THE YOLK PROTEINS OF *DROSOPHILA* ARE CONSERVED THROUGH DIPTERAN EVOLUTION

ALBERTO MARTINEZ

Thesis presented for the Degree of Doctor of Philosophy

University of Edinburgh

1991

DECLARATION

I declare that this thesis was composed by myself, and that the work described is my own, unless otherwise stated.

ABSTRACT

The yolk proteins (YPs) in *Drosophila* have been shown to enter the oocyte via receptor mediated endocytosis (Mahowald, 1972; Butterworth *et al.*, 1991). The YPs are sorted and specifically stored in the α-granules of the oocyte (Mahowald and Kambysellis, 1980). The aim of this project is to determine the YP domains involved in the specific interactions resulting in uptake into the oocyte. An evolutionary approach was chosen to attempt to determine these domains, since the YPs were assumed to have few evolutionary constraints and thus diverge readily. A functional assay was designed to investigate the uptake of foreign YPs in *D. melanogaster* and used to test YPs from different *Drosophila* and non-*Drosophila* species. The results indicated that the functional domains involved in YP uptake were conserved for up to 100 MYR (Million years) of Dipteran evolution. The molecular cloning and DNA sequence determination of *Calliphora yp* genes demonstrated that the YPs were very well conserved. Not only had the *Calliphora* YPs failed to diverge as expected but these also displayed the similarity to the vertebrate lipases shown in *Drosophila* YPs (Bownes *et al.*, 1988; Perssons *et al.*, 1989). Therefore suggesting that this is a structural constraint within the YPs of Diptera. The conservation of the amino acid sequence of the YPs did not suggest any domains which may have been involved in uptake; however, computer based database searches with five different

regions of the protein suggested putative sites of importance for YP uptake. Their origin and possible functions are discussed.

The *Calliphora yp* genes were also found to be expressed in a coordinate manner by both carcasses and ovaries, between stages 8 and 10b of oogenesis. These *yp* genes were found to be expressed in the follicle cells of the oocyte thus suggesting that the regulation of the *yp* gene expression in *Calliphora* and *Drosophila* is also conserved and thus possibly regulatory sequences and factors will be similar in these two species.

ACKNOWLEGMENTS

I would like to thank my supervisor Dr. Mary Bownes for her help and encouragement during my stay in her laboratory. I also thank past and present members of the Bownes group for their support and patience with the noise: especially Susan, Roger, Brenda, Gregor, Niels, Elaine, Nian, Claire, Alan, Ann, Jackie, Kathleen and Debbie.

I also thank the media and wash-up room workers, specially Ruby and Joan for never making a morning dull. A special thanks to Sandra Grant in Zoology for always having the flies ready and for her enthusiasm and attention to detail. I thank Graeme Brown and Frank Johnstone for their efforts to produce good quality photographs; and Mairead Blair for the beautiful figures.

I thank Susan Liddell from the bottom of my heart for deciphering my English into something resembling the real thing, and Mairead Blair, Roger Slee and Brenda Grimes for final proof reading.

I also thank Roger, Anna, Sybille, Craig, Susan, Greg, Gregor, Steve, Brenda, Chris, Ana, Derek, Dave, Nik, Carol, Mel for helping me see that there is more to the world than meets the eye.

I thank Roger for his madness and the many noises that made life in the lab something to remember. Jocky, Desmond, Melucha and las pulgas will always be remembered.

Finally, I specially thank my parents, more than they will know, without whom this thesis would not have been possible, Anna for her support and cheerfulness, and Mairead for being there when I needed her the most (always).

Para Papa and Mama un abrazo y mis gracias por el soporte que me han dado durante los ultimos nueve anos. A la Malucha por tu paciencia con este hermano.

ABBREVIATIONS

Adh	Alcohol dehydrogenase
amp	Ampicillin
APS	Ammonium persulphate
apprx	Approximately
ATP	Adenosine-5'-triphosphate
bp	Base pair
BCIG	5-bromo-4-chloro-e-indoyl-galactoside
BSA	Bovine serum albumin
°C	Degrees centigrade
cDNA	Complementary deoxyribonucleic acid
Ce	*Calliphora erythocephala*
Ci	Curies
cm	centimetre(s)
cpm	Counts per minute
CTP	Cytidine-5'-triphosphate
(d)dATP	2'(3'-di)deoxyadenosine-5'-triphosphate
(d)dCTP	2'(3'-di)deoxycytidine-5'-triphosphate
(d)dGTP	2'(3'-di)deoxyguanosine-5'-triphosphate
(d)dTTP	2'(3'-di)deoxythymidine-5'-triphosphate
(d)dNTP	2'(3'-di)deoxynuclotide-5'-triphosphate
DEAE	Diethylaminoethyl
Dm	*Drosophila melanogaster*
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dH_2O	Distilled water

dsH_2O	Distilled, sterile water
DTT	Dithiothreitol
EDTA	Diaminoathanetetra-acetic acid
EGTA	N,N,N',N'-(Ethyleneglycol-bis(--aminoethyl Ether) tetraacetic acid
g	Gram(mes)
glycine	Amino ethanoic acid
GTP	Guanosine-5'-triphosphate
HEPES	N-2-hydoxyehtylpiperazine-N'-2-ethanesulphonic acid
IPTG	Isopropyl- -D-thiogalactoside
kb	Kilobase
kD	KiloDalton
Klenow	Large fragment of DNA polymerase I
krpm	kilorevolutions per minute
L	Litre
LSP	Larval Serum Protein
M	Molar
mA	Milliampere(s)
mCi	Millicurie(s)
mg	Milligram(mes)
ml	Millilitre(s)
mm	Millimetre(s)
mM	Millimolar
mmol	Millimole(s)
min	Minutes
MOPS	Morpholinopropanesulphonic acid
mRNA	Messenger ribonucleic acid

ng	Nanogram(mes)
nmol	Nanomole(s)
OD	Optical density
OLB	Oligo labelling buffer
^{32}P	βemitting isotope of phophorous
PCP	Pupal Cuticle Protein
PEG	Polyethylene glycol
%	Percentage
pg	Picogram(mes)
pH	Log_{10} (hydrogen ion concentration)
PMSF	Phenylmethylsulphonyl fluoride
polyA$^+$ RNA	Polyadenylated ribonucleic acid
RNA	Ribonucleic acid
RNase	Ribonuclease
rRNA	Ribosomal ribonucleic acid
rpm	Revolutions per minute
^{35}S	βemitting isotope of sulphur
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
TEMED	NNN'N'-tetra-methyl-1,2-diamino-ethane
Tris	Tris(hydroxymethyl)-amino-methane
Triton-X-100	Octylphenoxypolyethoxyethanol
TCA	Trichloroacetic acid
U	Unit(s)
UTP	Uridine-5'-triphosphate
UV	Ultraviolet
uCi	Microcurie(s)

ug	Microgram(mes)
ul	Microlitre(s)
uM	Micromolar
umol	Micromole(s)
V	Volt(s)
v/v	volume per volume
w/v	weight per volume
yp$_1$	Yolk Protein gene 1
yp$_2$	Yolk Protein gene 2
yp$_3$	Yolk Protein gene 3
YP$_1$	Yolk Protein 1
YP$_2$	Yolk Protein 2
YP$_3$	Yolk Protein 3

CONTENTS

CHAPTER 1

INTRODUCTION

The specific uptake of proteins into cells is a universal problem. The systems used to study the specific uptake of proteins rely upon the understanding of the interactions between the proteins in question and their receptors. Once the receptor is characterised, further studies on the nature of the uptake can be undertaken. Furthermore, in the case of insects, the end result could be the design of a species specific pesticide to block the receptor, where the choice of the protein and receptor complex is critical. For example, by blocking the specific entry of yolk protein into developing oocytes female flies could be rendered sterile.

The aims of the studies discussed here are to determine the domains of the yolk proteins (YPs) of *Drosophila melanogaster* which are required for their specific uptake into the oocyte. The approach chosen to answer this question was an evolutionary one, whereby an assay for the uptake of foreign yolk proteins into *D. melanogaster* ovaries was developed. The results from this assay would enable one to choose a good candidate species whose YPs could be endocytosed but in which the polypeptide sequence is likely to have diverged from that of *D. melanogaster*. The *yp* genes from the chosen candidate would be cloned, their DNA sequence determined, and the derived protein sequence used in sequence comparisons with those of *D. melanogaster*. Domains important for receptor recognition should be conserved and

identifiable by this approach.

This introduction will describe oogenesis in insects, the yolk proteins, receptor mediated endocytosis, and the evolutionary relationship between the Dipteran species used in this study. The divergence in sequence between other proteins which have been studied in different Diptera will be discussed.

Oogenesis

The ovaries of Diptera are of the polytropic meroistic type in which a cluster of nurse cells and the oocyte are included within the follicular epithelium. The nurse cells and the oocyte are germ cell derived. The nurse cells are responsible for the synthesis of most of the RNA which is transferred to the oocyte (Mahowald and Kambysellis, 1980)

A common oviduct connects the two ovaries of adult female *Drosophila* flies. The oviduct leads to the uterus where the mature eggs are fertilised by sperm stored in the ventral receptacle and the spermathecae. Each ovary contains between 15 to 20 ovarioles in parallel, joined at their tips by terminal filaments. Each ovariole is surrounded by an epithelial sheath containing tracheoles and smooth muscle, which contracts rythmically to push the developing egg chamber from the germarium towards the

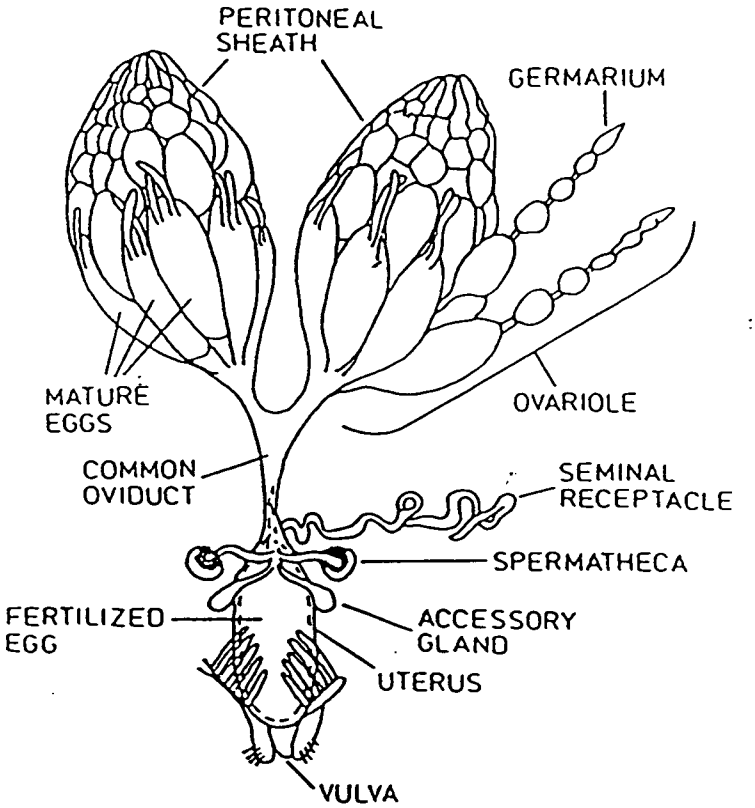
oviduct (Figure 1.1A) (Bownes and Dale, 1982). A convenient staging system based on morphology has been described to identify the developing egg chambers (Figure 1.1B) (King *et al* ., 1956). This staging system has been applied successfully to other fly species such as *Musca* (Goodman *et al*., 1968).

In the germarium, division of the stem cell gives rise to two daughter cells, one of which remains a stem cell while the other continues to divide to form a cluster of 16 cells. The resulting cluster of cells arises through four mitotic divisions with incomplete cytokinesis. The cells are thereby connected by intercellular bridges (Mahowald and Strassheim, 1970). One of the sixteen cells develops into the oocyte while the others become nurse cells. This cluster of cells is surrounded by profollicle cells of somatic origin, and together these cells comprise the egg chamber. The nurse cells provide the oocyte with essential components such as lipid droplets (King, 1970) and many RNA species (Mahowald and Tiefert, 1970). Quantitatively most of the RNA is ribosomal (Mahowald and Tiefert, 1970), although mRNA species critical to determining subsequent embryonic development are also present (Akam, 1987).

The oocyte grows 10^5 fold during oogenesis (King, 1970). Yolk is not the only cause of the increase in size of the oocyte and other nurse and follicle cells products

A.

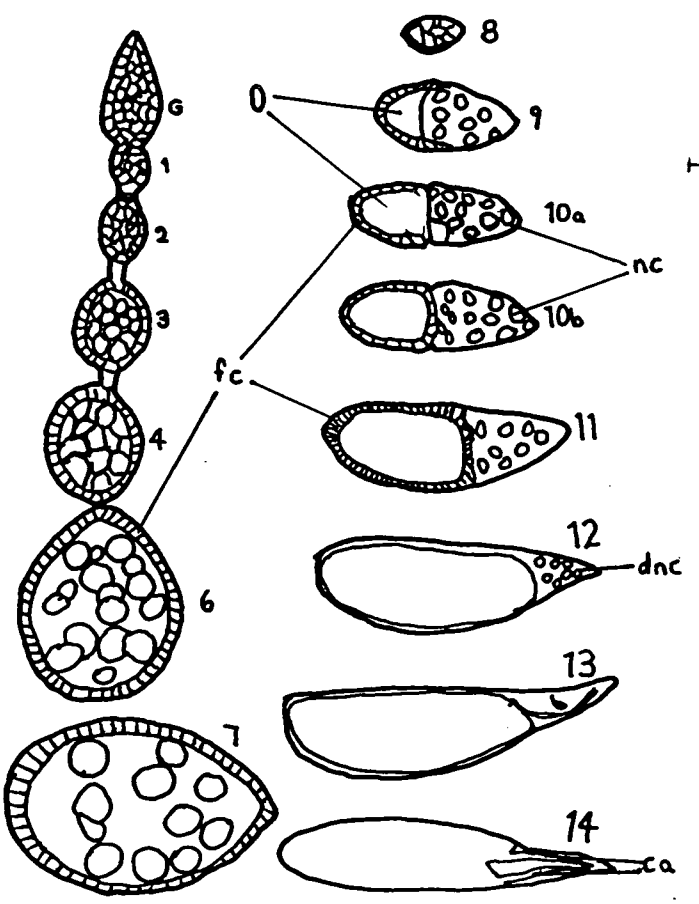
200μ



B.

10μ

100μ

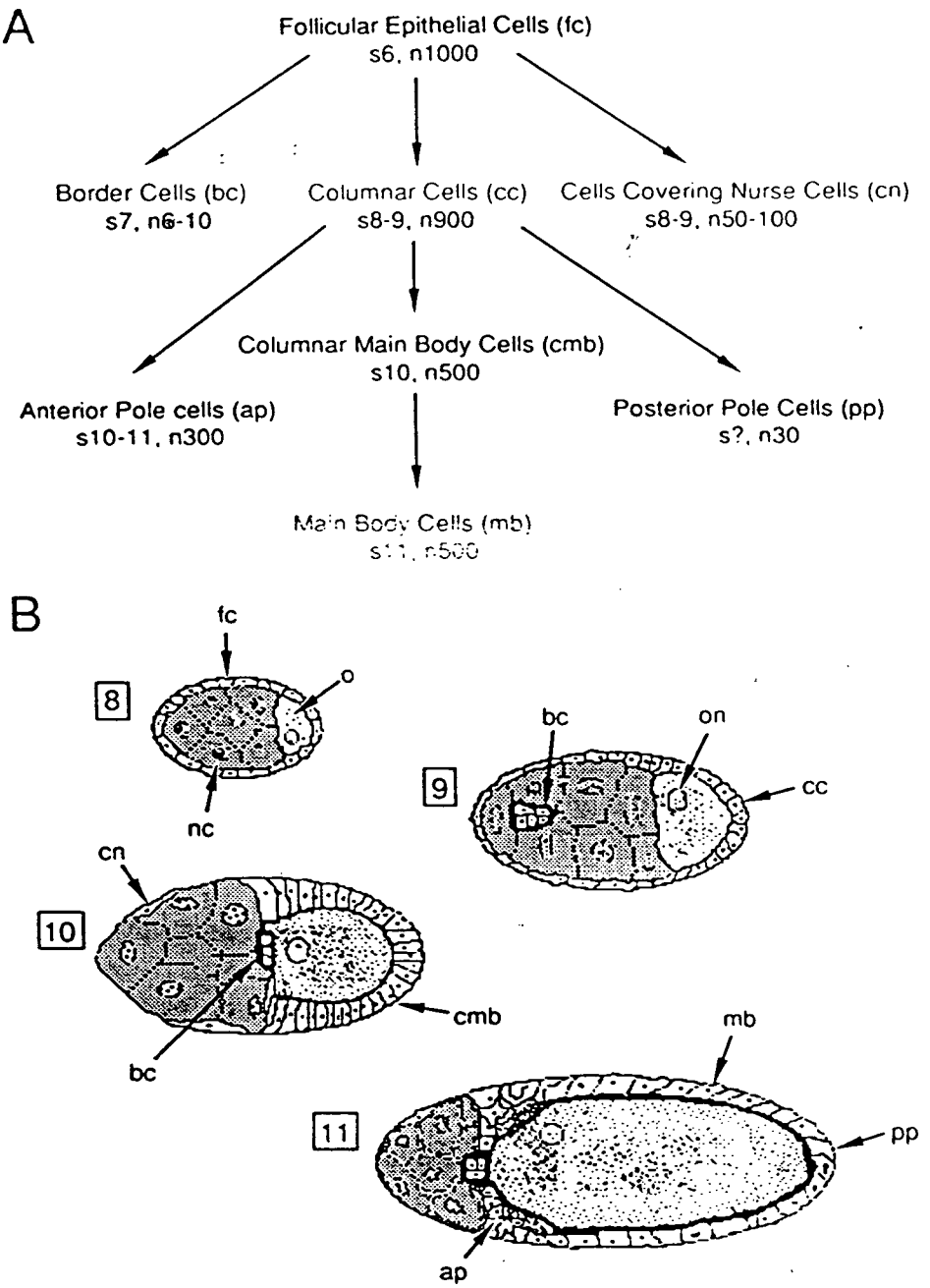


are also actively involved. The involvement of the nurse cells is first observed with the process of polyploidisation as soon as the egg chamber leaves the germarium (Jacob and Sirlin, 1959). A gradient of polyploidisation develops between anterior and posterior nurse cells whereby, by stage 6, the homologues do not show somatic pairing, hence these polyploid chromosomes are not visible. It was suggested that polytene chromosomes can not form because of the high transcriptional activity during oogenesis (Schultz, 1956); however, in *Calliphora* a strain can be induced to form polytene chromosomes yet it still retains the ability to produce RNA at the same constant rate (Bier, 1959). Furthermore, Stalker (1954) showed that 16 species belonging to 8 families revealed banded polytene chromosomes in ovarian nurse cells in a study that covered 190 species in 40 families of Diptera. In *D. melanogaster* the recessive female sterile mutation *fes* or *fs(2)B* produces banded polytene chromosomes in the nurse cells (Koch and King, 1964). It is more probable, therefore that the mechanism involved in somatic pairing in the wild type nurse cells of *D. melanogaster* is not functioning. By stage 10 the anterior nurse cell nuclei have reached a ploidy of 1048 whereas the DNA of posterior nurse cell nuclei has replicated only twice (Schulz, 1956).

The polyploid nurse cells are involved in producing RNA to be transferred into the oocyte. In *Musca*, it has

been shown that the RNA transfer takes place in two phases. In nurse cells, up to stage 10, there is rapid synthesis of RNA which accumulates in the cytoplasm, resulting in an increase in nurse cell size, such that at stage 10B, they are similar in size to the growing oocyte. At the end of stage 10B a massive influx of nurse cell cytoplasm into the oocyte takes place and with it the RNA accumulated in the nurse cell cytoplasm (as determined by autoradiography) (Bier, 1963). In contrast, autoradiography experiments in *Drosophila* (Zalokar, 1965; Mahowald and Tiefert, 1970) and *Calliphora* (Bier *et al.*, 1967) have shown that a very low rate of RNA synthesis is observed in the oocyte through oogenesis.

The other component of the egg chamber consists of the follicle cells. As stated before, the nurse cells and the oocyte are enveloped by profollicle cells. Little is known about these except that by stage 2 of oogenesis there are 80 of them (King, 1970; Bownes and Dale, 1982). From stage 2 to 5 the cells undergo four mitotic divisions to reach approximately 1200 follicle cells (King and Koch, 1963). During stages 7 to 13 a series of cell movements take place (Figure 1.2A and B). The first set of movements takes place between stage 7 and 10, where most follicle cells, or border cells, move over to surround the expanding oocyte and a few cells are left covering the nurse cells. The second movement results from a group of anterior follicle cells moving between the nurse cells to


A
Follicular Epithelial Cells (fc)
s6, n1000
Border Cells (bc)
s7, n6-10
Columnar Cells (cc)
s8-9, n900
Cells Covering Nurse Cells (cn)
s8-9, n50-100
Columnar Main Body Cells (cmb)
s10, n500
Anterior Pole cells (ap)
s10-11, n300
Posterior Pole Cells (pp)
s?, n30
Main Body Cells (mb)
B
fc
o
8
nc
bc
on
9
cc
cn
10
cmb
bc
mb
11
pp
ap

the anterior surface of the oocyte and becoming located near to the oocyte nucleus. These cells are involved in vitelline membrane deposition and in the formation of the micropyle. The last movement, during stages 10 and 11, involves cells at the anterior end of the oocyte migrating centripetally between the nurse cells and the oocyte. These cells are involved in vitelline (stages 9 to 11) and chorion (stages 12 to 14) membrane deposition as well as the formation of chorionic appendages. After stage 14 the synthesis of the chorion stops and the follicle cells degenerate. It is important to note that the cytoplasm of the follicle cells has an elaborate endoplasmic reticulum network which is intercalated with Golgi regions, in which electron dense vesicles are formed (Mahowald and Kambysellis, 1980). This morphology is indicative of a high level of protein synthesis, coupled with secretion.

Vitellogenesis

Vitellogenesis is the process of yolk protein (YP) synthesis, specific uptake and accumulation which occurs during stages 8 to 10 (inclusive) of oogenesis. In *Drosophila* and other Diptera, the yolk proteins do not seem to be processed once internalised into the oocyte. The oocytes of *Drosophila* have three types of yolk spheres, namely protein spheres (or alpha spheres), glycogen spheres (or beta spheres) and neutral lipid droplet spheres (King, 1970). The alpha spheres are

composed of a crystalline protein which has been shown to be the storage state of the YPs by immunogold labelling using antibodies raised against the YPs of *D. melanogaster* (Butterworth *et al.*, 1991). It is also possible to see some alpha spheres that are not crystalline and may be in the stage before the formation of the mature yolk spheres (King, 1970). The formation of the alpha yolk spheres is continous between stages 8 to 10 (inclusive). In adult female flies the deposition of the yolk begins a few hours after eclosion. The beta spheres are formed mainly during stages 13 and 14 of oogenesis, and are recognised by their association with mitochondria. The lipid droplets are contributed by the nurse cells from stage 8 onwards (King, 1970).

The YPs are synthesised in the fat body (Gelti-Douka *et al.*, 1974; Kambysellis, 1977; Bownes and Hames, 1978; Postlethwait and Kaschnitz, 1978) and ovaries of adult female *Drosophila* (Bownes, 1980, Brennan *et al.*, 1982; Issac and Bownes, 1982). The follicle cells surrounding the oocyte have been shown to synthesise YPs only at specific stages of vitellogenesis (stages 8, 9, 10A and 10B) (Brennen *et al.*, 1982; Issac and Bownes, 1982). In most other insects the vitellogenins are known to be synthesised exclusively in the fat body (Kunkel and Nordin, 1985). Likewise, in *Xenopus* (Tata, 1976) and chicken (Van het Schip *et al.*, 1987) their synthesis takes place in the liver (the equivalent organ in vertebrates to

the fat body of insects). In *Sarcophaga bullata* it has been shown by immunocytochemistry that the ovarian follicle cells synthesise the YPs during the vitellogenic stages of oogenesis (Geysen *et al.*, 1987). In *Calliphora erythrocephala* the use of radiolabelled DNA probes and autoradiography demonstrated the involvement of ovarian follicle cells in YP synthesis (Rubacha *et al*, 1988); however the stages at which expression takes place was not investigated.

Yolk proteins synthesised in the fat body are secreted into the haemolymph (Hames and Bownes, 1978), whereas the proteins synthesised in the ovaries are unidirectionally secreted into the intercellular spaces between follicle cells and the oocyte (Bownes and Butterworth, personal communications). Regardless of the site of synthesis the YPs are taken up specifically by the oocyte through receptor mediated endocytosis. The YPs appear in the haemolymph of female adult flies shortly after eclosion, and are first seen in the oocyte 10 hours later (Bownes and Dale, 1982). Yolk protein synthesis correlates very closely with the level of yp mRNA which is first detected just after eclosion, and plateaus 24 hours later (Barnett and Wensink, 1981, Issac and Bownes, 1982).

The yolk proteins do not appear to differ in biochemical properties regardless of their site of

synthesis. However, it has been shown that the rate of YP_3 synthesis relative to YP_1 and YP_2 is up to 90% lower in the ovary than in the fat body (Brennen *et al.*, 1982; Issac and Bownes, 1982). For further detail, page 19 of this thesis. SDS-PAGE analysis of the products of a cell free translation of fat body and ovarian mRNA has shown that the yp_3 message is under represented in the ovary relative to the message for yp_1 and yp_2 (Issac and Bownes, 1982). Northern blots using RNA from fat body and ovary comfirmed the above results. It was demonstrated that no amplification of the *yp* genes took place in either tissue (unlike the chorion genes which are amplified in the follicle cells prior to their transcription (reviewed by Orr-Weaver, 1991)) and therefore the low levels of yp_3 transcripts were not explained by selective amplification of yp_1 and yp_2. Furthermore, the rate of transcription was found to be the same for yp_1, yp_2 and yp_3 (Williams and Bownes, 1986). Therefore, the authors conclude that the yp_3 transcript may be less stable in the ovaries than in the fat body (Williams and Bownes, 1986).

Vitellogenesis is central to the process of oogenesis, and is therefore very well studied. Although vitellogenesis is dependent upon the correct tissue, stage, and sex specific expression of the YPs, other factors affect it and these are dicussed below.

Hormonal regulation

Juvenile hormone (JH) and ecdysone are the two main hormones involved in vitellogenesis in *D. melanogaster*. Juvenile hormone has been implicated in regulating the synthesis of yolk proteins (Postlethwait and Handler, 1979; Postlethwait and Shirk, 1981; Bownes and Blair, 1986; Bownes *et al.*, 1988) as well as in rendering the ovaries competent for yolk protein uptake (Posthlethwait and Handler, 1979; Giorgi, 1979; Bownes, 1980). In *Aedes aegypti*, Juvenile hormone has been shown to mediate the development of the endocytotic machinery in oocytes. Moreover, the oocytes that develop this machinery are the only ones capable of protein uptake (Raikhel and Lea, 1985). The fat body also requires JH to be able to undergo yolk protein synthesis. The production of JH is triggered by the brain in response to a blood meal and JH acts on the *yp* genes in the fat body (Bownes, 1986).

In *Drosophila* there is a brain factor which causes the synthesis of JH in the *Corpus allatum* after eclosion of the fly. The removal of the source of JH by ligation of abdomens stops egg maturation. The resumption of YP synthesis in the fat body takes place after injection of either ecdysone or JH into the isolated abdomens; whereas resumption of YP synthesis in the ovary is only triggered by JH (Jowett and Postlethwait, 1980). Vitellogenesis is prevented by mutations in the *Corpus allatum* affecting JH levels (eg. *apterous*^{4} and *fs(3)A1*; Postlethwait and

Handler, 1978); allactectomy, or treatment with precocene (an anti JH hormone drug) (Wilson *et al.*, 1983).

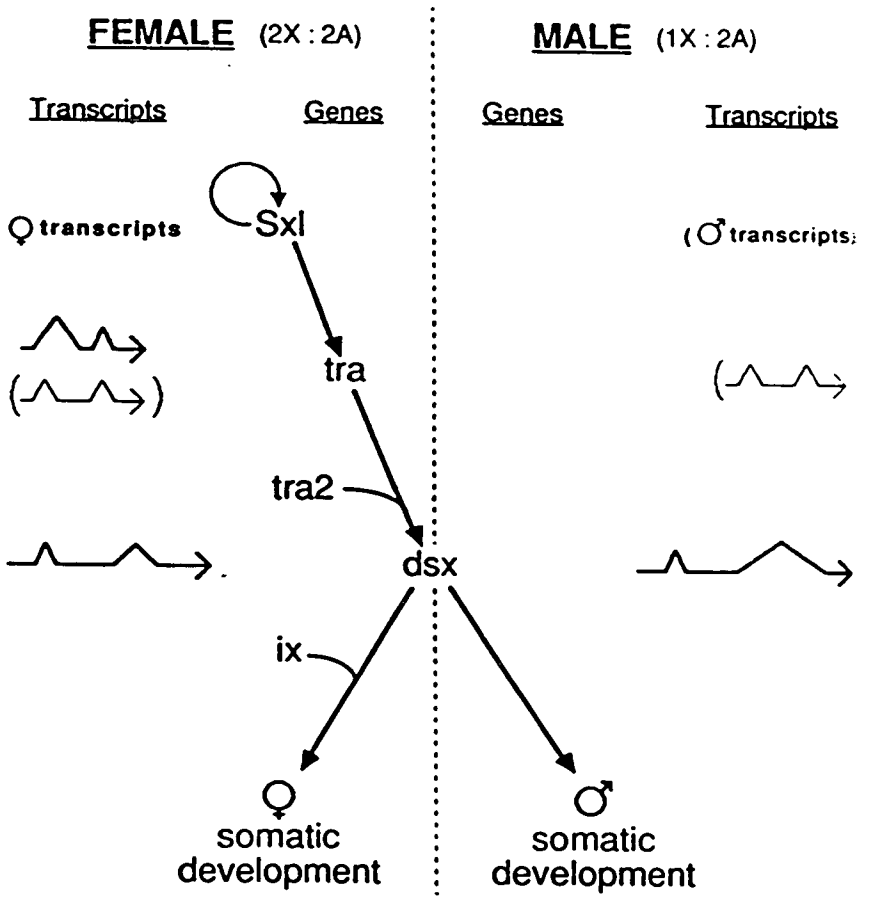
The steroid hormone ecdysone has been shown to affect the synthesis of the *yp* genes. Injection of ecdysone to isolated abdomens causes an increase in YP levels (Postlethwait and Jowett, 1980), however, YP synthesis is not induced in ovaries after flies were injected with ecdysone (Bownes *et al.*, 1983). Ecdysone induces YP production in male fat bodies (artificially high levels are required for an effect to be visible) (Postlethwait *et al.*, 1980) and female fat bodies (Bownes *et al.*, 1983). The effect of ecdysone in males is transient and inhibited by cycloheximide (a protein synthesis inhibitor), thereby indicating the requirement of other gene products for the ecdysone to result in activation of transcription of yps in males (Bownes *et al.*, 1987). The induction of the YP synthesis by ecdysone in males has been observed in *D. melanogaster* sibling species, *D. virilis*, *D. funebris* (Kozma and Bownes, 1986a) *Musca* (Adams *et al.*, 1985), *Calliphora*, *Phormia*, *Lucilia and Sarcophaga* (Huybrechts and De Loof, 1982).

Nutrition

Vitellogenesis in *D. melanogaster* is affected by diet; whereby feeding flies with a sugar diet results in a reduced rate of egg laying compared to flies reared with a

complete diet (Bownes and Blair, 1986). The restoration of the diet brings a slow recovery in egg laying. However, the treatment of starved flies with ZR515 (a Juvenile hormone homologue) results in a rapid but temporary resumption of vitellogenesis as judged by ovary morphology. On the other hand, treatment of starved flies with 20-hydroxyecdysone has no effect on vitellogenesis (Bownes and Blair, 1986). The effect of starvation on *yp* gene synthesis can be partly overcomed by Juvenile hormone (JH), ecdysone, casein or a sugar diet, but complete recovery is brought about only by a complete diet (Bownes *et al.*, 1988). Both hormones increase the yp transcript levels in the fat body of starved flies. This is not an effect on general transcription rates since the α-tubulin transcripts were unchanged (Bownes and Blair, 1986). The effect of starvation is not due to under-development of the fat body or a general effect on protein synthesis (Bownes *et al.*, 1988).

In other insects such as the Mosquito *Aedes aegypti*, oogenesis is arrested in previtellogenic stages, until a blood meal triggers vitellogenesis (Raikhel and Lea, 1985). For vitellogenesis to proceed, feeding has to be coupled with digestion (Van Handel and Lea, 1984). Similar observations have been made in *Musca domestica* (Morrison and Davis, 1964). In *Phormia regina* (Pappas and Fraenkel, 1977), *Protophormia terrae-novae* (Harlow, 1956) and *Calliphora erythrocephala* (Fraenkel, 1940;


FEMALE (2X : 2A)
MALE (1X : 2A)
Transcripts
Genes
Genes
Transcripts
♀ transcripts
Sxl
(♂ transcripts)
tra
tra2
dsx
ix
♀
somatic development
♂
somatic development

Strangeways-Dixon, 1971) autogeny (ie. vitellogenesis is initiated and completed without protein feeding) has not been observed while in *Sarcophaga bullata* only incipient autogeny (ie. sugar initiates vitellogenesis but does not complete it) has been observed (Pappas and Fraenkel, 1977). On the other hand autogeny has been observed in *Sarcophaga argyrostoma* (Denlinger, 1971) as well as in three other tropical *Sarcophaga* species (Chadha and Denlinger, 1976). It is important to note however, that full or incipient autogeny has been shown to vary with species, strain, larval rearing and general adult condition (Strangeways-Dixon, 1961; Pappas and Fraenkel, 1977).

Sex genes

In *Drosophila*, the primary determinant of sex is the number of X chromosomes relative to autosomes (A) in the cell, X:A ratio. When the ratio is 1 (2X:2A), a female cell is produced, and if the ratio is 0.5 (1X:2A) a male cell is formed. The assessment of the X:A ratio and determination of the sex of the cells takes place during embryonic development after formation of the blastoderm (Baker and Belote, 1983). The X:A ratio exerts control over a hierarchy of regulatory genes, namely sex-lethal (*Sxl*), transformer (*tra*), transformer-2 (*tra-2*), double sex (*dsx*) and intersex (*ix*) (Figure 1.3). The action of several of these genes in sex-determination is regulated

at the level of differential processing of the primary transcript in males and females. Female specific splicing of the *Sxl* transcript results in an active protein that dictates the female specific *tra* splicing pattern. The resulting *tra* protein directs the female specific *dsx* splicing pattern. Conversely, *Sxl* male specific splicing truncates the reading frame resulting in an inactive protein product. In the absence of *Sxl* product, the *tra* transcript is spliced in a non-sex specific mode resulting in a truncated and inactive protein produced. Equally, the lack of *tra* product results in the *dsx* transcript being spliced in a male specific mode. The male and female specific *dsx* transcripts both encode active proteins with opposite roles in sex determination (reviewed in Slee and Bownes, 1990).

Sxl is the central switch for sexual dimorphism and is only active in females. If turned off by mutation (even in chromosomally female cells) it leads to male development (reviewed in Slee and Bownes, 1990). The *Sex-lethal* protein product is known to have an RNA binding domain and is involved in directing the female specific splicing of its own transcript and that of *tra*. Bell *et al.* (1991), showed by ectopically expressing a female *Sex-lethal* cDNA in male flies, that the female determined state was maintained by *Sxl* positive autoregulation, while the male state was determined by default. The female specific *tra* protein product along with a product of *tra-2*

direct the female specific splicing of the *dsx* transcript. Recently, *tra-2* protein has been reported to bind to the second intron of the *dsx* pre-mRNA to direct the female specific splicing and polyadenylation of the *dsx* transcript (Hedley and Maniatis, 1991). DSX^F (ie. the double sex female specific protein product) together with the product of *ix* represses male terminal differentiation genes resulting in female sexual phenotype (Slee and Bownes, 1990). In males, in absence of functional *tra* protein, the *dsx* transcript is spliced in the default mode resulting in a male specifc product (DSX^M) which represses female differentiation genes (Slee and Bownes, 1990).

Sxl activation requires maternally provided products of two genes, daughterless (*da*) and Daughterkiller (*Dk*). *da* and *Dk* show the same phenotype when mutant resulting in maternal-effect female embryonic lethality. The phenotype can be rescued by a zygotic constitutive *Sxl* mutation. *da* has been cloned and the protein sequence shows similarities with the *hairy* and *bicoid* gene products, both of which are involved in regulating *Drosophila* segmentation.

In other Diptera sex determination has not been studied in such detail as in *D. melanogaster*. The best understood species is *Musca domestica* where three different pathways for sex determination have been observed. In one pathway, males are determined by the Y

chromosome which has a male determining factor, M. The second system has M on an autosome so that XX male or female flies are possible (Franco *et al.*, 1982; Inoue and Hiroyoshi, 1986). The presence of M, on the Y chromosome or autosome, may be regarded as a polymorphism (Inoue and Hiroyoshi, 1986). In the third system both males and females have M in a homozygous state and it is a female determining factor (F) that determines sex. Therefore, the activity of F is epistatic over M. In addition to these sex factors *Musca* also has maternal effect genes, Arrhenogenous (*Ag*) and transformer (*tra*), that behave in a similar fashion to each other (Nothiger and Steinmann-Zwicky, 1985; Inoue and Hyroyoshi, 1986). It has been proposed (Table 1.1) that sex determination in *Musca* has many similarities to sex determination in *D. melanogaster*; where *Ag* and *tra* are equivalent to *da* and *Dk*, *M* is equivalent to the X:A ratio and *F* is equivalent to *Sxl*. Genes equivalent to *tra*, *tra-2*, *dsx* and *ix* in *D. melanogaster* have not yet been found in *Musca* (Nothiger and Steinmann-Zwicky, 1985a; Inoue and Hyroyoshi, 1986). In other flies such as *Chironomas tentans* (Thompson and Bowen, 1972), and *Chysomya rufifacies* (Ullerich, 1984), the *F* gene has been identified. In *Chironomus* species (Martin and Lee, 1984; Hagele, 1985), the sex determining gene *M* appears to behave as a transposable element as it may be found in different chromosomes.

The generation of sex specific somatic transcripts in

Table 1.1 The proposed parallels between Musca and Drosophila Sex determination

Species	Maternal genes	Primary signal	Key gene	Subordinate control genes	
				tra-2	ix
Drosophila	da, Dk	X:A	Sxl	tra	dsx
Musca	tra. Ag	M	F	to be determined	

Based on Nothinger and Steinmann-Zwicky, 1985.

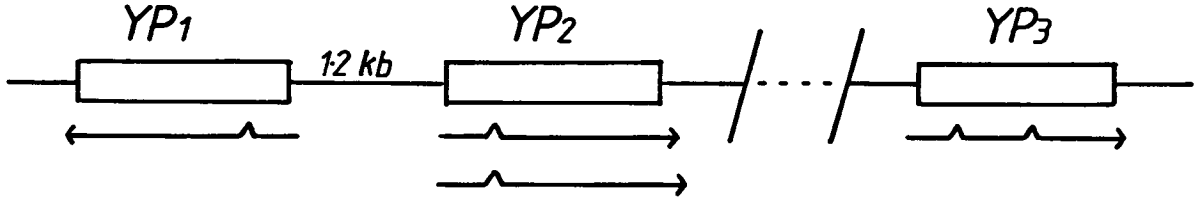
Drosophila can be achieved in at least three different manners. Firstly, the regulatory genes direct the formation of sex specific tissues where subsequent gene expression is governed by tissue specific factors and is independent of the regulatory hierarchy (Wolfner, 1988). Such is the case for the ovarian expression of the yolk proteins (Bownes *et al.*, 1990). Bownes *et al.* (1990) showed that the ovarian expression of the YPs is dependent upon tissue specific factors by using XX; *tra-2*^{ts} mutant flies, in which chromosomal females at 16°C (permissive temperature) develop as females and produce YPs but at 29°C (restrictive temperature) develop as males and do not synthesise YPs. The XX; *tra-2*^{ts} flies were raised at the permissive temperature and shifted to the restrictive temperature 1-2 days or 0-2 hours before eclosion. The ovaries were found to be vitellogenic even when transplanted into normal female or male hosts kept at the restrictive temperature, showing that ovarian YP synthesis continues in the ovary at the restrictive temperature. The second type of sex specific somatic gene regulation takes place when the production of a sex specific transcript is dependent upon the continous function of the regulatory hierarchy (Wolfner, 1988). For example, the fat body expression of yolk proteins is regulated in this way as evidenced by its continual dependence on a functional *tra-2* gene (Belote *et al.*, 1985; Bownes *et al.*, 1987). The continual dependence upon *tra-2* of the YP fat body expression has been shown by temperature shift

experiments using a *tra-2* temperature sensitive mutant *tra-2*^{ts1}. Belote *et al.* (1985), showed that *tra-2* was required for initiation of yp transcription and was likely to be needed for maintenance of their synthesis. The requirement of *tra-2* for maintenance of YP synthesis was demonstrated by shifting *tra-2*^{ts1} flies to 29°C and once YP synthesis was greatly reduced back to 18°C, where yp transcription was resumed (Bownes *et al.*, 1987). Finally, the third mode of sex specific somatic gene regulation may involve the interactions between different cell types (Wolfner, 1988). This is illustrated by experiments showing that the formation of male specific muscle in the fifth abdominal segment (A5) requires interactions with neurons of the correct sex and genotype and is not influenced by the genotype of the myoblast that contributes to the A5 male muscle (Lawrence and Johnston, 1986).

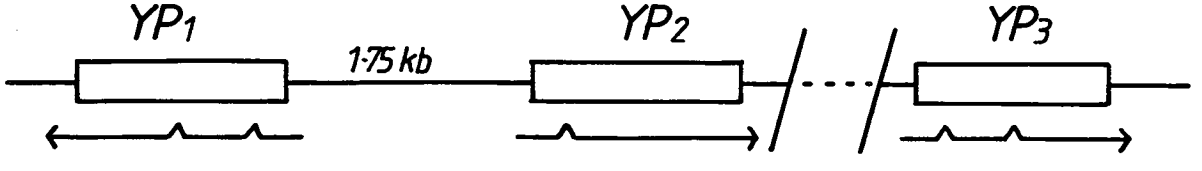
The Yolk Proteins

The yolk proteins in *D. melanogaster* have been the subject of intense molecular studies. The three *yp* genes are located on the X chromosome and each is found in a single copy (Barnett *et al.*, 1980; Postlethwait and Jowett, 1980). The *yp*$_1$ and *yp*$_2$ genes are closely linked (separated by a 1.2 kb intergenic region) and divergently transcribed (Hung *et al.*, 1982) (Figure 1.4A). The *yp*$_3$ gene is located about 1000 kb away from the *yp*$_1$ and *yp*$_2$

D.melanogaster



D.grimshawi



genes (Barnett *et al.*, 1980, Bownes, 1982) The *yp* genes in other *Drosophila* species have also been localised to the X chromosome, namely *D. yakuba*, *D. teissieri*, *D. erecta*, *D. simulans*, *D. mauritiana*, *D. sechellia*, *D. funebris* and *D. virilis* (Kozma amd Bownes, 1986a). In *D. grimshawi* (Figure 1.4B), the three *yp* genes are also located on the X chromosome, where yp_1 and yp_2 are divergently transcribed and have an intergenic region spanning 1.75 kb (Hatzopoulus and Kambysellis, 1987). The yp_1 and yp_2 genes of *D. melanogaster* each have one intron (Hung and Wensink, 1983), whereas the yp_3 gene has two introns (Garabedian *et al.*, 1987). In *D. grimshawi* the yp_1 and yp_3 genes have two introns while yp_2 has one (Hatzopoulus and Kambysellis, 1987). Hatzopoulus and Kambysellis (1987), use the similarities in *yp* gene organisation and the intron organisation within the *yp* genes of *D. melanogaster* and *D. grimshawi* as evidence for these genes originating from duplication events through evolution. Yan *et al.* (1987), also argue that the *yp* genes in *D. melanogaster* arose from duplication events.

The three *yp* genes of *D. melanogaster* have been cloned and sequenced (Hovemann et al., 1981; Hung and Wensink, 1981; 1983; Garabedian *et al.*, 1987; Yan *et al.*, 1987). The overall sequence identity between the derived protein sequences of the three proteins is only 43%; however, the C-terminal 2/5 of the sequence has 57% identity compared to 29% for the 3/5 N-terminal end

(Garabedian *et al.*, 1987) (Figure 1.5). As described below, the yolk proteins have a significant similarity to vertebrate lipases (Bownes *et al.*, 1988; Perssons *et al.*, 1989). Despite the three YPs of *D. melanogaster* not having strong identity, antibodies raised against them cross-react with the YPs of other species, namely *D. simulans*, *D. mauritiana*, *D. virilis*, *D. orena*, *D. erecta*, *D. teissieri* and *D. yakuba* (Kozma and Bownes, 1986b).

At the biochemical level much work remains to be done. A brief description of the biochemistry of the YPs is given below. The YPs of *D. melanogaster* are known to form oligomers, and have been proposed to be hexamers (P. Issac, Ph.D Thesis, 1982) or tetramers (Gingeras *et al*, 1973; Fourney *et al.*, 1982). The difficulty in assessing the number of units in the oligomers is due to the very unstable nature of the YP complexes of *D. melanogaster*. Fourney *et al.* (1982) were able to demonstrate that the YPs of *Calliphora erythrocephala* oligomerise into tetramers. However, it is not known if oligomers are formed of a single YP species or a mixture of YP_1, YP_2 and YP_3.

The YPs of *D. melanogaster* (Brennan *et al.*, 1980; Minoo and Postlethwait, 1985) and *Sarcophaga bullata* (Cardeon *et al.*, 1986) have been shown to have signal peptides. In both cases *in vitro* translation of yp mRNA

```
           1                                                   50
 yp1.gap   MNPMRVLSLL AC.LAVA... .ALAKPNGRM DNSVNQALKP SQWLSGSQLE
 yp2.gap   MNPLRTLCVM ACLLAVAMGN PQSGNRSGRR SNSLDNVEQP SNWVNPREVE
 yp3.gap   MMSLR..... ICLLATCLLV AAHA....SK DAS.NDRLKP TKWLTATELE
Consensus  M---R----- -C-LA----- ---------- --S------P --W------E

           51                                                  100
 yp1.gap   AIPALDDFTI ERLENMNLER GAELLQQVYH LSQIHHNVEP NY..VPSGIQ
 yp2.gap   ELPNLKEVTL KKLQEMSMEE GATLLDKLYH LSQFNHVFKP DYTPEPSQIR
 yp3.gap   NVPSLNDITW ERLENQPLEQ GAKVIEKIYH VGQIKHDLTP SFVPSPSNVP
Consensus  --P-L---T- --L-----E- GA------YH --Q--H---P -----PS---

           101                                                 150
 yp1.gap   VYVPKPNGDK TVAPLNEMIQ RLKQKQNFGE DEVTIIVTGL PQTSETVKKA
 yp2.gap   GYIVGERGQK IEFNLNTLVE KVKRQQKFGD DEVTIFIQGL PETNTQVQKA
 yp3.gap   VWIIKSNGQK VECKLNNYVE TAKAQPGFGE DEVTIVLTGL PKTSPAQQKA
Consensus  -------G-K ----LN---- --K-----FG- DEVTI---GL P-T-----KA

           151                                                 200
 yp1.gap   TRKLVQAYMQ RYNLQ..QQR QHGKNGNQDY QDQSNEQRKN QRTSSEEDYS
 yp2.gap   TRKLVQAYQQ RYNLQ.PYET TDYSNEEQSQ RSSSEEQQTQ RRKQNGE..Q
 yp3.gap   MRRLIQAYVQ KYNLQQ...L QKNAQEQQQQ LKSSDYDYT. ...SSEE.AA
Consensus  -R-L-QAY-Q -YNLQ----- -------Q-- ---S------ ------E---

           201                                                 250
 yp1.gap   EEVKNAKTQS GDIIVIDLGS KLNTYERYAM LDIEKTGAKI GKWIVQMVNE
 yp2.gap   DDTK...T.. GDLIVIQLGN AIEDFEQYAT LNIERLGEII GNRLVELTNT
 yp3.gap   DQWKSAKAAS GDLIIIDLGS TLTNFKRYAM LDVLNTGAMI GQTLIDLTN.
Consensus  ---K------ GD-I-I-LG- -------YA- L-----G--I G-------N-

           251                                                 300
 yp1.gap   LDMPFDTIHL IGQNVGAHVA GAAAQEFTRL TGHKLRRVTG LDPSKIVAKS
 yp2.gap   VNVPQEIIHL IGSGPAAHVA GVAGRQFTRQ TGHKLRRITA LDPTKIYGKP
 yp3.gap   KGVPQEIIHL IGQGISAHVA GAAGNKYTAQ TGHKLRRITG LDPAKVLSKR
Consensus  ---P---IHL IG----AHVA G-A----T-- TGHKLRR-T- LDP-K---K-

           301                                                 350
 yp1.gap   KNTLTGLARG DAEFVDAIHT SVYGMGTPIR SGDVDFYPNG PAAGVPGASN
 yp2.gap   EERLTGLARG DADFVDAIHT SAYGMGTSQR LANVDFFPNG PSTGVPGADN
 yp3.gap   PQILGGLSRG DADFVDAIHT STFAMGTPIR CGDVDFYPNG PSTGVPGSEN
Consensus  ---L-GI-RG DA-FVDAIHT S---MGT--R ---VDF-PNG P--GVPG--N

           351                                                 400
 yp1.gap   VVEAAMRATR YFAESVRPGN ERSFPAVPAN SLQQYKQNDG FGKRAYMGID
 yp2.gap   VVEATMRATR YFAESVRPGN ERNFPSVAAS SYQEYKQNKG YGKRGYMGIA
 yp3.gap   VIEAVARATR YFAESVRPGS ERNFPAVPAN SLKQYKEQDG FGKRAYMGLQ
Consensus  V-EA--RATR YFAESVRPG- ER-FP-V-A- S---YK---G -GKR-YMG--

           401                                                 450
 yp1.gap   TAHDLEGDYI LQVNPKSPFG RNAPAQKQSS YHGVHQAW.. .NTNQDSKDY
 yp2.gap   TDFDLQGDYI LQVNSKSPFG RSTPAQKQTG YHQVHQPWRQ SSSNQGSRRQ
 yp3.gap   IDYDLRGDYI LEVNAKSPFG QRSPAHKQAA YHGMHHAQN. ..........
Consensus  ---DL-GDYI L-VN-KSPFG ---PA-KQ-- YH--H----- ----------

           451
 yp1.gap   Q
 yp2.gap   .
 yp3.gap   .
Consensus  -
```

combined with pancreatic microsomes resulted in cleavage of the signal peptide in the three YPs. The cleavage of the signal petide of the YP_2 protein was also observed in mouse fibroblast L cells transfected with the *yp*$_2$ of *D. melanogaster* (Friederich *et al.*, 1988). The location of the signal peptide cleavage sites on the YPs of *D. melanogaster* are not known. However, using *in vivo* labelling of the *Drosophila* YPs with ^{35}S-cysteine, Minoo and Postlethwait (1985) attempted to locate the site for signal peptidase activity. YP_1 has only one cysteine at residue position 8, YP_2 has two sites at 8 and 12 and YP_3 has four residues at 7, 12, 103 and 312. The only protein found to be labelled after signal cleavage was YP_3, indicating that at least 12 residues are cleaved off by signal peptidase in YP_2 and and 8 residues in YP_1. The site for signal peptidase activity has been predicted to be between residues 19 and 20 in all three yolk polypeptides (Yan *et al.*, 1987; S. Liddell, Ph.D Thesis, 1990).

The YPs were found to be phosphorylated *in vivo*, at both sites of synthesis, and to an equal level, for the three YPs (Brennan and Mahowald, 1982; Minoo and Postlethwait, 1985). The YPs can also be phosphorylated *in vitro* with a cAMP independent protein kinase (Di Mario and Mahowald, 1987). The phosphorylation *in vitro* is unequal, where the YP_2 protein is phosphorylated more than the YP_1 or YP_3. The reasons for this are unknown but it

is suggested that it may be due to the phosphorylation of previously unphosphorylated sites rather than simple phosphate exchange (Di Mario and Mahowald, 1987).

Glycosylation of proteins is thought to stabilise protein interactions and/or to provide recognition sites for receptor recognition (Paulson, 1989). Glycosylation may take place in two different modes, either N-glycosylation or O-glycosylation. In the latter, glycosylation takes place in serine or threonine residues along the protein. The site for potential N-glycosylation is defined as Asn-X-Ser(Thr) where the asparagine is the acceptor site for the carbohydrates. Minoo and Postlethwait (1985) injected female *D. melanogaster* flies with ^{14}C-mannose and ^{14}C-glucosamine and showed that the YPs remained labelled when treated with endoglycosidase-H indicating that the YPs were not glycosylated via N-glycosylation. The authors therefore concluded that the YPs were O-glycosylated. However, Minoo and Postlethwait (1985) did not show that the radiolabel affixed to the YPs was solely due to the injected radiolabelled monosaccharides rather than to the recycling of the monosaccharides resulting in *de novo* synthesis of radiolabelled amino acids which in turn may have been incorporated into newly synthesised YPs. Mintzas and Kambysellis (1982) determined the carbohydrate content of the YPs purified from *D. melanogaster* eggs and found that only YP_2 had detectable amounts of carbohydrate.

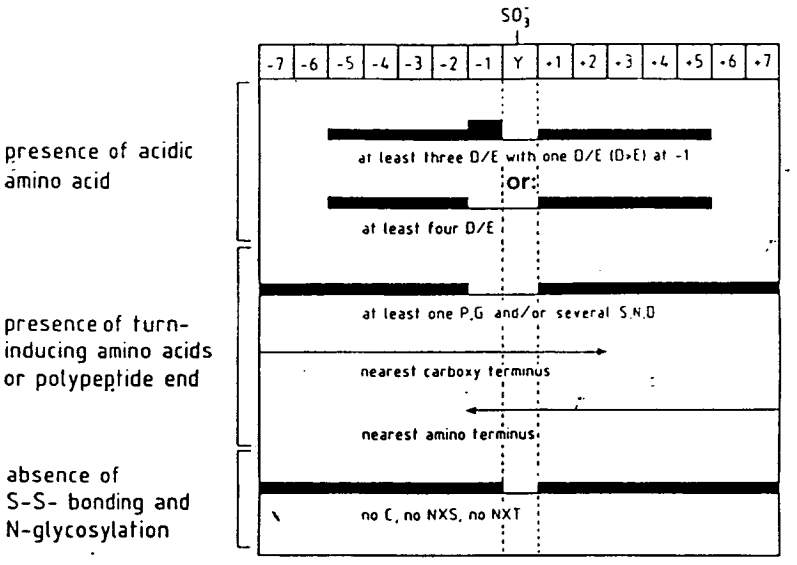
Moreover, the protein sequence of the three YPs of *D. melanogaster* reveal that only YP_2 has a potential N-glycosylation site at residue 25 (Asn^{25}-Arg-Ser) (Hung and Wensink, 1983; Garabedian *et al.*, 1987; Yan *et al.*, 1987). In summary, it appears there is sufficient evidence for YP_2 to be glycosylated. On the other hand, more detailed studies should follow which may describe the actual site and types of glycosylation in the YPs.

The vitellogenins of insects such as *L. migratoria*, *P. americana*, *B. germanica*, *B, discoidailis* (Nordin *et al.*, 1984) *Nauphoeta cinerae* (Imboden *et al.*, 1987) and *Manduca sexta* (Osir *et al.*, 1986) are glycosylated with high mannose oligosaccharides. The glycosylation of the vitellogenins of *B. germanica* have been shown to be involved in the binding of vitellogenin to oocyte membrane preparations (Konig *et al.*, 1988) and the stabilisation of vitellin trimers (Gochoco *et al.*, 1988). The selective uptake of vitellin into the oocyte of *Manduca* is not impaired by treatment of purified vitellin with endoglycosidase-H (Osir *et al.*, 1986). Therefore, it appears that in *Manduca* glycosylation is not important for the recognition of vitellin.

One post-translational modification which has been studied in more detail in YPs is sulfation of tyrosine residues. The sulfation of tyrosine in proteins is known to take place in the Trans-golgi network of cells (Lee and

Huttner, 1985, Baeuerle and Huttner, 1987). It is a highly specific reaction catalysed by the enzyme tyrosylprotein sulfotransferase and is a non-reversible modification, thus having long lasting effects in the secreted protein (Huttner and Baeuerle, 1988). The analysis of sequences known to have sulfated tyrosine residues yielded a consensus sequence for tyrosine sulfation competence (Figure 1.6A) (Hortin *et al.*, 1986; Huttner and Baeuerle, 1988). The latter was used by Huttner and Baeuerle (1988) to predict the tyrosine sulfation sites in YP_1 and YP_2. Here, the same rules were used to predict tyrosine sulfation in YP_3 (Figure 1.6B). The YPs have been shown to be sulfated (Baeuerle and Huttner, 1985) and latterly YP_2 was reported to be sulfated at a single site, tyrosine 172 (Baeuerle *et al.*, 1988), in agreement with the predicted site (Huttner and Baeuerle, 1988). Experiments carried out in mouse fibroblast L cells transfected with a wild-type *yp*$_2$ and a mutated *yp*$_2$ (changing Tyr^{172} to Phe^{172} hence destroying the sulfation site) indicated that sulfation promotes the secretion of the YP_2 protein (Friederich *et al.*, 1988). Although these studies showed that the YP_2 protein was processed as *in vivo*, the importance and significance of the sulfation of the YPs still remains to be studied *in vivo* in *D. melanogaster*. In other studies, the secretion of α-2-Antiplasmin and C4 (ie. fourth component of complement) was not altered after tyrosine sulfation was inhibited in both proteins (both proteins are known to be

A.Consensus constraints used to assess the competence of tyrosine residues to sulfation.



B.Putative tyrosine sulfation sites in the *D. melanogaster* YPs.

YP_1	171	GKNGNQD<u>Y</u>QDQANEQ
	190	RTSSEED<u>Y</u>SEEVKNA
	400	AHDLEGD<u>Y</u>ILQVNPK
YP_2	172	QPYETTD<u>Y</u>SNEEQSQ
YP_3	173	QQLKSSD<u>Y</u>DYTSSEE
	175	LKSSDYD<u>Y</u>TSSEEAA
	384	YMGLQID<u>Y</u>DLRDYIL
	390	IDYDLRD<u>Y</u>ILEVNAK

sulfated normally) (Hortin and Graham, 1988). The authors conclude that tyrosine sulfation is required for protection of exposed tyrosines from extracellular oxidants and in some case may have other secondary functions such as a marker for protein secretion.

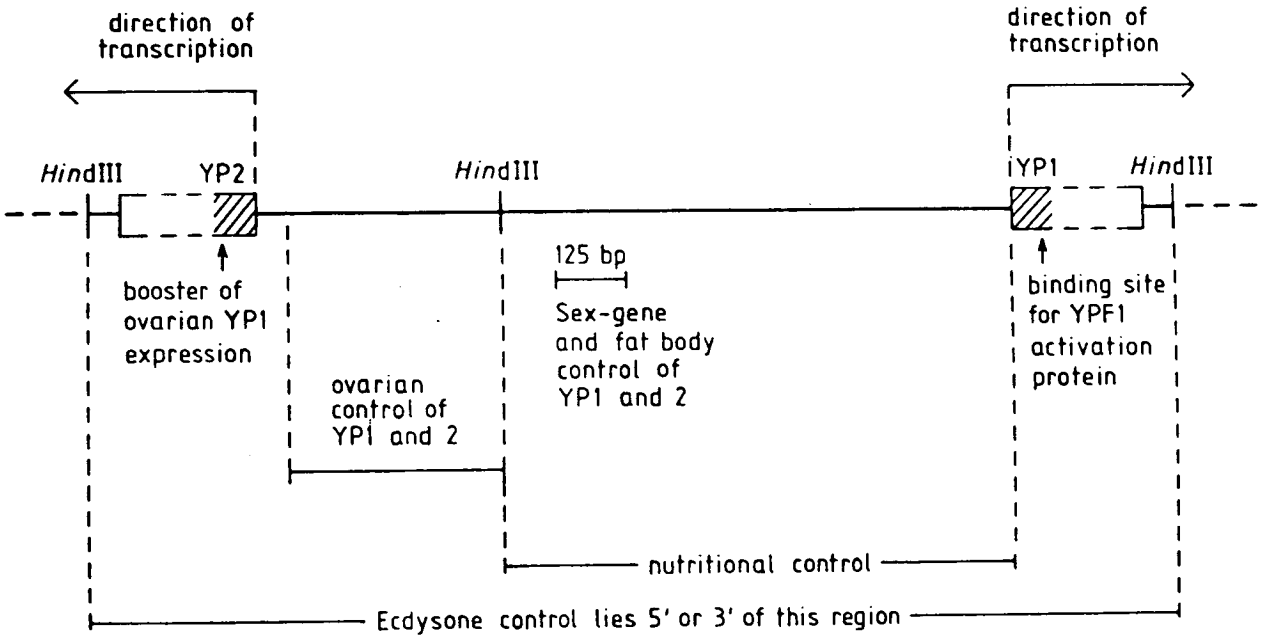
Although little is known about the precise nature of the post-translational modifications of the YPs in *Drosophila*, it is clear that these polypeptides have a signal peptide, and that they undergo phosphorylation, glycosylation and sulfation. Although the importance of all these modifications has not been investigated it is essential to analyse their effect so as to fully understand the targeting and storage of the YPs.

Regulation of the *yp* genes of *D. melanogaster*

In contrast to the limited understanding of the biochemistry of the YPs, the regulation of the *yp* genes has been well characterised. The intergenic region between the yp_1 and yp_2 genes of *D. melanogaster* has been studied in great detail, in order to determine the sequences involved in the regulation of the genes by tissue specific, sex specific and nutritional factors, and hormones. A summary of the data collected is given below in Figure 1.7, which illustrates the fragments within the 1.2 kb intergenic region that have been identified to be of importance in regulation of the *yp* genes.

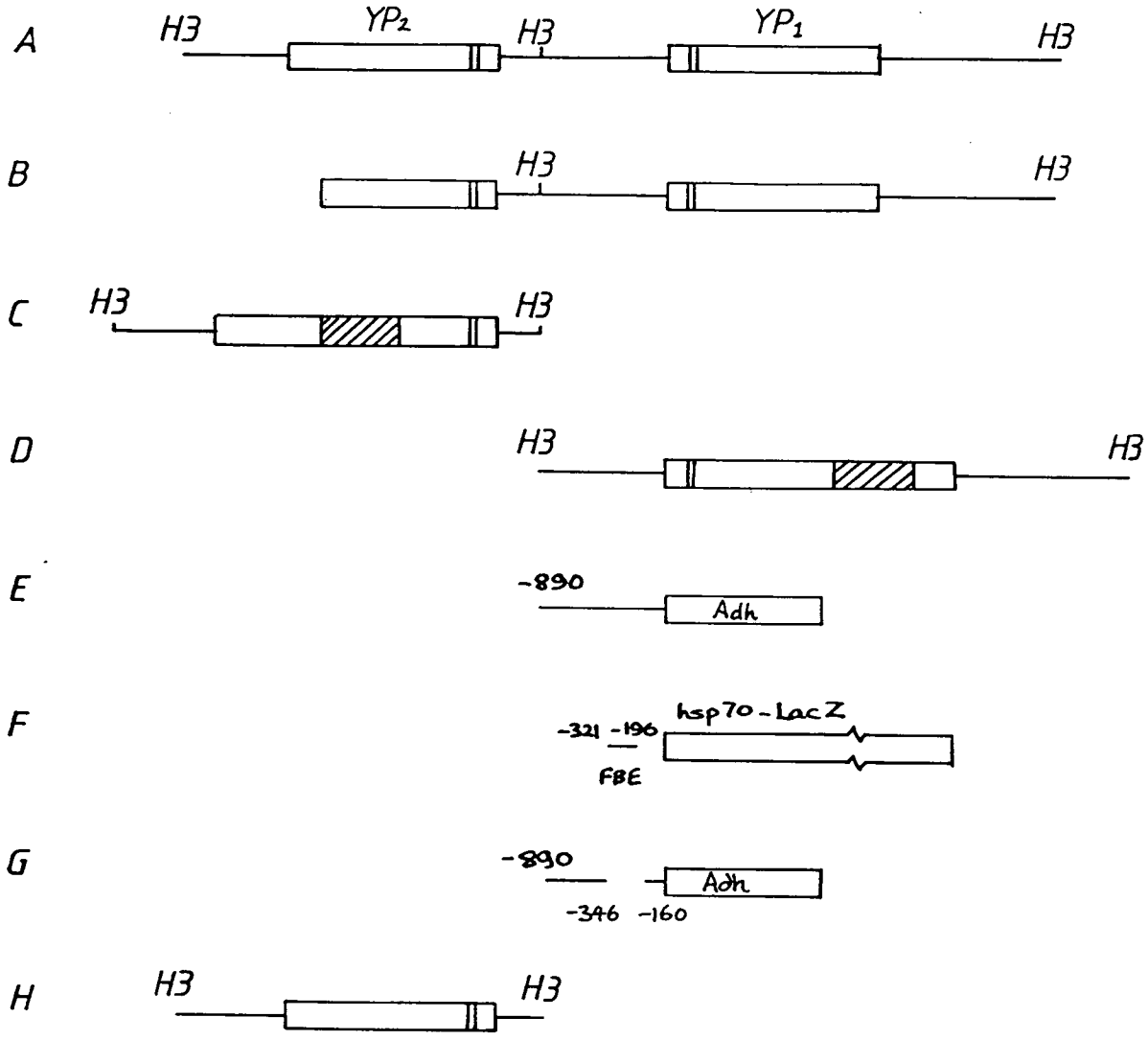
Figure 1.7

A. Summary of the regulatory sequences position in the intragenic region of the yp_1 and yp_2 genes of *D. melanogaster*.


direction of transcription
direction of transcription
HindIII
YP2
HindIII
YP1
HindIII
booster of ovarian YP1 expression
125 bp
Sex-gene and fat body control of YP1 and 2
binding site for YPF1 activation protein
ovarian control of YP1 and 2
nutritional control
Ecdysone control lies 5' or 3' of this region

The majority of the studies carried out to determine the location of *cis*-regulatory elements have involved germ line transformations using DNA upstream of the *yp* genes fused to either *yp* genes tagged with M13 bacteriophage sequences or to different reporter genes such as alcohol dehydrogenase (*Adh*) and β-galactosidase (*lacZ*).

Tamura *et al.* (1985), showed that a 5 kb region encompasing the yp_1 and yp_2 genes and the intergenic region was sufficient for correct tissue, stage and sex specific expression of the two *yp* genes (Figure 1.8B). Garabedian *et al.* (1985), reported that dividing the intragenic region at the HindIII site, such that yp_2 retained 342 bp of upstream region (Figure 1.8C) and yp_1 890 bp (Figure 1.8D), resulted in yp_2 being expressed only in the ovaries and yp_1 only in the fat body. The constructs used in those experiments had the *yp* genes tagged with M13 sequences. The 890 bp upstream fragment of yp_1 fused to the *Adh* coding sequences (Figure 1.8E) also confers fat body, sex and stage specific expression (Shirras and Bownes, 1987); moreover, injection of 20-hydroxyecdysone into males transformed with this construct did not induce expression of the *Adh* gene, even though the native *yp* genes were induced. Therefore, separation of DNA sequence elements mediating the ecdysone inducibility and sex specific activation was achieved (Shirras and Bownes, 1987).


A
H3
YP2
H3
YP1
H3
B
H3
H3
C
H3
H3
D
H3
H3
E
-890
Adh
F
-321
-196
hsp70-LacZ
FBE
G
-890
Adh
-346
-160
H
H3
H3

The search for the enhancer region giving fat body and sex specificity continued using the 890 *bp* fragment. 5'end deletions were fused to the yp_1 gene tagged with M13 and 3'end deletions were fused to an *hsp70-lacZ* gene fusion (Gabaredian *et al.*, 1986). The different deletions were transformed into flies (via P-element transformation) and were analysed *in vivo* for fat body specific expression. The 5'deletions up to -89 from the *yp1* cap site did not show sex or fat body expression; the 3' deletions up to -196 conferred normal expression while deletions up to -321 abolished expression. Further experiments showed that the DNA between -196 and -321 is sufficient for fat body expression and was named fat body enhancer region (FBE). The FBE was fused to a *hsp70-lacZ* fusion gene (Figure 1.8F), and was shown to give fat body and sex specific expression (Garabedian *et al.*, 1986). More recently, four DSX binding sites were mapped in the FBE fragment (Burtis *et al.*, 1991). The intriguing result from Burtis *et al.* (1991), was that DSX^M and DSX^F bind to the same sites on the FBE with the same affinity. Therefore, intersex may play a role in distinguishing between DSX^M and DSX^F to give *yp* expression in females but not in males (Burtis *et al.*, 1991).

The regulation of ovary specific expression by the 342 bp upstream of the yp_2 cap site was dissected. Logan *et al.* (1989), demonstrated that two separate DNA

sequences are required for *yp*$_1$ specific expression in the ovaries. The first is 301 bp long, located from -43 to -342, the second is located within the first 105 bp of the first exon in *yp*$_2$ (Logan and Wensink, 1990).

Abrahamsen *et al.* (1991), have found that the deletion of the 127bp fatbody enhancer region from within the -890 bp upstream region of the *yp*$_1$ gene does not result in the loss of tissue and sex specificity. The construct used to test this was a *yp*$_1$-*Adh* fusion gene where the 890 bp upsteam of the *yp*$_1$ cap site minus the 127 bp FBE were fused to the coding sequences of *Adh*. The total deletion of the fat body enhancer (-104 to -346 and -160 to -346) (Figure 1.8G) or partial deletion of the FBE (-206 to -346 leaving 10bp of FBE, and -267 to -346 leaving 73 bp of FBE) showed the same results. These data indicate that the FBE is not the only element within the 890 bp fragment which is able to confer fat body specific activity.

The ovarian enhancer regions were also reported to direct fat body specificity, using a plasmid in which the *yp*$_2$ gene and 343 bp of upstream sequence to *yp*$_2$ (Figure 1.8H) was introduced into flies and crossed into $YP_2^{12-1245}$ background (these flies lack YP_2 in their haemolymph and ovaries (Tamura *et al.*, 1985)) (Abrahamsen *et al.*, 1991). Previously, the same fragment had been reported to confer ovarian specific expression but not fat

Figure 1.8

Diagram showing the gene fusions used to determine the sites of *cis*-acting elements in the intergenic region between yp_1 and yp_2 genes of *D. melanogaster*. The experiments are discribed in the text.

A.The genomic organisation of the *yp* genes of *D. melanogaster*.

B.Diagramatic representation of construct used by Tamura *et al.*, 1985.

C.yp_2 gene tagged with M13 sequences ().

D.yp_1 gene tagged with M13 sequences ().

E.*Adh* coding region used as a reporter for the upstream regulatory sequences of the yp_1 gene.

F.The FBE fused to an *hsp70-lacZ* gene fusion.

G.*Adh* coding sequences fused to the 5'upstream region of the yp_1 gene lacking the FBE.

H.The yp_2 gene including 343bp upstream from the transcriptional start site. This construct did not contain M13 sequences within the coding region.

body specific expression (Garabedian *et al.*, 1985). However, Garabedian *et al* (1985) and Logan *et al.* (1989) when analysing the 343 bp fragment upstream of the yp_2 gene used the yp_2 tagged by M13 sequences to detect the expression pattern of the construct. It is therefore probable that the M13 sequences within the yp_2 gene may be interfering to upset the expression of the yp_2 gene in the fat body. The discrepancy in the results may be explained by the DSX binding motif being found, by computer data base searches, in clusters in the intergenic region as well as the coding regions of both yp_1 and yp_2 genes (Abrahamsen *et al.*, 1991). The results may reflect the flexibility of the biological system whereby a set of back-up elements may be in place should the one used by default fail. The redundancy of the system may also be explained by the need for several elements to be present to confer the approprate level and pattern of expression.

Bownes *et al.* (1988), using a yp_1-*Adh* fusion gene (which comprises the 890 bp upstream of the yp_1 gene fused to the coding sequences of the *Adh* gene) reported that the 890 bp fragment is required for diet enhanced transcription. The activity of JH hormone was also investigated with this construct, and the authors concluded that JH does not exert its activity via the 890 bp fragment of DNA (Bownes *et al.*, 1988).

It has been suggested that in males sex specific

post-transcriptional controls exist to limit *yp* gene expression after yp mRNA induction by ecdysone in fat bodies of male flies (Bownes *et al.*, 1983). Kraus *et al.* (1988) introduced into the *D. melanogaster* genome via P-element mediated transformation a construct containing a *hsp70-yp*$_1$ gene fusion which when induced produced an authentic *yp*$_1$ transcript (ie. identical to the endogenous yp$_1$ mRNA expressed in females). This construct was induced in males and the yp$_1$ mRNA was found to be normally spliced and translated and the YP$_1$ protein product secreted at comparable levels to those of the wild-type female flies. Thus this experiment demonstrated that the sex-specific regulation of the *yp*$_1$ gene is limited to the level of transcription and does not induce any post-transcriptional processes.

The *yp*$_3$ gene has been studied in less detail, partly due to the fact that it was cloned and sequenced a few years after *yp*$_1$ and *yp*$_2$ (Garabedian *et al.*, 1987). Liddell and Bownes (1991a), have investigated the *cis*-acting sequences required for expression of the *yp*$_3$ gene. The authors found that a 3.1 kb fragment containing the *yp*$_3$ gene was enough to confer ovary, fat body and sex specific expression of the *yp*$_3$ gene. These experiments were carried out using a YP$_3$ mutant (YP3^{a1}) which does not secrete the native YP$_3$ protein from the fat body or follicle cells due to a mutation in the *yp*$_3$ gene (Liddell and Bownes, 1991b).

The regulation of vitellogenesis and *yp* gene expression has been studied in detail in *D. melanogaster*. Despite all the effort invested in understanding the molecular mechanisms involved in *yp* regulation, it still remains at times a confusing area of study. It appears that choosing the yp_1 and yp_2 genes to dissect the molecular mechanisms may not have been ideal since the two genes appear to share the same enhancers within an intergenic region, thus creating difficulties. It is likely, however, that the studies using yp_3 will provide further insight into *yp* gene regulation. The understanding of the control of the *yp* genes would be significantly advanced by investigating the regulation of the *yp* genes in other insects and identifying conserved features of the regulatory mechanisms amongst them.

The Lipases

The lipases are a family of proteins (Komaromy and Schotz, 1987; Persson *et al.*, 1989; Enerback and Bjursell, 1989), with sequence similarity ranging from 77% (human hepatic triglyceride lipase and rat hepatic triglyceride lipase) to 35% (human hepatic triglyceride lipase and porcine pancreatic lipase) (Martin *et al.*, 1988). The functions of these proteins include the hydrolysis of di- and triacylglycerols, phosphatidylcholine and phosphatidylethanolamine contained in intermediate density

lipoproteins and high density lipoproteins (Martin *et al.*, 1988) and hydrolysis of triacylglycerols in very low density lipoproteins and chylamicra (Enerback and Bjursell, 1989). The homology shared by the lipases is concentrated in a central domain which contains the putative lipid-binding region (Komaramy and Schotz, 1987). The similarity of the lipases to the YPs also encompasses the central region (105 amino acids), which shows 24 amino acid identities between the aligned lipases (ie. human lipoprotein lipase, human hepatic triacylglyceride lipase and porcine pancreatic triacylglyceride lipase) and both YP_1 or YP_2, somewhat lower than that reported for porcine pancreatic lipase and human hepatic lipase (37 identities) (Figure 1.9A) (Persson *et al.*, 1989). Comparing the aligned polypeptide sequences of the YPs and triacylglycerol lipase produces 34 identities (Figure 1.9B) (Bownes *et al.*, 1988). Perssons *et al.* (1989), reported that 8 of the 40 identities are glycine residues, which may indicate distant relationships between different protein families (Jeffery *et al*, 1987), and may suggest comformation similarities between lipases and YPs (Perssons *et al.*, 1989). In fact the comparison of hydrophobicity plots covering the central region of homology illustrates large areas of similarity in the pattern of hydrophobicity (Perssons *et al.*, 1989). The central region is involved in lipid binding and contains the catalytic serine; however, the latter is not conserved in YPS (Bownes *et al.*, 1988; Perssons *et al.*, 1989;

Figure 1.9

A. Alignment of the vertebrate lipase proteins with the YP_1 of *D. melanogaster* (as reported by Perssons *et al.*, 1989). *, refers to an amino acid conserved in all five sequences.

. indicates a conservative substitution.

YP_1= Yolk protein 1

YP_2= Yolk protein 2

LPL= Human lipoprotein lipase

HL= Human hepatic lipase

PL= Pig pancreatic lipase

B. Alignment of the YPs of *D. melanogaster* with the rat hepatic triacylglycerol lipase (as reported by Bownes *et al.*, 1988). Capital letters refer to amino acids that are identical in all four sequences. Lower-case letters are shown when two out of the three YPs have an identical residue to the lipase. *, indicates a conservative substitution.

YP1, YP2 and YP3= Yolk proteins 1, 2 and 3

rhl= Rat hepatic triacyglycerol lipase

A

```
YP1   240 GEIIGNRLVELTNTVNVPQEIIHLIGSGPAAHVAGVAGRQ  280
YP2       GAKIGKWIVQMVNELDMPFDTIHLIGQNVGAHVAGAAAQE
LPL       GQDVARFINWMEEEFNYPLDNVHLLGYSLGAHAAGIAG--
HL        GKEVAALLRWLEESVQLSRSHVHLIGYSLGAHVSGFAG--
PL        GAEVAYFVEVLKSSLGYSPSNVHVIGHSLGSHAAGEAG--
          *  ..  .    ..     . .*..* . ..*..* *.

YP1   281 FTRQTGHKLRRITALDPTKIYGKPEERLTGLARGDADFVD 319
YP2       FTRLTGHKLRRVTGLDPSKIVAKSKNTLTGLARDGAEFVD
LPL       --SLTNKKVNRITGLDPAGPNFEYAEAPSRLSPDDADFVD
HL        SSIGGTHKIGRITGLDAAGPLFEGSAPSNRLSPDDASFVD
PL        --RRTNGTIERITGLDPAEPCFQGTPELVRLDPSDAKFVD
                . *.*.**.          .  .*  ..* ***

YP1       AIHTSAY------GMMTSQRLANVDFFPNG 343
YP2       AIHTSVY------GMGTPIRSGDVDFYPNG
LPL       VLHTFTRGS-PGRSIGIQKPVGHVDIYPNG
HL        AIHTFTREH-MGLSVGIKQPIGHTDFYPNG
PL        VIHTDAAPIIPNLGFGMSQTVGHLDFFPNG
          ..**          .      .. .*. ***
```

B

```
Lipase   95 KVESVNCICVDWKGGSR-TGYTQ-ASQNIRIVGAEVA-YFVEVLKSSLGYSPSNVHVIGHSLGSHAAGEAGR---R
YP 1    198 KTQSGDIIVIDL--GSKLNTYERYAMLDIEKTGAKIGKWIVQMVNE-LDMPFDTIHLIGQNVGAHVAGAAAQEFTR
YP 2    199 DTKTGDLIVIQL--GNAIEDFEQYATLNIERLGEIIGNRLVE-LTNTVNVPQEIIHLIGSGPAAHVAGVAGRQFTR
YP 3    207 KAASGDLIIIDL--GSTLTNFKRYAMLDVLNTGAMIGQTLID-LTNXKGVPQEIIHLIGQGISAHVAGAAGNKYTA
            k--s-*-I-*d---Gs----*---A--*i---Ga-**--*v*-l----------*H*IG----*H*AG-Ag----r

            -TNG-TIERITGLDPAEPCFQGTPE--LVRLDPSDAKFVDVIHTDAAPIIPNLGFGMSQTVGHLDFFPNG-GKQMPG 236  Lipase
            LT-GHKLRRVTGLDPSKIVAKSKNT--LTGLARGDAEFVDAIHT-S--VY-GMGTPI-RS-GDVDFYPNGPAAGVPG 338  YP 1
            QT-GHKLRRITALDPTKI-Y-GKPEERLTGLARGDADFVDAIHT-SA--Y---GMGTSQRLANVDFFPNGPSTGVPG 339  YP 2
            QT-GHKLRRITGLDPAKV-LSKRPQ-ILGGLSRGDADFVDAIHT-ST--F-AMGTPI-R-CGDVDLYPNGPSTGVPG 347  YP 3
            -T-G--*-RiTgLDPa-------p---L--L---DA-FVD*IHT-*-------G-------g-*Df*PNG----*PG
```

Enerback and Bjursell, 1989). The YPs have been shown to bind ecdysteroid conjugates but not to have any lipase activity (Bownes *et al.*, 1988). Therefore, the YPs have a lipid binding capability which correlates to the similarities observed between lipases and YPs. Insects, including *Drosophila*, are known to store ecdysone as conjugates in embryos (Redfern, 1989). Furthermore, in *Locusta* polar conjugates are bound to vitellin which when degraded release free ecdysone and this coincides with the secretion of cuticle in embryos (Lageoux *et al.*, 1979). The release of free ecdysone prior to cuticle synthesis in *D. melanogaster* embryos has also been demonstrated (Kraminshy *et al.*, 1980).

Receptor Mediated Endocytosis

The phenomenon of proteins or ligands entering the cell via a specific interaction with a receptor from the extracellular space is defined as receptor mediated endocytosis (Shepherd, 1989). There are several types of sorting once the receptor-ligand is internalised (Figure 1.10);

(i) receptor and ligand separate, where the receptor recycles to the membrane and the ligand is either delivered to lysosomes or storage vesicles (eg the low density lipoprotein (LDL) receptor),

(ii) both receptor and ligand recycle to the membrane (eg. Transferrin receptor),

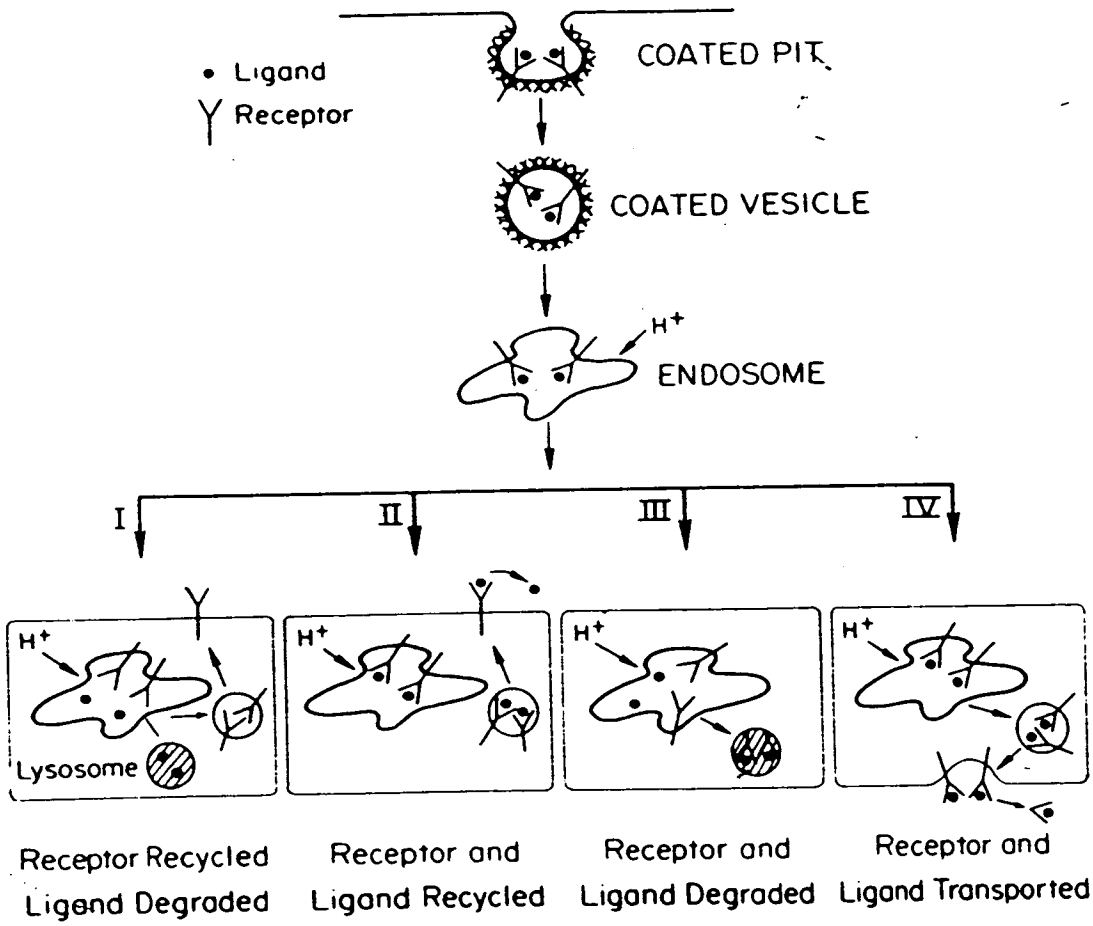
(iii) both receptor and ligand are transported to lysosomes (eg. down regulation of receptors in cell surface)

(iv) receptor and ligand are delivered to the opposite membrane in polarised cells, and the ligand is released into the extracellular space (Carvenberg and Howell, 1989; Shepherd, 1989; Diaz *et al.*, 1989).

The bulk of the discussion of receptor mediated endocytosis will be concentrated on the first form of sorting (LDL receptors), since (as described below) the YPs and their receptors behave in a similar fashion to the LDLs and their receptors. However, in all cases the ligand interacts with the appropriate receptor which in

Figure 1.10

The four pathways of receptor-mediated endocytosis. Receptor-ligand complexes enter cells via coated vesicles. Following coat removal receptors continue in one of the four complexes "I-IV". (Based on Goldstein *et al.*, 1985).


• Ligand
Y Receptor
COATED PIT
COATED VESICLE
H+
ENDOSOME
I
II
III
IV
H+
Lysosome
Receptor Recycled
Ligand Degraded
Receptor and
Ligand Recycled
Receptor and
Ligand Degraded
Receptor and
Ligand Transported

turn results in internalisation of the complex.

The precise mechanism by which membrane trafficking takes place is under intensive study. Two models of membrane trafficking in endocytosis have been formulated (Helenius *et al.*, 1983). The first one is vesicular transport (Palade, 1975), where early and late endosomes and lysosomes are pre-existing organelles each interconnected via vesicular transport. This model predicts the presence of resident molecules in each compartment. The second model or the maturation model explains the organelle formation as a progression, where the early endosome is formed by continually fusing endosome vesicles. The early endosome is translocated into the cell and golgi components are added to it resulting in a late endosome. This process continues such that the late endosome eventually becomes a lysosome. In this model, no resident proteins are expected in each of the organelles.

Coated Pits and Coated Vesicles

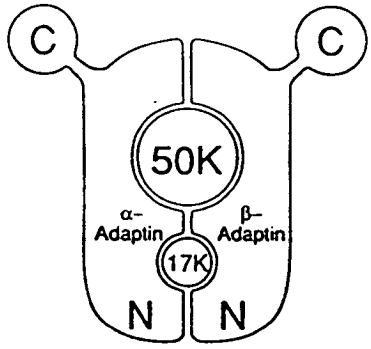
Coated pit formation depends upon the association of clathrin and clathrin associated proteins at the plasma membrane (Gruenberg and Howell, 1989). Some receptors are clustered in the coated pit (eg. LDL receptors), whereas others migrate to the coated pit after interacting with the target molecule (eg. hormone and growth factor

receptors) (Goldstein *et al.*, 1985). Coated vesicles were first observed and described by Roth and Porter (1964) when analysing yolk uptake in mosquito oocytes; their role however, was not recognised as part of receptor mediated endocytosis until studies showed that LDL receptors were concentrated in the coated pit and that clathrin was the protein coating the vesicle (Pearce, 1975). Coated vesicles not only exist in the endocytotic pathway but also in the secretory pathway, where they are involved in the sorting of newly synthesised lysosomal and secretory proteins (Hurtley, 1991). The central role of clathrin *in vivo* depends upon adaptor proteins which interact with clathrin and the cytoplasmic domains of the receptors on the membrane. Clathrin is made up of three heavy chains and three light chains (Ungewiskell and Branton; 1981; Kirchhaven and Harrison; 1981). The light chains are polymorphic (Jackson *et al.*, 1987) and may be interchangable. The light chains have been implicated in the central role of regulating the different aspects of clathrin function (Brodsky *et al.*, 1991).

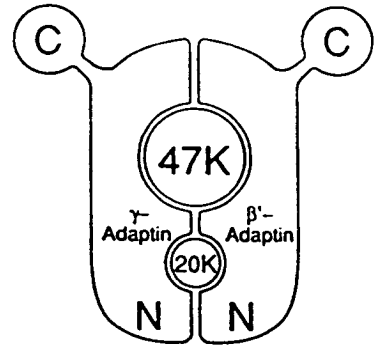
Two classes of adaptor molecules, HA-1 and HA-2, were separated by hydroxy-apatite chromatography (Pearce and Robinson, 1984; Keen, 1987), Figure 1.11 shows a diagramatic representation of the components of each class of adaptor molecule. *In vitro*, the adaptors are required for low concentrations of clathrin to form coats (Zaremba and Keen, 1983). An important observation is that the

Figure 1.11

A diagramatic representation of the HA-2 adaptor in comparison to the HA-1 adaptor.



Plasma membrane adaptor HA-2



Golgi adaptor HA-1

adaptors seem to be restricted to different sets of coated vesicles. HA-2 adaptor molecules (stained using antibodies against γ- adaptin) have been found in plasma membrane coated pits but not in golgi region coated pits (Robinson and Pearce, 1986). In contrast, the HA-1 adaptors (stained using monoclonal antibodies against αadaptin) were found in golgi region coated pits but not those found in the plasma membrane (Ahle *et al.*, 1988). Monoclonal antibodies raised against β and β' adaptins stain in both sites giving a staining pattern similar to that given by anti-clathrin antibodies (Ahle *et al.*, 1988). The latter observation not only comfirms the distribution of HA-1 and HA-2 but also suggests that the vesicles are uncoated when fusing to another compartment (Robinson and Pearce, 1986). Although an uncoating ATPase (a 70 kD protein related to the Heat shock proteins, Hsc70) removes clathrin from vesicles, it is thought that other factors are also involved since the ATPase does not remove the adaptors from vesicles (Schlossman *et al.*, 1984). The Hsc70 protein requires not only ATP but also the presence of the clathrin light chains. Clathrin light chains have a calcium binding site which when calcium is bound causes a comformational change enabling Hsc70 to bind and initiate the decoating process (DeLuca-Flaherty *et al.*, 1990).

This thesis will describe briefly the Mannose-6-phosphate receptor (M6PR) because it is sorted

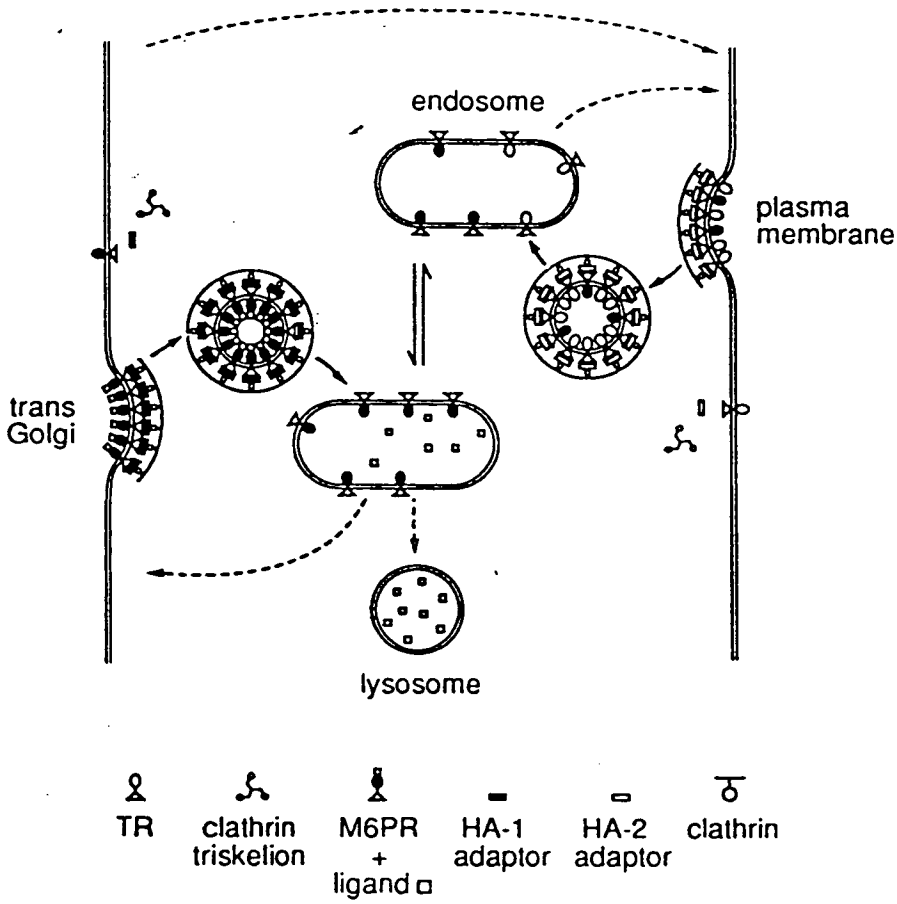
in clathrin vesicles originating from both plasma membrane and golgi. The majority of the M6PR receptor molecules are in the Trans-golgi network (TGN) and the prelysosomal endosome (Figure 1.12). In both the golgi and the plasma membrane, the receptors are concentrated into the coated pits. In the trans-golgi, proteins that bind to the M6PR are segregated from secretory protein and transported to the prelysosomal compartment, where the ligands dissociate. From here the M6PR is recycled back to the TGN; however, some of the receptor is recycled via the plasma membrane, i.e. the M6PR is sorted to the plasma membrane where it is taken back to the endosome via clathrin coated vesicles (Pearce and Robinsom, 1990)

As stated above, the adaptor molecules form a linking "bridge" between clathrin and the vesicle, where there must be an interaction between the adaptor and the cytoplasmic domain of the receptor protein. A tyrosine residue in the cytoplasmic domain of the LDL receptor (Davis *et al.*, 1986), large cation M6PR (Lobel *et al.*, 1989) and transferrin receptor (Jing *et al.*, 1990), is involved in this interaction; as demonstrated by mutants where the tyrosine is mutated to other residues. Moreover, a series of mutants in the cytoplasmic tail of transferrin demonstrated that a 10 amino acid stretch was required for endocytosis, and addition of the tail to the tailess mutants restores endocytosis (Jing *et al.*, 1990). However, at no time has the receptor failed to internalise

Figure 1.12

General scheme of two cycles of coated vesicle formation: the endocytic cycle at the plasma membrane involving clathrin and HA-2 adaptors (▭); and second sorting cycle at the trans-golgi membrane involving clathrin and HA-1 adaptors (▬).

This diagram shows the two recycling events taking place. The first one comprises membrane traffic between the plasma membrane and the endosome as a result of recycling receptors. The second one involves (B) the traffic between the trans-golgi and the prelysosomal compartment. (Adapted from Gruenberg and Howell, 1989)


endosome
plasma membrane
trans Golgi
lysosome
TR
clathrin triskelion
M6PR + ligand
HA-1 adaptor
HA-2 adaptor
clathrin

in all these mutants, even though the actual rate is approximately 10% of that of the wild-type receptor. This implies that the mutant receptors are not excluded from the coated vesicles (Pearce and Robinson, 1990) nor LDL (Davis *et al.*, 1987) or transferrin receptor (Jing *et al.*, 1990) recycling impeded. Pearce (1985), has shown that the mannose-6-phosphate receptor (M6PR) in the presence of clathrin and adaptor can reassemble coats incorporating the receptor. Furthermore, wild-type cytoplasmic tails of M6PR compete out HA-1 and HA-2 bound to a matrix column. The mutant cytoplasmic tail of M6PR passed through the same column only competes out HA-1 adaptor. This implies that the tyrosine residue is not the only portion of the cytoplasmic tail to interact with the adaptors (Glickman *et al.*, 1989; Lobel *et al.*, 1989). However, the competition experiment with the matrix column does provide evidence for the interaction between receptor and adaptor and adaptor sorting specificity (Pearce and Robinson, 1990).

Non-clathrin coated vesicles have also been studied and found to have a similar set of proteins associated with them. However these vesicles have been implicated in the secretory pathway. One of the components of the non-Clathin vesicles β-COP (reported to have strong similaries to the adaptins) is associated with the golgi non-clathrin coated vesicles (Duden *et al.*, 1991; Serafini *et al.*, 1991)

Endosomes

The clathrin coated vesicles, once formed from the plasma membrane, travel inwards in the cell. The budding of the clathrin vesicles is a process that requires ATP (Smythe *et al.*, 1989; Heuser, 1989). Furthermore, other cellular activities may be required for budding of coated pits from the plasma membrane as illustrated by the temperature sensitive mutation in the gene *shibire* of *Drosophila* in which at the restrictive temperature (29°C) the formation of coated vesicles is inhibited (Tsuruhura *et al.*, 1990). The target of these vesicles is the endosome or prelysosomal endosome (Gruenberg and Howell, 1989). Morphological evidence (Roth and Porter, 1964; Goldstein *et al.*,1985) shows that the clathrin coated vesicles lose their coat before fusing to the endosome. Sequential studies have shown that the internalised protein appears first in the early endosome, then in the late endosome and eventually in the lysosome (Gruenberg and Howell, 1989).

The kinetics of budding of the plasma membrane, fusing to endosomes and recycling of the receptor back to the plasma membrane is very rapid ($T_{1/2}$=5min), regardless of the method used to trace the reaction (Goldstein *et al.*, 1985; Haylett and Thilo, 1986; Drayo *et al*, 1987; Swanson and Silversten, 1988). This rapid recycling

suggests that it takes place at the early endosome stage. For recycling to take place it is obvious that uncoupling of the receptor molecule and the ligand is essential. The pH of the early endosome has been reported to be more acidic than the extracellular medium, this acidity has been implicated in the uncoupling of receptor and ligand. pH sensitive fluorescent probes have been used to show that the early endosomes are the least acidic compartments compared to late endosomes (pH5.5) and lysosomes (pH4.5) (Mellman *et al.*, 1986). The acidification of the early endosome is thought to be due to a counter effect of recycling (Na^+, K^+) ATPase, whereas in the lysosome and late endosome is as a result of a vacuolar ATPase (Mellman *et al.*, 1986). Acidification is not a universal method for sorting ligands in the cell. It appears that there are acidic insensitive ligands such as transferrin and macrophage lymphocyte Fc receptor which are sorted differently from the pathway outlined above. Transferrin remains bound to the receptor and returns to the plasma membrane intact (Klausner *et al.*, 1983; Daultry-Varsat *et al.*, 1983). This may be due to acidification causing release of the iron, which in turn causes the transferrin to change comformation such that it ensures it remains bound to its receptor. The macrophage lymphocyte Fc receptor is targeted to the lysosomes and degraded (Mellman and Plutner, 1984).

As mentioned before, the sorting of receptors takes

place very quickly. The ligand can be followed from endosome with receptor to endosome minus receptor (Mueller and Hubbard, 1986; Schmid *et al.*, 1988). The movement of material between the early and late endosome depends on cytosol and ATP (Diaz *et al.*, 1989) as well as microtubules. The microtubule interactions are also required for delivery from late endosomes to lysosomes. This appears to be the reason why a decrease in fusiogenic activity is observed in *in vitro* experiments between early endosomes, late endosomes (Wolkoff *et al.*, 1984; Mueller and Hubbard, 1986) and lysosomes (Dunn *et al.*, 1980), compared to the fusiogenic activity between plasma membrane vesicles and early endosomes (using the same conditions) (Gruenberg and Howell, 1989). The lower fusiogenic activity is also probably due to larger vesicles travelling between early and late endosomes as well as the requirement for microtubules (Gruenberg *et al.*, 1989). The transfer of material from late endosomes to lysosomes depends upon the presence of lysosomes, supporting the idea that late endosomes do not mature into lysosomes (Mullock *et al.*, 1989). Intraspecific cell fusions have shown rapid exchange of lysosomal contents and membrane in a microtubule dependant manner (Ferris *et al.*, 1987; Deng and Storric, 1988) which suggests that the lysosomes have a high fusiogenic activity in the presence of microtubules.

It is interesting to note that the late endosome is

an intermediate step between the lysosome and the TGN since acid phosphatase and other lysosomal hydrolases have been found within it (Storric *et al.*, 1984, Gruenberg *et al.*, 1989). However, acid phosphatase, in contrast to typical lysosomal hydrolases, is synthesised as an integral membrane precursor and found in the late endosomal compartment (Waheed *et al.*, 1988). The acid phosphatase, as well as other hydrolases, originates from the TGN and is transported to the late endosome via the cation independent mannose-6-phosphate receptor (CI-MPR), which in turn recycles back to the TGN after delivery of the ligand (Goda and Pfeffer, 1988). Hence, in addition to the recycling of receptors from early endosome to the plasma membrane there is a second recycling process between the TGN and the late endosome (Gruenberg and Howell, 1989). The transfer from endosome to the TGN is inhibited by GTP-γ-S (a potent inhibitor of GTP binding proteins) (Goda and Pfeffer, 1988). This is the first documented evidence of GTP binding protein involvement in the endocytotic membrane trafficking (Gruenberg and Howell, 1989). The pathway is summarised in Figure 1.12.

Is membrane traffic by vesicular transport or a process of maturation?

The existence of the early and late endosomes as separate identities with specific functions has been reported, using density gradient centrifugation (Kindberg

et al., 1984; Wall and Hubbard, 1985; Branch *et al.*, 1987), immuno-isolation (Muller and Hubbard, 1986; Gruenberg *et al.*, 1989) and free-flow electrophoresis (Schmid *et al.*, 1988). The endosomal entities can be separated by their sorting functions and acidification properties, their heterogeneity being reflected by the fact that early and late endosomes differ in protein composition (Mueller and Hubbard, 1986; Schmid *et al.*, 1988). It is important to note that in only one case have intermediates between early and late endosomes been isolated. This intermediate or large vesicle has been shown to have early endosome markers and to fuse with late endosomal compartments (Gruenberg *et al.*, 1989). Schmid *et al* (1988) isolated five different fractions in which early and late endosomal compartments were purified in one fraction each. The other fractions were reported to contain a mixture of both early and late endosomes, however the authors did not show supporting data. Therefore those fractions may contain endosomes in the process of maturation. Moreover, it has been recently shown that early and late endosomes are accessible to endocytotic vesicles with the same kinetics (Stoorvogel *et al.*, 1989; 1991). Stoorvogel *et al.* (1991) demonstrated by immunogold labelling that early as well as late endosomes have transferrin receptors, but that the transferrin receptors are found in lower concentrations in the late endosomes. The authors also report an increase in M6PR from early to late endosomes thus illustrating

that early endosomes mature into late endosomes and indicating the involvement of the TGN vesicular transport in this process. This in turn may explain the unique protein content observed in the late endosomes as reported by Schmid *et al.* (1988). Finally, Stoorvogel *et al.* (1991), suggest that the decrease in fusiogenic activity between early and late endosomes *in vitro* (as described above) is a result of limited vesicular transport between early and late endosomal compartments rather than the lack of microtubules. It therefore appears that membrane trafficking in endocytosis follows the maturation model which is dependant on vesicular transport at various stages of compartment maturation.

Can Receptor mediated endocytosis account for everything that goes into the cell?

The entrance of foreign proteins into cells has always been assumed to be by receptor mediated endocytosis and therefore this process has been studied in great detail. There is controversy as to whether coated pits account for everything pinocytosed. Some studies suggest coated pit mediated endocytosis accounts for all uptake (Marsh and Helenius, 1980), whereas other studies provide evidence for alternative pathways (Hopkins *et al.*, 1985). More recently, West *et al.* (1989), reported that upon epidermal growth factor (EGF) stimulation of human epidermoid carcinoma A431 cells, there was a marked

increase in pinocytosis, which was not due to an increase in the rate of coated pit formation because transferrin receptors were internalised at the same rate as in non-stimulated cells. More importantly, the increase in pinocytosis was inhibited specifically by Na^+/H^+ exchange activity being blocked, whereas coated pit formation remained unaffected. Other observations have showed that pinocytosis was inhibited by increased Ca^{++} or cAMP levels in the cytoplasm of human epidermal carcinoma KB cells (Miyata *et al.*, 1989). It is important to note that this pathway is not utilised by EGF for its own uptake, nor by transferrin (West *et al.*, 1989). Thus, however there seem to be other routes of pinocytosis which are independant of classical coated pit receptor mediated endocytosis, and may be the direct effect of stimulation by a hormone or growth factor.

How does vitellogenin enter the oocytes of insects?

A. Morphological studies

At the morphological level, Roth and Porter (1964) showed the presence of coated pits in the membrane of mosquito oocytes. Those experiments suggest that vitellogenins or yolk proteins entered the oocyte by receptor mediated endocytosis; Mahowald (1972), reported that receptor mediated endocytosis occured in oocytes of *D. melanogaster*. Confirmation that yolk enters the

oocyte in mosquitos by receptor mediated endocytosis came with immunolocalisation of yolk in receptors concentrated in the coated pits. The coated pits were internalised and then lost their coats. The yolk dissociated from the receptors soon after the vesicles fused into the endosomes, where yolk coalesces to form the transitory yolk body (TBY). The TBY acts as a sink for further yolk deposition and probably recycles receptors back to the membrane. The yolk crystallises in the yolk bodies by which time they are termed mature yolk bodies (MYB) or alpha spheres. The pathway lacks lysosomal hydrolases, and hence the nature of the MYB is equivalent to a prelysosomal endosome specialised for long term storage of yolk protein (Raikhel, 1984). Raikhel and Lea (1986), utilised horse-radish peroxidase (HRP) as a marker for fluid phase pinocytosis. The HRP was taken up but only yolk proteins and their receptors were specifically accumulated. The HRP was seen to segregate in the TYB, and in the MYB formed inclusion bodies whose fate may be the lysosomes (Raikhel and Lea, 1986). The exposure of *Xenopus* oocytes to HRP also resulted in the segregation of HRP from the yolk (Opresko *et al.*, 1980). The HRP is expected to behave as a marker for fluid phase pinocytosis, however it was reported to be saturable (Raikhel and Lea, 1986) and to bind to the oocyte membrane (Bock, 1972; Hezo and Regio, 1980; Raikhel and Lea, 1986); thereby, also giving an index of endocytosis.

In *Drosophila*, the utilisation of mutants has provided support for morphological evidence which indicated that yolk protein uptake is mediated by receptors. In the mutant *yolkless* (*yl*$^{-}$), lack of yolk in the oocyte is due to the reduced number of coated pits in the oocyte membrane (Di Mario and Mahowald, 1987). Alleles of *yolkless* were used in a morphological study of the phenotype. A direct correlation between sterility, yolk protein deposition and coated pit number was made, indicating that the entry of YPs into the oocyte is dependant upon receptor mediated endocytosis (Di Mario and Mahowald, 1987).

The other mutant studied is a temperature sensitive mutation in a gene called *shibire*. The mutant shows a build up of coated pits at the restrictive temperature (29°C) (Tsuruhara *et al.*, 1990). If the flies are kept at restrictive temperature, accumulation of coated pits on the surface of the oocyte is observed and this results in lack of yolk deposition in the oocyte. Once the flies are placed at 18°C, the coated vesicles form and yolk deposition is observed (Tsuruhara *et al.*, 1990). Both mutants therefore illustrate the dependance of yolk deposition on receptor mediated endocytosis.

B. Biochemical studies

The mutants in *Drosophila*, in addition to the

detailed morphological studies in mosquito, support receptor mediated endocytosis of yolk proteins as the mechanism involved in the deposition of yolk in the oocyte. To date, the receptors involved in the interactions with the yolk proteins have not been isolated; however, in many insects biochemical studies have shown the existance of a specific and saturable binding activity. Experiments to show the specific binding of vitellogenin to putative receptors in the oocyte membrane have been performed in two different ways. The first of these involves the isolation of oocyte membranes and testing for the binding of vitellogenins. This method was used with *Nauphoeta cinerea* (Konig and Lanzrein, 1985), *Blatella germanica* (Konig *et al.*, 1988a), *Manduca sexta* (Osir and law, 1986) and *Locusta migratoria* (Rohtkasten and Ferenz, 1986) all of which were described to have specific and saturable binding activities. Osir *et al.* (1986), reported that *Manduca* unglycosylated vitellogenin interacted with the oocyte membranes from *Manduca* to the same extent as native glycosylated vitellogenins. The unglycosylated vitellogenins of *Locusta* and *Blatella* showed a reduction in binding specificity to their respective oocyte membranes (Konig *et al.*, 1988a). Konig *et al.* (1988b), showed that binding of vitellogenins to oocytes membranes involves two different binding sites, both of which are affected by calcium concentration, where a low intracellular calcium level results in reduced vitellogenin endocytosis. The

addition of EGTA (a divalent ion chelating agent) or heat denaturation of the vitellogenin had deleterious effects on the interaction of vitellogenins and their receptors in *Locusta* (Rohrkasten and Ferenz, 1986). Furthermore, the specific binding of vitellogenins to oocyte membranes in *Manduca* was also found to be affected by calcium ion depletion (Osir and Law, 1986).

The second biochemical approach involved using intact ovaries from insects and determining if a specific affinity to the vitellogenins was observed. Purified vitellogenins from mosquito (*Aedes aegypti*) (Noah-Koller *et al.*, 1989) and *Hyalophora cecropia* (Kulakosky and Telfer, 1988) were used to assay binding to intact ovaries from their respective species. The uptake of vitellogenins was found to be saturable and to have characteristics of receptor mediated endocytosis (Kulakosky and Telfer, 1987; Noah-Koller *et al.*, 1989). In the case of *Hyalophora*, arylphorin and flavoprotein (haemolymph proteins not internalised by oocytes) were found concentrated in the basement lamina of the follicle, and very little of either arylphorin or flavoprotein was actually detected in the yolk bodies (Kulokosky and Telfer, 1987). In the intact follicles of *Nauphoeta* Calcium was shown to be necessary for the uptake of vitellogenins (Kindel *et al.*, 1990). In *D. melanogaster* the uptake of YPs is pH dependant *in vitro* where a pH of 6.5 or lower decreases the rate of uptake and pH6.8 or

greater increases the rate of uptake (Di Mario and Mahowald, 1986). The pH also affects the uptake of vitellogenin in *Nauphoeta* (Konig and Lanzrein, 1985).

The evidence amassed to date indicates that the uptake of vitellogenins or yolk proteins in insects occurs via receptor mediated endocytosis. This process has been shown in a variety of insect systems to involve a receptor in the oocyte membrane which is internalised once the yolk protein interacts with it. The interaction between yolk protein and receptor is specific, selective, ion and pH dependant. Moreover, the injection of foreign proteins (eg. horse-radish peroxidase) into mosquitos or *Xenopus* has shown that the accumulation of yolk proteins in the oocyte storage granules is specific for yolk. Thusfar, the receptors involved in binding the yolk proteins have not been purified nor cloned, however to further understand the mechanism of the very specific internalisation of the yolk proteins into the oocyte it is essential to isolate the receptor.

Drosophila Evolution

The purpose of this section of the introduction is to give an overview of the Drosophilidea and some of the Muscamorpha. It also will introduce studies carried out with other proteins pertaining to the species used in the experiments described in this thesis. In no way is this section intended as a complete review of the systematics of the Muscamorpha since there is still a great deal of controversy in the area and, more importantly, because the data derived in this thesis is not used to formulate a phylogenetic relationship.

Taxonomy of the Diptera

The Diptera encompass the suborders *Nematocera* and *Brachycera*. *Nematocera* are frail flies with long segmented antennae and aquatic larvae (eg. midges and mosquitos). The *Brachycera* include three infra-orders, namely Tabanomorpha, Asilomorpha and Muscamorpha, where the Tabanomorpha and Asilomorpha have short antenna, reduced larval head and free pupal stage (eg. horseflies and beeflies). The Muscamorpha have two divisions, the *Aschiza* (eg. hoverflies and scutterflies), and *Schizophora*. The Muscamorpha are considered the most advanced Diptera and within them the *Schizophora*. The division *Schizophora* have two series, the Acalyptratae and the Calyptratae, the latter are recognised by a complete

suture across the dorsal thorax, a cleft in the second antenna and the presence of developed flat-like appendages (called calypters) in the base of the wing. The Calyptratae include the Muscidae (eg. housefly), Calliphoridae (eg. blowfly) and Sarcophagidae (eg. fleshfly) among others. The Acalyptratae include the superfamily Ephydroidae within which the Drosophilidae are found. The Acalyptratae have been avoided by entomologists because of the large number of small species that lack economic value and are regarded as confusing and featureless. The closest family to the Drosophilidae are the Camillidae, found in the burrows of small mammals; the Ephydridae, small black flies associated with water; the Chloropidae, blackish flies commonly found in grasses; the Cryptochetidae, whose larvae are internal parasites in coccids; and others, not well studied, such as Diastatidae, Campichoetidae, Millichiidae and Tethnidae (Wheeler, 1981; Ashburner, 1989). The Drosophilidae family has over 3000 different species, most of which are not well known (Ashburner, 1989). It is thought that the drosophilids were probably a saprophytic group, specialised in fermenting substances (Throckmorton, 1975). However, these flies have exploited other ecological habitats ranging from leaf litter to fruits, flowers, fungi and running water (Ashburner, 1981; 1989).

The Drosophilidae family consists of two sub-families namely the steganinae (containing 444 species) and the

drosophilinae (containing 2575 species) which are not distinguishable by a single morphological character. The *Drosophila* species are grouped under the sub-family drosophilinae in the genus *Drosophila*, which encompasses most of the species used by experimental biologists. The *Drosophila* genus includes the subgenus *Antopocerus*, *Dorsilopha*, *Drosophila*, *Hirtodrosophila*, *Scaptodrosophila* and *Sophophora*, among others. The *Antopocerus* subgenus were once classified in their own genus, and are Hawaiian in origin. The *Dorsilopha* subgenus members (ie. *D. busckii*) are easily cultured. The *Drosophila* subgenus is the largest of the sub-family. Throckmorton (1975), divided the subgenus into three radiations namely *virilis-repleta*, *immigrans-tripunctata* and *hirtodrosophila*. There are many species groups, amongst which the *funebris* species group (eg. *D. funebris*), *repleta* species subgroup (eg. *D. hydei*, *D. arizonensis*), *immigrans* species group (eg. *D. nasuta*), and *virilis* species group (eg. *D. virilis*, *D. littoralis*, *D. lummnei*) are classified. Only one fossil has been found for the subgenus *Hirtodrosophila*, in the Miocene amber of the Dominican Republic. The *Scaptodrosophila* subgenus is a dominant feature of Australian fauna, their eggs generally have 6 to 8 chorionic filaments (eg. *D. lebanonensis*). The *Sophophora* subgenus includes the *melanogaster* species group, *obscura* species group and *willstoni* species group among others, The *melanogaster* species group (eg. *D.*

melanogaster, *D. mauritiana*, *D. yakuba*) contains 150 species and is divided in three lineages (*ananassae*, *montium* and *melanogaster-takahaskii-suzukii-eugracilis-ficusphila-elegans*) whose origin is mostly oriental. The *obscura* species group (eg. *D. azteka*, *D. pseudoobscura*) has colonising tendencies whereby *D. subobscura* has been found in Chile and North America. Other genera include *Grimshawomyia* (endemic to Hawaii), *Miomyia* (known only from fossils in Dominican Republic Miocene amber). *Mycodrosophial* (feed on fungi) and *Scaptomyza* (may be of Hawaiian origin but are widely distributed).

The Evolutionary time scale covering the *Drosophila* species and some of the Calyptratae.

The time of origin of a species is always difficult to assess as it depends on the fossil evidence in existence. In the case of *Drosophila* the paleontological evidence comes from a very limited number of fossils, originating from amber deposits. The data obtained is constantly being reviewed and used to complement the information gathered using other methods, in recent years, particularly molecular studies on protein and gene divergence. A phylogenetic tree derived from molecular evidence is only a framework, and hence a tool to further understanding of the mechanisms of evolution (Beverley and Wilson, 1982).

The value of molecular studies can only be maximised by understanding the limitations of a particular method; for instance, using electrophoretic variants of polypeptides can be useful in assessing differences between individuals within a species or species group, such as used by Lineruth and Lindberg (1988), when studying electrophoretic variants of follicle specific proteins in *D. melanogaster* natural populations. If a wider analysis is required to estimate the degree of sequence divergence among distant taxa, different methods such as amino acid sequence determination of the targeted protein should be used (Beverly and Wilson, 1982). Other methods such as DNA-DNA reassociation studies have been used to determine rates of nucleotide substitution; however this method is limited in that it only gives a rate for the conserved portions of DNA within the whole genome of the species being compared (Sullivan *et al.*, 1990).

The use of protein sequences to construct evolutionary trees has been successful in vertebrates (using cytochrome C) and has yielded trees that branch in approximate agreement with trees constructed using anatomical features (Wilson *et al.*, 1977). On the other hand, studies comparing cytochrome C protein sequences in insects appear to generate an evolutionary tree that does not compare favourably with the classical phylogenetic relationships (Inoue *et al.*, 1985). Beverley and Wilson (1982;

1984) use the microcomplement fixation technique to study immunological identities of larval serum proteins (LSPs) in different species. The success of this technique depends on the choice of the protein, in that the protein should be present in the insects being studied, the chosen protein should be functionally significant, and hence the protein should ideally be evolving slowly (Collier and MacIntyre, 1977). The advantage of this method (as used by Beverley and Wilson (1982; 1984), where the antisera raised against the LSPs of the different species were pooled together) is that polymorphisms within a species are undetected and therefore will not add to the background. The results obtained using the micro-complementation fixation technique showed that the Hawaiian *Drosophila* species do not have an accelerated rate of protein evolution (Beverley and Wilson, 1984; 1985). The results are summarised in Table 1.2 and appear to correlate well with those obtained using alpha-Glycerophosphate dehydrogenase (data re-evaluated by Beverley and Wilson, 1984 from Collier and MacIntyre, 1977) and Alcohol dehydrogenase (*Adh*) gene coding sequence comparisons for the *Drosophila* species (Sullivan *et al.*, 1990; Villaroya and Juan, 1991) and can be summarised as follows. The Tephritidae, Otitidae, Sarcophagidae, Calliphoridae, Muscidae and Drosophilidae families arose in the Cretaceous period (68-130 Million years (MYR)); however, the divergence within the Drosophilidae began in the Cretaceous and continued in the Cenozoic era. Thus,

Table 1.2 Molecular estimates of divergence times in higher Diptera and Drosophila

Lineages compared	Time (MYR)
Tephiritidae-Drosophilidea	123
Calliphoridae-Drosophilidae	99
Sophophora-Scaptodrosophila	62
Virilis-Repleta	42
Obscura-Melanogaster	46
Robusta-Repleta	35
Willistoni-Melanogaster	53
D.melanogaster-D.simulans	0-5

Based on Beverley and Wilson, 1984.

the *Drosophila* radiation began by about 62 MYR ago and the *melanogaster* species group appeared at least 26 MYR ago (Beverley and Wilson, 1984).

The fossil evidence does indicate that the Cyclorrhapha is not likely to be older than 135 MYR and the divergence of Calyptratae and Acalyptratae happened at least 70 MYR ago (there is only one fossil available of a pupa assigned to a Calyptratae mucoid family, possibly Calliphoridae (McAlpine, 1970)). The Drosophilidae have been placed between 40 MYR and not beyond 80 MYR ago. The upper limit has been derived from the fact that Australia and New Zealand land masses separated about 80 MYR ago and Australia has a distinct endemic drosophilid fauna whereas New Zealand does not (Beverley and Wilson, 1984). Throckmorton (1975) suggested (based mostly on biogeographical data) that the Drosophilidae originated in the Eocene (38 to 55 MYR ago) or before, with the major radiations having occurred by mid-Oligocene (about 30MYR ago). It is therefore interesting that the data obtained by fossil evidence and that from the protein comparisons appears to correlate closely.

The determination of species divergence may be misleading if the rates of protein evolution are assumed to be constant (La torre *et al.*, 1986). Beverley and Wilson, (1984) illustrated the case by showing that the LSP-1 rate of divergence is slower compared to that of

LSP-2. Bewley *et al.* (1989) compared the sn-glycerol-3-phophate dehydrogenase, in which a high degree of amino acid sequence identity has been maintained, from *D. melanogaster*, *D. virilis*, rabbit and mouse. They concluded that the rate of evolutionary divergence is very slow, and closely similar to that of LSP-1. A more dramatic difference (one order of magnitude) was reported when comparing the rate of evolution of Arginine Kinase (using the micro-complement fixation assay) with that of the Larval Serum Proteins (Table 1.3) (Collier, 1990). Furthermore, Martin et al (1988) found that different genes evolve with different rates and that a dramatic variation in evolutionary rates is seen near the 68C glue protein cluster in different *Drosophila* species. The variation in evolutionary rates from different subregions is dependent upon different local factors that influence the underlying mutation rate; the turnover of DNA (including gene conversion, unequal crossing over and slippage during DNA replication), and the selective forces (by removing and fixing mutations) (Martin and Meyerowitz, 1988b)

Genes, Proteins and Evolution

There are a few examples of amino acid sequence data from specific proteins in different *Drosophila* and sometimes other insects which have been used to study evolutionary divergence. In *Drosophila* most of the

Table 1.3 Evolutionary rates of seven Drosophila proteins

Protein	UEP (MYR)
Acid phosphatase	3
Alcohol dehydrogenase	3.5
Larval Serum Proteins	4
Superoxide desmutase	11
6-Phosphogluconate dehydrogenase	25
Glycerophosphate dehydrogenase	36
Arginine kinase	59

Based on Collier, 1990.

studies used DNA or amino acid sequences generated from *D. melanogaster* and *D. virilis*. It appears that the 60 MYR separating these two species (according to Beverley and Wilson, 1982; 1984) are sufficient to identify important regulatory sequences for gene expression. However, the same time span does not appear to be enough to observe amino acid sequence divergence which may suggest domains of functional importance. In this section a discussion emphasising insect sequence divergence will consider recent examples from genes whose sequence determination have illustrated different aspects of divergence.

The evolution of organisms may be achieved in two main ways, namely by varying the sequence of genes or changing the pattern of gene expression. In the first instance sequence variability may result in the alteration of the protein sequence and thus the specificity of function. In the second case, the alteration of regulatory processes may result in different temporal, spatial or sex activation thus changing the pattern of expression. Studies of genes have been carried out in which both modes are represented.

The chorion protein genes have been studied in detail because they offer an opportunity to study chromosome replicons in eukaryotes as well as developmental control of genes. The chorion genes in *D. melanogaster* are present in two clusters which are amplified selectively

(situated at 7F1 in the X chromosome and 66D11-15 in the third chromosome). The autosomal amplification and characteristic order of expression of the chorion genes occurs in four different *Drosophila* species, namely *D. melanogaster*, *D. subobscura*, *D. virilis* and *D. grimshawi* (Martinez-Cruzado *et al.*, 1988; Fenerjian *et al.*, 1989; Swimmer *et al.*, 1990). The DNA sequences of the chorion genes from these four species also reveal striking similarities (Orr-weaver, 1991); although there is extensive diversification in the intron and some in the coding sequences, the sequences of importance for regulation and amplification of the chorion genes appear to be conserved (Martinez-Cruzado *et al.*, 1988; Fenerjian *et al.*, 1989; Swimmer *et al.*, 1990). An illustration of the conserved sequences is a 71 bp fragment, ACE3 (or Amplification control element of the third chromosome), in *D. grimshawi* which was shown to functionally substitute for the equivalent element in *D. melanogaster* via P-element transformation (Swimmer *et al.*, 1990). Moreover, the hexamer TCACGT, a promotor element, has been found to be located in position -60 bp from the transcription start site of all major chorion genes in *D. melanogaster*, the known autosomal chorion genes of *D. subobscura*, *D. virilis*, *D. grimshawi* (Martinez-Cruzado *et al.*, 1988; Fenerijian *et al.*, 1989; Swimmer *et al.*, 1990) and the Mediterranean fruit fly (*Ceratitis capitana*) (Konsolaki *et al.*, 1990). The Mediterranean fruit fly has also been shown to amplify chorion sequences in the egg

chambers (Konsolaki *et al.*, 1990; Tolias *et al.*, 1990); however, the silkworm *Bombyx mori* uses a different approach by having multiple copies of each gene in the genome (Spoerel *et al.*, 1989). It was therefore of interest that the *B. mori* chorion genes are transcribed with the correct developmental pattern after being introduced into the *Drosophila melanogaster* genome via P-element mediated transformation.

Amino acid sequence similarity was observed when the *D. melanogaster* (Burtis *et al.*, 1990), *D. virilis* and *D. pseudoobscura* (Jones *et al.*, 1991) DNA sequence was determined for the ecdysone inducible gene, E74A. The C-terminal end sequence is 95% identical in all three genes (Jones *et al.*, 1991) but more importantly shows identity to the DNA binding domain found in the *ets* proto-oncogene superfamily (Burtis *et al.*, 1990). This region has since been shown to be involved in DNA binding as demonstrated by its affinity to early and late ecdysone induced polytene chromosome puffs (Urness and Thummel, 1990). The N-terminal end of the E74A protein is less similar but maintains the acidic composition (Burtis *et al.*, 1990; Jones *et al.*, 1991) which is correlated to transcriptional activator domain functions (Ptashne, 1988).

Evolutionary approaches have been taken to try to define amino acid sequences which may be responsible for a

function not present in a homologous protein of a different species, such as is the case for xanthine dehydrogenase (Xdh). In *Drosophila melanogaster* xanthine dehydrogenase is required for the oxidation of pterin into xanthopterin which is the precursor of eye pigment (Dulton and Chounick, 1988). In *Calliphora* and other Diptera the eye pigment consists of xanthomatin which is synthesised via a different pathway. Therefore, the molecular features required for eye pigment synthesis by *Drosophila Xdh* may not be present in the *Calliphora Xdh* (Houde *et al.*, 1989). The DNA sequence determination of the *Calliphora Xdh* (Houde *et al*, 1989) and the *D.psuedoobscura Xdh* (Riley, 1989) and subsequent translation to amino acid sequence, showed that *Calliphora* had 76% identity to *D. melanogaster* Xdh protein sequence despite the divergence in nucleotide sequence. The *D. pseudoobscura* Xdh has 141 amino acid substitutions when compared to the *D. melanogaster* Xdh (Riley, 1989) of which 102 were also found to be substituted when comparing *Calliphora* Xdh to *D. melanogaster* Xdh. Thus the differences observed do not appear to identify sequences which may represent functional differences (Houde *et al.*, 1989). It must be noted that the lack of the same eye pigment synthesis mechanism in *Calliphora* and *Drosophila* does not necessarily indicate a difference in Xdh protein activity. This difference may be a result of the *Calliphora Xdh* having an altered pattern of expression.

A more clear case of altered expression through evolution has been observed with the cloning of *hunchback* in both *D. melanogaster* (Tautz *et al.*, 1987) and *D. virilis* (Treier *et al.*, 1989). Although the amino acid sequence is said to have diverged by 20%, the *D. virilis* protein is recognised by antibodies raised against the *D. melanogaster hunchback* protein. In both species, the pattern of expression is the result of two transcripts (P1 and P2), whereby the P2 transcript gives rise to the same primary pattern of expression in the embryos of both species. The P1 transcript shows a secondary pattern of expression which differs between *D. melanogaster* and *D. virilis*. The alteration in the secondary pattern of expression appears to be due to a difference in the regulation of the P1 transcript in each species, which in turn may add to the divergence of morphological characters between the two species. Therefore, the modification of the expression pattern of a gene may be considered a good target for evolutionary change (Treier *et al.*, 1989). The degree of divergence between the *hunchback* proteins of both *D. melanogaster* and *D. virilis* is consistent with that found in engrailed (Kassis *et al.*, 1986) and the *mastermind* locus (Newfield *et al.*, 1991).

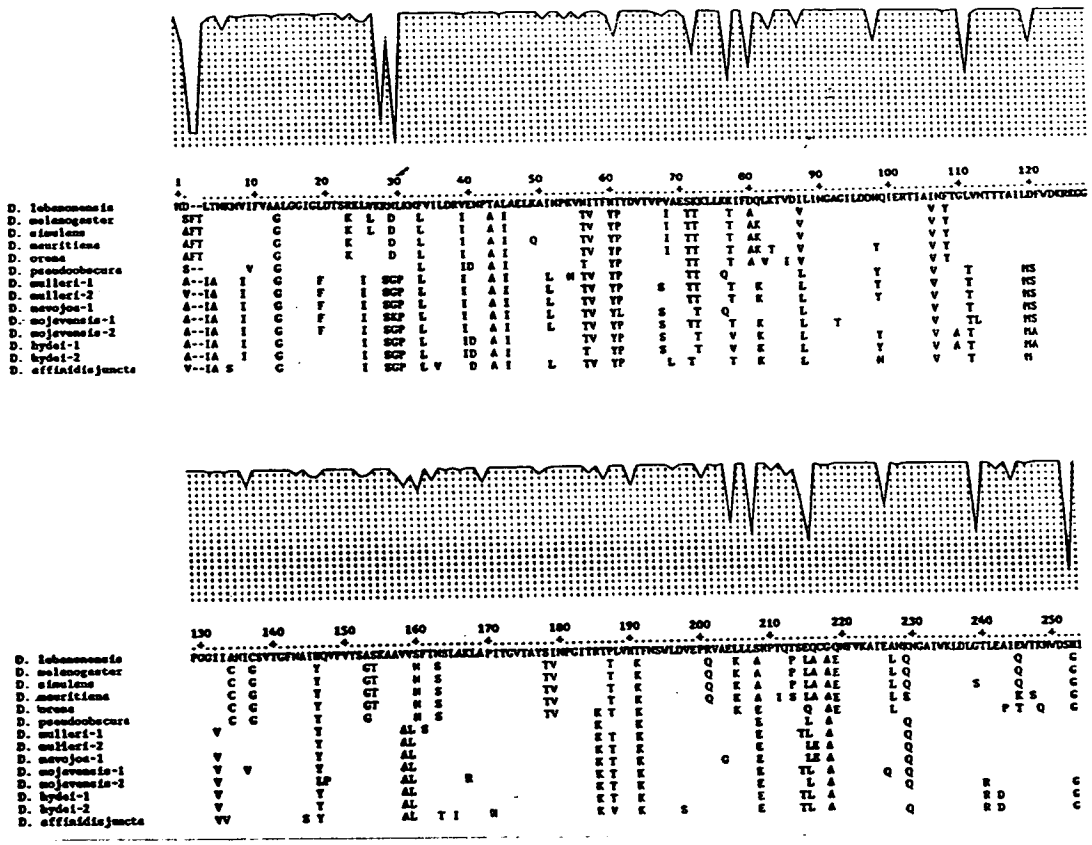
The difference in spatial and temporal expression has been widely observed in the expression of the alcohol dehydrogenase (*Adh*) gene in *Drosophila* (reviewed by Sullivan *et al.*, 1990). However despite this altered

expression, the Adh amino acid sequences from the species subgroups *melanogaster*, *obscura*, hawaiian picture-wing, *mulleri* and *hydei* show 170 amino acids out of 254 share identity in all species (67% average amino acid identity). The rate of amino acid substitution has been constant along the length of the protein sequence (hence no gradient in amino acid sequence identity). (Figure 1.13). Predictions of hydrophobicity and alpha helix formation, based on the deduced amino acid sequences of *D. melanogaster* Adh and *D. mojavensis* Adh1, suggests a similar pattern of folding for the two proteins. (Sullivan *et al.*, 1990; Villaroya and Juan, 1991). This is also the case for xanthine dehydrogenase in *Calliphora* and *Drosophila* (Houde *et al* ., 1989),

Regardless of the sequence and expression divergence observed in some genes through *Drosophila* evolution it appears that the molecular structure of these genes is conserved. Such is the case illustrated by the conservation of the nested gene structure of the pupal cuticle protein (PCP) within the *Gart* locus of *Drosophila* (Henikoff and Eghtedarzadeh, 1987). The PCP gene is thought to have arisen from the larval cuticle protein (LCP) genes via a duplication event. The divergence of the two genes (ie. PCP compared to LCP) is reported to have occurred 70 MYR ago (Moriyama and Gorobori, 1989) and since PCP and LCP genes in *D. pseudoobscura* have the same structure (Henikoff and Eghtedarzadeh, 1987) it is

Figure 1.13

Alignment of *Drosophila* ADH protein sequences referred to *D. lebanonensis* showing the dissimilarities. Figures of merit for each site in the alignment, where the higher the dotted line the more conserved is the amino acid residue in that position.



conceivable that the nested gene structure arose at the time of the duplication event (Moriyama and Gorobori, 1989). The conservation of the genomic structure has also been observed for hunchback (Treier *et al.*, 1989) and chorion genes (Orr-Weaver *et al.*, 1991). However, this is not true for all genes as shown by *Adh* where duplication events have given rise to two *Adh* genes in *D. mojavensis* (Sullivan *et al.*, 1990).

It is interesting to note that the non-coding regions (ie. upstream, downstream and intron) of the genes described thus far have diverged dramatically. The only cases in which these show a degree of conservation is when regulatory sequences (ie. TCACGT hexamer in the chorion genes) or putative regulatory sequences have been found. This is also true in the introns of *engrailed* where sequence conservation has been observed and thought to be of importance in gene regulation (Kassis *et al.*, 1986).

The analysis of gene sequences in a different species has provided in many cases a hold on amino acid sequences that may be of importance in protein function (eg. DNA binding finger in *hunchback* (Treier *et al.*, 1989)) and nucleotide sequences that may be of regulatory importance but which by and large remain to be investigated. This therefore, is another purpose of investigating sequences from different species beside studying sequence divergence and rates of evolution. The description of the chorion

genes, alcohol dehydrogenase, xanthine dehydrogenase, pupal cuticle proteins, ecdysone inducible gene E74A and hunchback in different systems illustrates the modes of divergence possible which may act on protein sequence or on alteration of the developmental expression of a protein. It is important to note therefore that the use of evolutionary approaches to find sequences involved in regulation of gene expression or protein domains with an important function have been successful and may yet provide many other clues to the mechanisms of gene evolution.

The Vitellogenins and Their Evolution

The divergence or evolution of yolk proteins in *Drosophila* has been studied at a very superficial level. Most of the research on YPs has been directed at the investigation of *yp* gene regulation using deletion studies rather than evolutionary studies. The evolutionary studies carried out thus far have involved investigating egg formation after ovary transplants have been performed from one *Drosphila* species to another (Srdic *et al.*, 1978; Bownes, 1980; Lamnissoa and Zouros, 1990). In the case of Srdic *et al.* (1978) and Bownes (1980), the studies were carried out to assess the relatedness of the yolk proteins and to determine if foreign YPs could functionally substitute for the native YPs. These studies lend support to ovarian synthesis of YPs, and to the fact that other

factors such as hormonal levels in the host flies may be important for ovarian development and therefore YP uptake. Finally, the YPs were shown (in the cases where species incompatibilities were minimal) to have some degree of similarity and complementary characters. In the study by Lamnissou and Zouros (1990), the ovarian transplants were used to assess the degree of evolutionary relatedness or compatibility between the different *Drosophila* species. The experiments were based on the premise that in order to observe YP uptake by transplanted ovaries there were more requirements than the receptor recognition of the YPs, as was demonstrated by Srdic *et al.* (1979) and Bownes (1980). Lamnissou and Zouros (1990), concluded that their transplant results were in good agreement with the intra-group phylogenetic relationship established by other criteria (ie. fossil evidence). However, these studies are qualitative and development of a quantitative assay would aid the assessment of taxonomic relationships as well as to provide a measure of evolutionary relatedness or compatibility between different species.

The vitellogenins of *Xenopus*, chicken and *C. elegans* have been studied in detail and shown to be related. These, however do not appear to share any similarity with the YPs of *Drosophila*. The *yp* gene structure between *Xenopus* and chickens is well conserved since both genes have 35 exons which appear to have conserved sites. *C. elegans* has only 5 exons but their boundaries are also

well conserved when compared to *Xenopus* and/or chicken *yp* genes (Figure 1.14A) (Whali, 1988; Bryne *et al.*, 1989). A comparison of the vitellogenin polypeptide sequences from chicken, *Xenopus* and *C. elegans* vit-2, vit-5 and vit-6 showed that a 20 amino acid segment was highly conserved and may play a role in oocyte receptor recognition (Figure 1.14B) (Spieth *et al.*, 1991). Moreover, Spieth *et al.* (1991) found conserved cysteine residues throughout the YP alignment, suggesting the maintenance of vitellogenin structure over a long evolutionary distance. The apparent independent evolution of the Dipteran *yp* genes to that of other insects, nematodes and vertebrates does not appear to be unique to the YPs. The Adh proteins of *Drosophila* are also strikingly different to the analogous proteins of yeast, higher plants or animals (Table 1.4) (Place *et al.*, 1986).

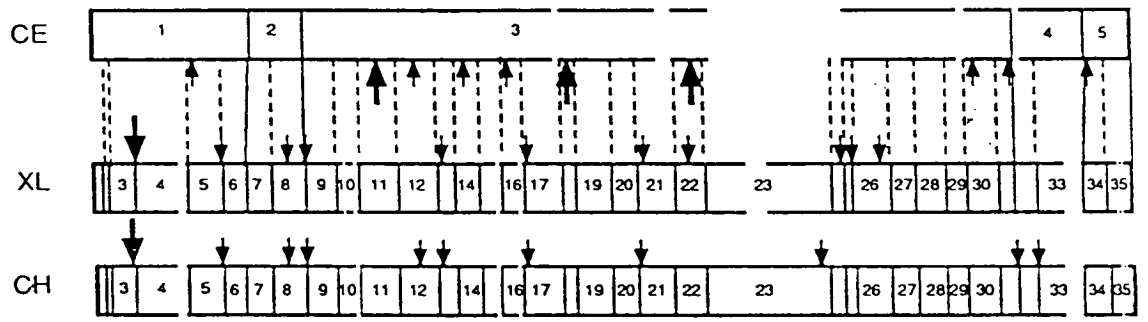
The YPs of *Drosophila* (Kozma and Bownes, 1986) and some of the other Diptera, namely *Protophormia terrae-novae*, *Sarcophaga bullata*, *Calliphora erythrocephala*, *Lucilia servicata* (Huybrechts and De loof, 1982), *Ceratitis capitana* (Rina and Mintzas, 1987) and *Dacus olea* (Zongza and Dimitriatis, 1988) are within the same molecular weight range (44 kD to 50 kD). The yolk proteins of the *Drosophila* siblings, *D. virilis* and *D. funebris* have been shown to cross react with antisera raised against the YPs of *D. melanogaster* (Kozma and Bownes, 1986a) and their genes cross-hybridise with the *yp*

Figure 1.14

A.Schematic representation of the exon organisation of the vitellogenin genes between of chicken, *Xenopus* and *C. elegans*. Gaps of one and two amino acids (small arrows) and gaps of three and five amino acids (large arrows) are indicated; longer gaps are indicated by interruption of the solid lines in the scheme. The dashed lines compare exon junctions in the *Xenopus* gene to the *C. elegans* gene (Adapted from Wahli, 1988).

B.Amino acid sequence of the highly conserved fragment between the vitellogenins of Chicken, *Xenopus* and *C. elegans* (Adapted from Spieth *et al.*, 1991).

A



B

Gene	Position	Consensus amino acid																		
		Y	E	K	X	L	A	L	K	T	L	G	N	A	G	X	D	X	S	V
vit-2	506	–	–	–	I	–	–	–	–	–	I	–	–	–	–	L	–	I	–	–
vit-5	505	–	–	–	V	–	–	–	–	–	–	–	–	–	–	I	–	L	–	–
vit-6	534	–	–	–	V	–	–	–	–	–	–	A	–	–	–	L	–	L	–	–
Xenopus A2	474	E	D	I	A	–	–	–	–	A	–	–	–	–	–	Q	P	E	–	I
Chick	474	D	K	M	K	–	–	–	–	C	I	–	–	M	–	E	P	A	–	L

Table 1.4 Comparison of Drosophila, Plant and Animal alcohol dehydrogenase proteins

Parameter	Drosophila Adh	Plant and Animal Adh
Molecular Wt.	55,644	80,000-140,000
Subunits	2	2-4
Residues	255	347-374
Metal content	None	2-4 atoms per subunit

Based on Place et al., 1986

genes of *D. melanogaster* (Kozma and Bownes, 1986b; R. Kozma Ph.D Thesis, 1986). At the structural level conservation of the *yp* genes is also observed in *Drosophila* as illustrated by the *yp* gene organisation of *D. grimshawi* and *D. melanogaster* (Hatzopoulus and Kambysellis, 1987). The mapping of the *yp* genes in polytene chromosomes of *Drosophila* sibling species shows their cytogenic conservation (Kozma and Bownes, 1986). The cytogenic conservation and the maintenance of the developmental expression pattern is also observed in the evolution of the chorion genes (Orr-Weaver *et al*, 1991), and the genes encoding the pupal cuticle protein (Moriyama and Gorobori, 1989) and the E74A protein (Thummel *et al*., 1990).

The *yp* genes in the *melanogaster* species subgroup appear to be well conserved. In the wider evolutionary spectrum the *yp* genes are expressed with the same developmental stage, tissue specificity and ecdysone inducibility in males. Furthermore, the yolk proteins are secreted from the fat body and ovarian follicle cells and accumulated specifically in the oocyte. Finally, there is evidence for a novel carrier function for YPs since *D. melanogaster* YPs have been shown to be carriers of ecdysteroids and include a strong sequence similarity with the lipid binding site of vertebrate lipases. Since other summary proteins subject to functional constraints (ie. chorion proteins, engrailed, hunchback, E74A) have been

shown to be well conserved through *Drosophila* evolution, and the yolk proteins are also under a number of constraints (ie, sites of synthesis and secretion, specific uptake and sorting and ecdysone carrier function) a high degree of conservation of *yp* gene and protein structure through *Drosophila* and Dipteran evolution would be predicted.

Aims of the Project

The yolk proteins enter the oocyte via receptor mediated endocytosis and are stored until further use during embryogenesis. The aim of this project was to identify the domains of the YPs required for their uptake into oocytes of *Drosophila melanogaster*. An evolutionary approach was used in order to identify a *Drosophila* species whose YPs were taken up by developing oocytes of *D. melanogaster* but which was distantly related to *D. melanogaster*. This approach ensures that the YP uptake function is maintained whilst the sequence divergence between the species tested and *D. melanogaster* is likely to increase as more distantly related species to *D. melanogaster* are tested. Once a candidate was identified the aim was to isolate and determine the DNA sequence of its *yp* gene or genes. The protein sequence derived from the newly sequenced gene would be compared to the YPs of *D. melanogaster* and similarities determined. Such a

comparison should provide a great deal of information regarding the conservation of protein sequence and specifically for this project should help to identify putative protein domains of importance in YP uptake. A precedent for the identification of domains of interest exists in the demonstration that the YPs have a significant similarity to mammalian lipases (Bownes *et al.*, 1988; Perssons *et al.*, 1989). Computer data base searches with the identified conserved domains would be carried out in order to attempt to assign putative functions to these domains. Another advantage of cloning a *yp* gene from a different species is that putative DNA elements responsible for the regulation of the YPs may be identified. Putative regulatory elements of different genes have been identified using this approach, such as in the case of *engrailed* (Kassis *et al.*, 1989), chorion genes (Martinez-Cruzado *et al.*, 1988; Fenerijian *et al.*, 1989; Swimmer *et al.*, 1990) and *hunchback* (Treier *et al.*, 1989). Moreover, the new sequences can be used to confirm the presence of the putative fatty acid binding domain in the YPs.

The cloned *yp* gene would be used to analyse the sites and pattern of expression of the YPs in the selected species. The conservation of temporal and spatial *yp* gene expression between species may indicate that the regulation of the YPs in the different species is achieved in a similar manner and thus that the *yp* gene DNA sequence

is likely to have conserved regulatory sequences.

This data and the *yp* gene DNA and YP amino acid sequences would provide information on the degree of yolk protein divergence (functional and sequence) and add to knowledge of protein sequence evolution in Diptera.

CHAPTER 2

MATERIALS AND METHODS

2.1 Solutions and Media

All chemicals were obtained from BDH, Fisons, Sigma and Aldrich. Radioisotopes were obtained from Amersham. The ^{35}S-methionine used in later experiments was obtained from New England Nuclear (NEN).

Restriction Enzymes were obtained from BCL and NBL. T_4 DNA Polymerase, T_7 DNA Polymerase, Klenow fragment of DNA polymerase, DNase I, RNase I and Proteinase K were obtained from Pharmacia and BCL.

The Nitrocellulose membranes were obtained from Schleider and Schuell. Hybond-C, Hybond-N in circles and rolls were obtained from Amersham.

Standard solutions were made with sterile distilled water in baked glassware and later sterilised by autoclaving (15 psi/15 to 20 min) or by passing the solution through a 0.45 μm nitrocellulose filter. The solutions and media most commonly used are described below.

TE Buffer	10mM Tris-HCl pH 8.0, 1mM EDTA
Sequencing TE Buffer	10mM Tris-HCl pH 8.0, 0.1 mM EDTA

TM Buffer	0.1M Tris-HCl pH 8.0, 50mM $MgCl_2$
PBS	130mM NaCl, 10mM Sodium Phosphate pH 7.2
Phage Buffer	3g KH_2PO_4, 7g Na_2HPO_4, 5g NaCl, 10ml 0.1M $MgSO_4$, 10ml 0.1M $CaCl_2$, 1ml 1% (w/v) gelatin per litre
10x High Salt Buffer	1M NaCl, 0.5M Tris-HCl pH 7.5, 0.1M $MgCl_2$, 10mM DTT
10x Medium Salt Buffer	0.5M NaCl, 0.1M Tris-HCl pH 7.5, 0.1M $MgCl_2$, 10mM DTT
10x Low Salt Buffer	0.1M Tris-HCl pH 7.5, 0.1M $MgCl_2$, 10mM DTT
10x DNA Loading Buffer	50% (v/v) glycerol, 0.25% (w/v) Bromophenol blue
RNA Loading Buffers	
FSB (formaldehyde sample (buffer)	50% (v/v) formamide, 25% (v/v) formaldehyde (at 4% (w/v)), 25% (v/v) 10 x MOPS

FDE (ficoll, dye, EDTA)	50% (v/v) 0.2M EDTA pH 7.0, 3% (w/v) Ficoll (type 400), 0.25% (w/v) Bromophenol Blue
10x TBE	0.89M Tris borate, 0.89M Boric acid, 10mM EDTA
10x MOPS	0.2M Na-MOPS pH 7.0, 50mM Sodium Acetate, 10mM EDTA
Denaturing Solution	0.5M NaOH, 1.5M NaCl
Neutralising Solution	0.5M Tris-HCl pH 7.2, 3M NaCl
DNA Extraction Buffer	0.5M Tris-HCl pH 8.0, 0.015M EDTA, 0.15M NaCl
RNA Extraction Buffer	0.1M Tris-HCl pH 7.5, 10mM EDTA, 0.3M NaCl, 2% SDS, 7M Urea
20x SSC	3M NaCl, 0.3M Na-Citrate
Salmon Sperm DNA	A 5mg/ml stock was sonicated extensibly to reduce viscosity. The solution was phenol extracted (2.5.1.1)

	and stored at -20 °C.
OLB (oligo labelling buffer)	Solution O: 0.125M $MgCl_2$, 1.25M Tris-HCl pH 8.0 Solution A: 0.95ml Solution O, 18ml 2-mercaptoethanol, 25µl 20mM dATP, 25µl 20mM dTTP, 25µl 20mM dGTP Solution B: 2M Hepes pH 6.6 Solution C: 90 OD Units/ml of Hexadeoxyribonucleotides (Pharmacia) in TE OLB was made by mixing Solutions A, B and C in the ratio 2:5:3 and stored at -20 °C
10x Denhardts Solution	0.2% (w/v) BSA, 0.2% (w/v) polyvinylpyrrolidine, 0.2% (w/v) Ficoll
X-gal	20mg/ml 5-bromo-4-chloro-3-indodyl-B-galactoside in dimethyl formamide and stored at 4 °C

IPTG	10mg/ml Isopropyl-B-D-thiogalactoside stored at -20 °C
Ringers Solution	3.2g NaCl, 3.0g KCl, 1.8g $MgSO_4$, 0.69g $CaCl_2.2H_2O$, 1.79g Tricine, 3.6g glucose, 17.1g sucrose. Make up to a litre with dH_2O, adjust to pH 6.95, filter sterilise and store at 4°C.
40% Acrylamide Stock (for DNA sequencing)	38% (w/v) Acrylamide, 2% (w/v) N-N-methylene bis-acrylamide in dH_2O, filtered and stored in dark at 4°C.
30% Acrylamide Stock (for protein gels)	29.2% (w/v) Acrylamide, 0.8% (w/v) N-N-methylene bis-acrylamide in dH_2O, stored in the dark at 4°C.
APS	15% or 25% (w/v) ammonium persulphate

Protein Gel Solutions

1x Polypeptide Sample Buffer	50mM Tris-HCl pH 6.8, 1% (w/v) glycerol, 2% (w/v) SDS, 0.5% (v/v) 2-mercaptoethanol, 0.01% (w/v) Bromophenol blue

Separating Gel Solutions

	20%	7%	10.5%
3M Tris-HCl pH 8.8	3ml	3ml	5ml
10% (w/v) SDS	0.2ml	0.2ml	0.4ml
Acrylamide gel stock	13.3ml	4.6ml	14.45ml
80% (w/v) sucrose	3.4ml	-	-
dH_2O	0.7ml	12.7ml	21.3ml
15% APS	35µl	35µl	120µl
TEMED	5.5µl	5.5µl	25µl

Staking Gel Solution

0.5M Tris-HCl pH6.8	1.25ml
Acrylamide gel stock	1.25ml
10%SDS	0.1ml
dH_2O	7.3ml
15% APS	75µl
TEMED	7µl

10x Running Buffer	1.5M glycine, 0.25M Tris, 1% (w/v) SDS

Western Blotting Solutions

Transfer Buffer	0.20M glycine, 25mM Tris, 0.1% (w/v) SDS, 20% (v/v) Methanol
TS Buffer	20mM Tris-HCl pH 7.4, 0.9% NaCl
TSS Buffer	20mM Tris-HCl pH 7.4, 0.9% NaCl, 0.1% TWEEN 20
Block Solution	3% (w/v) BSA in TS Buffer or 5% Milk (Marvel) in TSS Buffer
Developing Solution	120mg 4-chloro-1-naphtol dissolved in 40ml Methanol, add 200mls TS Buffer and H_2O_2 to 0.01% (v/v)

Media

All media (listed in Table 2.1) was sterilised by autoclaving. Vitamins, antibiotics and sugars were made in sterile distilled water, and filter sterilised. Where

Table 2.1 Media

L broth	10g Difco Bacto Tryptone, 5g Difco Bacto Yeast extract, 5g NaCl per liter, pH7.2
2 x TY broth	16g Difco Bacto Tryptone, 10g Difco Bacto Yeast extract, 5g NaCl per liter
L agar	As L broth with the addition of 15g Difco agar per liter
BBL agar	10g Baltimore Biological Laboratories trypticase, 5g NaCl, 10g Difco agar per liter
BBL top agar	As BBL agar except only 6.5g Difco agar per liter
Spitzen minimal salts	10g $(NH_4)_2SO_4$, 20g K_2HPO_4, 30g KH_2PO_4, 5g tri-sodium citrate, 1g $MgSO_4$ per liter
Minimal medium	80ml spritzen minimal salts, 4ml 20% (w/v) glucose, 0.2ml 1mg/ml vitamin B1 per 400ml
Minimal agar	As minimal medium with the addition of 6g Difco Bacto agar per 400ml

appropiate, antibiotics were added to medium.

2.2 Microbial Strains

All bacterial strains, plasmid vectors, bacteriophage vectors, phage vectors and other plasmids are described in Table 2.2. The bacterial stocks were kept in plates at 4°C or frozen in the presence of 15% glycerol. Bacteriophages and phages were kept at 4°C in phage buffer with a few drops of chloroform.

2.3 Fly Stocks and Maintenance

All the *Drosophila* species used were obtained from the Umea *Drosophila* Stock Centre. Corn meal medium was (Table 2.3) was used to maintain *Drosophila melanogaster, D. mauritiana, D. busckii and D. nasuta* at 18°C. The remaining species, namely *D. funebris, D. hydei, D. virilis, D.littoralis, D. lebanonensis, D. lummnei, D. azteka, D. pseudoobscura and D. yakuba* were maintained with Instant *Drosophila* food (North Carolina Laboratories) and kept at 18°C.

The non-*Drosophila* muscomorpha flies, *Calliphora erythrocephala, Lucilia sericata, Musca domestica, Protophormia terrae-novae and Sarcophaga argystoma* were obtained from the Department of Zoology. The flies were

Table 2.2

Host	Genotype	Comments	Reference
NM522	hsd ($M^-S^-R^-$) lac, supE, thi, /FproA$^+$ B$^+$, lagI^q, lacZ M15, traD36	Used as a host for M13 Bacteriophage and plasmids (e.g. pUC8)	Gough and Murray (1983)
TG1	(lac,pro), supE, thi, hsdD5/F'tra D36, proA$^+$B$^+$, laqI^q lacZ M15	Used as a host for M13 bacteriophage	Amersham International (unpublished)
C600	F$^-$,thi-1, leub6 lacY$_1$, tonA21, supE44,-$^-$	Used for growth of Lambda	Appleyard (1656)
NM514	Hfl, hsdR	Selection for cI$^-$ recombinants	Murray(1983)

Plasmids	Genotype	Comments	Reference
pUC8/19	Amp^r	Vector for subcloning *Drosophila* and *Calliphora* sequences	Messing (1983)
pBlue-script	Amp^r	Vector used for subcloning and sequencing of *Calliphora* sequences	

Table 2.2 (Continued)

Plasmids	Genotype	Comments	Reference
pDM238	Amp^r	pBR322(Sutcliffe,1978) containing a 12Kb EcoRI fragment of *Drosophila melanogaster* ribosomal DNA	Roiha *et al.*,1981
pYP1/2/3	Amp^r	pBR322 (Sutcliffe, 1978) containing a HindIII (yp and yp2) or SalI-HindIII (yp3) fragment encoding *Drosophila melanogaster* yp1, 2 and 3	Barnett *et al* (1980)
pGemAdh	Amp^r	Gemini-1 containing an XbaI fragment of the *Drosophila* alcohol dehydrogenase gene	Shirras and Bownes, 1987
pGem --tubulin	Amp^r	Gemini-2 containing a 1.7 kb --tubulin DNA fragment.	Kalfayan and Wensink, 1982

Bacteriophage	Comment	Reference
M13mp18/19	Vectors used for subcloning and production of single stranded DNA for sequencing	Yanish-Perron (1983)

Table 2.2 (Continued)

Phage	Comment	Reference
NM1149	Vector used to contruct HindIII genomic library from *D. funebris*	Murray(1983)
Calliphora EMBL4	Random *Calliphora* genomic library cloned in the BamHI unique cloning site	Rubacha *et al*, 1988

Table 2.3 Media used for Propagating *Drosophila*

Adh Food	100g Flake yeast, 100g sugar, 16g agar and distilled water to one liter
Cornmeal Food	25g Cornflour, 50g sugar, 17.5g Yeast pellets, 10g Agar and distilled water to one liter

kept at 25°C in three cubic feet cages and fed on sugar, raw meat and water. Vitellogenic females were chosen for all experiments.

2.4 Methods

2.4.1 General Methods

The most commonly used methods during these studies are described below.

2.4.1.1 Phenol extraction

Redistilled phenol equilibrated in water was obtained from Rathburn Chemical Limited (Walkerburn, Scotland). The phenol was equilibrated in TE buffer and 8-hydroxyquinoline was added to 0.1% (w/v). Phenol was stored at 4°C. Solutions of nucleic acids were deproteinised by gentle shaking (DNA extractions) or vortexing (RNA extractions) with an equal volume of phenol or phenol:chloroform (1:1). The phases were separated under centrifugation in a Sorvall RC-5B centrifuge (5Krpm, 10 min) or a microfuge (12Krpm, 3 min) and the aqueous layer carefully removed to a fresh tube. This procedure could be repeated up to three times depending on the type of extraction.

2.4.1.2 Chloroform extraction

The chloroform extraction was used to remove the remaining contaminating phenol after phenol or phenol: chloroform extractions. An equal volume of chloroform equilibrated with TE buffer was added to a solution of nucleic acid which was mixed or vortexed. The phases were separated by centrifugation (2.4.1.1) and the aqueous layer placed in a fresh tube.

2.4.1.3 Ether extraction

Di-ethyl ether extractions were used to remove contaminant phenol. Di-ethyl ether was equilibrated with TE buffer and an equal volume added to nucleic acid solutions. The solutions were shaken and the phases allowed to settle. The ether was removed carefully, and any remaining ether was removed by placing the solution at 65°C for 2 minutes.

2.4.1.4 Precipitation of nucleic acids

Nucleic acids were precipitated from solutions by adjusting the solution to 0.3M Sodium Acetate pH 5.5 and adding 2.5 volumes of pre-chilled absolute ethanol, or 0.6 volume of iso-propanol. The samples were incubated at -20°C for at least 2 hrs. The precipitate was recovered by

centrifugation in a Sorvall RC-5B centrifuge (8Krpm, 10 min) or microfuge (12Krpm, 10 min). The pellet was washed with pre-chilled 70% Ethanol and dried under vacuum. The dried pellet was dissolved in TE buffer or water.

DNA was also precipitated from solutions containing nucleoside triphosphates by adding an equal volume of 4M ammonium acetate pH 6.6 and 4 volumes of ethanol. The DNA solution was incubated for 5 minutes at -20°C and then recovered as described above. This precipitation carried out two times in succession removes up to 90% of the unincorporated labelled triphosphates from the reaction (Okayama and Berg, 1982).

2.4.1.5 Restriction endonuclease digestion of DNA

DNA was digested with a 2 to 4 fold excess of enzyme for a period of at least 2 hrs in reaction conditions recommended by the manufacturer. Completion of digestion was tested by running a fraction of the digestion in a 1% agarose gel.

2.4.1.6 Isolation of DNA Fragments

2.4.1.6.1 Isolation by GeneClean®II

GeneClean®II is a trademark of Stratech Scientific (Luton). The isolation of a particular fragment of DNA was completed by using the procedure included in the Kit. The mixture of fragments of DNA was run in a 0.75% to 1.25% agarose gel, depending on the size of the fragment of interest. Once the fragment was clearly separated from any other contaminating fragments it was excised with a sterile scapel blade from the gel and placed in a pre-weighed eppendorf tube. The tube was weighed again and the difference obtained. Assuming that 1 gramme of fragment is equivalent to 1 ml of NaI solution, the following were added to the tube: 0.5 vol of TBE equilibrator and 4.5 vol of NaI. The tube was shaken gently and placed in a water bath at 50°C. The eppendorf was inverted once every minute until the fragment was dissolved. At this point if the total amount of DNA is less than 5µg, then 5µl of the glass milk were added to the tube and mixed, followed by an incubation on ice of 10 minutes. The tube was gently shaken every minute, then spun for 8 seconds at 12krpm and the supernatant discarded. The pellet was re-spun for a minute to get rid of the remaining solution. The pellet was resuspended gently in 500µl of NEW solution (This is a concentrated stock of NaCl, EDTA and Tris HCl, provided by the manufacturer, which is made to a working solution by adding 14ml of the stock to 280 ml of dsH_2O and 310 ml of absolute ethanol) and spun for 8 seconds (12krpm), after which the supernatant was discarded. The procedure was

repeated twice and the tube re-spun for a minute (12krpm), the resultant supernatant being descarded. The remaining pellet was resuspended in sterile water (10µl) and placed at 50°C for 2 to 3 minutes to allow the DNA to become free from the glass beads. The tube was spun for a minute and the supernatant was collected into a fresh eppendorf. The procedure was repeated and the resulting supernatant was collected. A fraction of the solution was separated on a 1% agarose gel.

2.4.1.6.2 DEAE Paper

The digested DNA was fractionated in an agarose gel (2.4.3.2). The DEAE Paper was equilibrated in NET (150mM NaCl, 0.1mM EDTA, 20mM Tris HCl pH 8.0) and placed in front of the desired DNA fragment. The gel was run further until the fragment could not be seen in the gel anymore (100volts, 10min). The paper was taken off the gel and placed in an eppendorf tube with 300µl of HNET(1.0M NaCl, 0.1mM EDTA, 20mM Tris HCl pH 8.0). The tube was incubated at 65°C for 20 minutes and every 5 minutes vortexed. This enabled the DNA to detach itself from the paper. The solution was replaced with fresh HNET and the same procedure followed. Both fractions were pooled, phenol extracted (2.4.1.1), ether extracted (2.4.1.3), ethanol precipitated (2.4.1.4) and resuspended in sterile distilled water. A fraction of the solution was separated on a 1% agarose gel to estimate the yield.

2.4.1.7 Ligations

The vector and fragment DNA were cut to completion with restriction endonucleases. The vector DNA was treated with 0.01 units of calf intestine phosphatase (Boeringer Mainheim) for 15 to 20 minutes to prevent religation. The ligations used a DNA concentration of fragment to vector of 3:1 molar ratio. The ligations were carried out in 10 µl volumes in the presence of 1.5% PEG 6000 and 100µM ATP (in the buffer provided by the manufacturer) at 4°C overnight. In adhesive end ligations 0.01 units of T_4 DNA ligase was used. The ligations were transformed into *E. coli* (2.4.1.8).

2.4.1.8 Transformations

L-broth (25ml) was incubated with 0.25ml of an overnight culture of the relevant bacterial strain and grown with shaking until the OD_{650} reached 0.45 to 0.55. The culture was chilled on ice for 10 minutes and the cells were then pelleted by centrifugation (4K rpm,5 minutes,4°C) and gently resuspended in 15 ml of 50mM $CaCl_2$. The suspension was incubated in ice for 15 minutes and the cells re-pelleted as before. The pellet was resuspended in 2ml of ice cold 50mM $CaCl_2$. Aliquots of 0.2ml were added into 5ml glass tubes with ligation mixtures for each transformation and incubated in ice for

30 minutes. The cells were then heat shocked at 42°C for 2 minutes.

In the case of plasmid DNA transformations with antibiotic selectable genotypes, 0.9 ml of L-broth was added to each mixture and the cells incubated at 37°C for 1 hour. Samples of 0.2ml were plated onto L-agar plates supplemented with the appropiate antibiotic. The plates were inverted and incubated at 37°C overnight.

In the case of M13 DNA transformations, 3ml of molten BBL-top agar at 42°C, 25ul of BC1G (25μg/ml), 25ul of IPTG (25μg/ml) and 0.2ml of stationary TGI cells were added to each tube. The mixture was poured onto minimal medium plates and allowed to solidify before being inverted and incubated at 37°C overnight. In the case of pUC plasmid vectors the same procedure was follwed after the addition of L-broth and the incubation at 37°C with shaking for 1 hour.

2.4.1.9 Plasmid DNA extraction

2.4.1.9.1 mini-preparation

Cultures of chosen colonies were setup in L-broth supplemented if necessary with the required antibiotic. 1.5ml of the overnight culture was poured into an

eppendorf tube and spun for 10 minutes at 12Krpm (microfuge, Eppendorf model 5414). The supernatant was discarded and the pellet was gently resuspended in 0.3ml of 50mM Tris HCl pH8.0, 10mM EDTA and 400 µg RNAse A per ml. Inmediately after this, 0.3ml of 0.2M NaOH and 1% SDS were added and the tube shaken gently. The tube was incubated for 5 minutes at room temperature after which 0.3ml of 2.5M Potassium acetate pH4.8 was added, followed by gentle mixing and a 15 minute spin at 12Krpm. The supernatant was collected, placed in a fresh tube and ethanol precipitated (2.4.1.4). The pellet was resuspended in 50µl of TE buffer, 10µl of which was used for restriction endonuclease assay.

2.4.1.9.2 maxi-preparation

An overnight culture of the desired bacterial strain was added to 500ml of fresh prewarmed L-broth. The medium was supplemented with the required antibiotic. The culture was grown overnight at 37°C with vigorous shaking. The following morning the bacteria were placed in two 250ml buckets and spun in the GSA rotor (RC-5B, Sorvall centrifuge) at 4K rpm for 5 minutes. The supernatant was discarded and the pellet resuspended carefully in 5ml of solution 1 (50 mM glucose, 25mM Tris HCl pH8.0, 10 mM EDTA) to which freshly prepared lysozyme was added to a final concentration of 5µg/ml. The suspension was transfered to 50ml plastic tubes, incubated for 5 minutes

on ice and 12 ml of solution 2 was added (0.2M NaOH, 1%SDS). The tubes were mixed gently and incubated for a further 10 minutes on ice. 9ml of solution 3 (3M Sodium acetate pH4.8) was added to the mixture, mixed gently and incubated for 20 minutes before being centrifugated for 30 minutes at 20K rpm at 4°C (SS-34 rotor, RC-5B Sorvall centrifuge). The pellet was discarded and the supernatant transferred to a fresh tube where 2ml of 2.5M KCl was added. The solution was mixed and the precipitates collected at 15K for 10 minutes at 4°C (SS-34 rotor, RC-5B Sorvall centrifuge). The supernatant was collected and 10ml of iso-propanol added, this solution being incubated at -20°C for 1 hour. The precipitates were spun down at 12k rpm for 10min at 4°C (SS-34 rotor, RC-5B Sorvall centrifuge). The supernatant was discarded and the pellet washed with 70% ethanol and spun again as above. The pellet was resuspended in 10ml of TE buffer, 10g of CsCl was added in addition to 1ml of 10mg/ml of ethidium bromide. The mixture was loaded into a Ti50 tube and balanced to within 0.02g. The tube was sealed and spun in the ultracentrifuge for 48 hours at 38Krpm, after which it was removed from the rotor gently and placed under UV light. This revealed two bands of which the bottom one was taken. The top one was avoided as it contains sheared DNA. The bottom band was then placed in a glass tube where TE buffer with 1g/ml CsCl saturated isobutyl alcohol was added. The alcohol was used to eliminate ethidium bromide and up to five changes of the alcohol were used

before all residues of ethidium bromide were extracted. The remaining solution was dyalised against four changes of one litre of cold TE Buffer. The dyalised solution was collected and diluted to 10 ml with TE buffer and then iso-propanol precipitated (2.4.1.4). The pellet was resuspended in 05-1ml of TE buffer.

2.4.1.10 Phage DNA preparation

2.4.1.10.1 plate lysate

Plate lysates were used to make phage stocks. A single colony of the necessary plating cells was picked and grown overnight. The cells were spun and resuspended in 10mM $MgSO_4$. A 0.2 ml aliquot of this suspension was mixed with 50 µl of phage suspension and placed at 37°C for 20 minutes. The suspension was then added to 3 mls of molten agarose (0.7% w/v) and plated on a 90 mm petri dish. The plaques were allowed to grow until they touched each other (6 to 8 hours) and then 5 mls of phage buffer were added to the plate. The plate was placed at 4°C and mixed gently for a few hours. The suspension was collected and another 1 ml aliquot was added to the plate, which was placed at an angle for 20 minutes. This enabled the collection of the remaining suspension. Both collections were pooled together and 0.1 ml of chloroform was added. The mixture was vortexed and spun at 3000 x g for 10 minutes (SS-34 rotor, RC-5B Sorvall centrifuge).

The supernatant was collected, stored in a sterile vessel at 4°C and titred.

2.4.1.10.2 liquid lysate

Small scale liquid lysates were used routinely. The method used was a variation from Grossberger (1987). A plaque of the required clone (EMβL 3) was added to 0.2 ml of an overnight of C600 cells and 0.3 ml of 10mM $CaCl_2$ and 10 mM $MgCl_2$. The mixture was incubated for 10 minutes at 37°C after which it was added to 10 mls of L-broth supplemented with 10 mM $MgCl_2$ and 0.1% glucose. The culture was grown overnight after which lysis was observed. The culture was spun at 5K for 5 minutes at 4°C (SS-34 rotor, RC-5B Sorvall centrifuge) and the supernatant incubated at 37°C in the presence of 1ug/ml of DNase and RNase A for 30 minutes. The suspension was then spun at 30k for 1 hour in a swing out rotor (Tst 41.14), the supernatant discarded and the pellet resuspended in 0.4ml TM (mM Tris pH 8.0, mM $MgSO_4$) to which 1µg/ml of proteinase K was added. The suspension was incubated for 1hour at 65°C, phenol extracted (2.4.1.1), and ether extracted (2.4.1.3). The DNA was ethanol precipitated (2.4.1.4) and The washed and dried pellet resuspended in 50 µl of TE, of which 5 µl was used in the appropriate restriction endonuclease assay.

2.4.1.11 Radiolabelling of fly haemolymph and tissues

2.4.1.11.1 injection

Injection of ^{35}S-Methionine (400 Ci/mmol, New England Nuclear) was done with 50 µl capillary tubes heat pulled to form a fine end. The end was broken off with a pair of tweezers to form a very fine and sharp needle. The anaesthetised flies were injected between the fourth and fifth abdominal segments with 2 µCi of ^{35}S-Methionine. The flies were incubated for 4 hours at 25°C after which time the ovaries and haemolymph was extracted.

2.4.1.11.2 feeding

Vials containing 3% agar medium, baker's yeast, 100µl of 20% sucrose and 2µl of ^{35}S-methionine (24 µCi) were incubated at 25°C overnight. Female donor flies were placed in the vial and incubated for another 24 hours at 25°C. The ovaries were dissected and the yolk protein isolated as described below (2.4.2.1.3). This same procedure was used to label the haemolymph of donor flies for haemolymph transplant experiments (2.4.1.12).

2.4.1.11.3 ovary in vitro culture

The ovaries of flies were dissected out in Ringer's solution and rinsed two times. The ovaries were placed in an eppendorf tube with 0.1 ml of Ringer's solution and

10μCi of ^{35}S-Methionine, incubated for two hours at RT after which they were spun and the suspernatant collected and stored at 20°C. The ovaries were rinsed twice wth fresh Ringer's solution and stored at 20°C.

2.4.1.12 Haemolymph transplants

Female donor flies were injected with ^{35}S-Methionine as described in 2.4.1.11.1. The haemolymph was collected from the individual fly and inmediately injected into a host female fly. The flies were incubated at 25°C in a humid chamber for a period of 18 to 22 hours. The haemolymph was collected and the ovaries dissected out.

It is important to note that the collection of haemolymph from donor flies using the injection apparatus yielded less haemolymph than when capillary activity alone is used to recover the haemolymph. Furthermore, the injection of the ^{35}S-methionine results in dehydration of the flies which makes the recovery of haemolymph difficult. This was solved in part by placing the vials with injected flies in a sealed beaker with water.

In the experiments that required egg collection, males were placed in the same vials as the injected host flies to ensure fertility and thus egg laying.

2.4.1.13 Radioactive labelling of DNA

Fragments of DNA were radioactively labelled by Random Priming (Feinberg and Volstein, 1983; 1984). The DNA fragment was denatured by boiling for 5 minutes followed by rapid cooling on ice for another 5 minutes. Between 50 to 200 ng of DNA was labelled in a reaction volume of 50 µl containing 10µl OLB (2.1), 20 to 50 µCi ^{32}P-dCTP, 1µl BSA (20mg/ml) and 1µl of 1 Unit/ul of Klenow fragment of DNA polymerase 1. The reaction was incubated for 1 to 4 hours at 37°C.

To measure the percentage incorporation of the label in the DNA, 1µl of the solution was placed on a glass filter (Whatman). Another 1µl aliquot was placed in 200µl of dsH_2O containing 5µl of 20mg/ml BSA. To this, 50µl of 50% TCA was added and the mixture was passed through a glass filter. This was washed with 5% TCA after which both filters were dried and placed in a scintillation vial with 1ml of scintillation fluid (ecoscint, National Diagnostics). The ratio between the spot and acid insoluble counts was obtained, giving the percentage incorporation. Specific activity was also calculated by multiplying the total volume by the acid insoluble counts, the product then divided by the amount of DNA. The specific activity used in hybridisation was between 10^8 and 10^{10} cpm/µg DNA.

2.4.1.14 Autoradiography and fluorography

Autoradiography of dried DNA sequencing gels, dried protein gels and nitrocellulose or nylon membranes was performed using Cronex 4 (Dupont) X-ray film and cassettes. Films were developed in an Agfa 1 automatic film processor.

2.4.2 Preparation of Specimen Samples

2.4.2.1 Protein samples

Samples collected from flies, carcasses, ovaries and haemolymph were placed in 20 to 50 µl of 2 X Laemmli's buffer (Laemmli, 1970). The buffer was made up of, 6 mls of dsH_2O, 6 mls 0.5M Tris HCl pH6.8, 4.8 mls of glycerol, 9.6 mls 10% (w/v) SDS, 2.4 mls of 2-mercapto ethanol and 0.1 g of Bromophenol blue. The samples were stored at -20°C and before use were denatured at 100°C for 5 minutes.

2.4.2.1.1 Haemolymph extraction

Etherised flies were placed on double sided sticky tape on a slide. Glass needles were prepared by pulling 50ul glass capillaries (Drummond Scientific) in a vertical pipette puller (David Kopt Instruments, model 700G, Tujunga, California). The haemolymph was collected with a glass needle through the ventral surface of the thorax,

placed in polypeptide sample buffer, and stored at-20°C. Five to ten haemolymphs were collected per sample.

2.4.2.1.2 Tissue sample preparation

Etherised female flies were placed on a concave slide with Ringer's fly solution. The fly was submerged in the Ringer's and the ovaries dissected out using watchmaker forceps. The ovaries were collected in 50ul of polypeptide sample buffer and stored at -20°C. The ovaries were disrupted by repeated freeze/thawing. This process was aided by homogenising with a pasteur pipette the tip of which had been melted into a ball to fill the bottom of an eppendorf tube.

2.4.2.1.3 Yolk protein Isolation

The isolation of the Yolk protein was performed using a method obtained from Peter Isaac's Ph.D Thesis (1982). The method was devised for the isolation of Yolk proteins from embryos and used in this thesis to isolate yolk proteins from adult or ovaries of adult *D. melanogaster*. All stages of the procedure were done at 4°C. Two hundred adult flies were homogenised in 10ml of extraction buffer (400mM MaCl, 20mM $MgCl_2$ and 10mM Sodium orthophosphate pH7.0. The homogenate was spun at 10Krpm for 15 minutes (SS34 rotor, Sorvall RC-5B centrifuge), the supernatant was collected and filtered through four layers of sterile

muslin. The filtrate was then passed through a glass filter under vacuum (Whatman GF/C). After the filtration was completed the volume was made up to 30ml with extraction buffer and 30µl of freshly made 100mM PMSF. The following set requires slow addition of ammonium sulphate to 50% saturation (30g), the solution is stirred for an additional 30 minutes after the addition of the ammonium sulphate. The resulting solution was spun for 10 minutes in the HB4 rotor at 9Krpm. The supernatant was dialysed against 4.5 litres of 10mM sodium orthophosphate pH 7.0 overnight at 4°C, followed by dyalysis against dsH_2O for 6 to 8 hours. The resultant solution was transferred into three 100ml round bottom flasks each with approximately 30mls of fluid. These were freeze dried overnight, and the protein collected and stored at 4°C. A fraction of the protein was dissolved in 20% glycerol to 1mg/ml. Each flask normally yielded 15mgs of protein (see results Chapter 3 for resulting purity of yolk protein, Figure 3.2)

2.4.2.2 Nucleic acid sample preparation

2.4.2.2.1 Fly genomic DNA extraction

A rapid method for DNA extraction from single flies was used, essentially that of Marcus (1985). Care was taken to prevent shearing of DNA. The method was scaled up for up to 10 *Drosophila* flies or one non-*Drosophila*

fly. The same method was used in either case. Flies were etherised, collected into eppendorf tubes and frozen in liquid nitrogen. The flies were stored in liquid nitrogen or at -70°C. The flies were homogenised in 350ul of DNA extraction buffer on ice. A further 350µl of the same buffer, supplemented with 0.04% SDS, was added, followed by Proteinase K to 50µg/ml. The homogenate was mixed and placed at 65°C for 1 hour, then phenol chloroform extracted (2.4.1.1) twice, chloroform extracted (2.4.1.2) once and isopropanol precipitated (2.4.1.4). The nucleic acid was resuspended in water or TE buffer at 4°C.

The genomic DNA was digested for 2 hours with 10 units of a chosen restriction endonuclease, in the appropriate buffer conditions, RNase A (to 0.1µg/µl) was added to the reaction mixture.

2.4.2.2.2 RNA extraction from adult tissues and flies

Usually 10 to 20 carcasses of either whole males or females, or tissues (ie. ovaries) of *Drosophila* flies were frozen in liquid nitrogen in the presence of RNA extraction buffer. For the non-*Drosophila* flies, 1 to 3 carcasses and 2 to 5 ovaries were used and treated in the same manner. Homogenisation was performed quickly on ice and phenol chloroform extractions (2.4.1.1) followed inmediately. The phenol/chloroform extract was back extracted and the resulting aqueous layers pooled and

chloroform extracted (2.4.1.2) and then ethanol precipitated (2.4.1.4). The pellet was resuspended in sterile distilled water and 2 volumes of 3M LiCl in order to selectively precepitate RNA from the nucleic acid solution. The sample was incubated for 2 hours at -20°C followed by centrifugation and the pellet washed in 70% ethanol, dried and resuspended in water.

2.4.3 Gel Electrophoresis

2.4.3.1 SDS-PAGE of proteins

Samples for SDS-PAGE analysis (2.4.2.1) were collected in 50µl of polypeptide sample buffer, frozen and stored at -20°C. Before loading the samples onto the gel, they were boiled for 5 minutes and centrifuged briefly (12Krpm, 1 minute).

Acrylamide gels to separate proteins were either linear (10.5% (w/v)) or gradient (7 to 20% (w/v)). Gradient gels were poured using a two chamber pouring device aided by a peristaltic pump (P3-Pharmacia). The linear gels were run at 65 volts for 16 to 18 hours and gradient gels at 140 volts for a similar period of time.

After electrophoresis, radioactive gels were fixed in 10% (v/v) Acetic acid for 30 minutes followed by Amplify (Amersham) for another 30 minutes. The gels were then

dried at 65°C for 2 hours and exposed directly onto X-ray film (Dupont).

Gels for staining were either Coomassie blue stained or silver stained. Prior to staining the gels these were fixed for 30 minutes in 10% (v/v) Acetic acid followed by staining with 0.1% (w/v) Coomassie blue in 10% (v/v) Acetic acid and 40% (v/v) Methanol. After staining the gel for at least two hours the gel was destained with several changes of 10% (v/v) Acetic acid and 40% (v/v) Methanol until the gel was clear and showed distinct bands.

The gels for silver staining were Coomassie blue stained and destained as described above. The gels were then prefixed in 5% (v/v) Methanol and fixed with 10% (v/v) gluteraldehyde each for 30 minutes. After the fixing the gel was washed thoroughly with repeated changes of dH_2O for at least one hour. The gel was placed in a solution of 5µg/ml of DDT for 30 minutes followed by 30 minutes in 0.1% (w/v) $AgNO_3$. The gel was washed in dH_2O and then twice in 3% (w/v) Na_2CO_3 to equilibrate the protein gel. The gel was then placed in 3% (w/v) Na_2CO_3 with 100ul of 37% (v/v) formaldehyde per 100mls of solution. After 15 minutes (apprx.) of staining the reaction was stopped with 5mls of 2.3M Citric acid. The gel was washed thoroughly in dH_2O and was stored in 0.03% (w/v) Na_2CO_3 before a photograph was taken.

2.4.3.2 Agarose gel electrophoresis of DNA

Agarose gels were made and run in 1 x TBE gel buffer. Ethidium bromide was added at a concentration of 0.5mg/ml to the gel.

Mini-gels were cast and run in a Cambridge Biotechnology minigel apparatus (model CB1000). These gels were run at 50 to 80 volts for 1-2 hours. Larger gels (16x10x1 cm) were run submerged in a home made gel kit at 20 -100 volts for 3 - 18 hours. After electrophoresis the DNA bands were visualised on a UV trans-illuminator viewing system (254nm short, 265 nm long) Chromatovue (model C-70G UV).

2.4.3.3 Agarose gel electrophoresis of RNA

RNA samples (5-10µg total RNA in dH_2O) were added to an equal volume of formaldehyde sample buffer (FSB) and heated to 65°C for 5 minutes, then immediately cooled on ice before adding Ficoll-dye-EDTA (FDE). The samples were loaded onto a 1.3% (w/v) agarose gel containing 1 x MOPS, 17.3% (v/v) formaldehyde and 0.5mg/ml ethidium bromide. The gels were run submerged in home-made kits at 20-100 volts for 3-18 hours. After electrophoresis the RNA was visualised on a UV trans-illuminator viewing system (254nm short, 265nm long) Chromatovue (model C-70G UV).

2.4.3.4 Polyacrylamide gel electrophoresis for sequencing

Sequencing gels (40cm x 30cm x 0.4mm, 45cm x 18cm x 0.2mm or 60cm x 18cm x 0.2mm) were 6% (w/v) acrylamide (19:1 acrylamide : bis-acrylamide) in TBE gel buffer containing 7.7M urea. All gels were poured with 1 x TBE acrylamide gel mix.

Sequencing gels were run at a constant power of 40 watts for plates 40cm x 30cm x 0.4mm, for times ranging between 2 and 9 hours. For plates 45cm x 18cm x 0.2mm the power used was 35 watts whereas for plates 60cm x 18cm x 0.2mm gels were run at 45 watts, both for 2 to 9 hours. After electrophoresis, gels were fixed in 10% (v/v) acetic acid and 10% (v/v) methanol for 20-30 minutes, transferred onto blotting paper (Ford Goldmedal Blotting Paper), dried down on a gel drier (BioRad model 583) at 80°C for 2 hours, and autoradiographed (2.4.1.14) for 14-20 hours at room temperature.

2.4.4 Transfer onto Membrane Filters

2.4.4.1 Western blot

Protein gels were run as described (2.4.3.1) and the transfer of the electrophoresed proteins to a

nitrocellulose membrane was performed electrophoretically in a BioRad Trans-Blot Cell. The nitrocellulose was cut to the same size as the gel and a sandwich of blotting paper / nitrocellulose / gel / blotting paper was made between two scotchbrite pads. The blotting paper, nitrocellulose and scotchbrite pads were pre-soaked in transfer buffer. The sandwich was inserted into the cell so that the nitrocellulose filter was placed between the gel and the anode. The proteins were electrophoresed onto the nitrocellulose filter at 60 volts for 4 hours. After the transfer was completed, the proteins were visualised by staining with a solution of 0.3% (w/v) Ponceau S in 3% (w/v) TCA. The filter was destained with distilled H_2O.

The filter was placed in block solution for 2-4 hours. For transfers using anti-YP antibodies the block solution was 3% (w/v) BSA in TS buffer, however; for transfers using anti-BSA antibodies the block solution was 5% (w/v) Marvel (Cadbury's) in TS. The same procedure was followed for both sets of transfers unless otherwise stated. The block solution was substituted by fresh block solution containing primary antibody (rabbit anti-YP antibody diluted to 1:500 or mouse anti-BSA antibody diluted to 1:1000) and shaken at 4°C for 16-18 hours. The filter was then washed in five changes of TS over a period of 30 minutes, placed in block solution containing the secondary antibody (goat anti-rabbit IgG-horse raddish peroxidase conjugate or goat anti-mouse IgG-alkaline

phosphatase conjugate) and shaken at room temperature for 3-4 hours. The filter was washed again in 5 changes of TS over a period of 30 minutes before developing. If the secondary antibody had a horse rad'ish peroxidase conjugate the filter was developed in 200ml of developing solution, however; if the secondary antibody was conjugated to alkaline phosphatase the developing solution contained 140ul of BCLG () and 70ul of NBT (). The colour reaction was stopped by washing the filter in distilled water. The filter was dried and stored in the dark.

2.4.4.2 Southern blot (Southern, 1975)

Genomic DNA samples were run in 1% (w/v) agarose gels (2.4.3.2). Once electrophoresis and photography of the gel was completed, the gel was placed in denaturing solution for 30-60 minuted followed by neutralising solution for the same period of time.

A blotting paper wick was placed on a platform supported above a reservoir of 20X SSC in which the ends of the wick were immersed in the buffer. The wick was thoroughly wet with 20X SSC and the gel laid on top. A 2X SSC prewetted sheet of Hybond-N membrane filter (Amersham) cut to the same dimensions as the gel was placed on top of the gel. Three sheets of blotting paper cut to the same size as the gel were soaked in 2X SSC and placed on top of the filter. At every succesive step care was taken to

avoid air bubbles being trapped. Dry piece of blotting paper were laid on top of the wet blotting paper, followed by paper towels to a thickness of 3-4cm. A sheet of perspex was laid on top of the paper towels and a weight place upon this (500-1000 grams). Transfer was allowed to take place for 16-18 hours after which time the membrane was removed and rinsed in 2X SSC dried, and UV fixed for 8-10 minutes with the DNA side facing the UV light source. The membrane was latter prehybridised and probed as described below (2.4.5.1)

2.4.4.3 Northern blot (Thomas, 1980)

Gels for Northerns were prepared and electrophoresed as described above (2.4.3.3). After electrophoresis and photography, the gel was transferred to a platform with a blotting paper wick prewet in 20 x SSC. The remainder of the transfer procedure was carried out as described for Southern blotting (2.4.4.2).

2.4.4.4 Plaque lifts

Lambda phage were plated on BBL plates. An overnight culture of plating cells (NM514 for NM1149 or C600 for EMBL4) was placed in eppendorf tubes and spun in a microfuge for 5 minutes. The pellet was resuspended in half the original volume with 10mM $MgSO_4$. The required amount of phage (10ul of an adequate dilution) was mixed

with 100ul of plating cells and incubated for 20 minutes at 37°C. To this 3ml of molten L-agarose (0.7% w/v) was added and the mixture plated on the predried BBL plate. After the medium settled the plates were inverted and grown overnight at 37°C. In the case of M13 plaques, the same procedure was followed except that to the molten agar 100µl of plating cells (TG1 cells), 30µl of 25mg/ml IPTG and 30µl of 25mg/ml BCIG were added, before the phage and cell mixture was combined with the molten agar. The molten agar was then plated on minimal plates.

Plates with recombinant lambda or M13 plaques were placed at 4°C for 1 hour. Later, circular Hybond-N filters (Amersham) were placed on the surface of the plate taking care to avoid formation of bubbles. The filter was oriented and left on the surface of the plate for 1 minute. The filter was lifted gently and placed plaque side up on a pad of blotting paper, prewet with denaturing solution, for 7 minutes. The filters were tranferred to another pad of blotting, paper this time prewet with neutralising solution for 6 minutes. Finally the filters were rinsed for 1minute in 2 x SSC and blotted dry before UV fixing for 8 to 10 minutes. The filters were then prehybridised and probed as described below (2.4.5.1).

2.4.5 Hybridisation of Membrane Filters with DNA Probes

2.4.5.1 Hybridisation of southern blots

Filters were prehybridesed in heat sealable polythene bags (Krups Vacupac) for a minimum of 30 minutes at 42°C in 10ml of hybridisation solution. The hybridisation solution was 0.4M disodium hydrogen phosphate pH7.0, 7% (w/v) SDS and 50% (v/v) formamide. The prehybridisation solution was removed and replaced with 10ml of fresh hybridisation solution. The probe was heat denatured, cooled quickly and added to the bag. The hybridisation was allowed to proceed for 14-18 hours at 42°C after which the filter was washed in a solution of 1 x SSC, 0.1% (w/v) SDS twice for 20 minutes at 65°C. The filter was blot dried, wrapped in clingwrap (Vitafilm, Goodyear Tyre and Rubber Co, Ltd, Staffordshire, GB) and autoradiographed (2.4.1.14).

The blot may be stripped and reprobed. The stripping of the blot was carried out by allowing water to come to a boil. After the water has boiled the heat source is turned off and SDS is added to 0.1% (w/v). The filter is placed in the water after which the water is allowed to cool down to room temperature. Once the water reaches room temperature the filter is taken out and rinsed in 2 x SSC prior to prehybridisation.

2.4.5.2 Hybridisation of northern blots

The procedure of hybridisation of Northern blots was performed in the same manner as that for the hybridisation of Southern blots. The northern filters were also stripped and reprobed as described for southern blots.

2.4.6 Hybridisation to Whole Mounts

2.4.6.1 Fixation of tissues

Ovaries were dissected from *Calliphora* in the same manner as described for *Drosophila* (2.4.2.1.2). The method of fixing the ovaries was essentially the same as that described in Tautz and Pfiefle (1989). The ovaries were dissected in fixing solution (0.1M Hepes pH 6.9, 2mM $MgSO_4$, 1mM EGTA and 5% (w/v) paraformaldehyde. The paraformaldehyde stock solution (20% w/v) was dissolved by heating to 65°C, neutralised with NaOH and stored at -20°C. The dissected ovaries were placed in a sterile Universal flask for 20 minutes. Heptane was added to the later ovary stages (stages 11 and beyond) so that the fix could penetrate the chorion. After gentle shaking 10ml of methanol were added. The fix solution was discarded and replaced by 1ml of ME (90% (v/v) Methanol and 10% (v/v) 0.5M EGTA pH 8.0). The ovaries were stored at -20°C for up to two weeks. The ovaries were then refixed and dehydrated through a series of steps consisting of ME and PP (4% (v/v) paraformaldehyde in PBS). The first step in 7/3 ME/PP, the second step 1/1 ME/PP, the third step 3/7

ME/PP each for 5 minutes. The final step was in PP for 20 minutes. The ovaries were then washed in PBS twice for 10 minutes. All subsequent stepswere done in 1.5ml eppendorf tubes at room temperature on a revolving wheel. The ovaries were first washed 3 times for 5 minutes each in PBT (PBS plus 0.1% (v/v) Tween 20). They were then incubated for 3 minutes in 50 µg/ml of Proteinase K in PBS. The digestion was stopped by incubating in 2mg/ml of glycine in PBT for 2 minutes. The ovaries were then washed 2 times for 5 minutes each in PBT, refixed for 20 minutes with PP and finally washed 3 times for 10 minutes each in PBT.

2.4.6.2 Probe labelling

Probes were made with genomic fragments belonging to the 3' coding end of *yp* genes, labelled according to the protocol supplied with the Beohringer Mannheim Kit for non radioactive labelling of DNA with Digoxigenin. The DNA (250ng/reaction) to be labelled was purified (2.4.1.6), denatured at 100°C for 10 minutes and rapidly cooled on ice for a further 5 minutes. The buffer containing unlabelled nucleotides dATP, dGTP, dCTP and digoxygenin labelled dUTP in the correct salt and primer conditions were added, followed by the addition of 1 unit of the Klenow fragment of DNA Polymerase 1. The reaction was incubated at 37°C. The labelling was done by random priming (Feinberg and Vogelstein, 1983). The probes were

precipitated with an equal volume of 4M Ammonium Acetate pH 6.6, 10ug of Salmon sperm DNA and four volumes of absolute ethanol incubated at -20°C for 1 hour. The labelled fragments were recovered by centrifugation for 10 min, washed with 70% ethanol, dried and resuspended in hybridisation buffer.

2.4.6.3 Hybridisation and washing

The hybridisation buffer (HS) consisted of 50% (v/v) formamide, 5 X SSC, 50µg/ml heparin, 0.1% (v/v) Tween 20 and 100µg/ml sonicated salmon sperm DNA. The buffer was stored at 20°C. The tissues were washed for 20 minutes in 1:1 PBT/HS followed by washing for 60 minutes with HS. The tissues were then incubated at 45°C for 1 hour prior to addition of the heat denatured probe. The HS was discarded and the freshly heat denatured labelled DNA probe was added, the tube sealed and place horizontally in the water bath at 45°C. The final probe concentration was 0.5µg/ml. The hybridisation was done overnight followed by washes at room temperature of 20 minutes each. The first wash was HS, followed by 4/1 HS/PBT, 3/2 HS/PBT, 2/3 HS/PBT, 1/4 HS/PBT and two washes of PBT of 20 minutes each.

2.4.6.4 Signal detection

The antibody-conjugate solution (supplied with the

Boehringer Kit) was used to detect the digoxigenin. The tissues were placed in blocking solution of 5% sheep serum in PBT for 1 hour at room temperature, before the solution was discarded and replaced by fresh blocking solution with freshly added antibody to a 1:2000 dilution. The tissues were incubated at 4°C overnight. The antibody solution was discarded and the ovaries were washed 4 times for 20 minutes each in PBT, followed by 3 times for 5 minutes each in 100mM NaCl, 50mM $MgCl_2$, 100mM Tris HCl pH 9.5, 1mM Levamisole (a potent lysosomal phosphatase inhibitor) and 0.1% (v/v) Tween 20. The antibody signal was detected using a colour reaction by adding 1.3 µl of NBT (nitroblue tetrazolium salt, 75mg/ml in 70% (v/v) dimethyl formamide) and 0.9µl of X-phospate solution (5-bromo-4-chloro-3-indolyl phosphate, toluidinium salt, 50µg/ml in dimethyl formamide) both supplied in the Boehringer Kit. The colour developed after 2 hours in the dark. The development was stopped by incubating the tissues in PBT. The tissues were mounted in Gurr's Water Mounting Medium and flattned by placing a brass weight on top of the coverslip. The tissues were photographed using Kodak film.

2.4.7 Construction of a Library

The *D. funebris* genomic DNA was extracted using the method described in section 2.4.2.2.1, and HindIII digested. The lambda DNA used was from the immunity phage

NM1149. The DNA was obtained from Prof. N.E. Murray (ICMB, University of Edinburgh) and was digested with HindIII. 6.0 µg of HindIII digested target genomic DNA were mixed with 4.0 µg of HindIII digested lambda DNA, phenol extracted (2.4.1.1), ethanol precipitated (2.4.1.4) and later resuspended in 7.5 µl of dsH_2O. The ligation reaction was performed by adding to the DNA solution 1 µl of 10X ligation buffer, 1 µl of 10mM ATP and 0.1U of T_4 DNA ligase. The tube was spun, incubated at 4°C overnight and the ligation stopped by heating to 70°C for 5 minutes. The tube was allowed to equilibrate to room temperature and later placed on ice. The packaging extracts (obtained from Amersham International) stored at -70°C were thawed slowly on ice as soon as these were thawed the DNA was added to extract A (prepared from *E. coli* strain BHB2688) to which extract B (prepared from *E. coli* strain BHB2690) was immediatel y added and the tube was gently mixed. The tube was spun and incubated at room temperature for 2 hours. After the incubation 0.5 ml of phage buffer and 10ul of chloroform were added. The phage stock was stored at 4°C and later titered (2.4.4.4). A control library was made using DNA provided by the manufactures of the extracts (Amersham International).

2.4.8 Screening of Libraries

Two libraries were screened to isolate yp genes. The first library was made in Lambda NM1149, which has a

unique HindIII and EcoRI cloning sites, and *D. funebris* DNA digested with HindIII. The screening for recombinants in this library was achieved by plating the library with NM514 cells. The second library was obtained from Dr. K. Beckingham (Rice University) and was made by partially digesting the *Calliphora erythocephala* DNA with Sau 3A and ligating into the BamHI site of Lambda EMBL4 (Rubacha *et al.*, 1989). This library was plated with C600 cells.

In primary screens, the libraries were plated in either small plates (90mm)or large plates (245mm X 245mm) with the appropiate plating cells to a density of 5000 pfu (plaque forming units) per small plate or 50000 pfu/large plate. The plates were incubated overnight at 37°C after which they were placed at 4°C for 1 hour, preventing the top layer on the plates from peeling off when taking the plaque lifts (2.4.4.4). The secondary screen of possible positive phages, ie those that may contain the yp gene sequences, was performed on small plates with a phage density below 250 pfus/small plate. The plates were grown and treated in the same manner as above. In some cases further purification of a phage was required for which purposes small plates were used with a phage density not exceeding 50 pfus/small plate. Once the phage was pure a plaque was collected and placed in 1ml of phage buffer with one drop of chloroform. The phage stock was kept at 4°C and used to grow phage to isolate the DNA for further characterisation.

2.4.9 DNA Sequencing by the Dideoxynucleotide Chain Termination Method (Sanger *et al.*, 1977).

Single and double stranded DNA sequencing was used to sequence the two *yp* genes from *Calliphora erythrocephala*. Whilst both procedures are basically the same their differences will be explained for each step of the procedure.

2.4.9.1 Preparation of template DNA

Template DNA for M13 sequencing (single stranded DNA)was prepared as follows. The M13 phage were freshly plated on minimal media plates, a plaque was picked and placed in 1.5ml of 2 x TY medium with 15 ul of a fresh overnight of TG1 cells. The culture was allowed to grow for 51/2 hours, and was then was transferred to an eppendorf tube and spun for 10 minutes at room temperature. The supernatant was tranferred to a new tube and spun again for, a futher 10 minutes. The supernatant was again transferred to a fresh eppendorf tube, 150ul of 20% (w/v) PEG 8000 and 2.5 M NaCl were added, the tube was incubated for 15 minutes at room temperature and spun for 10 minutes. The supernatant was discarded and the tube again spun for 5 minutes in order to get rid of excess PEG used to precipitate the phage particles, which inhibits the sequencing reactions. The supernatant was carefully

taken off, making sure that none was left behind. The pellet was redisolved in 100µl of TE and the solution phenol extracted with 50ul of phenol (2.4.1.1). The mixture was spun for 2 minutes and 80ul of the aqueous solution was transferred to a fresh eppendorf tube and ethanol precipitated (2.4.1.4). The DNA pellet was redissolved in 30µl of TE. 10µl of this solution is required for the annealing of the primer (2.4.9.2).

In double stranded sequencing, before being able to sequence the DNA, it is necessary to denature it. The plasmid DNA was obtained from a maxi-prep as described above (2.4.1.9.2). The amount of DNA required for each sequencing reaction (2-4 µg of pure plasmid DNA) was placed in 8 µl of sdH_2O to which 2µl of 2M NaOH was added, and the solution incubated for 10 to 15 minutes at room temperature. After the incubation, 3µl of sodium acetate pH 4.8, 7µl of sdH_2O and 60µl of chilled absolute Ethanol are mixed in and the tube incubated on dry ice for 15 minutes. The DNA was spun for 10 minutes, washed and resuspended in 10 µl of sdH_2O. The redissolved DNA was then ready for the next step.

2.4.9.2 Sequencing reactions: annealing; primed synthesis

The denatured single stranded templates were annealed to the primer by adding 2µl of annealing buffer (0.1M

Tris-HCl pH 8.0, 50mM $CaCl_2$) and 2µl of primer. The commercial primers were at a concentration of 80µM in water, the primers synthesised by OSWEL were used at a dilution of 0.1 OD units. The mixture was spun and incubated at 65°C for 10 minutes, after which the tube was placed at room temperature for at least 10 more minutes.

The plasmid DNA was treated in the same manner with the exception of incubating the DNA:primer mixture at 37°C for 20 minutes followed by 10 minutes at room temperature.

The annealed templates (single or double stranded) were treated in the same manner. The tubes with primer annealed to the DNA were placed in a rack and 1µl of 0.1mM DTT, 1µl of labelling mix (10mM DDT, 0.25mM of each dNTP), 2µl of T_7 DNA polymerase at 1.5U/µl, and 1µl of ^{35}S-dATP (10µCi/µl) were added, the tube was spun and incubated at 37°C for 5 minutes. To each of four tubes previously labelled T,C,G and A, and containing 2.5µl of dideoxy T,C,G and A nucleotide respectively (see Table 2.4), 2.5µl of the above reaction were added, the tubes were spun and incubated again for 5 minutes at 37°C. The reactions were terminated by the addition of 5µl of termination mix (0.3% (w/v) Xylene cyanol, 0.3% (w/v) Bromophenol blue, 10mM EDTA in 100% deionised formamide). The DNA strands of the polymerised reactions were denatured by boiling for 2 minutes, immediately before loading on to a polyacrylamide sequencing gel (2.4.3.4)

Table 2.4 Composition of ddNTP solutions (all volumes in microliters)

	T mix	C mix	G mix	A mix
0.5 mM dTTP	25	500	500	500
0.5 mM dCTP	500	25	500	500
0.5 mM dGTP	500	500	25	500
10 mM ddTTP	40	-	-	-
10 mM ddCTP	-	6	-	-
10 mM ddGTP	-	-	16	-
10 mM ddATP	-	-	-	0.6
TE	935	969	959	500

CHAPTER 3

RESULTS

Studies on the functional conservation of the Yolk Proteins in *Drosophila* and other Diptera

The yolk proteins (YPs) of different *Drosophila* species are remarkably similar in molecular weight. As described in the Introduction, the YPs appear to be regulated in a similar fashion by ecdysteroids in all *Drosophila melanogaster* species subgroup where this aspect has been examined (Kozma and Bownes, 1986). These observations are extended to other insects (reviewed by Bownes, 1986) and beyond that to *Xenopus* and chickens (reviewed by Wallace, 1985). The interesting question to address is whether the YPs of *Drosophila* have been conserved functionally over *Drosophila* evolution. This in turn will allow the determination of peptide domains of the YPs that are conserved and which have an important function. The yolk proteins have been thought to have few evolutionary constraints because their function has been assumed to be to provide an amino acid supply for the developing embryo.

As stated in the Introduction one major aim of this thesis was to determine the protein domains in the YPs that are required for the yolk proteins to be taken up specifically into the oocyte. The aims of the experiments described in this chapter were to first develop a method to assay functional conservation of the uptake by the oocyte of YPs in other species. Here, functional

conservation is defined as the ability of the foreign yolk protein from a non-*Drosophila melanogaster* species to behave like the YPs of *D. melanogaster* with respect to uptake into the oocyte. Secondly, to assay different *Drosophila* species and find species closely related to *D. melanogaster* whose yolk proteins were not functionally conserved, and likewise to find distantly related species to *D. melanogaster* whose yolk protein uptake was conserved. Comparison of these results will allow the identification of suitable species whose genes may be isolated and characterised.

The Yolk Protein Uptake Assay

In the first instance, to study the yolk proteins for functional conservation the assay has to involve the injection of YPs from one species to another. The solution would be to isolate large amounts of YPs from *D. melanogaster* and inject them into the test species. However to ensure that the assay is devised to evaluate the similarities within the YPs it is necessary to use a single set of conditions which will apply throughout the investigation. For this reason the YPs of different species should be assayed in one species rather than the YPs of a single species being evaluated in different species. One initial problem is that when YPs are separated by SDS-PAGE (Figure 3.1A and C; Table 3.1), some of the species have very similar profiles of yolk proteins

Figure 3.1

Yolk proteins from *Drosophila* and non-*Drosophila* species used in the experiments described in this chapter.

A.Panel of a coomassie blue stained SDS-PAGE in which the ovarian YPs of the *Drosophila* species have been fractionated. nas= *D. nasuta*, lum= *D. lummnei*, lit= *D. littoralis*, leb= *D. lebanonensis*, bus= *D. busckii*, azt= *D. azteka*, ari= *D. arizonensis*, vir= *D. virilis*, psd= *D. pseudoobscura*, hyd= *D. hydei*, fun= *D. funebris*, yak= *D. yakuba*, mau= *D. mauritiana* and mel= *D. melanogaster*.

B.The equivalent gel (as in Figure 3.1A) was western transferred and detected using anti-YP antibodies raised against the *D. melanogaster* YPs and anti-rabbit IgG conjugated to horse raddish peroxidase.

C.Panel of a coomassie blue stained SDS-PAGE containing the ovarian YPs of the *D. melanogaster*, *D. funebris* and non-*Drosophila* species, namely mDm= male *D. melanogaster*, fDm= female *D. melanogaster*, Df= *D. funebris*, Sa= *Sarcophage argyrostoma*, Ce= *Calliphora erythrocephala*, Pt= *Protophormia terrae-novae*, Md= *Musca domestica* and Ls= *Lucilia servicata*.

D.A western blot of the ovarian proteins of the species in Figure 3.1C, detected withthe anti-Yp antibodies raised against the YPs of *D. melanogaster* and the anti-rabbit-IgG antibodies conjugated with horse rad ish peroxidase.

A Coomassie blue

B Western blot

C Coomassie blue

D Western blot

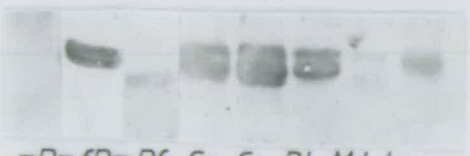


Table 3.1 Number of yolk proteins and their apparent molecular weights in different Dipteran species.

Species	YPs	Molecular Weights (K Daltons)		
D. lebanonensis	3	47.4	46.8	45.0
D. busckii	2	48.2	45.4	-
D. pseudoobscura	3	48.8	47.0	-
D. azteka	3	47.0	45.2	44.7
D. melanogaster	3	47.0	45.7	44.7
D. mauritiana	3	47.0	45.7	44.7
D. yakuba	3	47.0	45.7	44.7
D. lumnei	3	47.4	46.8	45.0
D. littoralis	3	47.4	46.8	45.0
D. virilis	2	47.3	44.7	-
D. arizonensis	2	47.0	44.7	-
D. hydei	3	48.6	47.0	44.7
D. nasuta	3	47.4	46.4	44.5
D. funebris	1	44.0	-	-
Musca domestica	3	47.4	45.6	44.0
Sarcophaga argyrostoma	3	47.2	45.7	44.8
Calliphora erythrocephala	3	46.0	45.7	45.0
Lucilia servicata	3	47.2	45.7	44.8
Protophormia terrae-novae	3	47.0	45.7	45.0

The apparent molecular weight of the yolk proteins were estimated from gradient SDS polyacrylamide gels. This table does not imply that the largest protein is always most similar to YP1 and the next to YP2 ect., of D. melanogaster. This will only be established by sequencing the genes encoding these polypeptides.

and thus cannot be distinguished from one another on the basis of molecular weight alone. The problem can be solved by radiolabelling the proteins with ^{35}S-methionine prior to isolation and investigating their uptake in injected flies using autoradiography. Initially, the isolation of YPs was carried out with ovaries of *D. melanogaster* (2.4.2.1.3) in order to assess the purity of the YP preparation and the viability of using such preparations to assess uptake of foreign YPs into oocytes of a host species. To assess the purity of the YP preparation samples were taken at each of the fractionation steps and subjected to SDS-PAGE. Silver staining of such gels (2.4.3.1) revealed that even though the YPs are enriched many other polypeptides have been co-purified and thus the resultant preparation is relatively impure (Figure 3.2). Moreover, the lyophilised protein was dissolved in 20% glycerol, a medium that proved difficult to inject successfully into flies. The low survival rate was compounded with the insolubility of the YPs at a concentration higher than 5ug/ul. The experiments were repeated using radioactive ovaries from flies injected with ^{35}S-methionine as a source of radiolabelled YPs in the hope of obtaining a cleaner preparation; however, the end result was similar to that obtained before. Therefore, while the method seemed attractive the low purity of the YPs and their insolubility made it unsuitable. A crude extract of some kind may be less labour intensive and as useful as the YPs

Figure 3.2

Silver stained polyacrylamide gel containing samples from the YP fractionation procedure (2.4.2.1.3).

Track 1, sample of female *D. melanogaster* haemolymph proteins.

Track 2, sample of the supernate after centrifugation at 10 krpm for 15 minutes.

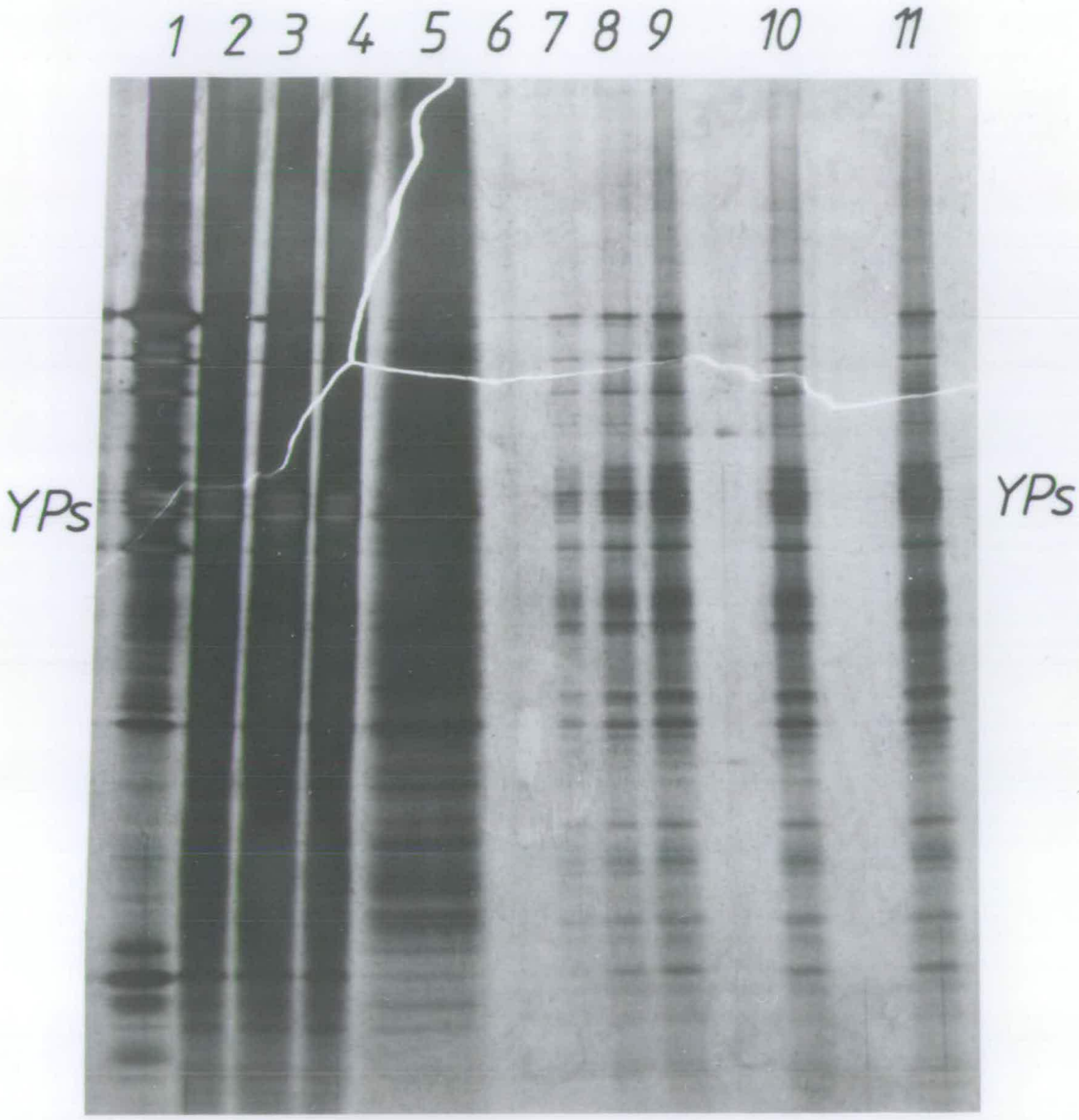
Track 3, sample from muslin filtrate.

Track 4, sample from the glass fiber filter filtrate.

Track 5, sample of supernate after centrifugation at 9 krpm for 10 minutes.

Track 6, sample from dialysed protein suspension.

Tracks 7, 8, 9, 10, and 11 contain 10, 20, 30, 40 and 50ug of protein from the resuspended freeze dried purified protein.


1 2 3 4 5 6 7 8 9 10 11
YPs
YPs

prepared from isolated ovaries.

A different approach was taken, which was to radiolabel the polypeptides in the haemolymph of the donor flies and inject this radiolabelled haemolymph into host flies. Such a method had been used succesfully and described by Giorgi *et al.* (1986). The use of haemolymph has several advantages. The YPs are secreted into the haemolymph from the fat body and therefore are easily radiolabelled. Secondly, while female haemolymph contains other proteins the major components are YPs. Moreover, haemolymph is easily extracted from flies and may be immediately injected into a host (test) fly. Therefore, haemolymph is not only easy to manipulate but it also provides a good method to obtain radiolabelled YPs from different species to test in a host species. The first problem was to find the most efficient way of radiolabelling the donor haemolymph. Two different approaches were taken, namely feeding donor flies with radiolabelled amino acids and injecting donor flies with ^{35}S-methionine.

Previous experience in the laboratory had shown that injection of ^{35}S-methionine into female *D. melanogaster* flies resulted in total depletion of unincorporated ^{35}S-methionine by 4 hours after injection. The initial experiments used the strains Or^{R} and *fs(2)A17/SM1*. The latter was used because the YPs produced in the fat body

are accumulated in the haemolymph as a result of defective ovarian development (Postlethwait and Handler, 1978). The accumulation of YPs in the haemolymph of *fs(2)A17/SM1* therefore offers a good source of radiolabelled YPs if one were to transplant *D. melanogaster* YPs into other species. Female flies of each strain were injected with ^{35}S-methionine and allowed to recover at 25°C for 4 hours. After this period the flies were re-etherised, the haemolymph extracted and immediately injected into a host Or^{R} female fly. This was repeated until 10 to 15 female flies had been injected with radioactive haemolymph. The host flies were then incubated for 18 to 22 hours at 25°C, after which the haemolymph, ovaries and carcasses were collected into Protein sample buffer and stored at -20°C. A similar experiment was carried out with donor flies fed with radioactive yeast (2.4.1.11.2). The results of the above experiments are shown in Figure 3.3, and contrasts the different methods of labelling the YPs for the purposes of transplants. Comparison of the two approaches shows that the radiolabelling of proteins in the haemolymph and ovaries of donor flies is more efficient when the label is injected into the flies. As a result the host ovaries (ie, from the flies which have been injected with the radioactive proteins of the haemolymph) in the transplant from radioactive ^{35}S- methionine injected Or^{R} showed a marked difference with respect to radiolabelled YP uptake when compared to transplants using ^{35}S-methionine fed Or^{R} flies. The results using

Figure 3.3

Autoradiograph of a SDS-PAGE containing samples from Or^R female flies injected with ^{35}S-methionine radiolabelled haemolymph proteins. The donor proteins were radiolabelled by either feeding or injection into either Or^r or *fs(2) A17/SM1*.

Track 1, haemolymph of donor Or^R flies.

Track 2, ovaries of donor Or^R flies.

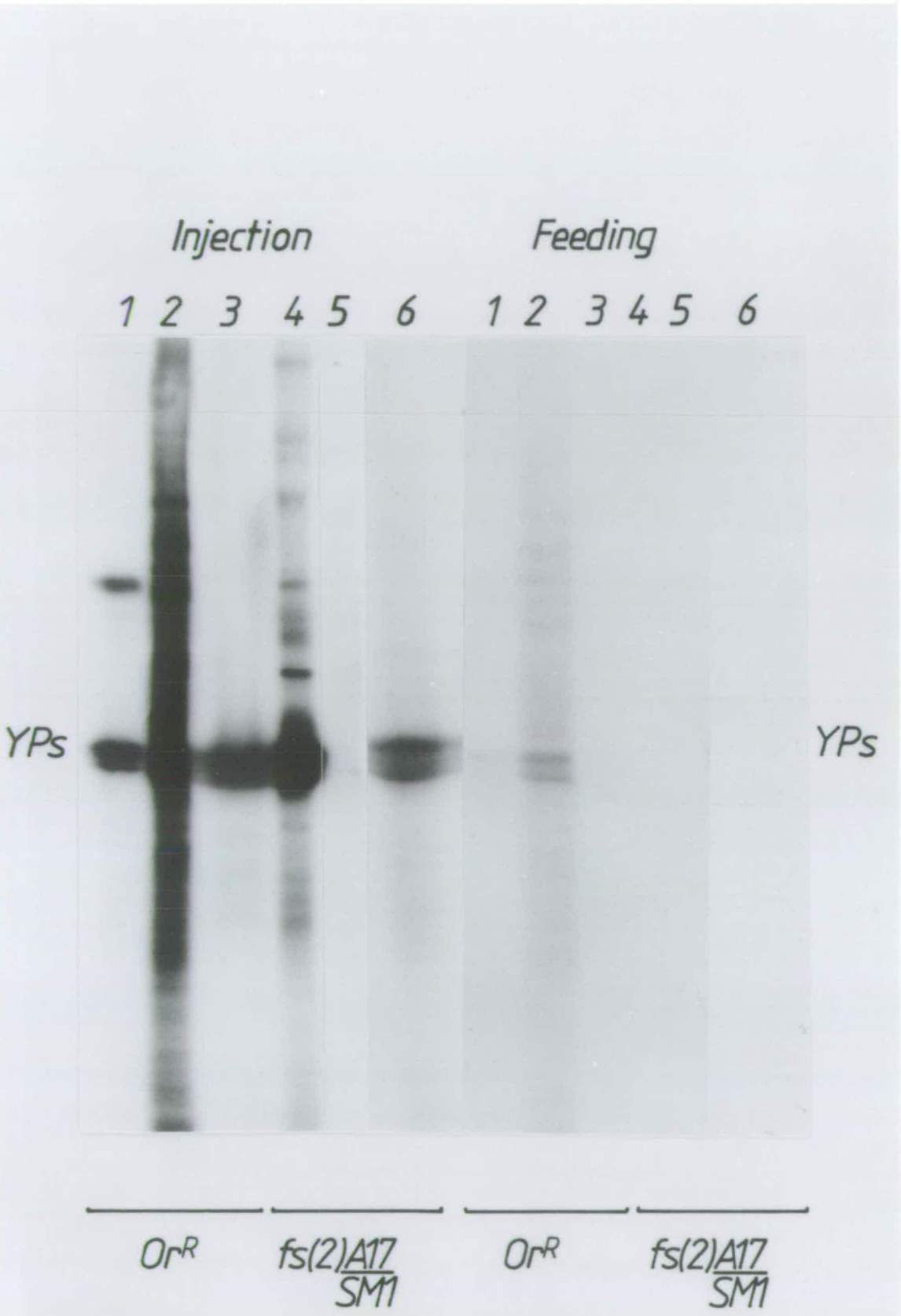
Track 3, ovaries of the host Or^R flies injected with haemolymph from donor Or^R flies.

Track 4, haemolymph of donor *fs(2)A17/SM1* flies.

Track 5, ovaries of donor *fs(2)a17/SM1* flies.

Track 6, ovaries of host Or^R flies injected with haemolymph from donor *fs(2)A17/SM1* flies.

YPs, indicate position of the yolk proteins in the autoradiograph.


Injection
Feeding
1 2 3 4 5 6
1 2 3 4 5 6
YPs
YPs
Or^R
fs(2)A17/SM1
Or^R
fs(2)A17/SM1

fs(2)A17/SM1 female flies as donors of radioactive haemolymph showed that residual free ^{35}S-methionine was available to host female flies resulting in a high background level of newly synthesised protein. Moreover this comfirmed the results obtained with Or^R where injecting ^{35}S-methionine into flies proved to be a better approach to radiolabelling the YPs in the haemolymph. Thus the ^{35}S-methionine injected Or^R female flies provided better haemolymph which in turn resulted in a stronger and clearer result. In any case, *fs(2)A17/SM1* would not be used in the transplant experiments since the aim of the assay is to test the uptake of YPs of different species into the oocytes of *D. melanogaster*.

The haemolymph transplant technique is advantageous to use because it provides a medium in which the YPs are easily radioactivelly labelled. Moreover, although the YPs are relatively pure in the haemolymph, the other proteins present can be used as internal controls to demonstrate the specificity of YP uptake by oocytes. Giorgi *et al.* (1986) showed that YPs obtained in this manner are stable and functionally intact since they are targeted and sorted in the same way as native YPs. Therefore haemolymph transplantation is a good way of injecting YPs into host flies. However, this technique has limitations, namely the amount of YPs injected into the host is variable and generally small. Ideally the best solution would be to purify radiolabelled YPs but the

YPs are difficult to dissolve once purified. For example, *in vivo* the YPs are soluble in the haemolymph but as they are taken up into the oocyte and stored they coalesce into the yolk spheres. The insolubility of YPs then appears to be pH and concentration dependant. Furthermore the purified YPs have not been shown to be functionally intact. Finally the isolation of YPs from many different species would be a labour intensive exercise.

Analysis of Foreign YP uptake into the ovaries of *D.melanogaster*

A. Haemolymph transplants using *Drosophila* species

Species used in the following experiments were chosen to cover a wide range of evolutionary distance within the Drosophilidae family. The flies ranged from *D. melanogaster* to *D. funebris* and are listed in Table 3.1.

The haemolymph transplants were carried out as described in methods and materials (2.4.1.12). The test flies had their haemolymph proteins radiolabelled by injecting 0.2ul (approx) of ^{35}S- methionine. The volumes of haemolymph extracted to transplant into *D. melanogaster* varied between individual flies; moreover, the smaller species of donor flies (e.g. *D. yakuba*, *D. azteka*)

yielded less haemolymph than the larger species (e.g *D. hydei*, *D. funebris*, *D. lummnei*). When the smaller species were used a great deal of care had to be taken with the injection of radiolabel and usually 15 female flies were used, while with flies of the larger species 10 females were used. Figure 3.4, shows results obtained with all the *Drosophila* species studied here. For each species, the samples were collected and separated by SDS-PAGE. The first track shows the haemolymph proteins remaining in the donor flies after the transplants. This track shows the successful labelling of the donor YPs and their profile. The next track shows the haemolymph of the host *D. melanogaster* flies which was used to determine the success of the transplant (ie. whether or not foreign YPs were succesfully injected). The ovaries of the surviving host females flies were loaded onto the third track. The yolk proteins seen in this track are compared to those seen in the haemolymph track of the donor species. The success of the YP uptake assay depends on whether or not the donor YPs are found in the host ovaries and on the lack of other radiolabelled proteins in the same ovaries. Finally, the last track contains the carcasses of the host flies, which are used to determine whether any host proteins have been radiolabelled. Usually, donor YPs are observed in this track and may be explained by the presence of some haemolymph remaining in the carcass. The results in Figure 3.4 (summarised in Table 3.2) show that the transplants were successful, and that the foreign YPs

Figure 3.4

Autoradiograph of a SDS-PAGE containing samples from radiolabelled haemolymph transplant experiments performed on *D. melanogaster* female host flies. The donor species whose haemolymph proteins were injected into *D. melanogaster* were *D. mauritiana*, *D. funebris*, *D. virilis*, *D. lebanonensis*, *D. pseudoobscura*, *D. lumnei*, *D. littoralis* and *D. busckii*, and their names are below the panel showing the particular transplant.

Track 1, haemolymph from ^{35}S-methionine injected doner female flies.

Track 2, haemolymph from host *D. melanogaster* flies 18 to 20 hrs after transplant.

Track 3, ovaries from host *D. melanogaster* flies 18 to 20 hrs after transplant.

Track 4, carcasses of host *D. melanogaster* flies 18 to 20 hrs after transplant.

YPs, indicate the position of the donor yolk proteins in each transplant.

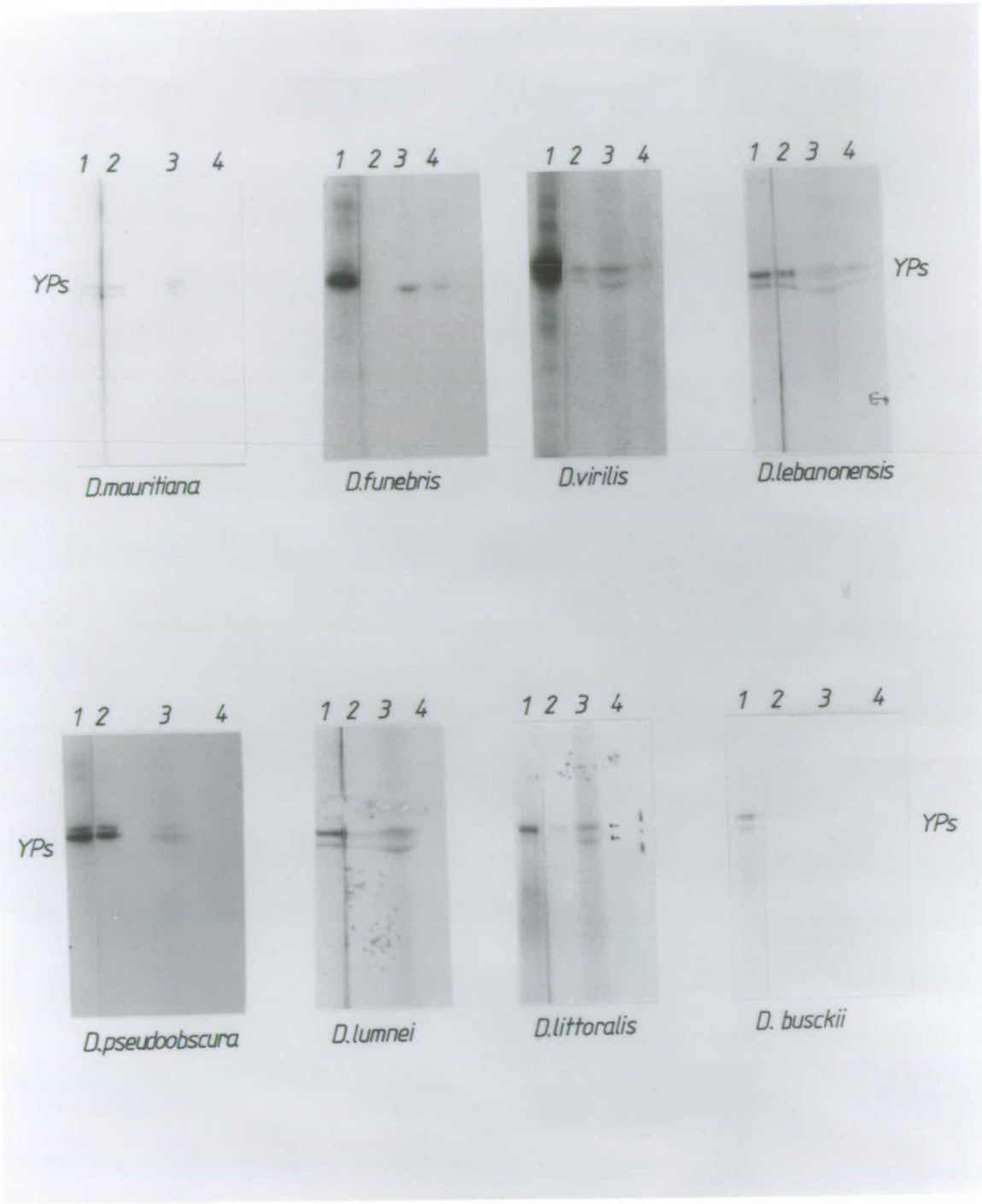

1 2 3 4
YPs
D.mauritiana
1 2 3 4
D.funebris
1 2 3 4
D.virilis
1 2 3 4
YPs
D.lebanonensis
1 2 3 4
YPs
D.pseudoobscura
1 2 3 4
D.lumnei
1 2 3 4
D.littoralis
1 2 3 4
YPs
D. busckii

Table 3.2 Similarity between yolk proteins and conservation of receptor mediated endocytosis amongst Dipteran species

Donor species	Host species	Uptake	immunological cross-reactivity to anti-YP antibodies
D. melanogaster	D. melanogaster	+	+
D. lebanonensis	D. melanogaster	+	+
D. busckii	D. melanogaster	+	+
D. pseudoobscura	D. melanogaster	+	+
D. azteka	D. melanogaster	+	+
D. mauritiana	D. melanogaster	+	+
D. yakuba	D. melanogaster	+	+
D. lumnei	D. melanogaster	+	+
D. littoralis	D. melanogaster	+	+
D. virilis	D. melanogaster	+	+
D. arizonensis	D. melanogaster	+	+
D. hydei	D. melanogaster	+	+
D. nasuta	D. melanogaster	+	+
D. funebris	D. melanogaster	+	+
D. melanogaster	D. funebris	+	+
C. erythrocephala	D. funebris	+	+
L. servicata	D. funebris	+	+
P. terrae-novae	D. funebris	+	+
M. domestica	D. funebris	+	+
S. argyrostoma	D. funebris	+	+

Cross-reactivity was determined by using polyclonal antisera raised against the yolk protein of D. melanogaster. The antibody was used to detect YPs in Western blots.

from *D. mauritiana*, *D. funebris*, *D. virilis*, *D. pseudoobscura*, *D.lebanonensis*, *D. littoralis*, *D. lummnei* and *D. busckii* were taken up by the host ovaries since in all cases foreign YPs were found in the ovaries of the host fly (*D. melanogaster*) and other haemolymph proteins were excluded. This result (summarised in Table 3.2) was unexpected as the yolk proteins have always been assumed to be under limited evolutionary constraints. Moreover, *in vitro* experiments showed that vitellogenins from *Nauphoeta cinerae* bind oocyte membrane preparations of *Leucophaea maderae*, yet the reciprocal experiment did not result in binding of vitellogenins of *L. maderae* to oocyte preparations of *N. cinerae* (Konig and Lanzrein, 1985). In this case the only difference was the glycosylation of each protein, thus illustrating that the receptors are very finely tuned to a specific vitellogenin.

B . Haemolymph transplants using non-*Drosophila* species

The results presented above indicated that the YPs in *Drosophila* evolution are functionally well conserved with respect to their specific uptake and therefore more distantly related species to *Drosophila* were assayed. Huybrechts and DeLoof (1982) showed that *Lucilia servicata*, *Sarcophaga argyrostoma*, *Protophormia terrae-novae* and *Calliphora erythrocephala* have YPs with similar molecular weights to those of *D. melanogaster*, suggesting that the YPs of these non-*Drosophila* species are more related to

the *Drosophila* YPs than to the vitellins of other insects (Figure 3.1C) (as discussed in the Introduction). One of the most striking differences between *Drosophila* and the species named above is size, where one *M. domestica* is equivalent to 10 *D. melanogaster* flies and one *C. erythrocephala* is equivalent to 20 *D. melanogaster* flies. For this reason labelling time was reduced and more label was injected to ensure high efficiency of labelling of the YPs. The transplant experiments were carried out in *D. funebris* since this species has only one YP band in protein gel profiles which migrates at a lower apparent molecular weight than all other YPs tested (Figure 3.1A and C, track fun). Before using *D. funebris* as a host in the haemolymph protein transplants with non-*Drosophila* species, haemolymph polypeptide transplants from *D. melanogaster* (donor) to *D. funebris* (host) were carried out to test the reciprocity of the system and the usefulness of *D.funebris* as a host. The results (Figure 3.5) clearly show that the *D. melanogaster* YPs have entered the ovaries of *D. funebris*, and hence *D. funebris* appears to be a good candidate to use in the haemolymph protein transplant experiments with the non-*Drosophila* species. More importantly, it shows that the YPs and YP receptors between these two species are well conserved in function and probably in structure.

Figure 3.5

Autoradiograph from a SDS-PAGE containing samples from a transplant experiment with radiolabelled *D. melanogaster* haemolymph proteins into *D. funebris* host flies.

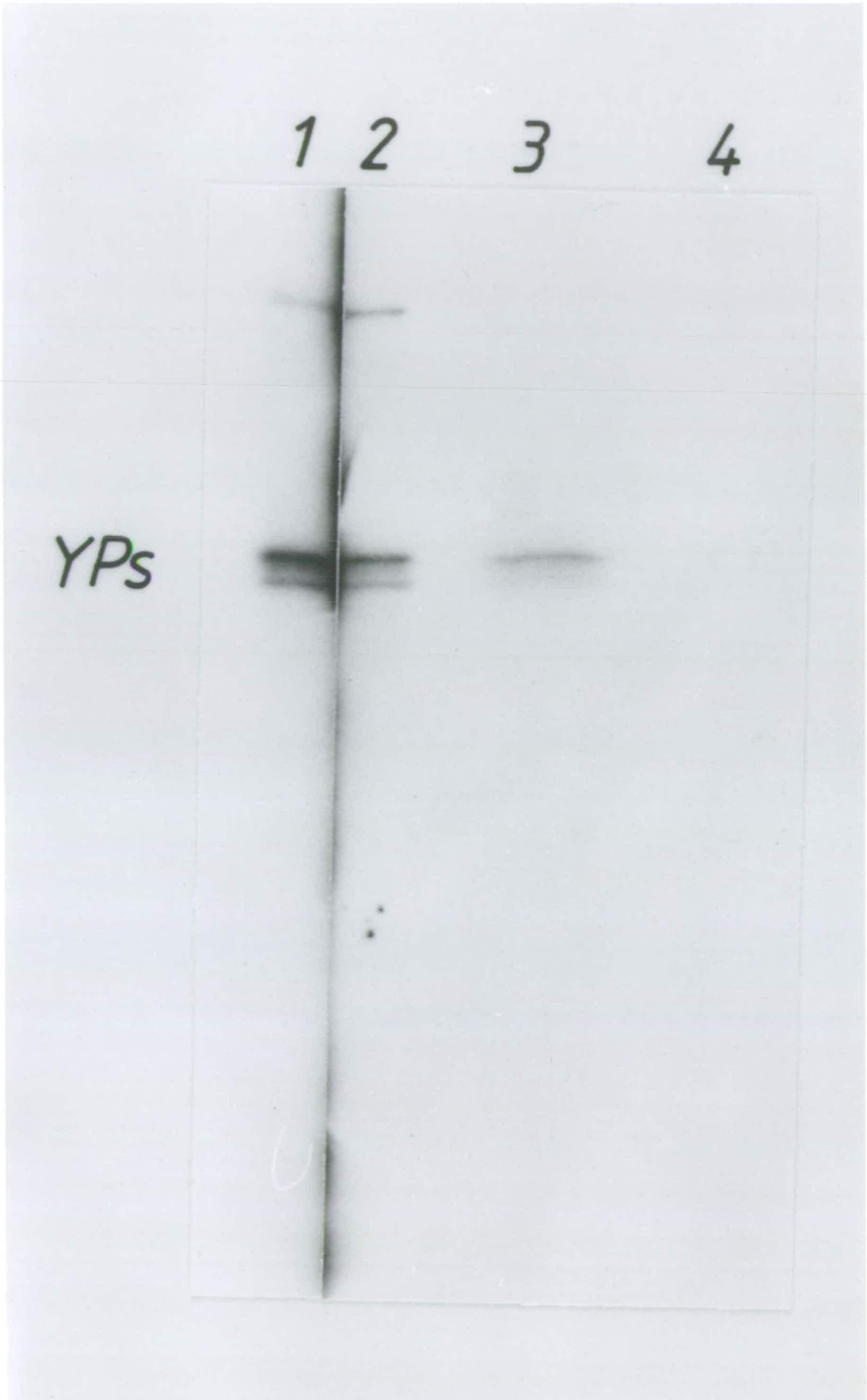
Track 1, haemolymph proteins from *D. melanogaster* donor flies 4hrs after being injected with ^{35}S-methionine.

Track 2, haemolymph from *D. funebris* host flies 20 hrs after injection with *D. melanogaster* radiolabelled haemolymph proteins.

Track 3, ovaries from *D. funebris* host flies 20 hrs after injection with *D. melanogaster* radiolabelled haemolymph proteins.

Track 4, carcasses form *D. funebris* flies 20 hrs after injection with *D. melanogaster* radiolabelled haemolymph proteins.

YPs, indicate the position of the *D. melanogaster* yolk proteins in the autoradiograph.


1 2 3 4
YPs

The experiments carried out with the non-*Drosophila* species were essentially the same as those done with the *Drosophila* species. The only changes were to decrease the labelling time to one and a half hours and double the amount (ie. instead of using 0.2uCi/fly, 0.4uCi/fly was used) of label being injected. Even though great care was taken to make sure that the donor flies were vitellogenic, the ovaries of the donors were dissected out after their haemolymph had been extracted to ensure that the host flies had received haemolymph from vitellogenic flies. Finally, only 3 or 4 female flies were injected since they yielded sufficient haemolymph to transplant into 10 or 15 host flies. The volumes of labelled haemolymph injected into *D. funebris* hosts were generally larger (ie. haemolymph volumes were estimated to be doubled) than when the *Drosophila* haemolymph transplant experiments were performed. Figure 3.6 shows the results obtained with *Sarcophaga argyrostoma*, *Calliphora erythrocephala*, *Lucilia servicata* and *Protophormia terra-novae*. Again in all cases the presence of foreign YPs in the ovaries of *D. funebris* is seen (Table 3.2) indicating that the uptake of non-*Drosophila* YPs takes place in oocytes of *D. funebris*.

In order to determine whether or not the foreign YPs are targeted into the oocyte and not just attached to the ovary or concentrated in the interfollicular spaces, eggs from *D. funebris* transplanted with labelled haemolymph proteins from *Calliphora* and *Sarcophaga* female flies were

Figure 3.6

Autoradiographs of a polyacrilamide gel containing samples from haemolymph transplant experiments using the non-*Drosophila* species as donors of haemolymph. The host species was *D. funebris*. Under each panel is the name of the donor species, namely *C. erythrocephala*, *L. servicata*, *P. terrae-novae* and *S. argyrostoma*.

Track 1, haemolymph from donor species 1 1/2 hrs after injection with ^{35}S- methionine.

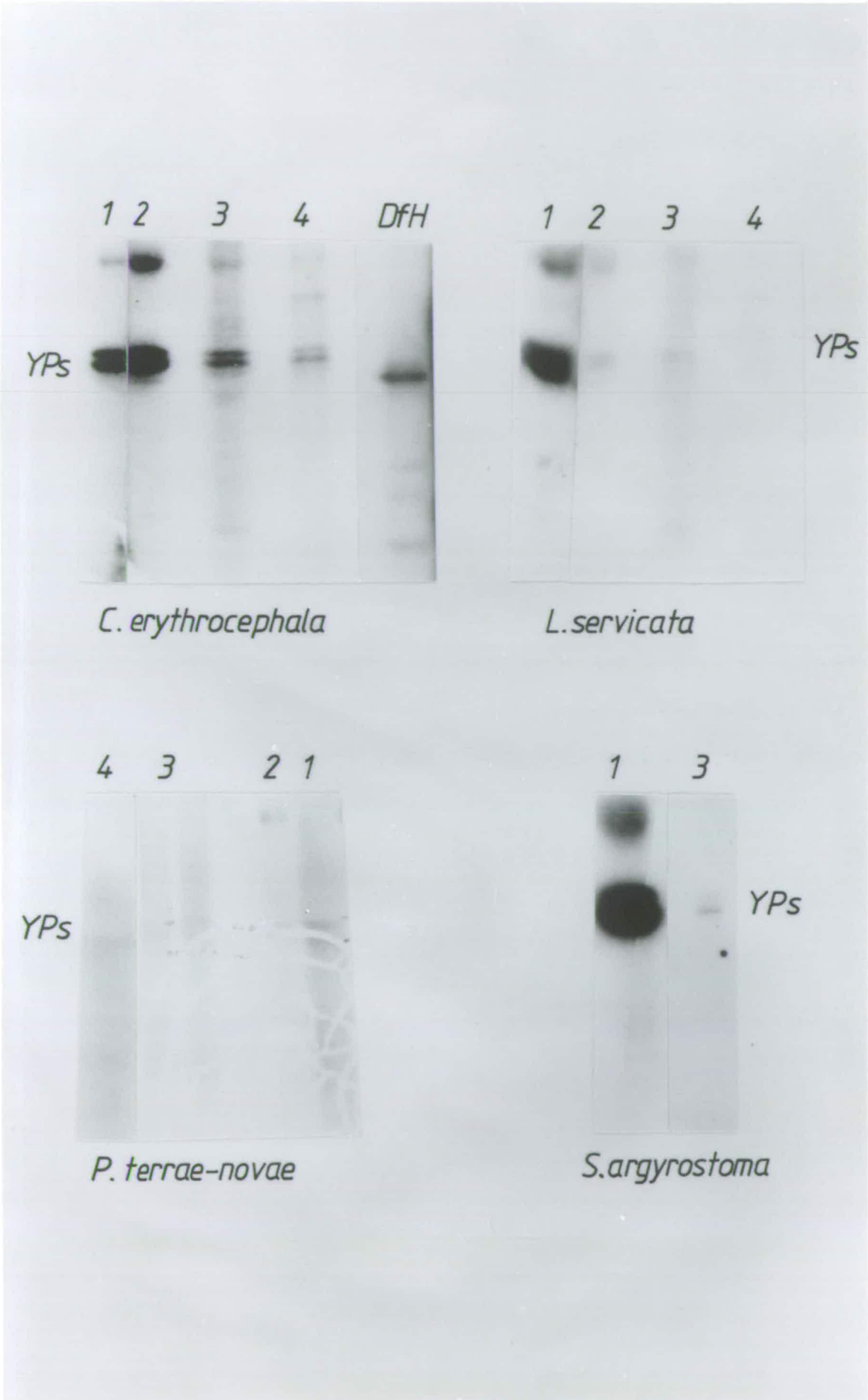
Track 2, haemolymph from *D. funebris* 20 hrs after donor haemolymph transplant.

Track 3, ovaries from *D. funebris* 20 hrs after haemolymph transplant.

Track 4, carcasses from *D. funebris* 20 hrs after haemolymph transplant.

Track DfH shows the radiolabelled haemolymph of *D. funebris*.

YPs, indicate the yolk proteins of the donor species in the autoradiograph.


1 2 3 4 DfH
YPs
C. erythrocephala
1 2 3 4
YPs
L. servicata
4 3 2 1
YPs
P. terrae-novae
1 3
YPs
S. argyrostoma

collected. The experiment was carried out as previously described; however, the host flies were incubated for 48 to 60 hours and the layed eggs collected every 12 hours. The host flies were placed in the presence of male flies to make certain of the fertility of the female flies which in turn would result in egg laying. The eggs were first dechorionated to ensure that none of the yolk proteins seen on the protein gel could be ascribed to yolk protein trapped between the vitelline and chorion membranes. The results (Figure 3.7) demonstrate the deposition of the foreign YPs in the host oocytes, proving that they are taken up by the oocyte.

On the Specificity of the uptake of the Foreign YPs by oocytes

The results reported in the previous sections show a high conservation of the uptake system of YPs through the evolution of *Drosophila* and non-*Drosophila* used in these experiments. In order to ensure that the uptake of the foreign YPs was due to receptor mediated endocytosis and high selectivity of the receptor to YPs, and not to non-specific uptake (ie, fluid phase pinocytosis), the following experiments was carried out. Hints of the selectivity of uptake are observable in Figures 3.5, 3.6 and 3.7, whereby in the injected haemolymph other proteins beside the YPs are observable, yet in the ovaries of the host flies only the YPs are seen. However, it is

Figure 3.7

Autoradiograph from a SDS-PAGE containing samples from either *C. erythrocephala* or *S. argyrostoma* haemolymph transplants into *D. funebris* host flies. The host *D. funebris* flies were incubated for 60 hrs and the layed eggs collected every 12 hrs.

Track 1, haemolymph proteins from donor species radiolabelled with ^{35}S-methionine.

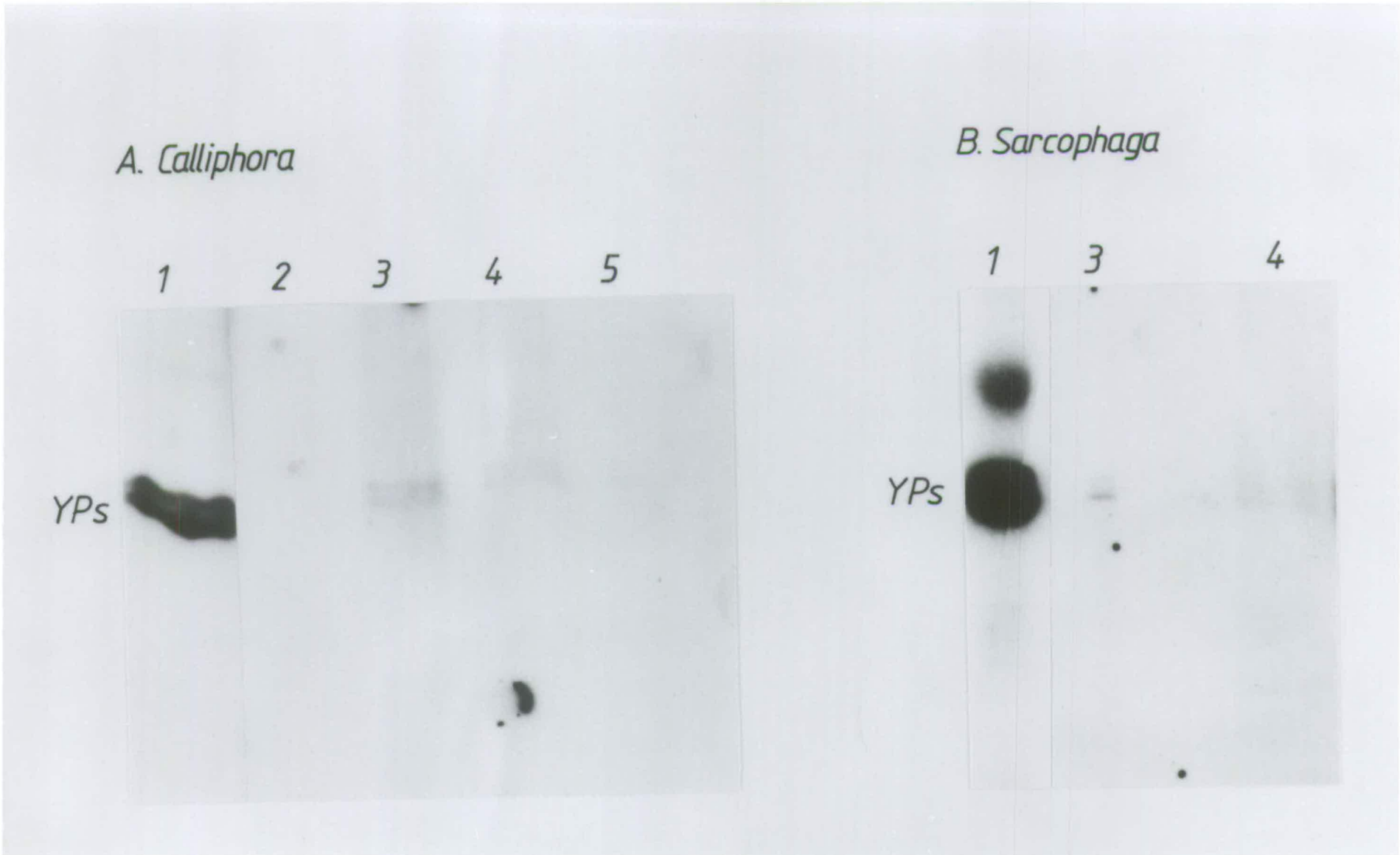
Track 2, haemolymph proteins of *D. funebris* (host flies) 60 hrs after haemolymph transplant.

Track 3, ovaries of *D. funebris* 60 hrs after haemolymph transplant.

Track 4, layed and dechorionated eggs of *D. funebris* 60 hrs after haemolymph transplant.

Track 5 carcasses of *D.funebris* 60 hrs after haemolymph transplant.

YPs, indicates the position of the donor yolk proteins in the autoradiograph.


A. Calliphora
1
2
3
4
5
YPs
B. Sarcophaga
1
3
4
YPs

important to note that the levels of these other labelled proteins in the donor haemolymph are lower than that of the YPs (ie. up to 80% of the labelled protein in the haemolymph is YPs), therefore it is difficult to assess their uptake into ovaries. To further test this, proteins in *D. melanogaster* male haemolymph were radiolabelled for 8hrs and then transplanted into *D. melanogaster* female flies. After an overnight incubation at 25°C, the haemolymph, ovaries and carcasses were collected and separated by SDS-PAGE. Figure 3.8 shows the male haemolymph proteins (track 1), and the female haemolymph proteins after the injection and incubation (track 2), the same protein profile is observed in both cases. The carcasses of the host female flies (track 4), contain only a low background of radiolabelled protein synthesis. The ovaries of the female flies (track 3) do not seem to show any proteins common to the male haemolymph. This experiment was carried out to test proteins that were commonly available in both male and female haemolymphs, with the exception of the YPs, for their uptake into the ovaries of females. The results showed that it was unlikely that proteins other than YPs in the haemolymph would be taken up selectively into the ovaries.

The limitation of the above experiment was that YPs form a large proportion of the protein in the haemolymph and hence it was important and interesting to test the uptake of a another abundant protein into the ovaries of

Figure 3.8

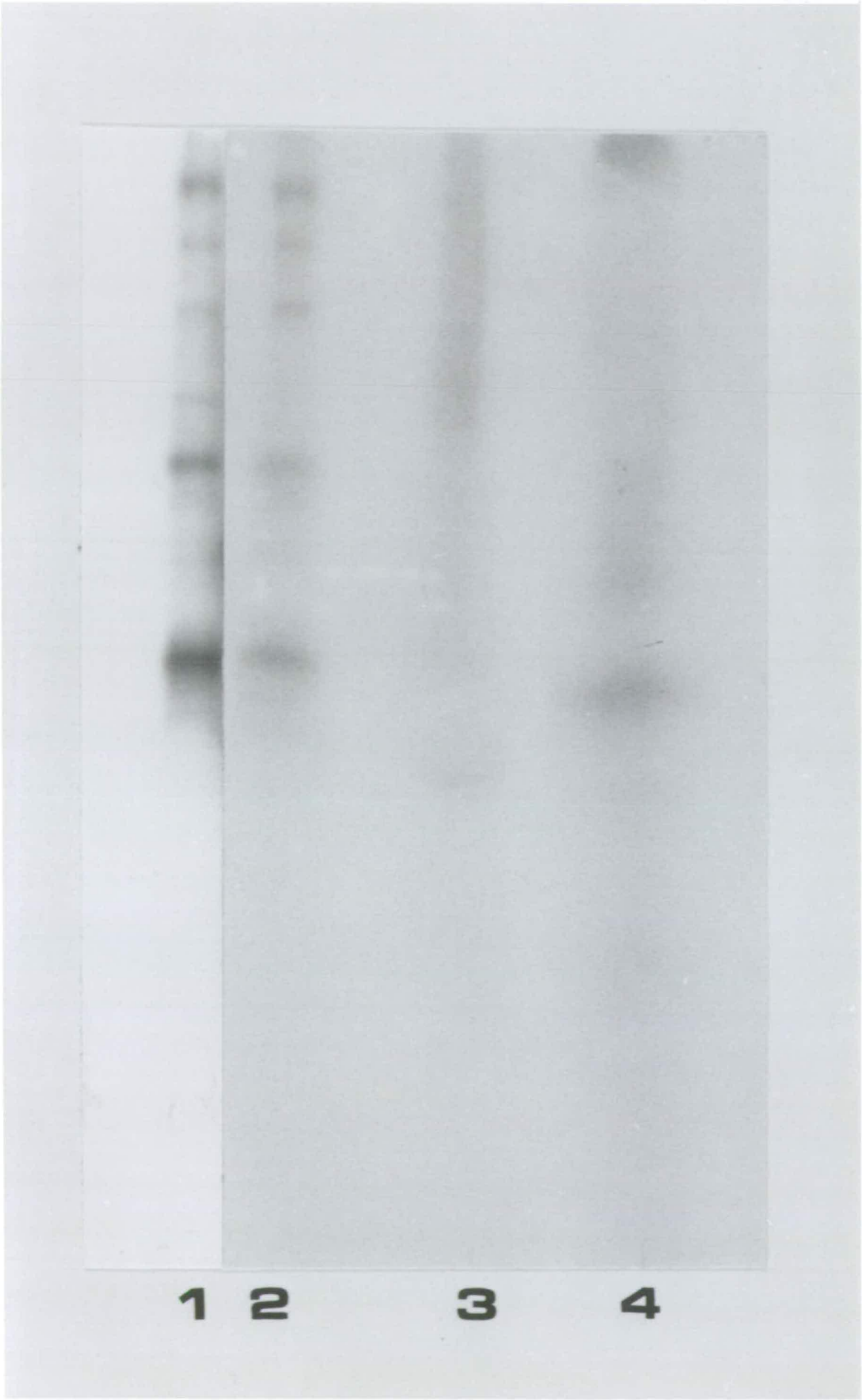
Autoradiograph from a SDS-PAGE containing samples from male *D. melanogaster* haemolymph transplant into female *D. melanogaster* host flies. The donor male flies were injected with ^{35}S-methionine and incubated for 8 hrs prior to haemolymph extraction and transplantation into female host flies.

Track 1, radiolabelled male *D. melanogaster* haemolymph proteins.

Track 2, haemolymph from female *D. melanogaster* 20 hrs after male haemolymph injection.

Track 3, ovaries from female *D. melanogaster* 20 hrs after male haemolymph transplant.

Track 4, carcasses from female *D. melanogaster* 20 hrs after male haemolymph transplant.


1 2 3 4

female *D. melanogaster* flies. Readily purified sources of *Drosophila* haemolymph proteins are not available. Therefore, to this end bovine serum albumin (BSA) was chosen because it is a foreign protein that is available commercially and which dissolves easily to make highly concentrated solutions. Initially, BSA was injected into the haemolymph of *D. melanogaster* flies which were then incubated for 0, 8 and 20 hrs before isolating haemolymph, ovaries and carcasses. The polypeptides in these samples were fractionated by SDS-PAGE and then Coomasie blue stained (2.4.3.1). Although the BSA was quite clear in the haemolymph, no conclusive evidence could be obtained as to whether or not any BSA had been taken up into the ovaries due to a number of co-migrating polypeptides in the region where BSA was expected. A commercial source (Sigma) of a monoclonal antisera against BSA was found and the experiments were repeated. The experiment was carried out using 10ug/ul of BSA in Ringer's solution. The injection of this solution (0.2 ul apprx.) resulted in an amount of protein comparable to that being injected in the haemolymph transplant experiments. In figure 3.9A, the results from the experiment in which 10ug/ul solution of BSA was injected into the haemolymph of female *D. melanogaster* flies are shown. The flies were incubated for 5, 8 and 24 hours post-injection. The antibody is shown to be highly specific to BSA since *Drosophila* protein (ie. Time 0hrs) from haemolymph, ovary and carcass remain undetected. At

Figure 3.9

Injection of BSA into *D. melanogaster* female flies. The BSA was detected using an anti-BSA monoclonal antibody and anti-mouse IgG antibody conjugated to acid phosphatase.

A.A western blot containing samples from the injection of 10ug/µl BSA solution into *D. melanogaster* flies. Samples were collected at 0, 5, 8 and 24 hrs after injection, separated by a SDS-PAGE and transferred onto nitrocellulose (2.4.4.1). The arrow heads indicate extra bands recognised by the anti-BSA monoclonal antibody in the ovary lanes.

H= haemolymph.

O= ovaries.

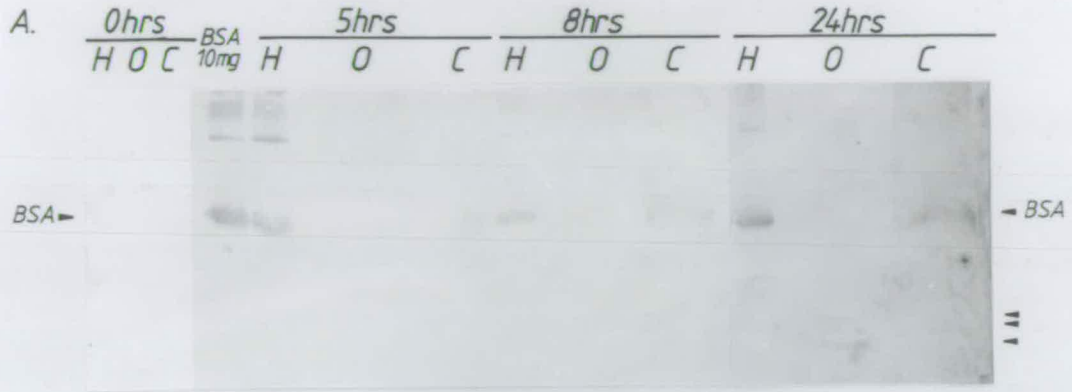
C= carcasses.

B.A western blot containing samples from the injection of 100ug/µl BSA solution into *D. melanogaster* female flies. Samples were collected at 0, 8 and 24 hrs. The arrow heads indicate bands not seen in the time 0 hrs nor in the track loaded with 100µg of BSA.

H= haemolymph.

O= ovaries.

C= carcasses.


A.
0hrs
H O C
BSA 10mg
5hrs
H O C
8hrs
H O C
24hrs
H O C
BSA
BSA


B.
8hrs
H O C
24hrs
H O C
Control
O
BSA 100mg
BSA
BSA

5, 8 and 24 hours the haemolymph and carcass tracks show a band comigrating with the BSA band of the control track (ie, 10ug BSA control); however, in the 5 and 8 hour ovary tracks the same band is not observed. Although at 24 hours the ovaries have an apparent lack of BSA, bands migrating near the bottom of the gel can be observed. These bands may be the products of BSA degradation since they are recognised by the monoclonal antibodies against BSA. To further investigate the fate of BSA levels in the ovaries, a solution containing 100ug/ul was injected and the flies incubated for 8 and 24 hours. Figure 3.9B shows that at 8 and 24 hours post-injection the haemolymph and carcass samples contain BSA. Important to note is the fact that less BSA appears to be present at 24 hours in haemolymph and carcasses, than at 8 hours. This result may be due to some degradation of BSA in the fly or less protein being injected into the 24 hours set of flies. The 8 hours ovaries show a band comigrating with BSA hence showing some BSA uptake; however, although they are faint other bands present (indicated by the arrow-heads) which may indicate products of degradation. The 24 hour ovaries do not appear to be accumulating BSA as would be expected if the BSA was entering the oocytes by receptor mediated endocytosis. Furthermore, the low molecular weight bands observed in 8 hour ovaries are also present but do not appear to have accumulated. Therefore, the depletion of BSA in the 24 hour haemolymph and carcasses may be a reflection of BSA degradation taking

place in the ovaries of the injected flies.

The BSA experiments demonstrate that BSA can enter into the ovaries, a result that is exacerbated at high concentrations of BSA in the haemolymph; however, this process of uptake does not mirror the process of YP uptake into the ovaries since BSA is not accumulated in the ovaries (Figure 3.9A and B). Therefore the most likely route of BSA entry into the ovaries is by fluid phase pinocytosis.

To investigate whether the foreign YPs accumulate in the ovary of *D.melanogaster* host or are degraded soon after uptake as observed with BSA, female *D. funebris* female flies were injected with ^{35}S-methionine and the haemolymph polypeptides labelled for 4 hours; their haemolymph was transplanted into *D. melanogaster* female flies which in turn were incubated for 2 and 24 hours at 25°C. Haemolymph, ovaries and carcasses were collected and the proteins separated by SDS-PAGE. The results (Figure 3.10) illustrate that, at 2 hours, haemolymph and carcass tracks have more labelled yolk protein than the ovaries at the same time point. A dramatic change is seen at 24 hours where haemolymph and carcasses have very little, if any, YP while the ovaries contain a large amount of radiolabelled YP. An important observation is that no other radiolabelled haemolymph proteins have accumulated in the ovaries. This experiment shows that

Figure 3.10

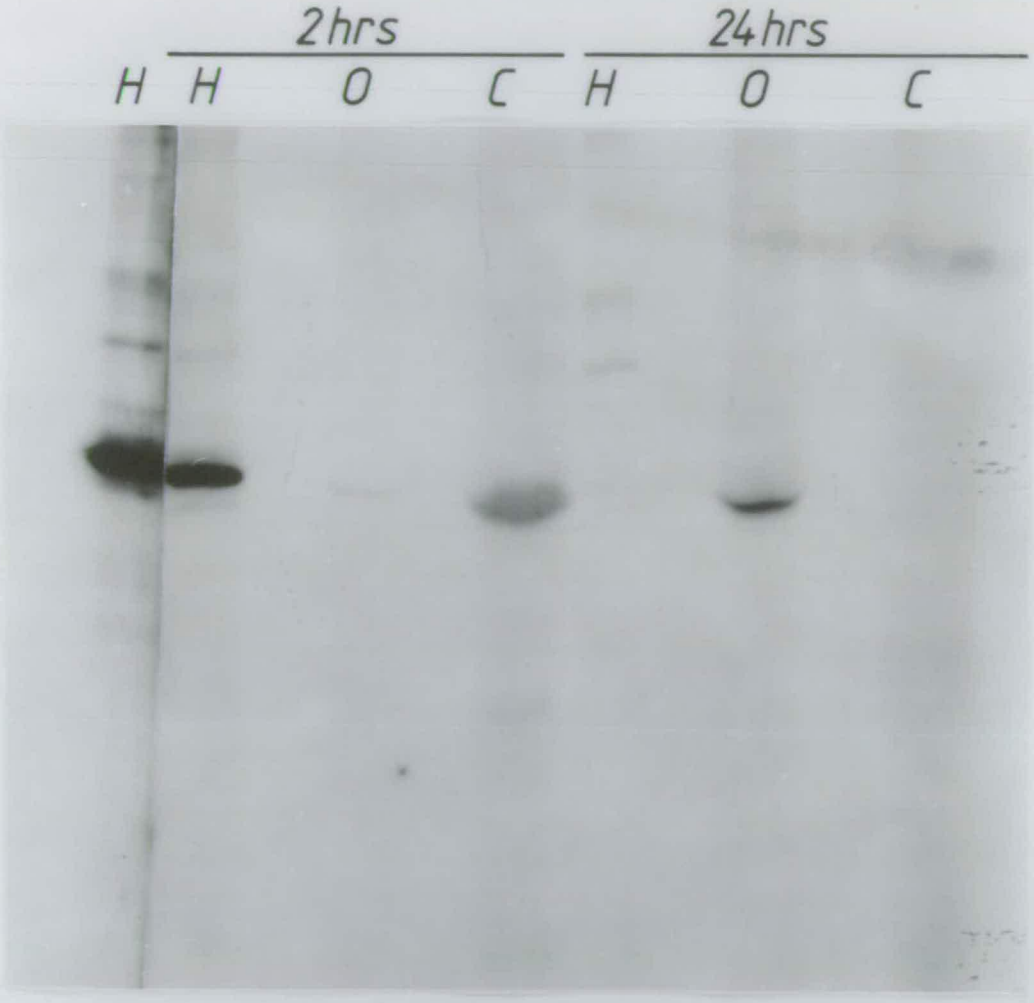
Autoradiograph of a polyacrilamide gel containing samples from haemolymph transplant carried out from *D. funebris* (donor) to *D. melanogaster* (host) flies. The samples were collected from *D. melanogaster* 2 and 24 hrs after injection of radiolabelled *D. funebris* haemolymph proteins.

H=*D.melanogaster* haemolymph at 2 or 24 hrs after injection of labelled donor haemolymph.

O=*D.melanogaster* ovaries at 2 or 24 hrs post-injection with labelled donor haemolymph proteins.

C=*D.melanogaster* carcasses at 2 or 24 hrs after injection with labelled donor haemolymph.

The left hand side track labelled H is haemolymph from *D. funebris* donor flies.


2hrs
24hrs
H H O C H O C

the foreign labelled yolk proteins are accumulated over time in the ovaries and not degraded. Conversely, other labelled proteins present in the haemolymph do not enter the ovaries at all. These results contrast with those obtained with BSA demonstrating that foreign YPs enter the oocyte specifically (ie. via receptor mediated endocytosis) resulting in their storage.

Investigation of the YP cross-reactivity to *D. melanogaster* YP antibodies

Experiments in which the ovarian yolk proteins of all *Drosophila* and non-*Drosophila* flies were tested for cross-reactivity to the polyclonal rabbit antibody raised against the yolk proteins of *Drosophila melanogaster* were carried out. These antibodies have previously been shown to cross-react with the YPs of *D. melanogaster*, *D. yakuba*, *D. virilis*, *D. erecta*. *D. mauritiana*, *D. orena*, *D. teissieri* and *D. simulans* (Kozma and Bownes, 1986). The ovaries of the species used in the experiments described above were dissected and placed in polypeptide buffer and immediately frozen at -20°C. The samples were fractionated by SDS-PAGE and a western blot performed on the gel (2.4.4.1). The YPs were detected with the aid of a rabbit polyclonal antibody against the YPs and an antirabbit IgG antibody conjugated to Alkaline phosphatase. All species tested cross-reacted with the anti YP antibodies (Figure 3.1 B and D). The quality of

the bands in the *Drosophila* tracks in the western blot are not as sharp as they may be because the reaction was allowed to continue for as long as possible to maximise the chance of detection of the YPs from the different species. However, the equivalents to YP_2 and YP_3 of *D. melanogaster* were not detected in *D. lebanonensis* nor was the YP_3 of *D. lummnei*, *D. nasuta* and *D. Hydei* (ie. these are normally seen on Coomasie blue stained gels as shown in Figure 3.1A, tracks leb, lum, nas and hyd). In the case of *M. domestica* the weak signal was due to underloading of the track (shown by Ponceau S staining of the western blot prior to antibody detection, and by equivalent Coomasie blue stained gel in Figure 3.1 C, track Md) rather than a weak interaction between YP and anti-YP antibodies. The cross-reactivity of the YPs of all species, with the noted exceptions, to the *D. melanogaster* YP antibodies is an indication of how well the YPs are conserved between the tested species.

Discussion

The yolk proteins of the species studied here were shown to have a similar pattern of migration after fractionation by SDS-PAGE (Figure 3.1A and C). This confirms the similarity of the molecular weights of the proteins, and therefore their usefulness in the experiments performed. The results obtained via haemolymph transplants showed that the YPs appear to be

more conserved than previously assumed and further lend support to the idea that these proteins are under greater evolutionary constraints than those expected for polypeptides whose sole purpose is providing fuel for embryogenesis.

The method chosen to assay the specific uptake of foreign YPs into the oocyte of the host species (*D. melanogaster* or *D. funebris*) involved injecting haemolymph from a donor species (containing a millieu of proteins) into the host species. As demonstrated in Figure 3.4, and 3.7 (summarised in Table 3.2) the foreign YPs were found in the ovaries of the host. At this time, it is important to note that such an assay cannot be performed in flies whose YP synthesis has been halted because this leads to the arrest of oogenesis. Furthermore, there is a lack of mutants which fail to synthesise YPs and have functional ovaries, and therefore the injected foreign yolk proteins will be competing with constantly high amounts of the host's YPs. This explains the presence of residual donor yolk proteins in the carcass and the haemolymph of the host after the overnight incubation. The fate of the foreign YPs was found to be within the oocyte as shown in Figure 3.7, thus demonstrating that the foreign YPs were targeted to the oocyte (as the native YPs) and not accumulated in between the chorion and the vitelline membrane. It is important to note that in all cases the uptake system was selective

to YPs since other labelled proteins present in the haemolymph were not found in the oocyte. Two further experiments showed that the male haemolymph proteins and BSA failed to be accumulated in the ovaries. In fact, in the case of BSA it was observed that injecting a solution of 100ug/ul resulted in uptake but that the protein was not accumulated. These results are similar to those observed by Raikhel and Lea (1986) where horse-rad ish peroxidase (HRP) is not accumulated in the yolk spheres of mosquito oocytes but appears to be segregated into exclusion bodies whose fate may be the lysosomes. In the experiments shown here, there are a number of smaller polypeptides recognised by the Anti-BSA monoclonal antibody (Figure 3.9 A and B). It is suggested therefore, that these polypeptides are the result of BSA degradation in the oocyte. The BSA experiments pose the question of whether or not the uptake of the foreign YPs in *D. melanogaster* oocytes results in accumulation of the YPs. Further experiments (Figure 3.10) demonstrated the accumulation of the foreign YPs in the ovaries of the host, but more importantly it also showed that the uptake was selective to the YPs (as the other haemolymph proteins remained in the haemolymph) and that degradation of the foreign YPs in the host ovaries was undetected.

A point which should not be overlooked is that the foreign YPs may enter the oocyte by forming complexes with the native YPs. As described in the Introduction, the YPs

in *D. melanogaster* are known to form quaternary structures whose composition is not known because of the instability of the complexes (Fourney *et al.*, 1982). If these structures are unstable there may be equilibrium of free YPs and oligomers in the haemolymph. Upon the addition of foreign YPs these may be incorporated into new complexes forming heteromers (YPs of different species, i.e. donor plus host YPs) which may result in uptake. However, for this to take place, the foreign YPs need to be similar in three dimensional structure to the native YPs so as to cooperate and interact together to form a structure suitable for uptake. Thus such a phenomenon would also indicate functional conservation of the YPs.

The methodology used in these experiments has a number of short falls. The method can only be used to detect an all or none uptake of the foreign YPs. It can be argued that for the purposes of this thesis that was all that was required since the aim was to find foreign YPs which were either endocytosed or not. However, it is probable that the efficiency of the uptake of the foreign YPs into the host ovaries is not identical. In view of the results described in this chapter it would be interesting to determine the binding constants for each set of foreign and native YPs to the receptor in the oocyte. However, this would first entail the development of a procedure to purify YPs to homogeneity. Secondly, an *in vitro* binding system using purified oocyte membranes to

test the purified YPs would have to be developed. This methodology has been used widely in other insects (Konig and Lanzrein, 1985; Osir and Law, 1986; Rohrkasten and Ferenz, 1986; Konig *et al.*, 1988) yet in *Drosophila* it has not been reported. In this laboratory the procedure has been tried and is still under development (Bownes and Howden, personal communications). The development of such a method would enable purified YPs to be tested on different oocyte membrane preparations or purified YPs from different species to be tested on *D. melanogaster* oocyte preparations. In the long term such a method can be used to test the binding properties of peptides designed from data obtained in sequence comparisons between the YPs of *D. melanogaster* and other species.

Quantitative measurements relating the efficiency of uptake of the foreign YPs to that of the native YPs can not be made using the method of haemolymph injections. The amount of YPs being injected is unknown since not every individual of a population will be producing the same amount of YPs and secreting them into the haemolymph. The amount of protein injected can vary from 1 to 5ug of which the largest proportion will be YPs. The variability in the amount of protein injected is compounded by the actual volume of haemolymph being injected so that an individual fly may yield enough haemolymph to inject two hosts whereas another individual in the same experiment may yield just enought to inject a single host.

The biggest problem, however, is that the injected foreign YPs will be competing with native YPs in the host fly. The competition and therefore the uptake of the foreign YPs will be dependant on similarities between the foreign and native YPs. Thus the efficiency of uptake will not only depend on the foreign YP interaction with the receptor but also on the stability of that interaction, whereby if the injected YPs are too dissimilar it will reduce the uptake even more. It may be argued that for the purposes of this thesis the assay is very suitable since in order to get uptake of foreign YPs these must share a large amount of similarity with the native YPs of the host. On the other hand, the presence of native YPs may enhance the probability of foreign YP uptake into the host ovaries via heteromer formation (as discussed earlier). In order to resolve this impass, the development of an *in vitro* system in which competitor protein can be added to study the effects on foreign YP interactions with the host receptors is needed. Although the interactions of the YPs with the receptor are not a direct index of endocytosis, it may indicate that the likelihood of endocytosis taking place is higher if the binding of the foreign yolk proteins to the receptor is not easily competed out by non-specific protein (ie, BSA).

Future experiments

The data presented in this chapter suggests that the foreign YPs may be sorted and targeted into the yolk granules in the oocyte along with the native YPs. The experiments with the non-*Drosophila* species have been repeated, the ovaries collected and thin tissue sections made to analyse them using autoradiography and microscopy to elucidate the final fate of the injected YPs. These experiments are being carried out in collaboration with Prof. F. Giorgi at the University of Pisa. Although the experiments were undertaken a year ago, the results of the analysis have not yet arrived from Pisa. A different approach may be taken to answer the same question in which the ovaries or eggs collected from flies transplanted with radiolabelled haemolymph proteins can be homogenised and the resulting homogenate separated via a sucrose gradient. The sucrose gradient is used to fractionate the different sorting compartments. This would allow the labelled protein to be followed though different stages of sorting until the eventual accumulation in the yolk granules. This approach was used successfully by Opresko *et al.* (1980) in *Xenopus* ovaries. The end result of both experimental procedures is the same and will require the fractionation of haemolymph, ovaries, eggs and carcass protein samples by SDS-PAGE to ensure that the labelled proteins being observed or followed (ie, depending on the method used, microscopy or sucrose gradient fractionation)

are the foreign injected YPs and not host proteins.

Conclusions

In summary, the experiments described in this chapter indicate that foreign YPs injected into the haemolymph of a host (*D. melanogaster* or *D, funebris*) accumulate in the host ovaries. This process of uptake and accumulation is specific to YPs and suggests that the YPs within *Drosophila* and *Musca, Calliphora, Sarcophaga, Lucilia* and *Protophormia* are conserved through their evolution.

The results obtained appear to complement other existing data. The molecular weights of the YPs in all the species used are within the same range (Figure 3.1A and C). Moreover, the YPs of all species cross reacted with the polyclonal antisera raised against the YPs of *D. melanogaster* (Figure 3.1B and D). In all species studied, the YPs are synthesised in the same sites (ie fat body and ovaries) as in *D. melanogaster* (Bownes, 1980; Barnett *et al.*, 1980; Isaac and Bownes, 1982; Huybrechts *et al.*, 1982). Synthesis of the YPs is induced in males by 20-hydroxyecdysone in *D. melanogaster* (Postlethwait *et al.*, 1980) and in sibling species (Kozma and Bownes, 1986) as well as in *Musca* (Adams *et al.*, 1985), *Sarcophaga, Lucilia, Phormia* and *Calliphora* (Huybrechts *et al.*, 1982). All of the above evidence indicates a similarity in the proteins, their synthesis and regulation. Moreover, the

YPs in *D. melanogaster* have a significant similarity to vertebrate lipases and have been shown to bind lipid conjugates (Bownes *et al.*, 1988), thus suggesting that the proteins have another function beside the nutrition of the developing embryo, and therefore providing further constraints on the YPs evolution. Also of great importance is the fact that YP mutants of *Drosophila* are not female sterile indicating that the YPs can substitute for each other (Bownes and Hodson, 1980; Williams *et al.*, 1987; Liddell and Bownes, 1991), albeit the production of eggs in these flies is not as efficient as in flies with the normal *yp* gene complement (Bownes *et al.*, 1991). Finally, the vitellogenins are known to enter the oocyte by receptor mediated endocytosis (Roth and Porter, 1964; Mahowald, 1972; Raikhel and Lea; 1986; Giorgi *et al.*, 1978; Butterworth *et al.*, 1991; Tsuhura *et al*, 1990) suggesting that the species which have conserved YPs may have a similar set of receptors to allow specific uptake and accumulation of YPs in the oocyte. Moreover, it may be assumed that the delivery of YPs into the oocyte is of great importance for embryo development and hence provides a great constraint in YP and receptor evolution.

It is interesting that a variety of proteins such as Adh (alcohol dehydrogenase) (Sullivan *et al.*, 1990), LSPs (Larval Serum Proteins) (Beverley and Wilson, 1982; 1984), PCP (Pupal Cuticle Protein) (Moriyama and Gojobori. 1989) and the chorion proteins (Orr-Weaver *et al.*, 1990) are

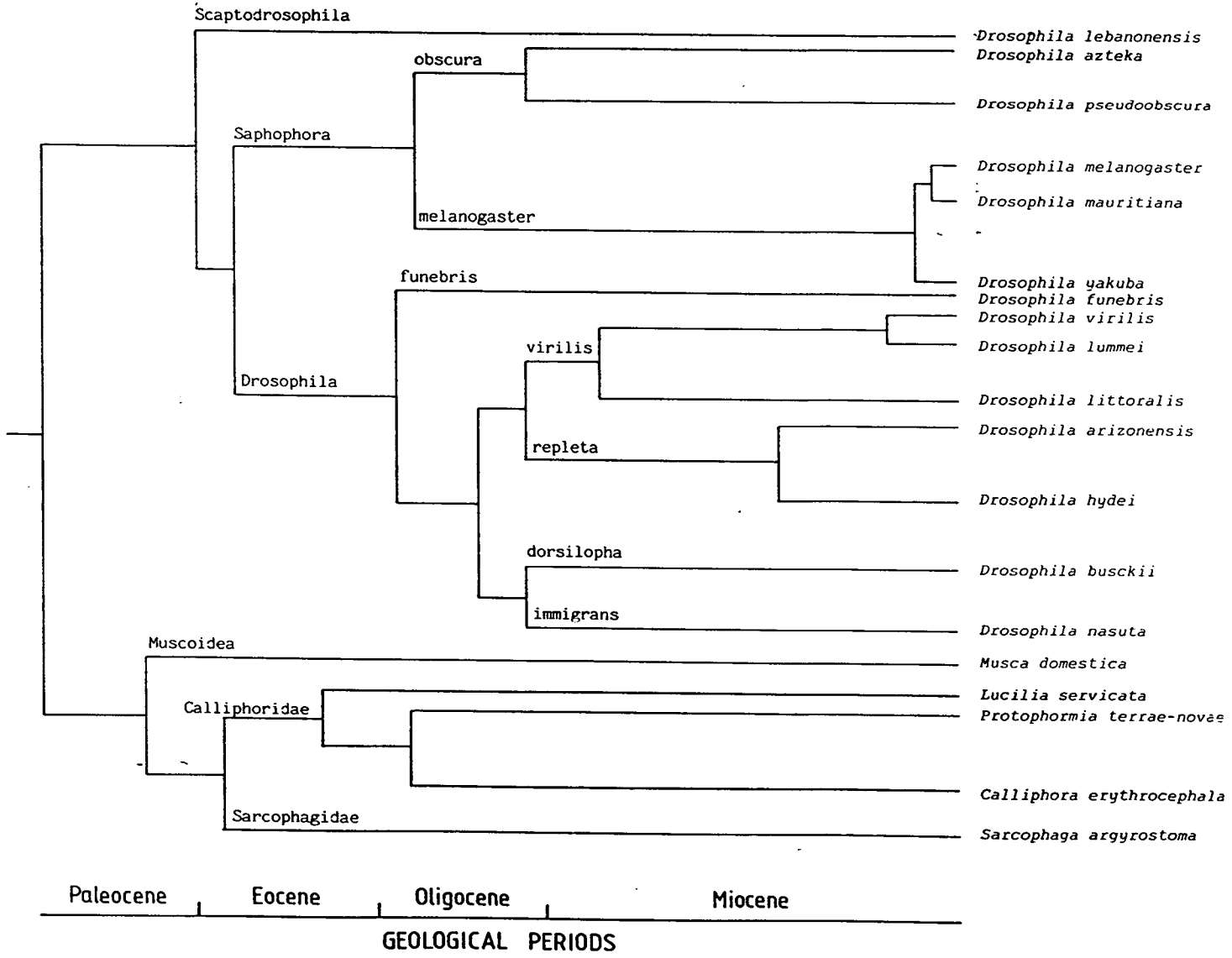
well conserved within *Drosophila* and therefore support the observed conservation of YPs with respect to molecular weight, recognition by anti-YP antibodies and uptake. In terms of gene structure more variability has been observed such that the *Adh* gene has duplicated through Drosophila evolution (Sullivan *et al.*, 1990) and the *LSP* equivalent genes in *Calliphora* (Schenkel *et al.*, 1985), *Lucilia* (Thomson *et al.*, 1976) and *Sarcophaga* (Tahara *et al.*, 1984) have very different genomic organisation to that found for the *LSP* genes of *D. melanogaster*. This diversity may be mirrored by the *yp* genes of *Drosophila* since in the species studied in this chapter there is a variability in the number of YPs (1 to 3). The *yp* genes within *Drosophila* probably do not have conserved gene organisation since in the species studied in this chapter there is species specific variability in the number of YPs (1 to 3). However, the fact that *D. grimshawi* has such a conserved gene structure and that the sibling species to *D. melanogaster* have a comparable chromosomal position for each *yp* gene, indicates that at least in some species there has been little divergence in *yp* gene structure, and poses the question of whether or not all species with three YPs (in *Drosophila* and non-*Drosophila* species) have the same gene organisation to *D. melanogaster* and *D. grimshawi*. Such conservation of gene structure has also been observed for the *PCP* genes between *D. melanogaster* and *D. virilis* (Moriyama and Goribori, 1989).

According to Beverly and Wilson (1982; 1984) and Collier *et al.* (1978) Calliphoridae and Sarcophigidae had a common ancestor with *D. melanogaster* 100MYR and t *D. grimshawi* separated from *D. melanogaster* 70MYR (Berverly and Wilson, 1985). Fossil evidence suggests that the Calyptratae (Calliphoridae, Muscidae) and Acalyptratae (Drosophilidae) diverged at least 70MYR (McAlpine, 1970) and places the Muscidae, Calliphoridae, Sarcophigidae and Drosophilidae in the cretaceous (68 to 130 MYR) (Beverly and Wilson, 1984); all of this is supported by the molecular evidence assembled so far. Therefore the data presented in this chapter suggests that the function of the YPs tested here has been conserved for up to 100MY of Dipteran evolution. Figure 3.11 shows a phylogenetic relationship concerned with the species used in the experiments described in this chapter, constructed from data obtained by McAlpine (1970), Thockmorton (1975), Collier *et al.* (1977), Beverley and Wilson (1982; 1984; 1985), Hackman and Vaisanen (1982), Ashburner (1981; 1989) and the advice of Dr. A. Saura at the University of Umea.

The conservation of genetic structure, protein function and protein characteristics in *Drosophila* appears to suggest that the conservation of YP uptake through Dipteran evolution is likely. Moreover, the YPs may have more functions than ensuring the supply of the correct dietary complement for embryo development, and therefore that the YPs have more evolutionary constraints which

Figure 3.11

Phylogenetic relationship between the species used in the experiments described in this chapter. (Based on McAlpine, 1970; Thockmorton, 1975; Collier *et al.*, 1977; Beverley and Wilson, 1982; 1984, Hackman and Vaisenen, 1982; Ashburner, 1980; 1990) and the advice of Dr. A. Saura at the University of Umea.


Scaptodrosophila
obscura
Saphophora
melanogaster
funebris
Drosophila
virilis
repleta
dorsilopha
immigrans
Muscoidea
Calliphoridae
Sarcophagidae
Drosophila lebanonensis
Drosophila azteka
Drosophila pseudoobscura
Drosophila melanogaster
Drosophila mauritiana
Drosophila yakuba
Drosophila funebris
Drosophila virilis
Drosophila lummei
Drosophila littoralis
Drosophila arizonensis
Drosophila hydei
Drosophila busckii
Drosophila nasuta
Musca domestica
Lucilia servicata
Protophormia terrae-novae
Calliphora erythrocephala
Sarcophaga argyrostoma
Paleocene
Eocene
Oligocene
Miocene
GEOLOGICAL PERIODS

limit the YP divergence in Diptera. The molecular cloning and DNA sequence determination of *yp* genes from other species will help to establish whether or not strongly conserved domains exist within the YPs. The results in this section indicate that candidates for cloning and sequencing of *yp* genes can be as distant as *Sarcophaga*, *Calliphora* and *Musca*. The isolation and characterisation of *yp* genes was targeted to distantly related *Drosophila* species to *D. melanogaster*, such as *D. funebris*, and if possible to *Calliphora erythrocephala* or *Sarcophaga argyrostoma* in which the DNA sequence may have diverged more with respect to the *yp* genes.

CHAPTER 4

RESULTS

Molecular Cloning of a *D. funebris yp* gene

Introduction

The previous chapter described experiments in which species belonging to the Diptera were screened for the ability of their yolk proteins to be internalised into the ovaries of *D. melanogaster* and *D. funebris*. The experiments were aimed at trying to find a species distantly related to *D. melanogaster* whose YPs were taken up by *D.melanogaster* ovaries, and a closely related species whose yolk proteins were not taken up. The results showed that the YPs are more similar to each other amongst the species tested than previously thought. The isolation and characterisation of a *yp* gene of *D. funebris* was chosen for a number of reasons. In the first instance the *D. funebris* YPs were taken up by *D. melanogaster* oocytes. Furthermore the reciprocal transplant (*D. melanogaster* to *D. funebris*, Figure 3.6) resulted in the internalisation of the *D. melanogaster* YPs by the *D. funebris* ovaries thus indicating that not only are the YPs similar but suggesting that the receptor mechanism used in YP uptake is also conserved. *D. funebris* is distantly related to *D. melanogaster* and hence divergence in the YP polypeptide sequences would allow the identification of important domains for YP uptake. Moreover, only one yolk protein from *D. funebris* is observed on protein gels indicating that *D. funebris* may only have one *yp* gene.

Finally, and of great practical importance, in genomic southern blots all three *yp* gene from *D. melanogaster* hybridise with discreet fragments of *D. funebris* genomic DNA (Kozma, Ph.D thesis; Kozma and Bownes, 1986b) but not with DNA from the non-*Drosophila* species. This showed that the *D. melanogaster yp* genes could be used as heterologous probes to obtain the *yp* genes from a *D. funebris* genomic DNA library.

Construction of a *D. funebris* genomic DNA library

The construction of a genomic DNA library (or any other kind of library) requires a careful choice of vector. In the case of the construction of the *D. funebris* genomic DNA library, the choice of cloning vector was determined by the length of the genomic DNA fragment known to encode the *D. funebris yp* gene. Another parameter to consider is the availability of a suitable restriction enzyme site in the vector enabling the cloning of the appropriate DNA insert fragment. In order to determine the size of the DNA fragment to be isolated, a DNA southern blot containing HindIII digested genomic DNA from *D. melanogaster* and *D. funebris* was hybridised with a *D. melanogaster* yp_3 DNA probe (as in methods and materials 2.4.5.1). The results of this experiment are shown in Figure 4.1. In the *D. melanogaster* track (track A) two HindIII genomic DNA fragments corresponding to *D. melanogaster* yp_3 (5.0 kb) and yp_2 (2.0 kb) sequences are

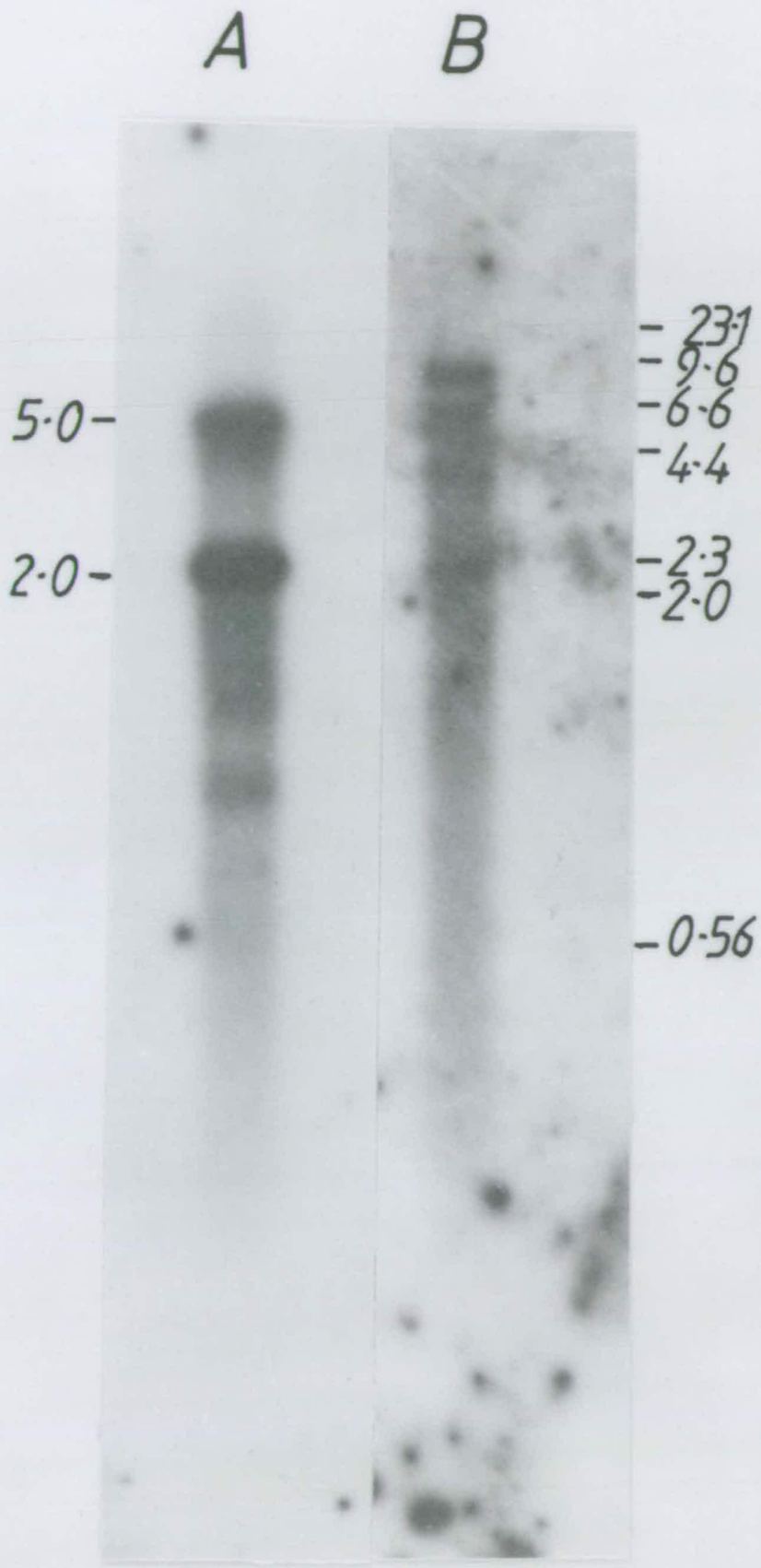
Figure 4.1

Autoradiograph of southern blot containing HindIII digested genomic DNA from *D. melanogaster* and *D. funebris* and hybridised with the radiolabelled yp_3 gene DNA.

Track A, *D. melanogaster* DNA.

Track B, *D. funebris* DNA.

The number on the left hand side of the panel indicate the size of the hybridised DNA fragments in the *D. melanogaster* genomic DNA (track A). The right hand side indicate the size of the lambda DNA HindIII digested markers.


A
B
23·1
9·6
6·6
4·4
2·3
2·0
0·56
5·0
2·0

detected. In the track containing *D. funebris* genomic DNA (track B) four HindIII DNA fragments are identified. These fragments are estimated to be 6.8 kb, 5.0 kb, 3.3 kb and 2.0 kb in length.

On the basis of the results from the genomic southern the phage cloning vector, λNM1149, was chosen which has a unique HindIII cloning site and is able to support up to 8 kb insert in length. λNM1149 is an immunity phage vector, The cloning site is found within the *cI* gene which allows the selection of recombinant (cI^-) over non-recombinant (*cI*) phages (Murray, 1983). The library was constructed as described in methods and materials (2.4.7). The λNM1149 vector DNA was obtained from Prof. N.E. Murray (ICMB, University of Edinburgh). The phage stock was stored at 4°C and later titred (2.4.4.4). The λNM1149 *D. funebris* library was titred using C600 cells which allow the growth of recombinant and non-recombinant phages, thus giving the total number of phage in the library. The phage were also titred in an *hfl* strain (NM514) which suppresses the growth of the non-recombinant phage. This way the recombinant phage titre is obtained (Table 4.1).

Screening the genomic *D. funebris* library

In the primary screen of the genomic DNA *D. funebris* library, an estimated 1 x 10^5 pfus were screened with the yp_3 DNA sequence of *D. melanogaster*. The yp_3 DNA insert

Table 4.1 Titre of the control and D. funebris genomic DNA libraries

Strain	Control	D. funebris
C600	-	$7X10^7$ pfus/ml
NM514	-	$1X10^7$ pfus/ml
ED5386	$10X10^{10}$ pfus/ml	-

Table 4.2 The size (kb) of the HindIII DNA insert in the recombinant phages obtained from the D.funebris library

Phage Name	Insert(s) size (kb)
1	3.4
2	2.0
3	4.3,2.6
4	6.8,6.1,4.5,3.9
5	2.6,2.7
6	3.9
7	5.0

fragment was purified from the plasmid by cutting with EcoRI and SalI, separated by gel electrophoresis and the agarose removed from the *yp*$_3$ DNA using Geneclean (2.4.1.6). After the secondary screen seven positive clones were identified and named λdf1 to λdf7. Table 4.2 shows the size of the *D. funebris* DNA inserts found in each of these clones. On the basis of the fragment size several clones were chosen for further analysis. λdf7 contains a 5.0 kb long HindIII DNA fragment which may correspond to the 5,0 kb fragment identified on the southern blot (Figure 4.1). Further characterisation of the clone was performed in two ways. The λdf7 DNA fragment was used as a probe in southern and northern blot analysis of samples obtained from *D. melanogaster* and *D. funebris* to show that it is in fact a yp gene fragment.

Southern blot analysis of the λdf7 phage.

A southern blot containing HindIII digested genomic DNA from *D. melanogaster* and *D. funebris* (Figure 4.2A), was probed with a 4.3 kb EcoRI/SalI fragment from the *D. melanogaster yp*$_3$ gene, which contains all of the *yp*$_3$ coding sequences. The results of the southern (Figure 4.2B) show that the *yp*$_3$ probe hybridises with the λdf7 insert (track 1) and a similar sized HindII fragment in the genomic DNA of *D. funebris* (Figure 4.2, track 2). In the *D. melanogaster* genomic DNA (track 1) three hybridisation signals are observed. The two larger DNA

Figure 4.2

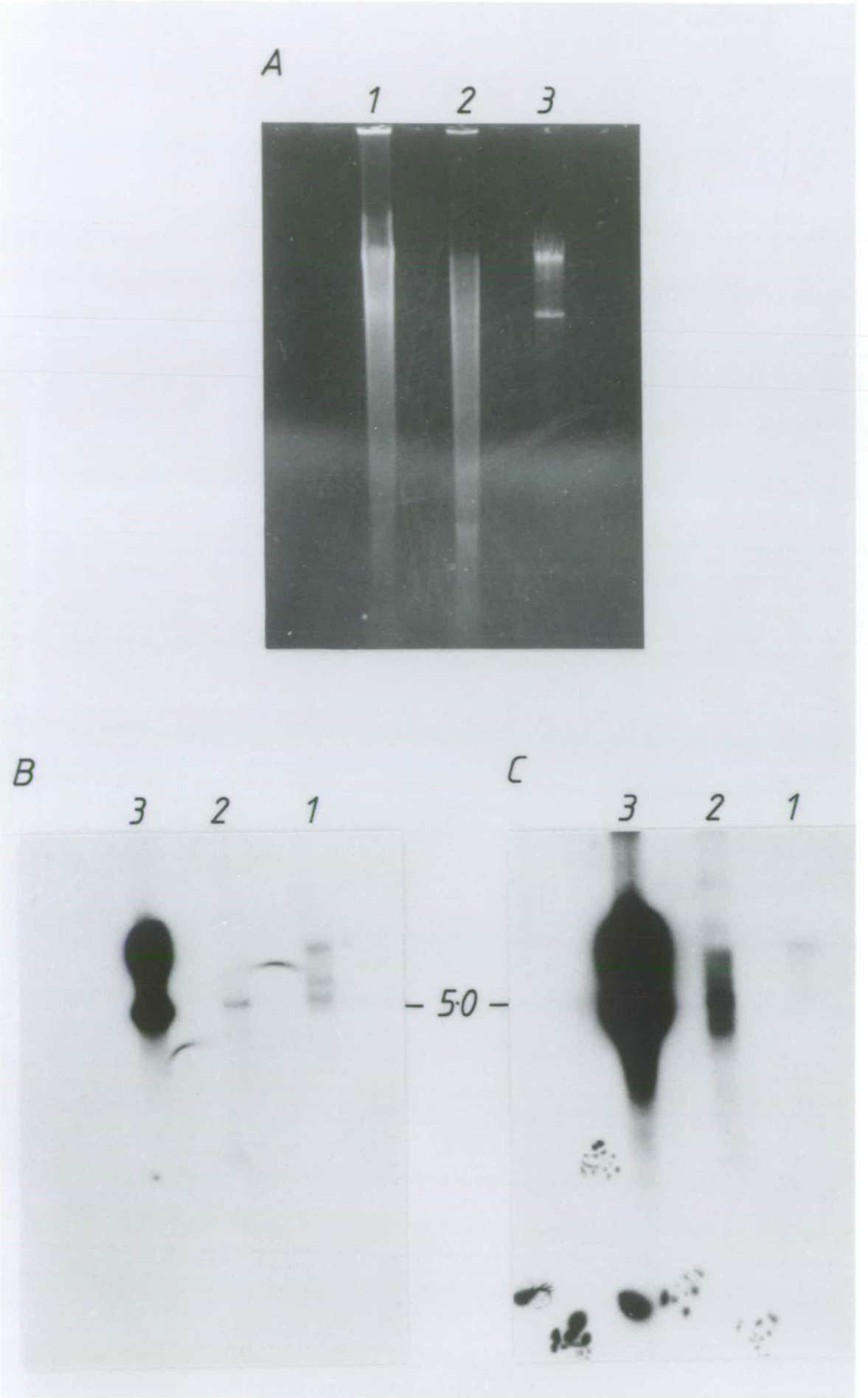
Southern blot analysis of the λdf7 phage insert.

A. Ethidium bromide stained agarose (1% w/v) gel containing genomic DNA digested with HindIII from *D. melanogaster* (track 1), *D. funebris* (track 2) and λdf7 (track 3).

B. Autoradiograph of the southern blot of gel in A hybridised with the *yp*$_3$ gene of *D. melanogaster*.

C. Autoradiograph of stripped southern blot in B and hybridised with the λdf7 insert.

Track 1, *D. melanogaster* HindIII digested genomic DNA.
Track 2, *D. funebris* HindIII digested genomic DNA.
Track 3, λdf7 HindIII digested genomic DNA.


A
1
2
3
B
3
2
1
– 50 –
C
3
2
1

fragments may have been detected as a result of incomplete digestion of the genomic DNA (see Figure 4.2A, track 1). More importantly however, a 5.0 kb HindIII genomic DNA fragment in the *D. melanogaster* DNA was detected by the *yp*$_3$ gene probe. When the same southern blot was stripped (2.4.5.1) and reprobed with the λdf7 insert, it also detected the 5.0 kb HindIII fragments in *D. funebris* and *D. melanogaster* genomic DNAs (Figure 4.2C, tracks 2 and 1), in addition to hybridising to itself (track 3). The results of this experiment indicates the likelihood that the λdf7 contains a HindIII DNA fragment which encodes for a *yp* gene of *D.funebris*.

Northern analysis of the λdf7 phage.

The λdf7 HindIII DNA insert fragment was used to probe a northern blot containing *D. funebris* male and female total RNA. The probe was found to cross hybridise with ribosomal RNA in both males and females. However, in the female RNA a further transcript similar in size to the *D. melanogaster* yp transcripts was observed (Figure 4.3A). The same result was obtained when a duplicate blot was probed with the 4.3 kb EcoRI/SalI *yp*$_3$ DNA probe from *D. melanogaster* (Figure 4.3A). Restriction analysis of the λDf7 phage indicates that the HindIII DNA insert fragment contains five EcoRI fragments. These have been named according to size, namely E$_1$ (1.9 kb), E$_2$ (1.5 kb), E$_3$ (0.85 kb), E$_4$ (0.79 kb) and E$_5$ (0.5 kb). E$_5$ contains

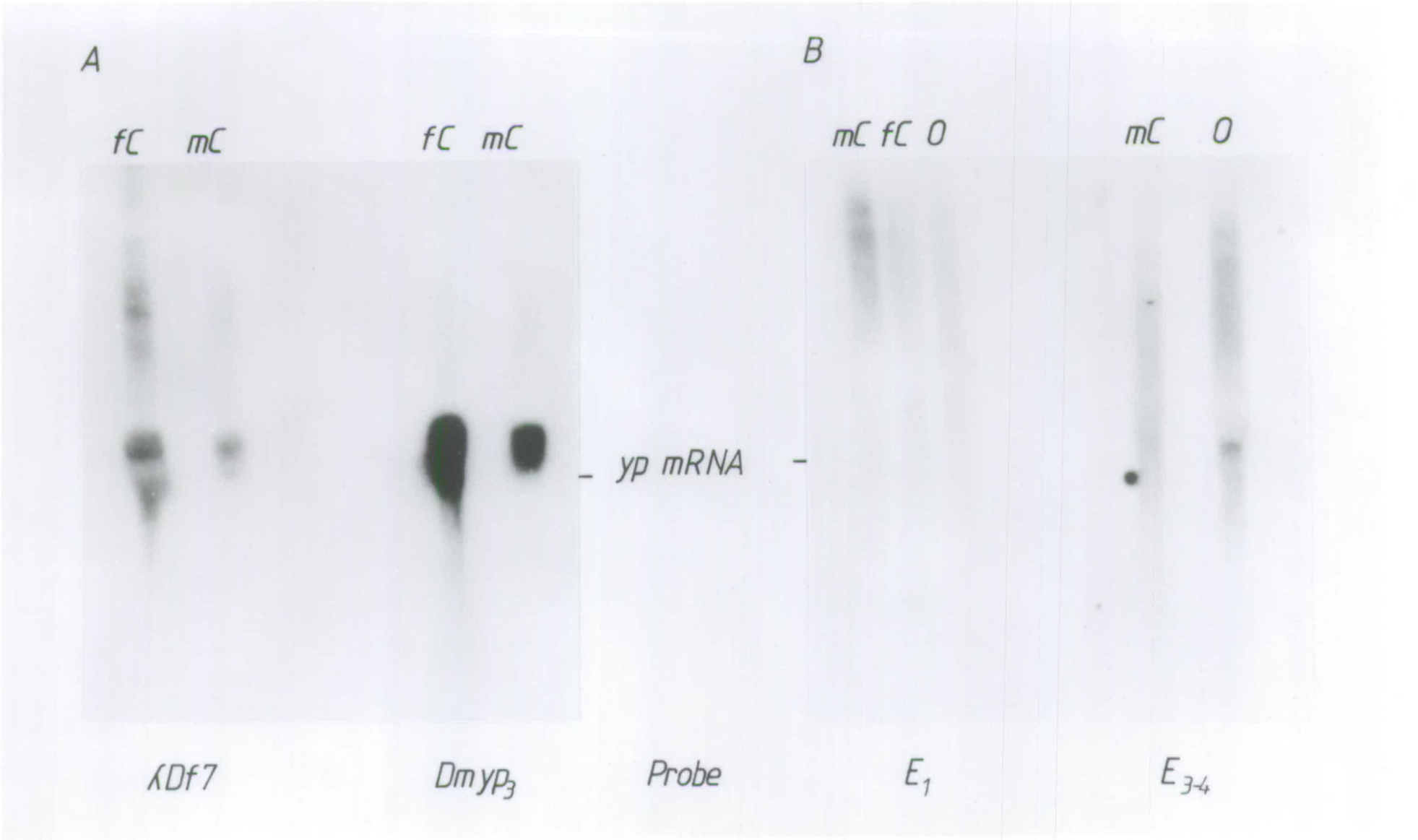
Figure 4.3

Northern blot analysis of the λdf7 phage insert.

A. Northern blot made with male (mC) and female (fC) total RNA from *D. funebris*. The northern blots were hybridised with the λdf7 insert and the *D. melanogaster* yp_3 gene.

B. Northern blots made with ovary (O), male (mC) and female (fC) carcass total RNA from *D. funebris*. The northern blot was hybridised with the E_1 and the E_3+E_4 (together) sub-fragments of λdf7.

ypmRNA, indicates the yp transcripts position in the northern blots.


A
fC
mC
fC
mC
yp mRNA
XDf7
Dmyp3
Probe
B
mC
fC
0
mC
0
E1
E3-4

lambda sequences belonging to the *cI* gene (Figure 4.4). DNA Fragments E_1 and E_3+E_4 (initially both were used together because of the difficulty of separating them) were used to probe total RNA prepared from *D. funebris* flies. The E_3+E_4 DNA fragments detected a female specific transcript estimated to be 1.6 kb in length (Figure 4.3B, track O) whereas no hybridisation signal was obtained with the E_1 DNA fragment (Figure 4.3B). Therefore, the E_3+E_4 DNA fragments contain sequences encoding a putative *yp* gene of *D. funebris*.

Discussion

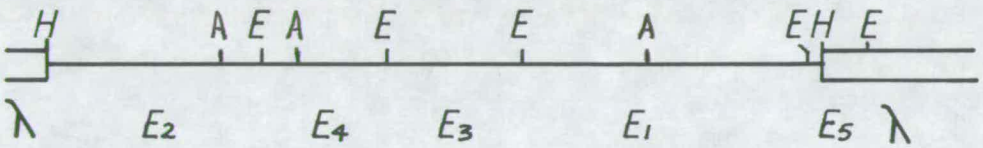
The aim of the experiments described in this chapter was to clone a *yp* gene from *D. funebris*. As stated before *D. funebris* has one detectable yolk protein band when the haemolymph or ovaries are fractionated by linear or gradient SDS-PAGE (Kozma and Bownes, 1986; this thesis Figure 3.1A and C). Kozma and Bownes (1986), probed southern blots of *D. funebris* genomic DNA with the three *yp* genes of *D. melanogaster* and concluded that *D. funebris* may have more than one *yp* gene in its genome. It is possible that more than one YP is synthesised by the *D. funebris* fat body and that the protein products have similar molecular weights which comigrate on linear and gradient polyacrylamide gels. Alternatively, the *D. funebris* genome may contain fuctionally inactive *yp* genes or pseudogenes, or the yp_3 DNA probe used may cross

Figure 4.4

Restriction map of the *D. funebris* insert DNA contained in the λdf7 phage.

E_1, E_2, E_3, E_4 and E_5 are the EcoRI sub-fragments of λdf7 insert DNA.

D. funebris



H - Hind III

A - AccI

E - EcoRI

λ - Lambda vector λNM1149

___ 500 bp

hybridise with non *yp* gene sequences.

The approach used to isolate the *D. funebris yp* gene (or genes) was to screen the *D. funebris* genomic library with the coding sequences of the *D. melanogaster* yp_3 gene. The yp_3 gene was chosen as a hybridisation probe since the YP of *D. funebris* comigrates with YP_3 of *D. melanogaster* and it is possible that the *D. funebris* YP is more similar to YP_3 than to YP_1 or YP_2 of *D. melanogaster*. A southern blot with HindIII digested genomic DNA from both species (*D. funebris* and *D. melanogaster*) was probed with the yp_3 gene and detected four HindIII fragments in *D. funebris* DNA. One of these fragments comigrated with the 5.0 kb HindIII fragment of *D. melanogaster* (Figure 4.1, track A). Kozma and Bownes (1986), probed *D. funebris* HindIII digested genomic DNA with the yp_3 gene of *D. melanogaster* and obtained hybridisation to one HindIII fragment 6.3 kb in length and which correlates with the 6.8 kb fragment observed in Figure 4.1. The other fragments observed in in the genomic *D. funebris* DNA (in Figure 4.1, track B) may be a result of cross-hybridisation to other putative *yp* genes or lipase genes. This is possible since the *D. melanogaster* yp_2 gene is reported to crosshybridise to *D. funebris* genomic HindIII fragments of 1.8, 5.5 and 6.9 kb (Kozma and Bownes, 1986) where 6.9 kb and 1.8 kb HindIII genomic fragments may compare to 6.8 kb and 2.0 kb observed in Figure 4.1.

When the *D. funebris* genomic DNA library was screened with the *D. melanogaster* yp_3 DNA probe a recombinant phage containing a 5.0 kb *D. funebris* DNA fragment was isolated (λdf7) (Table 4.2). It was shown by southern blot analysis that the λdf7 insert hybridised with the 5.0 kb HindIII fragment in *D. funebris* and *D. melanogaster* genomic DNA (Figure 4.2C). More importantly the same fragments were detected when the same southern blot was probed with the *D. melanogaster* yp_3 gene (Figure 4.2B), where the 5.0 kb *D. melanogaster* HindIII fragment corresponds to the yp_3 gene. The southern blot data is supported by the northern blot results (Figure 4.3). The λdf7 insert hybridises with a female specific transcript (1.6 Kb in length) in the size range expected for a *yp* gene message. Finally, only the two EcoRI sub-fragments, E_3 and E_4, from the λdf7 insert fragment detected the same 1.6 kb message as the whole fragment indicating that these fragments fully or partly encode a *D. funebris* *yp* gene sequences.

Further analysis of the λDf7 clone was problematic because the 5.0 kb HindIII fragment was difficult to subclone into the desired plasmid vector (pGEM-1). Such cloning difficulties have been observed with truncated *yp* gene sequences from *D. melanogaster* (P. Wensink, S. Liddell personal communications) and *Calliphora* (this thesis Chapter 5). Similarly, the cloning problems were encountered when attempts were made to sub-clone the λdf7

insert and its E_1, E_2, E_3, E_4 and E_5 sub-fragments into the M13mp18/19 cloning vectors. DNA sequence analysis of the recombinant M13mp18 and mp19 phages revealed that the inserts had undergone rearrangements and recombined readily. The lack of pGem or M13mp18 and mp19 recombinants also impaired the determination of the direction of transcription of the putative *yp* gene since strand specific probes could not be made to test on northern blots. Furthermore, the sub-cloning of the E_1, E_2, E_3, E_4 and E_5 sub-fragments was hindered by the low yields of insert fragments obtained when the λdf7 phage DNA was digested with EcoRi, since the insert DNA length only constitutes apprx. 10% of the length of the recombinant phage. The λdf7 fragment was subcloned eventually into pUC19 to make the plasmid pdf7. Attempts were made to subclone the E_3 and E_4 sub-fragments and to determine the DNA sequence of the pdf7 insert. The sequence analysis of the pdf7 clone was discontinued when sequence derived from *Calliphora erythrocephala yp* genes showed strong similarities to the *yp* genes of *D. melanogaster* (Chapter 5, this thesis). Moreover, the *C. erythrocephala* clones offered the opportunity to investigate the *yp* gene sequences of a Dipteran species more distantly related to *D. melanogaster* that any *Drosophila* species.

To further confirm that λdf7 contains *yp* sequences from *D. funebris* the distribution of the yp message in

ovaries and fat body may be analysed by *in situ* hybridisation to whole mounts of ovaries and carcasses. Finally, sequencing of the E_3 and E_4 fragments can be performed by subcloning these fragments into plasmid vectors rather than single stranded sequencing vectors, which should result in stabilisation of the putative *yp* sequences. Once the identity of the clone is confirmed the number of *yp* genes present in the genome of *D. funebris* may be assesed via genomic southern blots. The southern blot in Figure 4.2 shows the presence of other HindIII genomic fragments when the *D. funebris* putative *yp* gene was hybridised with *D. funebris* genomic DNA. However, it should not be assumed from this data that there is more than one *D. funebris yp* gene since this southern blot was washed at low stringency to ensure cross-hybridisation to the *D. melanogaster* genomic DNA. To determine the number of *yp* genes it will be necessary to probe genomic DNA digested with different restriction enzymes and to use more stringent post-hybridisation wash conditions than those used for inter species genomic DNA southerns.

The approach taken here to clone a *yp* gene from *D. funebris* using a heterologous probe proved to be difficult. In retrospect it may have been more advantageous to employ a differential screening procedures using cDNA derived from male and female polyA$^+$ RNA. As stated in the introduction the *yp* genes are expressed in a

tissue, temporal and sex specific fashion. The yp transcripts are stable and abundant in female carcasses and ovaries. The differential screen would involve using ^{32}P-dCTP labelled cDNA made from polyA$^+$ mRNA from female and male carcasses to screen a *D. funebris* genomic DNA library. Note that ovary tissues would be dissected out from female flies to prevent undesired isolation of the chorion genes or other ovary specific genes of *D. funebris*. This would enable a very stringent screening procedure to be performed whereby the likelihood of obtaining clones with *yp* sequences would be greatly enhanced and such techniques have been succesfully applied to the *yp* genes of *D. melanogaster* (Barnett *et al.*, 1980). Alternatively, there are two other techniques which may be used. One approach would involve making degenerate oligonucleotide primers from highly conserved sections of the YPs and using these to screen a cDNA or genomic library. The other approach is an extension of this whereby advantage is taken of the advances in the polymerase chain reaction (PCR) technology. In this case oligonucleotide primers could be synthesised from two different sites within the *yp* genes of *D. melanogaster* and used to make a *D. funebris* specific *yp* gene probe with which a cDNA or genomic library of *D. funebris* could be screened. This technique would also enable the direct sequencing of the probe to ensure that the PCR product is in fact what is expected. It is important to take into account the codon usage and genome nucleotide composition,

if possible, of the target species when designing the oligonucletides. This minimises the chances of generating contaminating PCR products.

Further characterisation of *yp* gene expression in *D. funebris*, together with the determination of *yp* gene number, gene organisation and sequence information should result in a better understanding of the evolution and divergence of the *yp* genes within *Drosophila*. Moreover the sequence data obtained may be used to search for regulatory DNA sequences common to *D. melanogaster* and *D. funebris yp* genes. The derived protein sequence may be used to attempt to define protein domains of importance in YP function.

The aim of this project was to isolate, clone and determine the DNA sequence of a *yp* gene which was encoded by a distantly related species to *D. melanogaster*. The *yp* genes of *C. erythrocephala* were isolated by Rubacha *et al.* (1989) and kindly gifted to this laboratory. Since *C. erythrocephala* is more distantly related to *D. melanogaster* than *D. funebris* further characterisation of the *Calliphora* putative *yp* gene sequences in preference to the *D. funebris yp* gene (Chapter 5) was carried out, thus ensuring a greater evolutionary distance and maximising for sequence divergence. However, the putative *yp* gene of *D. funebris* has been cloned and is available for future characterisation.

CHAPTER 5

RESULTS

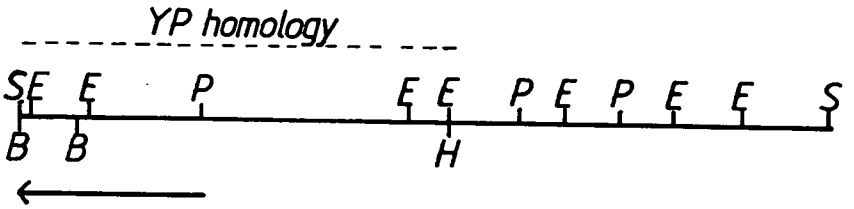
Characterisation and Sequencing of two *yp* genes of *Calliphora erythrocephala*

The availability of putative *yp* gene fragments from *C. erythrocephala* provided an excellent opportunity to sequence a *yp* gene out-with the *Drosophila* species. The sequence divergence between the *C. erythrocephala* and *D. melanogaster yp* genes is expected to be larger than within *Drosophila* species since there is more evolutionary distance between *Drosophila* and *Calliphora*. The YP amino acid sequences derived from *Calliphora yp* genes would be better suited to define protein sequences of importance for their specific uptake into oocytes than YP amino acid sequences derived from other *Drosophila* species. The *Calliphora yp* gene clones were isolated in Dr. K. Beckingham's laboratory (Rice University), from a *Calliphora* genomic library made in λEM6L4. The library was differentially screened using cDNA made from polyA$^+$ RNA extracted from previtellogenic ovaries and polyA$^+$ RNA obtained from vitellogenic ovaries. The initial goal was to find highly repetitive ribosomal DNA sequences, but the design of the screen also allowed the isolation of clones containing putative *yp* gene sequences (Rubacha *et al.*, 1988). The gift of the two recombinant clones, named A10B and B8I, containing *yp* gene sequences allowed further characterisation of these sequences. The restriction map organisation of A10B and B8I is shown in Figure 5.1. Rubacha *et al.* (1988) determined that these clones

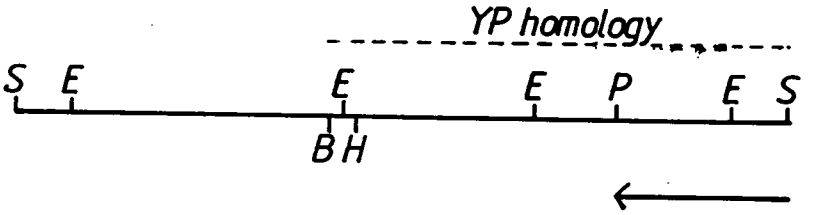
Figure 5.1

Restriction map and putative *yp* gene localisation within the *C. erythrocephala* DNA insert in phages A10B and B8I. The solid arrows indicate the direction of transcription of the mapped putative *yp* genes. The broken line (___) indicated the area which hybridised to the yp_1 gene pf *Drosophila melanogaster* (Based on Rubacha *et al.*, 1988).

Clone A10B



Clone B8I



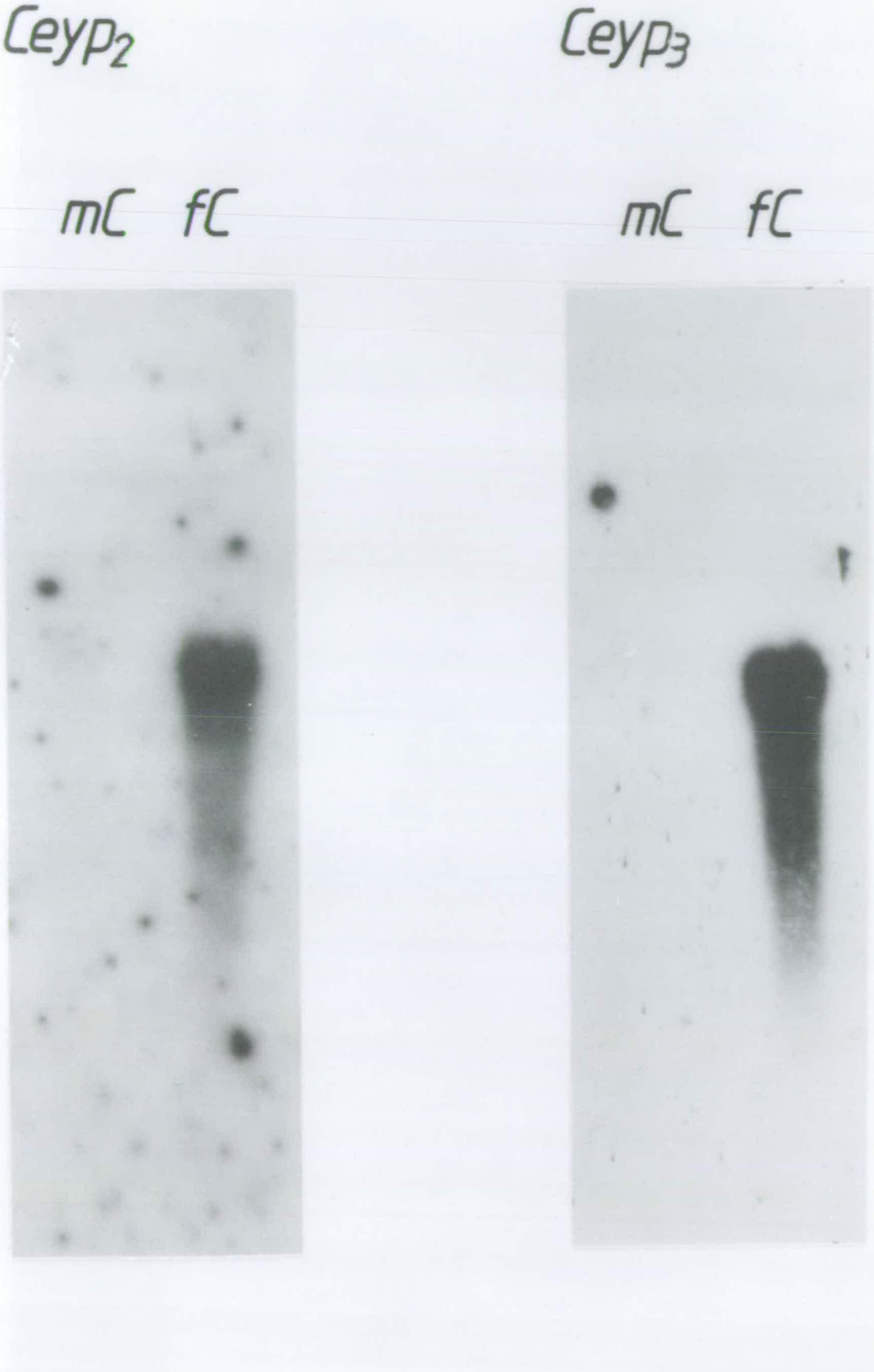
1 kb

contained putative *yp* gene sequences by showing that the insert fragments cross hybridised to the yp_1 gene of *D. melanogaster*. In addition to this, single stranded probes (Figure 5.1) were used to probe northern blots of total vitellogenic and non-vitellogenic ovary RNA showing that specific vitellogenic transcripts were obtained whose estimated molecular size correlated to the predicted yp transcript size. *In situs* to tissue sections of ovaries showed follicle cell specific expression (Rubacha *et al.*, 1988), which is a normal site of ovarian *yp* expression in *Drosophila* (Barnett *et al.*, 1980; Isaac and Bownes, 1982) and *Sarcophaga* (Geysen *et al.*, 1986).

In order to ensure that the clones obtained by Rubacha *et al.* (1988) contained putative *yp* gene sequences, RNA from males and vitellogenic female carcasses was extracted (2.4.2.2.2) and used in two northern blots (Figure 5.2). The northern filters were each probed with the B8I 1.6 kb PstI/EcoRI fragment and the A10B 0.6 kb BamHI fragment, resulting in hybridisation to an identically sized transcript present in female vitellogenic carcasses. This data comfirmed the identity of the clones since expression in the female carcass is an essential character of *yp* gene expression in *Calliphora* (Huybrechts and De loof, 1982).

Figure 5.2

Autoradiograph of northern blots containing total RNA from male (mC) and vitellogenic female carcasses (fC). The northern blots were hybridised with the Ce*yp*$_2$ and Ce*yp*$_3$ gene coding sequences (see text for details).


Ceyp2
mC fC
Ceyp3
mC fC

A. Sequencing the *yp* coding sequences in clone B8I

The DNA sequencing strategy is shown in Figure 5.3A. The 0.7 kb SalI/EcoRI and 1.6 kb PstI/EcoRI fragments were cloned into M13mp18 and mp19, resulting in the resolution of the DNA sequence of both clones in both orientations. The determination of the DNA sequence was carried out as described in methods and materials (2.4.8), and the nucleotide sequence is presented in Appendix 1. The nucleotide sequence is 1282 bp in length and contains an open reading frame.

Upon translating the DNA sequence into amino acid sequence and comparing it to the *D. melanogaster* YPs (Figure 5.3B), it became apparent that the YP sequence within the B8I clone was incomplete. The amino acid sequence lacks the first exon. The nucleotide sequence includes a small portion of the end of the intron (22 bp), thus illustrating that most of the intron is also missing. The DNA sequence reveals a polyadenylation site (AATAAA) 156 bp from the stop codon. The coding sequence extends for 1094 bp or 356 amino acids. Comparison between the amino acid sequence of B8I and the *D.melanogaster* YPs using the GCG Wisconsin package (Devereux *et al.*, 1984) (Table 5.1), indicated that the B8I protein sequence was more similar to the YP_2 sequence of *D. melanogaster*. The YP encoded by the B8I will now be referred to as $CeYP_2$ (Ce= *Calliphora erythrocephala*). The direction of

Figure 5.3

A.DNA sequence determination strategy of the fragments in clone B8I known to contain putative *yp* gene sequences of *C. erythocephala*.

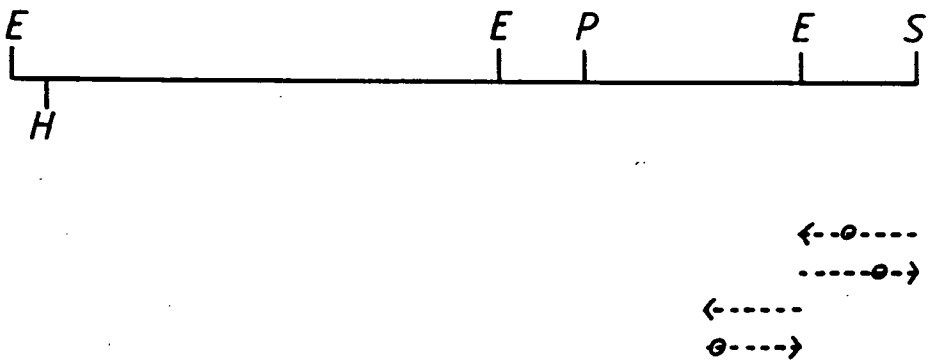
-----indicates DNA sequence determination by generation of single stranded templates (M13mp18.19)

---> indicates direction of sequence determination

O indicates utilisation of synthetic oligonucleotide primers designed from predetermined DNA sequence

B.Single letter amino acid sequence of the polypeptide derived from the B8I fragments containing *yp* gene sequences. The single letter amino acid key is in Appendix 3.

A



0·5kb

B Ce YP_2

```
  1  HLSQAGRAIE PSFVPKASEI PAFLITPDNK KVNFKLSELP KIAKEEKSFG
 51  DEEVTVYITG LPQKTETVKK ATRKLVQAYM QRYNGQAPER NSVRYEDDSS
101  EKRNPSSSEE DEWKSNSNKP SGNLVVIELG NVLTNMKNYV FLDVEKTGME
151  IGDILVQCTD KADVPQEIVH VIGSNIAAHV AGAVGRQFTR ETGHQLRRIT
201  GLDPSKIYAQ TQKSLRGLAR GDAEFVDVIH TSAYGWGTPT RCGDVDFYPN
251  GPSRGVPGAD NVVEASMRAI RYFAESVVPG NERNFPAVGA TSLKQYKEQN
301  GNGKRVYMGI NTDYDVEGDF ILQVNGKSPF GRSTPAQKQQ NYHNVHKPWK
351  MSSRDF
```

Table 5.1 B8I Protein sequence compared to Drosophila YP

Protein	Percent Similarity	Percent Identity
B8I X YP1	70.9	58.8
B8I X YP2	73.5	57.4
B8I X YP3	70.5	53.3

transcription as well as the identity of the *yp* gene equivalent in *D. melanogaster* were predicted correctly by Rubacha *et al.* (1988). Assuming that the CeYP$_2$ has a similarly sized N-terminal end to YP$_2$ in *D. melanogaster*, the predicted protein size would be 436 amino acids, which compares favourably with the 442 amino acid sequence of DmYP$_2$ (Dm = *D. melanogaster)*. In terms of structure the CeYP$_2$ and DmYP$_2$ have a conserved intron site, and a dot plot matrix comparing both sequences (Figure 5.4) shows very few gaps. Those gaps that are present are concentrated towards the N- terminal end of both proteins. Hence, as observed previously for *D. melanogaster* (Garabedian *et al.*, 1987) the yolk proteins are more similar at the C-terminal end than at the N-terminal end.

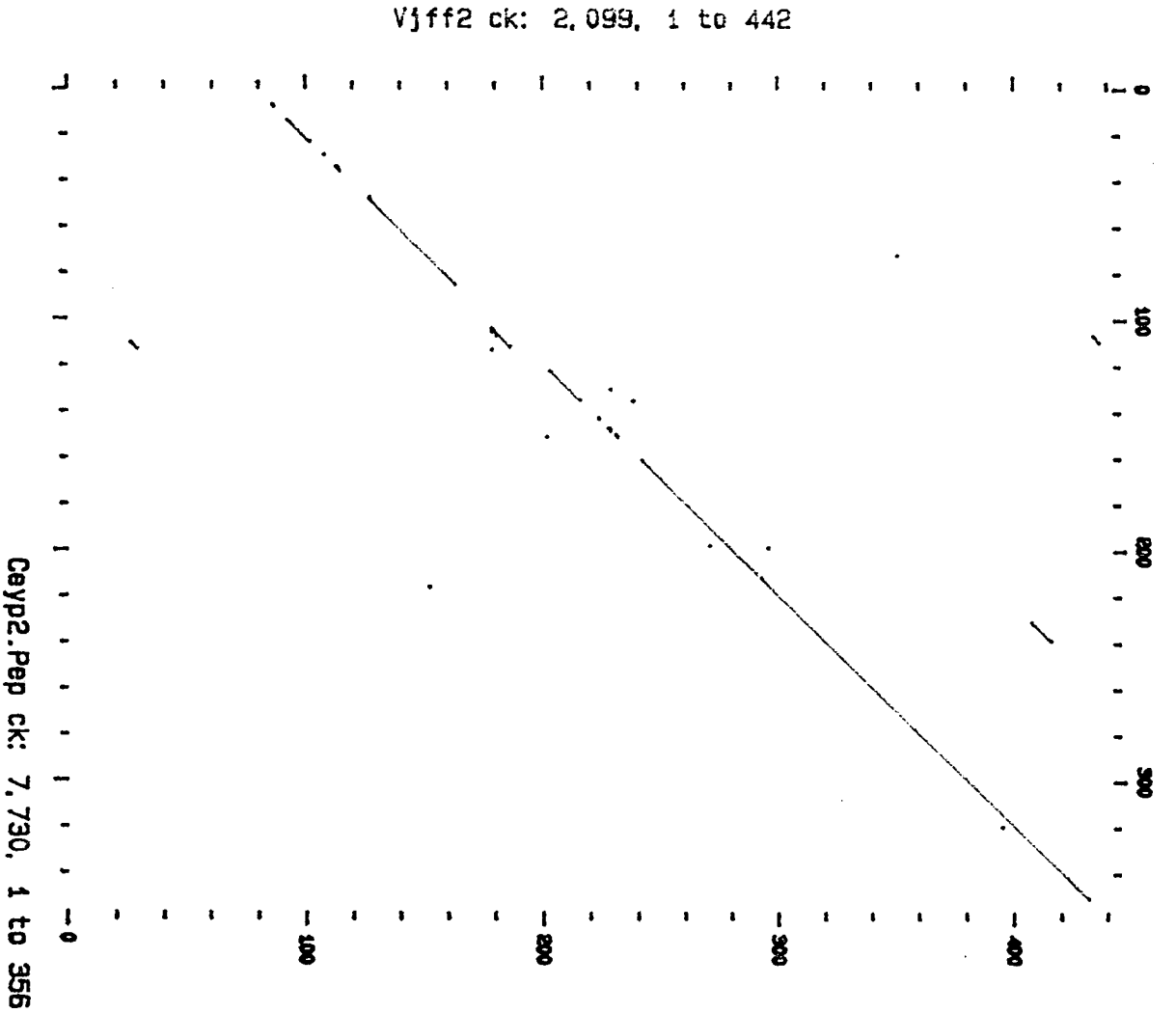
Attempts were made to clone the remaining portion of the Ce*yp*$_2$ gene. In the first instance the 0.6 kb SalI/EcoRI B8I fragment was used to probe a filter made from the plated λEMBL4 *Calliphora* genomic library (Rubacha *et al.*, 1988). This library has randomly generated inserts and is therefore expected to contain overlapping fragments. Thirteen clones were positive after the secondary screening procedure but were not further characterised due to the lack of time. It is possible that some of the clones isolated may encode other *yp* genes in addition to the remaining portion of the Ce*yp*$_2$. The second approach involved using the technique inverse polymerase chain reaction (PCR) (Ochman *et al.*, 1989)

Figure 5.4

Dot plot comparing the *D. melanogaster* and *C. erythrocephala* YP_2 amino acid sequences. The Dot plot was generated using the program "Compare" within GCG (Version 6.0) (Devereux *et al.*, 1984).

DOTPLOT of: Ceyp2.Pnt Density: 504.55 April 12, 1991 12:43

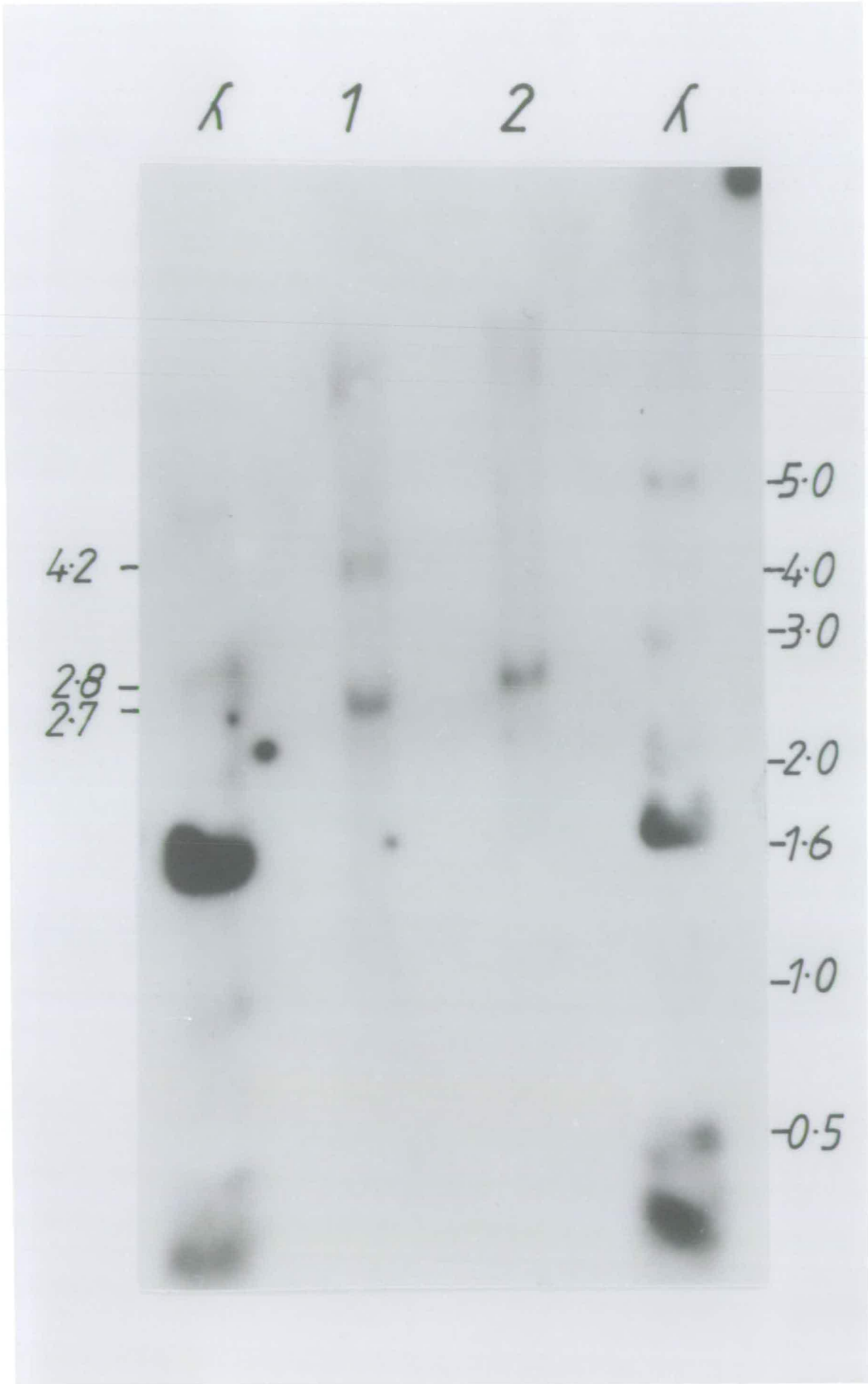
COMPARE Window: 12 Stringency: 9.0 Points: 301



whereby a genomic DNA fragment containing the 5' end sequences of the Ce*yp*$_2$ was recircularised and by using oligos to read into either side of the undetermined sequence a PCR reaction was carried out to amplify the remaining portion of the *yp* gene. This procedure is limited by the purity of the genomic DNA and the efficiency of the ligation prior to the PCR reaction. The size of the genomic fragment should not be smaller than 2.5 kb or larger than 4.0 kb in length to ensure efficient recircularisation, and to obtain an efficient result the sequence be to amplified should not exceed 1 kb in length (Ochman *et al.*, 1989). The size of the genomic DNA fragment was determined by cutting *Calliphora* genomic DNA with enzymes which would give a fragment size between 2.5 and 4.0 kb and which included at least 500 bp of unknown 5' sequence. EcoRI (Figure 5.5, track 1) and PstI (Figure 5.5, track 2) both yielded fragments of 2.7 and 2.8 kb respectively. Before performing the inverse PCR reactions, normal PCR reactions were carried out with the genomic DNA to ensure that it was of sufficiently high quality to permit amplification of specific sequences. These reactions were unsuccessful and require further testing to find the correct conditions. This technique will allow the isolation and sequencing of the remaining 5' coding sequence of the Ce*yp*$_2$ gene.

Figure 5.5

Autoradiograph of southern blot containing *C. erythrocephala* genomic DNA digested with EcoRI (track 1) and PstI (track 2). λ- signifies tracks loaded with a 1 kb DNA marker ladder (BRL). The numbers on the right hand side of the panel indicate the size of the markers. The size of the genomic DNA fragment to which the 0.6 EcoRI/SalI Ce*yp*$_2$ DNA fragment hybridises to are indicated in the left hand side of the panel, namely 2.7 kb EcoRI, 2.8 kb PstI and 4.2 kb EcoRI fragments.


λ
1
2
λ
5·0
4·2
4·0
3·0
2·8
2·7
2·0
1·6
1·0
0·5

B. Sequencing of the *yp* coding sequences in clone A10B

The DNA sequencing strategy of the *yp* gene coding sequences within A10B is depicted in Figure 5.6A. DNA sequence determination using M13mp18 and mp19 was limited to the 0.6 kb BamHI fragment. The rest of the sequence was derived using double stranded methods from plasmids containing 5.8 kb SalI/HindIII, 2.3 kb Sal/PstI and 1.6 kb PstI/EcoRI fragments (Figure 5.1) subcloned into pBluescript (Promega). The plasmid sequencing was carried out in triplicate using synthetic oligonucletides obtained from OSWEL DNA service (University of Edinburgh). The nucleotide sequence determined was 2104 bp in length and has one open reading frame and is presented in Appendix 2. The subcloning of the 1.6 kb PstI/EcoRI fragment into M13 bacteriophage proved to be very difficult, therefore single stranded sequencing was not used. The *yp* gene sequence, including introns, from the start codon to the stop codon is 1363 bp (Appendix 2). The TATA (ie. TATAAAA) box is located at -89 bp from the protein start site, and the polyadelynation site (ie. AATAAA) is at 1430 bp from the protein start site or 67 bp from the stop codon. Rubacha *et al.*, (1989) predicted the direction of transcription correctly. The deduced amino acid sequence is 422 amino acids long (Figure 5.6B). Upon comparison with the YPs of *D. melanogaster* (Table 5.2) it is difficult to decide if the equivalent protein is DmYP$_1$ (as suggested by Rubacha *et al.*, 1989) or DmYP$_3$. However

Figure 5.6

A.DNA sequence determination strategy of the fragment in clone A10B containing putative *yp* gene sequences of *C. erythrocephala*.

-----Indicates DNA sequence determination by the single stranded method.

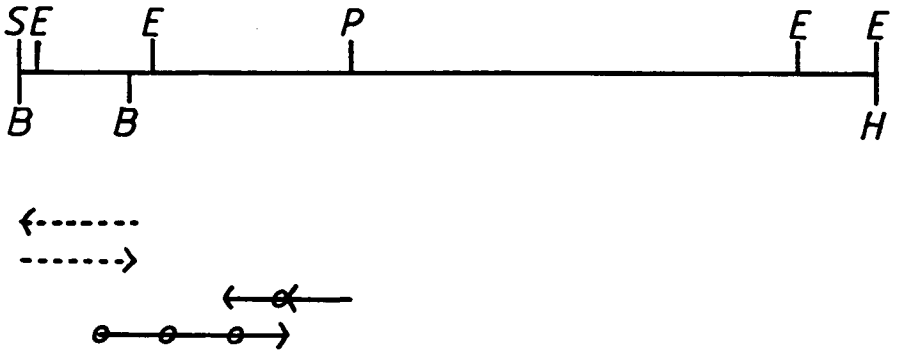
——— Indicates DNA sequence determination by the double standed methodology.

---> Indicates direction of sequence determination

O Indicates synthetic oligonucleotide primers designed from predetermined DNA sequence

B.Single letter amino acid sequence of the polypeptide derived from the DNA sequence of A10B fragments containing *yp* gene sequences. The single letter amino acid key is in Appendix 3.

A.



0·5kb

B. *Ce YP3*

```
  1  MNPLRIVCVA ALLLAAGSAN GNLSGLNKLR PSQWLSSSQL DKLPRSMEIS
 51  LQKLESMSVE KGAELMQKLY HLSQINNDLK PSFVPSSSNV PCYIVKPNGK
101  KVSTSLDKLA SACKQQPNFG EEEVTILITG LPATTETVRK ANRKLIDAYL
151  QRYSTKRQQP SKFDYSGEKM ARTSSEEDSN EWQNQQASSG NLVIIDLQNE
201  LNSFKRFSLL DVDETGAMIR SAIVEMTEKC DVADETIHVV AQGIAAHVAG
251  AAGNEFTRKT GRQLRRITAL DPSKILAKNP HSLTGLSRGD AEFVDAIHTN
301  VYGMGTIQRV GDIDFYPNGP SAAVPGAQSI VEASMRATRY FAESVRPGNE
351  RNFPAVAANS LKEYKNNDGF GKRAYMGINV DYDLEGDYIL EVNSKGPFGK
401  RAPVQKQNSY HGVQQFLEKE GQ
```

Table 5.2 A10B Protein sequence compared to Drosophila YP

Protein	Percent Similarity	Percent Identity
A10B X YP1	76.2	57.6
A10B X YP2	69.8	52.3
A10B X YP3	72.4	54.3

since the A10B protein contains two introns and is of similar size to DmYP$_3$ (420 amino acids) compared to one intron and 438 amino acids in DmYP$_1$, it is most likely that DmYP$_3$ is equivalent to the YP encoded within the A10B genomic fragment. On the other hand, the intron number of the *yp* genes in one species may not be used to decide the equivalent *yp* gene in another species, since in *D. melanogaster* the *yp*$_1$ gene has one intron and in *D. grimshawi* it has two introns. Sequence similarity at the DNA level between the coding regions of the *C. erythrocephala* and *D. melanogaster yp* genes indicates that Ce*yp*$_3$ is more similar to Dm*yp*$_1$ or Dm*yp*$_2$ (Table 5.3).

The intron splice sites between the Ce*yp*$_3$ and Dm*yp*$_3$ are in the same positions; however, the intron sizes differ slightly. Intron 1 in Dm*yp*$_3$ is 62 bp and Intron 2 is 72 bp long, compared to 61 bp and 64 bp respectively in Ce*yp*$_3$. The corresponding introns of each gene appear to have diverged in sequence since there is little similarity between them. This is not suprising since the *yp* coding sequences of *C. erythrocephala* have also diverged when compared to the *D. melanogaster yp* genes (Table 5.3). On the other hand, the amino acid composition of CeYP$_3$ is similar to that of the three *D. melanogaster* YPs (Table 5.4) and a dot plot matrix comparison between DmYP$_3$ and CeYP$_3$ (Figure 5.7) shows a few gaps. The gaps observed are concentrated in the N-terminal end of the protein, as was observed for *Drosophila* by Gabaredian *et al.* (1987)

Table 5.3 Comparison of the DNA coding sequence between C.erythrocephala yp3 gene and the yp gene D.melanogaster

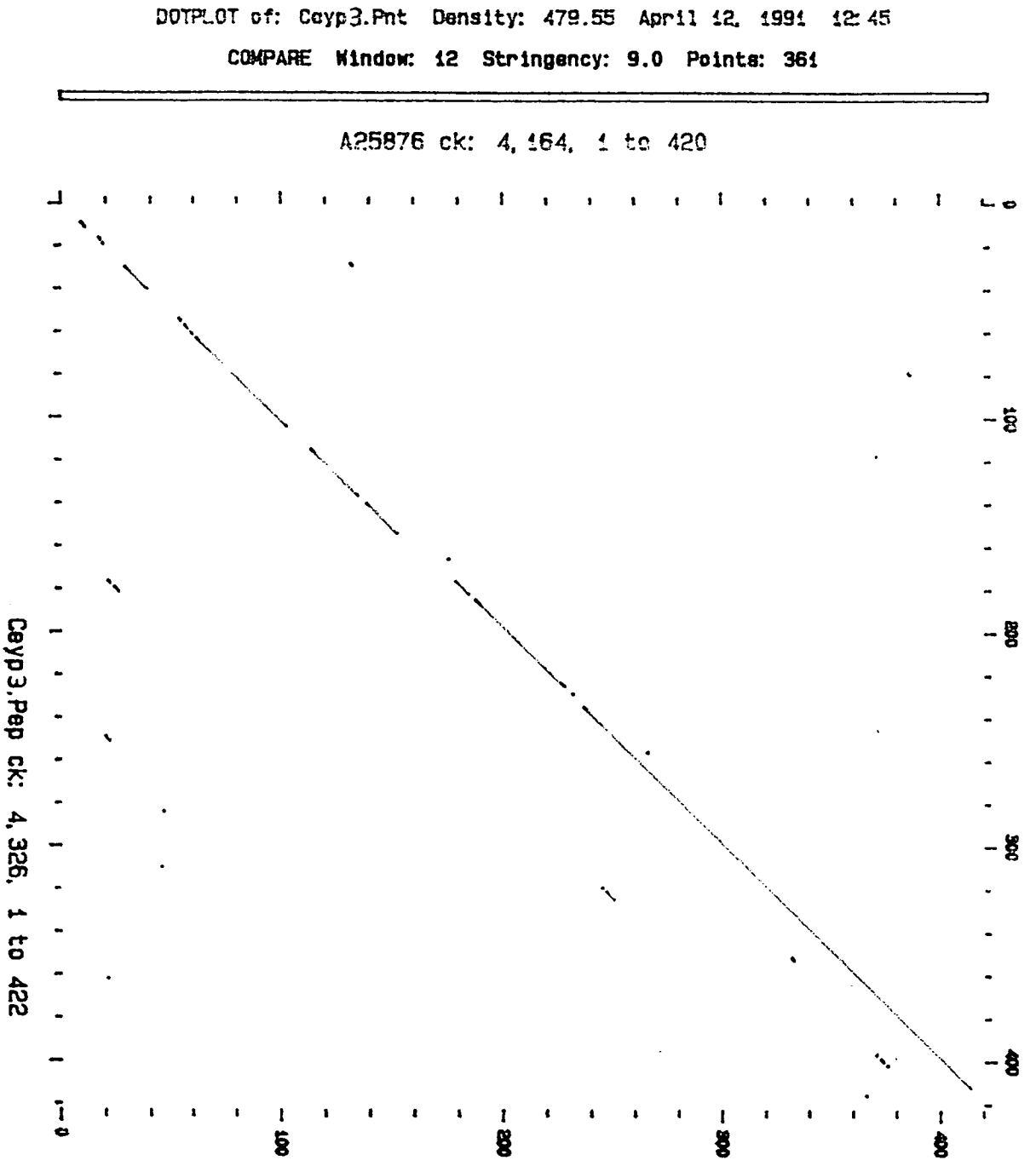
yp gene sequences	% similarity
Ceyp3 X Dmyp1	54.8
Ceyp3 X Dmyp2	53.6
Ceyp3 X Dmyp3	56.6

Table 5.4 Amino acid composition of the Drosophila and Calliphora yolk proteins

Amino acid	Yolk Proteins							
	Dm YP1	%	Dm YP2	%	Dm YP3	%	Ce YP3	%
Ala	38	8.65	26	5.88	41	9.76	37	8.77
Cys	1	0.22	2	0.45	4	0.95	4	0.95
Asp	24	5.47	19	4.5	22	5.23	20	4.74
Glu	24	5.47	32	7.23	20	4.76	27	6.40
Phe	10	2.27	14	3.17	10	2.38	13	3.08
Gly	33	7.51	36	8.14	32	7.62	31	7.34
His	11	2.51	8	1.81	11	2.61	6	1.42
Ile	20	4.55	19	4.50	25	5.95	21	4.98
Lys	26	5.92	21	4.75	28	6.67	28	6.63
Leu	32	7.29	32	7.23	36	8.57	36	8.53
Met	12	2.73	8	1.81	8	1.90	10	2.37
Asn	30	6.83	28	6.33	20	4.76	27	6.40
Pro	22	5.01	22	4.98	22	5.23	19	4.50
Gln	35	7.97	39	8.82	29	6.90	22	5.21
Arg	22	5.01	31	7.01	18	4.28	22	5.21
Ser	26	5.92	28	6.33	27	6.42	39	9.24
Thr	22	5.01	30	6.79	24	5.71	18	4.26
Val	31	7.06	28	6.33	24	5.71	27	6.40
Trp	3	0.68	2	0.45	4	0.95	2	0.47
Tyr	17	3.87	17	3.85	15	3.57	13	3.08
Total	439		442		420		422	

Figure 5.7

A dot plot comparing the YP_3 of *D. melanogaster* and *C. erythrocephala*. The Dot plot was generated using the program "Compare" in GCG (version 6.0) (Devereux *et al.*, 1984).


DOTPLOT of: Ceyp3.Pnt Density: 479.55 April 12, 1991 12:45
COMPARE Window: 12 Stringency: 9.0 Points: 361
A25876 ck: 4, 164, 1 to 420
Ceyp3.Pep ck: 4, 326, 1 to 422
0
100
200
300
400

and $CeYP_2$. Table 5.5, indicates that the *C. erythrocephala* codon usage is different (ie. favouring A/T rich codons) to that of *D. melanogaster* and therefore explains the similarity in amino acid composition and sequence between the *D. melanogaster* and *C. erythrocephala* YPs.

C. Post-translational sites in the *Calliphora* YPs.

i. Signal Peptidase site

The yolk proteins are secreted from the fat body into the haemolymph and from the follicle cells to the spaces between the follicle cells and oocyte. Regardless of the sites of synthesis, the nascent YPs are targeted into the Endoplasmic Reticulum (ER) resulting in the cleavage of the signal peptide and, ultimately, secretion. Yan *et al.* (1987) and Liddell (Ph.D Thesis. 1989), predicted that the signal peptidase site for all *D. melanogaster* YPs was between residues 19 and 20 (Figure 5.8A). In the case of $CeYP_3$, the signal peptidase cleavage site was predicted to be at the same residues as the *D. melanogaster* YPs. A signal peptide is composed of three regions, namely an n-region (positively charge N-terminal end), an h-region (central hydrophobic region) and a c-region (more polar than the h-region and defines the cleavage site). The n-region varies in length from 1 to 5 residues and has at least one charged amino acid (Lys or Arg) (Von Heijne,

Table 5.5 Codon usage frequency for Drosophila and Calliphora

Amino acid	Codon	Drosophila frequency	Calliphora frequency
Alanine	GCG	0.077	0
	GCA	0.083	0.05
	GCT	0.211	0.59
	GCC	0.629	0.36
Cysteine	TGT	0.317	0.17
	TGC	0.683	0.83
Aspartic Acid	GAT	0.477	0.79
	GAC	0.523	0.21
Glutamic Acid	GAG	0.888	0.25
	GAA	0.112	0.75
Phenylalanine	TTT	0.175	0.38
	TTC	0.825	0.62
Glycine	GGG	0.013	0
	GGA	0.317	0.07
	GGT	0.283	0.7
	GGC	0.317	0.23
Histidine	CAT	0.307	0.69
	CAC	0.693	0.31
Isoleucine	ATA	0.02	0.05
	ATT	0.256	0.63
	ATC	0.723	0.32
Lysine	AAG	0.884	0.52
	AAA	0.116	0.48
Leucine	TTG	0.11	0.47
	TTA	0.018	0.17
	CTG	0.595	0.06
	CTA	0.041	0.04
	CTT	0.077	0.15
	CTC	0.16	0.11
Methionine	ATG	1	1
Asparagine	AAT	0.179	0.63
	AAC	0.821	0.37
Proline	CCG	0.136	0
	CCA	0.159	0.1
	CCT	0.162	0.41
	CCC	0.543	0.49
Glutamine	CAG	0.852	0.28
	CAA	0.148	0.72
Arginine	AGG	0.087	0.02
	AGA	0.045	0.19
	CGG	0.075	0.02
	CGA	0.075	0.02
	CGT	0.25	0.52
	CGC	0.468	0.21
Serine	AGT	0.05	0.15
	AGC	0.278	0.22
	TCG	0.206	0.06
	TCA	0.05	0.09
	TCT	0.07	0.25
	TCC	0.345	0.23
Threonine	ACG	0.141	0
	ACA	0.073	0.03
	ACT	0.144	0.43
	ACC	0.642	0.54
Valine	GTG	0.432	0.21
	GTA	0.034	0.16
	GTT	0.155	0.5
	GTC	0.378	0.14
Tryptophan	TGG	1	1
Tyrosine	TAT	0.191	0.58
	TAC	0.81	0.42

Figure 5.8

A.The putative signal peptides of the *D. melanogaster* YPs (Yan *et al.*, 1987; Liddell, Ph.D Thesis, 1989). The arrows indicate the putative site of cleavage by the signal peptidase. The underlined amino acid residues comform to the -3 -1 rule.

B.The putative signal peptide of the *C.erythrocephala* YP_3. The arrows indicate the putative site of cleavage by the signal peptidase and the underlined amino acid comform to the -3 -1 rule.

C.Tyrosine sulfation putative sites in $CeYP_2$ and $CeYP_3$. (based on the consensus derived by Huttner and Baeuerle, 1988; Figure 1.6A). The number indicates the position of the tyrosine residue in the amino acid sequences for the YPs.

A.

DmYP$_1$ MNPMRVLSLLACLAVAALAKPNGRM

DmYP$_2$ MNPLSTLCVMACLLAVAMGNPQSGN

DmYP$_3$ MMSLRICLLATCLLVAAHASKDASN

B.

CeYP$_3$ MNPLRIVCVALLLAAAGSANGNLSG

C.

CeYP$_2$ 304 YMGINTDYDVEGDFI

CeYP$_3$ 383 YMGINVDYDLEGDYI

389 DYDLEGDYILEVNSK

1983; Perlman and Halvorsen, 1983). CeYP$_3$ has an arginine at residue position 5 which is followed by hydrophobic residues, thereby indicating the end of the n-region. The central hydrophobic region of the signal peptide is very distinctive, it may range in size from 7 to 20 residues and is composed of Phe, Leu, Ile, Val, Ala, Met and Trp. In CeYP$_3$ the h-region spans from residue 6 to residue 15 or 17. The c-region varies in size from 4 to 7 residues and the last three are the more easily predicted. Von Heijne (1983) and Perlman and Halvorsen (1983), proposed the -3, -1 rule whereby the -3 and -1 residues (relative to the cleavage site) should have small neutral side chains. The residue usually found in these positions are Ala, Ser, Gly, Cys, Thr, and Gln, although at position -3 Gln is uncommon. At residue -2, large aromatic side chains are preferred although hydrophobic or charged amino acids may be tolerated. Amino acids favouring B-turn structures (Pro, Ser amd Gly) are usually observed in the c-region. In the case of the CeYP$_3$ signal peptide (Figure 5.8B), there are two potential signal peptidase cleavage sites where residues (numbering from the Methionine start site) 17 (Gly), 19 (Ala) and 21 (Gly) amino acids comply with the characters required by the -3, -1 rule. At residue position 18 a Ser is observed, while at position 20 an Asp is found; both of which are tolerated at position -2. In both cases the two amino acids preceding the -3 residues are small neutral residues. The only feature that may make one site more preferable to the

other is that Ala is observed more frequently at position -1 and thus it may be inferred that the cleavage site in $CeYP_3$ is probably between amino acids 19 and 20.

ii. The N-glycosylation sites

The YP_2 protein in *D. melanogaster* has a potential N-glycosylation site at residue 25 (Asn^{25}-Arg-Ser) (Hung and Wensink, 1983). The $CeYP_2$ also has a potential site for N-glycosylation, at residue 181 (Asn^{181}-Pro-Ser) at the beginning of the second exon. There may be other N-glycosylation sites in the first exon of the $CeYP_2$ protein. In contrast to the *D. melanogaster* YP_3 protein (and $DmYP_1$), which does not have any potential N-glycosylation sites, the $CeYP_3$ has one potential site for N-glycosylation at residue 22 (Asn^{22}-Leu-Ser). It is not known if the yolk proteins of *Calliphora* are N-glycosylated or if they are glycosylated at all. However, after the removal of the signal peptide these proteins may be post-translationally modified as they pass through the secretory pathway.

iii. Tyrosine sulfation sites

As described in the Introduction the YPs of *D. melanogaster* are tyrosine sulfated (Baeuerle and Huttner, 1985). In the Introduction the possible sites of tyrosine sulfation for YP_1, YP_2 and YP_3 were discussed as well as

the concensus constraints derived by Hortin *et al.* (1986), and Huttner and Baeuerle (1988) for tyrosine sulfation. Here, the same concensus constraints were used to predict the potential sites of tyrosine sulfation within $CeYP_2$ and $CeYP_3$ (Figure 5.8C). In $CeYP_2$ only one, or perhaps two, sites can be found, in $CeYP_3$ there are 2 possible sites in close proximity to each other. In the concensus constraints rules described for tyrosine competence to sulfation by Hortin *et al.* (1986); Huttner and Baeuerle (1988), there is no evidence for an inhibitory effect on the enzyme's sulfation activity by a sulfated tyrosine in close proximity to another site.

D. Discussion

The main aim of this thesis was to clone and determine the DNA sequence of *yp* genes from a *Drosophila* species whose proteins were taken up specifically by *D. melanogaster* oocytes, but which was distantly related to *D. melanogaster*. In chapter 3, results showed that the YPs from not only *Drosophila* species but also other non-*Drosophila* species were taken up specifically by *D. melanogaster* or *D. funebris* oocytes. This result indicated that the YPs are a well conserved set of proteins for up to 100MYR of Dipteran evolution (this time scale is according to Beverley and Wilson, 1982; 1984). Thus it appeared that cloning and determining the DNA sequence of *yp* genes as distant as possible from *D.*

melanogaster would take advantage of as much sequence divergence as possible in order to determine the domains within the YPs that are required for specific uptake into oocytes.

The DNA sequence of the *yp* genes of *C. erythrocephala* confirmed that the *Calliphora* genome is 20% more A/T rich than the *Drosophila* genome (Houde *et al*, 1989). The general assumption that the main function of the YPs is to provide nutrients to the developing embryo may imply that because of the difference in genome composition the YPs of *Calliphora* may have a different amino acid composition. This however is not so (Table 5.4) since the *D. melanogaster* and *C. erythrocephala* YPs have similar amino acid composition. The maintenance of amino acid sequence similarity (as shown by the YP alignment in Figure 5.10) may be explained by the difference in codon usage between *Drosophila* and *Calliphora* (as shown in Table 5.5, where the *Calliphora* codon usage frequency was derived from the data obtained in this Thesis). There is no doubt that at the DNA level the genes have diverged since the DNA similarity found between the *yp* gene coding sequences for *Drosophila* and *Calliphora* is reduced (Table 5.3). However, the introns and *cis*-regulatory sequences are also very dissimilar. In fact, computer comparisons between introns of the *yp* genes of the two species did not yield any significant similarities, thus suggesting that the constraints in the coding sequence are real. In view of

the fact that the DNA sequence is so dissimilar it is puzzling why the YPs of *Calliphora* and *Drosophila* have not diverged more. To illustrate this point further, sequence comparison studies between gene and protein sequences from different *Drosophila* species have been used successfully to identify sequences of importance either in regulation of gene expression (ie. chorion genes (Orr-Weaver *et al.*, 1991)) or in protein domains whose functions may be of importance (ecdysone inducible protein E74A, Burtis *et al.*, 1990). For most studies *D. virilis* has been the choice of distantly related fly species to *D. melanogaster* and it appears that 60MYR (according to Berverly and Wilson, 1982; 1984) of divergence has been sufficient in those cases to identify sequences of importance. In this thesis the complete sequence of the yp_3 gene and partial sequence of the yp_2 gene of *Calliphora* presented may have diverged too far at the DNA level to identify regulatory sequences (discussed in further detail in chapter six). On the other hand, the protein sequence has been maintained very closely and thus the identification of putative domains involved in interactions with the receptor have been impaired. This in itself may indicate that there is a strong interaction across the whole sequence which maintains the structure of the conserved proteins.

The newly determined *yp* sequences of *Calliphora* show some interesting features when compared to other *yp* genes.

For instance, the intron-exon boundaries appear to be well conserved when compared to *D. melanogaster yp* gene sequences. In *D. grimshawi*, the intron sizes have been estimated by S1 mapping analysis and appear to be double the size of those in *D. melanogaster* and *Calliphora* (ie, 120 bp compared to 60 bp). As discussed in the introduction (Figure 1.4, summarised in Table 5.6) *D. grimshawi* has one intron in yp_2 and two introns each in yp_1 and yp_3 (Hatzopoulus and Kambysellis, 1987). Therefore intron number can not be adequately used to determine the *yp* equivalent to *D. melanogaster* of the *yp* sequences from *Calliphora*. The cloning of the yp_1 gene of *Calliphora* should provide a clue as to the number of introns the ancestral *yp* gene may have. Intron number and position are conserved in a number of other genes studied in different *Drosophila* species, such as the pupal cuticle protein genes (Moriyama and Gojobori, 1989) *engrailed* (Kassis *et al.*, 1986) *hunchback* (Treier *et al.*, 1989) and *hsp82* (Blackman and Meselson, 1986).

The isolation and DNA sequence of three *yp* genes in *D. melanogaster* (Barnett *et al.*, 1980; Hung and Wensink, 1982; 1983; Garabedian *et al.*, 1987), and isolation of three *yp* genes in *D. grimshawi* (Hatzopoulus and Kambysellis, 1987), as well as the number of *yp* gene sites in the polytene chromosomes in *D. melanogaster* sibling species (Kozma and Bownes, 1986) and *D. planitibia* (Dong *et al.*, 1985), has been used as evidence to support the

Table 5.6 Comparison between D. melanogaster, D. grimshawi C. erythrocephala number of introns in yp genes

Gene	D. melanogaster	D.grimshawi	C. erythrocephala
yp1	1	2	to be determined
yp2	1	1	1
yp3	2	2	2

idea that the three *yp* genes evolved from one ancestral gene though a series of duplication events early in *Drosophila* evolution (Yan *et al*, 1987). The sequence similarity between the three YPs, where YP_2 is more similar to YP_1 than to YP_3, was also used to support the proposed duplication. Therefore, the duplication of an ancestral gene gave rise to yp_3 and to the ancestor of yp_1/yp_2, and this in turn duplicated again to give rise to yp_1 and yp_2 (Yan *et al.*, 1987). Yan *et al.* (1987) support their claim by suggesting that only intron loss occurs during evolution, and thus, in *D. melanogaster* yp_1 and yp_2 arose from an ancestral gene which lost an intron prior to duplication. However, intron addition is thought to have occurred in the evolution of the *D. melanogaster* acetylcholinesterase gene since it contains six more introns than the vertebrate counterpart, but has the same number of introns as the homologue domain of the bovine thyroglobulin gene (Fournier *et al.*, 1989).

The generation of genes through duplication often gives rise to isozymes, or genes whose protein products are found at different specific stages or tissues during the lifetime of the organism. In *Drosophila* the *Adh* gene is a good example of duplication since there are isozymes present in the *mulleri* and *hydei* subgroups, whereas the other members of the *repleta* group (*repleta* and *mercatorum*) do not have isozymes. Thus the duplication event can be placed in the origin of the repleta group

(Batterham *et al.*, 1984). In the case of the *yp* genes the three genes code for proteins which are expressed in the same tissues with the same temporal and spatial distribution. Although there are *Drosophila* species with a range of YPs (ie, one to three proteins), the corresponding genes coding for this apparent number of proteins has not been studied. If a duplication event took place to generate the *yp* genes, it must have been early in *Drosophila* evolution since both *D. melanogaster* and *D. grimshawi* have the same genomic organisation (Figure 1.4). The fact that more distantly related Diptera have three and two yolk proteins may indicate that the evolution of these genes is complex and may have undergone different pathways in different orders. For instance, the origin of *yp* genes in the *melanogaster* subgroup may be as outlined by Yan *et al.* (1987), whereas for *D. grimshawi* the intron loss may have occurred after the second duplication event, thus leaving yp_1 and yp_3 with two introns and yp_2 with one. In any case, further studies on species closely and distantly related to *D. melanogaster* should follow, to try to understand the genesis of the three *yp* genes in *D.melanogaster*, *D. grimshawi* and *C. erythrocephala*.

At the protein level, the YPs in *D. melanogaster* and *Calliphora* show a number of conserved characters, besides the general sequence identity. The signal peptidase site appears to be conserved in YPs of both species. These

have been predicted to be between residues 19 and 20 of the nascent proteins. In the case of *Calliphora*, potential N-glycosylation sites were found in both YP proteins ($CeYP_3$ and $CeYP_2$); this is in contrast to only $DmYP_2$ having a potential N-glycosylation site. On the other hand it is not known if the *Calliphora* YPs are glycosylated. The glycosylation of vitellogenins in other insects is known to be important in uptake. This was shown when *Nauphoeta cinerae* vitellogenins bind oocyte membrane preparations of *Leucophaea maderae*. The reciprocal experiment did not result in binding of *Leucophaea maderae* vitellogenin to *Naphoeta cinerae* oocyte membranes (Konig and Lanzrein, 1985). The differences between the two set of proteins were in the glycosylation pattern. In this case it would appear that the carbohydrates are playing a role in receptor recognition. However, glycosylation may also play an important role in the stabilisation of the YPs as has been shown for the vitellogenins of *Blatella germanica* (Gochoco *et al.*, 1988) and *Manduca sexta* (Osir *et al.*, 1986). Finally, tyrosine sulfation sites have been predicted in the YPs of *Calliphora* (Figure 5.8C). The effect of the tyrosine sulfation in the YPs *in vivo* is not known and remains to be studied. However, *in vitro* the sulfation of the $DmYP_2$ appears to maximise the efficiency of secretion (Friederich *et al.*, 1988). The presence of tyrosine sulfation sites in the YPs of *Calliphora* and *D. melanogaster* may indicate that this post-translational

modification plays an important role in the secretion of the YPs. On the other hand, the tyrosine sulfation may have a role in the stabilisation of the YP structure or the protection of tyrosine residues in the extracellular medium.

In *Drosophila melanogaster* a number of YP mutants have been isolated. Some of these mutants have been characterised and appear to affect secretion of the YPs. YP_3^{51} is a mutant where the secretion of YP_3 is modified, the mutant YP_3^{51} protein is accumulated in the fat body of female flies. Molecular analysis of this mutation determined an amino acid substitution in the signal peptide which is thought to prevent the processing of the nascent YP_3 protein (Figure 5.9A) (Liddell and Bownes, 1991a). In *D. melanogaster* and *C. erythrocephala*, although the amino acid residue is not conserved, the character of the amino acids is the same and thus of importance to the functioning of the signal peptide. A mutation affecting YP_1, *fs(1) 1163*, has been characterised and found to have a single amino acid mutation which affects the secretion of the YP_1^{1163} (Figure 5.9B) (Saunders and Bownes, 1986). The mutation at this site changes the amino acid from a small uncharged residue (Ile^{92}) to a charged amino acid (Asp^{92}) thus having a destabilising effect. Two other mutations have been isolated which affect YP_2 secretion, *K313* and $YP_2^{12\text{-}1245}$. Molecular analysis of *K313* showed two amino acid

Figure 5.9

The mutations studies in YPs of *D. melanogaster* and their comparison to the aligned YP sequences of *C. erythrocephala* and *D. melanogaster*.

A.Mutation in YP_3 of the mutant YP_3^{s1} in which Ala^{10} has been mutated to Asp^{10} (Liddell and Bownes, 1991).

B.Mutation in the YP_1 protein of the mutant YP_1^{1163} in which residue 92 changed from Ile to Asn (Saunders and Bownes, 1986).

C.One of the two mutations affecting the YP_2 product of the mutant *K313*. The mutation changes amino acid residue 68 from Met to Leu (Williams *et al.*, 1987).

D.The second mutation affecting the mutant *K313* YP_2. The mutation is in residue 141 changing Pro for Leu (Williams *et al.*, 1987).

```
A. CeYP3    1 MNPLRIVCVA--LLLAA  15
   DmYP1    1 MNPMRVLSLLACLACAA  17
   DmYP2    1 MNPLRTLCVMACLAVAM  17
   DmYP3    1 MMSLRICLLATCLLVAA  17
               *   *   .   *. *

B. CeYP2   12 SFVPKASEIP  21
   CeYP3   82 SFVPSSSNVP  91
   DmYP1   86 NYVPSGIQV-  94
   DmYP2   91 DYTPEPSQIR 100
   DmYP3   81 SFVPSPSNVP  90
                ...*.  ..

C. CeYP3   52 QKLESMSVEK  61
   DmYP1   56 ERLENMNLER  65
   DmYP2   61 KKLQEMSMEE  70
   DmYP3   51 ERLENQPLEQ  60
                .* .   *

D. CeYP2   57 ITGLPQKTE  65
   CeYP3  127 ITGLPATTE 135
   DmYP1  130 VTGLPQTSE 138
   DmYP2  137 IQGLPETNT 145
   DmYP3  127 LTGLPKTSP 135
                . *** ..
```

substitutions which are thought to be the cause of the secretion defect observed (Williams *et al.*, 1987). The first mutation changes residue 68 from Met to Leu (Figure 5.9C). This mutation introduces a hydrophobic amino acid in a position where all other YPs sequenced have a hydrophobic residue. It is not clear what effect this change may have on the YP structure. The second mutation (Figure 5.9D) changes the amino acid residue 141 from Pro to Leu. Pro, at this position is conserved through all YPs sequenced and is situated amongst a cluster of well conserved amino acids. This mutation changes a strong alpha chain breaker to a strong alpha chain promotor and thus is expected to have a marked effect in the pattern of protein folding. Although $YP_2^{12-1245}$ has yet to be characterised, this mutant does produce a yp_2 transcript in both fat bodies and ovaries (Martinez and Bownes, unpublished); however, the YP_2 protein is not secreted into the haemolymph nor accumulated in the oocytes. Therefore this mutant may not synthesise a YP_2 protein since accumulation of the potential YP_2 product is not observed in the ovaries. The only mutants to have been characterised are those affecting secretion suggesting that mutations in the *yp* gene coding sequences affect the general protein structure.

These are readily identified by the lack of YPs in the haemolymph while mutants affecting uptake are not. This suggests, therefore, that the

mutations so far investigated may not only affect the general protein structure, but also secretion signalling domains.

A computer database search for conserved domains within the YPs

In general terms the amino acid sequence of YPs from both *D. melanogaster* and *C. erythrocephala* are very similar as shown by their alignment in Figure 5.10. It is also clear that the sequence identity is higher in the C-terminal 2/5 of the YPs (20.6%) compared to the N-terminal 3/5 (47.1%). This was also observed by Garabedian *et al* (1987) when comparing the sequences of the *D. melanogaster* YPs. The protein sequences of $CeYP_2$ and $CeYP_3$ are similar to $DmYP_2$ and $DmYP_3$ respectively when compared to each other (Figure 5.4 and 5.7). Furthermore, Pro residues (ie. 13 residues out of 19) and Gly residues (ie. 19 residues out of 32) through out the YPs are well conserved. Such position conservation of these amino acids has been reported to indicate a close maintenance of the secondary structure of other compared proteins (Jeffery *et al.*, 1987).

The evolutionary approach to determining domains within a protein which are of importance for their activity has been used succesfully, for example in the comparison of vitellogenins from nematodes and vertebrates

Figure 5.10

Sequence alignment of *D. melanogaster* and *C. erythrocephala* YPs. The $CeYP_2$ sequence is incomplete. The alignment produces a "*" for an amino acid residue which is identical in all sequences of the alignment. The dot (.) signifies a conservative substitution of the residue in one or more proteins of the alignment.

$CeYP_2$ = *C. erythrocephala* yolk protein 2.

$CeYP_3$ = *C. erythrocephala* yolk protein 3.

$DmYP_1$ = *D. melanogaster* yolk protein 1.

$DmYP_2$ = *D. melanogaster* yolk protein 2.

$DmYP_3$ = *D. melanogaster* yolk protein 3.

Gap fixed = 10 Gap vary. = 10 (program CLUSTAL)
* :=> match across all seqs.
. :=> conservative substitutions

Vitellogenin sequence multiple alignment

```
CeYP3   MNPLRIVCVA--LLLAAAGSAN-----GNL--SGLNKLRPSQWLSSSQLDKLPRDMEISL
CeYP2   ------------------------------------------------------------
DmYP1   MNPMRVLSLLACLAVAALAKPN-----GRMDNSVNQALKPSQWLSGSQLEAIPALDDFTI
DmYP2   MNPLRTLCVMACLLAVAMGNPQSGNRSGRRSNSLDNVEQPSNWVNPREVEELPNLKEVTL
DmYP3   MMSLRICLLATCLLVAAHAS----------KDASNDRLKPTKWLTATELENVPSLNDITW

CeYP3   QKLESMSVEKGAELMQKLYHLSQINNDLKPSFVPSSSNVPCYIVKPNGKKVSTSLDKLAS
CeYP2   ------------------HLSQAGRAIEPSFVPKASEIPAFLITPDNKKVNFKLSELPK
DmYP1   ERLENMNLERGAELLQQVYHLSQIHHNVEPNYVPSGIQV--YVPKPNGDKTVAPLNEMIQ
DmYP2   KKLQEMSMEEGATLLDKLYHLSQFNHVFKPDYTPEPSQIRGYIVGERGQKIEFNLNTLVE
DmYP3   ERLENQPLEQGAKVIEKIYHVGQIKHDLTPSFVPSPSNVPVWIIKSNGQKVECKLNNYVE
                          *..*  .  .*...*.  ..  ..    ..*.   *..

CeYP3   ACKQQPNFGEEEVTILITGLPATTETVRKANRKLIDAYLQRYSTKRQQP----SKFDYSG
CeYP2   IAKEEKSFGDEEVTVYITGLPQKTETVKKATRKLVQAYMQRYNGQAPERNS----VRYED
DmYP1   RLKQKQNFGEDEVTIIVTGLPQTSETVKKATRKLVQAYMQRYNLQQQRQHGKNGNQDYQD
DmYP2   KVKRQQKFGDDEVTIFIQGLPETNTQVQKATRKLVQAYQQRYNLQPYETT------DYSN
DmYP3   TAKAQPGFGEDEVTIVLTGLPKTSPAQQKAMRRLIQAYVQKYNLQQLQKNAQEQQQQLKS
          * .  **..***. . *** ..   .** *.*..** *.*. .           .

CeYP3   E-----KMARTSSEED-SNEWQNQQASSGNLVIIDLGNELNSFKRFSLLDVDETGAMIRS
CeYP2   DSSEKRNPSSSEEDEWKSN----SNKPSGNLVVIELGNVLTNMKNYVFLDVEKTGMEIGD
DmYP1   QSNEQRKNQRTSSEEDYSEEVKNAKTQSGDIIVIDLGSKLNTYERYAMLDIEKTGAKIGK
DmYP2   EEQSQRSSSEEQQTQRRKQNGEQDDTKTGDLIVIQLGNAIEDFEQYATLNIERLGEIIGN
DmYP3   SDYD------YTSSEEAADQWKSAKAASGDLIIIDLGSTLTNFKRYAMLDVLNTGAMIGQ
                    ..   .    .  .*....*.**. ... ...  *..   *  *

CeYP3   AIVEMTEKCDVADETIHVVAQGIAAHVAGAAGNEFTRKTGRQLRRITALDPSKILAKNPH
CeYP2   ILVQCTDKADVPQEIVHVIGSNIAAHVAGAVGRQFTRETGHQLRRITGLDPSKIYAQTQK
DmYP1   WIVQMVNELDMPFDTIHLIGQNVGAHVAGAAAQEFTRLTGHKLRRVTGLDPSKIVAKSKN
DmYP2   RLVELTNTVNVPQEIIHLIGSGPAAHVAGVAGRQFTRQTGHKLRRITALDPTKIYGKPEE
DmYP3   TLIDLTNK-GVPQEIIHLIGQGISAHVAGAAGNKYTAQTGHKLRRITGLDPAKVLSKRPQ
          ... ... ... ...*... . .*****......*  **..***.*.***.*. ..  .

CeYP3   SLTGLSRGDAEFVDAIHTNVYGMGTIQRVGDIDFYPNGPSAAVPGAQSIVEASMRATRYF
CeYP2   SLRGLARGDAEFVDVIHTSAYGWGTPTRCGDVDFYPNGPSRGVPGADNVVEASMRAIRYF
DmYP1   TLTGLARGDAEFVDAIHTSVYGMGTPIRSGDVDFYPNGPAAGVPGASNVVEAAMRATRYF
DmYP2   RLTGLARGDADFVDAIHTSAYGMGTSQRLANVDFFPNGPSTGVPGADNVVEATMRATRYF
DmYP3   ILGGLSRGDADFVDAIHTSTFAMGTPIRCGDVDFYPNGPSTGVPGSENVIEAVARATRYF
          * **.****.***.***.... **  * ...**.****. .***. ...**  **.***

CeYP3   AESVRPGNERNFPAVAANSLKEYKNNDGFGKRAYMGINVDYDLEGDYILEVNSKGPFGKR
CeYP2   AESVVPGNERNFPAVGATSLKQYKEQNGNGKRVYMGINTDYDVEGDFILQVNGKSPFGRS
DmYP1   AESVRPGNERSFPAVPANSLQQYKQNDGFGKRAYMGIDTAHDLEGDYILQVNPKSPFGRN
DmYP2   AESVRPGNERNFPSVAASSYQEYKQNKGYGKRGYMGIATDFDLQGDYILQVNSKSPFGRS
DmYP3   AESVRPGSERNFPAVPANSLKQYKEQDGFGKRAYMGLQIDYDLRGDYILEVNAKSPFGQR
        **** **.**.**.* *.* ..**...* *** ***.... *. **.**.** *.***..

CeYP3   APVQKQNSYHGVQQFL--E-----KEGQ
CeYP2   TPAQKQQNYHNVHKPWKMSSRD-----F
DmYP1   APAQKQSSYHGVHQAW--NTNQDSKDYQ
DmYP2   TPAQKQTGYHQVHQPWRQSSSNQGSRRQ
DmYP3   SPAHKQAAYHGMHH----------AQN
        .*..** .** ...
```

(Spieth *et al.*, 1991) and in the identification of some clathrin light chain motifs (reviewed in Brodsky *et al*, 1991). The data obtained from examining sequences of the same protein in different species has also been used to determine rates of divergence or amino acid replacement, as in the case of Adh in *Drosophila* (Villaroya and Juan, 1991). This thesis uses the same principal to identify sequences of domains which may prove important for uptake of the YPs. The alignment between the CeYPs and the DmYPs (Figure 5.10) does not appear to indicate specific domains of importance. A computer database search was carried out using the whole of the $CeYP_3$ protein sequence. The results from this search suggested five main regions namely, residues 19-95, 117-160, 170-215, 215-343 and 344-405 (the amino acid residues positions are relative to $CeYP_3$), in which to execute detailed database searches. Although the strongest degree of similarity is observed in the region of residues 344-405, all regions are well conserved (with the exception of the region 170-215). The main aims of these database searches was to find similarities to sequences of other proteins which may lead to identifying sites that are known to be involved in receptor/ligand interactions. However, other features important to YP function may be detected in these searches.

i. Region 19-95

The signal peptidase site has been postulated to be between residues 19 and 20. Therefore residues 1 to 19 are discounted from the searches as they are not of importance in the secreted YP and thus in the specific uptake of YPs by oocytes.

The database search found similarities to the C-terminal end of the hemagglutinin protein of A/Chicken/Scotland/59 influenza virus (residues 514 to 525) (Figure 5.11A), this area is well conserved between the hemagglutinin proteins of A/turkey/Ireland/83, A/chicken/Scotland/59 and A/chicken/Pennsylvania/83 (Figure 5.11B)(DeBarun *et al.*, 1988). The significance of this sequence in hemagglutinin is not known and therefore it does not suggest any possible functions in the YPs.

The search also identified close similarities to different alleles of the T-cell antigen receptor beta chain V region from mouse (Figure 5.12). The T-cell antigen receptor is composed of four regions, namely a leader, a variable region, a J region and a constant region. This structure resembles that of immunoglobulins (Hendrick *et al.*, 1984). The T-cell antigen receptor is known to interact with the $\alpha 1$-$\alpha 2$ domains of the major histocompatibility complex (MHC) class I (Mayer *et al.*, 1988). The identified portion of the T-cell antigen

Figure 5.11

Alignment between the YPs of *D. melanogaster* and *C. erythrocephala*, and hemagglutinin from the avian influenza virus, A/Chicken/Scotland/59 (HA59). A.Influenza virus vs YPs

B.Alignment of the hemagglutinin proteins from Influenza viruses A/Chicken/Scotland/59, A/Turkey/Ireland/83 and A/Chicken/Pennsylvania/83.

$CeYP_3$= *C. erythrocephala* yolk protein 3.

$DmYP_1$= *D. melanogaster* yolk protein 1.

$DmYP_2$= *D. melanogaster* yolk protein 2.

$DmYP_3$= *D. melanogaster* yolk protein 3.

HAS= hemagglutinin protein from A/Chicken/Scotland/59

HAP= hemagglutinin protein from A/Chicken/Pennsylvania/83

HAT= hemagglutinin protein from A/Turkey/Ireland/83

A. CeYP$_3$ 48 EISLQKLESMSV 59

DmYP$_1$ 51 DFTIERLENMNL 63

DmYP$_2$ 56 EVTLKKLQEMSM 68

DmYP$_3$ 46 DITWERLENQPL 58

HAS 514 EISGVKLESMGV 525

. . .*.. .

B. HAS 514 EISGVKLESMGV 525

HAT EIDGIKWESMGI

HAP EISGIKLESMGI

.*.* **.

Figure 5.12

Alignment of the YPs of *D. melanogaster* and *C. erythrocephala* with the T-cell antigen receptor beta chain V region

$CeYP_3$= *C. erythrocephala* yolk protein 3.

$DmYP_1$= *D. melanogaster* yolk protein 1.

$DmYP_2$= *D. melanogaster* yolk protein 2.

$DmYP_3$= *D. melanogaster* yolk protein 3.

T-CR= T-cell antigen receptor beta chain V region.

```
CeYP3      58 SVEKGAELMQKLYHLSQ-INNDLKPS-FVP 85
DmYP1      61 NLERGAELLQQVYHLSQ-IHHNVEPN-YVP 88
DmYP2      67 SMEEGATLLDHLYHLSQ-FNHVFKPD-YTP 94
DmYP3      57 PLERGAKVIEKIYHLGQ-IKHDLTPS-FVP 86
T-CR       56 SAEKPPELME-LYNLKQLIRNETVPSRFIP 88
                  *  .  .  .* * *       *    *
```

receptor similar to the YPs (amino acid residues 56 to 84 of $CeYP_3$) is in the variable region between the two cysteine residues that form a disulfide bond. Thus this portion of the T-cell antigen receptor protein may be antigenic or may be involved with ligand interactions. In the YP context, even though the chain is not supported by two cysteines it may behave in a similar fashion.

ii. Region 117-160

This region of the yolk proteins has 69% similarity and is thus well conserved. The database searches find a number of receptor molecules which all display similarity with the same segment within this region (ie, from 117 to 138 amino acid residues) (Figure 5.13A). Of interest is the fact that the human, chimpanzee and rhesus CD4 molecules are amongst the receptor molecules with similarities to $CeYP_3$. The region of the CD4 receptor with similarity to the YPs covers 16 residues (residues 90 to 106). The first seven amino acids (residues 90 to 97) of the CD4 receptor are part of the CDR3 which lie in the v-region- like domain known to interact with HIV (Camerini and Seed, 1990). Point mutations in residues 87 and 89 of CD4 or residues 3 and 5 of CDR-3 affect the formation of syncytium. However, in these mutations the static interactions between the CD4 and HIV were unaltered, thus suggesting that the interaction continues to take place while the induced conformational change in CD4 due to the

Figure 5.13

Sequence similarities observed within the region 117 to 160 of the YPs.

A.Alignment of the YPs from *D. melanogaster* and *C. erythrocephala* with the CD4 receptor molecule.

B.Alignment of the YPs with a receptor molecules which matched to the same region of the $CeYP_3$. The receptor with a match to the YPs were, namely Chicken neural cell adhesion protein (NAM), Rabbit Dihydropyridine receptor (DPR), *Bradyrhizobium japonicum* Heme binding protein (HBR), Rabbit hepatic Growth Hormone receptor (CBP), Rabbit hepatic Serotonin receptor (SR) and Bovine GTPase-activating protein(GTP-AP).

*= indicates conserved residue in all sequences.

.= indicates conservative substitution.

Bold letters indicate conserved residue between the particular protein the the YPs.

Underlined letters indicate a conservative substitution between the particular protein and the YPs.

A.

CeYP$_2$	46	EKSFGDEEVTVYITGLPQKTET	69
CeYP$_3$	116	QPNFGEEEVTILITGLPATTAT	137
DmYP$_1$	118	KQNFGEDEVTIIVTGLPQTSET	141
DmYP$_2$	125	QQKFGDDEVTIFIQGLPETNTQ	149
DmYP$_3$	115	QPGFGEDEVTIVLTGLPKTSPA	136
CD4	85	EVEDQKEEVQLLVFGLTANSDT	106

. . ..** . . ** .

B.

CeYP$_2$	46	EKSFGDEEVTVYITGL-PQKTETVK	71
CeYP$_3$	116	QPNFGEEEVTILITGL-PATTATVK	137
DmYP$_1$	118	KQNFGEDEVTIIVTGL-PQTSETVK	143
DmYP$_2$	125	QQKFGDDEVTIFIQGL-PETNTQVQ	151
DmYP$_3$	115	QPGFGEDEVTIVLTGL-PKTSPAQQ	138
NAM	431	EGTITISGLKPETTYSVR	448
DPR	1511	PPIGDDEVTV	1520
HBP	169	PYFGESIVTLL	179
CBP	120	PSFNQEDIYITTESLTTTAES	140
SR	550	PSFNQEDIYITTESLTTTAE	577
GTP-AP	159	PEYEEEEVAIPLTAPPT	175

. . . .

interaction with HIV does not take place (Camerini and Seed, 1990). Figure 5.13B, shows the receptor molecules identified by the same area of the CeYP$_3$. In none of these receptors is the function for that particular peptide known, it may be responsible for interactions with the respective ligand, or it may be involved in a specific type of protein folding. Thus, in the case of the YPs it may indicate a site for interaction with other YPs to form oligomers. Accordingly, a mutation in the DmYP$_2$ at residue 141 changing Pro for Leu results in a secretion defect mutation (Williams *et al.*, 1987). This mutation has an effect on the structure of the YP since Pro is an alpha chain breaker and Leu is an alpha chain promotor, thus suggesting that this portion of the protein is important for the overall structure of the YP structure.

iii. Region 170-215

Although the alignment shows a poor relationship of the YPs in this region, a conserved motif (SSEE) is found in all the proteins but is not easily aligned (Figure 5.10). This motif appears to be surrounded by sequences that are quite dissimilar. This region (residues 170-215), however, shows similarities to serine protease inhibitors and hormone regulated protease inhibitors in a region 9 to 13 residues away from the sequence responsible for the inhibitory activity (Tecce *et al.*, 1986; Le Cam *et al.*, 1987; Ohkubo *et al.*, 1991) (Figure 5.14A). The

Figure 5.14

Matches observed in the database searches using the YP region 170-217.

A.Alignment of the YPs of *D. melanogaster* and *C. erythrocephala* with a serine protease inhibitor. Only one sequence is shown since all protease inhibitors obtained share the same sequence.

B.Alignment of the YPs of *D. melanogaster* and *C. erythrocephala* with Calmodulin, Acetyl choline receptor and Nicotinic acid receptor.

C.The primary structure of calmodulin is Based on Klee *et al* (1980) and adapted by Kink *et al*. (1990). The amino acid sequence shows the four calcium binding pockets which are marked I-IV. The area of homology with the YPs is boxed.

$CeYP_2$= *C. erythrocephala* yolk protein 2.

$CeYP_3$= *C. erythrocephala* yolk protein 3.

$DmYP_1$= *D. melanogaster* yolk protein 1.

$DmYP_2$= *D. melanogaster* yolk protein 2.

$DmYP_3$= *D. melanogaster* yolk protein 3.

Cmd= Calmodulin

ACE= Acetyl choline receptor

NIC= Nicotinic acid receptor

SPI= Serine protease inhibitor

*= Indicates a conserve position in the alignment.

.= Indicates a conservative substitution in the position.

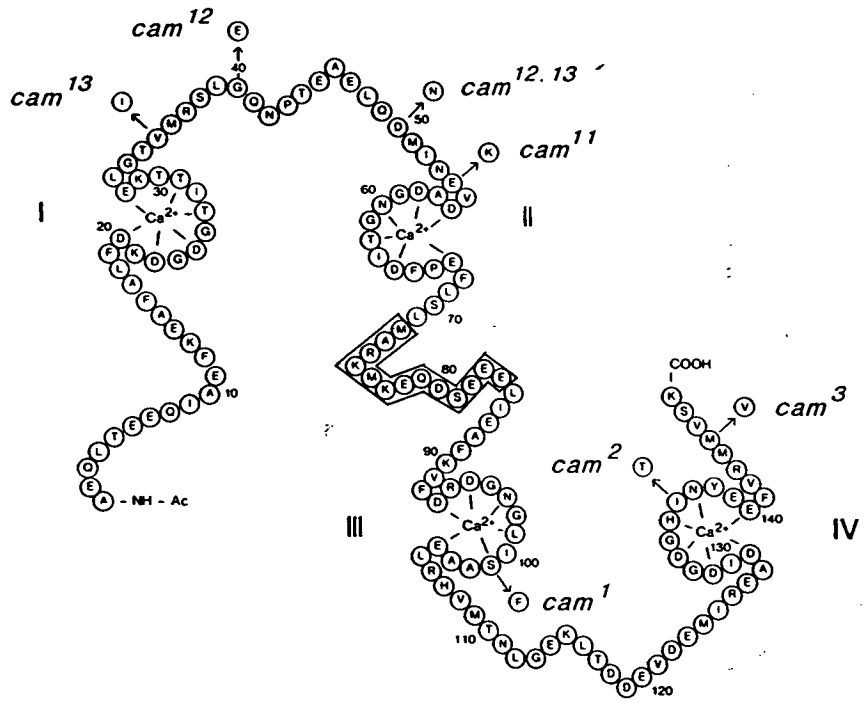
A.

$CeYP_2$	140	LDVEKTG	146
$CeYP_3$	210	LDVDETG	216
$DmYP_1$	221	LDIEKTG	227
$DmYP_2$	222	LDIEKLG	228
$DmYP_3$	212	LNILRTG	218
SPI		LDVDETG	
		*.. *	

B.

$CeYP_3$	170	MARTSSEEDSNE	181
Cmd	72	MARKMKEQDSEE	83
ACE	43	MSRADESEEEQP	54
NIC	267	MTRPTSDEENNQ	278
		* * . ..	

C.



serine protease inhibitors are closely related to the proteases (77% identity). The reactive site is made up of the lysine-serine dimer in which lysine has been found to be of importance in inhibiting trypsin activity (Le Cam *et al.*, 1987). The amino acid residues 210 to 216 of the YPs have similarity to the serine protease inhibitors; however, the YP alignment shows that there is not a conserved lysine-serine dipeptide and therefore the amino acid similarity may only have a structural function within the YPs.

The database searches in this region also identified a sequence similarity with calmodulin (Kink *et al.*, 1990) and calcium channel alpha I chain (dehydropyridine sensitive) (Mikami *et al.*, 1989) (Figure 5.14B). The similarity observed between amino acid residues 170 to 181 of CeYP$_3$ and calmodulin is not part of the calcium binding site in calmodulin (Figure 5.14C), nor is there any similarity to the whole of the putative divalent cation binding sites DXDXDGXXD or DXDXDGXXDXXE (Szebenyi *et al.*, 1981).

iv. Region 215-343

The YPs of *D. melanogaster* have been shown to have a domain with high similarity to the vertebrate lipases, which spans 150 amino acids from 198 to 348 (as observed by Bownes *et al.*, 1988; Perssons *et al.*, 1989). Figure

5.15, shows an alignment of the *Drosophila* YPs, *Calliphora* YPs and three lipases (human hepatic lipase, rat lipoprotein lipase and porcine pancreatic lipase). This alignment illustrates that the YP similarity to vertebrate lipases is maintained in *Calliphora*. The number of identities between the *D. melanogaster* YPs and the vertebrate lipases was of 34 residues (Bownes *et al.*, 1988) compared to 24 residues when both *Drosophila* and *Calliphora* YPs are considered.

The maintenance of the YPs similarities to vertebrate lipases (Figure 5.15) through at least 60MYR and up to 100MYR of Dipteran evolution (according to Beverley and Wilson 1982; 1984) lends support to this region of similarity being an important constraint in YP structure and protein sequence. Moreover, it may indicate that the proposed ecdysteroid binding by *D. melanogaster* YPs may also occur in YPs from *Calliphora*. This, however, cannot be assumed to be true until the *Calliphora* YPs are analysed for ecdysteroid contents and in a later date for their ability to bind ecdysteroid conjugates.

Yan *et al.* (1987) showed that the tripeptide Arg-Gly-Asp (RGD) is conserved in the three *D. melanogaster* YPs, and Figure 5.10 shows that this is also true for both *Calliphora* YPs (from amino acid residue 288 to 290 relative to CeYP$_3$). This tripeptide has been shown to have specific affinity to fibronectin receptor,

Figure 5.15

Sequence alignment of the *D. melanogaster* (Dm) and *C. erythrocephala* (Ce) yolk proteins (YPs) with the rat hepatic triacylglyceride lipase (Rhtl), pig pancreatic triacylglyceride lipase (Lipg) and human lipoprotein lipase (Lihul). This alignment includes 140 amino acid residues of each protein. The similarity between the *D. melanogaster* and lipase proteins has been reported by Bownes *et al.* (1988) and Perssons *et al.* (1989). The RGD motif is boxed.

*= Indicates that position is conserved in all proteins.

.= Indicated a conservative substitution in that position.

```
Dmyp1   196.......  NAKTQSGDII VIDLGSKLNT YERYAMLDIE KTGAKIGKWI
Dmyp3   205.......  SAKAASGDLI IIDLGSTLTN FKRYAMLDVL NTGAMIGQTL
Ceyp3   183.......  NQQASSGNLV IIDLGNELNS FKRFSLLDVD ETGAMIRSAI
Dmyp2   197.......  QDDTKTGDLI VIQLGNAIED FEQYATLNIE RLGEIIGNRL
Ceyp2   115.......  NSNKPSGNLV VIELGNVLTN MKNYVFLDVE KTGMEIGDIL
 Rht1   110.......  SRQSQPVNVG LVDWISLAYQ HYAIAVRNTR VVGQEVAALL
Lihu1   101.......  KREPDS.NVI VVDWLSRAQE HYPVSAGYTK LVGQDVARFI
  Lipg   95........  K..VESVNCI CVDWKGGSRT GYTQASQNIR IVGAEVAYFV
                        .   . .  ..... . .           .   .*  .   .

Dmyp1   VQMVNELDMP FDTIHLIGQN VGAHVAGAAA QEFTRLTGHK LRRVTGLDPS
Dmyp3   IDLTNK.GVP QEIIHLIGQG ISAHVAGAAG NKYTAQTGHK LRRITGLDPA
Ceyp3   VEMTEKCDVA DETIHVVAQG IAAHVAGAAG NEFTRKTGRQ LRRITALDPS
Dmyp2   VELTNTVNVP QEIIHLIGSG PAAHVAGVAG RQFTRQTGHK LRRITALDPT
Ceyp2   VQCTDKADVP QEIVHVIGSN IAAHVAGAVG RQFTRETGHQ LRRITGLDPS
 Rht1   LWLEESMKFS RSKVHLIGYS LGAHVSGFAG S..SMGGKRK IGRITGLDPA
Lihu1   NWMEEEFNYP LDNVHLLGYS LGAHAAGIAG S..L..TNKK VNRITGLDPA
  Lipg   EVLKSSLGYS PSNVHVIGHS LGSHAAGEAG ....RRTNGT IERITGLDPA
             .   .    . .*.. .  ..*..* ..        .  .  . *.*.***.

Dmyp1   KIVAKSKNTL TGLARGDAEF VDAIHT....  ..SVYGMGTP IRSGDVDFYP
Dmyp3   KVLSKRPGIL GGLSRGDADF VDAIHT....  ..STFAMGTP IRCGDVDFYP
Ceyp3   KILAKNPHSL TGLSRGDAEF VDAIHT....  ..NVYGMGTI GRVGDIDFYP
Dmyp2   KIYGKPEERL TGLARGDADF VDAIHT....  ..SAYGMGTS GRLANVDFFP
Ceyp2   KIYAQTGKSL RGLARGDAEF VDVIHT....  ..SAYGWGTP TRCGDVDFYP
 Rht1   GPMFEGTSPN ERLSPDDANF VDAIHT.FTR EHMGLSVGIK GPIAHYDFYP
Lihu1   GPNFEYAEAP SRLSPDDADF VDVLHT.FTR GSPGRSIGIQ KPVGHVDIYP
  Lipg   EPCFGGTPEL VRLDPSDAKF VDVIHTDAAP IIPNLGFGMS GTVGHLDFFP
            .   .        *  .**.* **..**        . *      .  *  *

Dmyp1   NGPAAGVPGA S.....236
Dmyp3   NGPSTGVPGS E.....339
Ceyp3   NGPSAAVPGA Q.....328
Dmyp2   NGPSTGVPGA D.....338
Ceyp2   NGPSRGVPGA D.....260
 Rht1   NGGSFQ.PGC H.....256
Lihu1   NGGTFQ.PGC N.....244
  Lipg   NGGKQM.PGC Q.....236
         **      **
```

vitronectin receptor, gpIIb/IIIa platelet glycoproteins, LFA-1 complement receptor type 3 and cell surface glycoprotein (p150,95) (Ruoslahti and Pierschbacher, 1986). It is important to note, however, that even though this motif is recognised by a wide range of cell receptor proteins, its interaction with each protein is specific. For instance, the fibronectin receptor will only recognise fibronectin and not vitronectin and viceversa. Thus the RGD motif is required for recognition and upon interaction this leads to a secondary site of recognition which gives the specificity of interaction (Ruoslahti and Pierschbacher, 1986). In this way the YPs may have used a widely available motif which is easily recognised by a number of receptors. The RGD motif is well conserved as well as the surrounding amino acid sequences within the Dipteran YPs (mostly or partly including the lipase similarity). However, this sequence is not conserved by the lipases (Figure 5.15), thus suggesting that it is not part of the lipid binding sequence motif and that it is particular to the YPs. The RGD sequence specificity to the YPs infers that it may be involved in interactions that may be important for other YP functions (eg, YP/receptor complex formation). The utilisation of the same region of the YP for two putative activities comfirms the idea that a non-conservative mutation within the *yp* coding sequence may have very wide repercussions to the whole protein structure and jeopardise its structure and function.

v. Region 344-415

The database searches found similarities to Ca^{++} binding related proteins such as the human collagen receptor alpha-2 chain precursor. This protein has seven putative calcium binding sites and the similarity to the $CeYP_3$ (amino acid residues 383 to 399) is at one of these, number VII (Figure 5.16A). This peptide has 10/17 amino acid identities with site VII sequence and somewhat fits the motif DXDXDGXXD and DXDXDGXXDXXE. Szebenyi *et al.* (1981) compared different calcium binding sites and found that aspartic acid (D) can be replaced by glutamic acid (E) or asparagine (N) however, the conserved glycine (G) was found to be predominant and only changes to serine infrequently. Tyrosine is a large bulky amino acid with a polar side chain which may disrupt the Ca^{++} binding site pocket. Significantly, it is known that in *Nauphoeta* (Gochoco *et al.*, 1988) and *Manduca* (Osir *et al.*, 1986) calcium is required for vitellogenins to interact with the receptors in oocyte membranes. Moreover, in *D. melanogaster* endocytosis of the YPs is known to be pH dependant (Di Mario and Mahowald, 1987) thus suggesting that ionic interactions have a role in ensuring YP uptake. The mode in which the divalent ions act on YP uptake may be by calcium ions interacting between the YPs and their receptors to form a YP/receptor complex or by the divalent ions interacting with the YPs to form YP oligomers. In any case, the complete divalent binding site may not be

Figure 5.16

A.Alignment of the yolk proteins (YPs) of *D. melanogaster* (Dm) and *C. erythrocephala* (Ce) with the collagen receptor type I (CR), *Mycobacterium leprae* 65 kD protein antigen (65K), the leukocyte common antigen (LCA) and boss protein of the fruit fly (boss).
Bold letters indicate conserved residues to the YPs.
*= Indicates conserved position in the alignment.
.= Indicates conservative substitution in that position.

B.Segments of the *D.melanogaster* and *C. erythrocephala* YPs within region II (residues 170 to 215) which show incomplete divalent ion binding sites (DIBS). The amino acid residues underlined indicate an acidic residue (D which may be replaced by E or N) that fits to the consensus sequence DXDXDGXXD.

C.Alignment of the putative divalent ion binding sites in region 344-405 of the YPs and the consensus sequence of the divalent ion binding site (DIBS).

A. CeYP$_2$ 315 DVEGDFILQVNGKSPFG 331

CeYP$_3$ 383 DLEGDYILEVNSKGPFG 399

DmYP$_1$ 395 DLEGDYILQVNPKSPKG 411

DmYP$_2$ 396 DLQGDYILQVNSKSPFG 412

DmYP$_3$ 385 DLRGDYILEVNAKSPFG 401

CR 627 DLNGDSITDV-SDGAFG 642

65K 213 LEDPYILIVSSK 224

boss 139 LEGDVVLSVLTKDP 152

LCA 260 LEEPYILLVSSK 271

. . . *. .

B. DBIS DXDXDGXXD

CeYP$_2$ 96 EDDSSEKRNPSSSEEDDEDWKSN 116

CeYP$_3$ 164 DYSGEKMARTSSEEDSNEWQN 184

DmYP$_1$ 166 NGNQDYQDQSNEQRKNQRTSSEEDYSEEVKN 196

DmYP$_2$ 171 DYSNEEQSQRSSSEEQQTQRRKQNGEQDDTKTGD 204

DmYP$_3$ 172 DYDYTSSEEAAD 183

C. DIBS DXDXDGXXD

CeYP$_2$ 313 DYDLEGDYILE 323

CeYP$_3$ 381 DYDVEGDFILQ 391

DmYP$_1$ 393 AHDLEGDYILQ 403

DmYP$_2$ 394 DFDLQGDYILQ 404

DmYP$_3$ 383 DYDLRGDYILQ 393

DIBS DXDXDGXXD

required as long as both YPs and receptors or YPs have enough of the binding site to allow interaction via divalent ions to take place. The similarity to calmodulin and calcium channel alpha chain protein in region 170-215 (ie. between amino acid residues 170 and 181) may also have a portion of a divalent ion binding site (Figure 5.16B) which in turn would provide two sites in the yolk protein capable of interacting with each other or with the receptor. Thus there may be two different incomplete divalent ion binding sites in the YPs (Figure 5.16B and C) which may be important in YP oligomer formation or in formation of receptor/YP complexes. The instability of the *D. melanogaster* YP oligomers may be a result of weak divalent binding ion site which may be easily disturbed by pH or Ca^{++} ion concentration fluctuations.

It is interesting to find a similarity to the Major Histocompatibility Complex (MHC) type I alleles in the $CeYP_3$. The sequence similarity encompases only 10 amino acid residues (amino acid residues 339 to 348) (Figure 5.17). The MHC type I allele serves as a recognition element for cytotoxic T-lymphocytes and is expressed as heterodimers of alpha (heavy) and beta (light) chains. The heavy chains has two polymorphic extracellular domains ($\alpha 1$ and $\alpha 2$), a conserved extracellular domain ($\alpha 3$), a transmembrane domain and a short cytoplasmic tail. The light chain is non-covalently bound to $\alpha 3$ whereas the $\alpha 1$ and $\alpha 2$ domains interact with the T-cell antigen receptor

Figure 5.17

Alignment of the yolk proteins (YPs) from *D. melanogaster* (Dm) and *C. erythrocephala* (Ce) with the Mayor histocompatibility Complex type I (MHC).

*= Indicates a conserved position of the alignment.

.= Indicated a conservative substitution at that position.

$CeYP_2$	271	RYFAESV-RPG	280
$CeYP_3$	339	RYFAESV-RPG	348
$DmYP_1$	351	RYFAESV-RPG	360
$DmYP_2$	352	RYFAESV-RPG	361
$DmYP_3$	341	RYFAESV-RPG	350
MHC	30	RYFSTSVSRPG	39
		*** . ** ***	

(Mayer *et al.*, 1988). The CeYP$_3$ sequence similarity to chimpanzee and human alleles of MHC type I lies at the N-terminal end of the -1 domain. In region 19-97 a match was found to T-cell antigen receptor molecule, however both the CeYP$_3$ similarities to T-cell antigen receptor and MHC type I alleles may not be of sufficient length to enable an interaction between them. This is not to say that these may not be involved in molecular interactions within the YPs or in YP/receptor complexes.

The last 24 amino acids of DmYP$_2$ (or last 17 residues relative to CeYP$_3$) have been shown to not be required for secretion, uptake, sorting and storing of DmYP$_2$. This was shown by transformation of mutant strains YP$_2^{12-1245}$ with a construct bearing the first 420 amino acid residues of Dm*yp*$_2$ fused to the yeast invertase coding sequences (Yan and Postlethwait, 1990). The YP$_2$-invertase fusion protein synthesised behaved in a manner expected of a YP. However, the embryos containing this fusion protein did not develop, suggesting that the last 24 amino acid of DmYP$_2$ are required for normal YP$_2$ utilisation during development or that the fusion protein is altered structurally resulting in a product stable to degradation during embryogenesis.

E. Concluding remarks

The likelihood of being able to identify domains required by the YPs to enter the oocyte via receptor mediated endocytosis is relatively low when the proteins have been so well conserved over Dipteran evolution. Moreover, the description of these domains and potential functions in the YPs is not aided by the lack of knowledge of their biochemistry. However, the little biochemical information that is available may be of some help to assess the significance attached to the sequence similarities described above.

In conclusion the database searches indicated five possible peptide sequences which may have important effects on YP protein structure (Figure 1.18). The peptide with similarity to MHC type I and T-cell antigen receptor may have the ability to interact together to enable the formation of YP oligomers or may be independantly involved in interacting with the YP receptor. Secondly, there are two peptides with partial similarities to divalent cation binding sites. These two sites are in the same regions as the putative sulfation sites of the YPs are located (ie. from amino acid residues 167 to 180 and 380 to 395 relative to $CeYP_3$) (Figure 5.17B and C). The post-translational modification of tyrosine sulfation takes place in the trans-golgi prior to secretion of the protein. For the modification to take

Figure 5.18

Summary figure of the matches found when the different YP regions were used in database searches. The YP used as a model was the YP_3 of *C. erythrocephala*. I - V indicate the five different regions used in the database searches and described in the text.

I= region I residues 19 to 97.

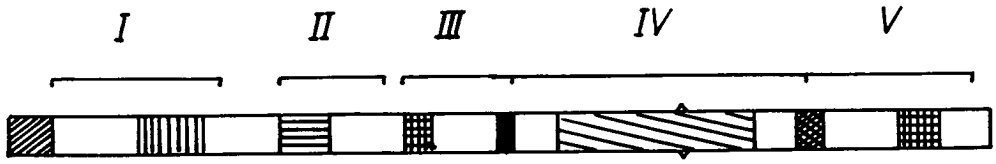
II= region II residues 117 to 160.

III= region III residues 170 to 216.

IV= region IV residues 216 to 343.

V= region V residues 344 to 405.

Ce YP3



- *Signal peptide (1-19)*
- *T-cell antigen receptor similarity (56-84)*
- *Receptor protein similarity (117-138)*
- *Putative Calcium binding site (170-181 and 383-399)*
- *Protease inhibitor similarity (210-216)*
- *Vertebrate lipase similarity (198 - 348)*
- *RGD motif (288-290)*
- *MHC type1 (339-348)*

place the tyrosine residues must be exposed on the surface of the folded protein. Since the tyrosine sulfation sites are within the same segments of the YPs as the putative divalent binding sites, this indicates that the putative divalent binding sites are also exposed. Therefore this would allow the divalent ion binding sites to be functional and capable of interacting with the YP receptor to form YP/receptor complexes or other YPs to form oligomers.

The RGD sequence in the YPs is found within the area of similarity to the vertebrate lipases (Figure 5.15). Moreover, the sequences surrounding the tripeptide are well conserved and may be a primary site for interaction with the YP receptor. Once the YP is recognised the secondary stage of specific recognition may take place. Thus this tripeptide may offer a fast yet specific interaction with the YP receptor resulting in specific uptake. There are other sequences that may have shown similarities to the YPs in these database searches which may infer that these proteins share structural similarities or that the sequence confers the same type of structure in both proteins. However, the importance of the sequence to uptake may be in their actual position within the protein and the context in which they are found. The peptide sequences discussed in this chapter may only be used as a guide to targetting portions of the YPs to determine the sites of interaction between YPs and

their receptors. In addition, the sequence suggests interactions with calcium ions and thus further biochemical studies to dissect the importance of divalent ions in YP interactions must be pursued.

In summary the YPs are functionally well conserved through Dipteran evolution. There appear to be many constraints that affect the structure of the YPs ensuring that YPs do not diverge through Dipteran evolution. These constraints may be summarised as follows. At the amino acid sequence level the similarity of the *D. melanogaster* and *C. erythrocephala* YPs ranges between 69% and 75%. The proteins have the same amino acid composition despite the divergence in DNA sequence. The similarity to vertebrate lipases is maintained. The proteins are secreted and specifically taken up into the oocyte where they are sorted and targeted to the α-granules. The YPs are known to form oligomers and the YP's structure to be affected by pH as demonstrated by pH dependance of uptake (DiMario and Mahowald, 1986). The protein may share domains such that the segment of protein involved in ecdysone binding may also interact with the receptor via the RGD motif. Therefore, the YPs appear to have structural constraints that may ensure conservation of the structure during evolution and suggests that mutations within the coding sequences of the *yp* gene may result in aberrant products which may not be secreted. At the same time the fusion of invertase to the C-terminal end of the $DmYP_2$ does not

affect secretion, uptake, sorting or storage (Yan and Postlethwait, 1990). However, the protein structure is affected in some way as the eggs containing the fusion product collapse. This indicates that not only do the YPs have to be targeted correctly but also that within their structure information is found concerning their utilisation. On the other hand the invertase portion of the fusion protein may have a stabilising effect on the YPs which prevents their usage during embryogenesis. Morphological and kinetic studies on the sorting, storage and usage of this YP/invertase fusion protein should indicate whether or not the fusion protein is targeted, stored and used at the same rate as the native YPs.

The studies carried out here in conjunction with the system established by Yan and Postlethwait (1990), may be used to investigate sequences which may be important for uptake of the YPs. A different approach was taken in this laboratory whereby deletions of segments within the DmYP$_3$ protein were carried out. These deletions concentrated in the 2/5 from the C-terminal end of the protein. The aims were to introduce these deleted genes into the genome of *D. melanogaster* and investigate their secretion, uptake, sorting and storage. Peptide domains found to be involved in uptake were to be analysed in detail using site directed mutagenesis. However, all of the deletions made were found to be unstable in the *D. melanogaster* genome and were not secreted (Lineruth and Bownes, unpublished).

A different approach can be carried out by using the *Lac-Z* gene as a reporter gene controlled by the *yp* upstream sequences and including test coding regions of the *yp* genes between these. The constructs would be introduced into the *D. melanogaster* genome and tested for their ability to direct the LacZ fusion product into the oocytes.

Finally, in order to test and confirm the results obtained in chapter three, where conservation of uptake, sorting and storage of foreign YPs into ovaries of *D. melanogaster* and *D. funebris* was observed, the *yp* coding region of the Ce*yp*$_3$ under the control of *D. melanogaster* promotor sequences can be introduced into the *D. melanogaster* genome via P-element mediated transformation. In this way the *Calliphora* protein synthesis, and its product secretion, endocytosis and sorting can be investigated in comparison to the *D. melanogaster yp* genes and their products. It is expected that since the YP proteins are so closely related the introduced *Calliphora* protein would behave in the same manner as the native YPs. One of the disadvantages of this methodology is that the test sequences may not be in the correct conformation to be presented to the receptors. Moreover, if uptake and storage of the fusion product is not observed, this result will not indicate that the tested peptide can not confer the information required for interaction with the YP receptor. It may be necessary to

make synthetic peptides corresponding to the chosen fragments of the YPs, and test *in vitro* their specific binding to *D. melanogaster* oocyte membrane preparations. Thus, the fusion protein experiments can be carried out with peptide sequences known to have a high specific activity to oocyte membrane preparations.

To conclude, this chapter shows that the YPs of *Calliphora erythrocephala* and *Drosophila melanogaster* are highly conserved, fully explaining the observed uptake of the CeYPs by ovaries of *Drosophila* (see chapter 3). The amino acid conservation also shows that the YPs appear to have many constraints to prevent divergence of the protein sequence through Dipteran evolution.

CHAPTER 6

RESULTS

Expression of the YPs in the Fat Body and Ovaries of *C. erythrocephala*

Introduction

As described in the previous chapter, Rubacha *et al.*, (1989) demonstrated that the yolk protein genes were expressed in *Calliphora* vitellogenic oocytes. Previously, Huybrecht and De Loof(1982), detected YPs in the haemolymph of *Calliphora erythrocephala* females and showed that synthesis of these proteins are induced in males by the injection of 20-hydroxyecdysone suggesting these are also expressed in the fat body. To confirm that the genes are expressed in the fat body of female flies northern blots containing male and female carcass RNA were carried out as described in chapter 5 (Figure 5.2). The data shows that the *yp* genes are expressed in the carcass and ovaries of female flies; however, their pattern of expression has not been throroughly investigated. Furthermore, vitellogenesis in *C. erythrocephala* females requires protein feeding (Fraenkel, 1940; Strangeways-Dixon, 1971). The food response is reminiscent to the requirement of a blood meal by mosquitos *Aedes aegypti* (Raikhel and Lea, 1985), and of protein feeding by *Musca domestica* (Morrison and Davis, 1964), *Protophormia terrae-novae* (Harlow, 1956) and *Phormia regina* (Pappas and Fraenkel, 1977) to trigger vitellogenesis.

The aims of this chapter are to determine the time of expression of the *yp* genes in the fat body and the ovaries of female flies. Furthermore, to determine if the expression in both types of tissues is synchronised, and investigate which cells are involved in the expression of the *yp* genes in the ovaries. *Calliphora* flies have polytrophic meroistic ovaries in which all oocytes develop at the same time in a particular cycle. This is in contrast to *Drosophila* where oocytes are constantly being produced resulting in ovaries which contain several stages of oogenesis. *Calliphora*, therefore offers a unique system to determine the timing of *yp* gene expression in ovaries and fat body. In order to realise the goals of this chapter it is necessary to obtain synchronous populations of flies.

Synchronising a population of *Calliphora*

The *Calliphora* flies were cultured in ICAPB (Institute for Cellular and Population Biology), formerly the Department of Zoology (University of Edinburgh). A one hour egg collection was carried out and the resulting embryos reared in the usual manner (2.3). The narrow band of egg laying was used to ensure that most resulting pupae would eclose within a short period of time. The pupae were allowed to eclose in a cage kept at 25°C with water and sugar in a 12 hour day/night cycle. The flies were

kept in these conditions for two days to ensure that all flies had eclosed before raw meat was placed in the cage (Time 0). The raw meat was changed every day, while water and sugar were replenished as needed. For the first 36 hours flies were collected every 12 hours after which collections were every 8 hours until 44 hours and then every 4 hours until the end of the experiment. Vitellogenic ovaries were not seen before 52 hours. The population of flies did not seem very synchronous since an assortment of ovaries at different stages of oogenesis was seen by 68 hours; however, as time progressed the majority of the ovaries were always found to be at an equivalent stage. Previtellogenic ovaries were again observed by 122 hours (or 5 days later) presumably from flies which had already layed their first batch of eggs. At every stage of the experiment the collected female flies were dissected and their ovaries carefully examined to determine the stage in oogenesis. The carcasses of the dissected flies were classified according to the stage of ovarian development because of the asynchrony of the fly populations. The collection of ovaries was carried out between previtellogenic (up to stage 7), post-vitellogenic stages (stage 12 and beyond) and then every ovary falling between stages 7 and 12 was classified and stored depending on whether it was being used for RNA extraction (2.4.2.2.2), *in situ* hybridisation to whole tissue mounts (2.4.6) or *in vitro* ovary culture (2.4.1.11.3).

Figure 6.1 shows examples of ovaries at different stages of oogenesis, where all oocytes are at the same stage of oogenesis. In Figure 6.2, the stages of oogenesis are shown in greater detail. These oocytes are representative of each stage of vitellogenesis and show the features used to identify the stages of the ovaries used in the experiments described below. The naming of the stages is analogous to the *Drosophila* staging system devised by King (1956), which has been used in other Diptera (ie. *Musca*) (Goodman *et al.*, 1968).

During which stages of oogenesis are the *yp* genes transcribed?

Ovaries and carcasses from female flies at various stages of oogenesis (previtellogenic, 8, 9, 10a, 10b, 11 and 12 or more) were collected into RNA buffer and the RNA extracted (2.4.2.2.2). Each track of the northern blots contains 10ug of total RNA. The filters were probed using the 0.7 kb BamHI DNA fragment (Ceyp_1) or the 1.6 kb PstI/EcoRI DNA fragment (Ceyp_2). These fragments were chosen since both cover the C-terminal end of YP_1 and YP_2 proteins, which is the most similar domain of the genes, resulting in a probe sensitive to all the *yp* genes. In both the ovarian and carcass northern filters the yp transcripts (Figure 6.3A and B) are present in flies of stages 8, 9, 10a and 10b and disappear by stage 11. In the previtellogenic and postvitellogenic stages (stage 12,

Figure 6.1

Photographs of *C. erythrocephala* ovaries in which all the oocytes are at the same stage of oogenesis.

A. shows ovaries at stage 9 of oogenesis.

B. shows ovaries at stage 10b of oogenesis.

C. shows ovaries at stage 11 of oogenesis. By this stage vitellogenesis has been completed and choriogenesis is under way.

D. shows ovaries at stage 12 of oogenesis.

magnification: X12

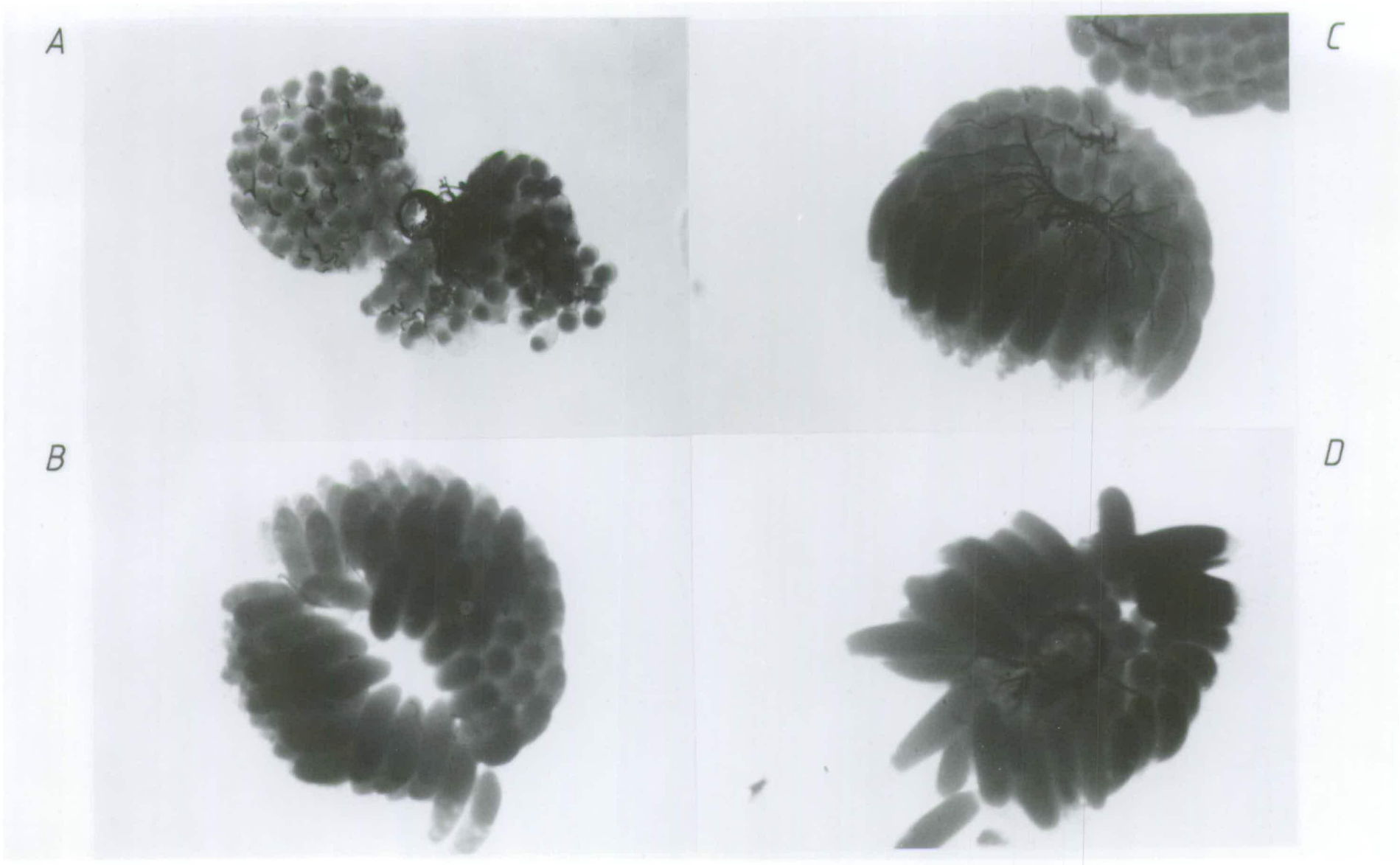

A
C
B
D

Figure 6.2

Photographs of the different stages of oogenesis. These are equivalent to those used to classify the ovaries and carcasses of female *C. erythrocephala* flies.

G= germarium

nc= nurse cells

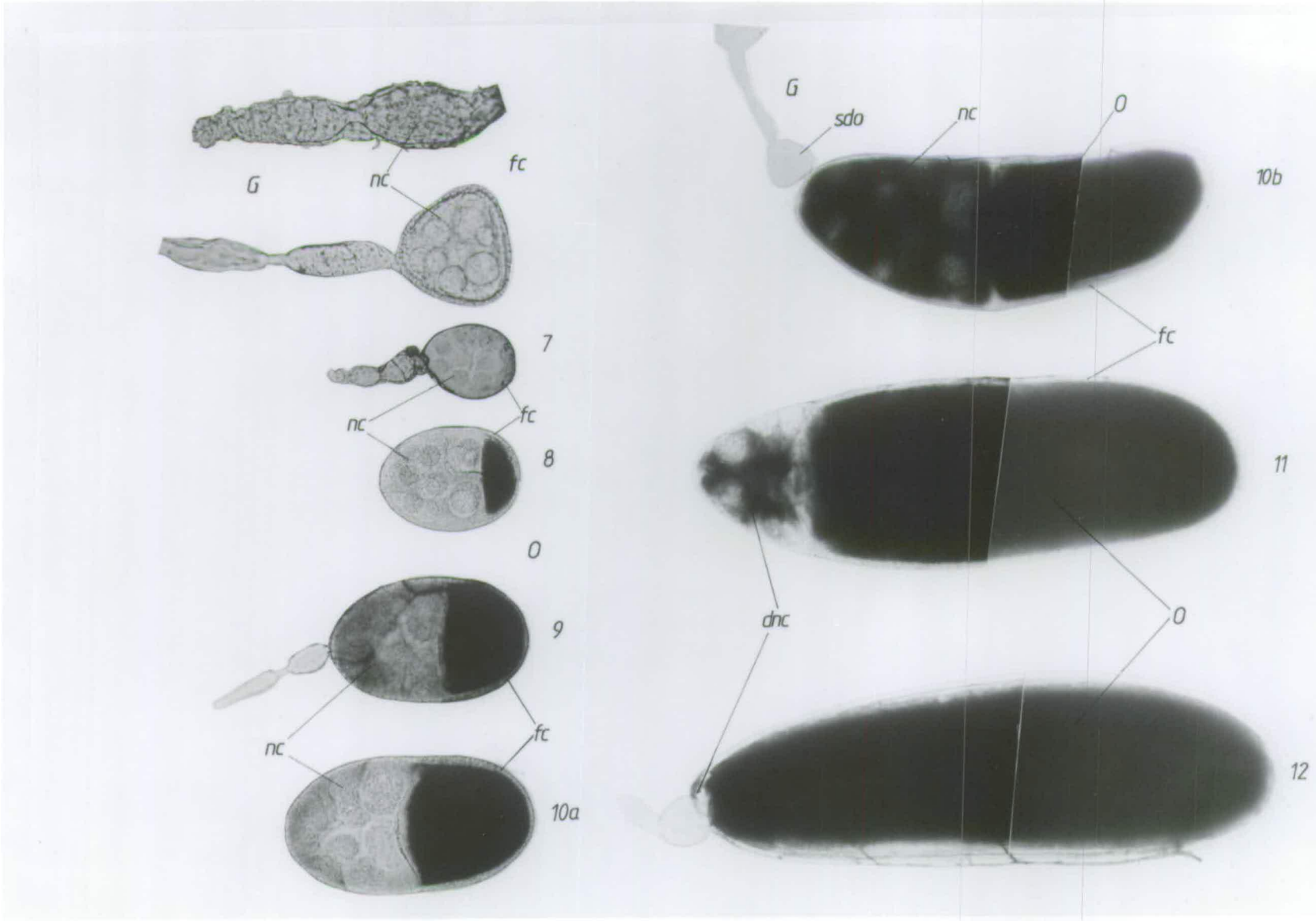
O= oocyte

fc= follicle cells

sda= stationary developing oocyte.

7-14= denote the stage of oogenesis.

dnc= degenerating nurse cells


G
nc
fc
7
nc
fc
8
O
9
nc
fc
10a
G
sdo
nc
O
10b
fc
11
dnc
O
12

13 and 14) and in males (fed and unfed) yp transcripts are undetected. A ribosomal probe was used after the northern filters were stripped to determine the RNA loading in each track. The results (Figure 6.3A and B) indicated ~~ loading of RNA in all tracks except for fat body at stage 12. The northerns blots were reprobed with *α-tubulin* and *Adh* gene probes from *D. melanogaster* to determine if the same loading pattern was observed. The results obtained using these probes were unclear as no specific binding was observed, suggesting that these genes have diverged too far to be sensitive enough to detect any homologous transcripts in *Calliphora*. However, the data obtained clearly indicates that the yp transcripts in *Calliphora* are present at stages 8 through to 11 of oogenesis.

Yolk protein synthesis in the ovaries.

The synthesis of yolk proteins in the ovary can be assayed by culturing ovaries *in vitro* in the presence of ^{35}S-methionine (2.4.1.11.3) to assess *de novo* synthesis of YPs. In the experiment described below all ovarian stages had 2 pairs of ovaries, except for the previtellogenic (ie, ovaries up to stage 7) and stage 12 (ie, ovaries from stage 12 to 14) where three pairs of ovaries were used. The ovaries were cultured in *Drosophila* Ringer's solution for 2 hours at 25°C. The culture medium was collected and subjected to SDS PAGE and the dried gel autoradiographed. The *de novo* synthesis of YPs by the ovaries mirrors the

Figure 6.3

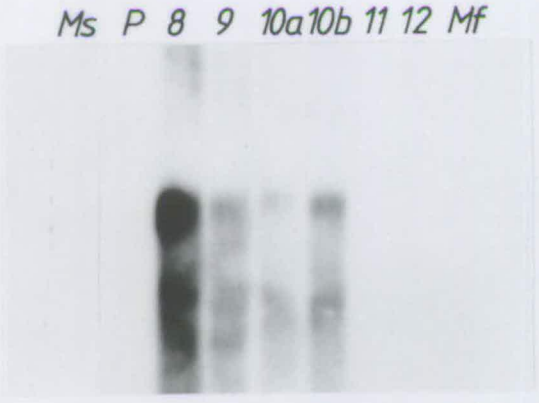
Northern blot analysis of temporal onset of *yp* gene transcription in the carcasses and ovaries of *C. erythrocephala*.

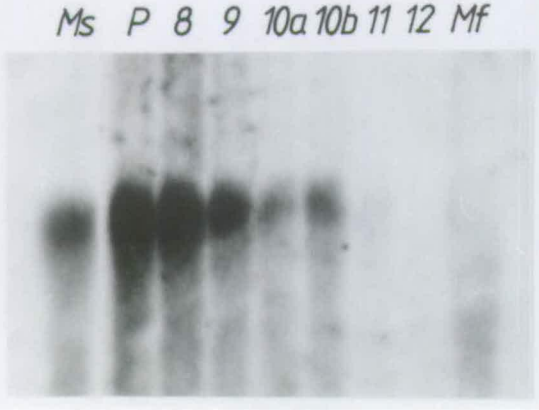
A.Northern blot containing total RNA from female carcasses at different stages of oogenesis. The top panel contains a autoradiograph of the northern blot hybridised with the Ce*yp*$_2$ or Ce*yp*$_3$ genes. The lower panel is an autoradiograph of the stripped northern filter and reprobed with *D. melanogaster* ribosomal gene coding sequences (pDm264)

B.Northern blot containing total ovary RNA at different stages of oogenesis. The top panel show the autoradiograph of the northern blot hybridised with the Ce*yp*$_2$ or Ce*yp*$_3$ genes. The bottom panel is the same filter hybridised with *D. melanogaster* ribosomal DNA.

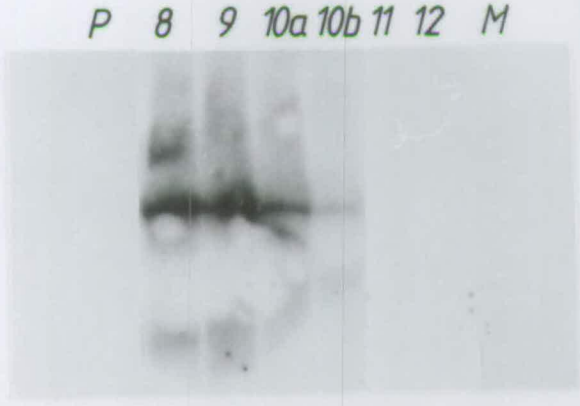
P= pre-vitellogenic carcasses or ovaries (stages 1 to 7).
8-11= Indicate stage of oogenesis of ovaries or carcasses.
12= Indicates stages 12 and beyond of oogenesis.
M= Indicates track loaded with male total RNA.

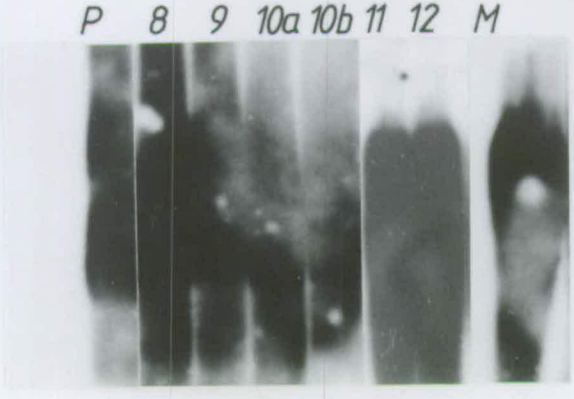
A. Carcass





B. Ovary





results obtained when the yp transcripts were examined by northern blotting (Figure 6.4). Synthesis of YPs by stage 8 oocytes is low and increases with time to stage 10a. At stage 10b, the oocytes are producing less yolk protein and by stage 11 the expression of YPs is undetected. On the other hand, by stage 11 the synthesis of other proteins is observed, and this increases in stage 12. The proteins synthesised in stage 11 and 12 probably correspond to the chorion proteins since these are the stages at which chorion synthesis occurs in *Drosophila* (Orr-Weaver *et al.*, 1991).

Ovarian follicle cells, a site for YP synthesis.

The yolk proteins are synthesised in the ovaries from stage 8 to stage 11. One way of determining the cell type involved in the synthesis of YPs in ovaries is by whole mount *in situ* hybridisations (Tautz and Pfeifle, 1989). In this case, a DNA probe is labelled with digoxigenin, and after the hybridisation the probe is detected using antibodies raised against digoxigenin. The probes used for these experiments are the 0.7 kb BamHI (Ce*yp*$_3$) and 1.6 kb PstI/EcoRI (Ce*yp*$_2$) DNA fragments since these are likely to be the most sensitive probes detecting all yp transcripts. The ovaries were collected, classified according to their stage in oogenesis and and subjected to the process of *in situ* hybridisation (2.4.6). Photographs were taken immediately the colour had developed and before

Figure 6.4

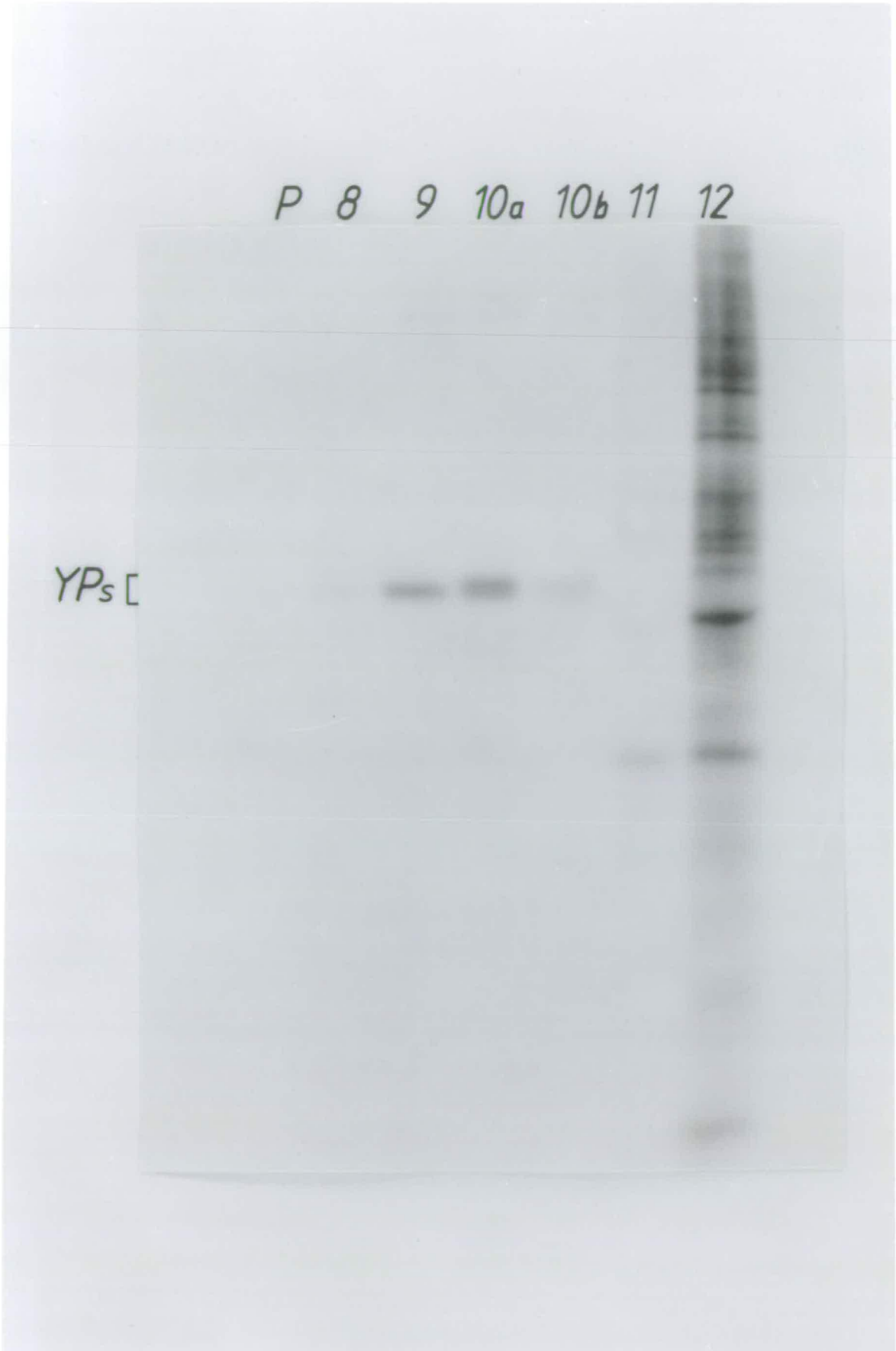
Autoradiograph of a SDS-PAGE containing the proteins in the culture medium of ovary *in vitro* cultures (2.4.1.11.3). The ovaries were dissected out and classified according to King *et al* (1956) and Figure 6.2.

P= pre-vitellogenic ovaries (Stages 1-7).

8-11= Indicate stage of oogenesis.

12= Indicates stage 12 or beyond.

YPs, indicates position of the yolk proteins in the autodiograph.


P
8
9
10a
10b
11
12
YPs [

background appeared in the oocyte. One problem with the method as used and described here is that diffusion of the colour reagent takes place. Leaving the ovaries in stop buffer or mounted in Gurr's mounting medium overnight resulted in a darkly coloured oocyte and an increase in background due to the diffusion of the colour reagent. However a dark band of colour still remains in the follicle cells. In stage 8 the colour is faint (Figure 6.5A) and difficult to detect but later stages (9, 10a and 10b) clearly show the colour stain in the follicle cells. In stage 9 ovaries colour stain is observed in the follicle cells that are in contact with the nurse cells. At stage 10a all follicle cells are in contact with the oocyte and are transcribing the *yp* genes. At stage 10b the follicle cells which have migrated between the oocyte and the nurse cells are also expressing the *yp* genes. Therefore, the data shows that the *Calliphora yp* genes in the ovaries are expressed in the same pattern as the genes of *D. melanogaster* (Barnett *et al.*, 1980; Isaac and Bownes, 1982) and *Sarcophaga bullata* (Geysen *et al.*, 1986)

Do the *Drosophila* and *Calliphora yp* genes share any regulatory sequences?

The *yp* genes of *Calliphora* are expressed in the ovaries and carcasses of female flies. Moreover, the expression in the ovaries is limited to the ovarian follicle cells as has been shown for *Drosophila* (Brennan

Figure 6.5

Photographs of ovary whole mount *in situ* hybridisations. The ovaries were treated as described in methods and materials (2.4.7) and hybridised with the Ce*yp*$_2$ and Ce*yp*$_3$ genes.

A.Stage 8 of oogenesis.

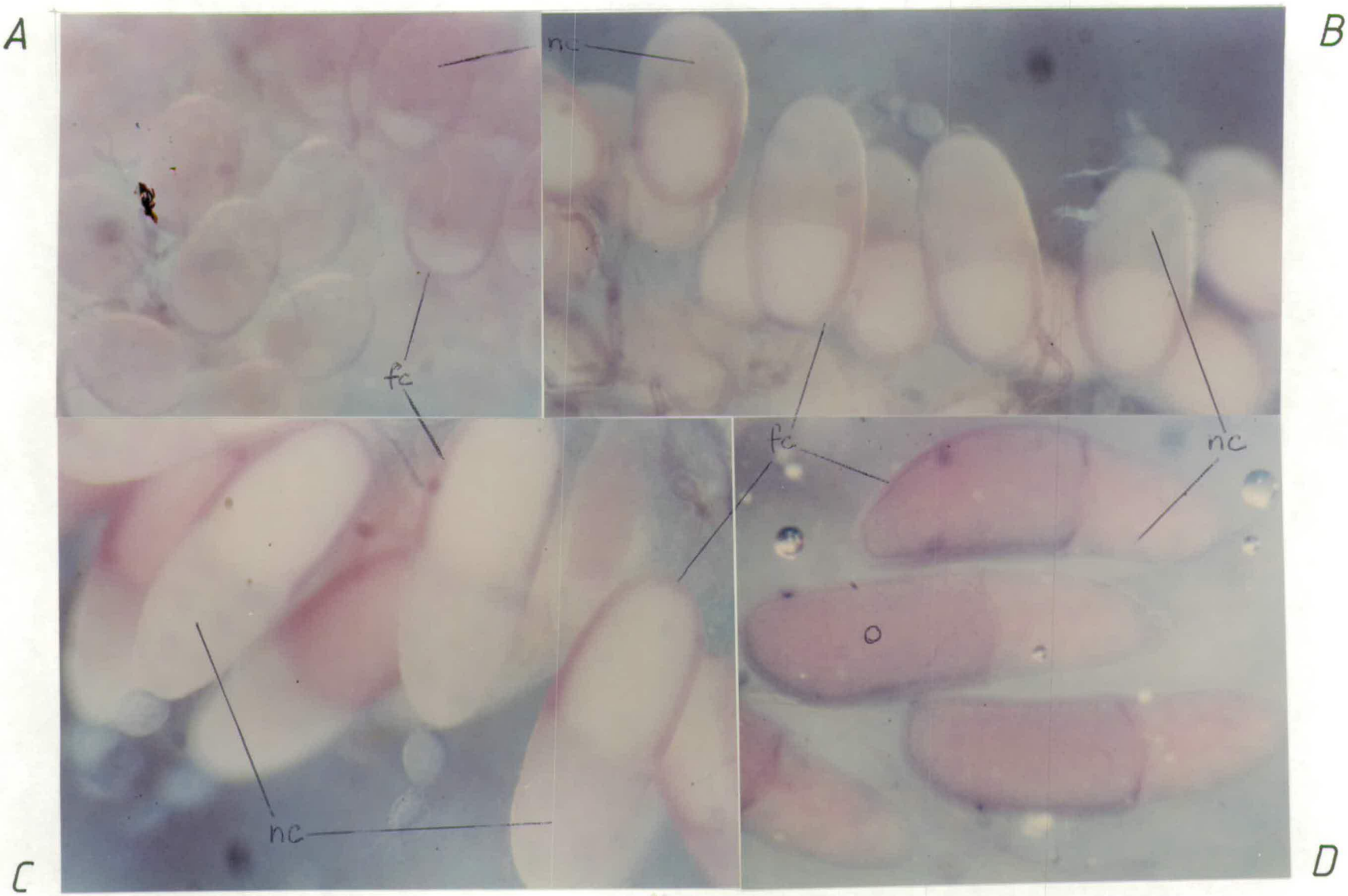
B.Stage 9 of oogenesis.

C.Stage 10a of oogenesis.

D.Stage 10b of oogenesis.

o : oocyte
nc : nurse cell
fc : follicle cell

magnification: x50


A
fc
B
nc
C
nc
D
fc
nc
o

et al., 1982; Issac and Bownes, 1982; Logan *et al*, 1989) and *Sarcophaga* (Geysen *et al.*, 1986). This may indicate a similarity in the regulatory sequences that govern the expression of the *yp* genes in *Drosophila* and *Calliphora*. A point of caution is that the two genomes have diverged enough so that sequence similarities at the DNA level are difficult to find (this thesis, Chapter 5) and secondly, the existance of sequence similarities does not indicate conservation of the element until these have been tested *in vivo*. However should conserved sequences exist these would be a starting point for further functional tests *in vivo*. Comparisons between the sequences known to confer fat body specificity (FBE, Garabedian *et al.*, 1986), ovary specific expression (OE1, OE2 Logan *et al.*, 1989) and those which bind the *double sex* protein (DSX) (Burtis *et al.*, 1991) were carried out in the *yp*$_3$ gene sequence of *Calliphora* (Ce*yp*$_3$) using the FastA program within the GCG package (version 6.0) (Devereux *et al.*, 1984). The similarities obtained did not cover the whole of the *Drosophila* template sequences. The fat body enhancer (FBE) sequence in *Drosophila* is between -321 and -194 from the transcription site of Dm*yp*$_1$ and is 127 bp long (Garabedian *et al.*, 1986). In *Calliphora* the longest similarity was found within the first exon of the Ce*yp*$_3$ gene, between -27 bp and +95 bp (base pair 1, indicates the first nucleotide of the methionine start codon site). This similarity is 122 bp long and 52.4 % similar to the FBE sequence (Figure 6.6A) The other two similarities

Figure 6.6

Sequence similarities at the DNA level between the Ce*yp*$_3$ gene and the fat body enhancer (FBE) region described for the *D. melanogaster* *yp*$_1$ and *yp*$_2$ genes. The searches were carried out using the program wordsearch and Fasta of the GCG package (version 6.0) (Devereux *et al.*, 1984).

A. The sequence similarity is 122bp long with 52.4% identity.

B. The sequence similarity is 69bp long with 60.8% identity.

C. The sequence similarity is only 33bp in length with 72.7% identity.

A.

```
Ceyp3  -27 GCGGTACTGACTGATAGCTTGAAAGCATGAATCCT.TTGAGAA   18
           || |  |  |||||  |||   ||||     |||  ||
FBE     18 GCTGACCCCACTGATGACTTAGCAGCA.....CCTGCTGCCGC  56

Ceyp3   19 TTGTTTGTGTGGCTGCTTTGTTATTGGC....TGCAGGTTCAG  58
           || || | |   |  |||||  ||   |  | |  ||| || |
FBE     57 CTGATTATTT...TAATTTGTAAT...CACTTTCTAGGCTCCG  92

Ceyp3   59 CTAATGGAAATCTTTCTGGTCTTAACAAGTTGAG.ACC   95
           || || | ||  ||    ||      |||||  | |||
FBE     93 CTGATTGCAACATT....GT.......AGTTGTGCAGG 119
```

B.

```
Ceyp3-190 TGACTATTATTATTTTAAAGAAAT...TTTCTAGCATC..AT -154
          ||| |||| | ||||    | |||   ||| |||  ||   |
FBE    57 TGATTATTTTAATTT....GTAATCACTTTGTAGGCTCCGCT  102

Ceyp3-153 GA..AGAAAATTTTATTAATTTCAGCAAAATATT -122
          ||    || ||| |    | ||  |||    |||
FBE   103 GATTGCAAGATTGT....AGTTGTGCACCGGATT  124
```

C.

```
Ceyp3-303 GACCAGAAAGATTTCCATTAGCTGAACCTGCAGCC -270
          ||||  |  |||  |  ||||| | |||||| |||
FBE    21 GACCCCACTGATGAC..TTAGCAGCACCTGCTGCC   53
```

found were shorter and within the 5' non-coding sequence of the Ce*yp*$_3$ (Figure 6.6 B and C). The ovary enhancer 1 (OE1) fragment is 301 bp long and is found between -343 bp and -42 bp from the Dm*yp*$_2$ transcription site. The best equivalent match in the *Calliphora* Ce*yp*$_3$ gene was 56 bp long and 62.7 % similar. It is located between -143 to -205 bp from the translation start site and is therefore in a similar position to OE1 (Figure 6.7A). Finally the equivalent site to ovary enhancer 2 (OE2) in *Calliphora* compares well with the one of OE2 in the first exon of the Dm*yp*$_2$. In *Calliphora* it is found between -11 bp to +62 bp and is 75 bp long with 60.7% similarity to OE2 (Figure 6.7B). This sequence however is located within the coding region of Dm*yp*$_2$ and directs Dm*yp*$_1$ ovary specific expression and thus may not be required in Ce*yp*$_3$ expression. Moreover, since the sequence similarity is in the first exon, it is possible that the homology is due to protein sequence conservation rather than to similarities in putative regulatory or binding sites.

Searches with the DNA binding motifs of the DSX protein (Burtis *etal.*, 1991) were done using Wordsearch and FastA programs within GCG (Devereux *et al.*, 1984). Only similarities that scored 7/9 nucleotide identities or better were considered since this is the observed divergence between the concensus sequence and the four identified DSX binding sites in the FBE (Burtis *et al.*, 1991). The results are summarised in Figure 6.8. The DSX

Figure 6.7

Sequence similarities between the Ce*yp*$_3$ gene sequence and the ovary enhancer regions described in the intragenic region between the *D. melanogaster* *yp*$_1$ and *yp*$_2$ genes.

A.Shows the similarity between OE1 and the Ce*yp*$_3$ gene. The similarity is 56bp long with 62.7% identity.

B.Shows the similarities between OE2 and the Ce*yp*$_3$ gene. This similarity is 75bp long with 70.6% identity.

A

```
Ceyp3-211 AATCTTAATAAAAAAAATGATGACTATTATTATTTTAAAGAAA -168
          |  |||||||||  ||||  |  |  |  | |  |    || ||||
OE1   130 ATACTTAATAACCAAAAAAAAAAAAAAAAAAAAGGGGAACGAAA  173

Ceyp3-167 TTTTCTAGCATCATGAAGAAAATTTTA -140
          ||||  ||| |||| |  |   ||  |
OE1   174 TTTTAGAGC-TCATCAGAAGTGTTGCA  199
```

B

```
Ceyp3  -12 AGCTTGAAAGC----ATGAATCCTTTGAGAATTGTTTGTGTGG  29
           | |||| ||||    |||||||||| || | |   |||| ||
OE2     40 AACTTGGAAGCCACAATGAATCCTCTGCGCACCCTTTGCGTTA 82

Ceyp3   30 CTGCTTTGTTATTGGCAGCAGGTTCAGCTAAT  61
             || |   |  |||| |  |     | ||||
OE2     83 TGGCCTGCCTTCTGGCGGTCGCCATGGGTAAT 114
```

Figure 6.8

Sequence similarities between the *LSP-2/yp* FBE region and the 5'sequence to the Ce*yp*$_3$ gene.

```
Ceyp3       -277 AAAATCAATGAAATCAAAA -259
                 ||||||||  ||||
LSP-2/yp       1 TGAATCAATGTAATCGTGC  19


Ceyp3         53 TCAGCTAATGGAAATCTTTC  70
                 |||   ||| | |||| | |
LSP-2/yp       1 TCAATCAAT.GTAATCGTGC  19


Ceyp3       -206 TTAATAAAAAAAATGATG -188
                 | ||| ||   ||| ||
LSP-2/yp       1 TCAATCAATGTAATCGTG  17


Ceyp3       -327 TCTATCAAT.TAAT -315
                 || |||||| ||||
LSP-2/yp       1 TCAATCAATGTAAT  14


Ceyp3         -8 TGAAAGCATG.AATCCT    9
                 | ||   ||| |||| |
LSP-2/yp       1 TCAATCAATGTAATCGT  17
```

binding site appears in upstream and coding sequences of the Ce*yp*$_3$ gene. The distribution of DSX binding sites has also been observed in genes not affected by the sex determination hierarchy (Abrahamsen *et al.*, 1991). This observation suggests that this site must be found in conjunction with other determinants or in a particular array to ensure the correct pattern of expression. Such an arrangement has been found within FBE where four DSX binding sites have been determined by DNA footprinting (Burtis *et al.*, 1991). In *Calliphora* only the shorter regions with FBE similarity (Figure 6.6B and C) appear to have putative DSX binding sites and thus may be better choices for putative FBE-like elements. On the other hand, clusters of putative DSX binding sites are observed in the rest of the upstream region of the Ce*yp*$_3$ gene.

Shirras and Bownes (1989) found a sequence between *yp*$_1$/*yp*$_2$, *yp*$_3$ and *lsp-2* upstream sequences with a high level of sequence similarity and suggested that it may be a tissue specific enhancer. This sequence is found within the FBE in the intergenic region between the Dm*yp*$_1$/*yp*$_2$. In the *Calliphora yp*$_3$ five sequence similarities were found to Dm*yp*$_{1/yp2}$ $_{lsp-2}$ sequence (Figure 6.8). The putative sequences obtained in this search appear to correlate to two of the sequences with FBE similarity (Figure 6.9). It should be noted that this *yp/lsp-2* has not been shown to confer any tissue specific activity. Therefore, the significance of Ce*yp*$_3$ sequence similarities

Figure 6.9

Summary figure of the location of the putative regulatory elements in the Ce*yp*$_3$ gene.

A.Diagram of the Ce*yp*$_3$ gene and the location of the putative *LSP-2/yp* FBE sequence, FBE, OE1 AND OE2 regulatory elements.

B.Diagram of the Ce*yp*$_3$ gene and the location of putative DSX binding sites. The sites indicated in this diagram had a similarity of 7 out of 9 nucleotides with the sites obtained by Burtis *et al.* (1991). The DSX binding sites used in the searches are:

A.CTACAATGT

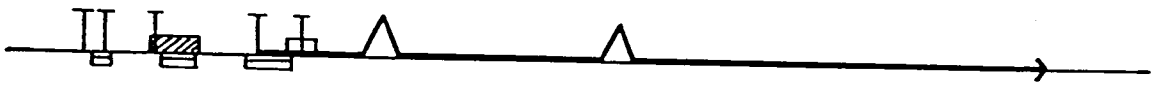
B:CTACAAAGT

C.TTACAAATT

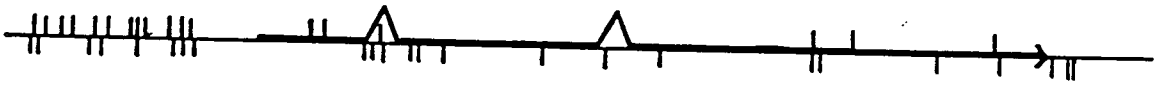
D.CTGCTAAGT

C. erythrocephala yp_3 *gene*

A



B



OE1

OE2

FBE

Dsx putative binding site

LSP2 and yp gene FBE region

100 bp

to the *yp/lsp-2* fragment is not known.

Discussion

The YPs of *Calliphora* are synthesised in the fat body and oocytes of female adult flies. The yp transcripts in both carcasses and ovaries first appear at stage 8 and are no longer present by stage 11 of oogenesis. The expression of the *yp* genes appears to be temporally co-ordinated in the two tissues. Oogenesis is not initiated until the female flies have been fed with raw meat and have mature males to fertilise the eggs. The yolk protein synthesis in the ovaries correlates well with the results obtained from the ovarian northern blot analysis, where the yp transcripts were expressed at stages 8, 9 10a and 10b. In *D. melanogaster* yp transcription is affected by nutritional factors (Bownes and Blair, 1986) and the synthesis of the YPs in the oocyte initiates at stage 8 through to stage 10b. Thus there must be a similarity in the regulation of these genes in the ovaries of both species. Although database searches using sequences found to regulate yp transcription in *D. melanogaster* do not appear to find Ce*yp*$_3$ sequences with significant similarities, these searches have highlighted areas of interest as summarised in Figure 6.6. The lack of high similarity to the *D. melanogaster* regulatory elements may be due to *Calliphora* having a different mode of regulation of the *yp* genes,

having a more A/T rich genome or perhaps not enough upstream DNA sequence has been determined. It appears that *Calliphora* may have a similar regulatory pathway of YP synthesis to *Musca* since in both species YP synthesis is induced in males by an injection of ecdysone and both species show the same response to a protein meal. In *Musca* the sex determination pathway has been proposed to be similar to that of *Drosophila* (Nothiger and Steinmann-Zwicky, 1985; Inoue and Hyroyoshi, 1986). In fact, a monoclonal antibody raised against the *D. melanogaster sex-lethal* (*Sxl*) protein product has been shown to cross-react with one *Musca* protein in a western blot of adult homogenate (R. Nothiger, personal communication). This experiment suggests that there is a homologous protein to Sxl in *Musca*. The same experiment could be carried out with *Calliphora* and may give an indication of how similar *Sex-lethal* in *Calliphora* is to that in *Musca*. If a *Calliphora* protein cross-reacts with the antibody raised against the *Drosophila Sxl* protein this may suggest a similar mechanism of regulation of sex determination and of the regulation of *yp* gene expression. The introduction of *yp* gene sequences from *Calliphora* into *D. melanogaster* via P-element mediated transformation, would determine whether or not the upstream regulatory sequences of the *yp* genes of *Calliphora* could direct the synthesis of *Calliphora* YPs in *Drosophila*. Similar experiments may be carried out by fusing the upstream sequences of the *Calliphora yp* genes to the coding

sequence of alcohol dehydrogenase or the *D. melanogaster yp* genes. Such experiments should indicate if a similar pathway of *yp* gene regulation is followed in *Calliphora* and *Drosophila*. Moreover it would indicate the degree of conservation of the factors involved in the regulation of expression. These types of experiments has been attempted with the chorion genes from *D. grimshawi* being transformed into *D. melanogaster*. The amplification control element of the third chromosome (ACE3) of *D. grimshawi* was shown to functionally substitute for that in *D. melanogaster* (Swimmer *et al.*, 1990). A more telling experiment was the introduction of *Bombyx mori* (silkmoth) chorion gene sequences into the *D. melanogaster* genome (via P-element mediated transformation); this resulted in the correct developmental expression of the silkmoth chorion genes in *Drosophila* (Mitsialis *et al.*, 1987). Therefore there appears to be a strong conservation of the chorion genes regulatory pathway and this may be the case between the *yp* genes of *Calliphora* and *Drosophila*.

A role for ovarian YP synthesis.

Vertebrates and some invertebrates synthesise the YPs or vitellogenins exclusively in the liver or fat body. It is therefore curious that the YPs should be synthesised in the ovaries of *Drosophila*, *Sarcophaga* and *Calliphora*. The chorion genes in the *Drosophila* species (Martinez-Cruzado *et al.*, 1988; Fenerjian *et al.*, 1989; Swimmer *et al.*,

1990) and the mediterranean fruit fly (Konsolaki *et al.*, 1990; Tolias *et al.*, 1990) are amplified to cope with the large requirements for their products during oogenesis. In other insects, such as *Bombyx mori*, the chorion genes are not amplified but several copies of each gene are present in the genome (Spoerel *et al.*, 1989). While the *yp* genes in *Drosophila* are not amplified nor have several copies of each gene in their genome, a large amount of YPs are required during vitellogenesis. This may explain the necessity of the two sites of synthesis for YPs (ie, fat body and ovaries) in *Drosophila*, *Calliphora* and *Sarcophaga*. Bownes *et al.* (1991) showed that the level of fertility of *D. melanogaster* is dependant upon the number of active *yp* genes, thus showing that a full complement of active *yp* genes in both ovaries and fat body is required for normal fertility.

Many insects store ecdysteroid conjugates in their oocytes (Hoffmann and Lagueux, 1985). Bownes *et al.* (1988) proposed that the YPs of *D. melanogaster* bind ecdysteroid conjugates which are then transported into the oocyte and stored. When the YPs are degraded for nutritional purposes the ecdysteroids are released. Thus the question is where and how do the YPs incorporate the conjugates of ecdysone? It would be reasonable to expect that the incorporation of ecdysone into YPs occurs in the sites of synthesis of the yolk proteins. In *Drosophila*, *Sarcophaga* and *Calliphora* that would be either the fat

body or the follicle cells of the oocyte. In *Locusta* and *Nauphoeta* the follicle cells of the ovary are involved in ecdysone synthesis (Goltzene *et al.*, 1978; Zhu *et al.*, 1983). The ovaries of *D.melanogaster* have been shown to synthesise ecdysteroids *in vitro* (Shwartz *et al.*, 1985), where stages 8 and 9 of oogenesis appear to be the main times of synthesis (Shwartz *et al.*, 1989). It is known that JH is required for further development of stages 8 and 9 of oogenesis (Wilson, 1982; Shwartz *et al.*, 1989). Although the fat body is under the influence of ecdysone (Bownes *et al.*, 1983), this tissue does not appear to synthesise it. It should be noted that the experiments carried out *in vitro* to measure ecdysone synthesis in different tissues have been contradictory, as data seems to vary according to culture medium and conditions used (Bownes, 1990b). Thus it is possible that the YPs synthesised in the ovaries incorporate ecdysone more readily that the YPs in the fat body. Such a proposal can be sustained since synthesis of the YPs and ecdysone in the ovaries is under the control of JH (Shwartz *et al.*, 1989) and because the amount of ecdysone present in the YPs is not equimolar to the number of YPs (Bownes *et al.*, 1988). Although the latter may be explained by one molecule of ecdysone being bound to one YP oligomer and thus result in 1:4 (or 1:6) ratio of ecdysone to monomer of YPs, it seems wasteful for ecdysone to occupy 1 site out of 4 (or, 1 out of 6) in the YP oligomer. It should be interesting to address this question by studying YPs

produced in the fat body and ovaries, and independently resolving the stoichiometry of apolar ecdysone complexes binding to YPs. This type of biochemical experiment may be more feasible in *Calliphora* than in *Drosophila* since in *Calliphora* it is possible to obtain tissues at specific stages of oogenesis. Moreover in *Drosophila* it is difficult to obtain enough material regardless of whether on not staging is required for the experiment.

A mechanism for the YP synthesis interruption at the end of vitellogenesis.

In *Drosophila* the YPs are synthesised continuously in the fat body whereas in the oocytes the synthesis of YPs is limited to stages 8 to 10b (inclusive) of individual oocytes. *Calliphora*, on the other hand, synthesises YPs in both tissues between stages 8 and 10b of oogenesis, at which time synthesis of the YPs stops until a new batch of oocytes has reached vitellogenesis. The regulation of this cycle and the events that lead to the cessation of YP synthesis in the fat body have been investigated in *Aedes aegypti* (Raikhel 1986). At the end of stage 10b of oogenesis two processes have been observed, transcription of the *yp* genes is interrupted and there is an increase in lysosomal activity specifically directed towards the degradation of the yolk proteins. Thus a concerted effort is launched by the fat body of the mosquito to stop the

synthesis and secretion of vitellogenins (Raikhel, 1989). In the fat body cells of *Calliphora* and *Sarcophaga* the same response may be envisaged at the end of vitellogenesis. The same changes may take place in *Drosophila*, *Calliphora* and *Sarcophaga* ovarian follicle cells since YP synthesis ceases and chorion protein synthesis proceeds. In any case, a signal must be given by the oocyte which causes the appearance of the degradative machinery. This however has to be a specific response to YPs. Although in *Drosophila* there is no evidence of fat body lysosomal activity directed towards the YPs since vitellogenesis is continous, a mutant, *fs(1)1163*, has been shown to secrete from the fat body cells the mutated YP_1^{1163} which coalesces and accumulates in the basal membrane of the fat body. This coalesced protein is re-absorbed into the fat body cells and the resulting vesicles fuse to lysosomes which in turn degrade this material (Butterworth *et al.*, 1991). This process appears to be specific to the YPs and a similar mechanism may be used by ovarian follicle cells at the end of vitellogenesis of individual egg chambers of *D. melanogaster*. Conversely, it is more likely that the regulatory mechanism in the ovaries turns off transcription prior to the end of vitellogenesis ensuring that all YPs are taken up prior to chorion secretion. Regardless of the actual mechanism there must be a regulatory signal involved in determining the timing of cessation of *yp* gene transcription and translation.

It has been proposed in *A. aegypti* that the YP titre increases at the end of vitellogenesis signalling the fat body to initiate interruption of YP synthesis and degradation of synthesised YPs (Raikhel, 1985). However, a mutant in *Drosophila*, *fs(2) A17/SM1*, is known to accumulate YPs in the haemolymph due to the lack of an ovarian sink yet YP synthesis does not appear to cease (Postlethwait and Handler, 1978), thus suggesting that ovarian development is required for a signal to be generated to interrupt *yp* gene transcription and initiate YP degradation in the fat body.

In summary, the *yp* genes of *Calliphora* and *Drosophila* have diverged sufficiently at the DNA level to impair easy recognition by computer searches of probable DNA regulatory elements shared by both *yp* genes of *Calliphora* and *Drosophila*. Despite this, the *yp* genes in both systems are regulated in an analogous fashion and thus must share a conserved mechanism of regulation.

The results obtained and presented in this thesis showed that YP uptake has been conserved though up to 100 MYR of Dipteran evolution. DNA sequence determination of the *Calliphora* yp_3 and yp_2 genes demonstrated a close relationship between *Calliphora* and *Drosophila* YPs. Comparison of the polypeptides illustrated the great conservation in amino acid sequence and thus in function between the YPs of the different species. These sequences explained the conservation of YP uptake through Dipteran evolution. However, the similarity in sequence hindered the identification of possible YP domains involved in their uptake into oocytes.

Database computer searches were carried out in order to identify proteins that may have a common domain with the YPs and whose function was known. In these searches five different sequence identities were found that may play a role in either YP/receptor interactions or in providing structural scaffolding to the YPs which will allow such interactions to take place. The sequence identity to the MHC type I and T-cell antigen receptor fall in the domains known to interact with each other; however, it may be that these similarities in the yolk proteins are not of sufficient length to result in any interactions. On the other hand these identities suggest that there may be specific sequences required for the

formation of the YP oligomers. The RGD motif in addition to the similarity to putative divalent ion binding motifs are possible sites for interaction with the YP receptor molecule. It is important to note that the putative tyrosine sulfation sites share the same sites as the putative divalent ion binding sites suggesting that these are exposed in the surface of the YPs and may therefore interact with the YP receptor.

The YPs are known to form oligomers, be ecdysone carriers, be secreted into the haemolymph and taken up by the oocyte, where they are sorted, stored and used during embryogenesis. These characteristics provide constraints in the polypeptide structure which ensure that all of the information that is required to fulfill the carrier and targeting functions is maintained and explain the conservation of the amino acid sequence (regardless of DNA divergence in species).

The *Calliphora* yolk protein genes are also expressed at specific stages of oogenesis in the carcass and the ovaries as shown by northern blots containing RNA from different stages of oogenesis. *In vitro* ovary culture showed that the YPs are synthesised in the same pattern shown for transcription of the *yp* genes. *In situ* hybridisation to *Calliphora* ovary whole mounts demonstrated that the *yp* genes are expressed in the follicle cells during stages 8 to 10b of oogenesis. All

this data indicates that the *yp* genes in *Calliphora* are expressed and synthesised in the carcass and follicle cells of ovaries between stages 8 and 10b of oogenesis. Thus not only the polypeptide sequence of the YPs is conserved in evolution but also the regulation of the expression of the *yp* genes shares conserved processes.

In the future there are a number of experiments which will provide information on a variety of areas concerning evolution and gene expression. In the first instance, the putative *yp* gene of *D. funebris* can be further characterised by DNA sequence determination and its specific pattern of expression investigated using *in situ* hybridisation to whole tissue mounts. Further, the identification and characterisation of other putative *yp* sequences in *D. funebris* and *C. erythrocephala* will help to understand the evolution of the *yp* genes through *Drosophila* and Diptera.

In a different set of experiments aimed at studying the likely domains of the YPs involved in uptake, the *Calliphora* yp_3 gene may be introduced into the *D. melanogaster* genome and its secretion, uptake, sorting, storage and usage investigated. These experiments would determine whether or not the *Calliphora* YPs can interact with *D. melanogaster* YPs and behave as the *Drosophila* YPs. These experiments also will be useful to investigate the expression and synthesis of an A/T rich foreign gene in

Drosophila.

The computer database searches identified putative domains of importance in the YPs which can be specifically amplified using PCR technology. Once these sequences are shown to be identical to the template, they can be fused to *lacZ* and introduced into the *D. melanogaster* genome. The construct will contain the signal peptide sequence to ensure that when the fusion protein is translated it is targeted into the endoplasmic reticulum (ER) and can thus be assayed for secretion, uptake, sorting and storage. These experiments can be carried out in conjunction with experiments analysing the upstream sequences of the *Calliphora yp* genes to determine if these sequences are recognised by the *Drosophila yp* regulatory machinery to give the pattern of expression characteristic of the *yp* genes. If this is the case, identification of the factors involved in *yp* gene expression and sex determination of *Calliphora* can be carried out and characterised.

Although the initial aims of this project were not fully achieved due to the YPs being so well conserved, a number of interesting avenues for research have been opened which should answer important questions pertaining to *yp* evolution, *yp* gene regulation, YP protein functions and sex determination in other Diptera.

REFERENCES

Abrahamsen, N., Martinez, A., Kjaer, T., Sondergaard, L. and Bownes, M. (1991) *Cis*-regulatory sequences leading to the fat body, sex specific, expression of the *yolk protein genes* 1 and 2 of *Drosophila melanogaster*. EMBO J. submitted.

Adams, T.S., Hagedorn, H.H. and Wheelock, G.D. (1985) Haemolymph ecdysteroid in the housefly, *Musca domestica*, during oogenesis and its relationship with vitellogenin levels. J. Insect Physiol. 31, 91-97

Ahle, S., Mann, A., Eichelsbacher, U.,and Ungewickell, E. (1988) Structure relationships between clathrin assembly proteins from the golgi and the plasma membrane. EMBO J. 7, 919-929.

Akam, M. (1987) The molecular basis for metameric pattern in the *Drosophila* embryo. Development 101, 1-22.

Appleyard, R.K (1956) Segregation of new lysogenic types during growth of a doubly lysogenic strain derived from *Escherichia coli* K12. Genetics 39, 440-459.

Ashburner M (1981) Entamophagous & other bizarre Drosophilidae In: Genetics & Biology of Drosophila ed. Ashburner M, Carson H L & Thomson J.N. Academic Press London Vol 3a pp 375-422.

Ashburner, M. (1989) *Drosophila*, a laboratory handbook. Cold Spring Harbour Laboratory Press, New York.

Baeuerle P A & Huttner W B (1987) Tyrosine sulfation is a trans-golgi-specific protein modification. J. Cell. Biol. 105 2655-2664

Baeuerle, P.A. and Huttner, W.B. (1985) Tyrosine

sulfation of yolk proteins 1,2 and 3 in *Drosophila melanogaster*. J. Biol. Chem. 260, 6434-6439.

Baker, B.S. (1989) Sex in flies: the spice of life. Nature 340: 521-524.

Baker, B.S. and Belote, J.M. (1983) Sex determination and dosage compensation in *Drosophila melanogaster*. Ann. Rev. Genet. 17: 345-393.

Barnett, T. and Wensink, P.C. (1981) Transcription and translation of yolk protein mRNA in the fat bodies of *Drosophila*. In: Developmental Biology using purified genes, ed. D. Brown, ICN-UCLA Symposium on Molecular and Cellular Biology. Academic Press, New York. pp 97-106.

Barnett, T., Pachl, C., Gergen, J.P. and Wensink, P.C. (1980) The isolation and characterisation of *Drosophila* yolk protein genes. Cell 21, 729-738.

Batterham, P., Chambers G.K Stamer W.T and Sullivan D.T (1984) Origins expression of an alcohol dehydrogenase gene duplication in the genus Drosophila Evolution 38 : 644-654.

Bell, L.R., Horabin, J.I., Schedl, P. and Cline, T.W. (1991) Positive autoregulation of *sex-lethal* by alternative splicing maintains the female determined state in *Drosophila*. Cell 65, 229-239.

Belote, J.M., Handler, A.M., Wolfner, M.F., Livak, K.J. and Baker, B.S. (1985) Sex-specific regulation of yolk protein gene expression in *Drosophila*. Cell 40, 339-348.

Bewley, G.C., Cook, J.L., Kusakabe, S., Mukai, T., Rigby, D.L. and Chambers, G.K. (1989). Sequence, structure and evolution of the gene coding for sn-glycerol-3-phosphate dehydrogenase in *Drosophila melanogaster*. Nucleic Acid Res. 17: 8553-8567.

Beverley, M.S. and Wilson, A.C. (1982) Molecular evolution in *Drosophila* and higher Diptera I.Micro-complement fixation studies of a Larval Haemolymph Protein.J. Mol. Evol.18: 251-264.

Beverley, M.S. and Wilson, A.C. (1984) Molecular evolution in *Drosophila* and the higher Diptera II. A time scale for fly evolution. J. Mol. Evol.*21*: 1-13.

Beverley, M.S. and Wilson, A.C. (1985) Ancient origin for Hawaiian Drosophilinae inferred from protein comparisons. P.N.A.S. USA 82: 4753-4657.

Bier, K. (1959) Quantitative untersuchungen uber die variabilitat der nahrzellkiernstrucktur und ihre beeinflussung durch die temperature. Chromosoma 10, 619-653.

Bier, K. (1963) Synthese, interzellurarer transport, und abbau von ribonukleinsaure im ovar der stubenfliege *Musca domestica*. J. Cell Biol. 16, 436-440.

Bier, K., Kunz, W. and Ribbert, D. (1967) Struktur and funktion der oocytenchromosomen und nucleolen sowie der extra-DNS wahend de oogenese panoistischer und meriostischer insekten. Chromosoma 23, 214-254.

Blackman, R.K. and Meselson, M. (1986) Interspecific nucleotide sequence comparisons used to identify regulatory and structural features of the *Drosophila hsp82* gene. J. Mol. Biol. 188, 499-515.

Bock, P (1972) Adsorption of horseraddish peroxidase to negatively charged groups. Acta Histochem 43 8-15.

Bownes, M. (1980) The use of yolk protein variations in *Drosophila* species to analyse the control of vitellogenesis. Differentiation 16: 109-116.

Bownes, M. (1982) Hormonal and genetic regulation of vitellogeniesis in *Drosophila*. The Quarterly Review of Biology 57: 247-274.

Bownes, M. (1986) Expression of the genes coding for vitellogenins (Yolk Proteins). Ann. Rev. Entomol. 31, 507-531.

Bownes, M. (1990a) The yolk proteins and their genes in *Drosophila*. Progress in Comparative Endocrinology. pp 336-342.

Bownes, M. (1990b) The nature and origin of ecdysteroids

during *Drosophila* metamorphosis. In Molting and metamorphosis. Ed. E. Ohnishi and H. Ishizahi. Japan Scientific Societies Press and Springer-Verlag, Tokyo. pp 107-120.

Bownes, M. and Hames, B.D. (1978) Analysis of the yolk proteins in *Drosophila melanogaster*. FEBS Lett. 96, 327-330.

Bownes, M. and Hodson, B.A. (1980) Mutant *fs(1) 1163* of *Drosophila melanogaster* alters yolk protein secretion from the fat body. Molec. Gen. Genet. 180: 411-418.

Bownes, M. and Nothiger, R. (1981) Sex determining genes and vitellogenin synthesis in *Drosophila melanogaster*. Mol. Gen. Genet. 182: 222-228.

Bownes, M. and Dale, L. (1982) Gametogenesis. In: Handbook of *Drosophila* development, ed. R. Ransom, Elsevier Biochemical Press, Oxford. pp 31-66.

Bownes, M. and Blair, M. (1986) The effect of a sugar diet and hormones on the expression of the *Drosophila* yolk protein genes. J. Insect Physiol. 32: 493-501.

Bownes, M., Dempster, M. and Blair, M. (1983) The regulation of yolk protein gene expression in *Drosophila melanogaster*. In Molecular Biology of Egg Maturation,

CIBA Foundation Symposium 98: 63-76.

Bownes, M., Smith, T. and Blair, M. (1985) The regulation of expression of the genes coding for the yolk proteins in *Drosophila*. Int. J. Invert. Rep. Dev. 8: 253-261.

Bownes, M., Scott, A. and Blair, M. (1987) The use of an inhibitor of protein synthesis to investigate the roles of ecdysteroids and sex-determination genes on the expression of the genes encoding the *Drosophila* yolk proteins. Development **101**: 931-941.

Bownes, M., Scott, A. and Shirras, A. (1988a) Dietary components modulate yolk protein gene transcription in *Drosophila melanogaster*. Development 103: 119-128.

Bownes, M., Steinmann-Zwicky, M. and Nothiger, R. (1990) Differential control of yolk protein gene expression in fat body and gonads by the sex-determining gene *tra-2* of *Drosophila*. EMBO J. 9, 3975-3980.

Bownes, M., Lineruth, K. and Mauchline,D. (1991) Egg production and fertility depends upon the yolk-protein gene copy number. Mol Gen. Genet, in press.

Bownes, M., Blair, M., Kozma, R. and Dempster, M. (1983) 20-hydroxyecdysone stimulates tissue-specific yolk-protein gene transcription in both male and female *Drosophila*. J.

Embryol. Exp. Morph. 78: 249-268.

Bownes, M., Shirras, A., Blair, M., Collins, J. and Coulson, A. (1988b) Evidence that insect embryogenesis is regulated by ecdysteroids released from yolk proteins. P.N.A.S.USA 85, 1554-1557.

Branch, W.J., Mullock, B.M, Luzio J.P. (1987) Rapid subcellular fractionation of the rat liver endocytic compartments involved in transcytosis of polymeric immunoglobulin A and endocytosis of asialofetuin. Biochem J. 244 311-315.

Brennan, M.D., Warren, T.G. and Mahowald, A.P. (1980) Signal peptides and signal peptidases in *Drosophila melanogaster*. J. Cell Biol. 87, 516-520

Brennan, M.D., Weiner, A.J., Goralski, T.J. and Mahowald, A.P. (1982) The follicle cells are a major site of vitellogenin synthesis in *Drosophila melanogaster*. Dev. Biol. 89, 225-236.

Brodsky, F.M., Hill, B.L., Acton, S.L., Nathke, I., Wong,D.H., Ponnambalam, S. and Parham, P. (1991) Clathrin light chains: arrays of protein motifs that regulate coated-vesicle dynamics. TIBS 16, 208-213.

Burtis, K.C., Thummel, C.S., Weldon Jones, C., Karim, F.D. and Hofness, D.S. (1990) The *Drosophila* 74EF early puff contains E74, a complex ecdysone-inducible gene that encodes two *ets*-related proteins. Cell 61, 85-99.

Burtis, K.C., Coschigano, K.T., Baker, B.S. and Wensink, P.C. (1991) *Drosophila doublesex* protein binds to a

sex-specific yolk protein gene enhancer. EMBO. J. in press.

Butterworth, F.M., Bownes, M. and Burde, V.S. (1991) Genetically modified yolk proteins precipitate in the adult *Drosophila* fat body. J. Cell Biol. 112: 727-737.

Byrne, B.M., Gruber, M. and Ab, G. (1989) The evolution of egg yolk proteins. Prog. Biophys. Mol. Biol. 53, 33-69.

Camerini, D. and Seed, B. (1990) A CD4 domain important for HIV-mediated syncytium formation lies outside the virus binding site. Cell 60, 747-754.

Cardeon, J., Huybrechts, R. and De Loof, A. (1986) Yolk polypeptide processing in the fat body of *Sarcophaga bullata*. Arch. Insect Biochem. Physiol. 7, 87-96.

Chadha, G.K. and Denlinger, D.L. (1976) Autogeny and endocrine involvement in the reproduction of tropical flies (Diptera, Sarcophagidae). Entomol. Exp. Appl. 20, 31-38.

Chandly, A.C. (1966) Studies on oogenesis in *Drosophila melanogaster* with ^{3}H-thymidine label. Exp. Cell Res. 44, 201-215.

Cherbas, L., Schulz, R.A., Koehler, M.M.D., Savakis, C. and Cherbas, P. (1986) Structure of the *Eip28/29* gene, an ecdysone-inducible gene from *Drosophila*. J. Mol. Biol. *189*: 617-631.

Collier, G.E. (1990) Evolution of arginine kinase within the genus *Drosophila*. J. Heredity 81, 177-182.

Collier, G.E. and MacIntyre, R.J. (1977) Microcomplement fixation studies on the evolution of --glycerophosphate dehydrogenase within the genus *Drosophila*. P.N.A.S. USA 74: 684-688.

Dautry-Varsat, A., Ciechanover, A and Lodish, H (1983) pH and the recycling of transferrin during receptor mediated endocytosis. P.N.A.S. USA 80 2258-2262

Davis, C.G., van Driel, I.R., Russell, D.W., Brown, M.S. and Goldstein, J.L. (1987) The low density lipoprotein receptor identification of amino acids in the cytoplasmic domain require for rapid endocytosis. J. Chem. Biol. 262, 4075-4082.

De Barun, K., Brownlee, G.G., Kendal, A.P. and [illegible] M.W. (1988) Complete sequence of a cDNA clone of the hemagglutinin gene of influenza A/Chicken/Scotland/59 (H5NI) virus: comparison with comtemporary Noth American and European strains. Nucleic Acid Res. 16, 4181-4182.

Delaney, S.J., Smith, D.F., McClelland, A., Sunkel, C. and Glover, D.M. (1986) Sequence conservation around the 5' ends of the Larval Serum Protein 1 genes of *Drosophila*

melanogaster. J. Mol. Biol. *189*: 1-11.

Della-Cioppa, G. and Engelmann, F. (1987) The vitellogenin of *Leocophaea maderae* synthesis as a large phosphorylated precursor. Insect Biochem. 17: 401-415.

DeLuca-Flaherty, C., McKay, D.B., Parham, D. and Hill, B.L. (1990) Uncoating protein (hsc70) binds a conformationally labile domain of clathrin light chain LC. to stimulate ATP hydrolysis. Cell 62, 875-887.

Deng, P. and Storrie, B. (1988) Animal cell lysosomes rapidly exchange mambrane proteins. P.N.A.S. USA 85, 3860-3864.

Denlinger, D.L. (1971) Autogeny in the fleshfly *Sarcophaga argyrostoma* Ann. Entomol. Soc. Amer. 64, 961-962.

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

Diaz, R., Mayorga, L.S., Mayorga, L.E. and Stahl, P. (1989) *In vitro* clustering and multiple fusion among macrophage endosomes. J. Biol. Chem. 264, 13171-13180.

DiMario, P.J. and Mahowald, A.P. (1987) *Female sterile(1)*

yolkless: A recessive female sterile mutation in *Drosophila melanogaster* with depressed numbers of coated pits and coated vesicles within the developing oocyte. J. Cell Biol. 105: 199-206.

DiMario, P.J., Warren, T.G. and Mahowald, A.P. (1987) The purification and *in vitro* phosphorylation of vitellogenin from *Drosophila melanogaster*. Insect Biochem. 17, 1187-1192.

Donk, K., Craddock, E.M. and Kambysellis, M.P. (1985) Conservation in the organisation of the vitellogenic genes in *Drosophila* species. Genetics 110, s56.

Draye, J.P., Quintart, J., Courtoy, P.J. and Boudhuin, P. (1987) Relations betwenn plasma and lysosomal membrane. I. Fate of covalently labelled plasma membrane protein. Eur. J. Biochem. 170, 395-403.

Duben, R., Griffiths, G., Frank, R., Argos, P. and Kreis, T.E. (1991) --COP, a 110 kD protein associated with non-clathrin-coated vesicles and the golgi complex, shows homology to --adaptin. Cell 64, 649-665.

Dunn, W.A., Hubbard, A.L. and Aronson, N.N.Jr(1980) Low temperature selectively inhibits fusion between pinocytic vesicles and lysosomes during heterphagy of ^{125}I asialofetuin by perfused rat liver. J. Biol. Chem. 255,

5971-5978.

Dutton, F.L. and Chovnick, A. (1988) Developmental regulation of the *rosy* locus in *Drosophila melanogaster*. In. Developmental Biology, Ed. L.W.Browder, Plenum, New York. Vol 5, 267-316.

Enerback, S. and Bjursell, G. (1989) Genomic organisation of the region encoding guinea pig lipoprotein lipase; evidence for exon fusion and unconventional splicing. Gene 84, 391-397.

Fang, X-M., Wu, C-Y. and Brennan, M.D. (1991) Complexity in evolved regulatory variation for alcohol dehydrogenase genes in Hawaiian *Drosophila*. J. Mol.Evol. *32*: 220-226.

Fenerjain, M.G., Martinez-Cruzado, J.C., Swimmer, C., King, D. and Kafatos, F.C. (1989) Evolution of the autosomal chorion cluster in *Drosophila* II. Chorion gene expression and sequence comparisons of hte *s16* and *s19* genes in evolutionaruly distant species. J. Mol. Evol. 29, 108-125.

Ferris, A.L., Brown, J.C., Park, R.D. and Storrie, B. (1987) Chinese hamster ovary cell lysosomes rapidly exchange contents. J. Cell Biol. 105, 2703-2712.

Fourney, R., Pratt, G., Harnish, D., Wyatt G. and White B (1982) Structure & synthesis of vitellogenins and vitellin from Calliphora erythrocephala Insect. Biochem. 12 311-321

Fournier, D., Karch, F., Bride, J-M., Hall, L.M.C., Berge,

J-B. and Spierer, P. (1989) *Drosophila melanogaster* acetylcholinesterase gene structure, evolution and mutation. J. Mol. Biol. 210: 15-22.

Fraenkel, G. (1940) Utilisation and digestion of carbohydrates by adult blowfly. J. Exp. Biol. 17, 18-29.

Franco, M.G., Rubini, P.G. and Vecchi, M. (1982) Sex-determinants and their distribution in various populations of *Musca domestica* L. of Western Europe. Genet. Res. Camb. 40: 279-293.

Friederich, E., Baeuerle, P.A., Garoff, H., Hovemann, B. and Huttner, W.B. (1988) Expression, tyrosine sulfation, and secretion of yolk protein 2 of *Drosophila melanogaster* in mouse fibroblasts. J. Biol. Chem. 263, 14930-14938.

Garabedian, M.J., Hung, M-C. and Wensink, P.C. (1985) Independent control elements that determine yolk protein gene expression in alternative *Drosophila* tissues. P.N.A.S. USA 82: 1396-1400.

Garabedian, M.J., Shepherd, B.M. and Wensink, P.C. (1986) A tissue-specific transcription enhancer from the *Drosophila* yolk prtoein 1 gene. Cell 45: 859-867.

Garabedian, M.J., Shirras, A.D., Bownes, M. and Wensink,

P.C. (1987) The nucleotide sequence of the gene coding for *Drosophila melanogaster* yolk protein 3. Gene 55: 1-8.

Gelti-Douka, H., Gingeras, T.R. and Kambysellis, M.P. (1974) Tolk proteins in *Drosophila*: identification and site of synthesis. J. Exp. Zool. 187, 167-172.

Geysen, J., Cardeon, J. and De Loof, A. (1987) Distribution of yolk polypeptides in the follicle cells during differentiation of the follicular epithelium in *Sarcophaga bullata* egg follicles. Development 101, 33-44.

Gingeras, T.R., Gelti-Douka, H. and Kambysellis, M.P. (1973) Yolk proteins in *Drosophila*. Dros. Imf. Serv. 50, 58.

Giorgi, F. (1979) *In vitro* induced pinocytotic activity by a juvenile hormone analogue in oocytes of *Drosophila melanogaster*. Cell Tissue Res. 203, 241-247.

Giorgi, F., Andolfi, P., Spinetti, I., Masetti, M. and Postlethwait, J. (1986) Vitellogenin endocytosis after transfusion of haemolymph from a female sterile mutant into normal *Drosophila*. J. Insect Physiol.

Glickman, J.N., Conibear, E. and Pearce, B.M.F. (1989) Specificity of binding of clathrin adaptors to signals onthe mannose-6-phosphate/insulin-like growth factor II

receptor. EMBO J. 8, 1041-1041.

Gochoco, C.H., Kunkel, J.G. and Nordin, J.H. (1984) The role of oligosaccharides in endocytosis of an insect vetellogenin by the oocyte. J. Cell Biol. 99, 372a.

Gochoco, C.H., Kunkel, J.G. and Nordin, J.H. (1988) Experimental modifications of an insect vitellin affect its structure and its uptake by oocytes. Archives of Insect Biochemistry and Physiology 9, 179-199.

Goda, Y. and Pfeffer, S.R. (1988) Selective recycling of the mannose-6-phosphate/IGF-II receptor to the trans golgi network *in vitro*. Cell 55, 309-320.

Goldstein, J.L., Brown, M.S., Anderson, R.W.G., Russell, D.W., and Schneider, W.J. (1985) Receptor mediated endocytosis: concepts emerging from the LDL receptor system, Annu. Rev. Cell Biol. 1, 1-39.

Goltzene, F., Lagueux, M., Charlet, M. and Hoffmann, J.A. (1978) Follicle cell epithelium of maturing ovaries of *Locusta migratoria*: new biosynthetic tissue for ecdysone. Hoppe Seylers Z. Physiol. Chem. 359, 1427-1424.

Goodman, T., Morrison, P.E and Davies D M (1968) Cytological changes with developing ovary of the horse fly fed milk and other diets Can. J. Zool 46 409-421

Grossberger, D. (1987) Minipreps of DNA from bacteriophage lambda. Nucleic Acids Res. 15,

Gruenberg, J. and Howell, K.E. (1989) Membrane traffic in endocytosis: insights from cell-free assays. Annu. Rev. Cell Biol. 5, 453-481.

Gruenberg, J., Griffiths, G. and Howell, K.E. (1989) Characterisation of the early endosome and putative endocytic carrier vesicles *in vivo* and with an assay of vesicle function *in vitro*. J. Cell Biol. 108, 1301-1316.

Gutzeit, H.O. (1986) Transport of molecules and organelles in meroistic ovarioles of insects. Differentiation 31, 155-165.

Hackman, W. and Vaisanen, R. (1982) Different classification systems in the Diptera. Ann. Zool. Fennici. 19, 209-219.

Hagele, K. (1985) Identification of a polytene chromosome band containing a male sex determiner of *Chironomus thummi thummi*. Chromosoma 91, 167-171.

Hames, B.D. and Bownes, M. (1978) Synthesis of yolk proteins in *Drosophila melanogaster*. Insect Biochem. 8, 317-328.

Harlow, P.M. (1956) A study of ovarian development and its relation to adult nutrition in the blowfly, *Protophormia terrae-novae* (R.D). J. Exp. Biol. 33,

777-797.

Hatzopoulos, P. and Kambysellis, M.P. (1987) Isolation ans structural analysis of the vitellogenin genes from *Drosophila grimshawi*. Mol. Gen. Genet. 206, 475-484.

Hatzopoulos, P. and Kambysellis, M.P. (1988) Comparative biochemical and immunological analysis of the three vitellogenins from *Drosophila grimshawi*. Comp. Biochem. Physiol. 89b, 557-564.

Hawkins, J.D. (1988) A survey on intron and exon lengths. Nucleic Acids Research 16, 9893-9905.

Haylett, T. and Thilo, L. (1986) Limited and selective transfer of plasma membrane of secondary lysosomes. J. Cell Biol. 103, 1249-1256.

Hedley, M.L. and Maniatis, T. (1991) Sex-specific splicing and polyadenylation of *dsx* pre m-RNA requires a sequence that binds specifically to *tra-2* protein. Cell 65, 579-586.

Helenius, A., Mellmann, I., Wall, D. and Hubbard, A. (1983) Endosomes. TIBS 8, 245-250.
Henikoff, S and Eghtedarzaden MK (1987) Conserved arrangement of nested genes at the Drosophila gart locus Genetics 117 711-725.

Hendrick, S.M., Nielsen, E.A., Kavaler, J., Cohen, D.I. and Davis, M.M. (1984) Sequence relationships between
Heuser, J. (1989) Changes in lysosyme shape + distribution correlated with changes in cytoplasmic pH. J. Cell. Biol. 108 855-864

putative T-cell receptor polypeptides and immunoglobins. Nature 308, 153-158.

Hertzog, V and Reggio H (1980) Pathways of endocytosis from luminal plasma membrane in rat exocrine pancreas. Eur J. Cell Biol. 21 141-150

Hopkins, C.R. (1985) The appearance and internalisation of transferrin receptors at the margins of spreading human tumor cells. Cell 10, 199-208.

Hopkins, C.R. (1983) Intracellular routing of transferrin and transferrin receptors in epidermoid carcinoma A431 cells Cell 35 321-330.

Hortin, G.L. and Graham, J.P. (1988) Sulfation of tyrosine residues does not influence secretion of α_2-antiplasmin or C4. Bioche. Biophys. Res. Comm. 151, 417-421.

Hortin, G., Folz, R., Gordon, R.I. and Strauss, A.W. (1986) Characterisation of sites of tyrosine sulfation in proteins and criteria for predicting their occurrence. Biochem. Biophys. Res. Comm. 141, 326-333.

Houde, M., Tireron, M-C. and Bregegere, F. (1989). Divergence ofthe nucleotide sequences encoding xanthine dehydrogenase in *Calliphora vicina* and *Drosophila melanogaster*. Gene 85, 391-402.

Hovemann, B., Galler, G., Walldorf, U., Kupper, H. and Bautz, E.K.F. (1981) Vitellogenin in *Drosophila melanogaster*: sequence of the yolk protein 1 gene and its flanking region. Nucleic Acids Res. 9, 4721-4734.

Hung, M. and Wensink, P.C. (1981) The sequence of

Drosophila melanogaster gene for yolk protein. Nucleic Acids Res. 9, 6407-6419.

Hung, M. and Wensink, P.C. (1983) Sequence and structure conservation in yolk proteins and their genes. J. Mol. Biol. 164, 481-489.

Hung, M., Barnett,T., Woolford,C. and Wensink, P.C. (1982) Transcript Maps of *Drosophila* Yolk Protein Genes. J. Mol. Biol. 154, 581-602.

Hurtley, JM (1991). Clathrin coated and non-clathrin coated vesicle adapters TiBS 16, 165-166

Huttner, W.B. and Baeuerle, P.A. (1988) Protein sulfation on tyrosine. Modern Cell Biology 6, 97-140.

Huybrechts, R. and De Loof, A. (1982) Similarities in vitellogenin and control of vitellogenin synthesis within the genera *Sarcophaga*, *Calliphora*, *Phormia* and *Lucilia* (Diptera). Comp. Biochem. Physiol. 72B, 339-344.

Imboden H, Kong R, Ott P, Lusha A, Kampfer U & Lanzrein B (1987) Characterization of the native vitellogenin and vitellin of the cockroach Nauphoeta cinerae and comparison with other species. Insect Biochem 17 353-365.

Inoue, H. and Hiroyoshi, T. (1986) A maternal-effect sex-transformation mutant of the housefly, *Musca domestica L.* Genetics 112, 469-482.

Inoue, S., Inoue, H., Hiroyoshi, T., Matsubara, H. and Yanamaka, T. (1986) Developmental Variation and Amino Acid Sequences of Cytochromes *c* of the Fruit Fly *Drosophila melanogaster* and the Flesh Fly *Boettcherisca peregrina*. J. Biochem. 100, 955-965.

Inoue, S., Matsubara, H. and Yanamaka, T. (1985) Complete Amino Acid Sequence of Cytochrome c from the Honeybee, *Apis mellifera*, and Evolutionary Relationship of the Honeybee to Other Insects on the Basis of the Amino Acid Sequence. J. Biochem. 97, 947-954.

Isaac, P.G. (1982) Molecular studies of the vitellogenins of *Drosophila melanogaster*. Ph.D Thesis, University of Edinburgh.

Isaac, P.G. and Bownes, M. (1982) Ovarian and Fat-Body Vitellogenin Synthesis in *Drosophila melanogaster*. Eur. J. Biochem. 123, 527-534.

Jackson, A.P., Seow, H.-F., Holmes, N., Drickamer, K. and Parham, P. (1987) Clathrin light chains contain brain-specific insertion sequences and a region of homology with intermediate filaments. Nature 326, 154-159.

Jacob, J. and Sirlin, J.L. (1959) Cell function in the ovary of Drosophila. I. DNA classes in nurse cells as determined by autoradiography. Chromosoma 10, 210-228.

Jeffery, J., Perssons, B., and Jornvall, H. (1987) Characteristics of alcohol/polyol dehydrogenases. Eur. J. Biochem. 167, 195-201.

Jing, S., Spencer, T., Miller, K., Hopkins, C. and Trowbridge, I.S. (1990) Role of the human transferrin receptor cytoplasmic domain in endocytosis: localisation

of a specific signal sequence for internalisation. J. Cell Biol. 110, 283-294.

Jones, C.W., Dalton, M.W. and Townley, L.H. (1991) Interspecific comparisons of the structure and regulation of the *Drosophila* ecdysone-inducible gene *E74*. Genetics, 127, 535-543.

Jowett, T. and Postlethwait, J.H. (1980) The regulation of yolk polypeptide synthesis in *Drosophila* ovaries and fat body by 20-hydroxyecdysone and a juvenile hormone analog. Dev. Biol. 80, 225-234.

Kambysellis, M.P. (1977) Genetic and hormonal regulation of vitellogenesis in *Drosophila*. Am. Zool. 17, 535-550.

Kassis, J.A., Poole, S.J., Wright, D.K. and O'Farrell, P.H. (1986) Sequence conservation in the protein coding and intron regions of the engrailed transcription unit. EMBO J. 5, 3583-3589.

Kassis, J.A., Desplan, C., Wright, D.K. and O'Farrell, P.H. (1989) Evolutionary conservation of homeodomain-binding sites and other sequences upstream and within the major transcription unit of the *Drosophila* segmentation gene *engrailed*. Mol. Cell. Biol. 9, 4304-4311.

Keen, J.H. (1987) Clathrin assembly proteins: affinity purification and a model for coat assembly. J. Cell Biol. 105, 1989-1998.

Kindle, H., Lanzrein, B. and Kunkel, J.G. (1990) the efect of ions, ion channel blockers, and ionophores on uptake of vitellogenin in cockroach follicles. Dev. Biol. 142, 381-391.

King, R.C., Rubinson, A.C. and Smith, R.F. (1956) Oogenesis in adult *Drosophila melanogaster*. Growth 21, 121-157.

King, R.C. (1970) Ovarian development of *Drosophila melanogaster*. Academic Press, New York.

King, R.C. and Koch, E.A. (1963) Studies on ovarian follicle cells of *Drosophila*. Q. J. Microsc. Sci. 104, 297-320.

Kingberg, G.M, Ford, T., Blomhoff R., Rickwood, D., Berg, T (1984). Separation of endocytic vesicles in mycodenz gradients. Anal. Biochem 142 455-462.

Kink, J.A., Maley, M.E., Preston, R.R., Ling, K-Y., Wallen-Friedman, M.A., Saimi, Y. and Kung, C. (1990) Mutations in *Paramecium* calmodulin indicate functional differences between the C-terminal and N-terminal lobes *in vivo*. Cell 62, 165-174.

Kirchhausen, T. and Harrison, S.C. (1981) Protein organisation in clathrin trimers. Cell 23, 755-761.

Klausner, R.D., van Renswoude, J., Aswell, G., Kempf, C., Schechter, A.N., Deang, A and Bridges, K.R (1983) Receptor mediated endocytosis of transferrin in K562 cells. J. Biol. Chem. 258 4715-4724

Koch, E.A. and King, R.C. (1964) Studies on the *fes* mutant of *Drosophila melanogaster*. Growth, 28, 325-369.

Koller, C.N., Dhadialla, T.S. and Raikhel, A.S. (1989) Selective Endocytosis of Vitellogenin by Oocytes of the Mosquito, *Aedes aegypti*: an *in vitro* study. Insect Biochem. 19, 693-702.

Komaromy, M.C. and Schotz, M. (1987) Cloning of rat hepatic lipase cDNA: evidence for a lipase gene family. P.N.A.S. USA 84, 1526-1530.

Konig, R. and Lanzrein, B. (1985) Binding of vitellogenin to specific receptors in oocyte membrane preparations of the ovoviviparous cockroach *Nauphoeta cinerae*. Insect Biochem. 15, 735-747.

Konig,R., Kindle, H., Kunkel, J.G and Lanzrein, B. (1988a) Vitellogenesis in the cockroach *Nauphoeta cinerea*: separation of two classes of ovarian binding sites and calcium effects on binding and uptake. Archives of Insect Biochemistry and Physiology 9, 323-337.

Konig, R; Nordin, J.H; Gochoco C.H and Kunkel J.G. (1988b) The role of the vitellin high mannose oligosaccharides in binding to vitellogenin receptors in follicle membrane preparations of the German cockroach Blatella Germanica Insect. Biochem 11. 395

Konsolaki, M., Komitopoulou, K., Tolias, P.P., King, D.L., Swimmer, C. and Kafatos, F.C. (1990) The chorion genes of the medfly, *Ceratitis capitata*. I: structural and regulatory conservation of the *s36* gene relative to two

Drosophila species. Nucl. Acids Res. 18, 1731-1737.

Kozma, R.S. (1985) Evolution and regulation of expression of the genes coding for *Drsophila* yolk proteins. Ph.D Thesis, University of Edinburgh.

Kozma, R. and Bownes, M (1986a) Yolk protein induction in males of several *Drosophila* species. Insect Biochem. 16, 263-271.

Kozma, R. and Bownes, M. (1986b) Identification and comparisons of the yolk polypeptides and the genes which code for them in *D.melanogaster* sibling species. Mol. Gen. Genet. 204, 302-309.

Kraminsky, G.P., Clark, W.C., Estelle, M.A., Gietz R.A., Sage, B.A., O'Conner J.D., Hognetts R.B (1980) Induction of translatable mRNA's for dopa decarboxylase in Drosophila : an early response to ecdysterone? PNAS USA 77 4175-4178

Kraus, K.W., Lee, Y., Lis, J.T. and Wolfner, M.F. (1988) Sex-Specific Control of *Drosophila melanogaster* Yolk Protein 1 Gene Expression Is Limited to Transcription. Molecular and Cellular Biology 8, 4756-4764.

Kulakosky, P.C. and Telfer, W.H. (1987) Selective endocytosis, *in vitro*, by ovarian follicles from *Hyalophora cecropia*. Insect Biochem. 17, 845-858.

Kunkel, J.G and Nordin J.H (1985) Yolk Proteins. In: Comprehensive Insect Physiology, Biochemistry & Pharmacology. eds. Gilbert L.I, Kerbur G.A. Pergamon Press, New York Vol 1 pp 83-111

Latorre, A., Moya, A. and Ayala, F.J. (1986) Evolution of mitochondrial DNA in *Drosophila sudobscura*. P.N.A.S.USA. 83, 8649-8653.

Laemmli, U.K. (1970) Cleavage of structural proteins during assembly of the head of bacteriophage T_4. Nature 227, 680-685.

Lagueux, M., Harry, P. and Hoffmann, J.A. (1981) Ecdysteroids are bound to vitellin in newly laid eggs of *Locusta*. Mol. Cell. Endocrinol. 24, 325-338.

Lamnissou, K. and Zouros, E. (1989) Interspecific ovarian transplants in *Drosophila*: vitellogenin uptake as an index of evolutionary relatedness. Heredity 63, 29-35.

Lawrence, P.A. and Johnston, P. (1986) The muscle pattern of a segment of *Drosophila* may be determined by neurons and not by contributing myoblasts. Cell *45*, 505-513.

Le Cam, A., Pages, G., Auberger, P., Le Cam, G., Loepold, P., Benarous, R. and Glalchenhous, N. (1987) Study of a growth hormone regulated protein secreted by rat hepatocytes: cDNA cloning, anti-protease activity and regulation of its synthesis by various hormones. EMBO J. 6, 1225-1232.

Lee, R.W.H. and Huttner, W.B. (1985) $(Glu^{62}, Ala^{30}, Tyr^{8})_n$ serves as high-affinity substrated for tyrosylprotein sulfotransferase: A golgi enzyme. P.N.A.S. USA 82, 6143-6147.

Liddell, S. (1990) Expression of the yolk protein 3 gene of *Drosophila melanogaster*. Ph.D Thesis, University of Edinburgh.

Liddell, S. and Bownes, M. (1991a) Characterisation, molecular cloning and sequencing of $yp3^{s1}$, a female yolk protein mutant in *Drosophila*. Mol. Gen. Genet. in press.

Liddell, S. and Bownes, M. (1991b) Investigation of *cis*-acting sequences regulating expression of the gene encoding yolk protein 3 in *Drosophila melanogaster*. Mol. Gen. Genet. in press.

Lineruth, K. and Lindberg, M. (1988) Electrophoretic variant forms of a set of follicle cell proteins in *Drosophila melanogaster*: Frequencies in wild-type strains and response in a selection experiment. Herieditas 108, 59-64.

Lobel, P., Fujimoto, K., Ye, R.D., Griffiths, G. and Kornfeld, S. (1989) Mutations in the cytoplasmic domain of the 275 kD mannose 6-phosphate receptor differentially alter lysosomal enzyme sorting and endocytosis. Cell 57, 787-796.

Logan, S., Garabedian, M.J. and Wensink. P.C. (1989) DNA regions that regulate the ovarian transcriptional

specifity of *Drosophila* yolk protein genes. Genes & Development 3, 1453-1461.

Logan, S. and Wensink, P.C. (1990) Ovarian follicle cell enhancers from the *Drosophila* yolk protein genes: different segments of one enhancer have different cell-type specifities that interact to give normal expression. Genes & Development , 613-623.

Lucchesi, J.C. (1978) Gene Dosage Compensation and the Evolution of Sex Chromosomes. Science 202, 711-716.

Mahowald, A.P. (1972) Ultrastructural Observations on Oogenesis in *Drosophila*. J.Morph. 137, 49-62.

Mahowald, A.P. and Strassheim, J.M. (1970) Intercellular migration of centrioles in the germarium of *Drosophila melanogaster*. An electron microscopic study. J. Cell Biol. 45, 306-320.

Mahowald, A.P. and Tiefert, M. (1970) Some structural changes in the *Drosophila* oocyte nucleus during a short period of RNA synthesis. Wilhelm Roux Arch. Entw. Org. 165, 8-25.

Mahowald, A.P. and Kambysellis, M.P. (1980) Oogenesis. In: The genetics and biology of *Drosophila*, ed. M. Ashburner and T.R.F. Wright. Academic Press, New York.

Vol 2d, 141-224.

Martin, C.H. and Meyerowitz, E.M. (1986) Characterisation of the boundaries between adjacent rapidly and slowly evolving genomic regions in *Drosophila*. P.N.A.S. USA 83, 8654-8658.

Martin, C.H. and Meyerowitz, E.M. (1988) Mosaic Evolution in the *Drosophila* Genome. BioEssays 9, 65-69.

Martin, C.H. Mayeda C.A and Meyerowitz EM (1988) Evolution of expression of the sgs-3 gene of Drosophila. J.Mol. Biol 201 273-287.

Martin, G.A., Bush, S.J., Meredith, G.D., Cardin, A.D., Blankenship, D.T., Mao, S.J.T., Rechtin, A.E., Woods, C.W., Racke, M.M., Schafer, M.P., Fitzgerald, M.C., Burke, D.M., Flanagan, M.A. and Jackson, R.L. (1988) Isolation and cDNA sequence of human post heparin plasma hepatic triglyceride lipase. J. Biol. Chem. 263, 10907-10914.

Martin, J. and Lee, B.T.O. (1984) A phylogenetic study of sex determiner location in a group of Australian *Chironomus* species (Diptera, Chironomidae). Chromosoma 90, 190-197.

Martinez-Cruzado, J.C., Swimmer, C., Fenerjian, M.G. and Kafatos, F.C. (1988) Evolution of the autosomal chorion locus in *Drosophila*. I. General organisation of the locus and sequence comparisons of the genes *s15* and *s19* in evolutionarily distant species. Genetics 119, 663-667.

Marsh, M. and Helenius, A. (1980) Adsorptive endocytosis of Semiliki Fortest virus. J. Mol. Biol. 142, 439-454.

Mayer, W.E., Jonker, M., Klein, D., Ivanyi, P., van, Seventer, G. and Klein, J. (1988) Nucleotide sequences of chimpanzee MHC class I alleles: evidence for trans-species mode of evolution. EMBO J. 7, 2765-2774.

McAlpine, J.F. (1970) First record of calypterata flies in the Mesozoic era. Entomol. 102, 342-346.

Medina, M. and Vallejo, C. (1989) The maternal origin of acid hydrolases in *Drosophila* and their relation with yolk protein degradation. Dev. Growth Diff. 31, 241-247.

Medina, M., Leon, P. and Vallejo, C. (1988) *Drosophila* cathepsin B-like preoteinase. A suggested role in yolk degradation, Arch. Biochem. Biophys. 263, 355-363.

Mehra, V., Sweetser, D. and Young, D. (1986) Efficient mapping of protein antigenic determinants. P.N.A.S. USA 83, 7013-7017.

Mellman, I. o Plutner, H (1984) Internalization and degradation of macrophage Fc receptors bound to polyvalent immune complexes. J. Cell. Biol. 86 1170-1177

Mellman, I., Fuchs, R. and Helenius, A. (1986) Acidification of the endocytic and exocytic pathways. Annu. Rev. Biochem. 55, 663-700.

Melton, D.A., Kreig, P.A., Rebagliat, M.R., Maniatis, T., Zinn, K. and Green, M.K. (1984) Efficient *in vitro* synthesis of biologically active RNA and RNA hybridisation probes from plasmids containing a bacteriophage SP6 promotor. Nucleic Acids Res. 12, 7035-7056.

Messing, J., Crea, R. and Seeburg, P.H. (1981) A system for shotgun DNA sequencing. Nucleic Acids Res. 9, 309-321.

Metz-Boutigue, M-H., Jolles, J., Mazurier, J., Schoentgen, F., Legrand, D., Spik, G., Montreuil, J. and Jolles, P. (1984) Human lactotransferrin: amino acid sequence and structural comparisons with other transferrins. Eur. J. Biochem. 145, 659-676.

Mikami, A., Imato, K., Tanabe, T., Nidome, T., Mori, Y., Takesshime, H., Narumiya, S. and Numa, S. (1989) Primary structure and functional expression of cardiac dihydropyridine-sensitive calcium channel. Nature 340, 230-233.

Minoo, P. and Postlethwait, J. (1985) Biosynthesis of *Drosophila* yolk polypeptides. Archives of Insect Biochemistry and Physiology 2, 7-27.

Mintzas, A. and Kambysellis, M.P. (1982) The yolk proteins of *Drosophila melanogaster*: isolation and

characterisation. Insect Biochem. 12, 25-33.

Mitsialis, S.A., Spoerel,N., Leviten, M. and Kafatos, F.C. (1987) A short 5'-flanking DNA region is sufficient for developmentally correct expression of moth chorion genes in *Drosophila*. P.N.A.S. USA 84, 7987-7991.

Miyata, Y, Nishida, E., Kayasu, S. Yahara, I. and Sakai H. (1989) Regulation by intracellular Ca^{2+} and cyclic AMP of the growth factor-induced ruffling membrane formation and stimulation of fluid-phase endocytosis and exocytosis. Exp. Cell. Res. 181 454-462.

Mori, L. and Perondini, A.L.P. (1980) Errors in the elimination of X chromosomes in *Scaria ocellaris*. Genetics 94, 663-673.

Morinaga, T., Fotedar, A., Singh, B., Wegmann, T. and Tamaoki, T. (1985) Isolation of cDNA clones encoding a T-cell receptor --chain from a beef insulin-specific hybridoma. P.N.A.S. USA 82, 8163-8167.

Moriyama, E.N. and Gojobori, T. (1989) Evolution of Nested Genes with Special Reference to Cuticle Proteins in *Drosophila melanogaster*. J.Mol.Evol. 28, 391-397.

Morrison, P.E. and Davis, D.M. (1964) Repeated ovarian cycles with ribonucleic acid in the diet of adult houseflies. Nature 201, 948-949.

Mueller, S.C. and Hubbard, A. (1986) Receptor-mediated endocytosis of asialoglycoproteins by rat hepatocytes: receptor positive and receptor negative endosomes. J. Cell Biol. 102, 932-942.

Mullock, B.M., Branch, W.J., van Schaik, M., GHilbert, L.K. and Luzio, J.P. (1989) Reconstruction of an endosome-lysosome interaction in a cell free system. J. Cell Biol. 108, 2093-2100.

Murray, N.E. (1983) Phage lambda and molecular cloning. In Lambda II. Ed. R.E. Hendrix, J.W. Roberts, F.W. Stahl and R.A. Weisberg. Cold Spring Harbour Laboratory. New York. pp 395-432.

Newfeld, J.S., Smoller, D.A. and Yedvobnick, B. (1991) Interspecific comparison of the unusually repetitive *Drosophila* locus *mastermind*. J. Mol. Evol. 32, 415-420.

Noah-Koller, C., Dhadialla, T.S and Raikhel A.S. (1989) Selective endocytosis of vitellogenin by the oocyte of the mosquito, Aedes aegypti; an in vitro study. Insect. Biochem 19 693-702

Norrander, J., Kempe, T. and Messing, J. (1983) Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene 26, 101-106.

Nothiger, R. and Steinmann-Zwicky, M. (1985) A Single Principle for Sex Determination in Insects. In Cold Spring Harbour Symposia on Quantitative Biology. Cold Spring Harbour Laboratory, New York. Volume L, 615-622.

Nothiger, R. and Steinmann-Zwicky, M. (1985) Sex determination in *Drosophila*. TIG July.

Ochman, H., Ajioka, J.W., Garza, D. and Hartl, D.L. (1989) Inverse polymerase chain reaction. In: PCR Technology, ed. H.A. Erlich. M. Stockton Press, London, pp 105-111.

Ohkubo, K., Ogata, S., Misumi, Y., Takami, N. and Ikehara, Y. (1991) Molecular cloning and characterisation of rat contrapsin-like protease inhibitor and related proteins. J. Biochem. 109, 212-250.
Okayama H and Berg P (1982) High efficiency cloning of full length cDNA. Mol Cell. Biol. 2 161

Opresko, L., Wiley, H.S. and Wallace, R.A. (1980) Differential postendocytotic compartmentation in *Xenopus* oocytes is mediated by a specifically bound ligand. Cell 22, 47-57.

Orci, L., Tagaya, M., Amherdt, M., Perrelet, A., Donaldson, J.G., Lippincott-Schwartz, J., Klausner, R.D. and Rothman, J.E. (1991) Brefeldin A, a drug that blocks secretion, prevents the assembly of non-clathrin-coated buds on golgi cisternae. Cell 64, 1183-1195.

Orr-Weaver, T.L. (1991) *Drosophila* chorion genes: cracking the eggshell's secrets. BioEssays 13, 97-105.

Osir, E.O. and Law, J.H. (1986) Studies on Binding and Uptake of Vitellogenin by Follicles of the Tobacco Hornworm, *Manduca sexta*. Arch. Insect Bioch. & Physiol. 3, 513-528.

Osir,E.O., Anderson, D.R., Grimes, W.J. and Law, J.H. (1986) Studies on the carbohydrate moiety of vitellogenin from the tobacco hornworm, *Manduca sexta*. Insect Biochem. 16, 471-478.

Palade, G.E. (1975) Intracellular aspects of the process of protein synthesis. Science 189, 347-358.

Pappas, C. and Fraenkel, G. (1977) Nutritional aspects of oogenis in the flies *Phormia regina* and *Sarcophaga bullata*. Physiological Zoology. 50, 237-246.

Parkhurst, S.M., Bopp, D. and Ish-Horowicz, D. (1990) X:A ratio, the primary sex-determining signal in *Drosophila*, is transduced by helix-loop-helix proteins. Cell 63, 1179-1191.

Pearse, B.M.F. (1975) Coated vesicles from pig brain-purification and biochemical characterisation. J. Mol. Biol. 97, 93-98.

Pearse, B.M.F. (1985) Assembly of the mannose-6-phosphate receptor into reconstituted clathin coats. EMBO J.4, 2457-2460.

Pearse, B.M.F. and Robinson. M.S. (1984) Purification and properties of 100kd proteins from coated vesicles and

their reconstitution with clathrin. EMBO J. 3, 1951-1957.

Pearce, B.M.F.and Robinson, M.S. (1990) Clathrin, adaptors, and sorting. Annu. Rev. Cell Biol. 6, 151-171.

Perlman, D. and Halvorson, H.O. (1983) A putative signal peptidase recognition site and sequence in eukaryotic and proyacyotic signal peptides. J. Mol. Biol. 167, 391-409.

Persson, B., Bergtsson-Olivecrona, G., Enerback, S., Olivecrona, T. and Jornvall, H. (1989) Structural features of lipoprotein lipase. Eur. J. Biochem. 179, 39-45.

Place, A.R., Anderson, S.M. and Sofer, W. (1986) Introns and domains coding regions in the dehydrogenase genes. In: Multidomain protein - structure and evolution, ed. D.G. Hardie and J.R. Coggins. Elsevier, Amsterdam. pp. 175-195.

Postlethwait, J.H. and Handler, A.M. (1978) Non-vitellogenic female sterile mutants and the regulation of vitellogenesis in *Drosophila melanogaster*. Dev. Biol. 67, 202-213.

Postlethwait, J.H. & Handler A.M (1979) The roles of juvenile hormone and ecdysterone during vitellogenesis in isolated abdomens of Drosophila melanogaster. J. Insect Physiol. 25 455-460.

Postlethwait, J.H. and Kaschnitz, R. (1978) The synthesis of *Drosophila melanogaster* vitellogenins *in vivo*, in culture, and in a cell-free translation system. FEBS Lett. 95, 247-251.

Postlethwait, J.H. and Jowett, T. (1980) Genetic analysis of the hormonally regulated yolk polypeptide genes in *D. melanogaster*. Cell 20, 671-678.

Postlethwait, J.H. and Shirk, P.D. (1981) Genetic and endocrine regulation of vitellogenesis in *Drosophila*. Am. Zool. 21, 687-700.

Postlethwait, J.H., Bownes, M. and Jowett, T. (1980) Sexual phenotype and vitellogenin synthesis in *Drosophila melanogaster*. Dev. Biol. 79, 378-387.

Ptashne, M. (1988) How eukaryotic transcriptional activators work. Nature 335, 683-689.

Raikhel, A.S. (1984) The Accumulative Pathway of Vitellogenin in the Mosquito Oocyte: A High-Resolution Immuno- and Cytochemical Study. J. Ult. Res. 87, 285-302.

Raikhel, A.S. (1986) Role of lysosomes in regulating of vitellogenin secretion in the mosquito fat body. Insect Physiol. 32, 597-604.

Raikhel, A.S. and Lea, A.O. (1985) Hormone-Mediated Formation of the Endocytic Complex in Mosquito Oocytes. Gen.& Compar. Endocrynol. 533, 422-433.

Raikhel, A.S. and Lea, A.O. (1986) Internalised proteins directed into accumulative compartments of mosquito oocytes by the specific ligand, vitellogenin. Tissue & Cell 18, 559-574.

Riley, M.A. (1989) Nucleotide sequence of the *Xdh* region in *Drosophila pseudoobscura* and an analysis of the evolution of synonymous codons. Molecular Biology and Evolution 6, 33-52.

Rina, M.D and Mintzas, A.C. (1987) Two vitellins-vitellogenins of the mediterranean fruit fly Ceratitis capitata : a comparative biochemical & immunological study. Comp. Biochem. Physiol. 86B: 801-808

Robinson, M.S. and Pearse, B.M.F. (1986) Immunofluorescence localisation of 100K coated vesicle proteins. J. Cell Biol. 102, 48-54.

Rognes, K. (1986) The systematic position of the genus *Helicobosca Bezzi* with a discussion of the monophyly of the calyptrate families Calliphoridae, Rhinophoridae, Sarcophagidae and Tachinidae (Diptera). Ent. Scand. 17 277-286.

Rohrkasten, A. and Ferenz, H-J. (1986) Properties of the vitellogenin receptor of isolated locust oocyte membranes. Int. J. Inv. Rep. Dev. 10, 133-142.

Roiha, H., Miller, J.R., Woods, L.C. and Glover,D.M. (1981) Arrangements and rearrangements of sequences flanking the two types of rDNA insertions in *D. melanogaster*. Nature 290, 749-753.

Roth, T.F. and Porter, K.R. (1964) Yolk protein uptake in mosquito *Aedes aegypti* L,. J. Cell Biol. 20, 313-332.

Rubacha, A., Tucker,M.A., De Valoir, T., Berlikoff, E.J. and Bechingham, K (1988) Genes with specific functions in the ovarian follicles of *Calliphora erythrocephala* (Diptera). Dev. Biol. 129, 449-463.

Ruoslahti, E. and Pierschbacher, M.D. (1986) Arg-Gly-Asp: a versitile cell recognition signal. Cell 44, 517-518.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-termination inhibitors. P.N.A.S. USA 74, 5463-5467.

Saunders, R.D.C. and Bownes, M. (1986) Sequence analysis of a yolk protein secretion mutant of *Drosophila melanogaster*. Mol. Gen. Genet. 205, 557-560.

Schelkel, H., Kejzlarova-Lepesant, J., Berreur, P., Maoreau, J., Scheller, K., Bregegere, F. and Lepesant J-A. (1985) Identification and molecular analysis of a

multigene family encoding calliphorin, the major larval serum protein of *Calliphora vicina*. EMBO J. 4, 2983-2990.

Schlossman,D.M., Schmid, S.L., Braell, W.A. and Rothman, J.E. (1984) An enzyme that removes clathrin coats: purification of an uncoating ATPasr. J. Cell Biol. 99, 723-733.

Schmid, S.L., Fuchs,R., Male, P. and Mellman, I (1988) Two distinct subpopulations of endosomes involved in membrane recycling and transport to the lysosomes. Cell52, 73-83.

Schultz, J. (1956) The relations of the heterochromatic chromosome region to the nucleic acids of the cell. Cold Spring Harbour Symp. Quant. 21, 307-328

Serafine, T., Stenbeck, G., Brecht, A., Lottspeich, F., Orci, L., Rothman, J.E. and Wieland, F.T. (1991) Identification of β-COP, a subunit of the coat of Golgi-derived (non-clathrin) coated vesicles, as a homologue of the clathrin-coated vesicle coat protein β-Adaptin. Nature

Sharp, P.M., Cowe, E., Higgins, D.G., Shield, D.C., Wolfe, K.H. and Wright, F. (1988) Codon usage patterns in *Escherichia coli, Bacillus subtilis, Saccharomyces*

cerevisiae, *Schizosaccharomyces pombe*, *Drosophila melanogaster* and *Homo sapiens*; a review of the considerations within-species diversity. Nucleic Acids Res. 16, 8207-8211.

Sharp, P.M. and Li, W-H. (1989) On the rate of DNA sequence evolution in *Drosophila*. J. Mol. Evol. 28, 398-402.

Shepherd, V.L. (1989) Intracellular pathways and mechanisns of sorting in receptor-mediated endocytosis. TiPS 10, 458-462.

Shirras, A.D. and Bownes, M. (1987) Separate DNA sequences are required for normal female and ecdysone-induced male expression of *Drosophila melanogaster* yolk protein 1. Mol. Gen. Genet. 210, 153-155.

Shirras, A. and Bownes, M. (1989) *cricket*: A locus regulating a number of adult functions of *Drosophila melanogaster*. P.N.A.S. USA 86, 4559-4563.

Shwartz, M.B., Kelly, T.J., Imberski, R.B. and Rubenstein, E.C. (1985) The effect of nutrition and methoprene treatment on ovarian ecdysteroid synthesis in *Drosophila melanogaster*. J. Insect Physiol. 31, 947-987.

Shwartz, M.B., Kelly, T.J., Woods, C. and Imberski, R.D. (1989) Ecdystroid fluctuations in adult *Drosophila melanogaster* caused by elimination of pupal reserves and synthesis by early vetellogenic ovarian follicles. Insect Biochem. 19, 243-249.

Singh, R.S. (1989) Population genetics and evolution of species related to *Drosophila melanogaster*. Annu. Rev. Genet. 23, 425-453.

Slee, R. and Bownes, M. (1990) Sex determination in *Drosophila melanogaster*. Quarterly Review of Biology 65, 175-204.

Smythe, E., Pypaert, M., Lucocq, J. and Warren, G. (1989) Formation of coated vvesicles from voated pits in broken A431 cells. J. Cell Biol. 108, 843-853.

Southern, E. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98, 503-517.

Spieth, J., Nettleton, M., Zucker-Aprison, E., Lea, K. and Blumenthal, T. (1991) Vitellogenin motifs conserved in nematodes and vertebrates. J. Mol. Evol. 32, 429-438.

Spoerel, N.A., Nguyen, H.T., Eickbush, T.H. and Kafatos, F.C. (1989) Gene evolution and regulation in the chorion

complex of *Bombyx mori*. J. Mol. Biol. 209, 1-19.

Srdic, Z., Beck, H. and Gloor, H. (1978) Yolk protein differences between species of *Drosophila*. Experientia 34, 1572-1574.

Stalker, J. (1956) Banded polythene chomosome in the ovarian nurse cells. J. Heredity 5, 259-264.

Stoorvogel, W., Geuze, H.J. and Stous, G.J. (1987) Sorting of endocytosed transferrin and asiloglycoprotein occurs immediately after internalisation in HepG2 cells. J. Cell Biol. 104, 1261-1268.

Stoorvogel, W., Strous, G.J., Geuze, H.J., Oorschot, V. and Schwartz, A.L. (1991) Late endosomes derive from early endosomed by maturation. Cell 65, 417-427.

Storrie, B. (1988) Assembly of lysosomes: perspectives from comparative molecular cel biology. Int. Rev. Cytology III, 53-105.

Storrie, B., Pool R.R. Jr., Sachdeva M, Mawrey, K.M and Oliver, C. (1984) Evidence for both polysosomal and lysosomal intermediates in endocytic pathways. J. Cell. Biol. 98 108-115.

Strangeways-Dixon, J. (1961) The relationship between nutrition, hormones and reproduction in the blowfly, *Calliphora erythrocephala* (Meig). J. Exp. Biol. 38, 225-235.

Sullivan, D.T., Atkinson, P.W and Starmer W.T. (1990) Molecular evolution of alcohol dehydrogenase genes in the genus Drosophila In: Evolutionary Biology Ed. Hecht M.K Wallace B. MacIntyre R.J. Plenum Press New York 24 107-147

Sutchiffe, J.G. (1978) Cold Spring Harbour Symposia for

Quantitative Biology 42, 4377-4390.

Swaunson, J.A. and Silverstein, S.C. (1988) Pinocytic flow through macrophages. In: Processing and presentation of antigens, ed. B. pernis, S. Silverstein and H. Vogel. Academic Press, London. pp 15-27.

Swimmer, C., Fenerjian, M.G., Martinez-Cruzado, J.C. and Kafatos, F.C. (1990) Evolution of the autosomal chorion cluster in *Drosophila* III. Comparison of the *s18* gene in evolutionarily distant species and interspecific controlof chorion gene amplification. J. Mol. Biol. 215, 225-235.

Szebenyi, D.M.E., Obendorf, S.K. and Moffat, K (1981) Structure of vitamin D dependent calcium binding protein from bovine intestine Nature 294 327-332.

Takada, Y. and Hemler, M.E. (1989) The primary structure of the VLA-2/collagen receptor α^2 subunit (Platelet GPIa): homology to other integrins and the presence of a possible collagen-binding domain. J. Cell. Biol. 109, 397-407.

Tamura, T., Kunert, C. and Posthethwait, J. (1985) Sex- and cell-specific regulation of yolk polypeptide genes introduced into *Drosophila* by *P*-element-mediated gene transfer. P.N.A.S.USA *82*, 7000-7004.

Tata, J.R. (1976) The expression of vitellogenin gene. Cell 9, 1-4.

Tautz, D. and Pfeifle, C. (1989) A non-radioactive *in situ* hybridisation method for the localisation of specific

RNAs in *Drosophila* embryos reveals translational control of the segmentation gene *hunchback*. Chromosoma 98, 81-85.

Tautz, D., Lehmann, R., Schnurch, H., Schuch, R., Serfert, E., Kienlin, A., Jones, K. and Jackle, H. (1987) Finger protein of novel structure encoded by *hunchback* a second member of the gap class of *Drosophila* segmentation genes. Nature 327, 383-389.

Tecce, M.F., Dozin, B., Magnuson, M.A. and Nikodenn, V.M. (1986) Transcriptional regulation by thyroid hormone of an mRNA homologous to a protease inhibitor Biochemistry 25, 5831-5834.

Thomas, P.S. (1980) Hybridisation of denatured RNA and small DNA fragments transferred to nitrocellulose. P.N.A.S. USA 77, 5201-5205.

Thompson, P.E. and Bowen, J.S. (1972) Interactions of differentiated primary sex factors in *Chironomus tentans*. Genetics 70, 491-493.

Throckmorton, L.H. (1975) The phylogeny, ecology and geography of *Drosophila*. In Handbook of genetics. Ed, R.C. King. Plenum Press, New York. Vol 3, 421-469.

Tolias, P.P., Konsolaki, M., Komitopoulou, K. and Kafatos, F.C. (1990) The chorion genes of the medfly *Ceratitis*

capitata. Devl. Biol 140, 105-112.

Treier, M., Pfeifle, C and Tautz, D (1989) Comparison of the gap segmentation gene *hunchback* between *Drosophila melanogaster* and *Drosophila virilis* reveals novel modes of evolutionary change. EMBO J.8, 1517-1525.

Tsuruhara, T., Koenig, J.H. and Ikeda, K. (1990) Synchronised endocytosis studied in the oocyte of a temperature-sensitive mutant of *Drosophila melanogaster*. Cell Tissue Res. 259, 199-297.

Ullerich, F-H. (1984) Analysis of sex determination in the monogenic blowfly *Chrysinya rufifacies* by pole cell transplantation. Mol. Gen. Genet. 193, 479-487.

Ungewickell, E and Branton, D (1981) Assembly units of clathrin coats. Nature 289 420-422

Urness, L.D. and Thummel, C.S. (1990) Molecular interactions within the ecdysone regulatory hierarchy: DNA binding properties of *Drosophila* ecdysone-inducible E74A protein. Cell 63, 47-61.

van Handel, E. and Lea, A.O. (1984) Vitellogenin synthesis in blood-fed *Aedes aegypti* in the absence of the head, thorax and ovaries. J. Insect Physiol. 30, 871-875.

Villarroya, A. and Juan, E. (1991) ADH and phylogenetic relationships of *Drosophila lebanonensis*

(Scaptodrosophila). J. Mol. Evol. 32, 421-428.

von Heijne, G. (1983) Patterns of amino acids near signal-sequence cleavage sites. Eur. J. Biochem. 133, 17-21.

van het Schip, F.D., Samallo, J., Bross, J., Ophuis, J., Mojet, M., Gruber, M., and AB, G. (1987) Nucleotide sequence of a chicken vitellogenin gene and derived amino acid sequence of the encoded yolk precursor protein. J. Mol. Biol. 196, 245-260.

Waheed, A., Gottschalk, S., Hille, A., Krentler, C. and Pohlman, R. (1988) Human lysosomal acid phosphatase is transported as a transmembrane protein to lysosomes in transfected baby hamster kidney cells. EMBO J. 7, 2351-2358.

Wahli, W. (1988) Evolution and expression of vitellogenin genes. Trends Gent. 4, 227-232.

Wall, D.A and Hubbard, A.C (1985) Receptor-mediated endocytosis of asialoglycoproteins by rat liver hepatocytes : Biochemical characterisations of the endosomal compartments. J. Cell. Biol. 101 2104-2112.

Wallace, R.A. (1985) In Developmental Biology, Vol 1. Ed. L.W. Browder, Plenum Publishing Corpotration. pp. 127-177.

Went, D.F. and Camenzind, R. (1980) Sex determination in the dipteran insect *Heteropeza pygmaea*. Genetica 52/53, 373-377.

West, M.A., Bretscher, M.S. and Watts, C. (1989) Distinct endocytotic pathways in epedermal growth factor-stimulated human carcinoma A431 cells. J. Cell Biol. 109, 2731-2739.

Wheeler, M.R. (1981) The Drosophilidae: A taxonomic overview. In The genetics and Biology of *Drosophila*. Ed. M. Ashburner *et al*. Academic Press, London. Vol 3e, 395-409.

Williams, J.L. and Bownes, M. (1986) Reduced stability of RNA coding for yolk polypeptide 3 in *Drosophila melanogaster* ovary. Eur. J. Biochem. 161, 95-101.

Williams, J.L., Saunders, R.D.C., Bownes,M. and Scott, A. (1987) Identification of a female-sterile mutation affecting yolk protein 2 in *Drosophila melanogaster*. Mol. Gen. Genet. 209, 360-365.

Wilson, A.C, Carlson, S.S and White, T.J (1977) Biochemical evolution. Ann. Rev. Biochem. 46 573-639

Wilson, T.G. (1982) A correlation between JH hormone deficiency and vitellogenic oocyte degradation in *Drosophila melanogaster*. Roux's Arch. Entw. Mech. Org. 191, 257-263.

Wilson, T.G., Lander, M.H. and Happ, G.M. (1983) Precocene I and II inhibition of vitellogenic oocyte development in *Drosophila melanogaster*. J. Insect

Physiol. 29, 249-254.

Wolfner, M.F. (1988) Sex-specific gene expression in somatic tissues of *Drosophila melanogaster*. Trends Genet. 4, 333-337.

Wolkoff, A.D., Klausner, R.D., Ashwell, G. and Harford, J. (1984) Intracellular segregation on asialoglycoproteins and their receptor: a prelysosomal event subsequent to dissociation of ligand-receptor complex. J. Cell Biol. 98, 375-381.

Yan, Y-L. and Postlethwait, J.H. (1990) Vitellogenesis in *Drosophila*: sequestration of a yolk polypeptide/invertase fusion protein into developing oocytes. Dev. Biol. 140, 281-290.

Yan, Y-L., Kunert, C. and Postlethwait, J.H. (1987) Sequence homologies among the three yolk polypeptide (Yp) genes in *Drosophila melanogaster*. Nucleic Acids Res. 15, 67-85.

Yanish-Perron,C., Vieira, J. and Messing, J. (1983) Nucleic Acids Res. 15, 67-85.

Zalokar, M. (1965) Etude de la formation de l'acide ribonucleique et des proteines chez les insects. Rev. Suisse Zool. 72, 241-261.

Zaremba, S. and Keen, J.H. (1983) Assembly polypeptides from coated vesicles mediate assembly of unique clathrin coats. J. Cell Biol. 97, 1339-1347.

Zhu, X.X., Gfeller, H. and Lanzrein, B. (1983) Ecdysteroids during oogenesis in the ovoviviparous cockroach *Nauphoeta cinerae*. J. Insect Physiol. 29, 225-235.

Zongza, V. and Dimitriadis, G.J. (1988) Vitellogenesis in the insect *Dacus oleae*. Isolation and characterisation of yolk protein mRNA. Insect Biochem. 18, 651-660.

Appendix I

DNA sequence of fragments containing *yp* gene sequences in clone B8I obtained from a -EMBL4 *Calliphora erythrocephala* genomic library (Rubacha *et al.*, 1988). The sequences underlined indicate the acceptor splicing end (GAT), the termination codon (TAA) and the polyadenylation site (AATAAA).

```
   1  AACCTGCAGG TCGACCTGCA GGTCAACGGA TCACTTGTCC CAGGCTGGTC
  51  GTGCTATTGA ACCCAGCTTT GTGCCTAAGG CTAGTGAAAT TCCTGCCTTC
 101  CTTATCACTC CCGACAACAA GAAAGTTAAC TTTAAATTAA GTGAATTGCC
 151  CAAAATTGCC AAGGAAGAAA AGAGCTTTGG CGATGAGGAA GTAACCGTCT
 201  ATATTACTGG TCTGCCACAA AAAACCGAAA CCGTAAAGAA AGCCACCCGC
 251  AAACTGGTGC AAGCCTATAT GCAGCGCTAC AATGGCCAGG CTCCCGAACG
 301  TAACAGTGTA CGCTATGAAG ACGATTCCAG TGAAAAGAGA AATCCCAGCT
 351  CTAGTGAAGA GGATGAGTGG AAGAGCAACT CCAACAAACC CTCTGGTAAT
 401  TTAGTGGTTA TTGAATTGGG CAATGTGTTA ACAAACATGA AAAACTACGT
 451  TTTCCTTGAT GTGGAAAAGA CTGGTATGGA AATTGGTGAT ATTTTGGTGC
 501  AGTGCACCGA TAAGGCTGAT GTACCCCAGG AAATCGTGCA TGTTATTGGC
 551  TCCAATATTG CTGCTCATGT GGCTGGTGCT GTTGGTCGTC AGTTTACCCG
 601  TGAAACTGGC CATCAGTTGC GTCGTATTAC CGGTTTGGAC CCCTCAAAAA
 651  TCTATGCTCA AACTCAGAAA TCTTTAAGAG GTTTAGCCCG CGGTGATGCT
 701  GAATTCGTTG ATGTCATCCA TACTTCCGCC TACGGTTGGG GTACTCCCAC
 751  CCGTTGCGGT GATGTAGATT TCTATCCCAA TGGTCCCAGT AGGGGCGTTC
 801  CTGGTGCTGA CAATGTTGTT GAGGCCTCCA TGCGTGCCAT CCGTTATTTC
 851  GCTGAATCTG TTGTACCCGG TAATGAACGC AATTTCCCTG CTGTTGGCGC
 901  CACCTCTCTC AAGCAGTACA AGGAACAAAA CGGCAATGGC AAACGTGTTT
 951  ATATGGGTAT CAATACCGAT TACGATGTTG AGGGTGACTT TATACTCCAG
1001  GTTAATGGCA AGAGTCCTTT CGGTCGCAGC ACTCCCGCTC AAAAGCAACA
1051  AAACTATCAT AATGTCCATA AGCCATGGAA AATGTCTTCT CGGGATTTTI
1101  AAGGGAACTT TAGCCAGCAG TTGGTAAAAT GAGTATATGG CTCATTTTGC
1151  TAAGATAATG AATTACTCTC CGATTTAGCA TAAACTTGTG ATTATTAAAT
1201  AGTTAAAAGA AATCAATTTT TGTATGTAGT AGACTAAGTA GAAACATTTT
1251  TTTTCTGTAA TAAACAACAA TTGCATTTTT ..
```

Appendix II

DNA sequence of fragments containing *yp* gene sequences (Figure 5.1 and 5.5) in clone A10B obtained from a -EMBL4 *Calliphora erythrocephala* genomic library (Rubacha *et al.*, 1988). The sequences underlined indicate TATA box (TATAAAA), Methionine start codon (ATG), 5'-donor splicing end (GGTAAGT), 3'-acceptor splicing end (GAT), stop codon (TAA) and the polyadenylation site (AATAAA). Nucleotide number 1 indicates the first nucleotide ofthe start codon.

```
-464  ..GTAAGCTA TTGTAAATAT GTATGTTGTT TACCTGATAA TTTAAAAATA
-414  CATTTTATTG ATTTTCTTAG ATTAAAAATA TTTCACGTCA CAGCATCATG
-364  GCTATAAATA TTGTCTCTTC AATGACTTTA AACATTTTCT ATCAATTAAT
-314  TTCTTATCAT CAAAAATATG TAATTAAACA AATTTCTAAA ATCAATGAAA
-264  TCAAAACAAA CAAAATAAAA GAAATTATTG TTATTTTAAT CTCAAAAATT
-214  GCTAAATCTT AATAAAAAAA ATGATGACTA TTATTATTTT AAAGAAATTT
-164  TCTAGCATCA TGAAGAAAAT TTTATTAATT TCAGCAAAAT ATTATTAAAA
-114  GTAAATAAAA ATGCAGAGTA TAAAAACCCC AGAATGGACC AACATCGACA
 -64  CAGTTCAGTT TTGTCTTCTG TTTAACGATA ACACTGTGCG GTACTGACTG

                    1
 -14  ATAGCTTGAA AGCATGAATC CTTTGAGAAT TGTTTGTGTG GCTGCTTTGT
  38  TATTGGCTGC AGGTTCAGCT AATGGAAATC TTTCTGGTCT TAACAAGTTG
  88  AGACCATCCC AATGGCTATC CTCATCTCAA CTTGACAAAC TCCCTCGATC
 138  GATGGAAATC AGTCTCCAAA AACTGGAAAG CATGTCGGTG GAAAAGGGTG
 188  CTGAATTAAT GCAAAAACTC TGTAAGTTTT AACAAGTAAT AAAAATGTAT
 238  TCGAGTAATT ATTATTAAAT TTGTTCTCCT AGATCACTTG TCTCAAATAA
 288  ATAACGATTT GAAACCCAGC TTTGTACCCA GCTCCAGTAA TGTGCCCTGC
 338  TACATTGTTA AGCCTAATGG CAAGAAGGTA TCCACCTCTT TGGACAAGTT
 388  GGCTTCAGCT TGCAAACAAC AACCCAACTT TGGTGAGGAG GAAGTAACCA
 438  TTCTAATCAC TGGATTACCC GCCACCACTG AAACTGTCAG AAAGGCCAAT
 488  CGTAAATTGA TTGATGCTTA CCTCCAACGT TACAGCACCA AGAGACAACA
 538  ACCCTCCAAG TTCGATTATT CTGGTGAGAA AATGGCACGC ACTTCCAGCG
 588  AAGAAGACTC TAACGAATGG CAAAATCAAC AAGCCTCCTC AGGCAATCTT
 638  GTGGTAAACT TAAATACATT TTTTAATTAT TTTCACAGCC TGGAATAAAT
 688  ATTTTTTCTA CCAATAGATC ATTGATTTGG GTAACGAATT GAACAGTTTC
 738  AAACGTTTTA GCCTTTTGGA TGTTGATGAA ACTGGTGCCA TGATTCGTTC
 788  CGCTATTGTT GAGATGACCG AGAAATGCGA TGTTGCCGAT GAAACCATCC
 838  ATGTTGTTGC TCAAGGTATT GCTGCTCATG TTGCTGGTGC CGCTGGTAAT
```

888 GAATTCACCC GCAAAACTGG TCGTCAATTG AGACGTATCA CTGCCTTGGA

938 TCCTTCCAAA ATTTTGGCCA AGAATCCTCA CTCTTTGACT GGTTTATCTC

988 GTGGTGATGC CGAATTTGTT GATGCCATCC ACACCAATGT CTATGGTATG

1038 GGTACCATTC AACGTGTTGG TGATATTGAT TTCTATCCTA ATGGTCCTTC

1088 GGCTGCCGTT CCTGGTGCTC AAAGCATTGT TGAGGCTAGC ATGCGTGCTA

1138 CCCGTTATTT CGCTGAATCT GTCCGCCCTG GCAATGAACG TAACTTCCCA

1188 GCTGTTGCTG CCAACTCATT GAAGGAGTAC AAGAACAATG ATGGTTTCGG

1238 CAAACGTGCT TATATGGGTA TTAATGTTGA TTATGATTTG GAAGGTGATT

1288 ACATTTTGGA AGTCAATTCG AAGGGTCCTT TCGGTAAGAG AGCCCCTGTT

1338 CAAAAACAGA ATAGCTACCA TGGAGTCCAA CAATTCCTTG AGAAAGAAGG

1388 ACAATAATTA AGATGTTTGG ACTTTGAAAT GAACTGATGA GATTTTTGTG

1438 TTTGTAATTT TTTAGTTATG TGAAAATAAA TGTTTTGCAA TTCTTTTAAA

1488 AAAGAATTAA TGAGTTTTAA TTAAGTTATT ACAGTTTTGC ATACCATAGC

1538 ACCAATATAG TTTGGAAGGG TTTTATTATT ATAACCTTCG CCCCAATTGT

1588 TAATACTATA TCTAGAATTC CATAAGACAT CATAAACAAA TTGACCTTGG

1638 ATCC

Appendix III

One letter symbols and three letter abbreviations for amino acids.

A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic acid
E	Glu	Glutamic acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
W	Trp	Tryptophan
Y	Tyr	Tyrosine