coding region. pST51 map data is in accordance with the restriction map predicted from the complete DNA sequence. Scale is in basepairs.

4.7.1 Multiple 3' Ends

Although polyA tail of mRNA may average several hundred basepairs, only a fraction of that length may be represented in a cDNA library This might be explained if there was competition for the prime site as shown in figure 4.18. Experience with the *eya* ZAPII library (section 3.9) suggests that the cloned polyA tail length in this library is very close to the length of the oligo dT primer used to construct the library. Hence differences in 3' end length of the pST51 related clones identified from the restriction maps are unlikely to be caused by variation in cloned polyA tail length alone.

Figures 4.17 and 4.19 suggests that the 3' ends of the cDNAs are of varying lengths, although clustered around a region between 200bp and 350bps of the 3' *Bam*HI site. This is clearly shown in figure 4.19a where the plasmids have been digested with *Xho*I and *Hind*III to liberate a single small 3' fragment presumably comprising the very 3' C-terminus of the predicted polypeptide coding region plus the 3' untranslated region and the cloned fragment of the polyA tail. The cDNA shows 3' ends that can be clustered into two loose groups; those of about 350bp and those of about 300bp. This variation in length could be due to variation in cloned polyA tail length and so to rule out this possibility sequencing runs were initiated on representative cDNAs from the vector based T7 primer. The results of this sequencing is shown in figure 4.19b.

The shorter 3' end, for example of pST51 and pST99, seems to be truncated versions of the longer 3' end of, for example, pST51r1. This truncation is within an A rich region of the longer 3' end, it seems possible that these are internal primings from the oligo dT first strand primer. The longer cDNA class does not introduce an extra polyA signal, although a near match (AATACA) occurs at -30 residues from the start of the presumed polyA tail.

It remains to be determined whether the observed difference in the 3'ends of the cDNAs is of biological significance. Certainly there have been a number of such 3' ends reported in other Na+/K+ ATPase β–subunit systems, often upon the basis of evidence such as described here. Since this experiment has highlighted this variation, further experiments can now be focused precisely on this issue. One useful technique would be to tail message from the head to introduce an anchor prime site that defines the very 3' end of the messages.

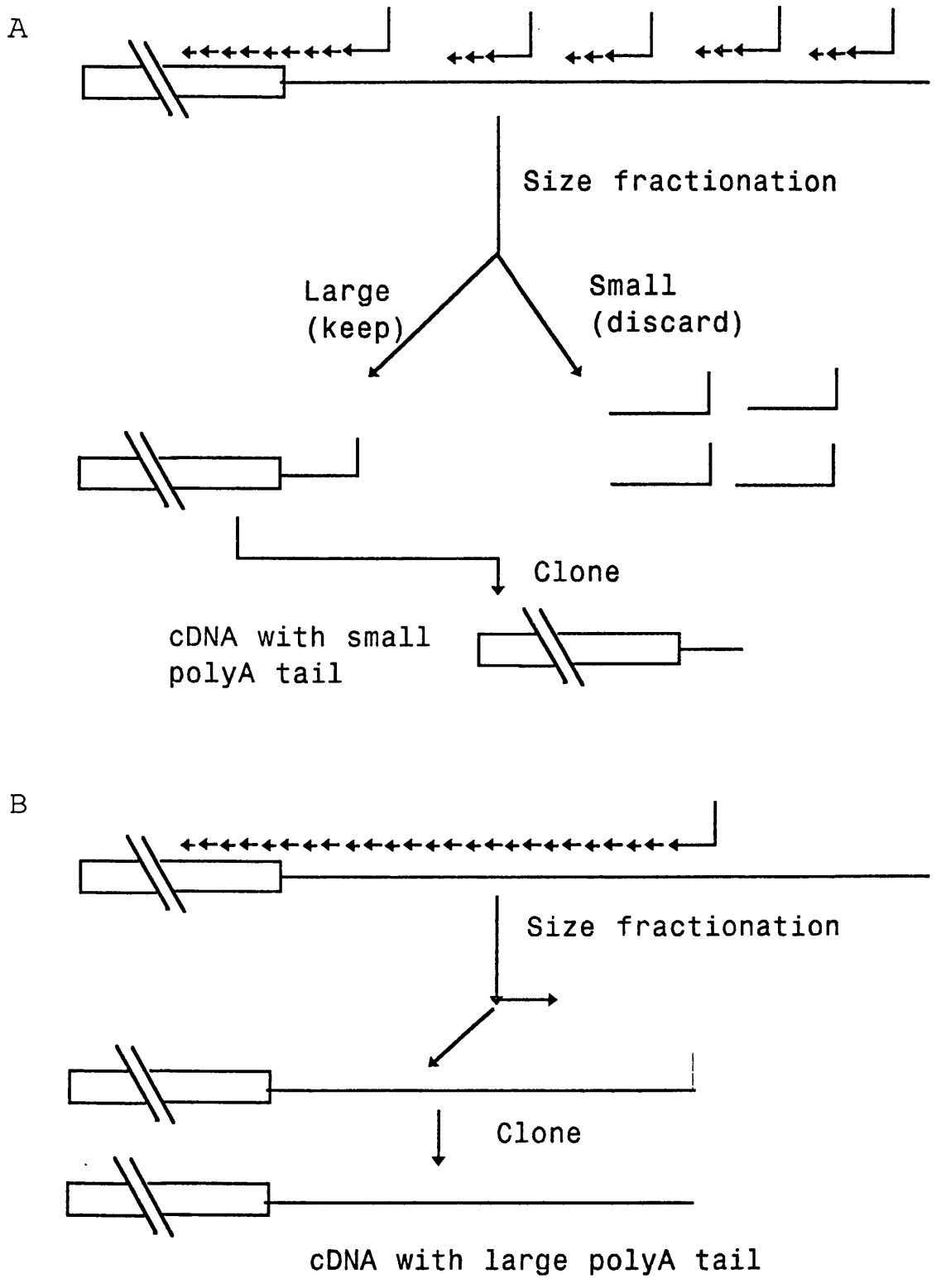
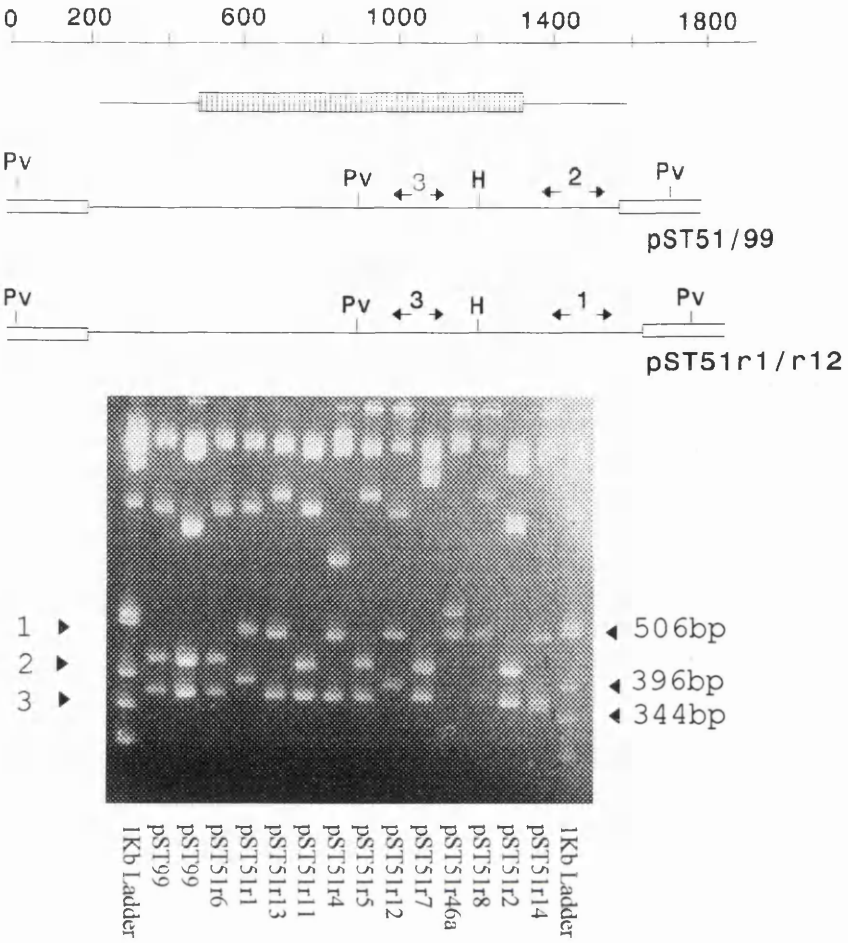


Figure 4.18 Competition for 3' Prime Sites
The schematic shows priming of the cDNA with the anchored oligo dT primer used in ZAPII library construction. The boxed region is the main body of the mRNA, the attached single line the polyA tail. **A** shows priming in the presence of excess primer, such that more than one oligo is able to prime per mRNA molecule. The effect is to produce the final cDNA from a prime site close to the junction of the polyA tail. **B** show what is expected to happen if the primers are able to prime once per molecule, producing a longer polyA tail than in **A**.

A



B

```
   pST51  ATAAAGGAAT GGAATATAGA AAAAAAAAAC TGACAACAAA ACCCACACAG
   pST99  ATAAAGGAAT GGAATATAGA AAAAAAAAAC TGACAACAAA ACCCACACAG
pST51r12  AAATAAGGAA TGGATATAGA AAAAAAAAAC TGACAACAAA ACCCACACAG
 pST51r1  AAATAAGGAA tGGATATAGA AAAAAAAAAC TGACAACAAA ACCCACACAG

   pST51  AATAGATACG AGAAAAATGT AGTAGCGCAG AAACGTCAGG TTAAAGCAAA
   pST99  AATAGATACG AGAAAAAtGT AGTAGCGCAG AAACGTCAGG TTAAAGCAAA
pST51r12  AATAGATACG AGAAAAATGT AGTAGCGCAG AAACGTCAGG TTAAAGCAAA
 pST51r1  .ATAGATACG AGAAAAA GT AGTAGCGCAG AAACGTCAGG TTAAAGCAAA

   pST51  AACAAAAACA AAAAAAAAAA AAAAA*
   pST99  AACAAAAACA AAAAAAAAAA AAAAAA*
pST51r12  AACAAAAACA AAAAAAAAAC AAAAAACGAA CGTGCGAAAT ACAGCAGCAA
 pST51r1  AACAAAAACA AAAAAAAAAC AAAAAACGAA CGTGCGANAT ACAGCAGCAA


pST51r12  CCAAAGAGAA AACTAACAAA AAAAAAAAAA AAAAA
 pST51r1  CCAAAGAGAA AACTAACAAA AAA
```

Figure 4.19 3' Termini Differences Between pST51 and Related Clones
A shows a partial restriction map showing each pST51 related clone digested with *Pvu*II and *Hind*III. The line diagram showing the origin of the fragments labelled **1**,**2** and **3** and is labelled as figure 4.17. The 1kb ladder marker was from Gibco-BRL. Size is in basepairs.
B shows sequence data at the 3' termini of selected clones. pST51 and pST99 have the smaller fragment **2**, pST51r1 and pST51r12 have the larger fragment **1**. The difference can be accounted for by an additional 60bp at the 3'end of each of the longer clones. The cloned polyA tail sequence is underlined and a star indicates those clones that have been sequenced over the cloning junction. The pST51r1 and pST51r12 sequences lack vector junction points but the position of the polyA tail is inferred from the distance to the sequencing prime sites.

The specific 3' termini can then be studied using a gene specific and the synthetic 3' terminal prime site using a PCR cloning strategy.

4.7.2 Multiple 5' Ends

From figure 4.17 it is clear that the cDNA clones show a wide variety of 5' termini. This is further demonstrated in figure 4.20a which shows a restriction digest of each clone with *Bam*HI. It is im[1]portant to be able to distinguish between biologically significant termini and the results of the cloning process itself. Several 5' termini were sequenced and show sequence consistant with being shorter versions of pST51 (data not shown). However, one sequence obtained from pST51r1 possesses an entirely different 50bps of sequence at the very 5' end of the cDNA when compared to pST51. This is shown in figure 4.20b. The unique 5' sequence could not arise from premature termination of the reverse transcription reaction, and is unlikely to have arisen by other rarer events such as the ligation of two cDNAs together at their common *Eco*RI sites. (Although the pST51r1 5' terminus is non-standard). Also, the extra 50bp is unlikely to be two clones ligated together because the cDNA library was size selected for fragments above 500bp prior to ligation. The unique sequence at the 5' end does not match any other sequenced parts of the pST51 cDNA group, or match anything in the published sequence databases (data not shown).

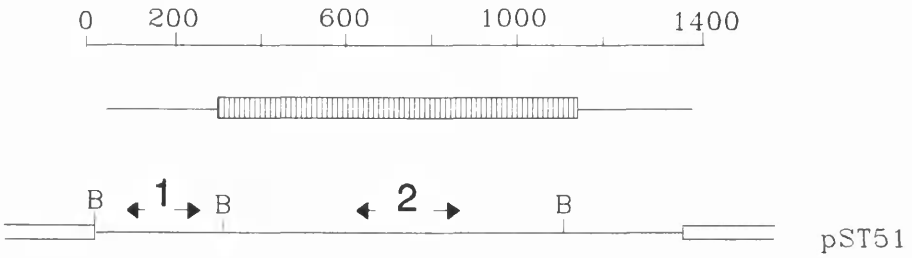
4.7.3 A Single Subtle Difference in the Coding Region

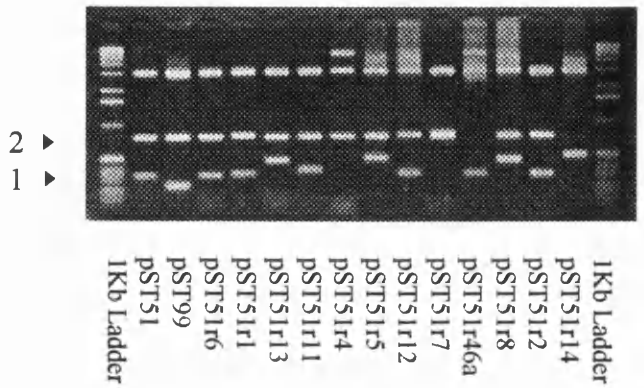
The analysis described here so far suggests that the cDNAs share common internal sequence and differ at both their 5' and 3' ends in a way that might generate a number of different cDNA species carrying a common core predicted amino acid coding sequence. No attempt has been made to predict the translation start and stop sites represented by the different cDNAs. However, during the simple restriction mapping a subtle difference in fragment size was observed. This small change can be seen in the reproduced gel photograph already shown as figure 4.19. From the figure it can be seen that band 3 is of a different size in pST51r1 and pST51r12 than in other clones. Since band is generated from coding sequence in pST51 then pST51r1 and pST51r12 must be copies of mRNA species that have the capacity to encode a protein distinct to that represented by pST51. The size difference was further localised using restriction mapping (data not shown), and suitable sequencing primers designed to sequence across the altered region in both strands. These

---

1

A





B

```
pST51r1   CCCCCGGGCTGCAGGTCTGGAGGCTAAA
          ||||||||||||||
pST51      CCCCCGGGCTGCAGGAATTCGGCACGAG
```

```
pST51r1 .CTATGTACTAGTAGAGGGGAAGCGTTTGAAGACGAAAAAACTTTCATTCTACGTTTAAA
                                                        |||||||||
pST51   .........GGCACGAGTTAAATGCAAGTCACAACAATACTATGGTTACAGACGTTTAAA
                        10        20        30        40        50

pST51r1 AATATATAAAAAAAAATTTTGCACCAGAACAGTTGATAAAAGAAACTTAAAATAACGACTT
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 pST51  AATATATAAAAAAAAATTTTGCACCAGAACAGTTGATAAAAGAAACTTAAAATAACGACTT
              60        70        80        90       100       110
```

Figure 4.20 The 5' Termini of pST51 Related cDNA Clones

A shows *Bam*HI digested pST51 related cDNA clones. The expected products of *Bam*HI digestion are shown schematically. 1kb ladder is a size marker from Gibco-BRL.

B shows aligned sequence from the 5' junction points of pST51 and pST51r1. Numbering is in basepairs and starts at the first base of the pST51 cDNA sequence. Vector derived sequences are underlined.

primers were also used in a PCR reaction to show more clearly the fragment size difference and to test the other non-sequenced cDNAs. Figure 4.21 shows the result of these experiments.

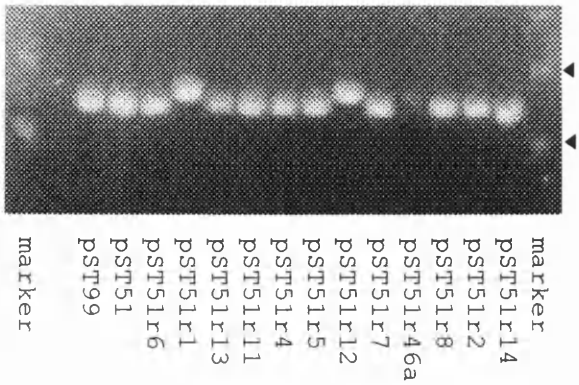
Clearly, the sequence of the cDNAs shows that pST51r1 and pST51r12 contain a 21 bp (in-frame) insertion relative to the coding region sequence of pST51. Interestingly, this insertion truncates one of the Dmbeta consensus glycosylation sites as shown in figure 4.21c. This site is conserved between the shrimp and *Drosophila* sequences (as shown figure 4.12).

Differences in glycosylation patterns between β–subunits from different tissues have been observed in other systems (Sweadner 1989), but this is the first time that any such change has been proposed to be derived from a subtle change in the mRNA amino acid coding region. Of course, this singles out pST51r1/12 as a clone of principle importance.

To look at the internal "*Bam*HI" fragment in more detail an experiment was performed, in which the clones were restricted with *Sau*IIIa and then the resulting DNA smear electrophoretically seperated on an agarose gel, blotted and probed with the ^{32}P-labelled *Bam*HI fragment from pST51. This fragment was isolated from a subcloned fragment to ensure no cross contamination from flanking sequences (data not shown). Since the core sequence of the flanking *Bam*HI sites is a *Sau*IIIA site then the *Bam*HI fragment probe should hybridise to only bands generated from internal, amino acid coding sequence. The first results of this experiment seem to indicate that there is an additional change in the pST51r1 cDNA in the 200bp of sequence surrounding the internal *Pvu*II site (data not shown). The cDNAs which appear to have a pST51-like internal *Bam*HI fragment do not show any distinctions from pST51 using this method.

The distinct pST51r1 and pST51r12 internal *Bam*HI fragment may be derived from a rearranged cDNA clone in the primary library that has subsequently been amplified and isolated on more than one occasion. If pST51r1 was derived from a rare rearrangement then it is unlikely to be represented in other head cDNA libraries. In order to test this a PCR primer was designed, designated pST51ol6, which was specific to the 21bp insert of pST51r1. This oligonucleotide was used in a PCR reaction between pST51ol4 and pST51ol2, figure 4.21. (pST51ol2 was one of the oligonucleotides used to sequence

A



B

```
                                                                pST51r1
CTAAGACTTTGCCAGATCATATGCCAGAGGAACTAAAACAGCACATCAAGGAAAAGCAAAGTCTTAGGCC
::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
CTAAGACTTTGCCAGATCATATGCCAGAGGAACTAAAACAGCACATCAAGGAAAAGCAAAGTCTTAGGCC
                                                                pST51        810

                                                                pST51r1
CAATGAATTATTTAAATACTACTTCCAGACAAATGTAGTTTGGGTCTCGTGCGAGGGAGAAAATCCCGCT
:::::::                      ::::::::::::::::::::::::::::::::::::::::::
CAATGAA                      ACAAATGTAGTTTGGGTCTCGTGCGAGGGAGAAAATCCCGCT
                                                                pST51        880

                                                                pST51r1
GATGTCGAGAACATAAAAGCACGCGACTACTATCCCCGAATGGGATTTCCTCGTTACTATTTTCCGTTTN
::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
GATGTCGAGAACATAAAAGCACGCGACTACTATCCCCGAAtGGGATTTCCTCGTTACTATTTTCCGTTTA
                                                                pST51        950
```

C

```
Dmbeta                                      RPCIFLKLNKIYNWIPEIYND
pST51r1                                     RPCIFLKLNKIYNWIPEIYND

Dmbeta     SKTLPDHMPEELKQHIKEKQSLRPNE-------TNVVWVSCEGENPADVE
pST51r1    SKTLPDHMPEELKQHIKEKQSLRPNELFKYYFQTNVVWVSCEGENPADVE

Dmbeta     NIKARDYYPRMGFPRYYFPFKNIQGYIPPIVAVQFTVET
pST51r1    NIKARDYYPRMGFPRYYFPF?NIQGYIPPIVAVQFTVET
```

Figure 4.21 Multiple Predicted Amino Acid Sequences

A shows PCR products (methods section 2.2.2) produced using 51ol2 and 51ol3 which can be used the generate a PCR product across the region of variation indentified from the restriction map. Two band sizes can be seen, that of pST51r1 and pST51r12 and the size present in pST51 and all other cDNAs. (NB pST51r46a is truncated at the 5' end and does not contain complete copies of both prime sites). Sizes shown are in basepairs and the DNA size ladder used was a 123bp ladder from Gibco-BRL.

B shows the results of a sequencing the variable region (methods section 2.13) using 51ol2 and 51ol3 as primers. The region shown has been sequenced in both strands for both pST51r1 and pST51r12 (not shown). Sequence has been obtained for the region contained within the PCR fragment (not shown) but the only region of variation detected is as shown in **B**. The 21bp insert is underlined.

C shows the predicted protein coding capacity of the sequenced region. A seven AA insert can be seen which disrupts one of the three consensus glycosylation sites. Consensus glycosylation sites are shown in bold

across the 21bp insert). Significantly, a characterisitic sized DNA band was detected in two independent libraries; one derived from *eya* heads and another from *Drosophila* wildtype, heads (data not shown). Thus the difference in this region between pST51r1 and pST51r12 is present in three independent libraries and so is unlikely to be a chance cloning artefact.

The difference between pST51 and pST51r1/12 could arise due to the persistance of a splicing intermediate in the tested libraries. Indeed, splicing intermediates are quite common in *Drosophila* cDNA libraries (Schwarz *et al.* 1990). However, the junction point between common sequence and the 21bp insert does not display consensus splice donor, splice acceptor sites and the insert is too small to be a conventional *Drosophila* intron (Mount *et al.* 1992). Attempts to PCR between pST51ol2 and pST51ol3 or pST51ol4 fail when genomic DNA is used as a template. Control reactions using the primers on diluted plasmid DNA supports the likely positioning of a large intron or several introns between the prime sites in the genome. The balance of the evidence is in favour of the presence of the capacity to encode a second β-subunit isoform within the mRNA pool used to construct the head cDNA library. The differential splicing of small exons to generate molecules proposed to be functionally distinct has been reported from a variety of *Drosophila* molecules including Fasciclin I (McAllister *et al.* 1992) and PS2α integrin (Brown *et al.* 1989).

From the analysis of the cDNAs so far, it appears that the coding capacity of this β–subunit gene includes the ability to produce mRNAs with multiple 5' and probably 3'ends. Unique to this locus out of all β–subunits known, the predicted mRNAs also differ in a way that would produce a second β–subunit isoform that is superficially similar to the first sequenced, but that different in a localised region(s). This change is highly likely to be biologically significant. No other Na+/K+ ATPase system has two identified subunits that vary in such a subtle way. With only one α-subunit primary polypeptide in *Drosophila* will be of great interest to see what subtle affects the different β–subunits identified here will have on the activity of the pump.

4.8 Localisation of pST51 mRNA in Head and Head/Body Sections.

In situ hybridisation to localised message in the head and body using digoxigenin labelled probes was performed as described in the methods section 2.17. Briefly, DIG labelled probes were generated from gel purified whole

insert of pST51 by random priming and hybridised to 10μm cryostat sections that had been fixed in gluteraldehyde. The hybridisation pattern was detected using the NBT/X-Phosphate colour reaction. Positive and negative controls were performed along side the pST51 *in situs* and for this linearised pBluescript (as negative) and pST41 insert (a opsin clone as a positive control) were used. Since a large screen employing the cryostat was in operation during these *in situs* access to the cryostat was limited. Several of the *in situs* shown in the next figure were produced from sections kindly prepared by M. Yang (Glasgow Genetics Department). Figure 4.22ab shows the *in situ* results with pST51 probe on sagital head sections..

The pattern of expression of pST51 RNAs is consistent with it being localised to all the major cortical regions of the fly's brain. Cortical localisation is often seen with *in situ* hybridisations to brain RNA. It may result from the mode of translation of the message; in neurons it is often the protein that is transported along the neuronal processes, translation most commonly takes place in the cell bodies (reviewed Steward and Banker 1992). The expression pattern certainly reminiscent of a gene that is expressed in abundance in neurons and does not show a glial type expression pattern which might be expected for type 2 β–subunits.

The RNA is localised in the eye and those parts of the brain responsible for visual input processing, again in a cortical pattern. Expression can also be observed in the lens region of the eye, although non-specific hybridisation is often observed in this region (but not in the negative control in this experiment). The overall pattern of expression is broadly consistent with that reported by Lebovitz *et al.* 1989 for the α-subunit protein.

Longitudinal cryostat sections were also prepared from whole flies. In these sections a problem was observed in that the gut regions of the adult produced a strong signal in the negative control. Gut phosphatases are not always blocked by lavamisol and it seems that the adult fly contains such a gut localised phosphatase. Several experiments were performed in an attempt to reduce this background by increasing the digestion times with the protease (to destroy the endogenous enzyme) but it seems that the structure of the tissue breaks down before the phosphatases (data not shown). Attempts are being currently made in the laboratory to switch to another detection system. Figure 4.22c therefore only shows the hybridisation pattern to the head and thoracic regions of the fly.

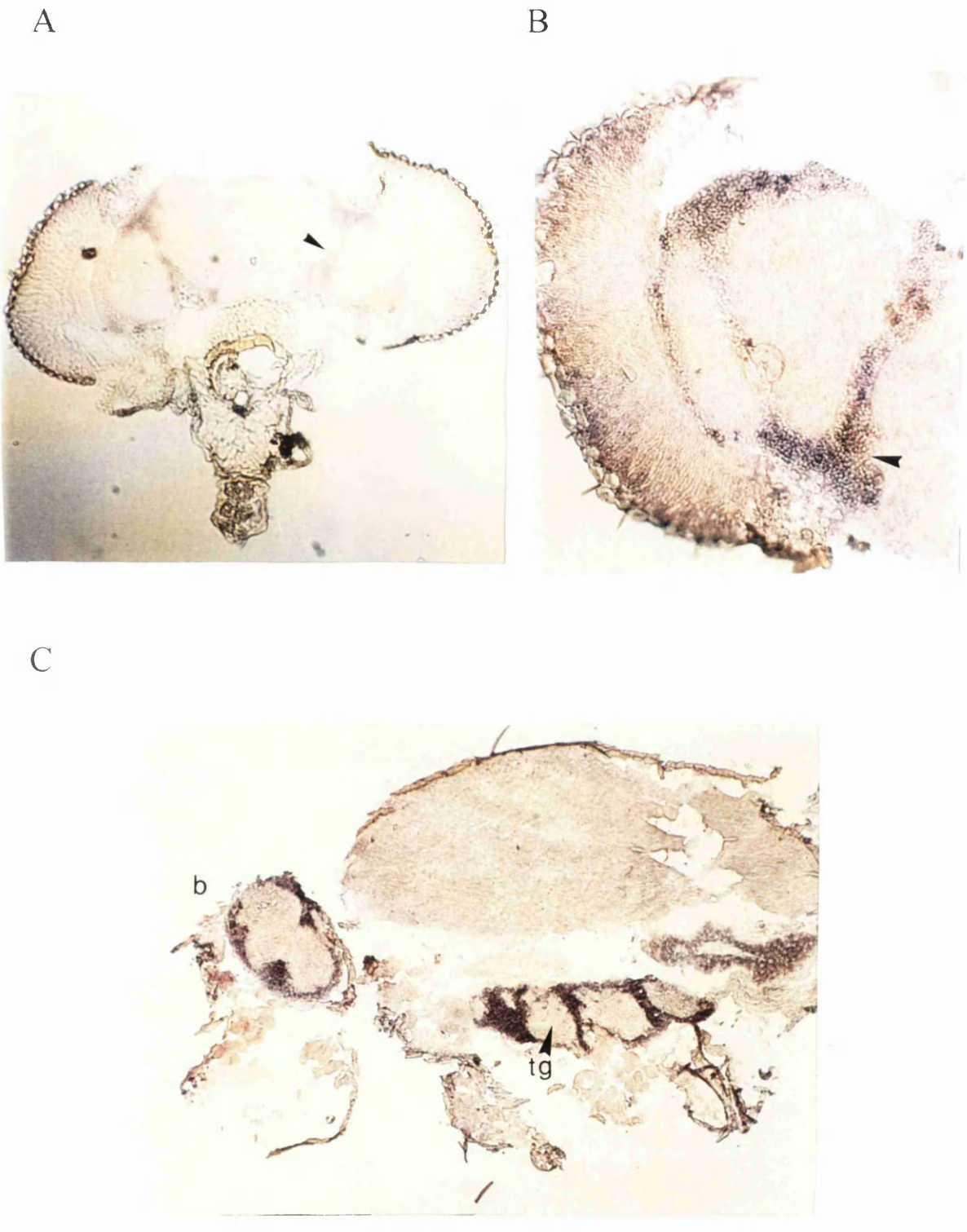


Figure 4.22 Localisation of RNA in Head and Thorax
In situ hybridisations were performed as described in methods section 2.18.
A shows the localisation of pST51 RNA in a saggital section through the *Drosophila* head. Signal to the cortical regions of the brain is arrowed.
B shows a closer view of one half of a sectioned fly head. Staining is arrowed.
C shows a section through the head and thorax of an adult fly. Staining can be seen in the central brain regions (**b**) and the three pairs of thoracic ganglia (**tg**). The gut staining observed (**g**) is also found in the negative control (not shown).

The abdominal pattern of expression cannot be determined using the current system and time was simply not available to develop a new one.

Figure 4.22c again shows strong expression in the head of the fly, but also in the thoracic ganglia. This pattern compares well with the published expression pattern of the α-subunit protein (Lebowitz *et al.* 1989). In the *in situ* the three paired thoracic lobes can clearly be seen. Expression in the leg muscle, reported for the α-subunit, is difficult to see in these sections.

Detection of expression in the thoracic ganglia is consistent with the detection of some expression in body tissue on the northern blot. The head messages are more abundant (about twenty times) than those of the body (section 4.6). Does that relationship hold for the *in situs*? The thoracic ganglia probably occupies about 1/30th to 1/40th of the total mass of the fly body, while the brain and optic lobes occupy about 70% of the available area, excluding cuticle. This gives a "guestimate" of a figure consistent with the northern blot.

The pattern of RNA expression described here is consistent with that of the α-subunit, further adding support to the hypothesis that this subunit is indeed that of the β–subunit of the Na+/K+ ATPase. In addition, the subunit is most abundantly expressed in regions of the nervous system in a pattern not consistent with wholesale glial expression i.e., it does not look like a class two β-subunit. It is not expressed in all cells and so does not appear like a β1 subunit. The expression of the related β–subunit RNA in shrimp can also be found in the nervous system, although it is less tissue specific than in the *Drosophila* case described here (Sun *et al.* 1992). It would be interesting to see whether the β–subunit is also expressed in the Malpigian tubules, since this is a region of localised expression for the *Drosophila* α-subunit (Lebovitz *et al.* 1989). The shrimp β–subunit is also principally involved in salt regulation, although this is perhaps of more physiological significance in a brine shrimp than *Drosophila.*

4.9 Expression in the Embryo

Whole mount embryonic *in situs* were performed as described in methods section 2.17 using electrophoretically purified whole insert of pST51. Embryos were approximately staged by collecting three eight hour collections from cages at 25°C and ageing some of the plates to give embryos 0-8, 8-16 and 16-24. In the first experiments these were then pooled and probed

together. Morphological features were then used to give the more accurate stages shown in figure 4.23 as described by Weischaus and Nusslein-Volhard 1986.

The central nervous system is formed from a specialised areas of the ectoderm called the neurogenic regions. From these regions, neuroblasts (neuronal precursor cells) first differentiate at stage 9. However, the progeny of these neuroblasts does not begin neuronal differentiation until stage 13. From figure 4.23 it can be seem that the first localised expression of pST51 RNA can be seen segmentally arranged in the developing nervous system at stage 13. Segmental arrangement of this region can be mophologically detected at stage 11 but it is only after the shortening of the germ band at stage 12 that expression of pST51 RNA can be detected. Thus the RNA is expressed in the developing nervous system just as the neuronal cells beginning to take on mature, differentiated characteristics. Expression of pST51 RNA continues to be detected in the central nervous system throughout later stages. It can also be seen in the central nervous system of first instar larvae (data not shown).

Expression in later stage embryos can also be detected in components of the peripheral nervous system. Exact identification of specific cell types requires the use cell specific markers which have not been used in this study. However, staining in the antenomaxillary complex and the larval photoreceptor (Bolwigs organ) can clearly be seen. Staining of segmentally arranged sensilli can be clearly seen in later (stage 16) stage embryos. The peripheral staining is also shown in figure 4.23d.

Staining is not observed in non-neuronal structures of the embryo. Specifically there is no staining that can be detected in the gut and none detected in the musculature. Faint expression is seen in the dorsal longitudinal tracheal trunks, this staining is only seen for some preparations and is not specific to a particular probe. It is a commonly observed background hybridisation caused by the trapping of the probe in the convoluted tracheal structure (E. Zador, D Cromptom pers. comms. Glasgow Genetics Department). Thus the pST51 probe hybridises predominantly to nervous tissue.

The staining of the nervous system is consistant with pST51 homologous RNA's proposed role as a Na+/K+ ATPase β-subunit. However, since it is reasonable to suppose that the Na+/K+ ATPase is required in other tissues why

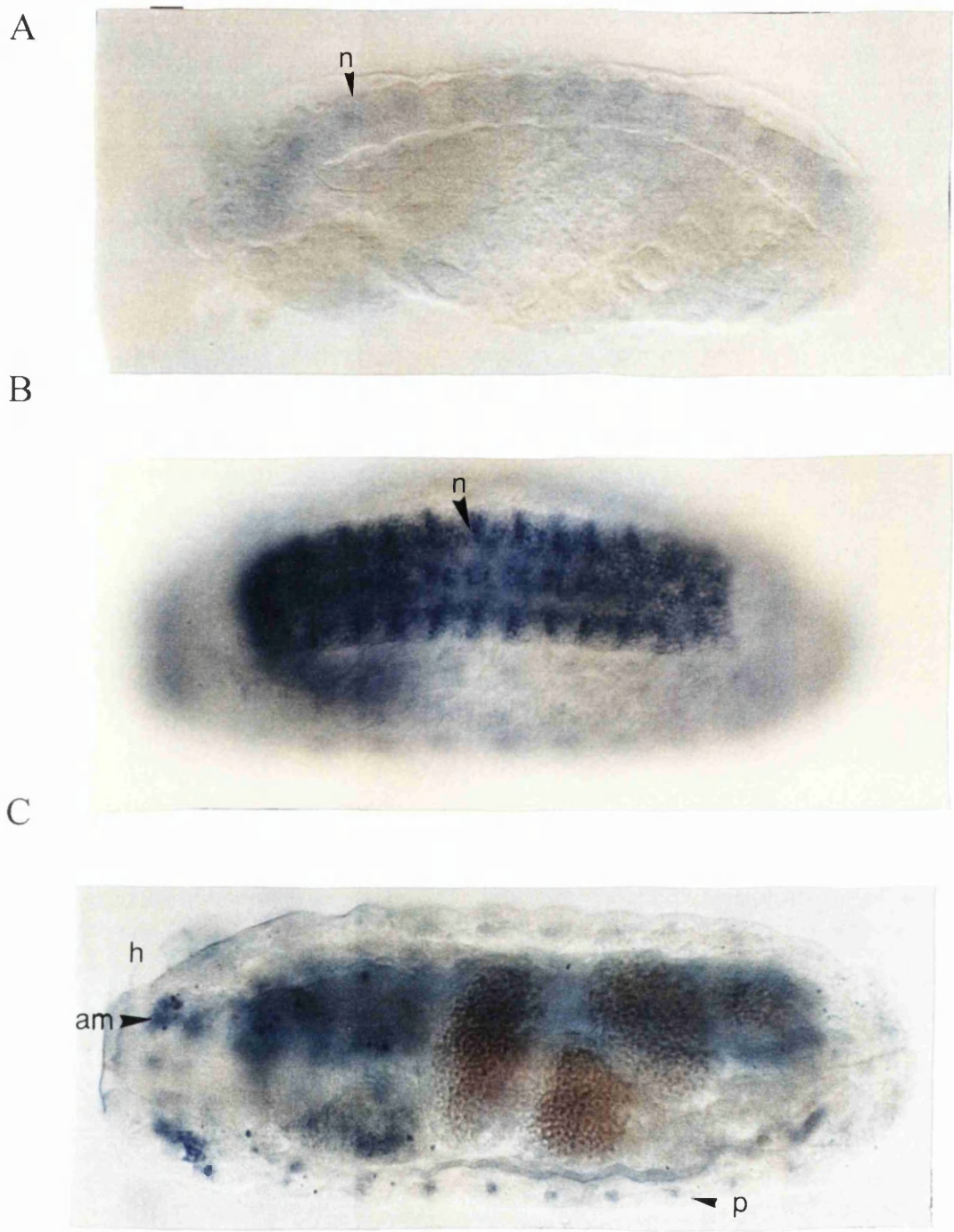


Figure 4.23 Expression in the Embryo

Whole mount *in situ* hybridisation to embryos was performed as described in methods section 2.17. Shown are examples of embryos derived from two independent experiments.

A shows a stage 13 embryo showing localised expression of the mRNA(s) in the developing nervous system (**n**). This is the earliest embryonic stage in which localised expression was detected.

B, and **C** show late stage embryos (stage 16) and demonstrate the different features of the late stage expression patterns. **B** shows the mackerel patterned staining of the central nervous system (**n**). **C** shows the staining thought in the ventral nerve cord and the developing head region (**h**) of the embryo. Staining can also be detected in the peripheral nervous system including the antennomaxillary (**am**) complex and seqmentally arranged components of the PNS (**p**).

don't those tissues also stain? One possibility is that other staining is below the level of detection and that the majority of tissues express the Na+/K+ ATPase at some low level. Although the detection technique can detect sensilli which usually comprise of one or two neurons and associated cells it is possible that cells or tissues that weakly express the subunit will go undetected. If the staining reaction is performed for a longer time period, weak, non-localised staining is observed. Non-localised staining occurs at all stages and in areas of the embryo that lack cellular structures and so is very unlikely to be cell specific. It also varies from experiment to experiment in its intensity. However, this effect would act to blot out lower level of expression. Further supporting evidence is required before weak expression in non-nervous tissues can be ruled out. For example, PCR or northern blotting might be used to search for transcripts in embryonic material harvested prior to the differentiation of the nervous system.

The *in situ* experiments presented here is consistent with message represented by pST51 having a specific role in the functioning of the nervous system. The experiments further support the hypothesis that product acts as a Na+/K+ ATPase β-subunit. The absence of detectable expression in non-nervous tissues suggests that pST51 could be a cloned neuron specific mRNA. This raises the possiblity of other functionally and structurally homologous subunits that perform a similar role in non-nervous tissues.

4.10 Genomic Southern and *In Situ* Localisation to Polytene Chromosomes

Genomic Southerns were performed as described in methods section 2.6 using the whole insert of pST51 as probe. The resulting blot is shown in figure 4.24. The transcribed region spans about 18kb. This estimate of genomic size may be an underestimate as several of the transcripts isolated in section 4.7 have flanking 5' and 3' regions not represented by the pST51 probe.

Polytene *in situs* were performed as described in methods section 2.19 using the whole insert of pST51 as probe. From the *in situs* shown in figure 4.25 a location was produced on chromosome two in band 41e or 41f. This region is close to the centromere on the right arm of chromosome two.

The region has not been studied in great detail (by *Drosophila* standards), although several deficiencies are located in the region of interest (Hilliker and Holm 1975, Hilliker 1976). There are also a few mapped P-element insertions

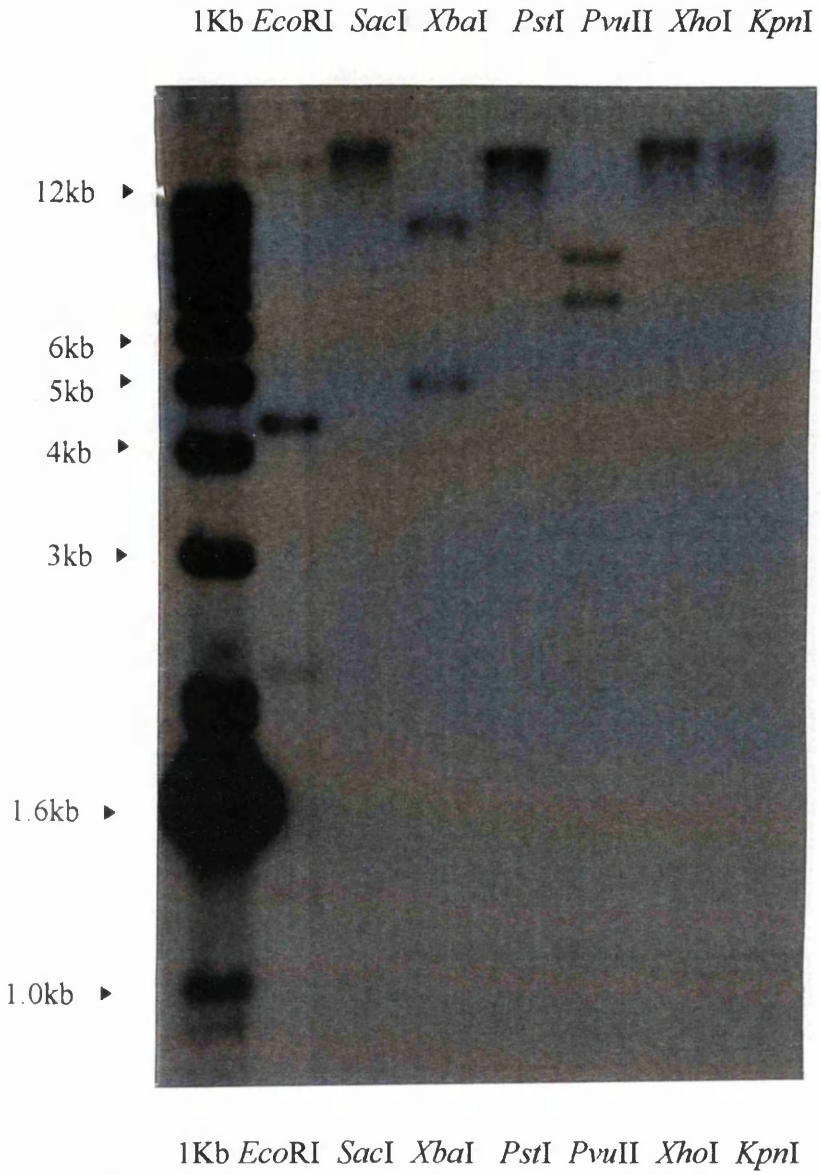


Figure 4.24 A Genomic Southern Using a pST51 Derived Probe
The genomic Southern was performed as described in methods section 2.6. The probe used was generated from the gel purified whole cDNA insert of pST51. DNA was restricted with *Eco*RI, *Sac*I, *Xba*I, *Pst*I, *Pvu*II, *Xho*I and *Kpn*I. The size marker used was 1kb ladder (Gibco-BRL) end-labelled as described in methods section 2.10.2.
Digestion with each of the restriction enzymes selected either yields a number of hybridising fragments which are larger than 15kb, and so the fragments are too large to resolve closely migrating bands. However, since digestion of both the cDNA and genomic DNA with *Pvu*II gives two bands it is tempting to speculate that the gene is single copy, or more precisely that it is not encoded by two or more genes dispersed throughout the genome. Tandem gene clusters cannot be distinguished from single genes by this type of analysis.

A

B

Figure 4.25 Polytene Chromosomal *In Situs* Using a pST51 Derived Probe
Polytene *in situs* were performed as described in methods section 2.19. Arrows indicate the position of the hybridisation signal.
A shows a set of polytene chromosomes displaying the chromosomal arms radiating from the chromocentre (**ch**).
B shows a different set of chromosomes displaying the hybridisation signal which is close to the chromocentre.
From these chromosome squashes and others (not shown) a putative location of the gene can be determined as 41e-f.

in this region. There are several useful insertions in the region that could be of great use in order to study the genetics of the pST51 defined locus. A P-element insertion, has been studied in the laboratory of D. Fambrough that is thought to disrupt the Na+/K+ ATPase α-subunit and produce a viable, wings down phenotype (Schubiger *et al.* 1993). The β–subunit might be expected to produce a similar phenotype in some of it's alleles, but no such phenotype is found in this region. Of course, although in the simplest case the β–subunits phenotype might mimic that of the α-subunit, a variety of possibilities exist where it will not, for example if multiple β–subunits occur specified by additional loci or if the gene disruption proved lethal.

Early physiological studies suggest mutants in the Na+/K+ ATPase may produce a bang-sensitive phenotype (Jan and Jan 1978, Ganetsky and Wu 1982). Schubiger *et al.* 1993 have performed a P-element screen to isolate further bang-sensitive mutants. One of these mutants mapped to the region of the polytene chromosomes known to encode the α-subunit. This mutant shows a bang-sensitive phenotype and an increased sensitivity to ouabain. Furthermore, the transcription pattern of the α-subunit gene is altered by this P-element insert. The insert has been mapped to the first intron of the α-subunit. This set of experiments clearly demonstrates the predicted link between the behavioural phenotype and the Na+/K+ ATPase. Furthermore, it demonstrates a ready means to isolate phenotypic inserts in the genes of the Na+/K+ ATPase.

4.11 Other Related Genes

Coincidentally, a screen had been previously performed by J.Davis (J.Davis Glasgow Genetics Department, pers. comm.) to clone Na+/K+ ATPase β–subunit genes. They were considered of special interest because of reports that they may specify an X-linked member of the bang-sensitive group of behavioural mutants. The cDNAs were isolated using a bovine, type 1 β–subunit in a low stringency hybridisation. Five non-cross hybridising genomic clones were isolated and these were then mapped to the polytene chromosomes. None of these loci mapped to the X-chromosome and so they were not studied further. J. Davis kindly provided 'lambda genomic clones from these loci for testing and they were cut with suitable restriction enzymes, blotted and probed with pST51 at both high and low stringency (methods section 2.11). The protocol used allowed for washing a single blot at a variety of stringencies and these blots are all shown in figure 4.26.

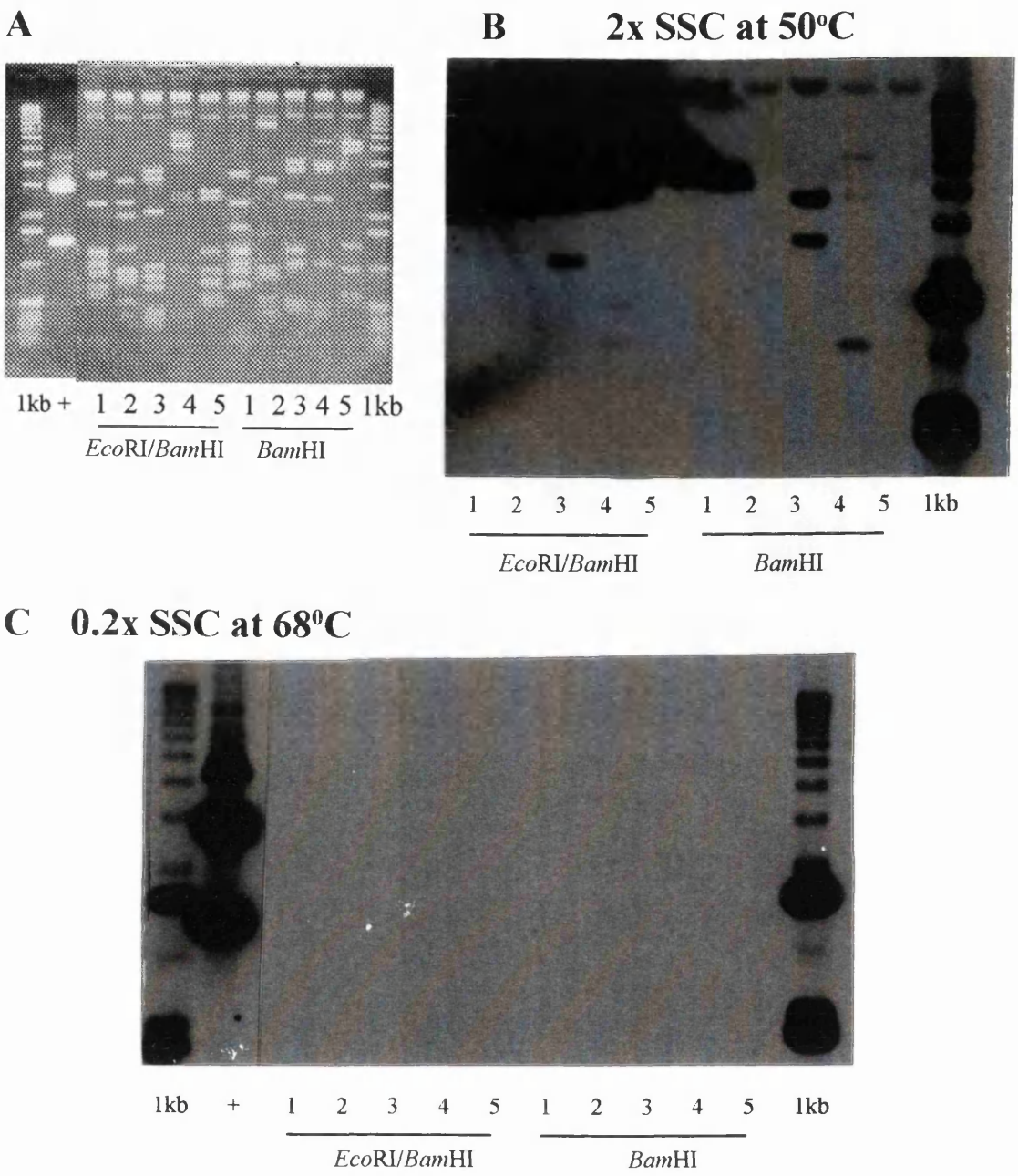


Figure 4.26 Relationship of the pST51 cDNA Insert to Lambda Clones Isolated by J.Davis

Genomic Clones were restricted with *BamHI*, and double digested with *Eco*RI/ *Bam*HI. The five independent genomic clones are numbered 1 to 5.

A shows an ethidium bromide stained gel prior to blotting. Two duplicate gels and blots were produced (methods section 2.6), only one of the stained gels is shown. All blots were hybridised with an identical probe generated by random priming a gel purified whole fragment (methods section 2.10.1) of pST51. **B** shows an autoradiograph of a blot hybridised at low stringency (methods section 2.11.2) and washed in 2x SSC at 50°C. This signal can be washed off the blot at high stringencies. **C** shows the blot hybridised at high stringency in 6xSSC (methods section 2.11.1) and washed in 0.2x SSC at 68°C.

Each filter also contained a positive control, *Xho*I/*Xba*I digested pST51 (+), which gave an intense signal at all stringencies (not shown for **B**). Radiolabelled 1kb ladder (Gibco-BRL) was prepared as section 2.10.2.

From the results it appears that hybridisation can be detected to specific bands at low stringency, but these wash off at higher stringency until they are no longer detected. The high stringency blot did not show any specific hybridisation other than to the positive control. Thus it appears that related loci do exist, but none of the genes isolated in the Davis screen are the one described in this chapter. The presence of related cDNAs is also supported by a low stringency re-probe of the high stringency genomic blot previously presented (data not shown).

The failure of the Davis screen to isolate a fragment the pST51 locus is not surprising if one compares the bovine β1 sequence to the pST51 cDNA sequence, no significant regions of homology can be found. Another screen for a *Drosophila* β–subunit by D.Fambrough's group failed to find any related loci at all (using a PCR based method; D Fambrough, John Hopkins University, pers. comm.) and so it will be very interesting to see what these positive clones prove to represent.

4.12 Chapter 4 Conclusions

4.12.1 Technical Considerations

The principle purpose of the analysis in this chapter was to confirm that the cloned gene in question had the properties that were expected of it from the screen. I believe that this chapter has satisfied that requirement; the gene is transcribed in the selected pattern as assayed by northern blotting, and shown to be nervous system localised by *in situ*- hybridisation to the head, body and embryo. The clone represents a molecule that is expected to be abundant in the nervous system. Finally, the clone has not been previously identified and the level of DNA homology suggests a reason why such screens would fail, and why, even in 1994 there is a role for differential screening for the finding of such molecules.

4.12.2 The Relationship to the Brine Shrimp Sequence

As described in several sections in this chapter, the preliminary analysis of this cDNA from the screen indicates that it is likely to encode a Na+/K+ ATPase β–subunit, most closely related to that of the shrimp. If one refers back to the phylogeny presented in figure 14b it is clear that this *Drosophila* subunit is still considerably distanced from the shrimp sequence. Indeed, it is further

distanced than the relationship between the mouse β2 and *Xenopus* β3 sequences. Because of the experiments performed using interspecific and interpump mixes of subunits, and my own preliminary analysis of tissue specific transcription pattern of this *Drosophila* β-subunit I have little doubt that this subunit is likely to form an active pump with the *Drosophila* α-subunit. It will also be interesting to see if it can do so with the shrimp β–subunit.

It is tempting to define this new arthropod class as β4, although the group would show a high level of evolutionary divergence. Because many of the subunits from other systems were isolated by DNA sequence homology to others it may be possible that a mammalian homologue of "β4" may yet be found. Even between shrimp and *Drosophila* β–subunits there are considerable differences in likely biological functioning of the Na+/K+ ATPase. The shrimp gene is very abundantly expressed in tissues involved in salt regulation as well as the excitatory tissues (Sun *et al.* 1992). This may reflect the strong physiological need for salt regulation in the brine shrimp, a requirement greatly diminished in terrestrial insects. Interestingly, it has been reported that the shrimp produces two β–subunit proteins that are subtly different in size (Baxter-Lowe *et al.* 1988). But also, in the brine shrimp there are also two α-subunit proteins (Peterson *et al.* 1982). Therefore, the similarity between the two arthropod β–subunit protein sequences is not completely extended to the specific biology of the subunits in the two systems.

4.12.3 Future Work

In this chapter I have presented the molecular properties of the cDNA clone. These include:-

- The full-length DNA sequence of one of the cDNAs
- The inferred amino acid sequence and its principle features
- The evolutionary relationship to other loci
- The presence of multiple cDNAs that suggests that the gene has the capacity to encode different isoforms of the protein.
- The tissue localisation of the RNA in embryos, and adult head and body.
- The copy number of the gene and its chromosomal localisation
- The presence of related loci

These experiments naturally suggest other experiments that should be performed. Principle among these are that I would like to see pST51r1/12 fully sequenced and transcript specific probes generated and used to study the

relative distributions of the two mRNA types. It would be productive to map the 5' ends using primer extension and to map the transcripts onto the genomic organisation of the gene. Genomic 'lambda clones could be isolated from the *Drosophila* genomic library described in section 3.2.1. It is also essential to show that the β–subunit protein can produce an active pump *in vitro* in combination with the *Drosophila* α-subunit. It would be very interesting to take an *in vitro* produced protein and raise antibodies which could be used in a whole range of studies, not least for *in situ* localisation of the protein(s). It is also possible to use isoform specific antibodies to be used to distinguish between isoforms.

4.12.14 Mutagenesis Screens

One of the purposes of this work was to define molecules that could be mutated by site-selected mutagenesis. In fact, even with the level of characterisation of this gene, at least four simple mutagenesis strategies present themselves. Firstly, one could use the deficiencies that map to the region and cover the locus with overlapping deletions and study the phenotype directly from this, or more subtly, use the deficiencies to target P-elements or any other mutagenic agents to the region.

Secondly, one could use site-selected mutagenesis. I have already designed primers for this purpose and used them to screen DNA stored from previous mutagenesis experiments (methods section 2.20). Although no P-element inserts have yet been found (data not shown), the number of flies screened was only about 5,000.

Thirdly, one could use the α-subunit insert in a supressor/enhancer screen. Some of these modifiers might reasonably be expected to map to the β–subunit (D. Fambrough, John Hopkins University pers. comm.).

Finally, one could perform further sets of behavioural screens to isolate mutants with a bang sensitive phenotype that co-map with the β-subunit locus. This is the approach that has recently proved successful for the α-subunit (Schubiger *et al.* 1993). Of course these strategies are not mutually exclusive, one might use the deletions to moved marked P-elements to the locality and then use local jumping and site-selected mutagenesis to move the P-element to the precise region required. I expect that in the next few years many of these experimental approaches will be pursued.

Chapter 5 A Novel Member of the Synaptobrevin Gene Family

5.1 Introduction

Small synaptic vesicles uptake, store and release a number of neurotransmitters at nerve termini. This process is central to the transmission of the nerve impulse across the chemical synapse and hence to the functioning of the vast network of neurons that make up the nervous system of any higher animal.

When the action potential reaches the synapse, the presynaptic membrane depolarises and this event causes, via an influx of calcium (Ca2+) ions, the release of the neurotransmitter, from the small synaptic vesicles that are docked close to the synaptic membrane. This fusion event is reversed by the reclaiming of the fused vesicle membranes by endocytosis. This process is shown in figure 5.1.

The mechanism of exocytotic vesicle release utilises a system independent of Golgi or trans-Golgi processing network and is capable of releasing neurotransmitter about 50 times per second. The quantal release of neurotransmitter is thought to be tightly regulated, both in terms of the cycling of the transmission system and also the regulation of the magnitude of response to a given signal (Zucker *et al.* 1989). The regulation of transmitter release is thought to be the target for a number of interesting phenomena such as long term potentiation; a property of neurons that is thought to underlie the process of memory formation (Malinow and Tsien 1990, Bekkers and Steven 1990). Transmission of a signal across the synapse is a central and essential process that must be understood at the molecular level in order to understand even the basic functioning of the nervous system.

The isolation of proteins found on the surface of small synaptic vesicles has been greatly facilitated because they are both abundant (they hold 6% of cellular protein by mass; Sudhoff and Jahn 1991) and can be isolated by sucrose density centrifugation in an essentially pure form. The individual proteins can be purified allowing a reverse genetic approach to be adopted to clone the individual genes using standard techniques. Using this strategy a

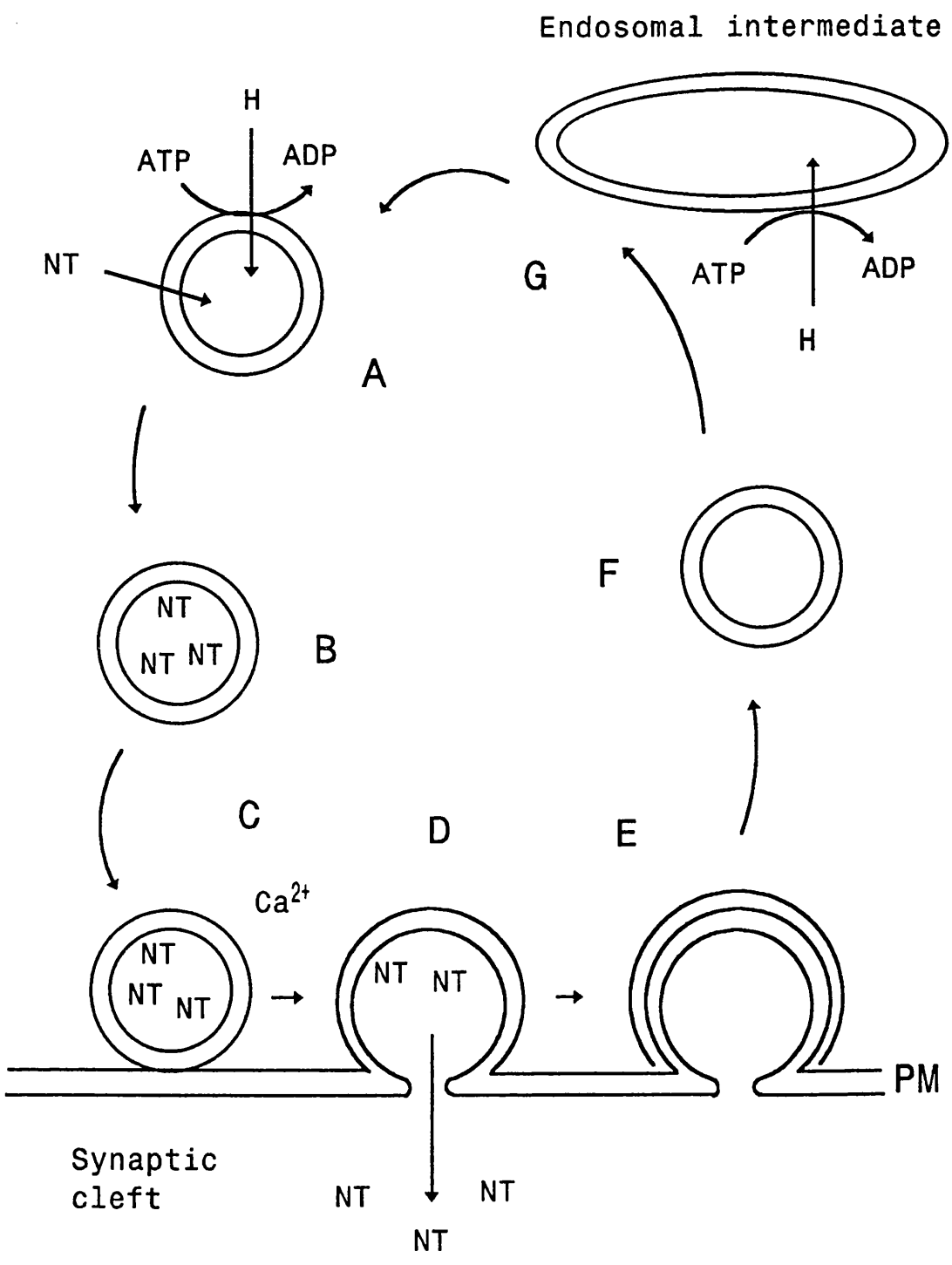


Figure 5.1 Synaptic Vesicle Movement at the Nerve Terminal

The synaptic vesicles are filled with neurotransmitter by active transport (**A**), are moved and docked (**B** and **C**) at the active zone (non-docked vesicles are thought to be held in a waiting state), calcium influx results in release of neurotransmitter (**D**) and then the synaptic vesicle membranes and marker proteins are endocytosed via clathrin-coated pits (**E**)and recycled (**F** and **G**). The E and G stages are not definatively proven. **PM** is the plasma membrane, **NT** is neurotransmitter. Re-drawn and paraphrased from Sudhoff and Jahn 1991.

number of proteins have been identified that fall into two basic groups; the molecules responsible for charging the vesicle with neurotransmitter and those thought to be involved in the cycling of the vesicle itself.

In mammalian systems, several of the "charging" molecules have been cloned. Principle among these are most of the 8-9 subunits of the vacuolar ATPase and a number of neurotransmitter transporters including those of acetylcoline and glutamate. The vacuolar ATPase acts as a proton pump to establish an electrochemical gradient across the vesicle membrane. It is to this gradient that the transporter molecules energetically couple the transport of the neurotransmitters. The vacuolar ATPase is not homologous the Na+/K+ ATPase discussed in the previous chapter. The vacuolar ATPase is thought to be almost solely confined to the vacuolar membranes (Sudhoff and Jahn 1991).

The second class of molecules are involved in "targeted movement and fusion" (Sudhoff and Jahn 1991). A number of such molecules have been studied and include synapsins synaptophysins, synaptotagmins, rab3As and synaptobrevins (also called VAMP). The basic properties of these proteins is described below and summarised in figure 5.2.

Synapsins- These were originally discovered as being the major phosphorylated protein in nerve termini (Johnston *et al.* 1972). They have been shown to bind components of the cytoskeleton *in vitro* (Bains and Bennett 1985, 1986, Sudhoff *et al.* 1989b). Four homologous proteins have so far been identified-synapsins Ia, Ib, IIa and IIb, encoded by two genes in mammals. The four proteins are differentially expressed in various regions of the adult brain (Sudhoff *et al.* 1989b). Synapsin Ia/Ib has 4 distinct protein kinase sites that are substrates for 4 different kinases. Synapsin IIa/IIb has a single site known to be regulated by two kinases. (recently reviewed Greengard *et al.* 1993). Phosphorylation of Ia/Ib induces a conformational change (Benfenati *et al.* 1990). Phosphorylation of Ia/Ib with Ca2+/calmodulin dependant protein kinase 2 modulates the binding to synaptic vesicles and to actin (Petrucci and Morrow 1987, Bahler and Greengard *et al.* 1987). The known biochemical properties of the synapsins suggest that they mediate the interaction of the cytoskeleton and synaptic vesicles. In doing so synapsins are likely to exert a key regulatory role in

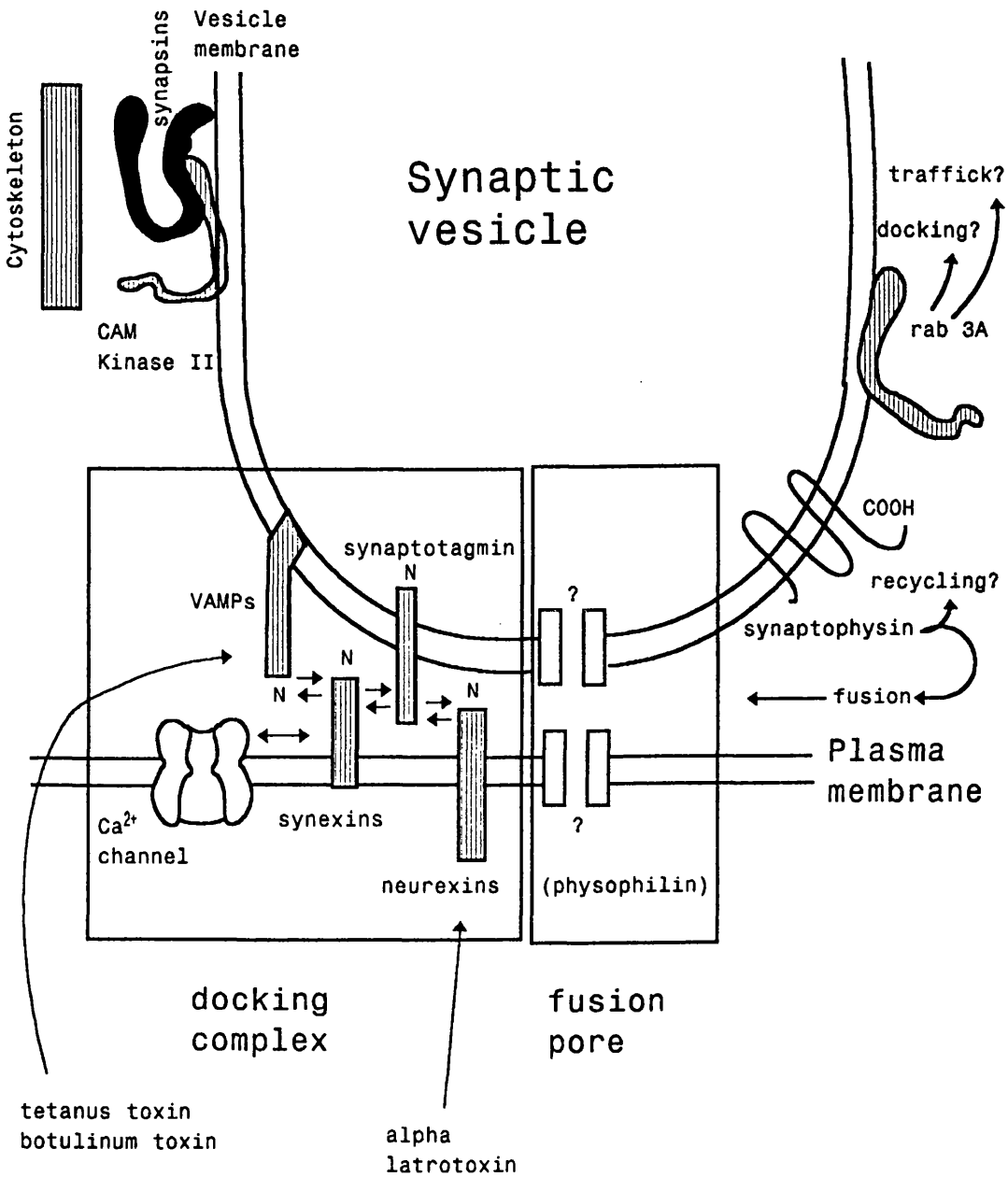


Figure 5.2 Synaptic Proteins and their Potential Functions
The figure depicts the most characterised synaptic vesicle proteins and their proposed roles in the functioning of the exocytosis system. The docking complex and fusion pore are shown as being distinct; this is for schematic purposes and following one of the current views on the actual structure of the system. Some proposed plasma membrane receptors are also shown in the diagram, as are the toxins known to act upon this system. Please refer to main text for further details. This diagram is redrawn and slightly modified from Jessell and Kandel 1993.

neurotransmitter release by controlling the availability of vesicles for exocytosis.

rab3A is a small GTP binding protein. It is thought to mediate the vesicle docking/fusion stage of the excretory cycle (Fischer V Mollard *et al.* 1990). This is by analogy to a class of yeast mutants deficient in the docking and fusion stages of secretion (*sec4*) and encode a small GTP binding protein. rab3A is specific to small synaptic vesicles.

Synaptotagmin (p65) may play a role in the formation of the docking complex formed at the vesicle/plasma membrane junction. The complexes release of neurotransmitter is directly dependent upon cellular Ca2+ levels and so a Ca2+ binding protein is expected to play a key role in the complex. Synaptotagmin possesses an evolutionary conserved protein kinase C2 domain, implicated in Ca2+ binding (Nishizuka 1989). It is possible that this domain is also involved in lipid binding (Perin 1990). The synaptotagmin protein is found as an aggregated complex (Perin 1991). Synaptotagmin may interact with components of the plasma membrane such as synexins (Bennett *et al.* 1992) and neurexins (Bennett *et al.* 1991).

Synaptophysin (p38) exists *in vivo* as a three or four subunit complex in association with another unidentified subunit (Buckley and Kelly 1987, Thomas *et al.* 1988). It is capable of forming homo-oligomeric channels in black lipid membranes (Thomas *et al.* 1988), although the physiological significance of this has been questioned (Sudhoff and Jahn 1991). It is proposed to form a "fusion pore" at the site of vesicle fusion, perhaps in association with plasma membrane components such as physophilin (Thomas and Betz 1990).

The molecules that have so far been discussed have putative roles in the excretory process. These proposals are largely based upon analogy or homology to other, more clearly understood systems. The studies have largely concentrated on the the most abundant molecular components of the synaptic vesicle membrane. Together they may account for 30% of the vesicular surface protein, although there are perhaps 50 proteins that can be identified from synaptic vesicles (Sudhoff and Jahn 1991). Another abundant component of the vesicle membrane that is thought to act alongside these

molecules is **synaptobrevin** (also called VAMP). Unlike the molecules so far discussed synaptobrevin does not possess sequence homologies that immediately suggest a role for the molecule in the exocytotic process. Considerable progress has recently been made in understanding the role this class of molecule plays in the exocytotic process. What has so far been learned of the role of synaptobrevins is discussed in detail below.

Synaptobrevins were originally discovered as major components of cholinergic synaptic vesicles from marine rays and independently as a rat brain synaptic vesicle protein (Trimble *et al.* 1988, Baument *et al.* 1989). The small 120AA, 18kDa protein has been remarkably well conserved in evolution and so has been isolated from a variety of organisms. Comparison of the conserved regions highlights the four domains structure of the molecule. Principle among these structures is the transmembrane domain that has been shown to span the membrane once in *in vitro* systems (Trimble *et al.* 1988, Sudhoff *et al.* 1989). The protein is orientated such that the small C-terminal domain, just two AA in bovine and *torpedo*, projects intracellularly into the vesicles interior. The most N-terminal cytoplasmic domain is rich in proline (26% in bovine) or asparagine, and is of variable length. The second cytoplasmic domain is very highly conserved with 75% invariant AA residues between synaptobrevins isolated from organisms as diverse as cattle, pacific ray and *Drosophila* (Sudhoff *et al.* 1989).

In mammals, synaptobrevins are found as an exclusive component of the small synaptic vesicles. They have not been identified as having homologues in other cellular excretory systems, even in the large dense particles that release neuropeptides. Hence, in mammals, synaptobrevins, along with several other identified components might act as specific trafficking markers for the vesicle. [The one mammalian exception to this has been reported by Cain *et al.* 1992 who has identified a form of synaptobrevin that plays a role in glucose transportation in adipocytes. This is very much "the exception that proves the rule" as the release of the GluT4 glucose transporter from these tissues occurs via small synaptic vesicles]. Other, non-specific components are likely to become directed by synaptobrevins. Because of the structure of synaptobrevins it has been suggested that they interact with a receptor on the plasma membrane. Attempts to identify synaptobrevin receptor proteins have so far failed (as described Sudhoff and Jahn 1991).

A characteristic feature of the docking of synaptic vesicles on the plasma membrane surface are the large multiprotein aggregates that are observed under the EM (Bennett *et al.* 1992). Interestingly, under certain conditions of purification, reversible aggregates of a number of synaptic vesicle proteins can be formed. Synaptobrevin is one of these proteins that can interact hydrophobically with both synaptophysin and synaptotagmin under different conditions of purification. This is shown in figure 5.3. Bennett *et al.* 1992 speculate that these are indeed the complexes observed under the EM.

Synaptobrevin homologues with the characteristic conserved four domain structure have recently been identified in yeast (Dascher *et al.* 1991). Two homologues were identified as plasmid clones able to suppress the loss of YPT1, a locus that encodes a small GTP binding protein. This protein is a member of the RAS superfamily and is thought to function in the protein transport machinery in the same way as rab3A. YPT1 is known to act in the transport process from the ER to the Golgi. Homologous RAS family members such as Sec4p function in the Golgi to plasma membrane transport system. Complementation by over-expression is thought to occur between the mutant locus and gene products in the same pathway to which it directly interacts (Dascher *et al.* 1991). Thus this work is directly suggestive of a direct interaction between a member of the RAS superfamily and a synaptobrevin family member.

A great increase in our understanding of the role of synaptobrevin can be expected to follow the discovery that tetanus and botulinum-B toxin specifically acts by cleaving synaptobrevin. In man, rat and yeast two synaptobrevin proteins have been identified, transcribed from two independent loci. Interestingly, tetanus and botulinum toxin cleave synaptobrevin-2 and not synaptobrevin-1 in rat. This is thought to be due to a single variant amino acid between the two potential cleavage sites (Schiavo *et al.* 1992).

Two distinct isoforms of synaptobrevin, encoded by two separate loci have now been reported in human (Archer *et al.* 1990), rat (Elferink *et al.* 1989) and yeast (Dascher *et al.* 1991). Comparative expression studies between synaptobrevin-1 and synaptobrevin-2 have been performed. In rat, for example, Trimble *et al.* 1990 have shown a clear difference in expression between the two transcripts. Synaptobrevin-1 (which they call VAMP-1) is expressed in brain nuclei and spinal cord lamellae known to be involved in

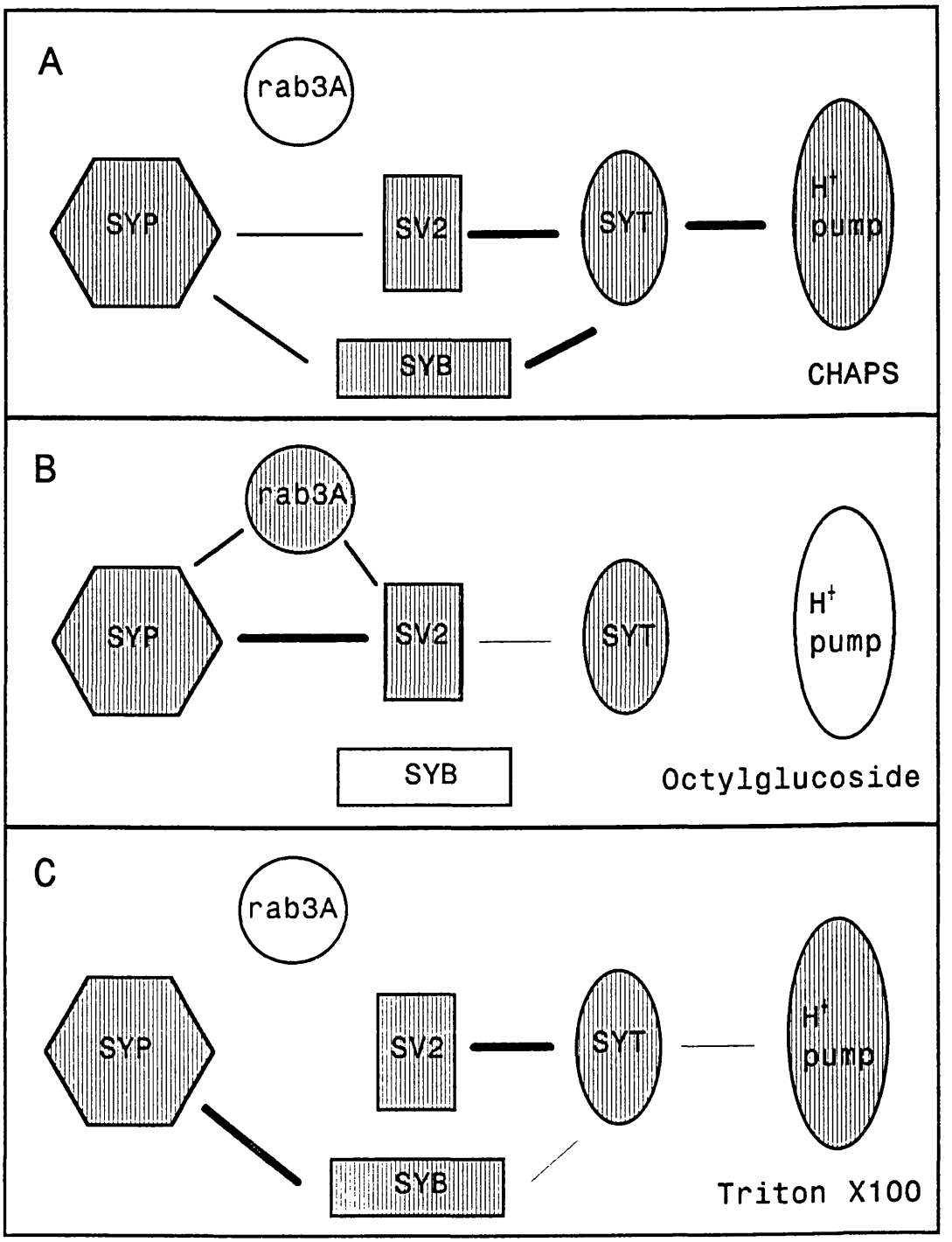


Figure 5.3 The Interactions Between Synaptic Vesicle Proteins after Detergent Solubilisation.

Shaded objects indicate interacting proteins, non-shaded proteins are independent of other markers. Thickness of line indicates the relative strength of interaction. Redrawn from Bennett *et al.* 1992.

somatonuclear function. Synaptobrevin-2 is more widely expressed in the central nervous system, but predominantly in nuclei responsible for autonomic motor, sensory and integrative functions.

Several synaptic vesicle proteins have identified homologues in *Drosophila*, including synapsin (Heimbeck *et al.* 1991), synaptotagmin (Perin *et al.* 1991) and synaptobrevin (Sudhoff *et al.* 1989). To date no mutants have been identified from *Drosophila* that co-map with these loci. However, one mutant is known from *Drosophila*, called *shibire*, that causes the accumulation of small synaptic vesicles near the pre-synaptic membrane surface. Not suprisingly, several lethal alleles have been characterised, although the most commonly studied allele is temperature sensitive. At non-permissive temperatures this defect causes paralysis that can be reversed by restoration to permissive conditions. There is considerable interest in this mutant and it has recently been cloned by several groups (Chen *et al.* 1991, Van der Bliek and Meyerowitz 1991). It show homology to dynamin, a GTPase (Shepetner and Vallee 1992) thought to be involved in microtubule binding (Chen *et al.* 1991, Van der Bliek and Meyerowitz 1991).

The synaptobrevin homologue was cloned in *Drosophila* by homology to the bovine sequence (Sudhoff *et al.* 1989). The inferred amino acid coding region again shows the characteristic four domain structure. The protein was shown to be highly abundant in head tissue. This is encoded in head tissue by two messages of 0.85 and 0.9kb. The synaptobrevin homologue was found to be produced from a single copy gene located at 47E-47F on the *Drosophila* polytene map. The locus was named *syb*. Unfortunately, no known mutations map to this region.

Recently additional sequence has been released for the genomic region of the *syb* gene. The authors, Chin *et al.* 1993, describe the locus as producing two distinct transcripts that produce different polypeptides. One of these corresponds to the Sudhoff's cDNA (with extra untranslated sequence) although there are several single base pair changes. One of these changes alters the tetanus toxin cleavage site. These changes are not discussed in the header comments to the sequence, although of critical importance (see later).

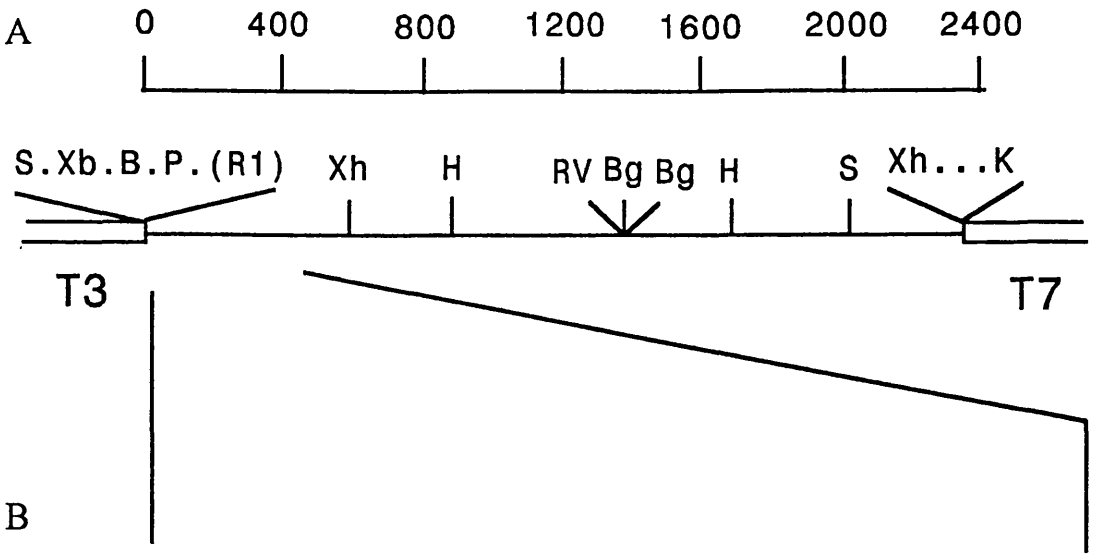
In *Drosophila*, only the one synaptobrevin locus has so far been published. However, the study of Sudhoff *et al.* 1989 on the *syb* locus reports the presence of other bands on a low stringency northern to head tissue that is strongly suggestive of the presence of a second locus. In this chapter I will discuss my own evidence that the *Drosophila* genome does indeed encode a second form of synaptobrevin, and that pST123 is a cloned fragment of one of the messages encoding this novel form.

5.2 Introductory Analysis of pST123

As described in chapter 3, pST123 was isolated as a cDNA for which parent transcripts were found abundantly in the *Drosophila* head but only very weakly in the *Drosophila* body. Subsequently, a sequencing run from the 5' end of the cDNA molecule was used to identify it as encoding a member of the synaptobrevin gene family. This cDNA was the first of the head elevated clones to be identified upon the basis of 5' and 3' sequence analysis. The incredibly high degree of DNA sequence conservation between this cDNA and cDNAs from the other synaptobrevin family members (75% over 250bp in the most conserved region) made identification using DNA sequence searching methods a relatively straightforward process.

Even with the 250bp of sequence data from one strand of the clone it was apparent that the cDNA was not the synaptobrevin cDNA clone that had been previously reported (discussed in chapter 3); there were considerable differences between the predicted amino acid coding regions of the two *Drosophila* cDNA sequences. If the first synaptobrevin protein encoding frame was superimposed over that derived from the pST123 insert (data not shown), and allowing for the presence of two frame shifts due to sequencing errors, changes between the subunits show many synonymous changes. Thus the differences between the sequences appeared to be due to evolution and not to experimental artefacts.

The pST123 insert was restriction mapped using standard restriction enzymes, the restriction map is shown in figure 5.4a. From this it can be seen that the pST123 insert is 2.3kb long, while the two synaptobrevin RNAs reported by Sudhoff *et al.* 1989 were 0.85 and 0.9kb. Again this supports the argument that this was a novel transcript from a second locus. The cDNA restriction map is as expected for the method of library construction used, apart from an



Synaptobrevin Homologue	Percentage Similarity	Percentage Identity	location 5' end
Human 1	90%	81%	-34
Human 2	90%	84%	-32
Drosophila	86%	83%	-48
Rat 1	88%	76%	-34
Rat 2	90%	84%	-32
Yeast (sncl)	66%	38%	-29
Torpedo	90%	81%	-34

Figure 5.4 A Restriction Map of pST123 and the Relationship to Other cDNAs from Different Species

A shows a restriction map of pST123 where **S**-*Sac*I, **Xb**-*Xba*I, **B**-*Bam*HI, **R1**-*Eco*RI, **RV**-*Eco*RV, **P**-*Pvu*II, **Xh**-*Xho*I, **H**-*Hind*III, **Bg**-*Bgl*II, **K**-*Kpn*I. Scale is in base pairs. The cloning vector is shown as an open box.

B shows DNA sequence produced by a using a vector based T3 primer together with the predicted peptide encoded by the sequence fragment. The 5' vector sequence is underlined. The two bases shown in bold were ambiguous in the first sequence read.

C shows the peptide sequence from **B** compared to that of known synaptobrevins. Also shown is the relative location, in amino acid residues, of the start methionine of each of the known synaptobrevins to the most 5' amino acid predicted from pST123.

odd 5' junction lacking a 5' *Eco*RI site. This is a common feature of isolates from this library; 30% of clones lack the 5' *Eco*R1 cleavage site. Sequencing of the 5' and 3' cloning junctions supports the restriction data and reveal a 3'polyA tail remnant. Thus the clone is correctly orientated in the vector and show no evidence of gross rearrangement.

The 5' junction point of the cDNA clone starts in domain two of the predicted amino acid coding region. No known synaptobrevin homologues begin in domain 2 and the clone corresponds to coding sequence up to the 5' junction point. The most likely explanation for this, supported by the northern analysis discussed later (in section 5.5.1), is that the clone is truncated at the 5'end (see figure 4b). The predicted structure of the complete cDNA based on analogy to other family members is shown as figure 4c.

5.3 Partial Sequencing of pST123

pST123 was the first cDNA to be "internally" sequenced as part of this project. Advice was sought from a number of sources as to which of the many sequencing strategies available was the most appropriate to use. In the absence of a consensus, several strategies were tried, including basic subcloning, the use of internal oligonucleotides and exoIII deletions. Selection of a final strategy was based upon this first hand experience gained from this trial run. Taking into account the local conditions, conventional subcloning didn't appear practical. Of the other two, oligonucleotide sequencing was quick, efficient but expensive while exoIII deletions were slower and much more practically intense, but less expensive. ExoIII deletions were chosen as the strategy of choice to sequence one strand of this clone.

ExoIII deletions are described in methods section 2.12 and also in chapter 4.3, I will not repeat these details here. However, in this first attempt to perform exoIII deletions incorporation of α-phosphorothioate nucleotides were used to block the "protected" end as described in the methods section 2.12.2 and illustrated in figure 5.5a. Figure 5.5b shows test exoIII deletions of the deletion template and several controls. Aliquots of the final deletion reaction were also electrophoretically sized through an agarose gel. A photograph of the ethidium bromide stained gel is shown in figure 5c. From these photographs of the ethidium bromide stained gels it is clear that this method

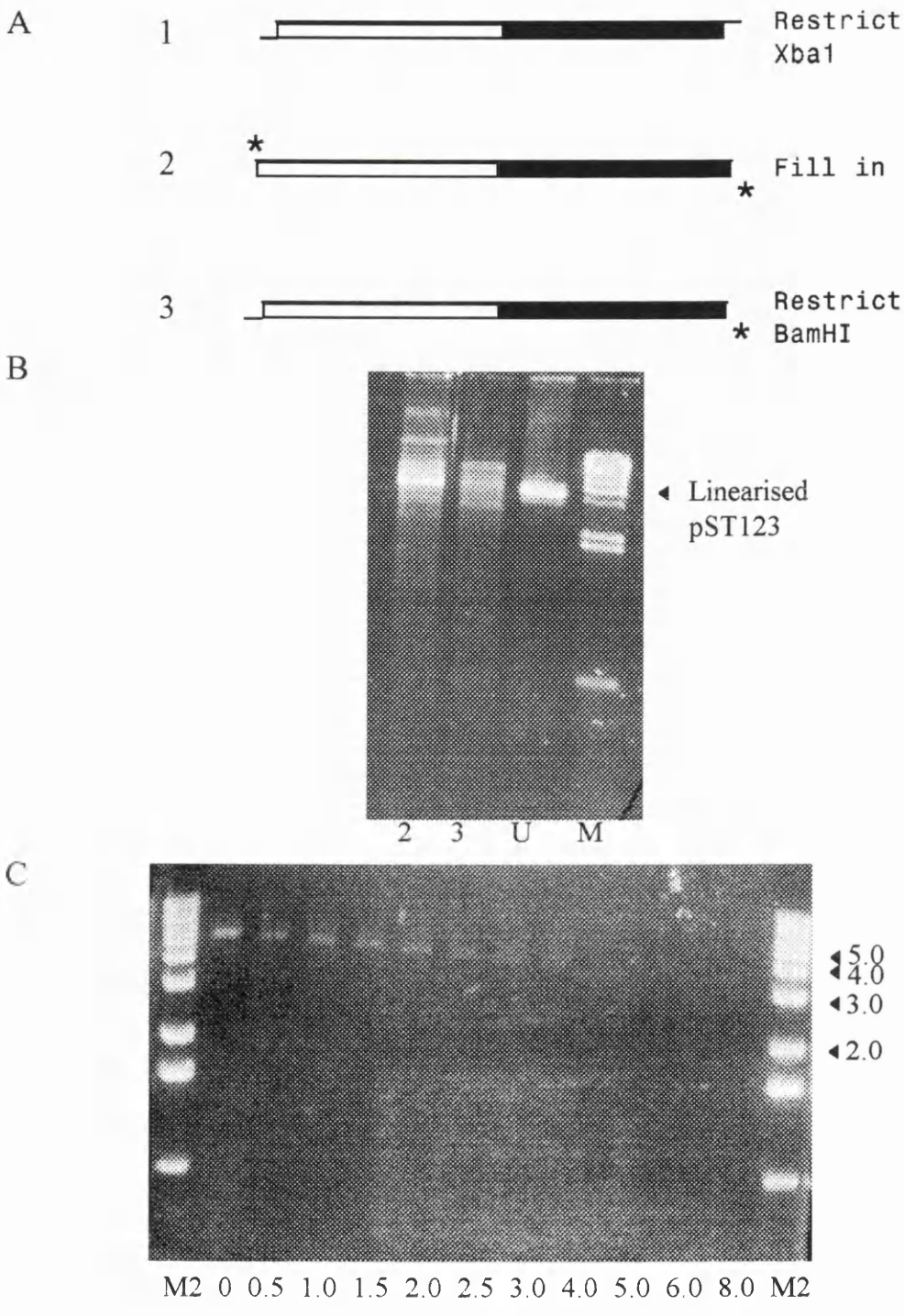


Figure 5.5 Exonuclease Deletion of pST123

A shows the α-phosphorothioate protection strategy used to generate the deletion template. The plasmid was first restricted with *Xba*1 (**1**) before filling in of the 5' overhang with Klenow and α-phosphorothioate bases (**2**). This molecule was then restricted with *Bam*HI to generate the final template (**3**).

B shows the results of agarose gel electrophoresis of template **1** (**1U**) together with the products of test digestion (5 minutes at 37°C, methods section 2.12) with exonucleaseIII(exoIII) of templates **2** and **3** (as labelled). Under these conditions most of the filled in template was protected from exonuclease digestion, while the double cut template was succeptable to digestion.

C show the large scale digestion of template 3 with exoIII and S1 nuclease. Samples of the reaction were taken at the times shown. Time is in minutes. **M2** indicates the position of the 1kb ladder (Gibco-BRL) size markers. Size is shown in kb.

can be successfully used to produce deleted templates. Figure 5.5c also shows how the amount of surviving template decreases over time; as discussed in section 4.3 for 3' overhang based "protection".

Samples from each timepoint were cloned as described in methods section 2.12.5. Five clones were selected from each timepoint, DNA extracted using plasmid minipreps (methods section 2.2.4.1) and a sample of this DNA restricted with *Sac*I and *Kpn*I. From this templates were selected and prepared for sequencing (methods section 2.2.4.2). Figure 5.6 shows the exoIII deletions (and a few subclones and oligonucleotides) used to sequence the single strand of this 2.3kb cDNA. The subclones shown in this figure were prepared from agarose gel isolated fragments and cloned into pBluescript II SK^- by gel purification of restriction digest fragments (methods section 2.2.5) followed by ligation into the vector (methods section 2.2.1).

Sequence data from this cDNA will be discussed in context of that from a second sequenced clone related to pST123 described later in this chapter. However, this part of the sequencing project confirmed several predictions from the first analysis. Extended homology to synaptobrevins could be demonstrated along to a potential stop codon. After this the cDNA seemed to have an untranslated region that extends some 2kb to the 3' polyA tail and T7 cloning site of the plasmid vector. Database searches with this region alone revealed no further homology to any known locus (data not shown). 3' untranslated regions between other synaptobrevins are not conserved. Importantly, the 3' untranslated region shows no evidence for vector related sequences that might remain if the clone had been rearranged or concatenated.

5.4 A Genomic Southern to Demonstrate that pST123 is Likely to be From a Single Copy Gene Distinct from the Locus of Sudhoff *et al.* 1989

The *syb Drosophila* synaptobrevin locus was the target of several rounds of site-selected mutagenesis (Ballinger and Benzer 1989, Kaiser and Goodwin 1990) performed by A. Ducanson (Glasgow Genetics Department). To perform these experiments PCR was used to isolate a genomic subclone containing much of the protein encoding region of this synaptobrevin (henceforth called S-SYB). This provided a readily available DNA probe of this original locus. The cloned PCR product was partially sequenced to confirm its identity. As a further check of this I have restriction mapped the clone with a number of six-cutter restriction enzymes. Recently, the genomic

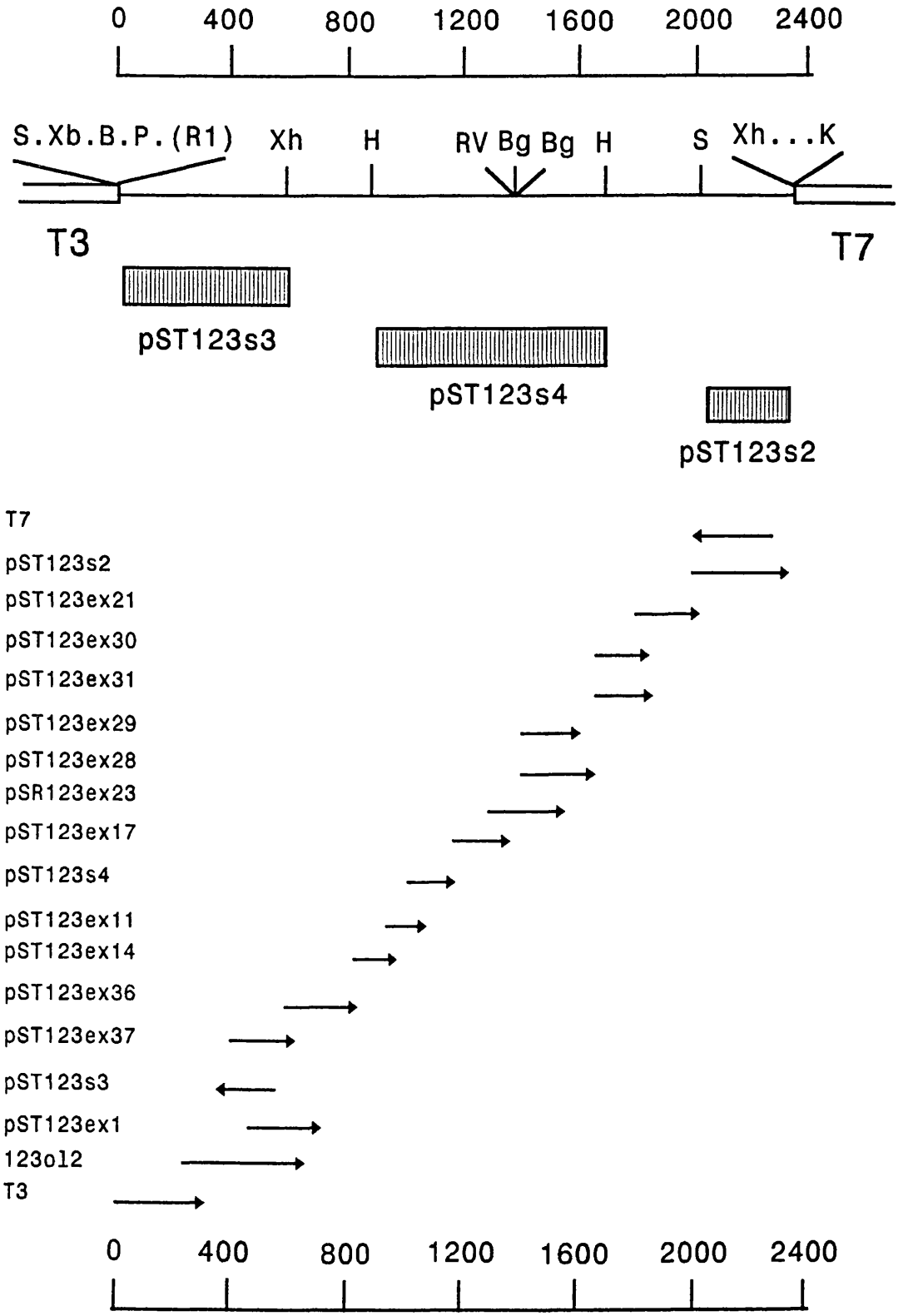
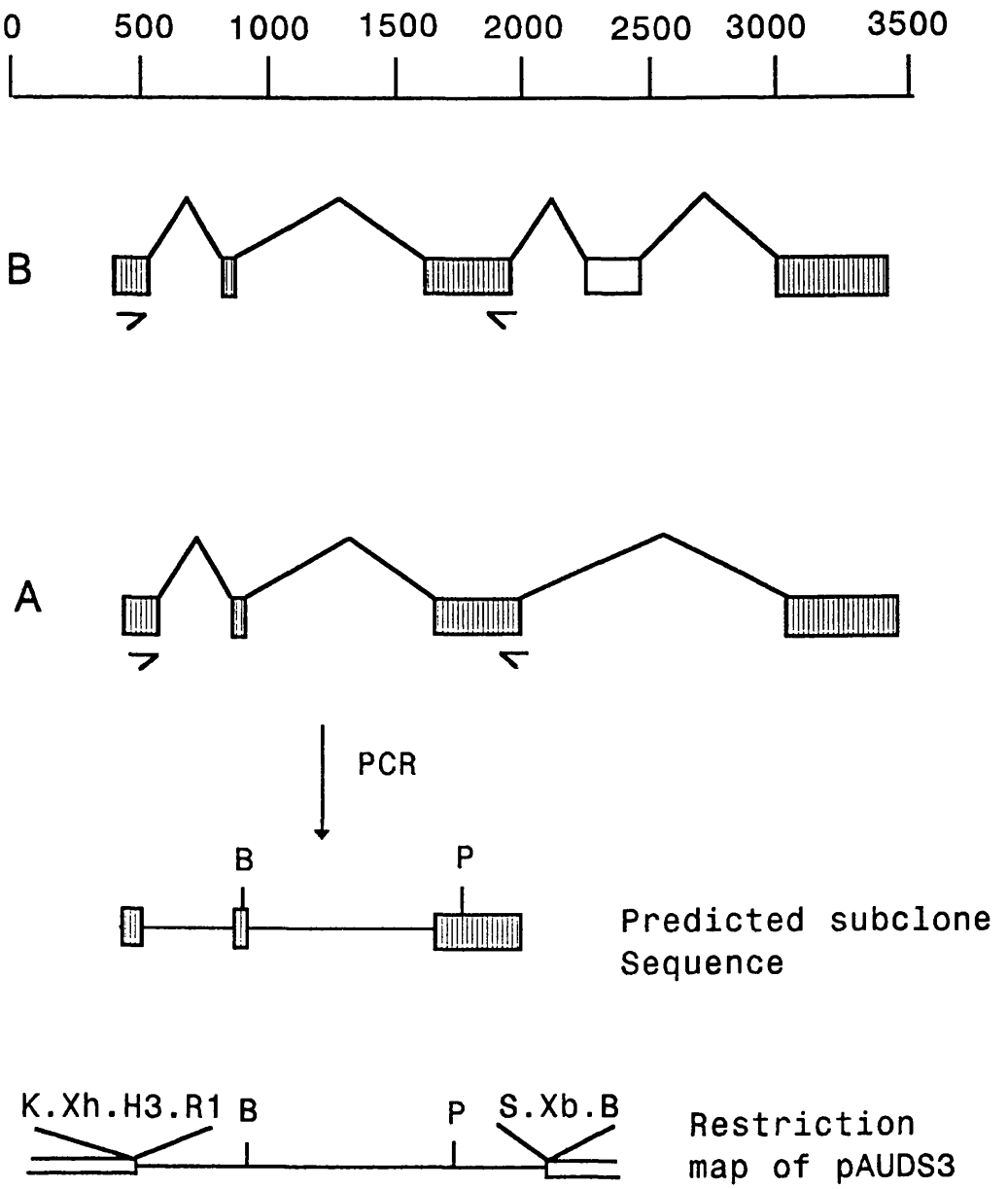


Figure 5.6 Sequencing Extensions Used to Sequence One Strand of pST123
Arrows indicate the direction of sequencing runs used to construct a contig of the sequence of pST123. The sequence prime site was the vector based T3 site of the exoIII deletion unless otherwise indicated. The restriction map of the clone is also shown above for comparison and labelled as figure 5.4. Shaded boxes indicate the location of three subclones used for sequencing and to produce probes used later in the chapter. Scale is in base pairs. The two close *Bgl*II sites cannot be distinguished by restriction analysis; the map shown is corrected using sequence data.



igure 5.7 The Genomic Subclone of the Sudhoff *et al. 1989* Synaptobrevin called AUDS3

he figure shows the genomic region of this locus as reported by Chin *et al.* 1993. Arrows ıdicate the two PCR primers used by A. Duncanson (Glasgow Genetics Department) to ıbclone the genomic region shown. The subclone has been checked by sequencing. Boxed :gions are exons. A and B are the two differentially spliced peptides reported by Chin *et al.* 993. Protein A seems to correspond to the Sudhoff *et al.* 1989 sequence, although the ıblished cDNA sequence seems to lack both 3' and 5' terminal sequence. Restriction sites :e **K**-*Kpn*I, **Xh**-*Xho*I, **H3**-*Hind*III, **R1**-*Eco*RI, **B**-*Bam*HI, **P**-*Pvu*II, **S**-*Sac*I, **Xb**-*Xba*I. Scale in basepairs.

sequence has been released for this region (Chin *et al.* 1993) and this, in relation to the pAUDS3 subclone, made by A. Duncanson, is shown in figure 5.7.

A probe specific for pST123 was prepared by gel purification of the insert fragment from subclone pST123s3 which contains the 0.5kb 5'*Xho*I fragment of pST123 (please refer figure 5.6). This fragment contained an analogous coding region to the S-SYB genomic fragment. The absence of most of the 3' untranslated region does not interfere with the comparison of the hybridisation signals from the two probes. Figure 5.8 shows the genomic Southern blots produced from OrR wild-type *Drosophila* genomic DNA restricted with the enzymes shown.

From the blots shown in figure 5.8a, it is clear that the two probes gave distinct banding patterns with a number of restriction enzymes and hence are likely to be produced from distinct loci. In addition this blot suggests that the pST123 related transcripts are produced from a single locus. Multiple loci might reasonably be expected to produce multiple bands in all lanes as flanking restriction sites to the probe are highly likely to be located in a non-conserved or unique chromosomal context.

5.5 Northern Analysis

5.5.1 Northern Blotting using Head and Body mRNA

A northern blot was prepared as described in methods section 2.9 using OrR (*Drosophila* wildtype) head and body mRNA and total RNA as a control for non-specific ribosomal hybridisation, a common artefact on RNA blots. The blot was probed with a gel purified 5'*Xho*I/*Bam*HI fragment. This blot is shown in figure 5.9 together with a re-probe of the blot using rp49 as a loading control. From this gel it can be seen that the head material contains two abundant transcripts, that are only faintly detectable in the body lane. By comparison to the rp49 (O'Connell and Rosbach 1981) load control the head RNA is approximately evenly loaded relative to that of the body. Size estimates of the two transcripts from the ribosomal bands on the ethidium bromide stained gel and the loading control estimates the bands to be 1.9kb and 3.5kb. Because of the nature of the screen used, and of pST123 itself, it is likely that pST123 corresponds to a truncated version of the transcript producing the larger band. This suggests, when the likely polyA tail length is

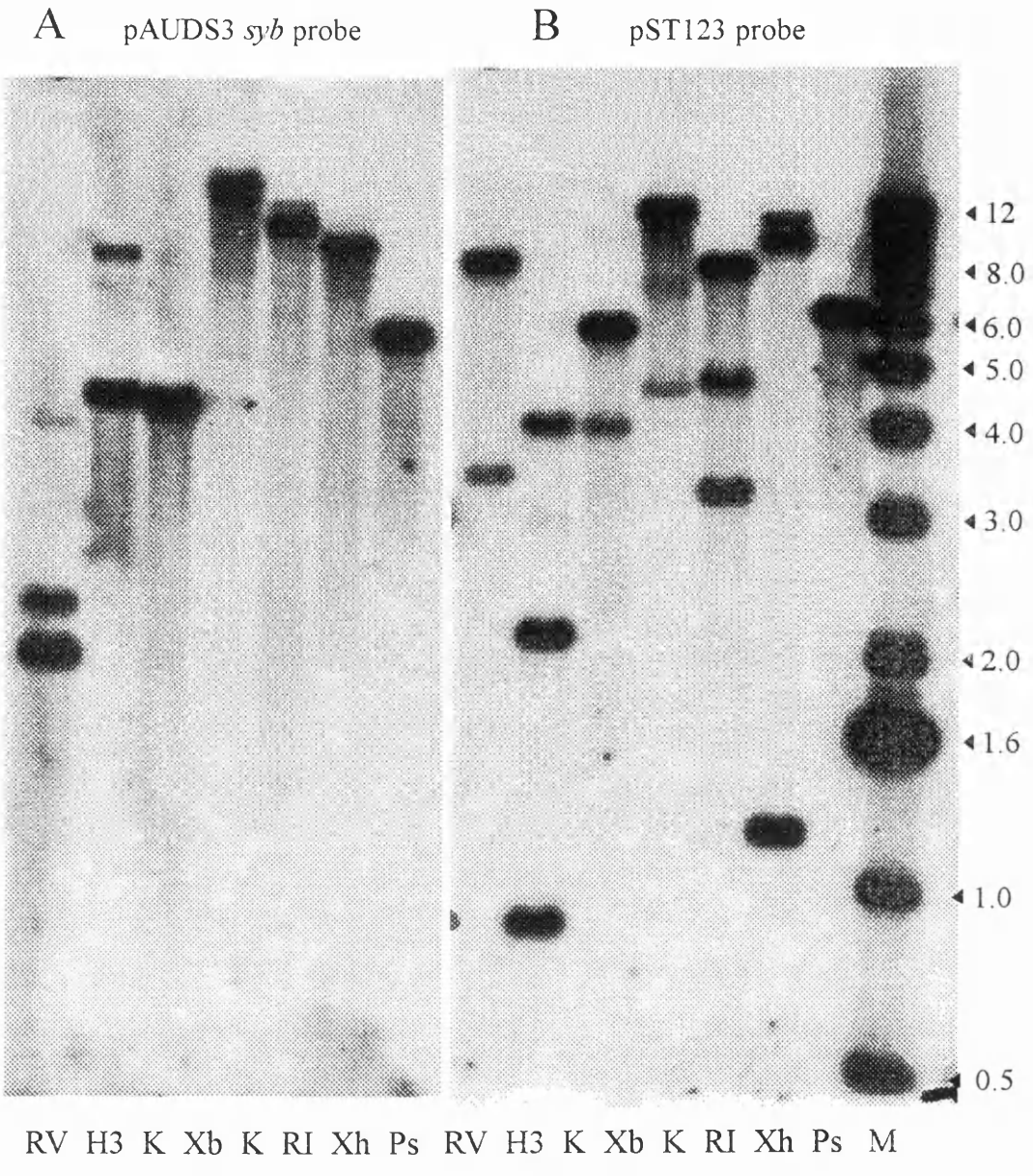


Figure 5.8 Tandem Genomic Southerns using a pAUDS3 and a pST123 Derived Probes

OrR (wildtype) *Drosophila* genomic DNA (methods section 2.5) was digested with a range of six-cutter restriction endonucleases and electrophoretically separated through a 1% agarose gel before blotting onto Hybond-N membrane. Duplicate blots were susequently probed with a ^{32}P labelled probe generated from pAUDs3 and pST123s3 (methods section 2.6).

A shows the autoradiograph produced using the pAUDs3 probe and **B** shows the autoradiograph produced using the pST123s3 insert probe. Restriction digestions were **RV**-*Eco*RV, **H3**-*Hind*III, **Xb**-*Xba*I, **K**-*Kpn*I, **RI**-*Eco*RI, **Xh**-*Xho*I, **Ps**-*Pst*I. M is a size marker generated from a 1kb ladder (Gibco-BRL) end labelled with ^{32}P using Klenow polymerase (methods section 2.10.2). Scale is in kilobase pairs.

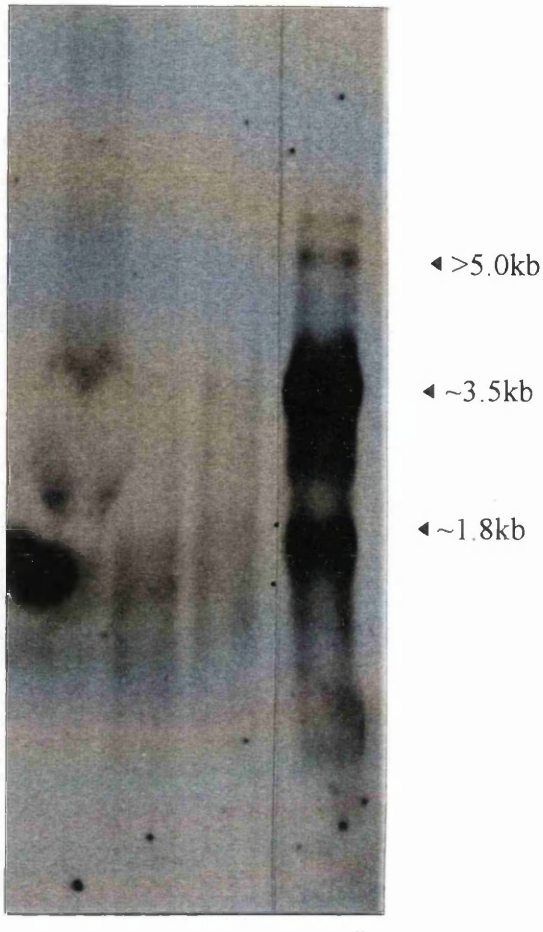


Figure 5.9 A Using a pST123 Probe Shows Trnascripts that are Found at Elevated Levels in the Head Compared to Body.

The blot was produced as described in methods section 2.9. Each track contains an estimated 1-2 μg of polyA mRNA. Total RNA (**T**) was used as a control against ribosomal hybridisation, a common artifact on northern blots. Also labelled on the blot are the lanes containing *eya* body mRNA (B+e), OrR (wildtype) body mRNA (B+o) and OrR head mRNA (H+).

The pST123 probe was prepared from *Sac*I digested pST123.. This fragment includes the protein coding region of pST123 together with ~1kb 3' untranslated region.. A probe with 30μCi of incorporated ^{32}P radioactivity was prepared using the random priming (methods section 2.10.1). The blot shown is from a two day exposure of the autoradiograph. An exact duplicate blot has been probed with RP49 (shown as figure 4.15) and the result indicates that the head track is about five-fold underloaded as compared with that of the body. Sizes shown are in kilobase pairs, estimated from ribosomal bands and marker.

taken into account, that pST123 lacks about 1kb of the 5'end, much of which might be expected to be 5' leader sequence.

5.5.2 A Developmental Northern Using a Probe From pST123

The preliminary analysis of the expression pattern detected using a pST123 derived probe presented in chapter 3.0 indicated that the clone seemed to have detectable expression not only in the head and body, but also in mixed embryos and using mid-pupal prepared probes. This expression appeared greater than was detectable using a body cDNA probe. To investigate these expression predictions further a developmental northern was prepared using mRNA generated from embryonic, first and third larval instars, mixed pupal (methods section 2.14.3) and from adult head and body OrR wild-type message (embryonic and L1 staged message was a kind gift from A.Griffin, Glasgow Genetics Department). This blot was probed with a gel purified fragment generated from pST123s3 and is shown as figure 5.10, together with an autoradiograph produced after the blot was re-probed using rp49 (O'Connell and Rosbach 1981) as a loading control.

Unfortunately in this experiment, the OrR head mRNA shows some indication of being slightly degraded and so in the long exposure migrates as a smear. However, two distinct bands are visible in the shorter exposure, estimated to be of the size produced with the *Xho*I fragment derived probe shown in figure 5.9. From this it seems unlikely that the locus produces any other abundant transcripts that are only homologous to the 3'untranslated region of pST123. The long exposure of the blot is sufficiently smeared to obscure any rare bands. However, it can be seen from the blot that the pST123 derived probe produces several developmentally regulated transcripts. For example, a large transcript of ~3.6kb can be seen as an abundant transcript in the L1 preparation, while it is less so in the embryonic preparation and is not clearly seen in the head or body preparation. Reference to figure 5.9 shows the very faint presence of a similar sized band in the head track but it is at least twenty fold under-represented when compared to it's relative abundance in the L1 preparation. This large band is also completely absent from the L3. Indeed the L3 stage shows little evidence of expression (refer rp49 control for loading). It is possible that a weak band may be obscured in part of the L3 track because of local partial transfer. The large 3.5kb band and indeed the

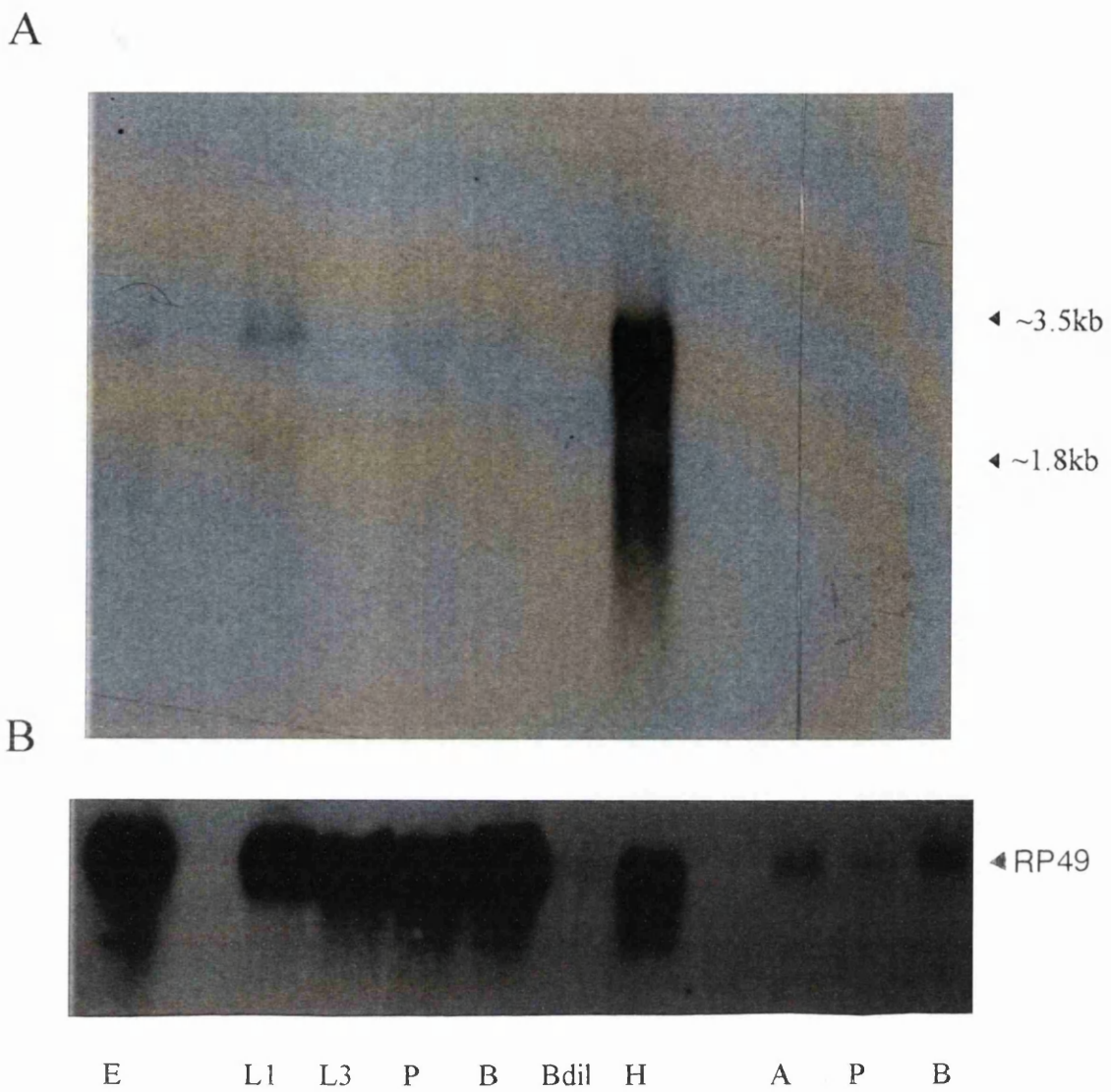


Figure 5.10 A Developmental Northern Probed with a pST123 Derived Probe
A shows a developmental northern prepared as described in methods section 2.9 probed with a ^{32}P labelled probe generated from the pST123s3 subclone fragment using random priming (methods section 2.10.1). Estimated size (from ribosomal bands) is shown in kb.
B shows the same northern blot reprobed with a RP49 derived probe as a loading control.
1μg of polyA+ mRNA was prepared from *Drosophila* OrR mixed embryo (**E**), first larval instar (**L1**), mixed third instar larvae(**L3**), mid-pupal (**P**), body (**B**) and head(**H**). 0.1μg Body mRNA was also loaded (**Bdil**). 5μg of total RNA was also loaded derived from mixed adult (**At**), pupal (**Pt**) and body(**Bt**).

other L1 and head abundant bands are outside this region. Several unique rarer bands may also be present in the mixed embryo preparation.

Time considerations prevented the repeating of this developmental northern to a standard completely to my satisfaction. But this preliminary blot raises a number of important questions. It would have been very informative to be able to compare the developmental expression to the Sudhoff *et al.* 1989 synaptobrevin (S-SYB); although this data has not been published. I would predict that this locus is abundantly expressed in L3s for example, which might reasonably be expected to be engaged in synaptic transmission. But why does the ST123 gene produce developmentally regulated transcripts? What may be distinguished here by these different transcripts are different neuronal types perhaps present at distinct stages in different proportions. Or perhaps the ST123 gene products are engaged in other cellular processes at these stages. Indeed Cain *et al.* 1992 have recently reported the involvement of synaptobrevins in a non-neuronal process. They demonstrated that synaptobrevins are also involved in the release of the GluT4 glucose transporter from adipocytes. This process also occurs via small vesicle based exocytosis.

To date the blot shown here is the first developmental transcript regulation reported for a member of the synaptobrevin gene family.

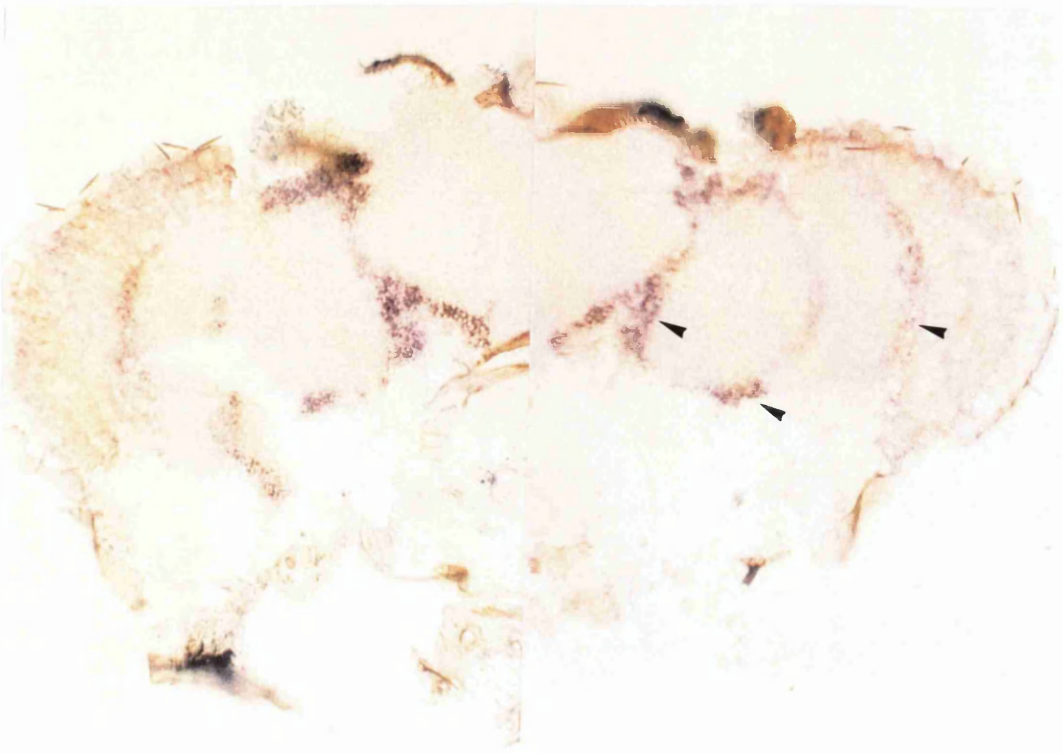
5.6 Localisation of the Novel Synaptobrevin RNA in Head Sections and in Mixed Embryos

Gene specific probes were prepared from pST123 and used to determine the tissue specific distribution of the parent RNA. Head sections were of particular interest because of the known patterns of distribution of the two synaptobrevin homologues in rat (Trimble *et al.* 1990). Embryonic *in situs* were performed to account for the detected expression in the embryo, especially to determine whether the encoding RNA exhibited only neuronal expression.

5.6.1 Expression in Cryostat Prepared Head Sections

In situ hybridisation was performed as described in methods section 2.18. A pST123 specific probe was prepared from *Sac*I digested related clone pST123r27 (refer figure 5.14). Representative 10μM sagital sections probed with these DIG labelled probes are shown as figure 5.11.

A



B C

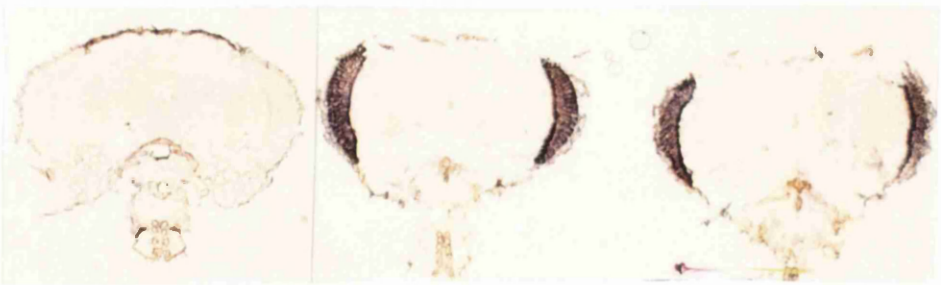


Figure 5.11 Localisation of RNA that Hybridises to a pST123 Probe

Probes were generated from the pST123 related clone pST123r27 (figure 5.16) digested with *Sac*I. Probes were prepared by 30 rounds of extension with a vector based T3 primer using *Taq* polymerase (as methods section 2.2.2) in the presence of DIG labelled nucleotide. An opsin probe was similarly generated from the clone pST41 and used as a positive control. A negative control was prepared from the anti-sense strand of the pST41 insert. (I acknowledge the assistance of S. Goodwin, Glasgow Genetics Department in setting up the probe synthesis reactions). Representative 10μM sagittal sections probed with these DIG labelled probe.

A shows the synaptobrevin hybridisation pattern, **B** shows the negative control and **C** shows the positive. The probe shows detectable expression in the major regions of the brain (arrowed) that contain the neuronal cell bodies.

Expression can be seen to be localised to the cortex of the central brain. Expression can also be seen in the cortical regions of the eye. No central brain expression can be detected. The central brain is composed of neurophil and should not hybridise to an RNA probe that is localised to the neuronal cell bodies. The RNA expression pattern shown is consistent with a neuron specific transcript that is expressed in most or all neurons of the nervous system. In this sense the pattern of expression is reminiscent of that found for rat synaptobrevin-2, in that it is abundantly distributed in the central nervous system. However, expression in nervous system specific non-neuronal cell types cannot be excluded by this analysis.

5.6.2 Expression in Embryos

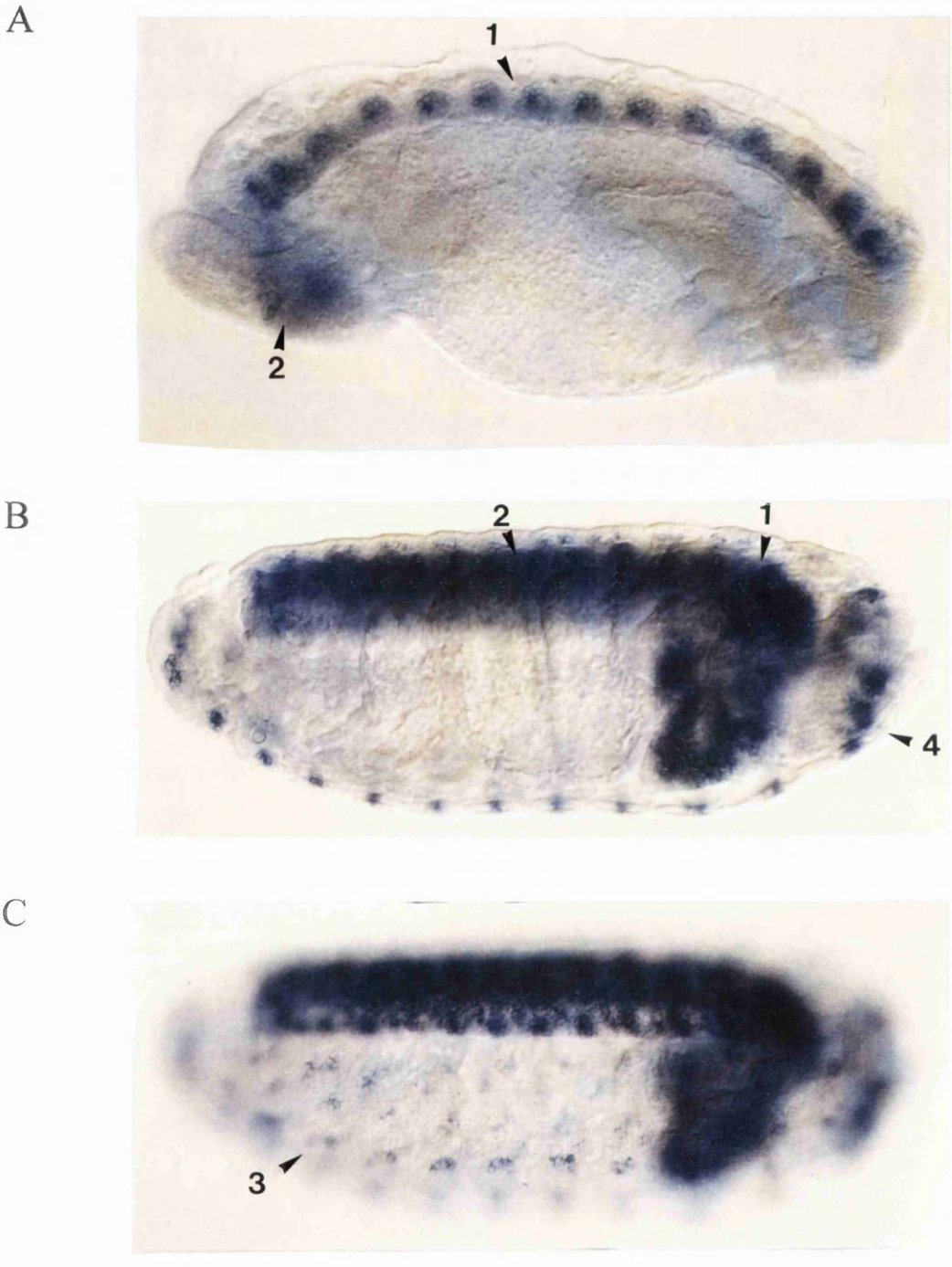
Whole mount *in situs* were performed on mixed embryo collections as described in methods section 2.17 using the pST123s3 subcloned fragment. The results of several such hybridisation experiments is shown in figure 5.12.

Localised expression starts at stage 13 and can be seen in the ventral nerve cord. Expression in later stages follows the contraction of the nerve cord and is almost exclusively nervous system specific. Expression does not appear to be regional within the central nervous system. Curious expression in regions located close to the cuticle is seen in each segment of later stage embryos. This is interpreted as expression in some components of the peripheral nervous system.

The presence of several transcripts in the embryo continues to raise the possibility that specific transcripts are spatially or temporally localised. NB. These transcripts are only weakly detected by the northerns blot (figure 5.10) but are detected strongly using the *in situ* protocol because the northern lane was generated from a mixed embryo population.

5.7 A screen to isolate a Longer Clone Related to pST123

Several strategies are commonly used to isolate further fragments of a known cDNA, the most common being to rescreen the cDNA library and select for suitable clones or to use a 5' pointing gene specific PCR primer to amplify DNA between it and a 3'pointing primer (5'RACE). In this investigation only the library screening approach was adopted. It was considered that the target cDNA would be common in the *eya* head library used and that a conventional



5.12 Localisation of pST123 Parent RNA(s) in Whole Mount *Drosophila* Embryos

Hybridisations were performed as described in methods section 2.17. Probe incorporation was assayed by trace labelling the synthesis reaction with ^{32}P. Two probes that gave negligible incorporation gave no detectable signal on *in situ* hybridisation and were used as negative controls (not shown).

Segmental expession can first be detected between stage 13 (**A**) and follows the development of the CNS. Staining can be detected in both the developing nerve cord (**1**) and the central brain (**2**). Segmentally arranged hybridisation, likely to be to components of the peripheral nervous system, can also be detected in later staged embryos(**3**), together with staining in the antenomaxillary complex(**4**). Contaminating first instar larvae show strong CNS staining, indicatiing that expression continues into the larval stages.

screening approach is a far more desirable route than PCR because of the problem of assigning a particular PCR product to the mRNA of interest. To do this it is often necessary to re-screen the library with the new fragment.

In order to increase the likelihood of finding long cDNAs the 5' *Xho*I fragment isolated from pST123s3 was used to prepare the probe. The library was screened as described in methods section 2.14.2 but with randomly plated plaques. The probe was prepared by random priming as described in methods section 2.10.1. This probe gave approximately 20 positive clones in 50,000. These were rescreened by replating each positive and re-screening. The clones were then excised as pBluescript II SK-'phagemid from the ZAPII clones (methods section 2.3.6), restricted with *Bam*HI and *Xho*I to liberate the inserts and also with *Hind*III, and electrophoresed through a 1% 1x TBE agarose gel used to screen the library. From the positives purified, no clones gave a *Xho*I/*Bam*HI fragment larger than that found in pST123 together with a characteristically sized *Hind*III fragment. As a confirmation a duplicate blot was probed with *Hind*III. Several clones were identified with distinct restriction maps, or that appeared to be truncated versions of the original pST123 (data not shown).

The results of the first screen were rather surprising. The cDNAs abundance in wildtype mRNA populations can be estimated from the northern blots and reverse Northerns by comparison to positive controls to be about 0.3%. No elevation of expression in the eye is observed from the *in situ* data and so one might reasonably expect to obtain several hundred positives in such a screen. The ten fold reduction in this number is suggestive of a rather large number of truncated versions of this cDNA in the library.

A second screening of the *eya* head cDNA library was later instigated using a probe prepared from the subcloned insert isolated from pST123s2. This insert encompasses all but the very most 3' 200bp of pST123. In this screen, 0.4% of plated clones appear to hybridise with this probe; not significantly different from the number of positives expected. Secondary positives from this screen were rescreened with the 5'*Xho*I fragment after digestion with *Xho*I/*Xba*I and *Hind*III. Once again, none of the positives show a strongly hybridising *Xho*I/*Xba*I fragment together with a characteristically sized *Hind*III fragment. However, two of the secondary positives hybridised strongly to the probe with an identically sized ~0.8kb fragment. They did not possess the *Hind*III

fragment that would be located 3' of the hybridising region if the clones were a longer version of pST123.

To assess the coding capacity of the two positives, designated pST123r17 and pST123r27, they were both used as templates in a PCR reaction. This employed a primer derived from the "T3" region of the pBluescript II SK minus vector (kindly supplied by C.Millagan, Glasgow Genetics Department) and 123ol2. pST123 was used as a positive control As shown in figure 5.13, PCR of these two positives gives a PCR band of 880bp compared to the band produced from pST123 of 300bp. This extra 580bp would be sufficient to accommodate ~200AA, much greater than would be reasonably expected to be located at the 5' end of pST123. As a final confirmation, sequencing runs were performed using the same two primers used above. The T3 primer identifies sequence not found in the other clone, while 123ol2 identifies sequence previously identified as related to that of other synaptobrevin cDNAs but that extends beyond the 5' cloning boundary of pST123. Hence, although this new class of cDNA is distinct from pST123 at the 3' end, carries more DNA sequence at the 5' end. pST123r27 is likely to encode a full protein coding region.

5.8 Preparation for Full-Length Sequencing of pST123r27

5' and 3' sequencing extensions from T3 and T7 primers reveal that both pST123r17 and pST123r27 have identical T3 and T7 junction points. This raises the possibility that the clones were not derived from independent cloning events during the construction of the primary library. pST123r27 was chosen as representative of the two, and restriction mapped (methods section 2.2.5) as a prelude to full-length sequencing. Figure 5.14 shows the restriction map of pST123r27 and, for comparison, pST123. Also shown is the predicted restriction map of S-SYB, based upon the published DNA sequence for this cDNA "hypothetically directionally cloned" into the pBluescript SK minus cloning vector. The hatched box in the figure shows the fragment that cross-hybridise to the 5'*Xho*I fragment. The position and direction of hybridisation of pST123ol2 is also shown. [pST123ol2 was designed for use in several experiments and was deliberately constructed so that it would prime from both pST123 and S-SYB DNA sequence].

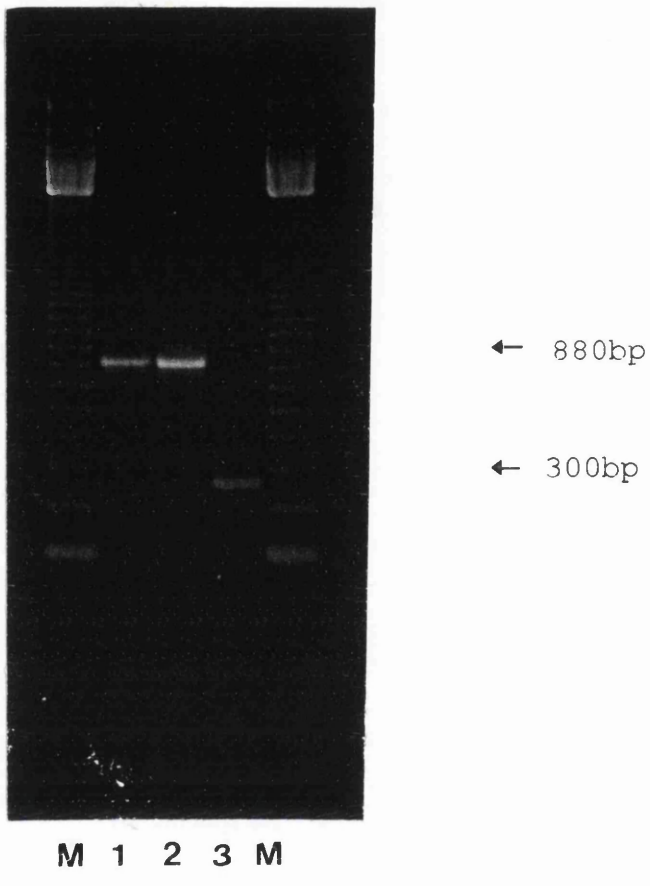


Figure 5.13 PCR of 5' Regions of pST123r17, pST123r27 and pST123.
50ng each plasmid was used per PCR reaction (methods section 2.2.2) using 123ol2 and a vector based, 3' pointing primer called T3 (kindly supplied by C. Millagan, Glasgow Genetics Department). The PCR products were electrophoretically separated using a 1% agarose gel and stained with ethidium bromide. **M** shows the position of the 123bp ladder size markers (Gibco-BRL) , **1** shows the products of the reaction using pST123r17, **2** using pST123r27 and **3** using pST123.
The gene specific primer 123ol2 suggested that pST123r17 and pST123r27 contain sequences homologous to synaptobrevins. The difference in size between the pST123 and pST123r27 bands indicates that pST123r27 extends 5' to pST123 by 580bp (880-300=580bp). This is sufficient to specify 190 amino acids. The expected "missing" 5' region was 29-48 amino acid residues (refer figure 5.4). Hence pST123r17 and pST123r27 are likely to contain the desired 5' coding capacity.

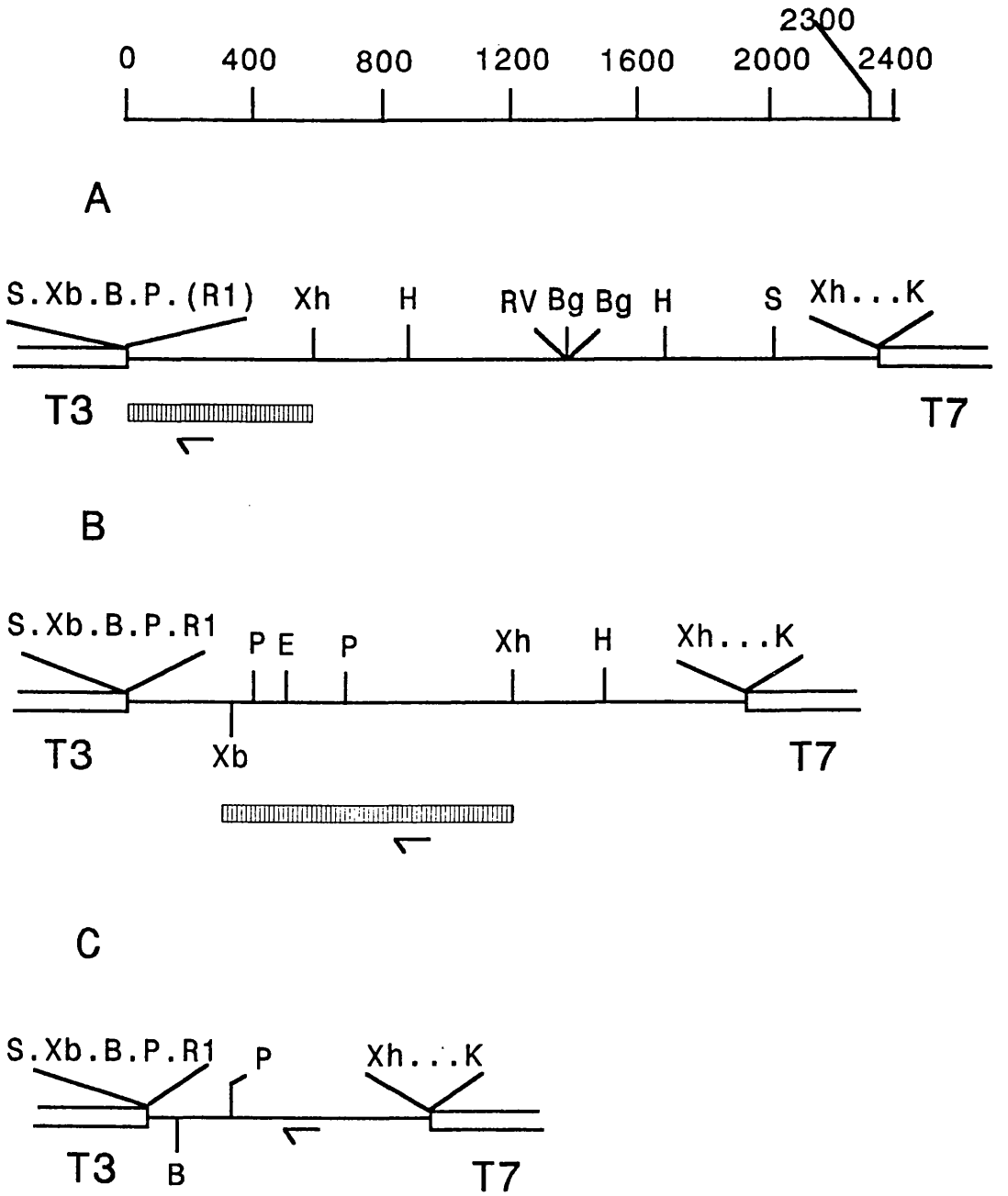


Figure 5.14 Restriction Maps of Synaptobrevin Clones pST123(A) and pST123r27(B) and that of Sudhoff *et al.* 1989 (C)

Where S-*Sac*I, Xb-*Xba*I, B-*Bam*HI, R1-*Eco*RI, RV-*Eco*RV, P-*Pvu*II, Xh-*Xho*I, H-*Hind*III, Bg-*Bgl*II, K-*Kpn*I restriction sites. Scale is in base pairs. Maps **A** and **B** are as generated by restriction mapping, and are supported by sequence data. Map **C** is a restriction map generated from published sequence (Sudhoff *et al.* 1989) as if it were directionally cloned into pBluescriptII SK-. Boundary vector sequence is shown open boxed, with the orientation of the pBluescript SK$^-$ vector indicated by either T3 or T7. Hatched boxes indicate the region of pST123r27 known to cross-hybridise with the 5*Xho*I/*Bam*HI fragment from pST123, (also shown hatched). For continuity, the T3 pointing oligonucleotide 123ol2 designed to hybridise to both pST123(**A**) and the S-SYB cDNA(**C**) is shown as an arrow indicating position and direction of priming on each clone.

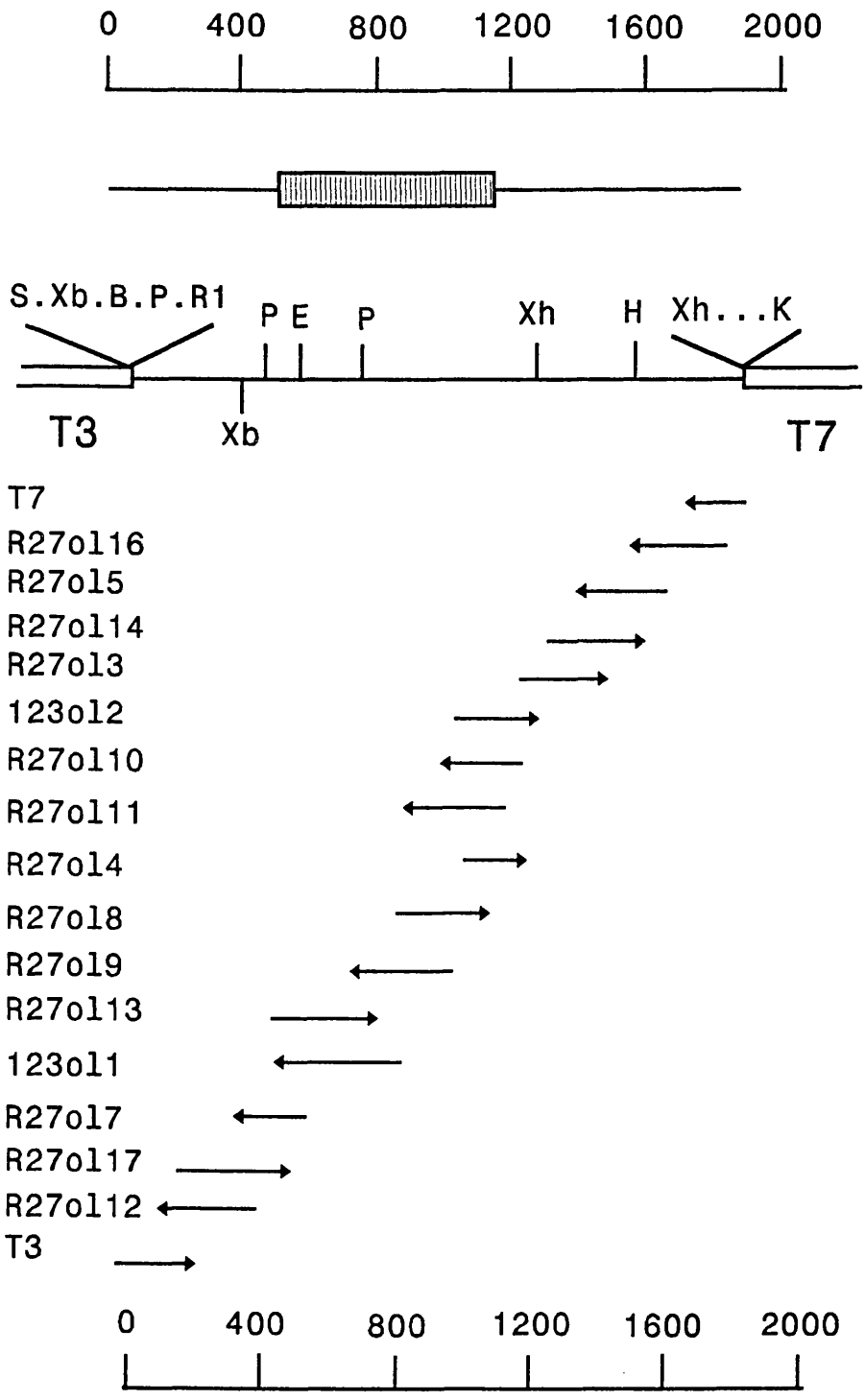


Figure 5.15 Sequencing Extensions used to Produce pST123r27 Consensus Sequence

Arrows indicate the length and direction of the sequencing extension. Extensions are labelled with the primers used, T3 and T7 are vector based primers (Promega). Scale is in base pairs. The restriction map of pST123r27 and the map of the predicted coding sequence are also shown for reference. The predicted amino acid coding region is shown boxed. The restriction enzyme cleavage sites are abbreviated as shown figure 5.14.

5.9 Sequencing of pST123r27 using Oligonucleotides

Oligonucleotide sequencing is described in methods section 2.13. The sequences derived from T3, T7, pST123ol2 and pST123ol1 (which was also found to prime pST123r27) were used as sequencing entry points for the sequential design of oligonucleotides. The oligonucleotides used to sequence pST123r27 are shown in figure 5.15. The putative 5' leader and 3' untranslated regions were sequenced in one strand, while the cloned amino acid coding region, was sequenced in both strands.

5.9.1 The relationship between pST123 and pST123r27

Striking sequence similarity is seen between the pST123r27 DNA sequence and that of the pST123 cDNA. The prediction first held from the restriction map is that pST123r27 extends considerably 5' of pST123, but contains 1.3kb less 3' untranslated region. Once both cDNAs were sequenced it is clear that the sequence of the cDNAs is identical in the overlapping region. The pST123r27 insert polyA tail remnant begins in an AT rich region of pST123 and so it is possible that pST123r27 insert is an internally primed version of pST123. However, since their are two major bands on the northern blot, at least in OrR flies, it is possible that the two cDNAs represent true mRNAs. This relationship is summarised in figure 5.16.

5.9.2 DNA Sequence Features

Figure 5.17 shows the full length DNA sequence of pST123r27. As stated in the previous section this work is yet to be completed ~400bps remain to sequenced in both strands. In figure 5.17 the cloning junction points are shown underlined. In accordance with the restriction map the 5' junction point of this clone is not as expected, it lacks the vector based *Eco*RI site. The sites vector side of the junction point are intact. Lack of this *Eco*RI site is a common feature of these clones as previously described. The 3' junction point appears as predicted; no clone yet sequenced (n~20) has shown a non-standard junction point at the 3' T7 termini. The vector based restriction sites of *Bam*HI and *Xho*I used to cut the insert out of the vector are shown in bold.

The predicted amino acid coding region is shown aligned to the DNA sequence in figure 5.17. Homology to the other known synaptobrevin sequences is strongly supportive of the predicted reading frame. However, the

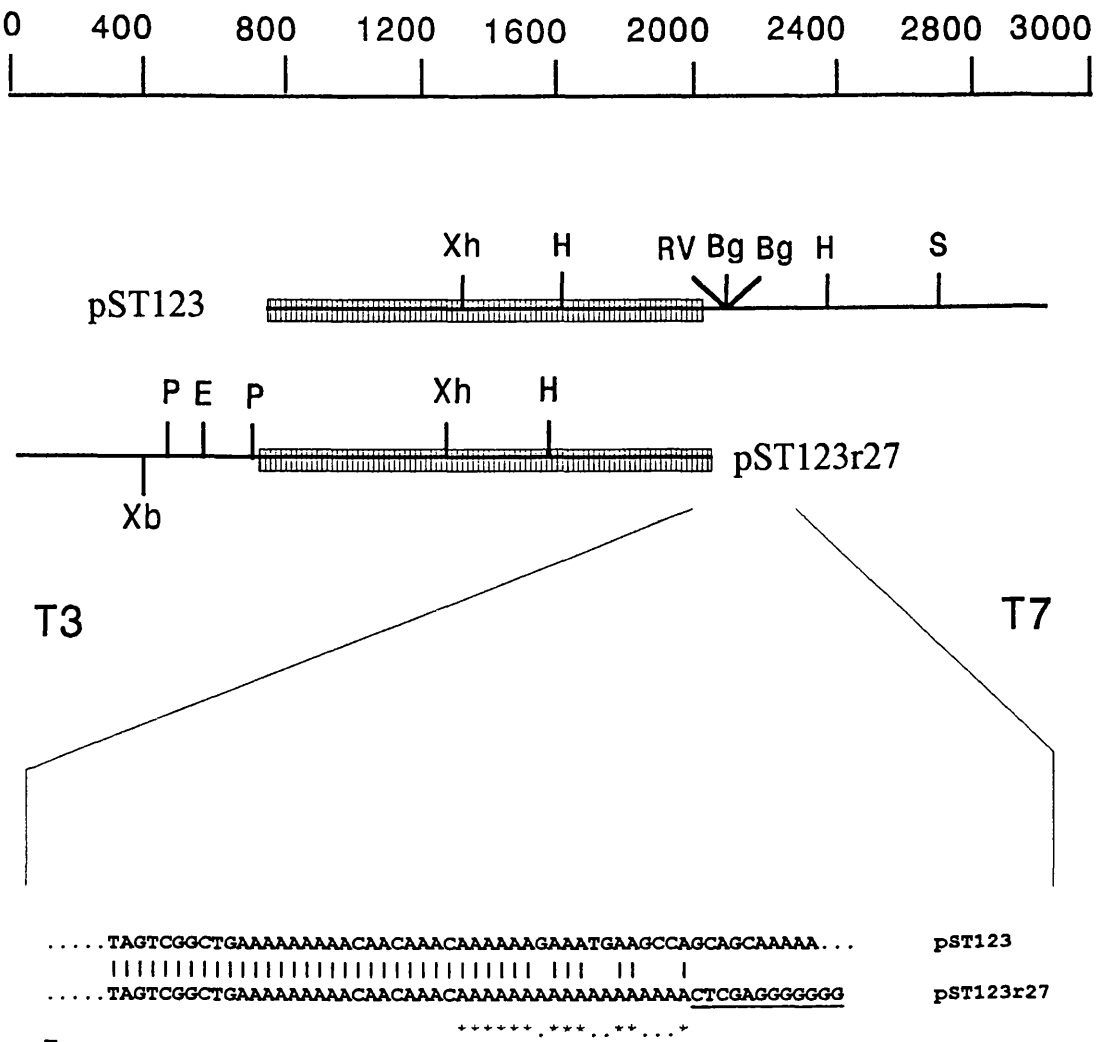


Figure 5.16 The Detailed Relationship between pST123 and pST123r27
The restriction maps of the two cDNAs are shown aligned. The boxed region is identical in both clones. pST123r27 extends 680bp further 5' than pST123 to include a putative translation start site missing from pST123. pST123 extends over 1kb further 3' than pST123r27, this region is presumably additional untranslated region.
The 3' oligo dT prime site of pST123r27 is located in an A rich region of pST123, as shown by the included sequence. The region underlined of the pST123r27 sequence shown is from the T7 end of the vector. Stars indicate bases that could base-pair if pST123r27 was internally primed from the same mRNA template that encoded pST123. Similarly, periods indicate those residues that could not base pair. Overall in the possible internal region there are a maximum of 12/18 base pairings.

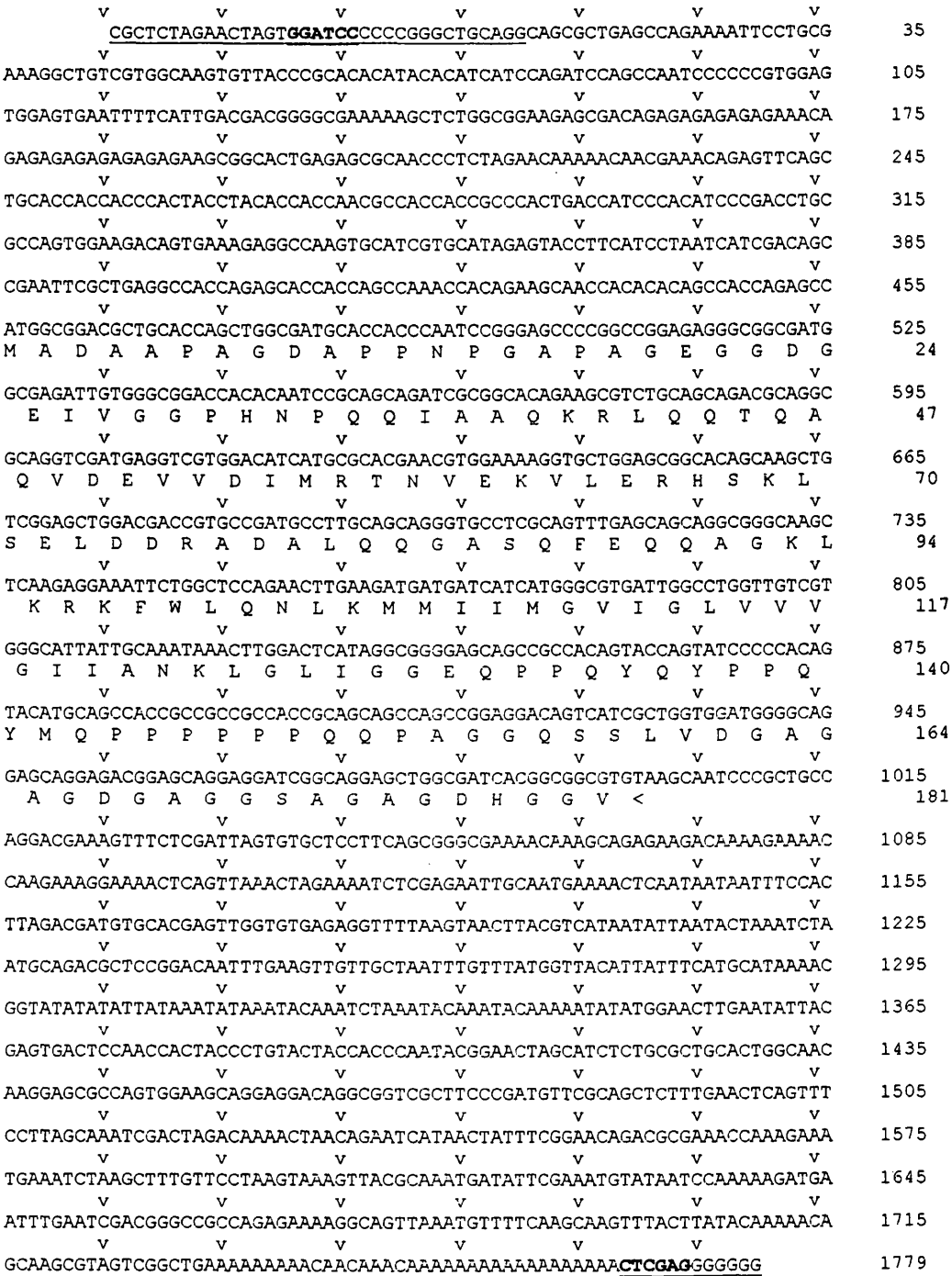


Figure 5.17 Nucleic Acid Sequence, and Predicted Amino Acid Sequence from pST123r27

5' and 3' junction points are shown underlined. Flanking vector sequence is pBluescript II SK$^-$. The expected 5' *Eco*RI site has been lost during cloning. The *Bam*HI site shown in bold was commonly used to remove the insert from the vector, together with the 3'*Xho*I site which is also shown in bold in the 3' polylinker. Numbering is of base pairs and starts at the most 5' base of the cDNA. The predicted amino acid sequence is also numbered. The sequence details shown here are putative.

N- and C-termini are not well conserved in known synaptobrevins in either their length or coding content. The ATG start of translation is the only methionine in-frame with the coding region. A second methionine is specified in another forward frame, but this looks an unlikely start site as it requires a frame shifting sequencing error to be used for the main protein coding region (data not shown). The sequence reading in this region has been double checked in both strands. At the carboxy terminal the reading frame progresses longer than any previously described synaptobrevin.

The regions flanking the putative coding regions is predicted to be non-coding using the testcode program (Fristensky 1982, methods section 2.22.11). The very 3' of the cDNA insert has an 18 residue long cloned poly A tail, but no consensus polyA signal is found upstream of this tail. Several sites are near matches and may be used as alternatives. It is also possible that this cDNA is internally primed as previously described, although the pST123 does not itself contain a consensus poly A signal. The absence of such a signal in either of these cDNAs is not a cause for concern, polyA signals are quite often absent from cDNA clones e.g. Sudhoffs synaptobrevin sequence (Sudhoff *et al.* 1989), and for sites close to the tail itself in both the Na+/K ATPase β-subunit sequence presented here and in the related shrimp sequence (although two are found in the *Drosophila* α-subunit sequence). Unfortunately, issues of polyadenylation and of 3' untranslated region are seldom addressed in primary cloning papers and are often not followed up in future work.

5.9.3 The Putative Amino Acid Coding Region

Figure 5.18 shows a ClustalV multiple alignment (methods section 2.22.13) of representative synaptobrevin family members and figure 5.19 shows this same alignment that has been shaded using the Boxshade program (methods section 2.22.13).

The significant internal feature of synaptobrevins is the transmembrane region. Hydropathy plots (not shown) indicate that the region underlined for pST123r27 on figure 4.18 is likely to be a transmembrane region. This aligns well with those predicted for other synaptobrevins, although the region is not as well conserved between family members as in the more N-terminal regions. Some synaptobrevin family members only possess a few residues on the side of the membrane thought to project into the vesicle. In this case however the molecule seems to project ~60AA into the vesicular space.

```
hum1        MSAPAQPPAEGTEG------TAPG--GGP------PGPPPNMTSNRRLQQ
hum2        MSATAATAPP----------AAPAGEGGP------PAPPPNLTSNRRLQQ
dro1        MENNEAPSPSGSNNNDFPILPPPPNANDNYNQFGDHQIRNNNAAQKKLQQ
rat1        MSAPAQPPAEGTEG------AAPG--GGP------PGPPPNTTSNRRLQQ
rat2        MSATAATVPP----------AAPAGEGGP------PAPPPNLTSNRRLQQ
snc1        MSSSTPFDPYALSEHD----------------EERPQNVQSKSRTAE
torc        MSAPPSGPAPDAQG------GAPGQPTGP------PGAPPNTTSNRRLQQ
R27         MADAAPAGDAPPNP------GAPAGEGGDGEIVGGPHNPQQIAAQKRLQQ
            *                                          .  ..  .

hum1        TQAQVEEVVDIIRVNVDKVLERDQKLSELDDRADALQAGASQFESSAAKL
hum2        TQAQVDEVVDIMRVNVDKVLERDQKLSELDDRADALQAGASQFETSAAKL
dro1        TQAKVDEVVGIMRVNVEKVLERDQKLSELGERADQLEQGASQFEQQAGKL
rat1        TQAQVEEVVDIIRVNVDKVLERDQKLSELDDRADALQAGASVFESSAAKL
rat2        TQAQVDEVVDIMRVNVDKVLERDQKLSELDDRADALQAGASQFETSAAKL
snc1        LQAEIDDTVGIMRDNINKVAERGERLTSIEDKADNLAVSAQGFKRGANRV
torc        TQAQVEEVVDIIRVNVDKVLERDQKLSELDDRADALQAGASQFESSAAKL
R27         TQAQVDEVVDIMRTNVEKVLERHSKLSELDDRADALQQGASQFEQQAGKL
             ** ... *.*.* *..** **  .*. . ..** *  .*  *   * ..

hum1        KRKYWWKNCKMMIMLGAICAIIVVVIVI----------YFFT------
hum2        KRKYWWKNLKMMIILGVICAIILIIIIV----------YFSS------
dro1        KRKQWWANMKMMIILGVIAVVLLIIVLVSV--------------------
rat1        KRKYWWKNCKMMIMLGAICAIIVVVIVI----------YIFT------
rat2        KRKYWWKNLKMMIILGVICAIILIIIIV----------YFST------
snc1        RKAMWYKDLKMKMCLALVIIILLVVIIVPIAV-------HF---------
torc        KRKYWWKNCKMMIMLGGIGAIIVIVIII----------YFFT------
R27         KRKFWLQNLKMMIIMGVIGLVVVGIIANKLGLIGGEQPPQYQYPPQYMQP
            ..  *  . ** . .. .  ... ..

hum1        ------------------------------------
hum2        ------------------------------------
dro1        --------------WPSSSDSGSGGGNKAITQAPPH
rat1        ------------------------------------
rat2        ------------------------------------
snc1        ----------------------------------SR
torc        ------------------------------------
R27         PPPPPQQPAGGQSSLVDGAGAGDGAGGSAGAGDHGG
```

Figure 5.18 ClustalV Multiple Sequence Alignment
Stars indicate identity to the corresponding residue of the pST123r27 (**R27**) predicted polypeptide. Periods indicate that some of the residues at that position are identical. The sequences used are human synaptobrevins 1 and 2,(**Hum1** and **2** from Archer *et al.* 1990), **rat 1** and **2** (from Elferink *et al.* 1989) pacific ray (**torc** from Trimble *et al.* 1988) Sudhoffs synaptobrevin (**dro1** from Sudhoff *et al.* 1989), and a synaptobrevin molecule from yeast (**snc1** from Dascher *et al.* 1991).

The predicted transmembrane region is shown underlined and the potential tetanus toxin/botulinum B cleavage region is shown in bold. Cleavage in rat takes place at the Q (Gln)76-F(Phe)77 bond of rat2, but not at the V(Val)78-F(phe)79 bond of **rat1**. (Schiavo *et al.* 1992) The peptide in bold corresponds to the homologous region to QFET that they used to block the action of the toxin by Sciavo *et al.* 1992

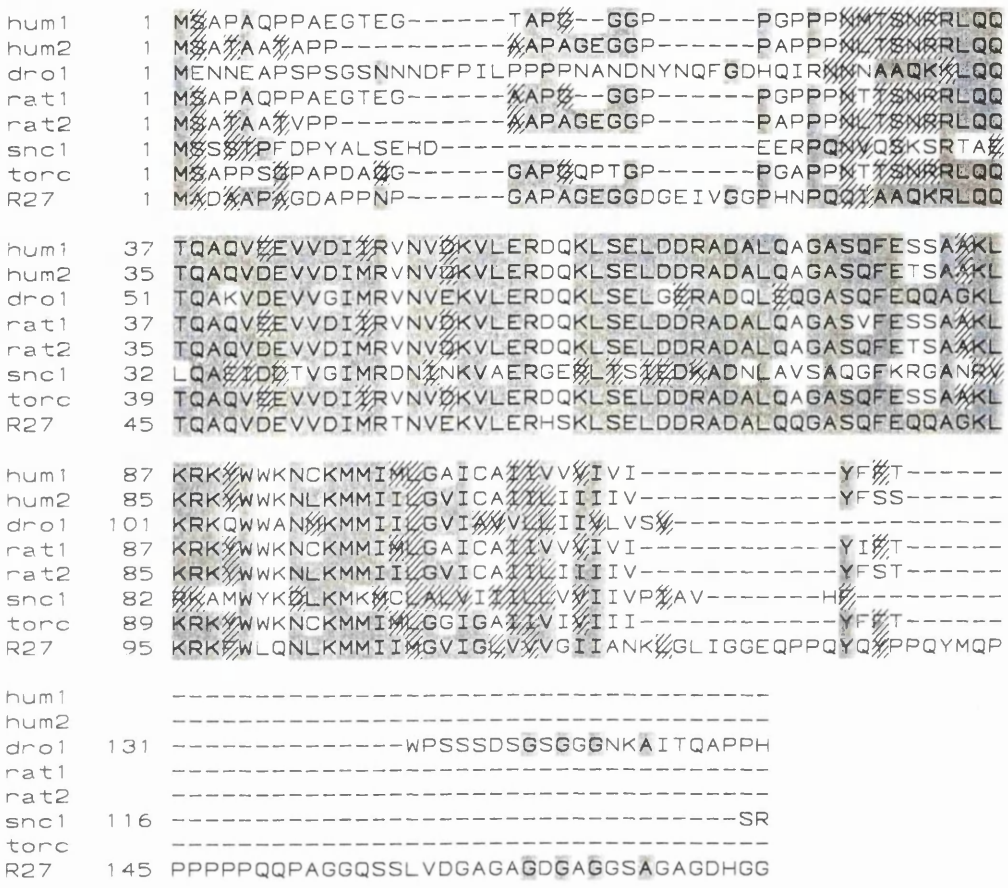


Figure 5.19 Boxed Multiple Alignment of Representative Synaptobrevin Proteins

This alignment is a shaded version of the ClustalV multiple alignment shown as figure 5.18, prepared using Boxshade (methods section 2.22.13). Those aligned positions with residues identical to the pST123r27 derived sequence are shown blacked and those residues that are similar to that of the R27 sequence are shown hatched.

The sequences used are human synaptobrevins 1 and 2, (**Hum1** and **2** from Archer *et al.* 1990), **rat 1** and **2** (from Elferink *et al.* 1989) pacific ray (**torc** from Trimble *et al.* 1988) Sudhoffs synaptobrevin (**dro1** from Sudhoff *et al.* 1989), a synaptobrevin sequence from yeast (**snc1** from Dascher *et al.* 1991) and **r27** derived from the pST123r27 sequence.

Apart from the probable location in the membrane, this predicted synaptobrevin amino acid sequence does not display any peptide motifs that might suggest a specific function for the molecule. This is a common feature of members of the synaptobrevin gene family. Pattern searching programs (GCG Motif and MacPattern, methods section 2.22.11) fail to detect any peptide motifs. In this case these programs also fail to identify this molecule as a synaptobrevin family member. The synaptobrevin pattern should now be updated.

The predicted amino acid sequence displays the four domain structure that is characteristic of synaptobrevins. Domain one is the most amino terminal and is bounded by the sequence QQT in the case of pST123r27. It is characterised by being of variable length, poorly conserved and rich in proline (marine ray 29% and bovine 26%) or aspartamine (26%) in the case of the S-SYB synaptobrevin sequence (data from Sudhoff *et al.* 1989). The synaptobrevin described here has 21% proline in this domain and 24% glycine, another similarly charged AA to proline. Domain 2 is from the QQT boundary to the transmembrane region and is highly conserved in all known family members (refer figure 5.19). Domain 3 is the predicted transmembrane region which is characteristically better conserved on the extravesicular side than on the intravesicular side. Domain 4 is the C-terminal end of variable length. In the pST123r27 protein this domain is composed largely of neutral or hydrophobic residues. This domain structure is shown diagrammatically in figure 5.20b.

5.9.4 Predicted Cleavage of the molecule with Tetanus Toxin/ Botulinum B Toxin

The predicted region of tetanus and botulinum B toxin cleavage is shown in figure 4.18. The cleavage site itself is thought to be in rat between Q(Gln)76 and F(phe)77 in the rat type 2 synaptobrevin (Schiavo *et al.* 1992). The toxins will not cut rat type 1 which has a valine instead of a glutamine as the first residue of the cleavage site. Sciavo *et al.* 1992 predict that both the human synaptobrevins will cleave. It seems likely from the data presented here that both the described *Drosophila* proteins will also do so.

5.9.5 Relationship of the Sudhoff *et al.* 1989 cDNA Sequence to the Genomic Sequence of Chin *et al.* 1993

Very recently genomic sequence has been released for the same locus as that of Sudhoff *et al.* 1989. They have made the very interesting observation that

the locus produces two mRNAs capable of encoding two distinct primary polypeptides. One of these cDNAs, designated dro-a by Chin *et al.* 1993 is almost identical to that of Sudhoff *et al.* 1989. The pair of cDNAs vary at their 5' and 3' termini, but also at a few residues internally (as shown figure 5.20). One of these residues alters the DNA sequence at the potential cleavage site. This would render this new sequence unlikely to cleave by tetanus toxin. The differences between these sequences has not been discussed elsewhere.

The second cDNA described by Chin *et al.* 1993 is longer than the first and produced by differential splicing as shown in figure 5.21a. The second predicted polypeptide is also longer than the first and differs in the very C-termini of the molecule. The products generated are shown in figure 5.21b. Only the Sudhoff *et al.* 1989 sequence has been used for the multiple alignments, because the second isoform does not significantly add to observations made in these sections. Similarly, only the protein sequence predicted by Sudhoff *et al.* 1989 has been used for the phylogenetic analysis presented in the next section.

5.9.6 A Phylogeny of the Synaptobrevin Gene Family

The phylogenetic relationships between the mammalian synaptobrevins were largely responsible for the assignment of molecules as being either type 1 or type 2. However reference to the multiple alignment should indicate that this simple classification breaks down as one looks at molecules from species more distantly related in evolutionary time than for the mammalian radiations. A simple binomial classification seems odd when there are such large differences between family members in domains 1 and 4 between more distantly related members. It is possible that domain 4 serves only to anchor the molecule in the plasma membrane- what residues are specified in this domain may be unimportant.

Domain 1 again is very variable. Proximal to the boundary with domain 2 there is evidence of conservation between different family members. For example the pST123r27 derived amino acid sequence possesses the peptide APAGEGG which is only found in the mammalian type 2 subunits. Does this mean that this is a type 2 molecule? Elsewhere the pST123r27 predicted protein looks more like the first *Drosophila* subunit (by default type 1) seen

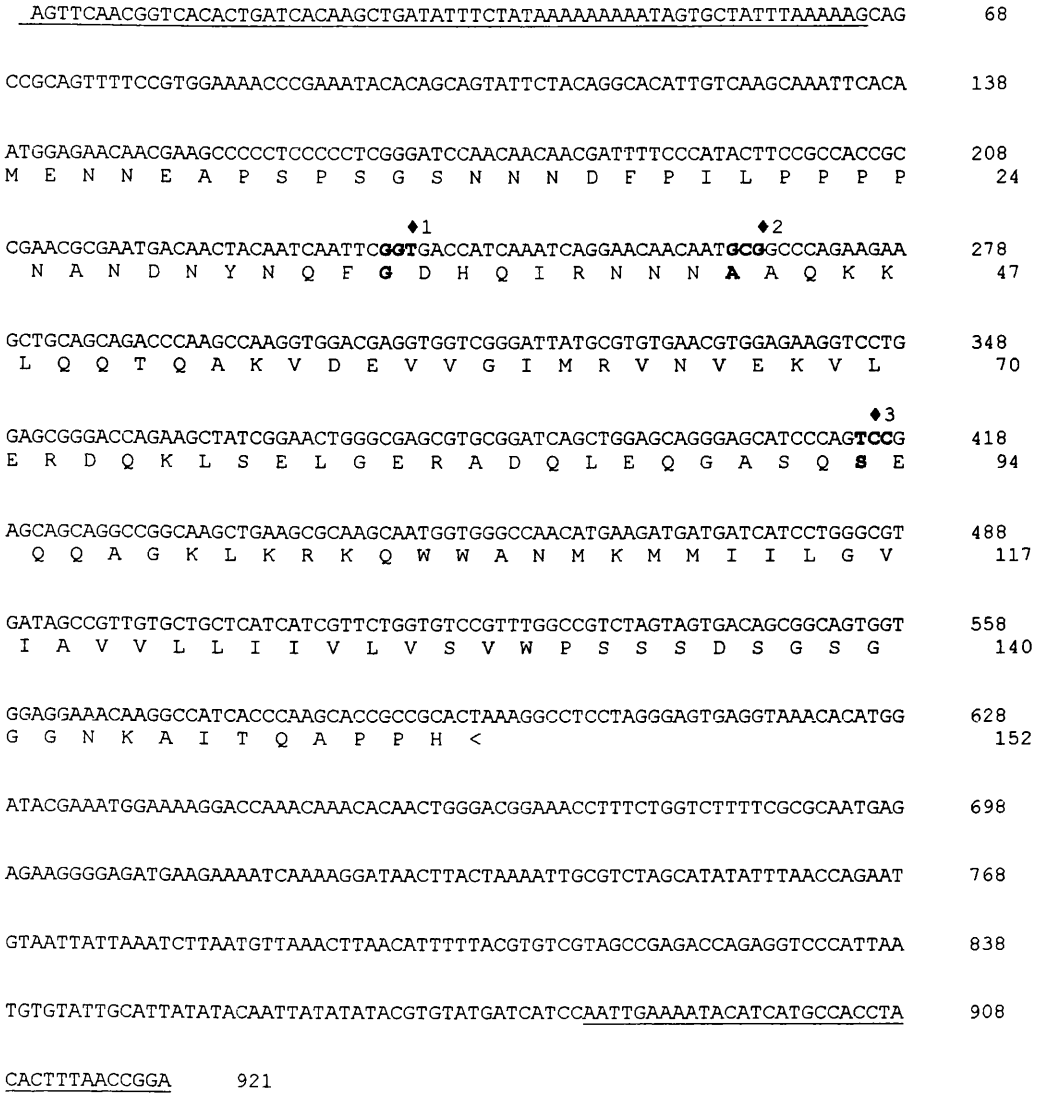
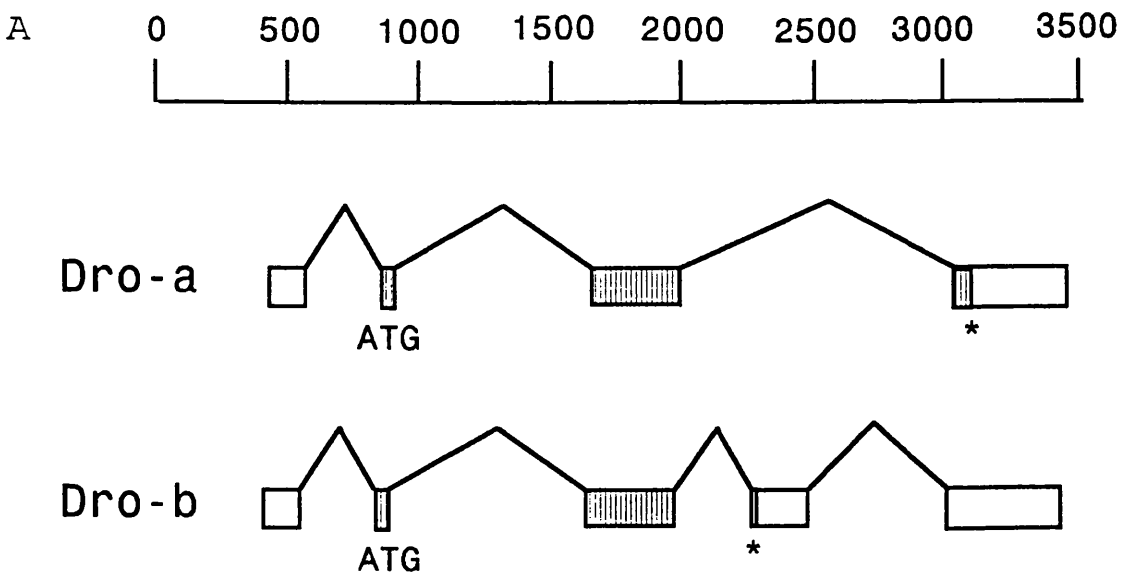


Figure 5.20 The relationship between the Chin *et al.* 1993 predicted mRNA and that the cDNA sequence of Sudhoff *et al.* 1989

The sequence of **syb-a** with the reported ORF shown (from Chin *et al.* 1993). The differences between this sequence and that reported of a cDNA from Sudhoff *et al.* 1989 are shown in bold or underlined. Underlined sequence is not present in the Sudhoff cDNA. The sequence and AA residues in bold are those that are changed in the new sequence compared to that of the cDNA. Changes labelled one and two are in third base positions and do not change the AA sequence. Change 3 however, results in the substitution of TCC S(Serine) for TTC (Phenylalinine).



B

```
Sud             MENNEAPSPSGSNNNDFPILPPPPNANDNYNQFGDHQIRNNNAAQKKLQQ
Syb-a           MENNEAPSPSGSNNNDFPILPPPPNANDNYNQFGDHQIRNNNAAQKKLQQ
Syb-b           MENNEAPSPSGSNNNDFPILPPPPNANDNYNQFGDHQIRNNNAAQKKLQQ

Sud             TQAKVDEVVGIMRVNVEKVLERDQKLSELGERADQLEQGASQFEQQAGKL
Syb-a           TQAKVDEVVGIMRVNVEKVLERDQKLSELGERADQLEQGASQSEQQAGKL
Syb-b           TQAKVDEVVGIMRVNVEKVLERDQKLSELGERADQLEQGASQSEQQAGKL

Sud             KRKQWWANMKMMIILGVIAVVLLIIVLVSVWPSSSDSGSGGGNKAITQAP
Syb-a           KRKQWWANMKMMIILGVIAVVLLIIVLVSVWPSSSDSGSGGGNKAITQAP
Syb-b           KRKQWWANMKMMIILGVIAVVLLIIVLVSLF----------------

Sud             PH*
Syb-a           PH*
Syb-b           -N*
                 .
CLUSTAL V multiple sequence alignment.
```

Figure 5.21 Relationship Between mRNA Species Described by Chin *et al.* 1993 and Sudhoff *et al.* 1989

A shows the exon structure described by Chin *et al.* 1993. Shaded regions in this figure are exons. As shown the cDNAs differ in 3' regions, the extra exon introducing an early in-frame stop. For clarity ATG marks the start codon and * marks the termination codon. Section **B** shows a clustalV alignment (methods section 2.22.12) of the Sudhoff (**Sud**) and Chin proteins(**Syb-a** and **Syb-b**). Underlined residues are those thought to comprise the membrane spanning region. The predicted amino acid sequence from the Sudhoff *et al.* 1989 cDNA clearly corresponds to **Syb-a**. The one difference between the sequences is shown in bold.

most clearly in the region of domain 2 from residue 70 to 95 of the polypeptide. There is a lack of experimental evidence to distinguish functionally between these domains.

A phylogenetic tree can be produced from the multiple alignments already presented. This was produced using the neighbour joining method of Saitou and Nei (1987) as implemented in the ClustalV sequence multiple alignment package (methods section 2.22.12). The phylogeny is shown as figure 5.22a. This tree is unrooted, and distances shown are percentage divergence. These are for guidance only, they are not corrected for multiple substitutions. This correction would tend to stretch the longest branches still further. From the phylogeny it should be clear that both *Drosophila* subunits look more like type 2 than 1, with the pST123r27 derived sequence falling more closely related than that of Sudhoff *et al.* 1989 to either of the mammalian subunits. The yeast subunit is evenly and distantly related to both *Drosophila* subunits.

If one roots the tree at the most distantly related subunit then the same data can be re-drawn as shown in figure 22b. This seems a reasonable rooting of the tree, the yeast subunit branchpoint from the others predates the formation of a nervous system. It is possible to test how well the data supports this tree over any other using bootstrapping (data not shown). All the internal branch points are well supported by the data, with the exception of the relative grouping of the two *Drosophila* subunits to the yeast. In other words, it cannot be reasonably shown as to which *Drosophila* subunit is the most closely related to the yeast sequence. The significance of the phylogeny is dependent upon the multiple alignment. Several alignment parameters have been tried, including the use of related regions of the different proteins (data not shown). The phylogenies that result are not significantly different to those shown.

It is informative to view the phylogenetic tree in the context of gene duplication. The groupings of subunits as type one or two is suggestive of a phylogenetic relationship such that a single gene duplication event took place at a time that predates the mammalian radiations. The duplicate genes would then follow separate evolutionary histories, being selected for two related but distinct roles. The result would be two related molecules in each species that were more closely related to molecules of the same type from other species (for example between rat1 and hum1, please refer figure 5.22) than to different

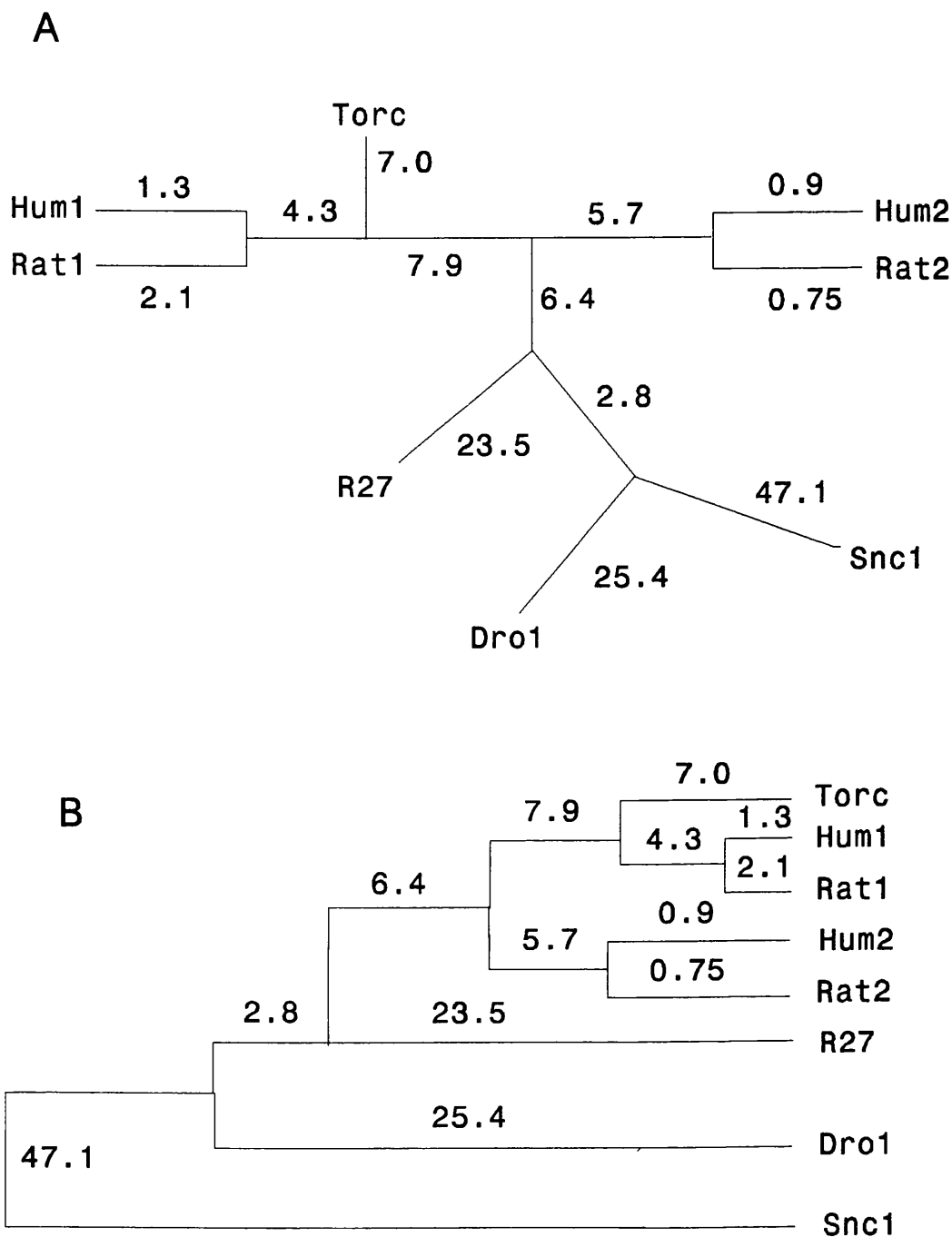


Figure 5.22a,b A Phylogenetic Tree Constructed Using the Neighbour Joining Method of Saitou and Nei 1987

The phylogenies shown are constructed from the alignment shown in figure 5.18 using ClustalV (methods section 2.22.13). The labels used for each species are as figure 5.18. The numeric values shown on the tree branches are a measure of distance as percentage divergence. This value is not corrected for multiple substitutions and a weighting score was used for regions in the alignment for which one or more of the sequences has a gap at that position. Tree **A** is unrooted, while tree **B** is the same data rooted from the longest branch, that from the yeast homologue labelled **snc1**.

types within the same species (e.g. hum1 and hum2). This is the relationship that is seen between the mammalian subunits. The studies on yeast synaptobrevins, (which also has two loci although the sequence is only available for one) suggests that the duplication took place prior to the formation of the nervous system.

But the relationship of the two *Drosophila* sequences suggests that they are more closely related to each other than to any of the mammalian subunits. Hence they appear to have diverged from each other later than from the mammalian/*torpedo* subunits. But the gene duplication event took place before the common ancestor of yeast. How can the *Drosophila* subunits remain more closely related? It is possible that common "*Drosophila* specific" selection pressures have maintained this relationship and we are looking at the effect of convergence, or that the genes have duplicated on more than one occasion. It will be interesting to see where on the tree the second yeast sequence falls. Apart from evolutionary considerations the relationships between these subunits will have considerable practical implications when one prepares to draw cross-subunit comparisons to intimate exact functional relationships.

5.9.7 Studies using a pST123r27 or a pST123 Derived Probe are Comparable

The probe used for the genomic Southern and developmental northern blot described in sections 5.4 and 5.5 was common to both cDNA groups and so predictions for pST123 hold for pST123r27.

5.10 Conclusions and Future work

The aims of this chapter were two-fold. Firstly to confirm the predictions made in chapter 3.0 about the expression of the cDNA, and secondly, to begin a study of the molecular biology of the locus.

The principle predictions from chapter 3.0 were that the gene was transcribed abundantly in the head and less so in the body. It was also predicted that reduced levels would be detected in eggs and larvae. These expression details have been confirmed both with the use of developmental northerns and *in situ* hybridisation to tissue sections. The first sequence data suggested that the cDNA encoded a novel synaptobrevin, probably from a second locus. This

prediction has been born out by more extensive sequencing, genomic Southern and polytene *in situ* analysis.

The molecular biological properties of the synaptobrevin gene family were described in the introduction. The properties of this new locus follow very closely that ascribed to other family members. The locus is abundantly expressed in nervous tissue in both the embryonic and adult stages are consistent with the locus as having a role in the nervous system. Expression in the embryonic stages shows a pattern consistant with expression in the PNS as well as the CNS.

The strong sequence homology between this locus and other members of the synaptobrevin gene family is overwhelming evidence that these cDNAs encode a novel *Drosophila* synaptobrevin homologue transcribed from a novel locus.

5.10.1 Future work

This analysis has highlighted a number of interesting properties of this new synaptobrevin family member that would be of great interest to study further. It would be very informative to do a more complete study of the expression of this locus. A comparative study between the products of the two loci and the individual cDNAs would be of particular interest. Further developmental work on S-SYB has yet to be published following the isolation paper in 1989, hence it may be necessary to perform such an analysis on both loci. The presence of multiple transcripts that are developmentally regulated suggests that this new subunit may have a more complex role in the nervous system than first suggested.

Comparative studies between the loci should employ antibodies specific to the two polypeptides. This will allow the testing of the hypothesis that the synaptobrevin proteins are expressed at the synapse. The isolation and characterisation of a full length amino acid coding sequence presented in this chapter will greatly facilitate this approach. This work has already been started in the laboratory of H.Bellen with the cDNA pST123r27 that I have supplied.

The possible cleavage with tetanus toxin is both biologically and technically useful. Since tetanus toxin is a peptide that works in an analogous way to

ricin it is possible to use tetanus to design a "nervous system-specific" ablation system in *Drosophila*. pST123r27 has been supplied to K. O'Kanes laboratory (Cambridge Genetics Department) for this purpose.

In the Kaiser laboratory there is considerable interest in using site-selected mutagenesis to insert P-elements into the locality of genes of interest. I have designed and tested a single primer common to both loci. This primer has been used in several rounds of mutagenesis. As yet no insertions have been detected using this primer or independently using a primer specific to S-SYB. This last experiment was performed by A.Duncanson (Glasgow Genetics Department). The work is ongoing in Glasgow.

In this chapter I have presented the evidence that pST123 and related cDNAs are derived from a gene encoding a novel form a synaptobrevin, a protein with a defined role in the nervous system. In the next chapter I will describe the properties of two further putative "head-specific" clones to which no defined function can yet be assigned.

5.11 Note Added in Proof

After the writing of this chapter DiAntonio *et al.* 1993 published a paper describing the identical locus. They call this synaptobrevin locus *n-syb* because of the predominant expression in the nervous system.

Key features of this locus include:-

- Head specific expression.
- A 181 amino acid coding region with identical 5' and 3' junctions and identical coding sequence except for 3 variant amino acids.
- Closely matching embryonic and adult *in situ* expression patterns.

The DiAntonio *et al.* 1993 screen was performed using a rat synaptobrevin 1 probe at low stringency to a *Drosophila* head cDNA library. Subsequent analysis of clones identified using the first clone as a probe identified the published coding region.

DiAntonio *et al.* 1993 show a northern blot that is very similar to that discussed in this chapter. They estimate that the large band in a doublet of 3.0 and 3.2kb. This doublet only resolves as a single band on the gel system that I have used to produce the northern blot described in this chapter. My own size estimate was based upon the size of the ribosomal bands and so the difference in predicted size is not unreasonable. Interestingly DiAntonio *et al.* 1993 also produce a band "near 2kb" which they describe as being an artefact from interference effects by the ribosomal bands contaminating the polyA preparation. They do not describe the larger bands I have observed if the blot on long exposures of the autoradiograph. This transcripts are several kb larger than the major band.

If the DiAntonio *et al.* 1993 assertions about the small ~2kb band being an artefact then the short cDNA pST123r27 sequenced in this chapter was produced by internal priming. The full length sequence could then be produced by adding using the identical overlapping region of pST123 to give a full length sequence of 2.95kb, close to that expected. DiAntonio *et al.* 1993 have sequenced about 1kb of the 2kb untranslated region. Since their library was random primed they may not have all the 3' untranslated region. By screening for the amino acid coding region they select for clones that have 5' regions of the message, but not necessarily 3' regions because the cDNA can sit down on the message anywhere 3' of the selected sequence. Hence pST123 contains 1kb of additional sequence to that found by DiAntonio *et al.* 1993.

The data that has led to the deduction that the ~2kb band was artefactual is not shown in the published work. But it is interesting that three independent RNA preparations (total adult and adult head from DiAntonio *et al.* 1993, and my own polyA preparation) all show this effect and yet data using similar preparations (also discussed in the DiAntonio *et al.* 1993 paper and my own data) do not show this effect. It is difficult to assess the extent of mapping and analysis of these cDNAs from the published data.

The *in situ* work using cDNA probes gives a near identical expression pattern. Most importantly DiAntonio *et al.* 1993 confirms the expression in the CNS and PNS and the antennomaxillary complex. They also describe staining in all the cell body layers of the adult head. Immunohistochemical studies of

DiAntonio *et al.* 1993 additionally reveal that expression of the synaptobrevin protein is synaptic.

The published predicted amino acid sequence differs in 3 AA but otherwise is identical. The differences may be strain specific or could be due to sequencing errors in either piece of work. DiAntonio *et al.* 1993 describe several different 5' regions found on cDNA clones. The upstream region of the published sequence is identical to my own in the region of overlap, although the pST123r27 sequence extends a further 350bp 5'. Predictions of the transmembrane region are identical and multiple alignments of the sequences are in accordance in both studies.

The work of DiAntonio *et al.* 1993 provides independent confirmation of the results presented in this chapter. In addition it provides further support for the role of the molecules discussed in this chapter in synaptic transmission. The most critical difference between the works is interpretational; whether the ~2kb band seen on all the head/body northerns and at other stages is real or an artefact. This issue will only be resolved once the 3' end of the messages have been further mapped.

Chapter 6.0 The "Unknown" cDNAs pST170 and pST162

6.1 Introduction

The stated aim of this project was to identify novel components of the nervous system. So far in the thesis I've described the isolation of candidate clones and the subsequent identification and/or characterisation of three of these clones. All these clones were identified by sequence homology and then further evidence sort to support and extend the sequence data. I have described a previously cloned opsin that matches the selected pattern of expression, a novel synaptobrevin gene and a cDNA derived from the first Na+/K+ ATPase β-subunit to be isolated from an insect. But what of the other clones?

Two clones, pST162 and pST170, have also examined. Both these clones are yet to be identified as having homologies to known *Drosophila* genes or to genes from other organism. But as I will describe in this chapter how these cloned genes fit the selected expression pattern and are of continuing interest in the laboratory.

6.2 pST170 Introductory Information.

pST170 is one of a group of three independently isolated cDNAs, pST133, pST170 and pST59 that are from related mRNAs. Because 3/18 of the isolated cDNAs were of this clone it seems likely that the message template of these cDNAs was extremely abundant in *Drosophila* head polyA$^+$ mRNA..

pST170 has been restriction mapped and the map is shown in figure 6.1. The predicted relationship to the other two cDNAs also shown. The pST170 insert is 600bp long making it the shortest of the cDNAs in it's class. pST170 was chosen as representative of the class because full length sequence data in one strand was available for the clone after the initial T3 and T7 runs described in chapter 3.

Since members of the pST170 group are so abundant it was expected that the parallel screen of C. Millagan (Glasgow Genetics Department) should also detect this cDNA class. Indeed, C. Millagan has two cDNAs that show homology to pST170 designated pC12 and pC13. However, these two clones will not be considered in this chapter.

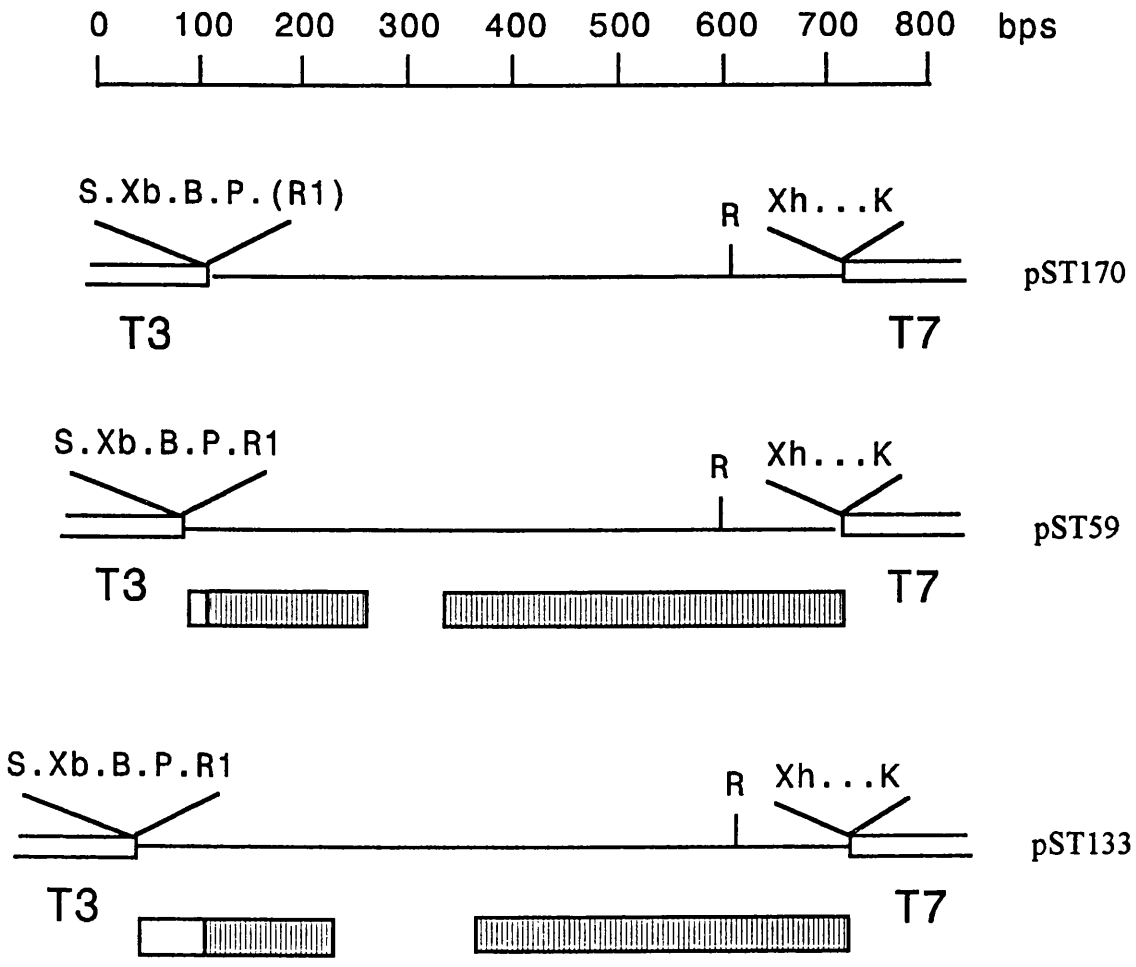


Figure 6.1 Restriction Map of the Three Related cDNA Clones pST170, pST133 and pST59.

Restriction sites shown are **S**-*Sac*I, **Xb**-*Xba*I, **B**-*Bam*HI, **P**-*Pst*I, **R**-*Eco*RI, **Xh**-*Xho*I and **K**-*Kpn*1. Plasmids were also digested with *Eco*RV and *Hind*III which do not cut within the flanking vector seqence. The *Eco*RI site shown bracketed should be present in the plasmid but was absent from this clone (by restriction and sequence data). All three maps are orientated around the common internal *Eco*RI site. The blocks show the region of the pST59 and pST133 that have been sequenced in one strand. Hatched areas represent the regions that are identical to pST170. (pST170 has been fully sequenced in both strands). At the 5' junction point, pST133 and pST59 are 80 and 15bp longer than pST170. 3' junction points only differ in the length of cloned polyA tail. The related clones pC12 and pC13 are not shown.

6.3 A Northern Blot Using the pST170 Insert as a Probe

The insert from pST170 was excised using *Bam*HI and *Xho*I and used to probe a northern blot made with both head and body polyA$^+$ mRNA. (This was prepared as described in methods section 2.9). The resulting autoradiograph is reproduced as figure 6.2. As can clearly be seem from this figure the probe hybridises to a single band of ~0.9kb. A head/body northern blot performed independently by C. Millagan (Glasgow Genetics Department) using a pC13 probe obtained an identical blot. An approximate abundance measurement for the the blots based upon rp49 hybridisation intensity indicates that this clone may account for 0.5-1% of total poly A$^+$ message.

6.4 Full Length Sequence of pST170

T3 and T7 sequencing runs discussed in chapter 3 together provide sequence of the whole cDNA insert in one strand. To finish the sequence in both strands, a scaled down exoIII deletion series was used with one timepoint in each strand (methods section 2.12). This method was chosen as the reactions were run alongside deletion series already discussed (data not shown). To complete the sequence a single oligonucleotide was constructed, and used to prime a sequencing reaction. The final sequencing map is shown as figure 6.3.

6.5 Sequence Features of pST170

No single reading frame runs through the cDNA. Conventional coding region reporting programs e.g. GCG testcode (methods section 2.22.11) need to use a particular window size to attain reasonable statistical significance. Unfortunately, the window sizes used are too large (150-200bp) to be of values in identifying actual reading frames in this case.

Initial database searches with this sequence were using DNA sequence through GenEMBL using FastA and predicted open reading frames through PIR and Swissprot (methods section 2.22.4 and 2.22.5). These protein frames were the three forward frames, with any stops removed. None of these searches identified significant homologies (data not shown).

The composite reading frames used in the protein search are not very sensitive because of the possibility that a linked "nonsense peptide" can cloud faint matches. To overcome this potential problem the full *Drosophila* translation start consensus sequence of Calvener 1987 was used to identify likely frames

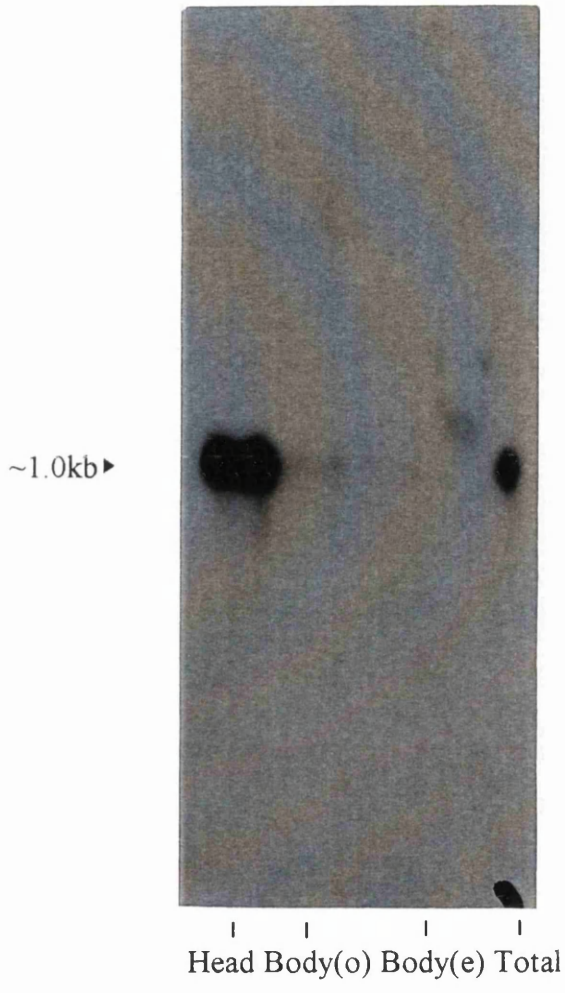
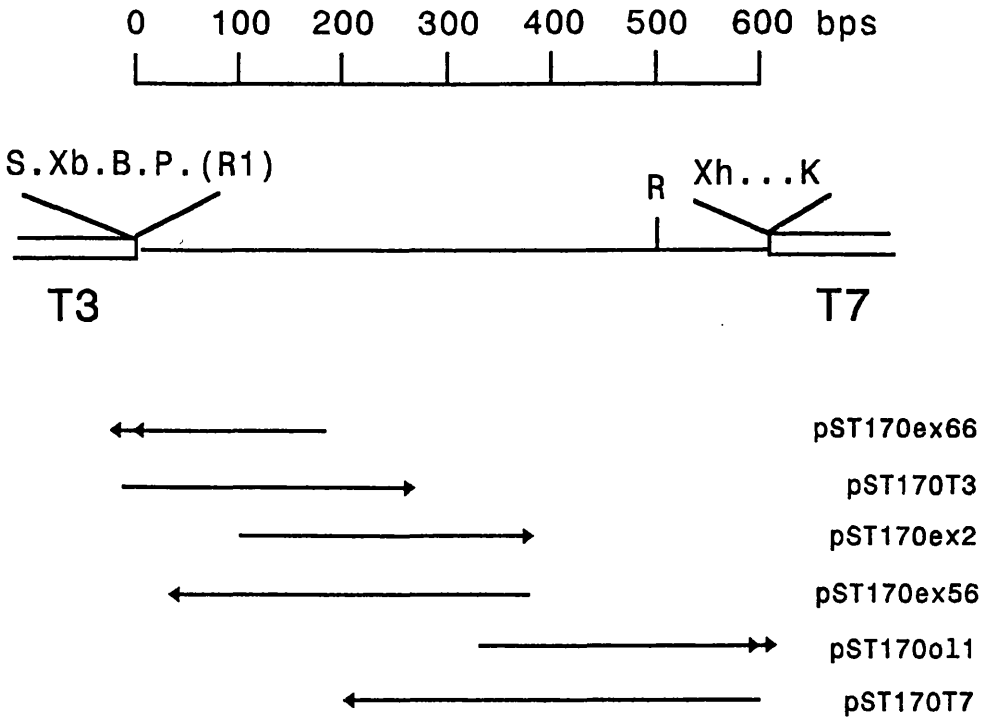


Figure 6.2 A Northern Blot using a pST170 Derived Probe
Northern bloting was performed exactly as described in methods section 2.9. RNA samples loaded on the gel were *Drosophila* wildtype head polyA+ mRNA (**Head**) and body polyA+(**Body(o)**), eya body polyA+ mRNA (**Body(e)**) and total RNA (**Total**). Size is shown in kilobases and is estimated from the migration of the 18s rRNA and the size of a known marker (RP49). The head track (**H**) and body track (**Bo**) are approximately evenly loaded as judged by RP49 hybridisation to a duplicate blot. An identical blot, with a cross-hybridising probe, gives a near identical result (C. Milligan, Glasgow Genetics Department pers. comm.).

A



B

CCCCCGGGCTGCAG**GAATCATC**AAAATCATCATCGTGTTATCAGGACGACCGCAAAAAATCTCTGAAA 46

ATCGTACAAAAAAATATTAAAACCTCAAAATGCATTGGTGTGTGTGTCCCGTTCGTGTAATGCCTGCGCG 116

CCACGCCCCCCGCCCCTCCCACTGTCACGCGCTCGCCACGCCCCCCAGTTGCCGCCCCATTAGCAGCGGG 186

CAGATGGACAACGTGTGGGTGGCCCACAAGGACATCTAGTAACACGACGCCCAACAGCAGCCGCAAGGTC 256

TGATACGCCGCCCGCCACGCCCATCGTGTTTGGGCGGCAGAGGAAGCGGTCCGGAAGCGGAAACGGAAAC 326

GGAAACGGGCGAGCAAAATGGTGGCGCGGTATCGCGGGCAAGGCGACGGCGCGTCCACAAAAAATAACCA 396

TAGAGACGTTTAAGGCAAATTCTAAATGAACAAAAGTATAAACCAAAAACAATTGTAACGAGAAAACAAA 466

CGAATTCAAAAAAATGAGCAATATGAGCAACAACTAATAATGGCAAAAAAAGCAAAAACGACAACTGCAA 536

ATTACGACACAACACCTTCGAAAAAGACCTCAGTAATAATTTTTTTTTTTTTTTTTTTTTTCTCGAGGG 563

GGGGCCC

Figure 6.3 Sequencing of pST170 and Completed Sequence

A shows the sequencing extensions used to construct the full length sequence of pST170. Arrows indicate the direction of priming. Reactions labelled T3 and T7 are from vector based T3 and T7 prime sites on the native plasmid, the reaction labelled pST170o11 was primed from a single oligonucleotide. All other reactions are from exoIII deletion plasmids. Double arrows indicate extensions used to sequence cloning junction points. The restriction map is shown for comparison and is labelled as figure 6.1.

B shows the completed full length sequence, vector sequence is underlined, potential restriction cloning sites are shown in bold. Numbering is of the insert only.

and these frames used in extensive searches of the GenEMBL, PIR and SwissPort databases. Again, no significant matches were found. Further searches of DNA and protein databases were implemented at Daresbury (Daresbury Laboratories, Warrington, England) using the BLAST algorithm through locally held databases also do not identify likely homologues in other organisms (data not shown).

6.6 Tissue *In Situ*s Using a pST170 Derived Probe

No *in situ* data is available as an indication of the adult expression pattern, but C.Millagan (Glasgow Genetics Department) has performed an embryonic *in situ* and shown that the cDNA seems to be weakly detected diffusely over the whole embryo, and specifically in segmentally arranged cells in the nervous system in later stage embryos. Expression becomes more localised to the CNS in later embryos (data not shown).

6.7 pST162 General Features

pST162 has been restriction mapped. The clone is 2.1kb long and characterised by a complete absence of restriction cleavage sites for the six-cutter enzymes with cleavage sites in the pBluescript II polylinker (refer figure 6.5).

6.8 pST162 Head/Body Northern

The insert was excised from pST162 and radiolabelled using random priming (methods section 2.10.1). This was used to probe a duplicate of the head/body northern already described. The resultant autoradiograph is shown as figure 6.4.

From figure 6.4 it was clear that hybridising transcripts to the pST162 probe are found abundantly in the head. They cannot be detected in the body, even after a two week exposure of the blot (30μCi incorporated activity into the probe). This makes this the most head specific locus discussed in this thesis. The size if the single transcript suggests very strongly that the cDNA clone is not a full length representation of the message. From the northern blot estimates of the band size is around 6kb, but the cDNA pST162 insert is just 2.1kb. pST162 is considerably short of being a full length copy of the message.

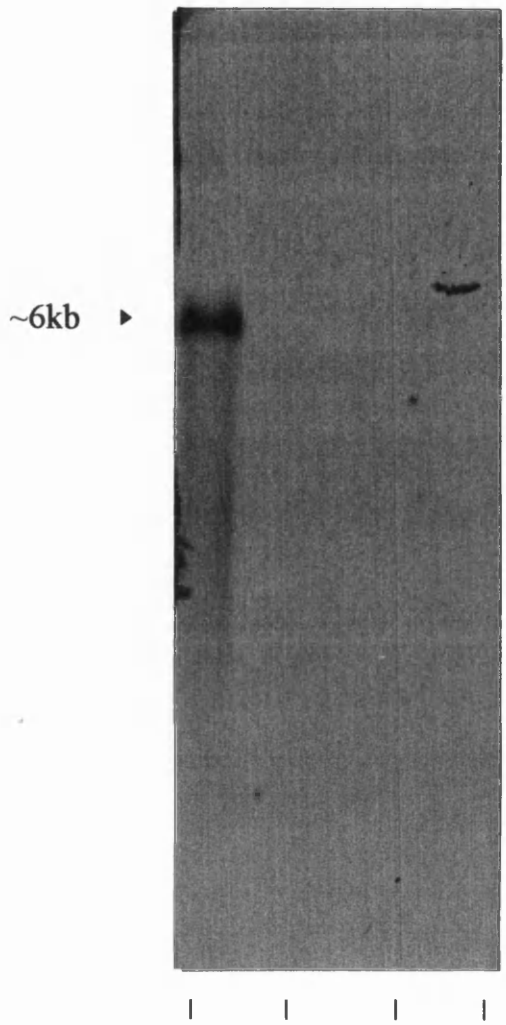


Figure 6.4 A Northern Blot using a pST162 Derived Probe
Northern bloting was performed exactly as described in methods section 2.9. RNA samples loaded on the gel were *Drosophila* wildtype head polyA+ mRNA (**Head**) and body polyA+(**Body(o)**), *eya* body polyA+ mRNA (**Body(e)**) and total RNA (**Total**). Size is shown in basepairs and is estimated from the migration of the rRNAs and confirmed by comparison to a duplicate blot probed with pST123 that gives bands of the sizes reported by DiAntonio *et al.* 1993. Loading of the head (**H**) and body (**B**) lanes is approximately even as judged by RP49 probing of a duplicate blot (figure 4.15).

6.9 Partial Sequencing of pST162

To extend the sequence information available, exoIII deletion reactions were performed on the parent plasmid (methods section 2.12). Although pST162 was not a full length representation of the message, the presence of a single band on a head northern ensures that a composite sequence covering the whole message can easily be constructed at a later date. ExoIII deletions were used to produce deletion clones across the whole cDNA in both strands (data not shown). Suitable sequencing templates were identified by restriction of candidate minipreped plasmids and those selected for sequencing used to prepare sequence grade template (methods section 2.2.4.2). Sequencing reactions were performed as described in methods section 2.13. Figure 6.5 shows the current extent of the pST162 sequence.

As shown in figure 6.5 a single contig has yet to be constructed for pST162. However, two contigs which account for ~2kb of sequence (93% of pST162 coding capacity) were used to search both sequence and protein databases as DNA and protein searches essentially as described for pST170. The results of these searches did not identify significant homologies (data not shown).

6.10 Localisation of the Message in Tissue Sections

In situ hybridisation experiments have been performed using a pST162 derived probe to hybridise to both saggital head sections and whole mount embryos (data not shown). The expression in these *in situs* is consistent with a gene that is specifically but uniformly expressed in the nervous system. The expression profile is very similar to that found when using a pST51 derived probe, although the relative staining in the peripheral components of the nervous system in the later stage embryo is relatively less intense.

6.11 Conclusions for pST162 and pST170

pST170 is a small cDNA that is yet to be identified as being a member of a known gene family. The most amazing feature of the parent mRNA is its abundance-it composes between 0.5 and 1% of polyA+ message from adult fly brain. Furthermore, it is strongly head elevated and expressed in the nervous system. Several transcripts have been identified in mammals that are highly nervous system abundant and have small cDNAs. One such class is the BC class of transcripts (e.g. BC1 from rat DeCiara and Brosius 1987). these transcripts are small (152 nucleotides), poorly conserved and thought to play a role in RNA transport within the neuron. However, the lack of

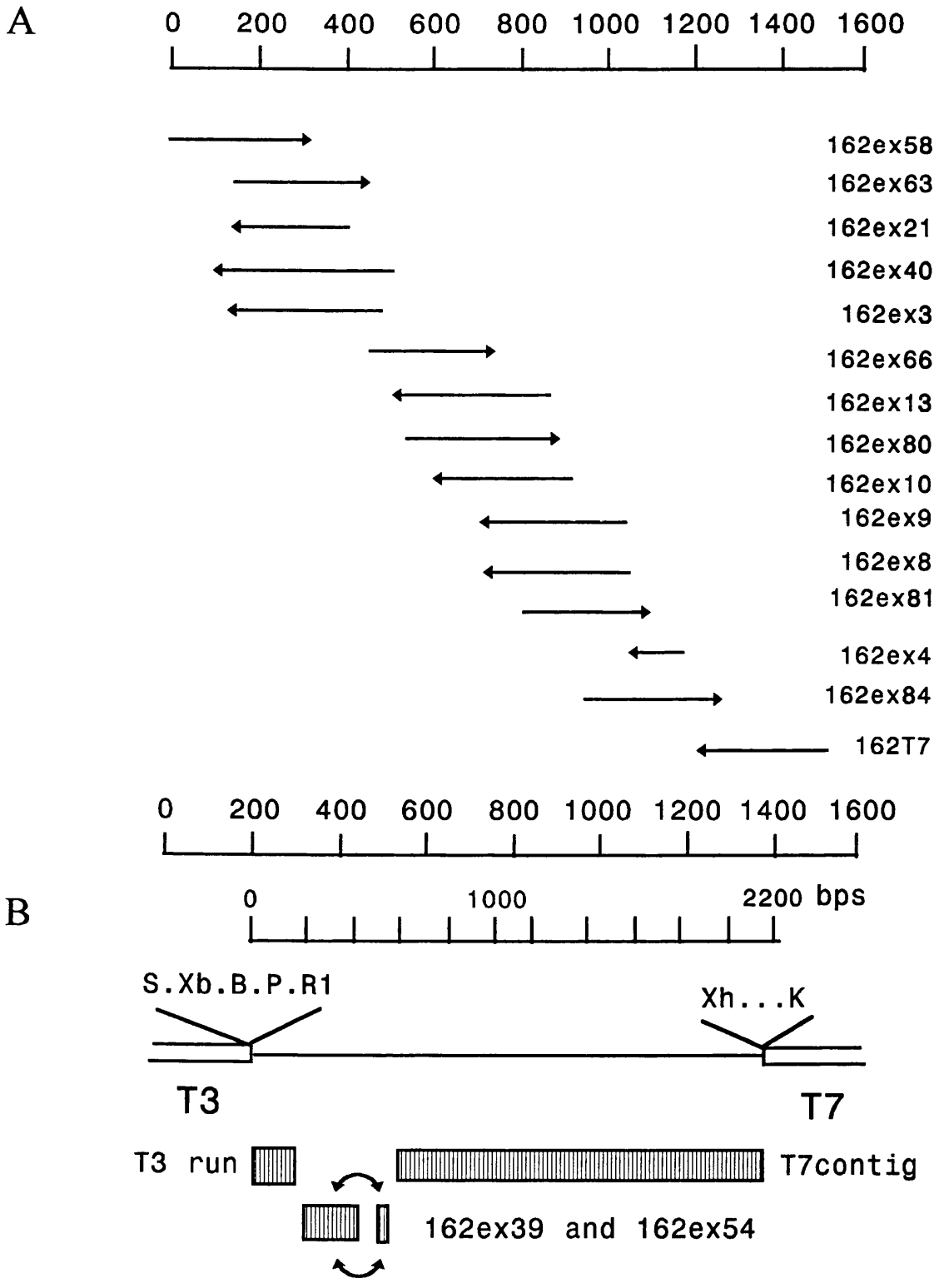


Figure 6.5 The Current Status of the pST162 Sequencing Experiment
A shows the largest contig that spans 1.5kb of the 2.1kb cDNA. This contig crosses over the T7 3' cloning border of the parent cDNA. Arrows indicate direction of priming, scale is in base pairs. **B** shows the current coverage of the whole cDNA. Restriction enzyme cleavage sitea are SacI (S), XbaI(Xb), BamHI (B), PvuII (P), EcoRI (R1), XhoI (Xh) and KpnI (K). Shaded boxes indicate the regions of the plasmid that have been sequenced. The double headed arrows indicate that the relative orientation of the sequencing runs from the exoIII deletion plasmids 162ex39 and 162ex54 has not been determined. In total 1.95kb of the 2.1kb plasmid has been sequenced in at least one strand.

identifiable homology of pST170 to BC1 and the presence of ORFs within pST170 rather argue against an analogous function. Another interesting class to which identity would be difficult to establish by database searching alone are the neuropeptides, which are only 9-12AA long. But again pST170 encoded ORFs does not correspond to known features of the so called FMRFamides (Schneider and Taghert 1990, Chin *et al.* 1990). The identity of pST170 is rather difficult to establish by sequence alone because of the small sizes of the cDNAs and the potential amino acid coding regions. A clearer role for this cDNA will hopefully be established once more extensive expression data is available and once the locus has accurately been mapped to the polytene chromosomes.

pST162 RNA is also abundant in head but cannot be clearly detected in body polyA$^+$ mRNA by northern blotting. From the *in situ* data available so far it appears that the single transcript is expressed in regions of the adult brain that contain the neuronal cell bodies. Expression in the embryo seems to be mainly localised to the CNS rather than to the PNS (e.g. by comparison to the expression pattern of pST123). The 2kb of sequence data available does not identify pST162 as being related to other proteins within the database; perhaps the coding region falls within the missing 5' region or simply is not well conserved between mammals and *Drosophila.* With the mammalian sequencing projects currently having sequence data from several thousand cDNAs isolated from neuronal cDNA libraries it seems unlikely that pST162 (or indeed pST170) defined genes have a well conserved mammalian homologue.

Initial characterisation of these two clones should continue by expanding the sequence and expression data presented here. Polytene map localisation, allowing access to the wealth of *Drosophila* genetic data will prove invaluable. Very recently, C. Millagan (Glasgow Genetics Department) has mapped the locus to polytene band 74A. No genetically defined genes maps to this region which have a neuronal phenotype, although work is continuing to more closely define this locus. Predictions, using such programs as GCG Testcode (methods section 2.22.11), suggest that pST162 contains amino acid coding sequence (data not shown). A full length predicted amino acid sequence should be all that is required to assign a putative function to this locus.

Chapter 7 Final Conclusions

In this thesis I have described the isolation of 30 candidate clones, their characterisation and the selection of four for further study. Out of the screening came cDNAs derived from genes encoding a previously isolated opsin, a novel form of synaptobrevin, the first insect Na+/K+ ATPase β-subunit to be cloned and studied and several other interesting, but as yet completely unidentified, molecules.

The screen described in chapter 3 is a rather general approach to the study of the nervous system. This reflects the direction of the laboratory which had only began to move into the study of the nervous system at the start of this project. I believe I have established ample precedent for such a screen in chapter 1 where I described the contribution that the Levy *et al.* 1982 screen has made to the study of the *Drosophila* nervous system.

The period of this PhD project has seen the exciting beginnings of the genome mapping and sequencing projects. The man and *Drosophila* (Adams *et al.* 1991) projects have begun mass sequencing of cDNAs isolated from neuronal cDNA libraries. Some of these libraries have been generated using subtractive methods. Neuronal libraries have been chosen because of the high complexity of the cDNA populations. Once generated these EST tags are used as to identify gene encoding regions of the genome for sequence entry points. The might of these projects is such that spectacular progress in the molecular identification of neuronal genes should be expected in the near future. The aim of my project was not to emulate the genome sequencing projects but to isolate a modest set of neuronal genes for further characterisation.

The first cycle of the screen presented in chapter 3 was very similar to the Levy *et al.* 1982 screen because it used a total head probe rather than a *eya* head probe. This part of the screen identified the *ninaE* cDNA. The initial intention was to rescreen these positives with an *eya* probe (*eya* head material is relatively difficult to obtain). This part of the screen was superseded when an *eya* head library became available. As suggested in chapter 3 many of these clones will also be of interest. Primitive expression profiles for these genes (data not shown) were obtained by hybridising cDNA arrays to developmental cDNA probes (kindly provided by A Griffin Glasgow Genetics

Department). These profiles suggest that the screen identified at least six highly abundant genes that displayed distinct developmental profiles. Several of these were expressed in embryonic stages which is not expected for a simple eye specific locus. Levy and Manning 1982 also describe positives with such expression patterns. It will be interesting to study these clones in more detail and to determine the exact relationship to the Levy *et al.* 1982 positives.

The more extensive characterisation of the four selected clone groups began with the independent confirmation that these clones were indeed from head specific genes. This was by using northern blotting and by the assessing of the extent of sequence similarity to known gene family members. Expression profiles were also supported by head and body *in situs* and in the case of pST123 by the independent work of DiAntonio *et al.* 1993. These expression results mark the point when the screen has been confirmed a technical success and when the main emphasis of the work switched from technical to biological.

Chapter 4 describes the initial characterisation of a cDNA with homology to a family of ATPase subunit genes that are often found abundantly expressed in the nervous system. Analysis include *in situ* hybridisation to a variety of stages of development. Expression followed that expected for a gene involved specifically in the functioning of neuronal cells. The analysis of genomic Southern and polytene *in situ* data indicates that the gene encoding the subunit is single copy and located near the heterochromatic border of 2R. It is not located on the X chromosome near known mutants of the bang sensitive class suggested to encode a subunit of the Na+/K+ ATPase (Jan and Jan 1978, Ganetsky and Wu 1982). Low stringency hybridisation indicates the presence of several related genes in the *Drosophila* genome. In addition, the locus produces a number of transcripts. Isolation of a set of cDNAs that hybridise at high stringency to pST51 suggests that different sized transcripts may be produced using a variety of 5' and 3' termini and that the gene may have the capacity to produce more than one polypeptide.

Recent work has shown that mutants of the α-subunit gene have a bang sensitive phenotype (Schubiger *et al.* 1993) strongly suggests that the β-subunit should also show such a phenotype. This also provides a ready means to isolate P element inserts in this β-subunit. Such lines would be invaluable

to genetically examine the role of the β-subunit in the functioning of the Na+/K+ ATPase.

The low level of DNA conservation suggests that a clone homologous to pST51 could never easily be isolated by low stringency hybridisation or similar method. Indeed if a mammalian cDNA probe was chosen (as for the J Davies, Glasgow Genetics Department, unpublished screen) then there is little chance of designing a suitable nucleic acid probe. Antibody screening of expression libraries is a possible cloning route, direct purification another, but these are not straight forward methods. Hence this clone demonstrates the value of this technique in order to clone differential highly expressed genes that are not highly conserved.

The synaptobrevin gene represented by pST123 was another gene one would expect to find using differential screening. It was rather surprising to identify a gene that had not at that time been previously reported, although the low stringency northerns reported by Sudhoff *et al.* 1989, certainly suggests the presence of further related genes. The discovery that synaptobrevin is the target of tetanus toxin cleavage (Sciavo *et al.* 1992), provides a pharmacological tool to study that role of synaptobrevins in synaptic transmission.

Since synaptobrevins are highly conserved it was not surprising when DiAntonio *et al.* 1993 reported the cloning by homology screening of the same gene as reported in chapter 5. This independent work serves to confirm the observations on the locus discussed in this thesis. Homology screening is a fast and more direct approach to this particular class of molecules because of the high degree of sequence conservation between family members. Even so, the first *Drosophila* synaptobrevin locus cloned was by Sudhoff *et al.* 1989 using homology screening and a bovine derived probe. This yielded a *Drosophila* synaptobrevin called *syb*. This locus was the most homologous to the probe used and was found in the *Drosophila* head library screened. But further published data on this gene (Chin *et al.* 1993) shows that it's expression is not limited to the nervous system but that this locus is abundantly expressed in non-neuronal cells (the gut), an expression pattern that was not predicted by analogy to the bovine sequence. [Indeed, this was so unexpected that the *syb* sequence is described as neuronal in database

records (accession number P18489) and the isolation paper was published in Neuron!] It took a further four years to find the neuronal synaptobrevin.

Although the cloning of a ATPase β-subunit cDNAs and a synaptobrevin gene family members vindicates the approach to the nervous system taken in this thesis it is for the cloning of cDNAs such as pST162 and pST170 that differential screening is of lasting value. pST162 does not represent the sort of molecule that encodes a part of the battery of proteins that are found in every cell that perform basic "house-keeping" roles. pST162 RNA has a specific function in most or all cells of the central nervous system. If pST162 RNA codes for a protein then full length sequence data (perhaps spanning a number of overlapping cDNAs) will enable predictions to be made as to the likely amino acid coding capacity. This can then be used to perform sensitive searches of protein and DNA databases.

The pST170 cDNA insert has been sequenced completely in both strands and yet the sequence does not immediately suggest why the parent mRNA is expressed. What is incredible about the gene is that it is so highly abundant in the adult head. It is probably *the* most highly expressed head specific gene that is not specific to the eye. This suggests it has an important role to play in the specific functioning of the nervous system. Further work will seek to define this role.

Appendix Summary of Plasmid Clones and Oligonucleotides

Listed are the plasmid clones used/constructed during this thesis. Exonuclease III deletion clones (which number several hundred) are not listed. However, these clones are systematically named so that pST123ex12 is an exoIII deletion clone derived from pST123(see below). Clones that are numbered 1-50 were deleted 5' to3' while clones numbered 3'-5' were deleted 3'to5'.

A Plasmid Clones

Plasmid	When made and by whom	Description and Comments
pST123s1	SRT 8/91	pST123 without 0.7kb *Hind*III fragment
pST123s2	SRT 8/91	pST123 3' *Sac*1 frag. 0.3kb
pST123s3	SRT 8/91	pST123 5' *Xho*I frag. 0.5kb
pST123s4	SRT 8/91	pST123 internal *Hind*III frag
pST42	SRT 6/91	HEC, SK-
pST51	SRT 6/91	HEC, SK-, fully sequence
pST59	SRT 5/91	HEC, SK- related to pST133 and pST170
pST72	SRT 5/91	HEC, SK-
pST78	SRT 6/91	HEC, SK-
pST99	SRT 6/91	HEC, SK- Related to pST51
pST116	SRT 05/91	HEC, SK-
pST123	SRT 5/91	HEC, SK- fully sequenced
pST133	SRT 05/91	HEC, SK-
pST134	SRT 5/91	HEC, SK-
p1ST34	SRT 05/91	HEC, SK-
pST135	SRT 05/91	HEC, SK- near identical to pST151
pST151	SRT 05/91	HEC, SK- near identical to pST135
pST157	SRT 05/91	HEC, SK-
pST162	SRT 05/91	HEC, SK- almost fully sequenced
pST163	SRT 05/91	HEC, SK-
pST170	SRT 5/91	HEC, SK- fully sequenced, related to pST133 and pST59
pST141	SRT 05/91	HEC, SK-
λ41	SRT 03/91	HEC. *Eco*RI/*Hind*III frag in NM1149
pSTs1r1	SRT 6/11/91	Related cDNA to pST123 by screening ZAPII library
rp49	SFG 20/11/91	0.6kb H/E fragment SK- subclone of RP49
pAUDs3	AD 28/11/91	1.5kb genomic fragment of *svb*
pST41-1	8/4/92	Cloning of λ41 known as pST41
pC12	CM 1993	HEC, SK- related to pST133, pST170, pST59 and pC13
pC13	CM 1993	HEC, SK- related to pST133, pST170, pST59 and pC13
α1 tubulin	NA	α1 3' subclone in pBR322 from Kalfeyan and Wensink 1982.
pST41-2	SRT 1/5/92	Cloning of λ41 believed identical to pST41-1
pST51r14	SRT	Related to pST51, in SK-, isolated from ZAPII
pST51r2	SRT	Related to pST51, in SK-, isolated from ZAPII
pST51r8	SRT	Related to pST51, in SK-, isolated from ZAPII
pST51r1	SRT	Related to pST51, in SK-, isolated from ZAPII
pST51r4	SRT	Related to pST51, in SK-, isolated from ZAPII

pST51r12	SRT	Related to pST51, in SK-, isolated from ZAPII
pST51r13	SRT	Related to pST51, in SK-, isolated from ZAPII
pST51r5	SRT	Related to pST51, in SK-, isolated from ZAPII
pST51r46a	SRT	Related to pST51, in SK-, isolated from ZAPII
pST51r11	SRT	Related to pST51, in SK-, isolated from ZAPII
pST51r6	SRT	Related to pST51, in SK-, isolated from ZAPII
pST51r7	SRT	Related to pST51, in SK-, isolated from ZAPII
pST51Bam	SRT	Related to pST51, in SK-, isolated from ZAPII
pST123r27	SRT	Related clone to pST123 in SK- Fully sequenced.
pST123r17	SRT	Related to pST51, in SK-, isolated from ZAPII

Where Sk- is pBluescriptII SK-, HEC one of the differential screen positives. Initial are SRT, Simon R Tomlinson, AD Audrey Duncanson and SFG Steven F Goodwin (all Glasgow Genetics department).

B Oligonucleotides

Oligonucleotide	Sequence 5' to 3'.	Comments
123ol1	TCACGCCCATGATGATCATCATC	Designed to prime both synaptobrevins
123ol2	AGCAGCCGCCACAGTAC	Sequencing primer to region of pST123
51ol1	TTGCACCAGAACAGTTG	Sequencing primer
51ol2	GACTTCTCCAGCTTCTCGCCATG	PCR across region variation pST51 and pST51r12
51ol3	AACGGTCGTGATTAATGTTGCGG	PCR across region variation pST51 and pST51r12
51ol4	AACTGGATCCGGTGCGTGCGTCCAAGG	5' pointing primer off pST51 for mutagenisis
51ol5	GTCTACTAGTACATAGTAGCTCC	5' pointing pimer off pST51r7 for mutagenisis
51ol6	CAATGAATTATTTAAATACTAC	Primer specific to pST51r12 insert region
r27ol13	AGCCAAACCACAGAAGC	pST123r27 sequencing primer
r27ol12	GAGCACCACGAGCCAAA	pST123r27 sequencing primer
r27ol14	GGTTACATTATTTCATG	pST123r27 sequencing primer
r27ol15	CAGGTAGTGGTTGGAGC	pST123r27 sequencing primer
r27ol11	CATTGCAATTCTCGAGA	pST123r27 sequencing primer
r27ol8	CAGCAAGCTGTCGGAGC	pST123r27 sequencing primer
r27ol10	ACCTCTCACACCAACTC	pST123r27 sequencing primer
r27ol9	CCTCCTGCTCCGTCTCC	pST123r27 sequencing primer
r27ol7	GACGACTTCTGTGCCGC	pST123r27 sequencing primer
r27ol5	TCTGGCGGCCCGTCGAT	pST123r27 sequencing primer
r27ol6	GCCTGACAGTCTGCTGC	pST123r27 sequencing primer
r27ol4	GCTCTGGCGGAAGAGCG	pST123r27 sequencing primer
r27ol3	GTGCTCCTTCAGCGGGC	pST123r27 sequencing primer
170ol1	CCAACAGCAGCCGCAAG	pST123r27 sequencing primer
r27ol8	CAGCAAGCTGTCGGAGC	pST123r27 sequencing primer
r27ol16	CCCGTCGATTCAAATTC	pST123r27 sequencing primer
r27ol17	GCTCTGGCGGAAGAGCG	pST123r27 sequencing primer

All primers listed were synthesised by SRT.

Bibliography

Adams M.D., Kelley J.M., Gocayne J.D., Dubnick M., Polymeropoulos M.H., Xiao H., Merril C.R., Wu A., Olde B., Moreno F., Kerlavage A.R., McCombie W.R. and Venter J.C., (1991). Complementary DNA sequencing: Expressed sequence tags and human genome project. Science, **252**, 1651-1656.

Archer B.T., Ozcelik T., Jahn R., Franke U. and Sudhof T.C., (1990). Structures and chromosomal localizations of two genes encoding synaptobrevin 1 and 2. J. Biol. Chem., **265**, 17267-17273.

Ashburner M., (1989). *Drosophila.* A laboratory handbook. Cold Spring harbour Laboratory Press. U.S.A.

Aviv H. and Leder P., (1972). Purification of biologically active globin messenger RNA by chromotography on oligothymidylic acid-cellulose. Proc. Natl. Acad. Sci., **69**, 1408

Bahler M. and Greengard P., (1987). Synapsin I bundles F-actin in a phosphorylation-dependent manner. Nature, **326**, 704-707.

Bains A.J. and Bennett V., (1985). Synapsin I is a spectrin-binding protein immunologically related to erythrocyte protein 4.1. Nature, **315**, 410-413.

Ballinger D.G., and Benzer S., (1989). Targeted gene mutation in *Drosophila.* Proc. Natl. Acad. Sci. USA., **86**, 9402-9406.

Baument M., Maycox P.R., Navone F., De Camilli P. and Jahn R., (1989). Synaptobrevin: an intergral membrane protein of 18, 000 daltons present in small synaptic vesicles of rat brain. EMBO. J., **8**, 379-384.

Baxter-Lowe L.A., Yohanan J.M. and Hokin L.E., (1988). In vitro biosynthesis of the β-subunit of the Na+/K+ ATPase in developing brine shrimp: glycosylation and membrane insertion. Biochim. Biophy. Acta, **943**, 343-348.

Beckman C.,(1970). *sk*: stuck. Drosophila Inf. Serv., **46**, 36.

Bekkers J.M. and Stevens C.F., (1990). Presynaptic mechanism for long term potentiation in the hippocampus. Nature, **346**, 724-729.

Benfenati F., Neyroz P., Bahler M., Masotti L. and Greengard P., (1990). Time-resolved florescence study of the neuron-specific phosphoprotein synapsin I. Evidence for phosphorylation dependent conformational changes. J. Biol. Chem., **265**, 12584-12594.

Bennett M.K., Calakos N., Kreiner T. and Scheller R.H., (1992). Synaptic vesicle membrane proteins interact to form a multimeric complex. J. Cell. Biol., **116**, 761-775.

Bennett M.K., Calakos N. and Scheller R.H., (1992). Syntaxin: a synaptic protein implicated in docking of synaptic vesicles at presynaptic active zones. Science, **257**, 255-259.

Bennett M.V.L., Barrio L.C., Bargiello T.A., Spray D.C., Hertzberg E. and Saez J.C., (1991). Gap junctions: new tools, new answers, new questions. Neuron **6**, 305,320.

Berger S.L., and Kimmel A.R., (1987). Guide to molecular cloning techniques. Methods in Enzymology **152.**

Bertorello A.M., Aperia A., Walaas S.I., Nairn A.C. and Greengard P., (1991). Phosphorylation of the catalytic subunit of Na+,K+ -ATPase inhibits the activity of the enzyme. Proc. Natl. Acad. Sci. (USA), **88**, 11359-11362.

Bhattacharyya K.K., Bergstrom E.E., and Hokin L.E., (1990). Molecular cloning of the β-subunit of the Na,K-ATPase in the brine shrimp, *Artemia.* FEBS. Lett., **269**, 233-238.

Bonini N.M., Leiserman W.M. and Benzer S., (1993). The *eyes absent* gene: Genetic control of cell survival and differentiation in the developing *Drosophila* eye. Cell, **72**, 379-395.

Bossy B., Ballivet M. and Spierer P., (1988). Conservation of neural nicotinic acetylcoline receptors from *Drosophila* to vertebrate central nervous system. EMBO J., **7**, 7763-7767.

Britten R.J., Graham D.E. and Neufeld B.R., (1974). Analysis of repeating DNA sequences by reassociation . Meth. Enzymol., **29E**, 363-380.

Broude N.E., Monastyrskaya G.S., Petrukhin K.E., Grishin A.V, Kiyatkin N.I., Melkov A.M., Smirnov Y.V Sverdlov V.E., Malyshev I.V. and Modyanov N.N., (1987). The primary structure of the beta subunit of Na&K ATPase from pig kidney. Biorg. khim. No 1, **13**, p14-19.

Brown N.H., King D.L, Wilcox M., and Kafatos F.C., (1989). Developmentally regulated alternative splicing of *Drosophila* integrin PS2-α transcripts. Cell, **59**, 185-195.

Brown T.A., Horowitz B., Miller R.P., McDonough A.A. and Farley R.A., (1987). Molecular cloning and sequence analysis of the ($Na^+ + K^+$)-ATPase beta subunit from dog. Biochim. Biophys. Acta., **912**, 244-253.

Buchner E., (1991). Genes expressed in the adult brain of *Drosophila* and effects of their mutations on behaviour: A survey of transmitter- and second messenger-related genes. J. Neurogenet., **7**, 153-192.

Buckley K.M., Floor E. and Kelly R.B., (1987).Cloning and sequence analysis of a cDNA encoding p38, a major synaptic vesicle protein. J. Cell. Biol., **105**, 2447-2456.

Bullock W.O , Fernandez J.M. , and Short J.M., (1987). XL1-Blue- A high efficiency plasmid for transforming recA *Escherichia coli* strains with beta-galactosidase selection. Biotechniques , **5**, 376.

Burnet B. and Connolly K.J., (1974) Activity and sexual behaviour of *Drosophila melanogaster. in* The Genetics of Behaviour. Ed. J.H.F. van Abkelen, Amsterdam, N. Holland.

Cain C.C., Trimble W.S., and Lienhard G.E., (1992). Members of the VAMP family of synaptic vesicle proteins are components of glucose transporter-containing vesicles from rat adipocytes. J. Biol. Chem., **267**, 11681-11684.

Caplan M.J., Forbush III,B., Palade G.E. and Jamieson J.D., (1990). Biosynthesis of the Na,K-ATPase in Madan-Darby canine kidney-cells-Activation and cell surface delivery. J. Biol. Chem. **265**, No 6, 3528-3534.

Cassilli J.A., Whitney M., Joazeiro C.A.P., Becker A. and Zucker C.S., (1991). Isolation of *Drosophila* genes encoding G protein-coupled receptor kinases. Proc. Natl. Acad. Sci. USA., **88**, 11067-11070.

Catsch A., (1944). Eine erbliche storung des berwegungsmechanismus bei *Drosophila melanogaster*. Z. Indukt. Abstammungs Vererbungsl., **82**, 64.

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

Chen M.S., Obar R.A., Schroeder C.C., Austin T.W., Poodry C.A., Wadsworth S.C. and Vallee R.B., (1991). Multiple forms of dynamin are encoded by *shibire*, a *Drosophila* gene involved in endocytosis. Nature, **351**, 583-586.

Chilbalin A.V., Lopina O.D., Pettukhov S.P. and Vasilets L.A., (1991). Phosphorylation of Na+, K+ -ATPase by protein-kinase-C and cAMP dependent protein-kinase. Biol. Membr., **8**, 1140-1141.

Chin A.C., Reynolds E.R., and Scheller R.H., (1990). Organisation and expression of the Drosophila FMRFamide related prohormone gene. DNA. Cell. Biol., **9**, 263-271.

Chin A.C., Burgess R.W, Wong B.R., Schwarz T.L. and Scheller R.H., (1993). Differential expression of transcripts from *syb*, a *Drosophila* VAMP gene that is abundant in non-neuronal cells. (GenEMBL accession number L14270.) Gene, **131**, 173-180.

Chirgwin J.M., Przybyla A.E., MacDonald R.J. and Rutter W.J., (1989). Isolation of biologically acive ribonucleic acid from sources enriched in ribonuclease. Biochemistry, **18**, 5294

Chomczynski P. and Sacchi N., (1887) Single step method of RNA isolation by acid guanidine thiocyanate phenol. Anal. Biochem., No 1, **162**, 156-159.

Crawford G., Slemmon J.F., and Salveterra P.M., (1982). Monoclonal antibodies selective for *Drosophila melanogaster* choline acetyltransferase. J. Biol. Chem., **257**, 3853-3856.

Dascher C., Ossig R., Gallwitz D. and Scmitt H.D., (1991). Identification and structure of four yeast genes (SLY) that are able to suppress the functional loss of YPT1, a member of the RAS superfamily. Mol. Cell. Biol. **11**, 872-885.

Dayhoff M.O., (ed.) (1969). Atlas of protein sequence and structure. vol **4** Silver Spring, Maryland.

Dean R.B., (1987). Reminiscences on the sodium pump. TINS., **10**, 451-454.

DeChiara T.M. and Brosius J., (1987). Neural BC1 RNA- cDNA clones reveal non-repetitive sequence content. Proc. Natl. Acad. Sci. USA., **84**, No 9, 2624-2628.

DiAntonio A., Burgess R.W., Chin A.C., Deitcher D.L. Scheller R.H. and Schwarz T.L., (1993). Identification and characterization of *Drosophila* genes for synaptic vesicle proteins. J. Neurosci., **13**, 4924-4935.

Dudai Y., Jan Y.N., Byers D., Quinn W.G. and Benzer S., (1976). *dunce*, a mutant of *Drosophila* deficient in learning. Proc. Acad. Sci. USA.., **73**, 1684-1688.

Elferink L.A., Trimble W.S. and Scheller R.H., (1989). Two vesicle-associated membrane protein genes are differentially expressed in rat central nervous system. J. Biol. Chem., **264**, 11061-11064.

Emanuel J.R., Garetz S., Stone L. and Levenson R., (1987). Differential expression of Na^+,K^+ -ATPase α- and β-subunit mRNA in rat tissues and cell lines. Proc. Natl. Acad. Sci. USA., **84**, 9030-9034.

Fambrough D.M., (1983). Studies on the Na+-K+ ATPase of skeletal-muscle and nerve. Cold Spr. H. Symp. Quant. Biol., **48**, 297-304.

Feinberg A.P., and Vogelstein B., (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem., **132**, 6-13.

Felsenstein J., (1985). Confidence limits on phylogenies: an approach using the bootstrap. Evolution, **39**, 783-791.

Fischbach K.F. and Heisenberg M., (1981). Structural brain mutant of *Drosophila melanogaster* with reduced cell number in the medulla cortex and with normal optomotor yaw response. Proc. Natl. Acad. Sci. USA., **78**, 1105-1109.

Fischer v. Mollard G., Mignery G.A., Baumert M., Perin M.S., Hanson T.J., Burger P.M., Jahn R. and Sudhoff T.C., (1990). rab3 is a small GTP-binding protein exclusively localized to synaptic vesicles. Proc. Natl. Acad. Sci. USA., **87**, 1988-1992.

Foster J.L., Higgins G.C. and Jackson F.R., (1988). Cloning, sequencing, and expression of the *Drosophila* cAMP-dependent protein kinase catalytic subunit gene. J. Biol. Chem., **263**, 1676-1681.

Fristensky B., Lis J., and Wu R., (1982). Portable microcomputer software for nucleotide sequence analysis. Nucl. Acids Res., **10**, No 20, 6451-6463.

Fujita S.C., Zipurski S.L., Benzer S., Ferrus A. and Shotwell S.L., (1982). Monoclonal antibodies against the *Drosophila* nervous system. Proc. Natl. Acad. Sci. USA., **79**, 7929-7933.

Ganetzhy B. and Wu C.-F., (1982). Indirect suppression involving behavioural mutants with altered neuronal excitability in *Drosophila melanogaster*. Genetics, **100**, 597-614.

Geering K., Theulaz I., Verrey F., Hauptle M.T. and Rossier B.C., (1989). A role for the-subunit in the expression of functional Na^+-K^+ -ATPase in *Xenopus* oocytes. Am. J. Physiol., **257**, C851-858.

Gloor S., (1989). Cloning and nucleotide sequence of the mouse

Na,K -ATPase beta subunit. Nucleic Acid Res., **17**, 10117.

Gloor S., Antonicek H., Sweadner K., Pagliusi S., and Frank R., Moos M. and Schachner M., (1990). The adhesion molecule on glia (AMOG) is a homologue of the subunit of the Na,K -ATPase. J. Cell Biol., **110**, 165-174.

Good P.J., Richter K. and Dawid I.B., (1990). A nervous system-specific isotype of the subunit of Na^+,K^+-ATPase expressed during early development of *Xenopus laevis*. Proc. Natl. Acad. Sci. USA., **87**, 9088-9092.

Greengard P., Valtorta F., Czernik A.J., and Benfenati F., (1993). Synaptic vesicle phosphoproteins and regulation of synaptic function. Science, **259**, 780-785.

Heimbeck G., Klagges G.O., Pflugfelder R., Reifegerste A., Hofbauer S., Buchner S. and Buchner E., (1991). A gene selectively expressed in the brain of *Drosophila* encodes a protein with partial homology to vertebrate Synapsin I. Abst. 12th EDRC Mainz.

Henikoff S., (1984). Unidirectional digestion with exonuclease III creates targeted breakpoints. Gene, **28**, no 3, 351-359.

Heppel L.L., (1940). Am. J. Physiol., **128**, 449-454.

Hermans-Borgmeyer I., Zopf D., Ryseck R.P., Hovemann B., Betz H. and Gundelfinger E.D., (1986). Primary structure of a developmentally regulated nicotinic acetylcholine receptor protein from *Drosophila*. EMBO. J., **5**, 1503-1508.

Hermans-Brogmeyer I., Hoffmans S., Sawruk E., Betz H., Schmitt B. and Gundelfinger E.D., (1989). Neuronal acetylcholine receptors in *Drosophila*: Mature and immature transcripts of the *ard* gene in the developing central nervous system. Neuron, **2**, 1147-1156.

Higgins, D.G. and Sharp, P.M., (1988) CLUSTAL: a package for performing multiple sequence alignments on a microcomputer. Gene, **73**, 237-244.

Hillicker A.J. and Holm D.G., (1975). Genetic analysis of the proximal region of chromosome 2 of *Drosophila melanogaster*. I. Detachment products of compound autosomes. Genetics, **81**, 705-721.

Hilliker A.J., (1976). Genetic analysis of the centromeric heterochromatin of chromosome 2 of *Drosophila melanogaster*: Deficiency mapping of EMS-induced lethal complementation groups. Genetics, **83**, 765-782.

Horisberger J.-D., Lemas V., Kraehenbuhl J.-P., and Rossier B.C., (1991). Structure-function relationship of Na,K-ATPase. Annu. Rev. Physiol., **53**, 565-584.

Horisberger J.D., Jaunin P., Reuben M.A., Lasater L.S., Chow D.C. Forte J.G., Sachs G., Rossier B.C. and Geering K., (1991b). The H,K-ATPase beta-subunit can act as a surrogate for the beta-subunit of Na,K-pumps. J. Biol. Chem., **266**, No. 29, 19131-19134.

Horowiz B., Eakle K.A., Scheinerbobis G., Randolph G.R., Chen C.Y., Hitzeman R.A. and Farley R.A., (1990). Synthesis and assembly of functional mammalian Na, K-ATPase in yeast. J. Biol. Chem., **265**, 4189-4192.

Inesi G. and Kirtley M.R., (1992). Structural features of cation transport ATPases. J. Bioeneg. Biomemb., **24**, 271-283.

Itoh J.J., Slemmon J.R., Hawke D.H., Williamson R., Morita E., Itakura K., Roberts E., Shively J.E., Crawford G.D. and Salavaterra P.M., (1986). Cloning *Drosophila* choline acetyltransferase cDNA. Proc. Natl. Acad. Sci. USA., **83**, 4081-4085.

Izquierdo M. and Bishop J.O., (1989). An analysis of cytoplasmic RNA populations in *Drosophila melanogaster*, Oregon R. Biochem. Genet., **17**, 473-479.

Jaissier F., Canessa C.M., Horisberger J.-D. and Rossier B.C., (1992). Primary sequence and functional expression of a novel ouabain resistant Na,K-ATPase. J. Biol. Chem., **267**, 16895-16903.

Jan Y.N. and Jan L.Y., (1978). Genetic dissection of short-term and long-term facilitation at the *Drosophila* neuromuscular junction. Proc. Natl. Acad. Sci. USA., **75**, No 1, 515-519.

Jessell T.M. and Kandel E.R., (1993). Synaptic transmission: A bidirectional and self-modifiable form of cell-cell communication. Cell, **72**, 1-30.

Johnson E.M., Ueda T., Maeno H. and Greengard P., (1972). Adenosine 3',5'-monophosphate-dependent phosphorylation of a specific protein in synaptic membrane fractions from rat cerebrum. J. Biol. Chem., **247**, 5650-5652.

Kaiser K., and Goodwin S.F., (1990). "Site-selected" transposon mutagenesis of *Drosophila*. Proc. Natl. Acad. Sci. USA , **87**, 1686-1690.

Kalfayan L. and Wensink P.C., (1982). Developmental Regulation of Drosophila alpha- tubulin genes. Cell, **29**, 91-98.

Kaplan W.D. and Trout W.E., (1969). The behaviour of four neurological mutants in *Drosophila*. Genetics, **61**, 399-409.

Kimura M. (1983). The Neutral Theory of Molecular Evolution. Cambridge University Press, Cambridge, England.

Kohne D.E., Levison S.A. and Byers M.J., (1977). Room temperature method of increasing the rate of DNA reassociation by many thousand fold: The phenol emulsion reassociation technique. Biochemistry, **16**, 5329-5341.

Konopka R.J. and Benzer S., (1971). Clock mutants of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA.., **68**, 2112-2116.

Kyte J. and Doolittle R.F., (1982). A simple method for displaying the hydropathic character of a protein. J. Mol Biol., **157**, No.1, 105-132.

Lebowitz R.M., Takeyasu K. and Fambrough D.M., (1989). Molecular characterization and expression of the (Na^+ + K^+)-ATPase alpha-subunit in *Drosophila melanogaster*. EMBO. J., **8**, 193-202.

Lee Y.-J., Dobbs M.B., Verardi M.L. and Hyde D.R., (1990). *dgq*: A *Drosophila* gene encoding a visual system-specific G molecule. Neuron, **5**, 889-898.

Levy L.S. and Manning J.E., (1981). Messenger RNA sequence complexity and homology in developmental stages of *Drosophila*. Dev. Biol., **85**, 141-149.

Levy L.S., Ganguly R., Ganguly N. and Manning J.E., (1982). The selection, expression, and organization of a set of head-specific genes in *Drosophila*. Dev. Biol., **94**, 451-464.

Levy L.S. and Manning J.E., (1982). Expression of a set of head-specific genes during *Drosophila* development. Dev. Biol., **94**, 465-476.

Levy W.B. and McCarthy B.J., (1975)., Messenger RNA complexity in *Drosophila melanogaster*. Biochem., **14**, 2440-2446.

Lewin B.L., (1987)., Genes III. John Wiley and Sons.

Ling L. and Cantely L. (1984)., The (Na,K)-ATPase of friend-erythroleukemia cells is phosphorylated near the ATP hydrolysis site by an endogenous membrane-bound kinase. J. Biol. Chem., **259**, 4089-4095.

Lingrel J.B., (1992). Na, K-ATPase: Isoform structure, function, and expression. J. Bioeneg. Biomemb., **24**, 263-270.

Ma J.y., Song Y.h., Sjostand S.E., Rask L. and Mardh S., (1991). cDNA cloning of the beta-subunit of the human gastric H,K-ATPase. Biochem. Biophy. Res. Comm., **180**, 39-45.

Malinow R. and Tsien R.W., (1990). Presynaptic enhancement shown by whole-cell recordings of long-term potentiation in hippocampal slices. Nature, **346**, 177-180.

Marshall P.J. and Hokin L.E., (1979). Biochem. Biophy. Res. Commun., **87**, 476-482.

Matsumoto H., Ozaki M., Isono L., Randall L., Florean M., Larrivee D., Pye Q. and Pak W.L., (1986). *Nina*: a class of mutations affecting photoreceptor-specific gene functions. In Molecular Neurobiology of *Drosophila*., 16 (Abst.), New York, Cold Spring Harbour Laboratory.

McAllister L., Rehm E.J., Goodman C.S. and Zinn K., (1992). Alternative splicing of micro-exons creates multiple forms of the insect cell adhesion molecule fasciclin I. J. Neurosci., **12**, 895-905.

McDonough A.A, Brown T.A., Horowitz B., Chiu R., Sclotterbeck J., Bowen J., and Schmitt C.A., (1988). Thyroid hormone coordinately regulates Na,K-ATPase alpha and beta subunit mRNA levels in kidney. Am. J. Physiol., **254**, C323-C329.

McDonough A.A., Geering K. and Farley R.A., (1990). The sodium pump needs its β-subunit. FASEB J., **4**, 1598-1605.

Mackensie C., Stewart B. and Kasier K., (1989). A device for creating and replicating ordered arrays of plaques or colonies. Technique, **1**, 49-52.

Montell C., Jones K., Hafen E. and Rubin G., (1985). Rescue of the *Drosophila* phototransduction mutation *trp* by germline transformation. Science, **230**, 1040-1043.

Montell C. and Rubin G., (1989). Molecular characterization of the *Drosophila trp* locus: A putative integral membrane protein required for phototransduction. Neuron, **2**, 1313-1323.

Mount S.M., Burks C., Hertz G. Stormo G.D., White O. and Fields C., (1992). Splicing signals in Drosophila: intron size, information content and consensus sequences. Nucleic Acid Res., **20**, 4255-4262.

Nishizuka Y., (1989). The family of protein kinase C for signal transduction. J. Am. Med. Assoc., **262**, 1826-1833.

Noguchi S., Noda M., Takahashi H., Kawakami K., Ohta T., Nagano K., Hirose T., Inayama S., Kawamura M., and Numa S., (1986). Primary structure of the-

subunit of *Torpedo californica* ($Na^+ + K^+$)ATPase deduced from the cDNA sequence. FEBS. Lett., **320**, 315-320.

Noguchi S., Mishina M., Kawamura M. and Numa S., (1987). Expression of functional ($Na^+ + K^+$)-ATPase from clones cDNAs. FEBS. Lett., **255**, 27-32.

O'Connell P. and Rosbach M., (1984). Sequence, structure and codon preference of the *Drosophila* ribosomal protein 49 gene. Nucleic Acid Res. **12**, 5495-5513.

O'Kane C. and Gehring W., (1987). Detection *in situ* of genomic regulatory elements in *Drosophila*. Proc. Natl. Acad. Sci. USA., **84** 9123-9127.

O'Tousa J.E., Baehr N., Martin R.L., Hirsh J., Pak W.L. and Applebury M.L., (1985). The *Drosophila ninaE* gene encodes an opsin. Cell, **40**, 839-850.

Palazzolo M.J., Hyde D.R., VijayRaghaven K., Mecklenburg K., Benzer S., and Meyerowitz E., (1989). Use of a new strategy to isolate and characterize 463 *Drosophila* cDNA clones corresponding to RNAs detected in adult heads but not in early embryos. Neuron, **3**, 527-539.

Pearson W., (1990). *in* Methods in Enzymology **183**, pp 63-98.

Perin M.S., Fried V.A., Mignery G.A., Jahn R. and Sudhoff T.C., (1990). Phospholipid binding of a synaptic vesicle protein homologous to the regulatory region of protein kinase C. Nature **345**, 260-263.

Perin M.S., Brose N., Jahn R. and Sudhoff T.C., (1991). Domain structure of synaptotagmin (p65). J. Biol. Chem., **266**, 623-629.

Perin M.S., Johnston P.A., Ozcelik T., Jahn R. Franke U. and Sudhoff T.C., (1991). Structural and functional conservation of synaptotagmin (p65) in *Drosophila* and humans. J. Biol. Chem., **266**, 615-622.

Peterson G.L., Churchill L., Fisher J.A. and Hokin L.E., (1982). Structure and Biosynthesis of [Na,K]-ATPase in developing brine shrimp nauplii. Ann. Acad. Sci. (New York), **402**, 185-206.

Peterson G.L. and Hokin L.E., (1980). Improved purification of brine shrimp (*Artemia saline*)(Na+ + K+)-activated adenosine triphosphatase and amino-acid and carbohydrate analyses of the isolated subunit. Biochem. J., **192**, 107-118.

Petrucci T.C and Morrow J.S., (1987). Synapsin I: an actin-bundling protein under phosphorylation control. J. Cell Biol., **105**, 1355-1363.

Pustell,J. and Kafatos, F., (1982). A high speed, high capacity similarity matrix: zooming through SV40 and polyoma. Nucl. Acids Res. **10**, 15, 4765-4782.

Ray K., and Ganguly R., (1992). The *Drosophila* G protein subunit gene (D-G1) produces three developmentally regulated transcripts and is predominantly expressed in the central nervous system. J. Biol. Chem., **267**, 6086-6092.

Renaud K.J., Inman E.M. and Fambrough D.M., (1991). Cytoplasmic and transmembrane domain deletions of the Na,K-ATPase β-subunit. J. Biol. Chem., **266**, 20491-20497.

Reuben M.A., Lasater L.S., and Sachs G., (1990). Characterization of a beta-subunit of the gastric H+/K+ transporting ATPase. Proc. Natl. Acad. Sci. USA., **87**, 6767-6771.

Roberts D.B., (1986). *Drosophila*, A practical approach. IRL Press.

Saitou N. and Nei M., (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol., **4**, 406-425.

Salvaterra P.M., Crawford G.D., Klotz J.I. and Ikeda K., (1985). Production and use of monoclonal antibodies to biochemically defined insect neuronal antigens. In Neurochemical Techniques in Insect Research. *ed* Breer H., and Miller T.A., Springer Verlag, Berlin, 223-242.

Sambrook J. , Fritsch E. F., and Maniatis T., (1989). Molecular Cloning: a laboratory manual. Cold Spring Harbour Laboratory Press. U.S.A.

Sanger F., Nicklen S., and Coulson A.R., (1977). DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA., **74**, 5463-5467.

Sawruk F., Hermans-Borgmeyer I., Betz H., and Gundelfinger E.D., (1988). Characterization of an invertebrate nicotinic acetylcholine receptor gene: The *ard* gene of *Drosophila melanogaster*. FEBS. Lett., **235**, 40-46.

Sawruk F., Schloss P., Betz H and Schmitt B., (1990). Heterogeneity of *Drosophila* nicotinic acetylcholine receptors: SAD, a novel developmentally regulated -subunit. EMBO. J., **9**, 2671-2677.

Schiavo G., Benfenati F., Poulain B., Rossetto O., Polverino de Laureto P., Das Guptal B.R. and Montecucco C., (1992). Tetanus toxin and botulinum-B neurotoxins block neurotransmitter release by proteolytic cleavage of synaptobrevin. Nature, **359**, 832-835.

Schloss P., Hermans-Borgmeyer I., Betz H., and Gundelfinger E.D., (1988). Neuronal acetylcholine receptors in *Drosophila*: The ARD protein is a component of a high affinity α-bungarotoxin binding complex. EMBO. J., **7**, 2889-2894.

Schneider L.E., and Taghert P.H., (1990). Organisation and expression of the Drosophila Phe-Met-Arg-Phe-NH2 neuropeptide gene. J. Biol. Chem., **265**, 6890-6895.

Schubiger M., Feng Y., Fambrough D.M. and Palka J., (1993). Characteristics of a hypomorphic mutation of the Na+,K+-ATPase α-subunit gene. Cold Spring Harbor *Drosophila* Neurobiology Meeting Abst.

Schwarz T.L., Papazian D.M., Caretto R.C., Jan Y.N. and Jan L.Y., (1990). Immunological characterization of K^+ channel components from the *Shaker* locus and differential distribution of splicing variants in *Drosophila*. Neuron, **2**, 119-127.

Shpetner H.S. and Vallee R.B., (1992). Dynamin is a GTPase stimulated to high levels of activity by microtubules. Nature, **355**, 733-735.

Short J.M , Fernandez J.M., Sorge J. A and Huse W.D., (1988). Lambda Zap: a bacteriophage lambda expression vector with *in vivo* excision properties. Nucleic Acid Res., **16**, 7583-7600.

Shull G.E, Lane L.K. and Lingrel J.B., (1986). Amino-acid sequence of the beta-subunit of the (Na^++K^+)ATPase deduced from a cDNA.

Shull G.E., (1990). cDNA cloning of the β-subunit of the rat gastric H,K-ATPase. J.Biol. Chem., **265**, 12123-12126.

Sive H.L. and St John T., (1988). A simple subtractive hybridisation technique employing photoactivatable biotin and phenol extraction. Nucleic Acid Res., **16**, 10937.

Skou J.C., (1957). The influence of some cations on an adenosine triphosphatase from periferal nerves. Biochem. Biophy. Acta, **23**, 394-401 (K22).

Slemmon J.R., Salveterra P.M., Crawford G. and Roberts E., (1982). Purification of choline acetyltransferase from *Drosophila melanogaster*. J. Biol. Chem., **257**, 3847-3852.

Smart J.L. and Deth R.C., (1988). Influence of alpha-1-adrenergic receptor stimulation and phorbol esters on hepatic Na+/K+ ATPase activity. Pharmacology **37**, 94-104.

Spector M., O'Neal S. and Racker E., (1988). Regulation of phosphorylation of the -subunit of the Ehrlich Ascites tumor Na+K+ -ATPase by a protein kinase cascade. J. Biol. Chem., **256**, 4219-4227.

States D.J and Botstein D., (1991). Molecular sequence accuracy and the analysis of protein coding regions. Proc. Natl. Acad. Sci. USA., **88**, 5518-5522.

Steward O. and Banker G.A., (1992) Getting the message from the gene to the synapse: sorting and intracellular transport of RNA in neurons. TINS., **15**, 180-186.

Sturzl M. and Roth W.K., (1990). PCR-synthesised single stranded DNA: A useful tool for "hyb" and "HAP" standardization for construction of subtraction libraries. TIG., **6**, 106.

Sudhoff T.C., Baumert M., Perin M.S. and Jahn R., (1989). A synaptic vesicle membrane protein is conserved from mammals to *Drosophila*. Neuron, **2**, 1475-1481.

Sudhoff T.C., Czernik A.J., Kao H.-T., Takei K., Johnston P.A., Horiuchi A., Kanazir S.D., Wagner M.A., Perin M.S., DeCamilli P. and Greengard P., (1989b). Synapsins: mosaics of shared and individual domains in a family of synaptic vesicle phosphoproteins. Science, **245**, 1474-1480.

Sudhoff T.C. and Jahn R., (1991). Proteins of synaptic vesicles involved in exocytosis and membrane recycling. Neuron **6**, 665-677.

Sun D.Y., Guo S.J.Z., Hartmann H.A. Uno H. and Hokin L., (1992). Differential expression of the alpha and beta messenger RNAs of Na,K-ATPase in developing brine shrimp as measured by in situ hybridisation. The Journal of Histochemistry and Cytochemistry **40**, 555-562.

Sved J., (1986). Eyes Absent (*eya*). *Drosophila* Inf. Serv., **63**, 169.

Sweadner K.J., (1985). Enzymic properties of separated isozymes of the Na,K-ATPase. Substrate affinities, kinetic cooperativity and ion transport stoichiometry. J. Biol. Chem., **260**, 11508-11513.

Sweadner K.J. and Gilkeson R.C., (1985). 2 isozymes of the Na,K -ATPase have distinct antigenic determinants. J. Biol. Chem. **260**, 9016-9022.

Sweadner K.J., (1989). Isozymes of the Na+/K+ -ATPase. Biochim. Biophy. Acta, **988**, 185-220.

Takeyasu K., Tamkun M.M., Siegel N.R. and Fambrough D.M., (1987). The expression of hybrid (Na&K)-ATPase molecules after transfection of mouse LTK cells with DNA encoding the beta-subunit of an avian brain sodium pump. J. Biol. Chem., **262**, 10733-10740.

Thomas J.B. and Wyman R.J., (1984). Mutations altering synaptic connectivity between identified neurons in *Drosophila*. J. Neurosci., **4**, 530-538.

Thomas L., Hartung K., Langosch D., Rehm H., Bamberg E., Franke W.W. and Betz H., (1988). Identification of synaptophysin as a hexameric channel protein of the synaptic vesicle membrane. Science, **242**, 1050-1053.

Thomas L. and Betz H., (1990). Synaptophysin binds to physophilin, a putative synaptic plasma membrane protein. J. Cell Biol., **111**, 2041-2052.

Trimble W.S., Gray T.S., Elferink L.A., Wilson M.C. and Scheller R.H., (1990). Distinct patterns of expression of two VAMP genes within the rat brain. J. Neurosci., **10**, 1380-1387.

Ussing H.H. and Zerahn K., (1951). Active transport of sodium as a source of electric current in the short-circuited isolated frog skin. Acta. Physiol. Scand., **23**, 110-127 (K20).

van der Bliek A.M., Meyerowitz E.M., (1991). Dynamin-like protein encoded by the *Drosophila shibire* gene associated with vesicular traffic. Nature, **351**, 411-414.

Váradi A., Gilmore-Heber M. and Benz E.J., (1989). Amplification of the phosphorylation site-ATP-binding site cDNA fragment of the Na+, K+-ATPase and the Ca2+ ATPase of *Drosophila melanogaster* by polymerase chain reaction. FEBS. Lett., **258**, 203-207.

Wadsworth S.C., Rosenthal L.S., Kammermeyer K.L., Potter M.B. and Nelson D.J., (1988). Expression of a *Drosophila melanogaster* acetylcholine receptor-related gene in the central nervous system. Mol. Cell. Biol., **8**, 778-785.

Weischaus E. and Nusslein-Volhard C., (1986). *in* D.M. Roberts. *Drosophila*, A practical approach. IRL Press.

Whittaker P. A., Cambell A.J.B., Southern E.M. and Murray N.E., (1988). Enhanced recovery and restriction mapping of DNA fragments cloned in a new lambda vector. Nucleic Acid Res., **16**, 6725-6736

Yanisch-Perron C., Vieira J. and Messing J., (1985). Improved M13 cloning vectors and host strains: nucleotide sequence of M13mp18 and pUC19 vectors. Gene, **33**, 103-119.

Zamofing D., Rossier B.C. and Geering K., (1988). Role of the Na.K-ATPase beta subunit in the cellular accumulation and maturation of the enzyme as assessed by glycosylation inhibitors. J. Membr. Biol., **104**, 69-79.

Zamofing D., Rossier B.C. and Geering K., (1989). Inhibition of N-glycosylation affects transepithelial Na^+ but not Na^+-K^+-ATPase transport. Am. J. Physiol., **256**, C956-C968.

Zoon R.A (1987) in Berger S.L., and Kimmel A.R., (1987). Guide to molecular cloning techniques. Methods in Enzymology, **152.**

Zucker C.S., Cowman A.F. and Rubin G.M., (1985). Isolation and structure of a rhodopsin gene from *D. melanogaster.* Cell, **40**, 851-858.

Zucker R.S., (1989). Short-term synaptic plasticity. Annu. Rev. Neurosci., **12**, 13-31.