Analysis of cDNA Libraries and Expressed Sequence Tags of Retinal Pigment Epithelial Cells

Thesis submitted in accordance
with the requirements of
the University of Liverpool
for the degree of
Doctor in Philosophy
by

Luminita Paraoan

June, 1998

ACKNOWLEDGMENTS

"As physicists did 40 years ago, contemporary biologists have invaded a domain of human innocence. (...) Can (...) life be described in terms of molecules? For many, such description seems to diminish the beauty of nature. For others of us, the beauty and wonder of nature are nowhere more manifest than in the submicroscopic plan of life..."
(Weinberg, 1985)

The work described in this thesis was supported by the Wellcome Trust (Wellcome Prize studentship) and was carried out at the School of Biological Sciences and Unit of Ophthalmology, Department of Medicine, University of Liverpool. The analysis of DNA sequences performed during this study benefited greatly from the contact with the UK Human Genome Mapping Project Resource Centre.

I wish to thank Professor Ted Maden and Professor Ian Grierson for their inspiring advice and constant support throughout this investigation, for always taking time to answer my questions and listen to my ideas. Their association created an inspiring climate for pursuing research. I am very grateful for their generous appreciation and critical assessment of my experiments and ideas.

My thanks go also to:

Dr. Dean Bok, for making available to me the cDNA library of human foetal RPE cells and for the informative discussions about the "versatile" RPE cells;

Dr. Adrian Timmers, for informative discussions about various procedures for isolating RPE cells;

Dr. Jim Haslam, for firstly inviting me to Liverpool;

Dr. John Hughes and Karen Pugh, for being my friends;

Angela Roisin, for assistance with the automatic sequencing and Maureen Wilde, for advice with photography;

all my friends and colleagues in the School of Biological Sciences and Unit of Ophthalmology for answering my questions and contributing to an enjoyable atmosphere.

Last, but most, I thank my parents and my husband for their love and tireless understanding and support; and my sons, Vlad and Silviu, for stubbornly showing me that the real wonder of nature is to be found not at submicroscopic level!

ABSTRACT

The retinal pigment epithelium (RPE) has a pivotal and complex role in supporting neuroretinal function and, very likely, also in the development of the neural retina. Visual processing and the functional integrity of the adjacent photoreceptor cell layer depend critically on functions carried out by the RPE.

The aim of this study was to reveal details of the profile of genes expressed in the RPE cells, and to identify proteins which may be important for RPE functions. The transcriptional picture of the RPE cells was studied by characterization of ESTs of randomly chosen clones from cDNA libraries.

The first cDNA libraries were constructed from bovine material. These yielded limited success due, at least in part, to limitations of starting material, due both to the limited number of intact RPE cells that can be isolated from freshly dissected eyes and to restrictions imposed as a result of the BSE epidemic. Consequently, a commercially constructed cDNA library of cultured human foetal RPE cells (an aliquot of which was a gift from Dr. Dean Bok, UCLA, USA) was used as the principal source of clones for sequencing.

A total number of 272 individual clones (30 bovine clones and 242 human clones) from the different cDNA libraries of RPE cells were subjected to single-pass sequencing. The partial sequences of 203 cDNA clones (25 bovine, 178 human) were determined. The average length of the ESTs thus generated and used for comparison with database sequences was about 350 bp. Following removal of rRNA, mitochondrial DNA, chimaeric, adaptor multimers- and repeats-containing clones, 166 informative ESTs (12 bovine and 154 human) were defined.

Of the 154 informative human RPE ESTs, 126 sequences (81.8%) showed significant similarity with sequences in public databases. However, a substantial contribution to this proportion (51 ESTs, i.e. 33.1%) was from sequences matching "anonymous" database sequences, i.e. sequences reported by cDNA/genomic sequencing projects, with no identification/function assigned. More than half of the analyzed ESTs represent unknown genes, including the sequences identical to anonymous sequences in databases, sequences showing partial homology to known genes or ESTs, and sequences presenting no significant homology with other sequences in current databases.

When 75 putatively identified RPE ESTs were sorted according to the possible cellular role of the represented gene/protein, approximately a third of assigned ESTs were found to represent genes with involvement in gene/protein expression. In particular, clones coding for proteins associated with proteolytic activity and its regulation appeared highly represented. Interestingly, the cDNA data indicate that one of the abundant transcripts in the RPE is that of a specific cysteine proteinase inhibitor, cystatin C, thus suggesting that this protein might have an important role in supporting functions of the RPE.

ABBREVIATIONS

$A_{260(280)}$	absorbance at 260 (280) nm
Amp	ampicillin
bp	base pairs
BSA	bovine serum albumin
BSE	bovine spongiform encephalopathy
c^7 dGTP	7-deaza deoxyguanosine-5'-triphosphate
c^7 dITP	7-deaza deoxyinosine-5'-triphosphate
cDNA	complementary DNA
CRALBP	cellular retinaldehyde binding protein
CRBP	cellular retinol binding protein
CSF	cerebrospinal fluid
ddNTP	dideoxynucleoside-5'-triphosphate
DEPC	diethylpyrocarbonate
DMF	dimethylformamide
DMS	dimethylsulphate
DMSO	dimethylsulphoxide
dNTP	deoxynucleoside-5'-triphosphate
ds	double stranded
ECM	extracellular matrix
EDTA	ethylene diaminotetraacetate
EST	expressed sequence tag
FITC	fluorescein isothiocyanate
GFP	green fluorescent protein
HGMP	Human Genome Mapping Project
IPM	interphotoreceptor matrix
IPTG	isopropyl β-D-thiogalactopyranoside
IRBP	interphotoreceptor retinoid binding protein
ISH	*in situ* hybridization
kb	kilobase pairs
kDa	kilo Dalton
MCS	multiple cloning site
MOPS	3-(N-morpholino)propane-sulphonic acid

mRNA	messenger RNA
OD	optical density
OS	outer segment
PBS	phosphate-buffered saline
PIPES	1,4-piperazine-diethanesulphonic acid
PVP	polyvinylpyrrolidone
PVR	proliferative vitreoretinopathy
RBP	plasma retinol binding protein
rDNA	ribosomal DNA (ribosomal RNA gene)
RPE	retinal pigment epithelium
rpm	revolutions per minute
rRNA	ribosomal RNA
RT	reverse transcriptase
SDS	sodium dodecylsulphate
ss	single stranded
STS	sequence tag site
TEMED	N,N,N',N'-tetramethylethylenediamine
Tris	tris(hydroxymethyl)aminomethane
tRNA	transfer RNA
UV	ultraviolet
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

CONTENTS

CHAPTER 1

RETINAL PIGMENT EPITHELIAL CELLS AND VISION: A VERSATILE RELATIONSHIP

1.1. THE EYE

1.1.1. INTRODUCTION

Vision, a feature of the animal kingdom, depends on the function of the light detecting organ - the eye. Despite the variety of forms and structures in which eyes occur in different organisms, they share a structural component, in the form of visual pigments contained by the photoreceptor cells, and a functional characteristic, which is coupling of light absorption with visual excitation.

The visual pigments are a set of evolutionarily related proteins, opsins, which are covalently linked to a chromophore or "prosthetic" group that contains a conjugated polyene system (11-*cis*-retinaldehyde). Functionally, the eye is a highly specialized photosensitive organ developed to convert incident light energy into neuronal signals, thus enabling analysis of form, light intensity, and colour reflected from objects. The actual site of visual excitation is located in the photoreceptor cells.

It would go far beyond the scope of this thesis to consider all aspects of vision. Therefore, I will refer only to some aspects of vision in the vertebrate eye.

1.1.2. ASPECTS OF EYE ANATOMY AND HISTOLOGY

The basic design of the vertebrate eye (Figure 1.1) consists of three concentric layers: an external layer formed by the sclera and the cornea; a middle, vascular (uveal) layer consisting of the choroid, ciliary body, and iris; and an inner layer composed of an outer retinal pigment epithelium and neural retina. The terms outer, i.e. external, and inner, i.e. internal are

conventionally used, with respect to the structure of the eye. The photosensitive retina extends forward in the eyeball to the ora serrata. The retina communicates with the cerebrum through the optic nerve.

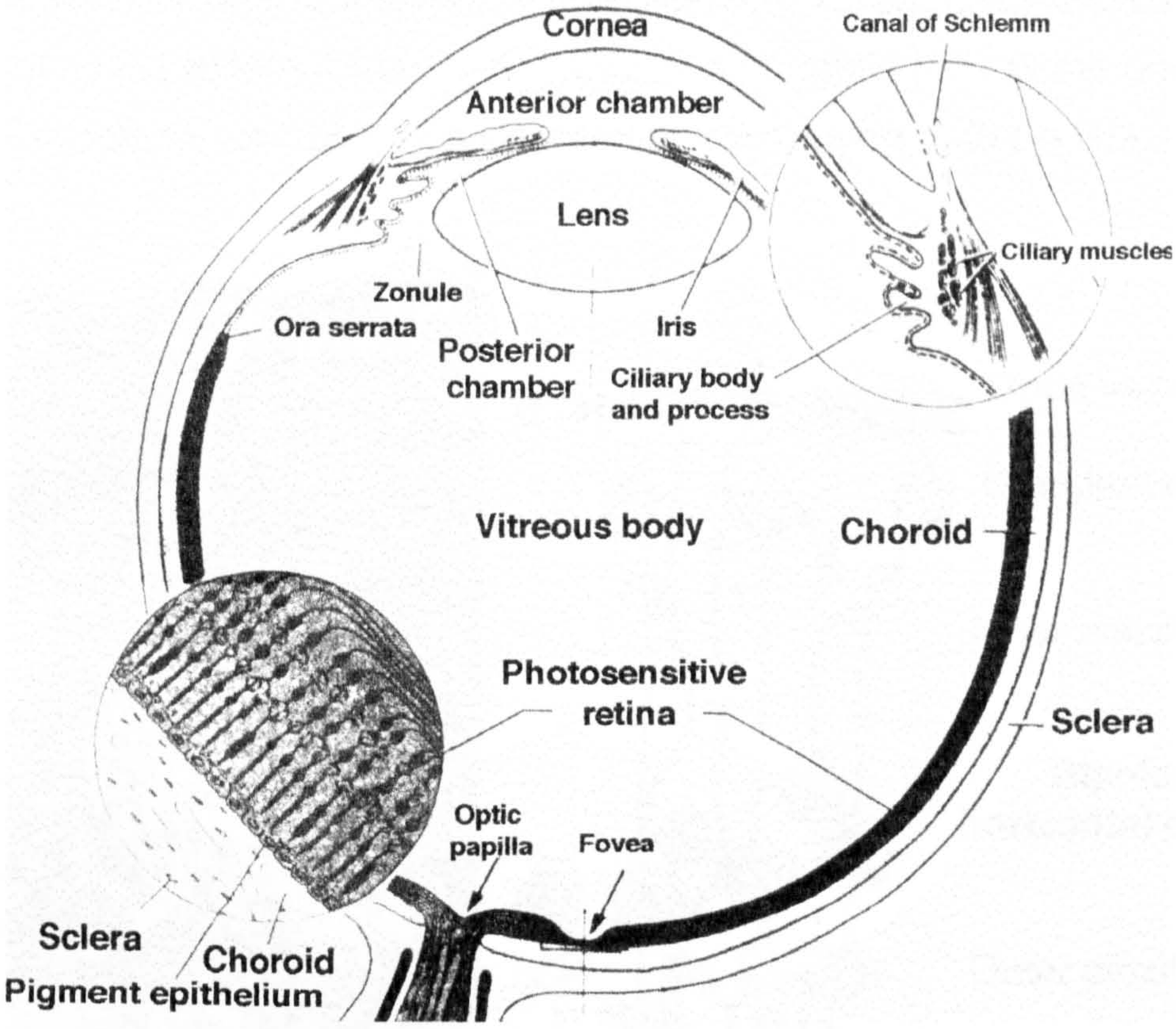


Figure 1.1 Schematic diagram of a cross section of the vertebrate eye, including enlarged diagrams of the retina and ciliary body. (Modified and reproduced from Junqueira *et al.*, 1992).

Three compartments are defined within the eye: the anterior chamber (between cornea, iris and lens); the posterior chamber (between iris, ciliary processes, zonular attachments and lens); and the vitreous space (behind lens and zonular attachments and surrounded by retina). The anterior and posterior chambers are filled with a fluid called the aqueous humor, while the vitreous space contains a gelatinous substance called the vitreous body.

Light entering the eye is refracted by the colourless and transparent media of the cornea and lens to form an image on the light sensitive retina.

1.1.2.1. The retina

The retina has a multistratified structure, composed of several cell types (Figure 1.2): ganglion cells, bipolar and horizontal cells, and photoreceptor cells. The plexiform layers contain synapses between the cells in the respective adjacent layers. The Muller cells extend their processes from the inner to the external limiting membrane and, analogous to the neuroglia's protective role for neurons, provide mechanical support to the retina.

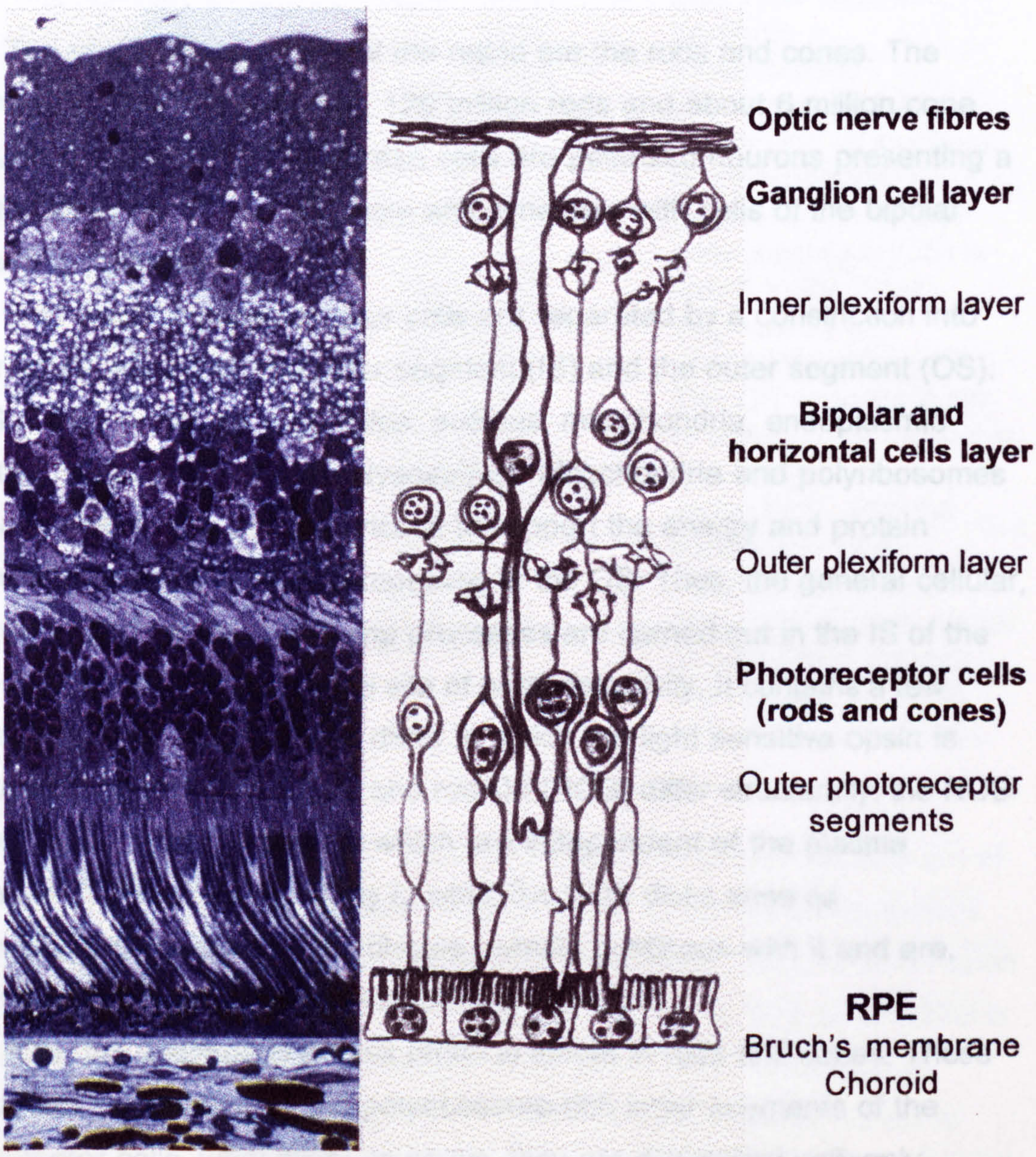


Figure 1.2. Light micrograph and schematic diagram of the neural retina and RPE in owl monkey (reproduced with permission from the collection of Professor Ian Grierson).

Light enters retina from the vitreous and traverses the cell layers towards the photoreceptor cells (rods and cones), where the actual photon absorption occurs. The neural stimulation generated by the incident light on rods and cones proceeds in the opposite direction.

The retinal pigment epithelial (RPE) cells border the photoreceptor cells distally, thus forming the outermost layer of the retina.

1.1.2.2. Anatomy of the photoreceptor cells

The photosensitive cells of the retina are the rods and cones. The human retina has approximately 120 million rods and about 6 million cone cells (Junqueira et al., 1992). These cells are polarized neurons presenting a photosensitive dendrite at one pole and synapses with cells of the bipolar layer at the other pole.

Both types of photoreceptor cells are separated by a constriction into two structural elements: the inner segment (IS) and the outer segment (OS). The IS contains the cell organelles: nucleus, mitochondria, endoplasmic reticulum, Golgi apparatus and lysosomes. Mitochondria and polyribosomes are accumulated near the constriction to support the energy and protein synthesis requirements of the processes in the OS. Thus, the general cellular, biosynthetic and energy-producing processes are carried out in the IS of the photoreceptor cell. The OS is the site of photosensitivity. It contains a few hundred flattened membranous disks in which the light sensitive opsin is embedded. The cone OS (COS) and rod OS (ROS) differ structurally: the ROS contain closed, separated disks which are independent of the plasma membrane, but are enveloped by it, while the COS disks arise as invaginations of the plasma membrane, remain continous with it and are, therefore, open to the extracellular space (Figure 1.3).

Also, the distribution of disk proteins differs in rods and cones. These proteins are synthesized in the polyribosome-rich inner segments of the photoreceptor cells, wherefrom, in cones, they are distributed uniformly throughout the COS invaginations; in rods, the corresponding proteins are incorporated into membranes, giving rise to flattened disks which gradually migrate towards the cell apex. These structures are progressively shed,

phagocytized, and digested by the cells of the pigment epithelium. This process is further discussed in section 1.2.1.

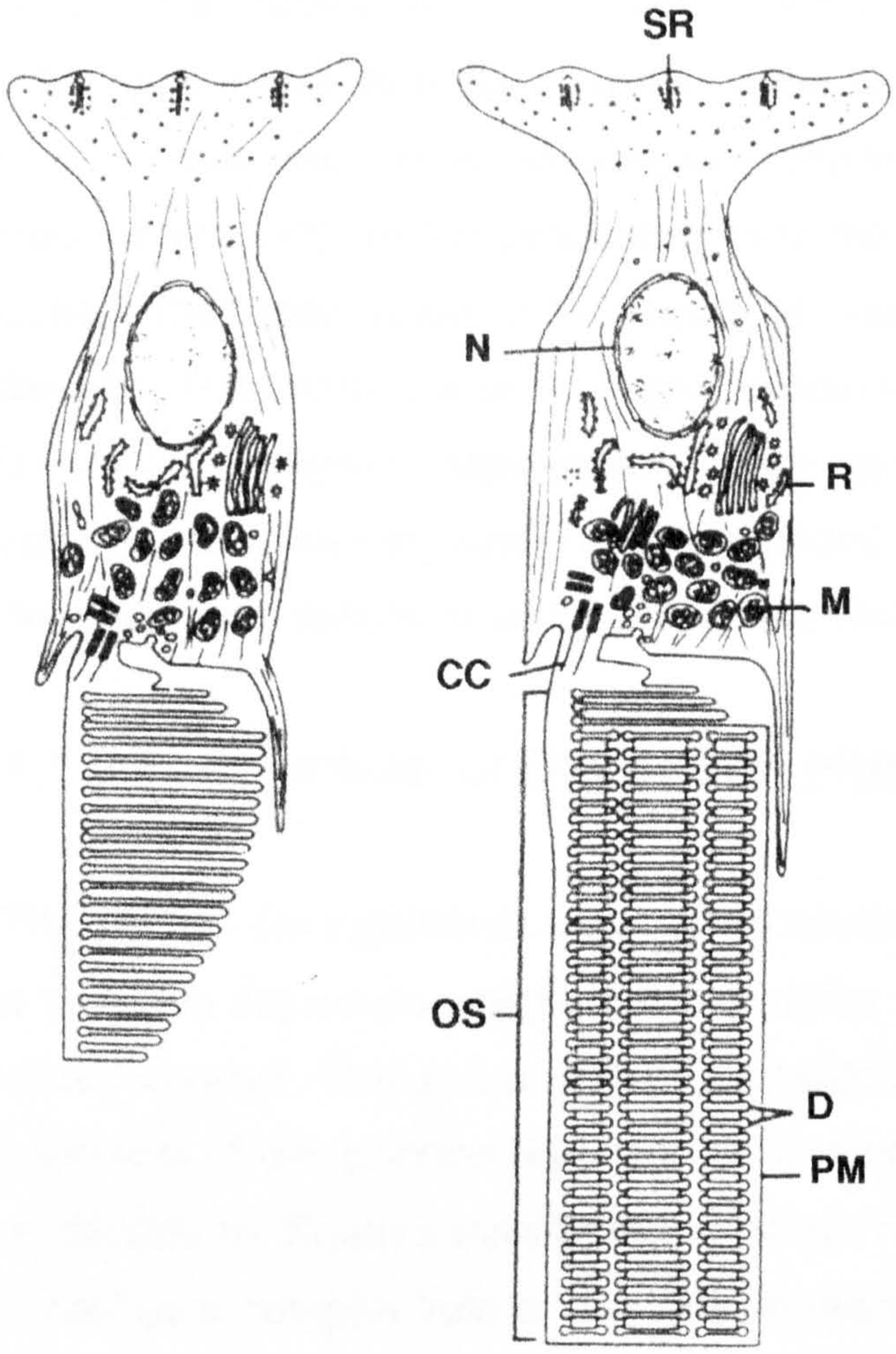


Figure 1.3. Schematic diagram of the photoreceptor cells (left, cone; right, rod). (reproduced after Bok, 1985) The inner segment presents a characteristic distribution of the cell organelles (N-nucleus; M-mitochondria; R-polyribosomes). The outer segment (OS) is linked to the inner segment by a connecting cilium (CC). The OS contains a pile of disks (D) in which the visual pigment is embedded. Plasma membrane (PM) surrounds the disks. The light-initiated signal is conducted through the synaptic terminal (SR-synaptic ribbons).

ROS contain rhodopsin as visual pigment. Cone photoreceptors are of three types with regard to the characteristics of the contained photopigment. The maximum sensitivity of these pigments is in the red, green, or blue region of the visible spectrum. Each of the four human visual pigment apoproteins is encoded by a separate gene (Nathans & Hogness, 1984; Nathans *et al.*, 1986). All visual pigments contain 11-*cis* retinal as chromophore, which is isomerized by light to all-*trans* retinal, thus activating the visual pigment and initiating the visual transduction cascade (reviewed by McNaughton, 1990). Rods and cones are complementary in their function as light receptors. Rods function in low levels of light, while cones have a much higher light signal

threshold. This is related to the fact that cones connect with fewer neurons and, as a consequence, provide better visual acuity.

Opsin, the apoprotein of rhodopsin, is a glycosylated polypeptide of 348 amino acids (Hargrave *et al.*, 1983), which has a seven-helix arrangement through the disk membrane; its C-terminal end is located at the cytoplasmic surface of the disk membrane, whereas the glycosylated N-terminal end is positioned intradiscally. In the process of light detection and signal transduction, rhodopsin undergoes photolysis, leading to opsin and all-*trans* retinaldehyde. Therefore, continous light detection requires, besides *de novo* biosynthesis of rhodopsin, regeneration of the apoprotein with the chromophoric group into an active visual pigment. The RPE plays a central role in this process, details of which are presented in section 1.2.2.3.

1.1.2.3. Structure of the retinal pigment epithelium

The RPE is the outermost layer of the retina, occupying a strategic position between the neurosensitive layers of the retina and the highly vascularized choroid. The apical side of the polarized RPE cells faces the outer segments of the photoreceptor cells. The basal membrane is separated from the choroid by Bruch's membrane to which RPE cells adhere firmly. This position implies a complex role of the RPE in retinal processes. For brevity, further description here of the general anatomical and morphological features of the RPE cells will provide the background information, while specific functions of the RPE will be discussed in more detail in section 1.2.

The RPE cells form a continuous monolayer of cuboidal cells which extends from the margins of the optic nerve head to the ora serrata (Figure 1.1), where it is continuous with the posterior pigment epithelium of the ciliary body (pars plana). The cells have a typical hexagonal epithelial morphology and measure about 10-14µm in height and 15-60µm in diameter, the size varying with age and location (Zinn & Benjamin-Henkind, 1979). Elaborate junctional complexes interconnect the RPE cells, which show a characteristic polarized structure (Figure 1.4.A). At the basal side, the cell membranes have numerous basal infoldings that increase the absorption interface with Bruch's membrane. This membrane is composed of a network of collagen fibres and

acts as a selective, permeable barrier between the choriocapillaris and the RPE (Figure 1.4.B). The cytoplasm of RPE cells presents a high density of membranous structures (endoplasmic reticulum and Golgi apparatus), mitochondria, lysosomes and phagosomes; melanosomes are concentrated in the apical region (Hogan *et al.*, 1971). The apical side of the cells extends into microvilli, which surround the outer segments of the photoreceptor cells.

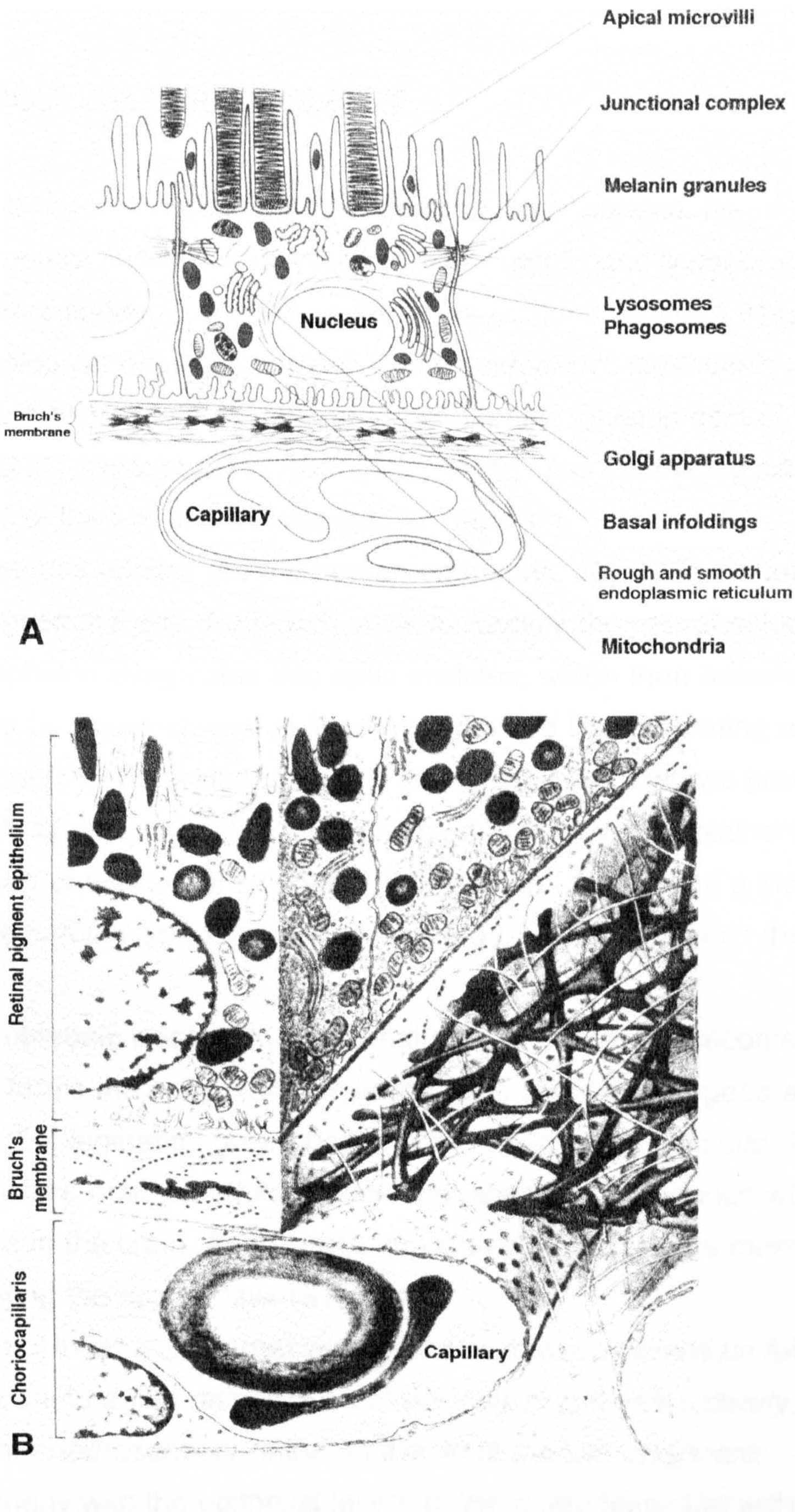


Figure 1.4. Schematic diagram of retinal pigment epithelium and its relation to photoreceptor outer segments, Bruch's membrane and choriocapillary layer.

A. Cross section (reproduced after Forrester *et al.*, 1996).

B. Three dimensional drawing (reproduced after Hogan *et al.*, 1971).

1.1.3. OCULAR EMBRYOLOGY

The eye is one of the first organs to develop in vertebrates. Its complicated structure requires a correspondingly complicated burst of cell division and differentiation. Complex series of interactions between cells of different embryological origins, neuroectoderm, endoderm, mesoderm and neural crest, are believed to contribute to the embryonic development of the eye (Forrester *et al.*, 1996; Mund & Rodrigues, 1979). The embryological origin of different parts of the eye are summarized in Table 1.1.

Optic primordia appear around human embryonic day 22 in the anterior region of the prosencephalic (forebrain) vesicle. Rapidly the neural ectoderm of the prosencephalon evaginates into optic vesicles, which then become partly enveloped by mesenchymal cells, mostly derived from migrating neural crest cells (Figure 1.5). Remarkable for this stage is the fact that two groups of proliferating cells of different ectodermic origins are closely apposed: a disk-shaped thickening of the neural ectoderm (future neural retina) and a localized thickening of the surface ectoderm, which later (day 27) develops into the lens placode.

During embryonic week 4, the proximal part of the vesicle becomes constricted and forms the optic stalk, while the optic vesicle undergoes a critical process of invagination giving rise to a double-layered optic cup. The resulting cavity of the optic cup (future subretinal space) is continuous with the ventricular space in the brain. As the optic cup grows, its two layers make contact, obliterating the space between them.

The external layer and internal layer of the optic cup differentiate fully by the end of week 8 into a thin retinal pigment epithelium and, respectively, a much thicker inner neurosensory retina. In the adult the two layers are continuous anteriorly with the epithelial layers of the ciliary body and with the posterior iris surface and adhere to each other across the subretinal space. This space can be re-established pathologically later in life in the form of different retinopathies.

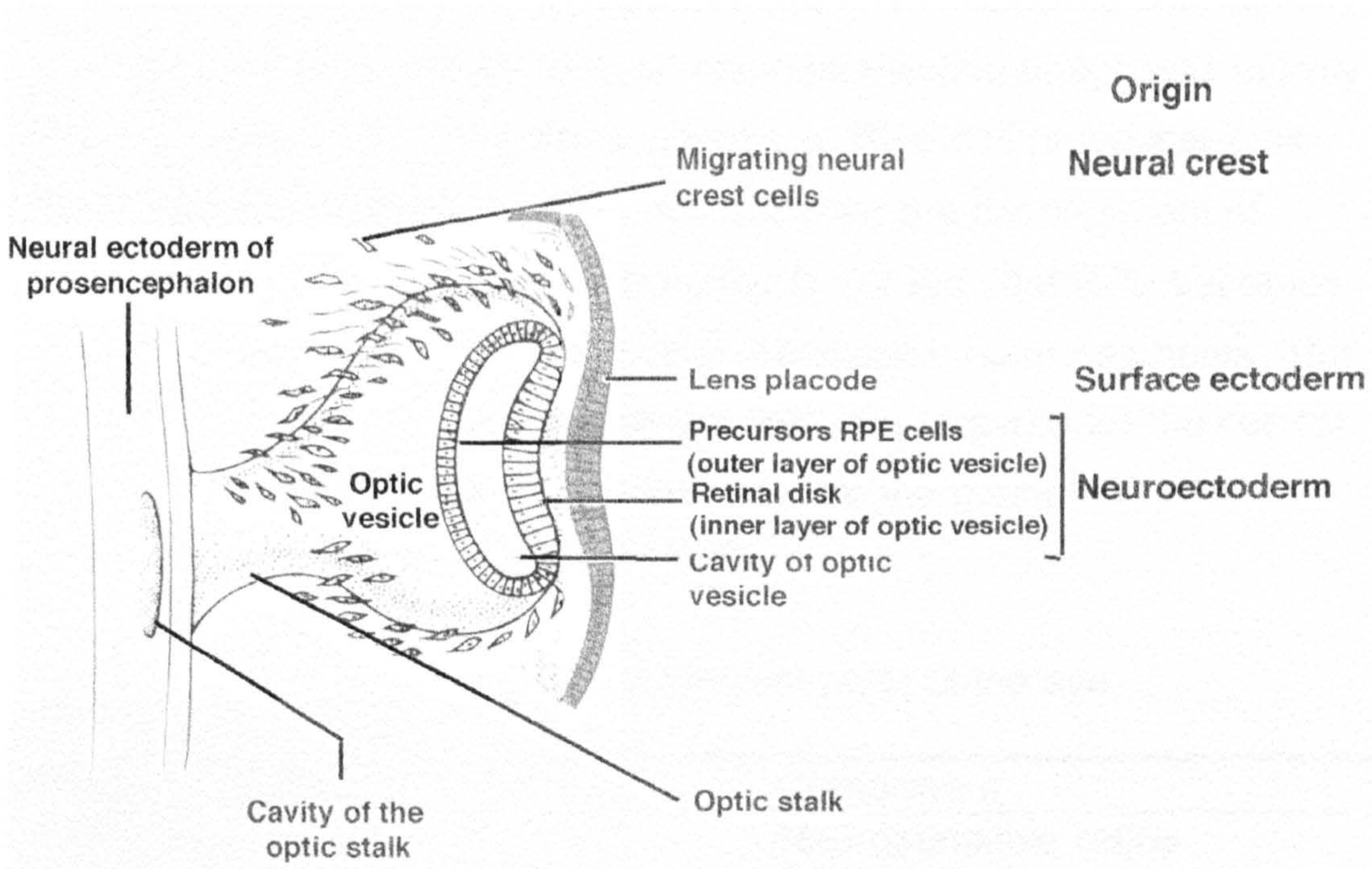


Figure 1.5. Schematic representation of the optic vesicle and stalk in a 27 day (5 mm) human embryo (reproduced after Forrester *et al.*, 1996).

A major landmark in eye development is represented by the appearance of melanin in the RPE at about 5-6 weeks of gestation (mature melanosomes are seen at 7 weeks of development); the process of pigmentation appears to be complete by 3 months. In contrast to skin, where melanin is only formed in melanocytes, melanosomes are distributed throughout every retinal pigment epithelial cell (Mund & Rodrigues, 1979). Also, RPE cells are unique among pigmented cells in being not derived from the neural crest lineage as other pigmented cells are, such as melanocytes of the skin, inner ear and choroid of the eye.

It is believed that the RPE induces surrounding mesenchymal tissue (mostly of neural crest origin) to organize into two layers (Hoar, 1982): the inner layer develops into the highly vascular choroid (homologous in its embryonic origin with the *pia mater* enfolding the brain) and the outer layer of the mesenchyme gives rise to sclera (homologous and continous with the *dura mater* around the optic nerve and brain). A series of (potential) roles have been ascribed to the RPE during embryonic development (Coulombre, 1979).

These are due, among others, to: an apparent transient inductive potency of the RPE for neuroretinal development; an early established phagocytic activity (prior to the appearance of rod outer segments or RPE cell processes); as well as a transport and trophic function, essential for the development of photoreceptor cell outer segments. Noteworthy is the fact that RPE becomes highly specialized before the formation of photoreceptor outer segments. The presence and the signalling mediated by the RPE are required for the correct morphogenesis of the neurosensitive retina and for the maintenance of its laminar structure (Raymond & Jackson, 1995).

Table 1.1. Embryological origin of different parts of the eye.

Origin	Component
Neural ectoderm	Neurosensitive retina RPE Epithelium of the ciliary body Posterior epithelium of the iris
Surface ectoderm	Lens Corneal epithelium
Neural crest and mesoderm (dense, outer layer of the mesenchyme)	Cornea (stroma) Sclera
Neural crest and mesoderm (vascular, inner layer of the mesenchyme)	Choroid Ciliary body Iris (stroma)

Thus, ontogenetically, the retina and RPE cells are closely related being almost simultaneous derivatives of the same basic embryonic layer as the brain (neuroectoderm). Yet, they evolve into two highly differentiated tissues which, nonetheless, maintain a complex functional interdependence. Through their neuroectodermic origin, RPE cells differ from other epithelial cells. In development, they even exhibit the capacity to transdifferentiate into neuroretinal cells or, with appropriate stimuli, into iris or lens epithelium (Buse *et al.*, 1993). It is not clear when the multipotential precursors of the RPE cells cease to be able to undergo the process of transdifferentiation (metaplasia) and become irreversibly commited to the pigment epithelial lineage (Coulombre, 1979). Moreover remarkably, although the RPE cells are

mitotically inactive in the adult eye, they do not lose the potential to divide and they are able to proliferate, either *in vivo*, in pathological conditions (such as proliferative vitreoretinopathies), or *in vitro*, in tissue cultures. Proliferation is part of complex behavioral changes that RPE can undergo and it is associated with phenotypic alterations and migratory properties (Grierson *et al.*, 1994).

The retina and RPE uphold an intimate relationship mediated by the contact of the RPE microvilli with the photoreceptor outer segments. The subretinal space between photoreceptors and RPE, which is the remnant of the embryonic cup vesicle, is kept of minimal size in the developed normal eye. When ocular mechanisms responsible for this function fail, it can re-open and lead to pathological conditions of retinal detachment (see 1.2.2.4). The interphotoreceptor matrix develops simultaneously with the differentiation of the outer segments and is likely to be involved in retinal attachment (Zauberman, 1979) and in the exchange of metabolites between RPE and retina.

1.2. THE RETINAL PIGMENT EPITHELIUM: BIOCHEMISTRY AND FUNCTIONS

The retinal pigment epithelium performs multiple physical, optical, biochemical and transport functions which make it an indispensable component of the vertebrate retina. The molecular characteristics of processes in which RPE plays an active role are far from completely unveiled. Nevertheless, agreement has been reached over the nature of the major functions performed by the RPE:

- daily phagocytosis of rod and, to a lesser extent, of cone outer segment fragments shed from their distal ends;

- providing a selectively permeable barrier between the choroid and neurosensory retina, which is a major part of the blood-retina barrier;

- maintaining adhesion of neurosensory retina;

- synthesis of interphotoreceptor matrix components (secretion and sulfation of glycosaminoglycans);

- transport of metabolites, to and from the retina, and uptake, processing and transport of vitamin A and its visual cycle intermediates;
- absorbtion of excess light and reduction of light scatter, thus enhancing visual acuity.

Most of these specialized activities begin early in life and must be continously carried out. Under normal circumstances the differentiated RPE does not renew itself by cell division. Therefore it must remain viable and fit to perform its functions throughout the life of the individual.

I will discuss in some detail the RPE functions for which the outcome of this study may have particular implications.

1.2.1. PHAGOCYTOSIS OF PHOTORECEPTOR OUTER SEGMENTS BY RPE

Phagocytosis of photoreceptor outer segments (OS) accounts for a large part of the metabolic activity of RPE (Cohen & Noell, 1965). An astonishingly large phagocytic load is handled by the RPE. It has been estimated that each cell ingests, per day, in frog about 600 disks, in rat 7500 disks (Bok & Young, 1979), in monkey in the range of 2000-4000 disks. These numbers follow from the facts that the apical processes of each RPE cell surround 25 to 45 rods, each of which assemble 80 to 90 new disks and shed a similar number of old disks per day (Young, 1971).

The ingestion and digestion capacities of the RPE are of vital importance for maintenance of the integrity of visual cells. The rods and cones exist in a "perpetual state of youth" (Bok, 1985), considering at least the disk membranes of their outer segments which are constantly being renewed throughout life (Young, 1967). New disks are formed at the base of the OS from protein synthesis taking place mostly in the inner segment (Young & Droz, 1968). This assembly of new disks is counteracted by the process of phagocytosis of groups of disks shed at the apex of the OS. Young and Bok (1969) demonstrated very elegantly that the rod OS tips, shed in a diurnal manner, are removed and degraded by the RPE cells in short bursts of phagocytic activity. Using electron microscopic autoradiographic techniques, they showed convincingly that frog ROS disks, containing newly synthesized

proteins labelled with ^{3}H-leucine and -phenylalanine, move down the OS, reach its apex and ultimately are detached and rapidly phagocytized by the RPE. The direct correlation between rod OS turnover and opsin turnover in frog retina has been established by Hall *et al.* (1969).

1.2.1.1. Circadian rhythm of disk shedding

The rhythmic (circadian) time scale of disk shedding was first described by LaVail (1976). In rats reared in cyclic light, the majority of the ROS disks are detached shortly after the light onset, when a marked increase in number of phagosomes is observed in the RPE (peak around 2 hours after beginning of light period). Subsequently, the phagosomes are rapidly degraded by the RPE. Similar results have been obtained in a study of frog retina (Basinger *et al.*, 1976). The event occurs even in continous darkness but at the time when animals are acustomed to light onset.

Rhythmic shedding of ROS appears to be a feature of all vertebrates examined thus far (reviewed by Bok, 1985). Shedding of cone OS disks occurs after beginning of the dark period in the majority of species examined, although exceptions have been described (Anderson & Fisher, 1975; Hogan *et al.*, 1974; O'Day & Young, 1978; Young, 1977a; Young, 1977b). The control mechanisms which govern the circadian clock are not fully understood.

1.2.1.2. Disk shedding mechanisms

Disk shedding is the process of detachment of packages of disks from the photoreceptor distal tip, which are subsequently phagocytized by the RPE. The major aspect of this process, for which an explanation at the molecular level is still to be uncovered, concerns the type of cell which is in control: photoreceptor or RPE cell. Evidence could be invoked in favour of initiation of shedding by either photoreceptor or RPE. The observations of Young (1971), in monkey photoreceptor-RPE interactions, suggested that groups of apical disks were separated from the rest by protrusion of ROS plasma membrane without an apparent involvement of the RPE. Experiments in *X. laevis* also support the active involvement of OS in this process (Matsumoto & Besharse,

1985; Cahill & Besharse, 1992). In retinas isolated prior to light onset, stained with fluorescent dyes which normally cannot permeate plasma membranes, the dye was observed at the point of incipient disk separation in OS. However, close association of photoreceptor and RPE cell has been described as required before the processs can take place (Williams & Fisher, 1987). Also, Besharse (1984) showed that activation of shedding is associated with formation of pseudopodia, which envelop ROS tips prior to disk detachment. These observations complemented those previously made by Spitznas and Hogan (1970). Moreover, cytochalasin, an actin inhibitor, inhibits not only the pseudopod formation (the pseudopod containing actin microfilaments), but also the incipient disk separation within the OS (Besharse & Iuvone, 1983). Therefore, in the light of all these and related observations, it is very likely that shedding involves active processes in both distal ends of OS and RPE cells.

1.2.1.3. Phagocytosis of ROS disks

Phagocytosis, the actual process of incorporation in the RPE cells of OS shed disk packages, can be divided in four distinct and consecutive stages: recognition and attachment, ingestion, phagosomal-lysosomal fusions, and degradation. Interruption of any of these steps leads to major disturbance of retina's health. A dramatic example of defective phagocytic process is found in the RCS strain of rats (Dowling & Sidman, 1962). In these dystrophic rats the phagocytic activity of RPE is impaired and, as OS elongate in the subretinal space, a layer of buildup membranes derived from OS disks accumulates as debris between the photoreceptor and RPE cell layers (Bok & Hall, 1971). This, in turn, causes an interruption of the flow of nutrients between the two cell layers and leads to photoreceptor death. Thus, the cause for the phagocytosis failure lies not in the OS, but in the incapacity of the RPE to complete the process. Chaitin and Hall (1983a) showed that recognition and attachment of the phagocytic load take place in these defective RPE cells, but a crucial step between these processes and ingestion is lacking. The nature of a possible transmembrane signal required for activation of ingestion is still to be discovered.

The recognition and attachment involve an interaction between ROS membrane and surface plasma membrane of the RPE cell (reviewed by Bok, 1988). The mechanisms of recognition and specific signaling are unresolved. Putative mediators are carbohydrate residues on the surface of the OS plasma membrane, which might serve as recognition sites for putative glycoprotein receptors on the apical surface of the RPE cell (Colley *et al.*, 1987). There is some evidence suggesting that mannose or sialic acid could be part of the phagocytic RPE receptor system (Mayerson & Hall, 1986; Tarnowski & McLaughlin, 1988). Though this matter remains still controversial because phagocytosis is not inhibited by free mannose, as might be expected for a mannose receptor-mediated process (Hall & Abrams, 1991).

Cultured RPE cells from several species including human are able to adopt two distinct phagocytic mechanisms, mediated by non-specific and specific receptors, respectively (Mayerson & Hall, 1986). *In vitro*, non-specific mechanisms are used for phagocytosis of latex beads and other inert particles, while specific mechanisms are used for selective phagocytosis of ROS in preference to latex beads (Philp & Nachmias, 1981). Given the specificity of receptors on RPE cell surface on which ROSs bind, the process of ROSs internalization by RPE is designated as phagocytosis (Hall & Abrams, 1987). However, the nature of the receptor involved in mediation of the specific mechanism of phagocytosis is unresolved.

Internalization of the incipient phagosomes, once triggered by a hypothetical transmembrane signalling process (Chaitin & Hall, 1983a), is driven by cytoskeletal components of the RPE cell. Subplasmalemmal actin plaques are formed at the ROS attachment site (Chaitin & Hall, 1983b) and, subsequently, internalized phagosomes are transported inside the RPE by cytoplasmic microtubules; inhibition of microtubule polymerization by colchicine leads to inhibition of phagosome transport and degradation (Besharse & Dunis, 1982).

Immediately after phagosomes are formed, the membranous material inside them begins to undergo chemical changes, as revealed by changed staining properties (Young & Bok, 1969). The phagosomes are gradually reduced in size, shifted towards the base of the the pigment epithelium and, finally, cleared from the cytoplasm of RPE cells.

The investigation of the sequence of events involved in the phagocytosis of OS, in particular the phagosomal-lysosomal fusion and digestion, highlighted interesting aspects of this process. Thus, the observations that cathepsin D, a lysosomal enzyme responsible for the degradation of opsin (Regan *et al.*, 1980) is found in the interphotoreceptor matrix (IPM) and that cultured RPE cells secrete this enzyme into the extracellular space (Adler & Martin, 1983) lead to the hypothesis that the digestion stage might occur or, at least, begin, before the OSs are ingested into the RPE. The cytochemical localization of another lysosomal enzyme (Mn-dependent pyrimidine 5'-nucleotidase) to the distal OSs before and during the early stages of ingestion (Irons, 1987) supported this hypothesis. However, Bosch *et al.* (1993) did not observe any early phagosomes containing cathepsin D, needed to sustain the hypothesis of cathepsin D entering the ROSs from the extracellular space and, consequently, a digestion of ROSs prior to ingestion stage. A small amount of cathepsin D (less than 5% of the total present in the cell) is contained in the apical processes of RPE cells. Despite the fact that this pool of enzyme is probably secreted into the IPM, its role in this extracellular compartment is not clear. Therefore, although the possibility of (partial) digestion beginning before the completion of ingestion stage can not be excluded, it is likely that the bulk of digestion occurs following full ingestion of phagosomes by RPE cells.

Moreover interestingly, the interaction of phagosomes with lysosomes is not that of a direct fusion. Details of this process have been examined by Bosch *et al.* (1993) by rapid freezing of rat RPE at different intervals with respect to light onset. Two physically independent processes develop simultaneously: phagosomes fuse with small lysosomes forming phagolysosomes, and lysosomes fuse with one another giving rise to large lysosomes (Figure 1.6). The phagolysosomes interact with large lysosomes without a complete fusion, but possibly through pore-like or bridge-like structures which allow the exchange of their content. Lysosomal fusion can be triggered by both light and circadian mechanisms. As the actual processing of the phagocytic load respects a circadian rhythm as well, it seems that putative circadian and light-triggered mechanisms control, independently and simultaneously, these two processes.

The degradation of the phagosome (described in detail morphologically by Bok & Young, 1979) involves, first, digestion of the plasma membrane surrounding the disk package, followed by progressive degradation of the membranous structure.

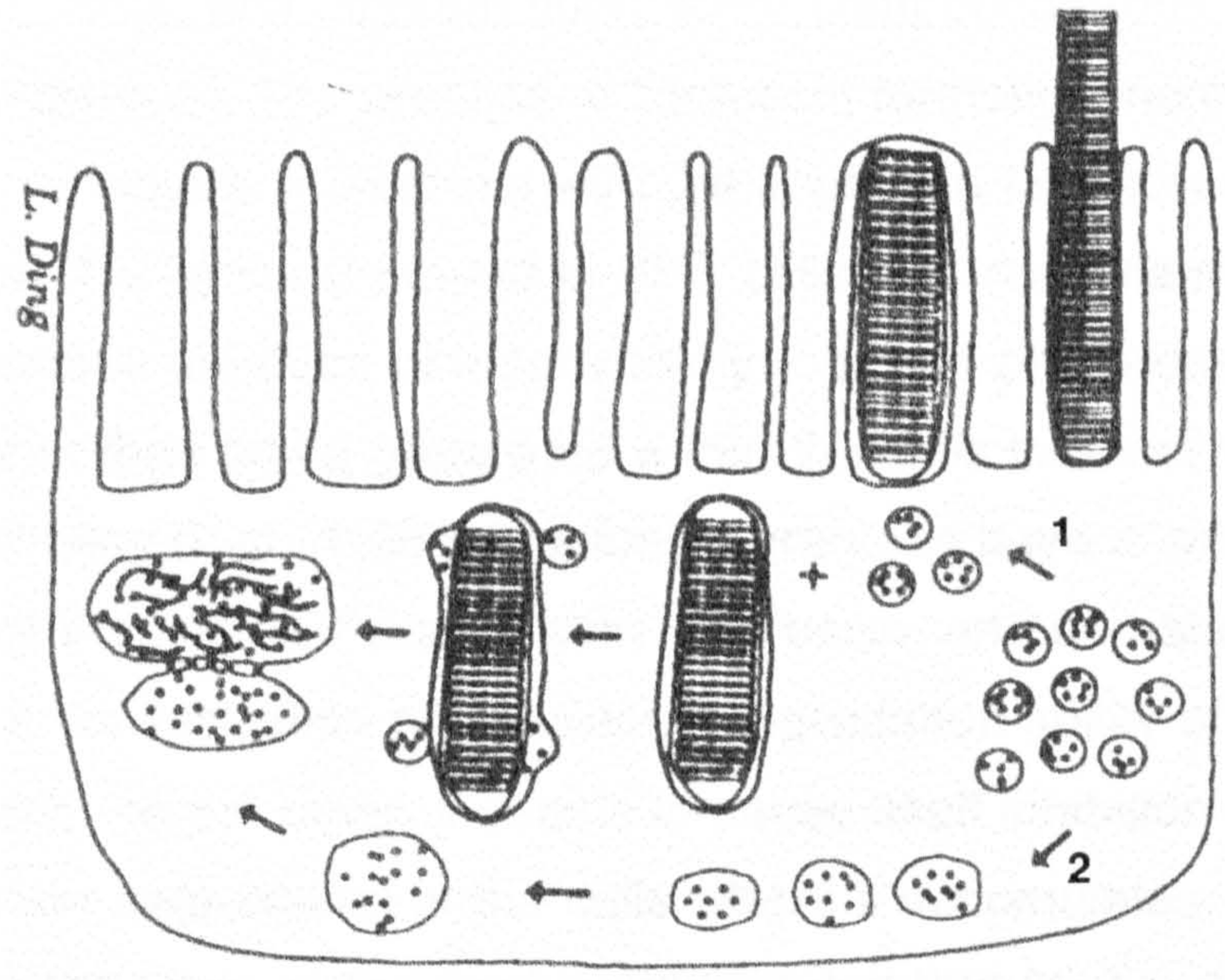


Figure 1.6. Model of phagolysosome interaction (reproduced after Bosch *et al.*, 1993). Digestion occurs in two steps, as observed from fixation and immunostaining of tissue from light-entrained animals at various times during the light-dark cycle. Initially, small lysosomes fuse with phagosomes, formed from ingested outer segment fragments (1). At the same time, small lysosomes fuse with one another to form larger lysosomes (2), which then interact with phagolysosomes, possibly by exchanging material through pore-like or bridge-like structures.

The phagolysosomal degradation leads to small peptides and amino acids which enter the regular RPE pool and are partially recycled to the photoreceptor cells, while another part of solubilized waste material is transported across the basal infoldings into the choriocapillaris. Free radicals and superoxides are generated during the degradation process and are inactivated by detoxifying enzymes such as superoxide dismutase and peroxidase (Hewitt & Adler, 1994).

Formation of lipofuscin

The residue left in the RPE cell assembles into lipofuscin granules, the so-called "age pigment", composed of a mixture of proteins (among which are opsin, peripherin and many proteins of unknown origin (Feeney-Burns & Eldred, 1985)), lipids, and different fluorescent compounds (one of which is a derivative of vitamin A). The process of lipofuscin formation increases with age and pathologic changes. In contrast with the lipofuscin that accumulates in other cell types, the lipofuscin found in RPE cells is formed mainly from the chemically modified residues of incompletely digested photoreceptor outer segments, rather than being generated within the cells themselves (for a review see Kennedy *et al.*, 1995). Lipofuscin granules have characteristic autofluorescent properties. Paradoxically, lipofuscin, on one hand, appears to be formed as a consequence of decreased degradation ability of the RPE (Katz, 1989), and, on the other, it seems to cause itself inhibition of the lysosomal protein degradation in the cells where it accumulates. One of the fluorophore components of the lipofuscin granules that inhibits lysosomal proteolysis has been identified as a quaternary nitrogen-containing cationic amphiphile, the bis-retinoid pyridinium salt A2-E (Eldred, 1995).

Aging of RPE and possibilities for therapy

The accumulation of lipofuscin during years of daily phagocytosis is associated with the aging RPE and, possibly, triggers events leading to age-related macular degeneration, the main cause of untreatable blindness in the elderly, in the Western world (Young, 1987). As the efficiency of the RPE decreases, its functions become compromised, causing the death of photoreceptor cells. A possible rescue might come from transplanted healthy RPE, as demonstrated in rat by Li & Turner (1988). Transplantation of RPE is possible due to its ability to become mitotically active when placed in culture or appropriate microenvironment. Development of methods suitable for this purpose is, therefore, of considerable current interest.

1.2.1.4. Hydrolytic activities of the RPE

The large digestive potential of the RPE cells arises from hydrolytic enzymes capable of degrading complex lipid-glycoprotein structures. Several such enzymes (Table 1.2) have been described having specific activities in RPE higher than in liver cells (Berman, 1979).

The lysosomal fraction of the RPE is more than three times as active as the equivalent fractions from other tissues in degrading the rhodopsin of ROS (Zimmerman *et al.*, 1983). It is thus conceivable that the lysosomes of the RPE contain hydrolytic enzymes specialized for the degradation of photoreceptor cell outer segments. Research efforts have been made during recent years to reveal the identity of enzymes and the regulatory factors to which this proteolytic activity is attributable. Important aspects remaining to be solved with regard to the degradation of the phagocytic load are: the enzymological details of the hydrolases that play key role(s), their localization, and the way in which they are delivered to the phagosomal particle.

Table 1.2. Hydrolytic enzymes in the RPE.

Enzyme	Reference
Acid phosphatase	Marshall, 1970; Hayasaka, 1983; Seyfried-Williams & McLaughlin, 1984;
α-fucosidase α-mannosidase N-acetyl-β-glucosaminidase β-galactosidase	Hayasaka & Shiono, 1982; Shiono *et al.*, 1983
Acid lipases Phospholipases A, B, C	Berman, 1971
Aspartic and cysteine proteinases (of the cathepsin D and cathepsin B type)	Hayasaka *et al.*, 1975a; Hayasaka *et al.*, 1975b; Zimmerman *et al.*, 1983; Tsung *et al.*, 1984

From among many proteolytic enzymes in the RPE, extensive studies have recently focused on cathepsin D, a lysosomal aspartic proteinase. This enzyme seems likely to be one of the major enzymes involved in the lysosomal digestion of photoreceptor OSs. The following observations are relevant. Expression of both 52 kDA procathepsin D and 34 kDa active cathepsin D was observed in cultured RPE cells by Rakoczy *et al.* (1997a). In the same study, the decrease in cathepsin D activity, caused by transfection with a construct containing the full-length sequence of the translated region of cathepsin D in antisense orientation, led to a slower processing of ROS. However, it did not eliminate the phagocytosing ability of the cells (Rakoczy *et al.*, 1997b). Moreover, the upregulation of cathepsin D activity, caused by the transfection with the construct containing the same sequence in sense orientation, did not determine acceleration in lysosomal digestion of ROS (Lai *et al.*, 1997). These two facts suggest that cathepsin D-mediated proteolysis is preceded by other digestive steps.

Changes in the efficiency of ROS processing determine accumulation of photoreceptor-derived, autofluorescent debris and lipofuscin. Rakoczy *et al.* (1996) indicated a possible correlation between the accumulation of such debris in cultured RPE cells and the presence of a multimer (maybe partially processed) form of cathepsin D.

Freshly isolated RPE cells from the central area of the retina appear to have higher cathepsin D activity compared with cells from peripheral retina (Boulton *et al.*, 1994; Burke & Twining, 1988). In human RPE cells cultured on nitrocellulose filter inserts, about 10% of the total cell-associated cathepsin D activity accumulates extracellularly, both in the apical and basal extracellular fluid (Wilcox, 1988). Interestingly, the distribution of the extracellular activity shows a marked age dependence: 70/30 to 10/90, basal/apical, for donors aged 9-93 years. In contrast, the distribution of extracellular activity of RPE-produced N-acetyl-β-glucosaminidase is constant regardless of the age of donor. This variation in the distribution of cathepsin D activity and the observation that cathepsin D activity increases with age (Boulton *et al.*, 1994) might suggest the implication of this enzyme, at least in culture conditions, in other proceses, e.g. remodelling of the extracellular matrix.

Evidence for the implication of other proteolytic enzymes in the degradation process caried out by RPE cells was provided by the following studies. Intravitreal injection of the cysteine proteinase inhibitor leupeptin caused a rapid accumulation of lipofuscin-like, autofluorescent bodies in the RPE of albino rats (Katz & Shanker, 1989). The accumulation of autofluorescent debris was also demonstrated in ROS-challenged RPE cell cultures in which the activity of cysteine proteinases was inhibited with leupeptin (Rakoczy *et al.*, 1994). A similar ROS-derived increase in autofluorescence was determined by antisense oligonucleotide-mediated inhibition of cathepsin S (cysteine proteinase). Thus, cysteine proteinases could have a key role, either directly in the ROS digestion process or in the activation of other lysosomal enzymes, e.g. cathepsin D, involved in this process.

1.2.2. RPE AS A SELECTIVELY PERMEABLE BARRIER AND MODULATOR OF ITS ENVIRONMENT

The strategic position occupied by the RPE within the eye, combined with the anatomy of these cells enables it to play a crucial role in protecting and nourishing the neural retina. RPE achieves this role by performing two inter-related functions: (i) it is part of the blood-retinal barrier and, hence, regulates the passage of molecules into and from the retina; and (ii) it modulates the composition of the subretinal fluid and extracellular matrix, including the interphotoreceptor matrix. These specialized functions and interactions with both the neural retina and the choriocapillaris are made possible by the structural characteristics and functional polarity of the RPE cells.

1.2.2.1. The role of RPE in the blood-retinal barrier

The blood-retinal barrier enables a mechanism of homeostatic control that prevents free diffusion of solutes from blood into the retinal space and regulates the supply and disposal of the necessary metabolites and catabolites (Steinberg & Miller, 1979). It has two components, an inner and an

outer barrier. The endothelial cells of the retinal vasculature are responsible for maintaining the inner barrier, while the outer barrier is enabled by the presence of tight junctions between RPE cells. The junctional complex in which RPE cells are implicated also facilitates the selective transport of nutrients and metabolites in the retina-RPE-choroid system.

In the apical zone, the cell junctions of the lateral RPE cell membranes are formed by continuous zonulae occludentes (impermeable, tight junctions) underneath a belt of zonulae adherentes (permeable, intermediate junctions). There are also desmosomes and gap junctions (Zinn & Benjamin-Henkind, 1979). Thus, the structure of the whole junctional complex of the RPE cells is such that it allows intercellular communication within the RPE (via gap junctions) while controlling diffusion of substances between the choriocapillaris and the subretinal space. The intercellular spaces in the RPE being firmly closed, so preventing the diffusion of even small molecules, most molecular movements must take place through the selective transcellular route. Therefore, the barrier, like the blood-brain barrier, is responsible for homeostasis of the neuroretina, sustaining its specific physiological environment and protecting it from toxic compounds (Cunha-Vaz, 1979). Damage to the junctional complex, as observed in the RCS rats, leads to increased permeability of the blood-retinal barrier and the redistribution of membrane components and gap junctions (Caldwell & McLaughlin, 1983; Hewitt & Adler, 1994).

1.2.2.2. Specific polarity of some membrane proteins in RPE

Besides their role in formation of the blood-retinal barrier, tight junctions are conceivably important for the maintenance of differentiation of the apical and basal RPE cell membranes. As a consequence of the fact that the apical and basal regions of this epithelial sheet are sealed off, the RPE plasma membrane presents a polarized organization, with molecules specifically distributed to apical and basal domains. The peculiar apical distribution of some membrane proteins may be related to the fact that the apical surface of RPE cells presents a unique feature among epithelia as it does not open into

a lumen, but establishes an intimate contact with the interphotoreceptor matrix (IPM) and photoreceptor OSs.

Three proteins that are basolateral in other transporting epithelia have been found to occur on the apical side of RPE plasma membrane *in vivo*. These proteins are: the Na^+,K^+-ATPase (discussed below in relation to movement of ions and water across RPE); CE-9/OX-47, a member of the immunoglobulin superfamily, recognized by the monoclonal antibody RET-PE2 (Finnemann *et al.*, 1997); and a transmembrane form of the neural adhesion cell molecule N-CAM (Gundersen *et al.*, 1993).

The recruitment of RET-PE2 antigen to the apical surface of RPE cells is a gradual process, that progresses simultaneously with the maturation of photoreceptor OS (in rats is fully established by postnatal day 14) and is maintained in the adult life. The antigen is relocalized to basolateral domain in RPE cell cultures (primary and rat cell line RPE-J). RET-PE2 antigen is abundant in the RPE, but its function is not yet known. It has been suggested that it may function as a receptor for extracellular matrix (Finnemann *et al.*, 1997).

Another molecule that is apically polarized in the RPE is one of the isoforms of the calcium-independent neural adhesion molecule, N-CAM (Gundersen *et al.*, 1993). This molecule plays key roles in the development of nervous system and in the organization of a multilayered retina (Hoffman *et al.*, 1986). Its specific apical distribution in the RPE, which is induced during eye development, is maintained during postnatal development. In contrast to this, in cultured RPE cells, N-CAM is distributed basolaterally, as it is in transfected epithelial MDCK cells. Since N-CAM is expressed in both RPE and photoreceptor OSs, its apical localization in the RPE, conditioned by the interaction with the neural retina, may indicate an important role for this molecule in the adhesion between neural retina and RPE.

The "reversed" apical polarity of some proteins, such as N-CAM, in the RPE cells *in vivo* appears to depend on the establishment of adult interactions between the RPE and the neural retina and/or the interphotoreceptor matrix. Possible mechanisms for these interactions include direct contact between these or alterations in the intracellular sorting patterns for these proteins in the RPE cells.

1.2.2.3. Transport mediated by RPE

The function of the RPE as a transporting epithelium is mediated by its elaborate basal plasma membrane infoldings and apical microvilli. Although the transport of various metabolites facilitated by the RPE is bidirectional, the bulk of it consists of movement from the retina to the choriocapillaris. Among the substances that have been shown to be actively transported by the RPE are: vitamin A (for a review, see Bok, 1993), amino acids (Miller & Steinberg, 1976; Pautler & Tengerdy, 1986; Sellner, 1986), ions and water (Steinberg & Miller, 1979). The studies concerned with the movement of solutes across RPE unveiled peculiar characteristics of the mechanisms by which this tissue controls the concentrations of solutes in the subretinal fluid. Here, some of these mechanisms will be briefly discussed.

Movement of organic solutes

The movement of amino acids across RPE determines a net flux in the direction retina to choroid, given by the greater unidirectional clearance of each such organic solute in the apical to basal than basal to apical direction (Sellner, 1986). This seems to contradict the presumed function of the RPE as a supplier of membrane precursors to the retina. Yet, retinal capillaries supply the inner retina from the inner nuclear layer to the optic nerve fiber layer, thus leaving the photoreceptors entirely dependent on extracellular fluid secreted by the RPE for the supply of amino acids and lipid components for continuous disks synthesis. Interestingly, RPE does not fulfil this function by "simply" transporting the required solutes from the blood to the retina, but controls the concentration of a particular substrate in the subretinal fluid (thus, probably preventing toxic levels from accumulating) and, consequently, establishes and maintains a concentration gradient for that substrate across the tissue. For example, the substantial gradient across the tissue for leucine (88.4 μM on the choroidal side vs. 11.8 μM on the retinal side) tends to drive a net flow of the amino acid toward retina (Sellner, 1986).

The gradient between blood and subretinal fluid is analogous to that between blood and cerebrospinal fluid (CSF) for which the blood-brain barrier is responsible. The concentration of leucine is approximately eight-fold lower in

CSF than that in plasma (Kruse *et al.*, 1985). In the eye, as the consumption of amino acids by the retina is not sufficient to account for the decline in concentration from blood to subretinal fluid, it must be concluded that the RPE, and not the retina, is responsible for continously controlling their subretinal concentration. It accomplishes this very efficiently by determining the transepithelial fluxes.

Transport of retinoids

The transport and processing of retinoids is one of the highly specialized functions of the RPE. Following the conversion by photoisomerization of 11-*cis* retinal to all-*trans* retinal and its release from opsin, the liberated all-*trans* retinal is reduced to all-*trans* retinol while still within the photoreceptor cell. The retinol is released into the interphotoreceptor matrix (IPM) and thence into the RPE. The purpose of the RPE-driven retinoid cycle is to regenerate 11-*cis* retinal and to redirect it to photoreceptors. 11-*cis* retinal is regenerated from all-*trans* retinol in the RPE (Bernstein *et al.*, 1987). The main events in this process, as known to date, are summarized in Figure 1.7. Retinol (vitamin A) is also delivered from the blood to the retina across RPE. The diet is the ultimate source of vitamin A, since it is not synthesized by vertebrate tissues. Although not all details of the molecular mechanisms underlying the visual cycle of retinol and its derivatives are resolved, important "players" have been described.

Retinol uptake occurs at both basolateral and apical surfaces of RPE by separate receptor-mediated processes. Plasma retinol binding protein (RBP), secreted by the liver (Kanai *et al.*, 1968), forms a complex with retinol (lipid-soluble molecule), thus increasing its solubility. RBP circulates in blood complexed with transthyretin, a thyroxin-binding protein, and mediates retinol delivery across RPE basolateral membranes, without itself entering the cell. Receptors for RBP are present on the basolateral surface of the RPE cells (Bok, 1990; Chen & Heller, 1977).

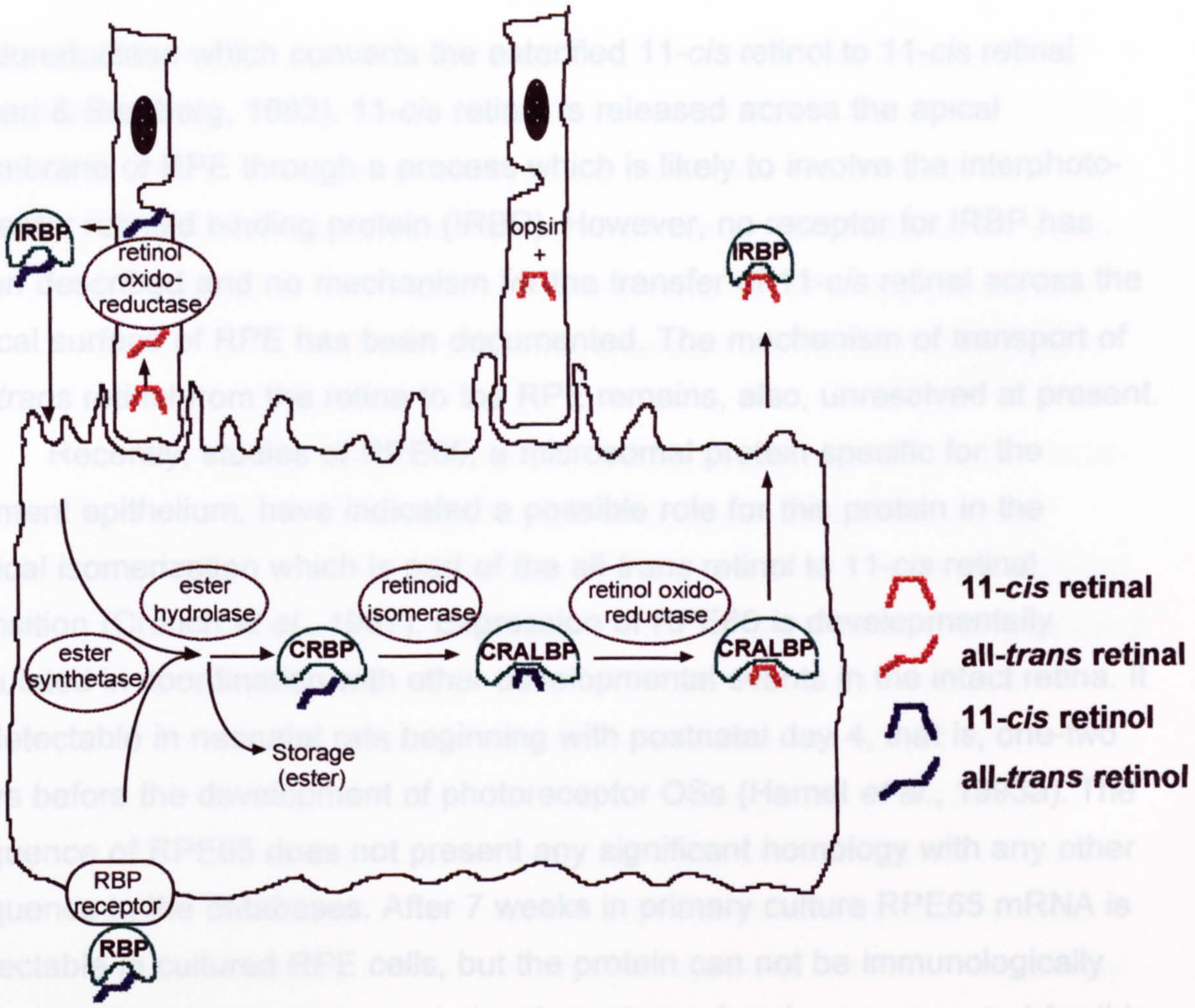


Figure 1.7. Diagram of conversion of retinoids in the visual cycle. All-*trans* retinol is converted, via various intermediates, into 11-*cis* retinal, the photopigment chromophore. RPE-specific enzymes play a major role in this process, alongside various extracellular and intracellular retinoid binding proteins. See text for further details.

Since free retinol is toxic for cells, the retinol undergoes esterification (probably bound to a putative plasma membrane receptor), mostly to palmitic acid (Ottonello *et al.*, 1987). The process of esterification is functionally coupled to that of retinol uptake. The fatty acid esters of retinol serve as storage form of retinol within the RPE. The intracellular carriers for retinol and its derivatives are the cellular retinol binding protein (CRBP, which has no amino acid sequence homology with RBP) and cellular retinaldehyde binding protein (CRALBP) (Saari *et al.*, 1982 and reviewed by Bok, 1993). A membrane-bound retinoid isomerase converts the all-*trans* retinoid complexed by CRBP to 11-*cis* retinol (Barry *et al.*, 1989). The 11-*cis* retinoid bound by CRALBP is the substrate for an

oxidoreductase which converts the esterified 11-*cis* retinol to 11-*cis* retinal (Saari & Bredberg, 1982). 11-*cis* retinal is released across the apical membrane of RPE through a process which is likely to involve the interphoto-receptor retinoid binding protein (IRBP). However, no receptor for IRBP has been described and no mechanism for the transfer of 11-*cis* retinal across the apical surface of RPE has been documented. The mechanism of transport of all-*trans* retinol from the retina to the RPE remains, also, unresolved at present.

Recently, studies of RPE65, a microsomal protein specific for the pigment epithelium, have indicated a possible role for this protein in the critical isomerization which is part of the all-*trans* retinol to 11-*cis* retinal transition (Crouch *et al.*, 1997). Expression of RPE65 is developmentally regulated in coordination with other developmental events in the intact retina. It is detectable in neonatal rats beginning with postnatal day 4, that is, one-two days before the development of photoreceptor OSs (Hamel *et al.*, 1993a). The sequence of RPE65 does not present any significant homology with any other sequence in the databases. After 7 weeks in primary culture RPE65 mRNA is detectable in cultured RPE cells, but the protein can not be immunologically detected; therefore post-transcriptional regulation has been suggested for this protein (Hamel *et al.*, 1993b). Further studies, including generation and analysis of RPE65 knockout mice, are currently being carried out by Redmond and coworkers (1997) in order to elucidate the role of RPE65 in the RPE.

Movement of ions and water across RPE

The junctional specialization and consequent polarity of the RPE account for the existence of an electrical potential difference resulting from ion transport between the two surfaces of the epithelium. Also, fluid is transported across the cells, from the apical surface towards the basal infoldings, into the choroid, thus maintaining the neurosensory retina in a proper state of dehydration and optical clarity. These mechanisms contribute to the critical role of RPE in subretinal fluid homeostasis and attachment of the retina to the choroid.

A peculiar feature of the RPE, shared only by the epithelium of the choroid plexus of the brain, is that Na^+,K^+-ATPase, the membrane-integrated enzyme required to set up the ion gradients that drive other ion transporters, is

localized to the apical plasma membrane, rather than to the basal plasma membrane (Bok, 1982; Ostwald & Steinberg, 1980). This specialization is first observed in day 11 in the chick embryo, at a time when the extracellular matrix of developing photoreceptor cells begins to be composed (Rizzolo & Heiges, 1991). Ankyrin and fodrin, the submembrane cytoskeletal proteins normally associated with Na^+,K^+-ATPase are, also, localized apically in the RPE cells (Gundersen *et al.*, 1991).

However, with two other known exceptions mentioned above, CE-9/OX-47 and N-CAM, other enzymes, such as aminopeptidase, have a similar localization to that in other epithelia. Also, the pattern of viral budding, dictated by the sorting of viral envelope glycoproteins, is the same as in other polarized epithelia. Thus, in cultured human and bovine RPE cells, the polarized budding of influenza and vesicular stomatitis virus from the apical, and respectively, basolateral surface is dictated by the targeting of influenza hemagglutinin and vesicular stomatitis glycoprotein to the respective side of membrane (Bok *et al.*, 1992). Therefore, the existence of a general reverse mechanism for membrane protein trafficking in the RPE could be reasonably excluded. Rather, the polarity of Na^+,K^+-ATPase could be dictated by unique domains in the primary sequence of its subunits, alongside with specific interactions with cytoskeletal elements and extracellular matrix components (Bok, 1993).

The apical configuration of the Na^+/K^+ pump is functionally important, as it ensures the return of Na^+ to the subretinal space, thus maintaining the Na^+ gradient needed for the photoreceptor function (Quinn & Miller, 1992) and stabilizing extracellular levels of K^+ following phototransduction (Joseph & Miller, 1991; Miller & Steinberg, 1982). Transport of K^+ by the RPE depends on the K^+ concentration on the retinal side and is performed in such a way as to keep minimal variations in this concentration.

Miller *et al.* (1982) have documented coupling of transport of water with movement of ions across RPE. The net fluid movement from the retina to the choroid across RPE occurs at a net rate of about 4-6 $\mu l/cm^2$/hour (Forrester et al., 1996) and is inhibited by increased intracellular concentration of cAMP. Hence, cAMP has been suggested to be a modulator of water and ion transport in the RPE.

It is very likely that RPE, due to its ability to actively transport ions and accompanying water from the subretinal space, contributes to the mechanisms of retinal adhesion/attachment. Subretinal fluid absorbtion is essential for keeping the space between photoreceptor OSs and RPE very thin and, consequently, for maintaining a proper communication between them.

1.2.2.4. Perturbation of outer blood-retinal barrier and accumulation of subretinal fluid

Failure or abnormalities of active transport or damage in the pigment epithelium can cause increased permeability and breakdown of the outer blood-retinal barrier, resulting in fluid accumulation in the subretinal space and subsequent retinal detachment.

The barrier is compromised in several retinopathies (Cunha-Vaz, 1979). For example, in serous retinal detachments, the primary pathology is caused by an impaired RPE fluid transport, while small RPE defects or leaks are secondary elements of this disease (Marmor, 1990). Since retinal function requires the existence of the blood-RPE barrier, active transport across RPE must play a major role in the mechanism of subretinal fluid control. Mechanisms are required both for preventing subretinal fluid to accumulate under normal physiological conditions and for removing it under pathological conditions.

Activities of several lysosomal enzymes, such as acid phosphatase, β-glucuronidase, and cathepsin D, have been observed in subretinal fluid (Hayasaka *et al.*, 1976). These activities increase with the duration of retinal detachment, as does the total protein content in subretinal fluid. Studies of enucleated pig eyes perfused through the ciliary artery with lysosomal enzymes have shown an increased permeability of the outer blood-retinal barrier for high molecular weight tracers, such as ferritin (Prunte & Kain, 1995). It is thus conceivable that lysosomal proteinases may participate in damaging this barrier. This action would be consistent with the correlation between degradation of the extracellular matrix by secreted proteinases and cell migration and proliferation in invasive growth and scar tissue formation.

The proteolytic activity in the subretinal fluid may contribute to the loosening and release of RPE cells from the monolayer into subretinal fluid, which is a phenomenon associated with retinal detachment and proliferative vitreoretinopathy. Other aspects of these retinopathies are discussed in relation to ECM modulation by RPE, in section 1.2.2.5.

1.2.2.5. The role of RPE in biosynthesis and modelling of extracellular matrix

RPE synthesizes components for two highly specialized and entirely different extracellular matrices (ECM) with which its apical and basal membranes come in contact: the interphotoreceptor matrix (IPM) and, respectively, the basal lamina. Besides synthesis and polarized deposition of ECM molecules, RPE also performs a systematic degradation of ECM components, as part of ECM remodelling (Campochiaro *et al.*, 1986). Abnormalities in these processes underpin structural and functional changes of the RPE.

Interphotoreceptor matrix

The IPM is an organized assembly of extracellular molecules, which normally fills the subretinal space surrounding completely the outer segments and a portion of the inner segments of photoreceptors and comes in direct contact with RPE cells. Interestingly, IPM serves as ECM between the apical surfaces of two types of cells, - an unusual situation considering that cell attachment to ECM or to other cells occurs normally at the basal side of cells.

IPM mediates the interaction of RPE with photoreceptor OSs, a key element for the function of this epithelium and an essential condition for normal retinal function. It is generally believed that the components of IPM play a role in mediating the exchange of substances between RPE and photoreceptors, in providing a developmental and trophic support for the retina, and in supporting the mechanism of retinal adhesion with the possibility for continuous repositioning connected with shedding and phagocytosis.

Research in the IPM has been directed first towards determining the nature of its components and then revealing which of these molecules possess enzymatic, adhesive, or ligand-carrying activities. A number of IPM components, including glycoproteins, proteoglycans, trophic and growth factors, and acid proteases have been identified, but their roles and their origin have not yet been fully characterized. Thus, even in the case of IRBP, which is the best characterized component of IPM and was the first to which a specific function was assigned, controversy still exists over the nature of the mechanism for transfer of retinoids (Ho *et al.*, 1989; Okajima *et al.*, 1989). The IPM composition is reviewed in detail by Hewitt and Adler (1994). IPM differs from other extracellular matrices in that it does not contain collagen and other attachment factors, such as N-CAM (Adler & Kluznik, 1982; Fliesler *et al.*, 1990).

IPM has an intricate level of organization with many components distributed heterogeneously within it and with distinct soluble and insoluble compartments. Some of its insoluble components, especially proteoglycans, show photoreceptor cell-specific distribution and form an elaborate system of sheaths surrounding individual photoreceptor OSs (Sameshima *et al.*, 1987). These structures may provide mechanical and physiologic insulation for photoreceptor OSs and/or could facilitate retinal attachment and phagocytosis of shed OSs.

Neural retina and RPE supply different components to IPM. For example, IRBP is synthesized and secreted into the IPM by photoreceptor cells (Politi *et al.*, 1989). On the other hand, lysosomal enzymes, such as β-galactosidase, β-glucuronidase, N-acetyl-glucosaminidase, cathepsin D,are secreted by RPE cells (Adler, 1989). These enzymes include the significant acid protease activity of the IPM. This has been assessed to represent aproximately 10% of the total such activity in the retina-IPM-RPE functional unit, and may be involved in processing of shed photoreceptor OSs prior to phagocytosis (Hewitt & Adler, 1994).

Bruch's membrane

RPE synthesizes and deposits molecules into a basement membrane to which it is firmly attached. The endothelium of the choriocapillaris also synthesizes a basement membrane. These two basement membranes are separated by a meshwork of structural proteins which constitute layers of the prominent Bruch's membrane. This structure contains: the matrix proteins collagen type IV and laminin (Hewitt & Adler, 1994), a mixture of proteoglycans, including collagen type I, III, VI, VII, elastin, heparan sulphate and chondroitin sulphate (Hewitt *et al.*, 1989; Pino *et al.*, 1982). These negatively charged proteoglycans probably confer molecular-sieving properties to Bruch's membrane, which can regulate the flow and nature of macromolecules from and towards the choroid circulation.

Bruch's membrane undergoes a constant turnover and remodelling throughout life. Changes in the RPE basement membrane and Bruch's membrane occur with age and are associated with various retinal and RPE pathological conditions, such as retinitis pigmentosa, age-related macular degeneration and drusen formation (Hewitt *et al.*, 1984; Sarks, 1980; Stramm *et al.*, 1989). It is possible that substances secreted by aging RPE cells initiate alterations in structure and elasticity of Bruch's membrane, which ultimately interfere with the transport of essential nutrients to the photoreceptors.

The nature of the factors controlling the steady state of structural and functional specializations of Bruch's membrane is unclear. Interesting interactions between RPE and choriocapillaris endothelium have been reported both *in vitro* and *in vivo*. *In vitro* studies of these interactions have led to a proposal of the existence of an RPE-produced inhibitor of neovascularization (Glaser *et al.*, 1985) and an endothelial-located chemoattractant for RPE cells (Campochiaro & Glaser, 1985). Such reports are complemented by observations of *in vivo* choriocapillary endothelial transformation and atrophy which occurs when RPE is destroyed (Korte *et al.*, 1984), thus indicating a role for the RPE in modulating choriocapillaris structure and function.

The role of RPE in modulation of the ECM has major implications in vitreoproliferative disorders. In proliferative vitreoretinopathy (PVR), the major cause of the failure of retinal reattachment surgery, RPE cells are thought to contribute, alongside retinal glia, macrophages and fibroblasts, to the formation of fibroelastic membranes consisting of proliferative cells blended with fibrillar proteins, such as collagen and elastin (Hiscott & Grierson, 1991; Hiscott *et al.*, 1984). The membranes attach to the retina and are progressively contracted by the cells, leading to retina detachment (Wiedemann & Weller, 1988). Changes in the phenotypic characteristics (metaplasia) of RPE cells have been observed in PVR. RPE dedifferentiation involves altered biologic properties and cytoskeletal changes, but the underlying molecular mechanisms are not fully understood (Grierson *et al.*, 1994; Grisanti & Guidry, 1995). Numerous factors that associate with, or stimulate migration and/or proliferation of RPE cells have been reported. Among these are: vitreous and serum-derived factors (Kirchhof *et al.*, 1989; Martini *et al.*, 1992); platelet-derived mitogens (Smith-Thomas *et al.*, 1996); adhesive ECM glycoproteins, such as thrombospondin (Hiscott *et al.*, 1992a) and fibronectin (Hiscott *et al.*, 1992b). Different agents are responsible for stimulating or influencing in part migration/adhesion/contraction and proliferation (Grierson *et al.*, 1996; Hackett *et al.*, 1989; Mazure & Grierson, 1992).

1.3. STUDY OF GENE EXPRESSION BY cDNA ANALYSIS. APPLICATION TO RPE

1.3.1. HIGHLIGHTS OF SOME STRATEGIES FOR GENOME ANALYSIS

1.3.1.1. Conception of the Human Genome Project

The development of molecular biology during the last decade has led to the conception of a scientifically most challenging project, that of systematically looking into and deciphering the human genome.This is a massive task. The human genes are spread through 3,000Mb of haploid DNA. They represent at the most 4% of the whole genome, considering the average lenghth of a human coding sequence of DNA of aproximately 1.2kb and the estimate of the total number of genes in the human genome of up to 100,000. It is the major goal of the Human Genome Project to identify, sequence and structurally and functionally characterize all of the human genes.

1.3.1.2. Gene mapping and comparative genomics

A concerted effort towards assembling a public-domain human gene map passed a milestone in October 1994 (meeting organized by the Wellcome Trust, in Washington), when the plans for coordinating the strategies pursued by various associations emerged. The primary repository for human genome mapping data generated by the Human Genome Project is the human Genome DataBase, GDB, developed and maintained by the Johns Hopkins University, USA. GDB provides extensive links to other databases with related information (e.g. mouse homologues (mgd), sequence data (EMBL, GenBank), genetic disorders (OMIM) and protein databases).

Once the genome sequence information is collected, the next challenge is to interpret it in functional terms. Consequently, other more direct approaches are being used, especially to identify coding sequences within large genomic regions of DNA. Some of these are based on the comparison

of human sequences with sequences from other organisms, assuming that conserved sequences have some function. Projects exploring other vertebrate genomes (one of which is that of Puffer fish, *Fugu rubripes*, currently developed at the UK Human Genome Mapping Project (UK HGMP) Resource Centre, in Hinxton) allow such comparative sequence analysis efforts. These projects were initially started due to the need for a simple genome approach to gene identification and characterization. In the Puffer fish genome (400Mb), for example, the proportion of coding DNA is much higher (around 20%) than in the human genome, owing to greatly reduced intron sizes, intergenic distances and repetitive DNA (Brenner *et al.*, 1993). Sequence comparisons between mammals or vertebrates have been used to identify functionally important domains of gene expression products (as in the case of Huntington's disease gene (Baxendale *et al.*, 1995)) and conserved regulatory elements (Aparicio *et al.*, 1995). The benefits of comparative sequencing with the mouse have recently been reviewed (Hardison *et al.*, 1997). Regions of up to 3 Mb of the mouse genome and syntenic (of conserved linkage) regions of human chromosomes are subjected to comparative sequencing analysis (Botcherby, 1997). An alternative approach is that of positional cloning projects directed towards identification of conserved linkage groups mostly amongst mammals but also with chicken (Shaper *et al.*, 1997; Womack & Kata, 1995).

A list of all bioinformatics centres working on the human or other species genome is available from the world wide web menu page of UK Human Genome Mapping Project (http://www.hgmp.ac.uk).

1.3.2. COMPLEMENTARY DNA SEQUENCING AND EXPRESSED SEQUENCE TAGS

1.3.2.1. cDNA analysis

Simultaneously with the idea of determining the complete nucleotide sequence of the human genome, the complementary conception of large-scale sequencing of cDNA, reverse transcribed from mRNA, was advanced (Brenner, 1990). The proposed advantages of cDNA sequencing are as follows. Since cDNA represents the protein-coding content of the human

genomic sequence, but only 3% of the total DNA, deciphering this information has the potential to lead more directly and less tediously to the functional gene products. Hence, genes could be identified much more quickly. Furthermore cDNA sequencing could reveal which genes are important for the function of specific tissues/cells, in health and disease. Limitations of cDNA sequencing lie in the difficulty of finding every mRNA in all cell types and in all developmental stages and in the inability to access the information from the intronic and intergenic regions which includes control and regulatory sequences.

Besides seeking to characterize the complexity of genetic information, analysis of cDNA aims to investigate the ordered and timely expression of this information. cDNA libraries are the starting material for analysis of cDNA.

1.3.2.2. Partial cDNA sequencing, characteristics of ESTs, and strategies for their use

Craig Venter and coworkers (Adams *et al.*, 1991) initiated a new approach for the rapid characterization of expressed genes by partial cDNA sequencing (see section 1.3.3.). This strategy is based on the analysis of Expressed Sequence Tags (ESTs).

ESTs are short identifying nucleotide sequences resulting from partial sequencing of cDNAs. Usually, the information from a single (one-pass) sequencing reaction is sufficient to allow meaningful comparisons with other sequences, which enable identification or assignment to a gene family when sufficient sequence similarity exists. ESTs, and implicitely cDNAs, can be mapped to a location in the genome by hybridization approaches or by PCR approaches. The sequence corresponding to an EST can be amplified with suitable primers by PCR from genomic DNA, converting the EST to a "sequence tagged site" (STS). These are short fragments of unique DNA sequences that define landmarks on the genome (Olson *et al.*, 1989).

ESTs differ from the traditional database entries both in qualitative and quantitative terms (Boguski *et al.*, 1993). The length of an EST is a function of current automated sequencing technology. The average lenghth of ESTs deposited in databases is about 300-400 nucleotides. Only very recently,

advances in this technology have increased the average lenghth of the sequence obtained through a single pass sequencing reaction to about 1,000 nucleotides. Also, EST sequences have a lower accuracy compared with the fully characterized clones, because only one strand is sequenced, and only on one occasion. Contamination with vector and other spurious sequences is possible. Moreover, it is not generally known, without further studies, if the fragments represent coding or non-coding regions of mRNA. It is also noticeable that a high proportion of ESTs (probably due to a combination of the facts mentioned above) do not present homology with other sequences at the time of submission. A high degree of redundancy (more than one EST corresponding to a gene) may be found. Nevertheless, ESTs represent an abundant source of new sequences.

Although sharing numerous features, in particular technical characteristics, projects involving ESTs are of two major strategies. Some, supported by large collaborative groups, aim to rapidly generate large numbers of such sequences which can then be used extensively in genomic sequencing projects (see section 1.3.2.3.). Other projects are concerned with unveiling the expression patterns of various cells and tissues. The latter ones, while, obviously, contributing to the general sequencing effort, most importantly enhance our knowledge about the biological functions of such cells or tissues. Some characteristics of this strategy are further described in section 1.3.3.

1.3.2.3. Contribution of ESTs analysis to the genome research

Following the impetus for starting large-scale EST work given by Craig Venter and others, the accumulation and analysis of ESTs has become an important component of genome research. The extensive use of ESTs in sequencing projects is based on the potential for early gene discovery associated with the long-term value for analysis of the entire genome.

In 1993, a new database of "public" data, dbEST, supported by the National Center for Biotechnology Information (NCBI), has been created specifically to accomodate the unique features of EST data. Some of the major

contributors to the information in this database are listed in Table 1.3. The most important of these, the collaboration between Washington University and Merck (and National Cancer Institute, since 1997) generated 280,000 human ESTs (Hillier *et al.*, 1996). EST sequences from dbEST database are available for homology searches from the BLAST network of NCBI (blast@ncbi.nlm.nih.gov), as well as from the EST division of GenBank.

Table 1.3. Major centres for ESTs analysis.

Project name/Location	**WWW site**
Washington University/Merck/ National Cancer Institute Sequencing Project	http://genome.wustl.edu/esthmog.html/
Sanger Centre/UK-HGMP	http://www.sanger.ac.uk/ http://www.hgmp.mrc.ac.uk/
The Institutre for Genomic Research (TIGR)	http://www.tigr.org/
Whitehead Institute/MIT	http://www-genome.wi.mit.edu/
Stanford University Genome Center	http://www-shgc.stanford.edu/
Genethon, Paris	http://www.genethon.fr/
Cold Spring Harbor Laboratory	http://www.cshl.org/

During the last years various international consortia have been initiated with the aim to minimize overlap and to integrate the access to information on cDNA library clones and derived ESTs. For example, the IMAGE (Integrated Molecular Analysis of Genomes and their Expression) Consortium, initiated in 1993 by four academic groups on a collaborative basis, coordinates a number of ongoing sequencing projects and places sequence, map, and expression data into the public domain (Lennon *et al.*, 1996). Its World-Wide Web site (http://www-bio.llnl.gov/bbrp/image/) is maintained at the Lawrence Livermore National Laboratory, University of California.

The number of ESTs reported by IMAGE Consortium to have been deposited in dbEST at the beginning of 1998 (over 800,000 of 5' and/or 3' sequences) indicates that, by now, a significant part of all human genes has

been tagged through EST sequencing. Defining how many different genes these represent is not easy and figures are rarely quoted. A number of technical issues are involved, such as: position along the transcript of the partial sequences, accuracy of sequencing, existence of alternatively spliced forms, as well as unforseen effects of cDNA library manipulations. An estimate of 60,000 distinct human genes has been indicated on the IMAGE's Web page in February 1998.

Further, ESTs are used extensively in the analysis of complex gene structures such as alternative splice sites, overlapping transcription units, and long introns. ESTs verify that predicted genes are transcribed and identify splice sites.

1.3.3. ASSESSMENT OF EXPRESSION PATTERNS BASED ON EST ANALYSIS

The EST strategy, based on analysis of single-pass, partial sequencing of cDNA clones from one or both ends, enables the access to many of the expressed genes of an organism. The importance of surveying expressed human genes is twofold. Firstly, it provides a key resource for gene mapping and a rapid means of gene discovery. This line of investigation has been pursued especially by those constructing physical maps of the human genome or searching human disease genes (Adams *et al.*, 1991; Houlgatte *et al.*, 1995; Kahn *et al.*, 1992). Secondly, the combination of data on gene expression and presumed gene functions inferred from sequence similarity allows the assessment of the transcriptional activity in various cells and tissues. The goal of this approach is to provide a description of the cDNA population by random sampling.

The latter idea was pioneered by Sutcliffe and coworkers in a study done in rat (Milner & Sutcliffe, 1983; Sutcliffe *et al.*, 1983). They isolated 158 rat brain cDNAs and compared their expression in brain, liver and kidney. The same approach was used in a study of mouse cerebellum, in which the complexity of cerebellar cDNA was compared with that of the whole brain (Kato, 1990).

The largest contribution to this area of investigation is that of Craig Venter and colleagues. Originally initiated to rapidly provide informative probes for the human genome, focusing on sequencing followed by chromosome localization of ESTs (Adams *et al.*, 1992; Adams *et al.*, 1991), their strategy evolved into analysis of diversity of transcripts in human brain (Adams *et al.*, 1993). The most striking result of this first large-scale analysis of ESTs (over 3,000 cDNA clones were partially sequenced) is that less than 15% of the sequences presented an exact human match and only about 20% of the ESTs could be assigned based on database matches. Thus, based on sequence similarities with genes characterized in different organisms and tissues, approximately 700 different proteins were shown to be expressed in brain. Over 55% of sequences showed no database match. Furthermore, out of these, a large number of probable coding sequences (over 50%) were not significantly similar to one another. This fact led to a new appreciation of the diversity of gene families. The high biochemical activity of the brain directed toward signalling and regulatory processes was indicated by the numerous receptors, transcription factors and other signal transduction proteins identified. Among the new genes identified based on significant similarities to genes in the database was that of a new neurotransmitter transporter gene, *Notch/Tan-1*.

This strategy grew in popularity and was applied by others to Caenorhabditis elegans (McCombie *et al.*, 1992), Arabidopsis (Newman, 1994), and rice (Sasaki *et al.*, 1994).

By 1995, Craig Venter's group at The Institute for Genomic Research released the most comprehensive survey to date of expression patterns of human genes (Adams *et al.*, 1995). The group of 85 co-authors reported the partial sequencing of over 170,000 cDNAs originating from 300 cDNA libraries constructed from 37 distinct organs and tissues. The dbEST database reached, following this major contribution, a total of almost 300,000 sequences. Less than 30% (under 90,000) from these were identified by the authors of this survey as distinct sequences. Remarkably, only 10,214 of these matched previously characterized entries in the databases, the rest showing no significant similarity with known sequences. The vast amount of data produced by this study and by the EST projects in other tissues, such as

brain (Adams *et al.*, 1993; Kahn *et al.*, 1992), liver (Okubo *et al.*, 1992), testis (Affara *et al.*, 1994), pancreatic islets cells (Takeda *et al.*, 1993), lung (Sudo *et al.*, 1994), and heart (Liew *et al.*, 1994) enabled for the first time a systematical classification of expressed human genes with respect to their functions and biological roles in cellular biology. The analysis of tissue distribution combined with that of transcript abundance (judged after the fraction of all ESTs corresponding to a single gene or a tentative human consensus sequence (THC) of the total profile of ESTs) identified potentially tissue-specific, abundant genes. Most of these genes that appear to be moderately or highly expressed in a tissue-specific manner have been previously characterized, for example: genes encoding digestive enzymes from the pancreas, histatins from the salivary gland, pulmonary surfactant associated proteins from the lung, and amyloid precursor-like protein from the brain.

Another large application of the expression profiling is the Anatomical Expression Database of human genes, "BodyMap". This project was initiated by Kosaku Okubo in 1991 (Okubo *et al.*, 1992) at the Institute for Molecular and Cellular Biology, Osaka University, Japan. In the BodyMap, gene transcripts are identified by the so called Gene Signature (GS) which is a sequence at the 3'-end of cDNA. The GS does not usually cover the coding region, as the primary aim of sequencing is not to identify the nature of the encoded protein. These 3'-end ESTs are generated from clones randomly isolated from 3'-directed cDNA libraries which are considered to be faithful representatives of the molar composition of the mRNA. Therefore, the frequency of appearance of any such GS in an expression profile is an indicative of the abundance of corresponding transcript and, hence, of the expression level of the corresponding gene in that tissue. It is the aim of this project is to compile expression profiles of different tissues in order to illustrate the relative expression level of diverse genes and to show the tissue-specific expression control of those genes (Okubo *et al.*, 1995). The BodyMap contains cDNA sequences that are not yet deposited in any other databank (the information is obtained from the www server at http://imcb.osaka-u.ac.jp/bodymap/). Currently, expression profiles of various complexities have been reported for a number of 38 cells and tissues. Among these are the following: lung, aortic endothelium, fibroblast cells, primary cultured keratinocytes, liver, colon

mucosa, cerebellum, hippocampus, thalamus, cultured fetal astrocytes, T cells, and granulocytes. About 20% of the GS collected can be identified as known genes, the remaining 80% being considered to potentially represent new genes.

1.3.4. STUDIES OF ESTs IN THE EYE

1932 ESTs derived from the whole eye were reported in the most comprehensive published collection of data on ESTs to date, The Genome Directory (Nature, Supplement to Volume 377, September 1995). The authors identify 547 distinct genes (Adams *et al.*, 1995). Putative roles were assigned to aproximately 75% of these, which have then been preliminarily classified into 7 groups of cellular roles (similarly with the classification of genes across the other 36 human tissues analyzed). The report does not contain references to specific tissues within the eye.

The BodyMap database (see section 1.3.3) contains information on expression profiles of genes within corneal epithelium (Nishida *et al.*, 1996) and retina (Okubo *et al.*, 1992). The expression profile of the retina contains 154 different GS; the highest frequency of appearance is that of the opsin GS (14 out of 331 clones). This data is extended by the information contained in the article of Shimizu-Matsumoto *et al.* (1997) which analyses 607 different sequences from neural retina. Transcripts of components that have a rapid turnover, such as those involved in the constant renewal of photoreceptor cells OSs were found recurrently in the profile. Thus, 290 sequences corresponded to 108 "very active genes", whose GS were found repeatedly in the profile. The assembled list of these includes, as the two most abundant transcripts, opsin (16 appearances) and Na^+K^+ ATPase $\beta 2$ subunit (11 appearances), alongside other components of phototransduction such as transducin γ and cone cGMP phosphodiesterase γ, each of which had 4 appearances. A secretory protein, transferrin, which has been regarded as a "serum" component of vitreous fluid, appeared relatively highly represented in the profile of retinal transcripts (3 appearances). Following comparisons with entries in dbEST database (currently containing about 3,000 3'-ESTs from retina) and with expression profiles of other tissues, 25 distinct sequences

out of the 108 "very active genes" have been identified as representing genes uniquely expressed in the retina. Six of these are known retina-specific genes identified in GenBank (two opsins, transducin-γ subunit, cone cGMP phosphodiesterase γ, guanylate cyclase activating protein, and cocaine amphetamine regulated transcript). The rest of 19 retina-specific GS characterize putative novel genes.

Another application of the EST analysis to the study of retina is the characterization of a human age-normalized cDNA library of the fovea, the highly specialized region of the retina responsible for acute vision (Bernstein *et al.*, 1996). The study represents a step towards analysing quantitative differences in gene expression in fovea *vs.* the rest of the retina. Single-pass sequencing was performed on 209 randomly isolated clones. A significant proportion (42.1%) of fovea ESTs thus generated represent new ESTs, having no identical matches to GenBank sequences. Interestingly, mitochondrial gene sequences represent a total of 18.7% of all clones screened, indicating a high level of transcription of these sequences in fovea. This suggests an elevated level of oxidative phosphorylation activity, which is consistent with the high energy requirement of the retina.

1.3.5. APPLICATION OF cDNA ANALYSIS BY CHARACTERIZATION OF ESTs TO RPE

The approach undertaken in the only other study to date of ESTs in the RPE was based on the use of a subtracted cDNA library constructed from a human RPE cell line (Gieser & Swaroop, 1992). A biotin-based subtraction procedure made use of the JY cell line, an EBV-transformed human B-lymphoblastoid cell line (Swaroop *et al.*, 1991). The rationale behind using a subtracted cDNA library was to identify RPE-specific genes; subsequent chromosomal mapping may help in detecting candidate genes for inherited eye disorders. Also, the new ESTs generated could provide new markers for physical and genetic mapping studies. ESTs were generated from the 3'-end of 108 randomly selected cDNA sequences and in each case the sequence near the poly(A) site was used as a reference point for comparing various ESTs. Among the 28 clones which were either identical to human genes or

similar to vertebrate genes, one appeared three times (myristoylated alanine-rich C kinase substrate) and six were isolated twice (apoferritin H, apolipoprotein E, cystatin C, DNA binding protein B, LL Rep3, profilin). Of the cDNAs with no homology to database sequences (61 clones), seven were subjected, in a later study (Agarwal *et al.*, 1995), to a comparative expression analysis by Southern and Northern blotting in the RPE and JY cell lines. This analysis revealed two clones preferentially expressed in RPE.

1.4. AIMS OF THIS STUDY

The molecular mechanisms underlying the processes in which RPE is involved are likely to require a broad spectrum of proteins. Some of these may be unique and specific to the RPE and some others may be of known function in other tissues and also responsible for essential activities in the RPE.

As has been highlighted in the previous sections, only parts of these molecular mechanisms are well documented. Knowing which genes are expressed in the RPE should contribute to a better understanding of the mechanisms that endow the RPE its unique status within the eye and its supportive function in the visual process. A transcriptional picture of these cells is, therefore, likely to be informative.

This study is an application of cDNA cloning to investigate the functions of RPE by identifying details of its transcriptional picture. The approach used for revealing characteristics of the profile of genes expressed in the RPE was the EST analysis of randomly selected clones from cDNA libraries. The analysis was aimed to be a step forward towards an overview of proteins found in the RPE and to help identify proteins which may be important for RPE cellular activities.

CHAPTER 2

MATERIALS AND METHODS

2.1. OUTLINE OF STRATEGY

As an approach to the identification of ESTs in the RPE, different RPE cDNA libraries were used as sources of clones for sequencing in this study.

At the start of the project, it was intended that the direction be towards the study of cDNA libraries of freshly isolated bovine RPE cells. It was believed that information which could be extracted from the analysis of cDNA libraries of freshly isolated RPE cells would have the potential to be more informative than that from cDNA from cultured cells, as it could reveal the gene species expressed *in vivo*. However, during the first year of the project, upon realising the difficulties in obtaining sufficient starting material (number of eyes available for dissection and, then, restrictions imposed by the BSE epidemic), the use of cDNA libraries of cultured RPE cells had to be considered. It was then found that the cDNA libraries constructed in-house did not have a high percentage of recombinants with long inserts. This led me to the further decision to make use of an aliquot of a commercially constructed cDNA library of cultured human foetal RPE cells, kindly made available to us by Dr. Dean Bok of UCLA, USA.

Since the strategy of the analysis was to ensure an efficient start for identifying genes of potential importance for RPE, the number and diversity of sources, i.e. cDNA libraries, of clones for sequencing was potentially advantageous. Technically, the libraries differ also with regard to the method of construction, i.e. priming of first strand cDNA synthesis, vector used, directionality of cloning. These characteristics and the way in which they were taken into account when analysing sequencing data are described in this thesis.

Most of the ESTs generated by this study originate from the commercially constructed cDNA library. Nevertheless, I include the description of the methodology used by me to construct cDNA libraries because I feel that the conclusions reached by pursuing this work could be useful for future reference.

Where appropriate in this chapter, in the interests of brevity, clarity and possible future utility, experimental protocols for straightforward easily reproduced methods are given in the present imperative.

2.2. MATERIALS AND REAGENTS

Unless otherwise indicated in the text, chemicals and antibiotics used in this study were purchased from BDH Chemicals Ltd, England and Sigma, UK; components of cell culture and bacterial culture media from Gibco BRL; disposable sterile materials from Sterilin and Nalgene Ltd, England; restriction and modifying enzymes from Boehringer Mannheim GmbH Germany, Pharmacia Biotech, and Amersham, UK; radioisotopes from ICN, Pharmaceuticals, Inc. The sources of antibodies are listed in Table 2.4. The manufacturers of specific kits used for various applications are indicated when the particular application is described.

2.3. MEDIA

2.3.1. MEDIA FOR BACTERIAL CULTURE

All bacteria were grown in LB (Luria-Bertani) medium. Special requirements of bacterial cultures used for phage manipulations are indicated in the appropriate protocols. Antibiotics were used when selection of plasmid-bearing strains was needed.

LB Medium	per liter:	final concentration
Bacto-tryptone	10g	1%
Yeast extract	5g	0.5%
NaCl	10g	1%

pH was adjusted to 7 with 5M NaOH. Media were sterilised by autoclaving immediately after preparation. When preparing plates or top agar, 15g or 8g respectively, of agar were added per liter of medium before autoclaving.

Stock solutions of thermolabile compounds, such as antibiotics and lactose, were filter-sterilised through Whatman filters (0.2µm pore size) and added to autoclaved media, previously allowed to cool to 55°C. The concentrations of antibiotics in media and stock solutions are presented in Table 2.1.

Agar plates were stored at +4°C and dried briefly before use in safety cabinets.

Antibiotics

All stock solutions of antibiotics were stored at -20°C in light-tight containers.

Table 2.1. Stock solutions and final working concentrations of antibiotics in bacterial culture media.

Antibiotic	Stock solution concentration (mg/ml)	Final concentration (µg/ml)
	in water:	
Ampicillin	100	50-100
Carbenicillin	25	50
Kanamycin	10	50
	in ethanol:	
Chloramphenicol	34	34
Tetracycline	5	12.5-50

2.3.2. MEDIA FOR BACTERIAL TRANSFORMATION

The following media (adapted after Miller, 1972) were used in the protocol for transformation of bacterial cells with plasmids.

SOB Medium

	per liter:	final concentration
Bacto-tryptone	10g	1%
Yeast extract	5g	0.5%
NaCl	5g	0.5%
KCl	0.18g	2.5mM
$MgSO_4$	4.92g	10mM

SOC Medium

SOB medium with glucose at a final concentration of 20mM.

Frozen Storage Buffer (FSB)

	per liter:	final concentration
CH_3COOK	0.98g	10mM
KCl	7.5g	100mM
$MnCl_2\ 4H_2O$	8.9g	45mM
$CaCl_2\ 2H_2O$	1.47g	10mM
$Co(NH_3)_6Cl_3$	0.8g	3mM
Glycerol	100ml	10%

The buffer (pH 6.4) was filter-sterilized and stored at 4°C.

2.3.3. RPE CELL CULTURE MEDIA

The culture media for bovine and human RPE cells consisted of Eagle's Minimal Essential Medium (MEM), containing 100units/ml penicillin, 100µg/ml streptomycin, 25µg/ml fungizone (amphotericin B), supplemented with 2mM L-glutamine and buffered with 2.2g/l sodium bicarbonate. pH was adjusted with 1M NaOH and either Foetal Calf Serum (FCS) was added to a final concentration of 20% for primary cultures, or Newborn Calf Serum (NCS)

to a concentration of 15% for passaged cultures. All media components were obtained from Gibco, UK.

2.4. BACTERIAL STRAINS

2.4.1. DESCRIPTION OF BACTERIAL STRAINS USED AND CULTURE CONDITIONS

The genotypes of *Escherichia coli* strains used in this study are presented in Table 2.2 and their culture conditions and applications are summarized in Table 2.3.

All *E. coli* strains were revived, when required, from frozen glycerol stocks by streaking small portions of the frozen stock on agar plates containing the appropriate antibiotic and allowed to grow overnight at 37°C. These primary streak plates could be stored at 4°C and, for up to 2 weeks, could provide a source of individual bacterial colonies used to inoculate liquid cultures. Fresh working stock plates were prepared at 2-week intervals.

The cell density in liquid cultures was determined spectrophotometrically (0.1 $OD_{600} \approx 10^7$ cells/ml).

2.4.2. STORAGE OF BACTERIA

For long term storage, bacterial strains and bacteria transformed with plasmid DNA were maintained as glycerol stocks.

Fresh glycerol stocks were made by adding sterile glycerol, to a final concentration of 40%, to fresh overnight cultures in LB medium with the appropriate antibiotic(s) for selection. Aliquots of the glycerol stocks thus prepared were stored at -70°C and should remain viable for several years.

Table 2.2. Bacterial host strain genotypes.

Strain	Genotype	Reference/Notes
DH5α	*supE44 endA1Δ (lacZYA-argF) (f80 lacZΔ M15) hsdR17 recA1gyrA96*	Hanahan, 1993
ER1647	F'*fhuA2ΔlacZr1 supE44 trp31 recD1014 mcrA1272::Tn10 rpsl104 xy17mt12 metB1 hsdS($r_{K12}^-m_{K12}^-$) Δ(mcrC-mrr)102::Tn10*	Woodcock *et al.*, 1989 Tn10 (Tet^R)
BM25.8	*supF thiΔ(lac proAB) (*F' *traD36 proAB$^+$ lacI^{q}ZΔM15)λ imm^{434}*(kan^R) P1(Cm^R) *hsdR($r_{K12}^-m_{K12}^-$)*	double lysogen of P1 (*cre*-recombinase) and *λ imm^{434}*(kan^R) (prevents lytic infection)
XL1-Blue	*supE44 endA1 lac thi-1 relA1 recA1 gyrA96 hsdR17 (F' lacI^{q}ZΔM15 proAB Tn10)*	Wood *et al.*, 1985 Tn10 confers resistance to tetracycline
BNN132	*supE44 endA1 thiΔ(lac-proAB) relA1 gyr96 hsdR17 (*F' *traD36 proAB$^+$ lacI^{q}ZΔM15) λ KC(kan-cre)*	Elledge *et al.*, 1991 Kanamycin-resistant, *cre*-producing lysogen
XL1-Blue MRF'	*supE44 endA1 lac thi-1 relA1 recA1 gyrA96 Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 (F' lacI^{q}ZΔM15 proAB* Tn*10)*	XL1-Blue with deletion of host restriction systems Tetracycline-resistant
SOLR	*Su$^-$ endA1 lac thi-1 relA1 gyrA96 e14$^-$(McrA$^-$)Δ(mcrCB-hsdSMR-mrr)171sbcC recB recJ uvrC umuC::*Tn5*λ ^{R}(F' proAB lacI^{q}ZΔM15)*	Kanamycin-resistant

Table 2.3. Host strain applications and media additives.

Strain	Application(s)	Stock plate
DH5α	plasmid DNA amplification/ storage, routine transformations	LB
ER1647	λMOS*Elox*-based library plating, screening and routine titering	LB/Tet(50µg/ml)
BM25.8	*cre*-mediated pMOS*ELox* plasmid subcloning	LB/Chl(34µg/ml); Kan(50µg/ml)
XL1-Blue	λTriplEX-based library plating and screening, blue/white (β-galactosidase) screening	LB/Tet(15µg/ml)
BNN132	*cre-lox*-mediated excision of pTriplEX from λTriplEX	LB/Kan(50g/ml)
XL1-Blue MRF'	Uni-ZAP™XR-based library plating screening, blue/white screening	LB/Tet(12.5µg/ml)
SOLR	*in vivo* excision of pBluescript phagemids from Uni-ZAP™XR clones	LB/Kan(50µg/ml)

2.4.3. PREPARATION OF PHAGE PLATING CELLS

To assure a high eficiency of phage infection, the bacterial cells used for phage growth need to be harvested in the logarithmic growth phase (thus minimizing the content of nonviable cells).

Protocol

1. Pick a single, isolated colony from the appropriate fresh stock plate of bacterial culture and use it to inoculate 10ml of LB medium, supplemented with 10mM $MgSO_4$ and 0.2% maltose (use maltose from a 20% stock solution, filter sterilised, kept at 4°C). Grow at 37°C for 4-6 hours, with vigorous shaking (250 rpm on an horizontal shaker gives good results) and monitoring of OD_{600}.

The use of Mg^{2+} and maltose in the growth medium induces phage receptors on the E. coli *surface* (Collins & Hohn, 1978). *The culture should not be allowed to grow past an OD_{600} of 1.0 in order to be substantially free from nonviable cells, as the phage can adhere to such cells as well, resulting in a decresed titer. Alternatively, grow overnight and use 1ml of this culture to inoculate 50ml of prewarmed same growth medium; incubate at 37°C until the cells have grown to an OD_{600} of 0.5-1.0.*

2. Centrifuge the cells for 10 minutes at 3000rpm, pour off the supernatant, and resuspend the cell pellet in half the original volume of sterile 10mM $MgSO_4$.

The cells may be stored at 4°C in 10mM $MgSO_4$, with no significant decrease in plating efficiency for up to 48 hours.

3. Dilute the cells to an OD_{600} of 0.5 with sterile 10mM $MgSO_4$ and use immediately following dilution.

2.5. RPE CELLS: ISOLATION AND CULTURE

2.5.1. ISOLATION OF BOVINE RPE CELLS

Bovine eyes were obtained from local abattoirs, transported on ice and subjected to dissection 2-4 hours post-mortem. The RPE cells were isolated using a technique adapted after Basu *et al.* (1983). Since most of these freshly isolated cells were used for mRNA preparations, needed for cDNA synthesis and Northern blotting analysis, special attention was given to optimizing conditions which can minimize degradation of RNA. The procedure is not easy, because the RPE must be separated both from the neuroretina and from Bruch's membrane.

Protocol

1. Remove extraocular muscles, orbital tissue and optic nerve. Immerse the eye globes in 70% ethanol for 20 seconds and then leave to soak for 10 minutes in sterile Ca^{2+}/Mg^{2+} -free phosphate buffered saline (Dulbecco's PBS: 200mg/l KCl, 8000mg/l NaCl, 200mg/l KH_2PO_4, 1150mg/l Na_2HPO_4) containing 100 units of penicillin and streptomycin.

2. Under sterile conditions, dissect the eye globes at the ora serrata and remove the anterior part of the eye, lens, vitreous and neuroretina.
Use fine forceps to tease out completely the retina, taking care not to pierce through the RPE monolayer into the choroid, as this could later release fibroblasts.

3. Rinse the interior of the eye cup with sterile PBS. Position three eye cups on a sterile petri dish lined with a sterile filter paper (to avoid slipage). Isolate the optic nerve heads (to avoid contamination with optic nerve head cells) with sterile plastic universal bottles (three of these being fixed on the lid of a petri dish). Place a weight of about 200g on top (Figure 2.1).
Alternatively, the optic nerve head can be excized and the resulting hole in the eye cup filled with wax. This also exposes a larger area of the RPE to the trypsin solution used next.

4. Fill each eye cup with 4 ml of a solution of 0.25% trypsin and 0.02% tetra-sodium ethylenediaminetetraacetic acid (EDTA-Na_4) in Ca^{2+}/Mg^{2+}-free PBS. Incubate at 37°C for 30-60 minutes.
Longer incubation times yield more RPE cells detached, but, in order to ensure that no cells are lysed at this stage (a requirement when the freshly dissected cells are to be used for RNA preparation), shorter incubation times are to be used.

5. Assist RPE cells to detach from Bruch's membrane by gently pipetting up and down. Aspirate the cell suspension from the eye cup and add at least 1ml of NCS to inactivate trypsin. When cells are used for direct RNA preparation, centrifuge the suspension at 4°C, 1200rpm, 5 minutes and pour off the supernatant. Wash the cell pellet with PBS at 4°C. Immediately freeze the cell pellet in liquid nitrogen and store at -70°C.
After trypsin has been inactivated, cells can be counted with a haemocytometer (the mean number of cells counted in 16 squares of the grid, i.e. in 0.1 mm^3 X 10^4 = cells/ml suspension). The morphology of cells was examined by light microscopy (section 2.6.1) and the purity of cells was ascertained by immunocytochemical staining (section 2.6.4).

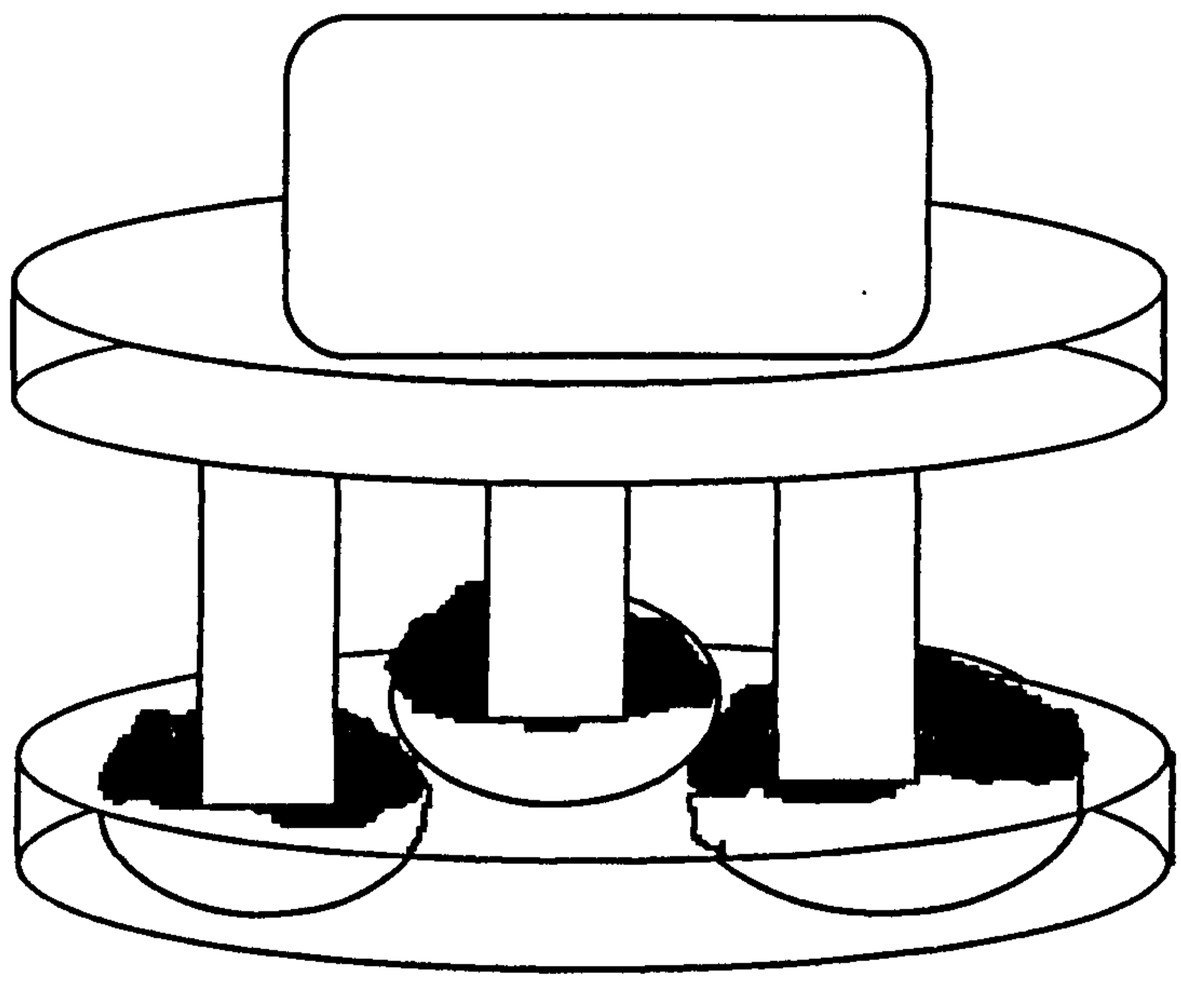


Figure 2.1. The system used to isolate bovine RPE cells.
The optic nerve head in each eyecup is covered by a sterile universal vial; the vials are fixed on a petri dish and a weight is placed on top. The eyecups are filled with trypsin solution, as described in text.

2.5.2. PRIMARY CULTURE OF BOVINE RPE CELLS

Cultures of bovine RPE cells were initially needed for immunohistochemical characterization (section 2.6.4). The protocol is a modification of the technique described by Edwards (Edwards, 1982).

Protocol

1. Transfer the RPE cell suspension obtained as described above from freshly dissected eyes into 25 cm^2 tissue culture flasks containing 1ml of NCS. Add 4ml of culture medium (see 2.3.3.) with 20% FCS to each flask.

2. Keep the cells in a humidified 5% CO_2 incubator at 37°C for about a week before examining.

During this time cells should attach and begin to proliferate.

3. Feed primary cultures twice a week by discarding most of the old medium and replacing it with 5 ml of fresh culture medium containing 20% FCS warmed at 37°C, until they reach confluence.

2.5.3. PASSAGING AND SUB-CULTURE OF BOVINE RPE CELLS

Confluent primary cell cultures were passaged into sub-cultures of bovine RPE cell cultures. These were used for experiments between the second and fourth passage.

Protocol

1. Discard the culture medium and wash cells twice with warm (37°C) PBS. To detach cells from the surface of the culture flask, add 2ml of PBS with 0.25% trypsin and 0.02% EDTA (disodium salt) to each flask. Stop trypsinization after 2 minutes with 1ml of serum (NCS or FCS).

During trypsinization the cells round up and go into suspension. Following inactivation of trypsin, cells can be counted and collected as described in section 2.5.1. Washed cell pellets can be directly resuspended in lysis buffer for RNA preparations.

2. Divide the cell suspension between an appropriate number of 75 cm^2 culture flasks (at a ratio of 1:3) containing 1 ml of NCS. Add fresh culture

medium with 15% NCS and incubate the cultures further in the same conditions as described in section 2.5.2.

For immunohistochemical applications, bovine RPE cells were grown in 8-chambered cell culture slides (LabTek, NUNC Inc., USA).

Protocol

1. Add to each chamber of the slide 0.1 ml of serum (NCS or FCS), then 7.5×10^3cells (resuspended in culture medium after inactivation of trypsin and centrifugation) and, finally, warm culture medium to a final volume of 0.8 ml.

2. Place LabTeks slides on moist filter papers in sterile petri dishes and incubate at 37°C in the same conditions as the cell cultures.

2.5.4. CULTURE OF HUMAN RPE CELLS

Human RPE cells (isolated from enucleated eyes, from St. Paul's eye bank, Liverpool) were revived from frozen stocks stored in liquid nitrogen. The cultures were maintained and passaged in the same conditions as described in section 2.5.3.

2.6. MORPHOLOGY, HISTOLOGY AND IMMUNOCYTOCHEMISTRY OF RPE CELLS

2.6.1. LIGHT MICROSCOPY

The morphology of bovine and human RPE cells in culture was routinely examined using an inverted phase-contrast light microscope (Nikon Diaphot, Japan). Fixed and stained cells were observed using bright field and differential interference contrast (DIC) microscopy (Polyvar, Reichert-Jung, Austria).

Phase-contrast and DIC are types of interference microscopy that make use of the different refractive indices/densities within the cells. In phase-contrast microscopy, the contrast of different regions of the cell is enhanced through the interference effect of two out-of-phase waves. DIC microscopy uses polarized light and gives a three-dimensional image of the cell.

2.6.2. HISTOLOGY OF FRESH TISSUE

Following treatment with trypsin solution for isolation of RPE cells (as described in 2.5.1.), portions of the eye cups (submersed in 10% formalin, i.e. formaldehyde-saline solution, at room temperature, overnight) were cut through sclera/choroid /RPE. The pieces of interest were processed on a Shandon 2LE processor and embedded in paraffin. The processing involved fixation in formalin (5 minutes), gradual dehydration in ethanol (60% ethanol, 90% ethanol and 3 X absolute ethanol, for 20 minutes in each bath), replacement of ethanol with xylene (4 baths, total 80 minutes), and, finally, infiltration of paraffin (60 minutes). Sections of the paraffin-mounted tissue were cut with a Shandon AS325 microtome (5µm), collected on glass microscope slides and rehydrated (reverse treatment of the above).

The sections were subjected to standard histological staining with haematoxylin/eosin (2 minutes immersion in each of the dyes, followed by washes in tap water, gradual dehydration in ethanol, clearing in xylene and mounting in DPX (RA Lamb Lab Supplies)).

2.6.3. FIXATION OF CULTURED CELLS

Protocol

1. Discard the culture medium from the LabTeks and remove the plastic frame of the wells. Wash the cells with PBS.

2. Fix cells by immersing the LabTeks in cold (-20°C) methanol for 5 minutes, followed by cold (-20°C) acetone for 2 minutes.

3. Air dry the slides. These can be either used immediately for immunocytochemistry, or stored at -20°C until required.

2.6.4. IMMUNOCYTOCHEMICAL STAINING

In order to ascertain that the cultures contained predominantly RPE cells, and not (for example) contaminating fibroblasts, the cells were examined immunocytochemically for cytoskeletal components that are characteristic of the RPE (see Results for more detailed description).

The immunocytochemical staining technique used was peroxidase-conjugated streptavidin immunostaining. In this indirect method, the secondary antibody reacts with the primary antibody bound to tissue antigen. The secondary antibody is biotinylated and the intensity of colour reaction, determined by the addition of a substrate-chromogen solution, is examined. Introduction of a secondary antibody increases the versatility of the staining method (a labeled secondary antibody can be used with different primary antibodies from the same species) and enhances sensitivity (several secondary antibodies can react with different epitopes on the primary antibody).

The antibodies used are presented in Table 2.4. All primary antibodies were monoclonal, derived from hybridomas produced by the fusion of mouse myeloma cells and splenocytes from an immunized mouse against intermediate filament (IF) proteins. Secondary antibodies were directed against mouse immunoglobulins. Primary and secondary antibodies were diluted to their appropriate titre with 1% NGS (normal goat serum) in PBS. The specificity of cytokeratin antibody used was investigated and reported by McKechnie *et al.* (1988).

Table 2.4. Antibodies used in indirect immunoperoxidase assay.

Antibody	Source	Dilution	Antigen
Mouse anti-cytokeratin(CK)8.13	Sigma	1:100	CK10, 11, and 18 (cells of epithelial origin)
Mouse anti-vimentin	Dako	1:100	Vimentin (cells of mesenchymal origin)
Mouse anti-glial fibrillary acidic protein (GFAP)	Sigma	1:400	GFAP (astrocytes specific IF protein)
Biotinylated-conjugated goat anti-mouse	Sigma	1:300 1:150	mouse immunoglobulins

Staining procedure

1. Incubate specimens in 0.3% hydrogen peroxide solution for 20 minutes, followed by wash in PBS. Block non-specific background staining by incubating the slides for 15 minutes in 1% NGS in PBS (pH7.5).

2. Apply optimally diluted primary antibody and incubate for 45 minutes. Add 1% NGS in PBS to a control well.

3. Wash the specimen for 5 minutes with 3 changes of PBS.

4. Add secondary antibody (biotinylated goat anti-mouse) and incubate for a further 45 minutes.

5. Apply peroxidase-conjugated streptavidin to the biotinylated secondary antibody and incubate for 30 minutes. Wash in PBS and incubate for 10 minutes in 0.3% H_2O_2 and 1mg/ml DAB (3,3' - diaminobenzidine tetrahydrochloride) in PBS. A dark brown, insoluble polymeric oxidation product appears associated with the antigenic sites. Dehydrate through graded alcohols, xylene, and mount in DPX.

2.7. ISOLATION AND CHARACTERIZATION OF RNA

Messenger RNA (mRNA) molecules contain the coding information for the synthesis of proteins. They are the products of transcription of genes followed by nuclear processing, including capping, polyadenylation and splicing of the primary gene transcripts. Since transcription and processing are the initial steps in gene expression, the nature and amount of mRNA molecules determine the differentiated and physiological status of the cell. Complementary DNA (cDNA) comprise the reverse-transcribed genetic information of mRNA, i.e. exclusively the coding regions of genes and their associated 5' and 3' untranslated regions. Hence, the study of gene expression and characterization of transcription pattern in any cell through analysis of cDNA requires, ideally, the quantitative isolation of undegraded mRNA.

A single typical sized mammalian cell has approximately 400,000 mRNA molecules, of about 10,000 - 20,000 different mRNA species present in different abundances (Alberts *et al.*, 1989). The actual amount of mRNA depends on the cell type and physiological state of the cell. The abundant messages can be present in over 10,000 copies per cell, thus each of these accounting for about

3% of the total mRNA. However, up to several thousand species of mRNA are present at low abundance with copy number less than 15 molecules per cell. These "rare" messages are believed to represent about 30% of the total mRNA population.

Moreover, the vast majority of the cellular RNA is represented by tRNA and rRNA molecules, and only 1-3% by mRNA. Due to this low fraction of mRNA in the total cellular RNA, enrichment for mRNA (achieved by removing the rRNA and tRNA from a total RNA preparation) is desirable if not essential for construction of cDNA libraries. Enrichment is usually achieved by selection of mRNA via its poly A tracts.

2.7.1. PREPARATION OF POLY A+ MESSENGER RNA FROM RPE CELLS

In this study mRNA was isolated directly from RPE cells without prior isolation of total cellular RNA. Two commercially available kits were used for this purpose, Oligotex-dT™ Direct mRNA kit from Qiagen and FastTrack™ kit from Invitrogen. Both procedures take advantage of the poly A tail of mRNAs and achieve isolation of poly A+ mRNA by hybridization with oligo-dT molecules. The main difference is in the solid phase matrix to which oligo-dT primers are coupled: polystyrene latex particles in the case of Oligotex and cellulose for the FastTrack kit. Since high salt concentration is necessary to allow hybridization, the poly A+ mRNA can subsequently be released by lowering the ionic strength which destabilizes dt:rA hybrids.

Other critical parameters for direct mRNA purification from cells are the lysis and homogenization conditions. These must ensure an efficient disruption of cellular membranes, denaturation of RNA/protein complexes, reduced lysate viscosity, and protection from endogenous ribonucleases released upon cell disruption. Oligotex Direct mRNA kit uses a lysis buffer containing guanidinium isothiocyanate and an SDS-containing wash buffer to create and maintain RNase-free conditions, while homogenization is achieved by passing the lysate through a QIAshredder spin column. The FastTrack procedure utilizes a detergent-based lysis buffer containing a RNase/Protein Degrader, followed by homogenization of cell lysate by repeated passing through a 18-21 gauge needle.

The respective manufacturer's protocol was used to isolate mRNA from up to $1x10^7$ RPE cells. mRNA was eluted in 5mM Tris-HCl (pH 7.5). Quantitation was performed as described in 2.7.2. and the integrity of mRNA was assessed by Northern blotting as described in 2.7.3.

Throughout isolation and handling of RNA, the RNase-free solutions and environment were created and maintained by DEPC (diethyl pyrocarbonate) treatment (added to a concentration of 0.2% to solutions, followed by an overnight incubation at 37°C and autoclaving) and, respectively, by direct application of RNase AWAY (Molecular *Bio*-Products, USA).

2.7.2. QUANTITATION OF mRNA

The concentration and purity of mRNA were determined by measuring the absorbance of a diluted sample at 260nm (A_{260}) and 280nm (A_{280}), in RNase-free silica cuvettes (500µl and 150µl), using a PYE UNICAM SP6-400 UV spectrophotometer and, respectively, an ATI UNICAM UV/VIS spectrometer UV2. The concentration was estimated using the following formula, assuming that an absorbance of 1 at 260nm coresponds to 40µg of RNA:

concentration of mRNA (µg/ml) = 40 x A_{260}x dilution factor

Total yield of mRNA (µg) = concentration x volume of stock (ml)

The purity of the mRNA was estimated based on the ratio A_{260}/A_{280} and was regarded as satisfactory if it was over 1.8.

When mRNA needed to be recovered after dilution, it was precipitated with 0.15 volume of 2M sodium acetate and 2.5 volume of 100% ethanol, then frozen at -70°C and spun in a MSE microcentrifuge at maximum speed (13,000rpm) for 15 minutes.

2.7.3. NORTHERN BLOTTING

This technique was used to assess the integrity of mRNA (by hybridizing the Northern blot with a probe specific for a message known to be present in the RNA pool) and to determine the presence of specific mRNAs. The protocol is adapted from a modification of the method described by Sambrook *et al.* (1989).

2.7.3.a. Denaturing formaldehyde gel electrophoresis of RNA

Protocol

1. Melt 1 -1.5g of agarose in 73ml of sterile water and 10ml of 10x MOPS buffer (see Appendix I). Cool the melted agarose solution to 50°C.

2. Add 17ml of 37% (v/v) formaldehyde. Mix the agarose solution and pour it into the gel tray. Submerge the solidified gel in 1x MOPS running buffer.

3. Prepare the samples by adding 2 volumes of formaldehyde gel loading buffer (see Appendix I), boil for 2 minutes, chill on ice and then load the gel. *The final concentration of formamide in the samples is thus 50%.*

4. Run the gel at 100V for about 2 hours.

Usually when bromophenol blue has run half the distance of the gel, the RNA has migrated sufficiently to allow examination of the size distribution relative to markers. By cutting off the gel the portion corresponding to the size markers and staining it with ethidium bromide (2μg/ml), the position of the size markers can be visualized under UV illumination. Alternatively, the marker RNA bands can be stained by submersing the gel in 1% toluidine blue solution in 1% CH_3COOH for 15 minutes, followed by destaining of the background with several changes of 1% CH_3COOH.

2.7.3.b. Blotting into nylon membranes

Northern blots were produced by capillary blotting the gels into nylon-supported, hydrophilic membranes (Hybond™-N, Amersham).

Protocol

1. Remove the formaldehyde from the gel by soaking it in two changes of 20x SSC buffer (see Appendix I), for 20 minutes each time.

2. Set up the capillary blot in a tray filled with blotting buffer (20x SSC) by covering a platform with a wick made of Whatman 3MM filter paper saturated with blotting buffer.

3. Place the gel inverted on the wick and cover it with a Hybond-N membrane of the exact size. Remove any trapped air bubbles using a glass rod.

4. On top of the Hybond-N membrane, place three sheets of Whatman 3MM paper cut to size, first prewetted in 20x SSC, the others dry. Stack on top about 4-5cm of absorbent paper towels cut to size.

5. Place a glass plate and a 500-750g weight on top. Allow the transfer to proceed overnight.

6. When dismantling the apparatus, mark the membrane with pencil to allow later identification of lanes.

7. Fix the samples to the membrane by oven baking at 80°C for 2 hours.

2.7.3.c. Probe labelling by random priming

Double-stranded probes were labelled by random priming of DNA synthesis in the presence of (α-^{32}P)dCTP. Both random primer labelling kits used for this purpose (Rediprime DNA Labelling System, Amersham, Life Sciences and Ready-To-Go™ DNA Labelling Beads, Pharmacia Biotech) are based on priming DNA synthesis with random sequence oligodeoxynucleotides at numerous sites along a denatured DNA template (Feinberg & Vogelstein, 1983; Feinberg & Vogelstein, 1984). The primer-template complex is a substrate for the "Klenow" fragment of DNA polymerase I.

The labelling reaction was carried out following the manufacturer's protocol using 15-25ng of DNA template (linearized plasmid or specific restriction fragments) and 5µl (50µCi) of 10mCi/ml (α-^{32}P)dCTP (specific activity 3000Ci/mmol). The unincorporated label was removed with Centriflex™ Primer/Probe Purification (Advanced Genetic Tech. Corp.) microcentrifuge cartridges. The single-stranded (denatured) DNA binds to the polymer membrane in these cartridges in high salt conditions and is subsequently released in TE buffer (see Appendix I).

When necessary, the specific activity of the labelled DNA probe was calculated using the following formulae:

$$\textbf{Mass of DNA (ng)} = \frac{13.2\text{x}(\mu\text{Ci added})\text{x}(\%\text{incorporation})}{\text{Specific activity of } (\alpha\text{-}^{32}\text{P})\text{dCTP}} + \text{starting template (ng)}$$

$$\textbf{Specific activity (dpm/µg)} = \frac{(2.2\text{x}10^{4})\text{x}(\mu\text{Ci added})\text{x}(\%\text{incorporation})\text{x}10^{3}}{\text{mass of DNA (ng)}}$$

The % incorporation is given by the ratio of acid-precipitable counts (i.e., counts corresponding to labelled DNA probe) over total counts. The incorporated counts were determined by spotting a known volume of the reaction mixture on to a

glass fibre filter and washing it with 5% trichloroacetic acid (TCA) to precipitate DNA, followed by an ethanol wash and scintillation counting in a Packard TRI-CARB 1900 TR liquid scintillation analyzer (Canberra Packard, UK).

2.7.3.d. Hybridization and washing conditions

All hybridizations and subsequent washes in this study were carried out in bottles in a Stuart Scientific hybridization oven/shaker SI 20H, at 6 rev/min, respectively 45 strokes/min, at the temperatures indicated in the protocol below.

Protocol

1. Make up the hybridization solution (see Table 2.5).

Table 2.5. Hybridization solution for Northern blots (see Appendix I).

For 20ml:

Solution	Volume (ml)	Final concentration
20x SSC	2	2x SSC
10% (w/v) SDS	1	0.5% SDS
100x Denhardt's solution	1	5x Denhardt's solution
Formamide (deionized)	10	50% Formamide
1M Tris, pH 7.5	1	50mM Tris
0.5M EDTA	0.2	5mM EDTA
Distilled H_2O	4.8	
10mg/ml sonicated salmon sperm DNA (ssDNA)	0.1	50µg/ml ssDNA

Prior to being added to the hybridization solution, the ssDNA needs to be denatured by heating to 100°C for 5 minutes, and then chilled on ice.

ssDNA, as well as Denhardt's solution, act as blocking agents and inhibit non-specific hybridization.

2. Prewet the filter membrane in 2x SSC, 0.5% SDS and prehybridize it in 0.2ml hybridization solution/cm^2 membrane at 42°C, for 1-2 hours.

3. Denature labelled probe by heating to 100°C for 5 minutes and mix it with hybridization solution. The use of the total probe generated in a labelling

reaction in the conditions described in 2.7.3.c. in 10ml of hybridization solution gives consistently good results.

4. Decant the prehybridization solution and replace it with the prewarmed hybridization solution and the probe.

5. Hybridize overnight at 42°C - 65°C.

6. Following hybridization, wash membranes with repeated changes of 2xSSC - 0.1xSSC, 0.1% (w/v) SDS, at 42°C - 65°C with shaking. After washing, keep the membranes wrapped in plastic sheets.

The stringency conditions for hybridization and washes are given by the temperatures and concentration of solutions and are decided as a function of specificity and length of probe.

The same procedure was also used for the hybridization with cDNA probes and subsequent washes of Human RNA Master Blot™ (Clontech Laboratories, Inc., USA). This blot is a positively charged nylon membrane to which poly A+ RNAs from 50 human tissues have been immobilized in separate dots, along with several controls (Figure 3.14).

2.7.3.e. Detection of hybridization signal and removal of probes

The detection of hybridization signal was performed either by autoradiography of membranes using *RX* Fuji Medical X-Ray film, in cassettes with intensifying screens, at -80°C, or by phosphor beta imaging using a Bio-Rad Molecular Imaging BI Screen (Bio-Rad Laboratories, USA). The BI (Beta Imaging) screen contains microscopic storage phosphor particles which are very sensitive in detection of both visible light and high energy β-emitting isotopes as ^{32}P. After exposing the screen to such a sample, the latent image was recovered by scanning the screen with infrared light in the Molecular Imager™ Scanner. The digitized image was processed with the Molecular Analyst™ software.

In order to allow re-hybridization with new probes, the Northern blot membranes (which were never allowed to dry during or after hybridization and washing) were stripped off by pouring over them a boiling solution of 0.1% (w/v) SDS and allowing it to cool to room temperature.

2.8. cDNA LIBRARIES CONSTRUCTION

2.8.1. SYNTHESIS OF cDNA

In order to generate libraries, double-stranded cDNA was synthesized from RPE mRNA. The protocol used followed the guidelines accompanying the cDNA Synthesis Module (Amersham Life Science) and the Time Saver® cDNA Synthesis kit (Pharmacia Biotech) and, therefore, it is not reproduced here. Rather, only the principle of the method and the particular details of its application to this study are described.

The cDNA synthesis procedure is outlined schematically in Figure 2.2. The method is based on the principle developed by Gubler and Hoffman (1983) where first strand cDNA is synthesized with reverse transcriptase (RT): Avian Myeloblastosis Virus (AMV) RT (Amersham) or Moloney Murine Leukemia Virus (MMLV) RT (Pharmacia). The cDNA synthesis is primed either with oligo dT or random hexanucleotide $(dN)_6$ primers.The anchored $(dT)_{25}$ is a mixed population of three primers, $(dT)_{25}$with dA, dC or dG at the 3' end. The 3' base ensures that the primers bind at the junction of the poly(A) tail and the mRNA sequence, thus avoiding transcription of long poly(A) streches. Priming with anchored $(dT)_{25}$ primers ensures the recovery of sequences near the 3' end of the poly(A) mRNA. The use of random hexanucleotide primers maximises the representation of the coding region of mRNA and produces shorter fragments representing various parts of each mRNA molecule. Also, by providing alternative priming sites (which may be less affected by secondary structure, for example) it may allow a more comprehensive representation of the mRNA sequence. The average length of the cDNAs depend on the relative concentrations of the primers and the mRNA. The mRNA strand of the mRNA-cDNA hybrid is nicked by ribonuclease H (RNase H) in a temperature-controlled reaction (proceeding at 12°C). The resulting nicks are used by the DNA polymerase I which replaces RNA with DNA by nick translation (Kornberg, 1992). Finally, the exonuclease activity of T4 DNA polymerase converts any protruding 3' and 5' termini in the double stranded (ds) cDNA to blunt ends.

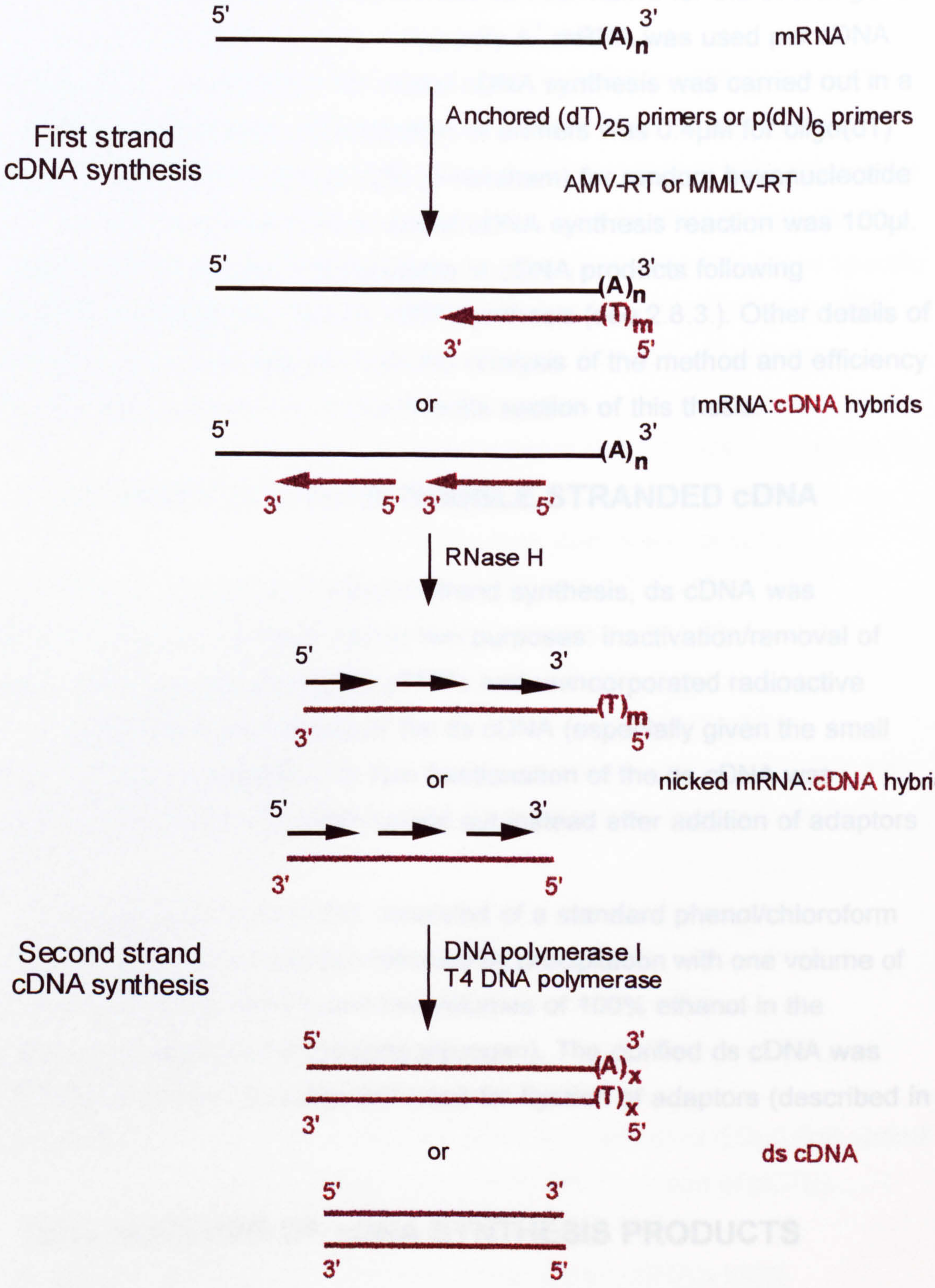


Figure 2.2. Schematic representation of double stranded cDNA synthesis.
The synthesis products when using either anchored (dT)25 or random hexanucleotide primers are indicated. See text for further explanation.

In this study, both priming strategies were used, as the cDNAs produced in each of these alternative approaches can be informative for the overall gene expression analysis. Between 0.5 - 1.2µg poly A^+ mRNA was used per cDNA synthesis reaction. A standard first strand cDNA synthesis was carried out in a total volume of 20µl and the concentration of primers was 0.4µM for oligo(dT) primers and 1µM (Pharmacia) or 7µM (Amersham) for random hexanucleotide primers. The volume of the second strand cDNA synthesis reaction was 100µl. The reactions were monitored by analysis of cDNA products following incorporation of radioactivity during cDNA synthesis (see 2.8.3.). Other details of variations are discussed together with the analysis of the method and efficiency of cDNA libraries construction in the Results section of this thesis.

2.8.2. PURIFICATION OF DOUBLE STRANDED cDNA

Following completion of second strand synthesis, ds cDNA was subjected to a purification step having two purposes: inactivation/removal of enzymes and removal of unreacted dNTPs and unincorporated radioactive label. In order to avoid any losses of the ds cDNA (especially given the small amounts of material available), no size fractionation of the ds cDNA was performed at this stage. This was carried out instead after addition of adaptors (see 2.8.4.).

The purification of ds cDNA consisted of a standard phenol/chloroform (1:1 v/v, see Appendix I) extraction followed by precipitation with one volume of 4M ammonium acetate pH 5.8 and two volumes of 100% ethanol in the presence of glycogen (1µl of 20mg/ml glycogen). The purified ds cDNA was resuspended in 5-10µl TE buffer and used for ligation of adaptors (described in section 2.8.4.).

2.8.3. ANALYSIS OF cDNA SYNTHESIS PRODUCTS

Examination of the efficiency of the cDNA synthesis reactions in terms of both yield and product size distribution was carried out by incorporating (α-^{32}P) dCTP (ICN) into both cDNA strands and subsequent determination of acid-precipitable counts and analysis by gel electrophoresis.

2.8.3.a. Incorporation of radioactivity during cDNA synthesis

Incorporation of (α-^{32}P) dCTP during cDNA synthesis and measurement of the incorporated radioactive label were performed as recommended in the Amersham's protocol of cDNA Synthesis Module. The labelling was achieved in separate reactions, i.e.-first strand and second strand labelled independently, and in sequential reaction, with the first strand being labelled to a lower specific activity than the second. The former method of labelling, with 10µCi of (α-^{32}P) dCTP added to each reaction, was used to monitor the yield of synthesis reaction (see below) as well as to vizualize, by agarose gel electrophoresis, separately the products of first and second strand cDNA synthesis. However, in order to avoid dividing the starting amount of mRNA, sequential reaction labelling was used more often to assess the size distribution of cDNA synthesis products by gel electrophoresis. In this case, 5µCi of (α-^{32}P) dCTP were included in the first strand cDNA synthesis reaction and 25µCi of the same label were added to the second strand synthesis.

The radioactivity incorporated in each strand was estimated by determining acid-precipitable counts, similarly as described in section 2.7.3.c. The amount and yield of each strand of cDNA synthesized were calculated using the following formulae:

Amount of cDNA synthesized (ng) = 350 x 4 x Z% x 10 where:

350g = the average molecular weight of 1 mole of dNMP

4 - accounts for all dNTPs

Z% = % incorporation in each strand

10nmol is the amount of unlabelled dCTP in a standard (20µl) first strand and (100µl) second strand reaction (2µl of 5mM solution of dCTP).

% of mRNA transcribed = $\frac{\text{(Amount of first strand cDNA)x100\%}}{\text{(Amount of mRNA used)}}$

% of second strand cDNA transcribed from the first = $\frac{\text{(Amount of second strand cDNA)x100\%}}{\text{(Amount of first strand cDNA)}}$

2.8.3.b. Analysis of cDNA by alkaline agarose gel electrophoresis

The radioactively labelled first and second strand cDNA synthesis products were subjected to gel electrophoresis in order to be visualized by autoradiography. Prior to electrophoresis, the samples were purified (see 2.8.2) to remove the unincorporated radioactive label, as this could obscure the detection of small cDNA fragments and resuspended in 1x alkaline electrophoresis buffer (Appendix I).

Protocol

1. Prepare a 1% agarose gel in 1x alkaline electrophoresis buffer by first dissolving the agarose in 90% of the total volume of water and then, after cooling at about 55°C, adding 10% of the final volume of 10x alkaline electrophoresis buffer (Appendix I). Mount the set gel in the electrophoresis tank.

The alkaline electrophoresis buffer ensures denaturing conditions (50mM NaOH in 1x electrophoresis running buffer).

2. Before loading, add 5x alkaline loading buffer (Appendix I) to the cDNA samples in the proportion 1 part buffer:4 parts samples.

3. Carry out the electrophoresis at voltages of up to 7.5V/cm gel until the dye has migrated approximately a third of the gel's length.

As bromophenol blue fades during the running of the gel, it becomes difficult to see it after about an hour. Running a 13cm gel at 100V for approximately 2 hours gives good separation. Since alkaline gels draw more current than neutral ones at the same voltage (having a lower electrical resistance), considerably more heat is generated during this type of electrophoresis.

4. Cover the gel with cling film and dry under vacuum on a gel drier.

If unlabelled markers are used, cut the corresponding lanes before drying the gel, soak in 1M Tris-HCl, pH 7.6 and 1.5M NaCl for 45 minutes and then stain with ethidium bromide.

5. Detect the cDNA by autoradiography (see 2.7.3.e.).

2.8.4. ADDITION OF EcoRI ADAPTORS AND SIZE FRACTIONATION OF cDNA

To prepare the cDNA for insertion into the EcoRI site of a λ phage-based cloning vector, the blunt-ends of the ds cDNA were converted to EcoRI cohesive termini through ligation with EcoRI adaptors. The size fractionation of cDNA was performed by spun column purification at the same time with the purification of "adapted" cDNA from unreacted adaptors. Then, the 5'-hydroxyl (OH) groups of "adapted" cDNA were converted to 5' phosphate (P) groups in the presence of T4 polynucleotide kinase.

Two types of EcoRI adaptors were used in this study. Both adaptors are blunt at one end and carry a 4-base non-phosphorylated EcoRI overhang at the other end. Their design (shown in Figure 2.3.) is different with regard to length, internal restriction sites and phosphorylation of 5' blunt end.

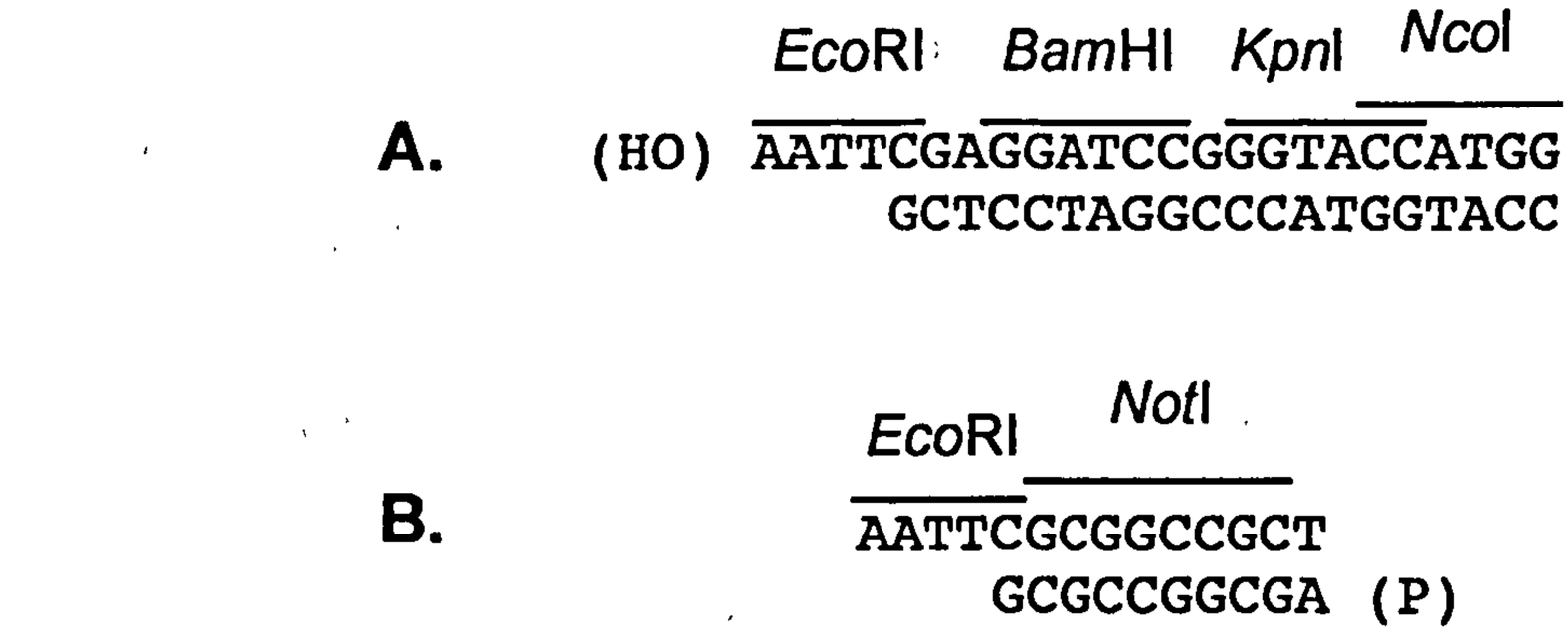


Figure 2.3. Design of adaptors used.
The adaptors are composed of two oligonucleotides with lengths of 24 and 20 bases (**A**.-Amersham), respectively 14 and 10 bases (**B**.-Pharmacia Biotech). Internal sites of restriction enzymes are indicated.

With both adaptors the respective manufacturer's protocol was followed. When using Amersham's adaptors, which have the 5' blunt end unphosphorylated in order to avoid adaptor self-ligation, the ligation reaction was performed in ATP-containing buffer and in the presence of a 100-fold molar excess of adaptors. With Pharmacia's adaptors the molar ratio adaptors:cDNA was of about 50:1.

The column purification and size fractionation of "adapted" cDNA was performed in SizeSep™-400 spun columns with Sepharose®CL-4B (Pharmacia Biotech) according to manufacturer's protocol. The columns were centrifuged in a Sorvall HB-4 swing-out rotor at 1550rpm (400xg). Following kinasing, a phenol/chloroform extraction (to remove all traces of the kinase) and an ethanol precipitation (to concentrate the cDNA solution) were performed prior to ligation with dephosphorylated EcoRI digested vector arms (see 2.8.5.).

2.8.5. LIGATION OF "ADAPTED" cDNA INTO CLONING VECTORS

Two λ phage-based vectors were used in this study to clone the RPE cDNA: λMOS*Elox* (Amersham Life Science) and λTriplEx™ (Clontech Laboratories, Inc.). Both vectors feature an inbuilt subcloning system for conversion of phage recombinants to plasmids *in vivo*. The multiple cloning sites (MCS) are located within a plasmid which is embedded in a λ phage genome and flanked by two 34bp *loxP* sites derived from bacteriophage P1 (Figures 2.4. and 2.5.). Transduction of recombinant phages into a *E. coli* strain expressing the P1 *cre* recombinase (such as BM25.8 (Amersham) or BNN132 (Clontech)), which recognizes the *loxP* sites, promotes *cre* recombinase-mediated release and circularization of a plasmid at the *loxP* sites. The cDNA clones could be in this way obtained in a plasmid vector (pMOSElox, respectively pTriplEx) that contains the ampicillin-resistance gene for selection and the origin for autonomous replication in *E. coli*. This system does not require a helper phage (see section 2.8.9 for the protocol for *in vivo* excision).

Both of these vectors allow screening of the corresponding cDNA libraries using both nucleic acid hybridization methods and expression screening with antibodies. As in other λ vectors, proteins are expressed fused with a short peptide encoded by the vector: T7 major capsid protein (encoded by *T7 gene 10*), respectively α-polypeptide of β-galactosidase (*lacZ'*). Although it was not used in this study, a feature of λTriplEx vector is notable: the vector incorporates two sets of translation initiation signals (including two ATG start codons) in different reading frames, followed by a transcription/translation slip site; this allows each insert to be translated in all three reading frames.

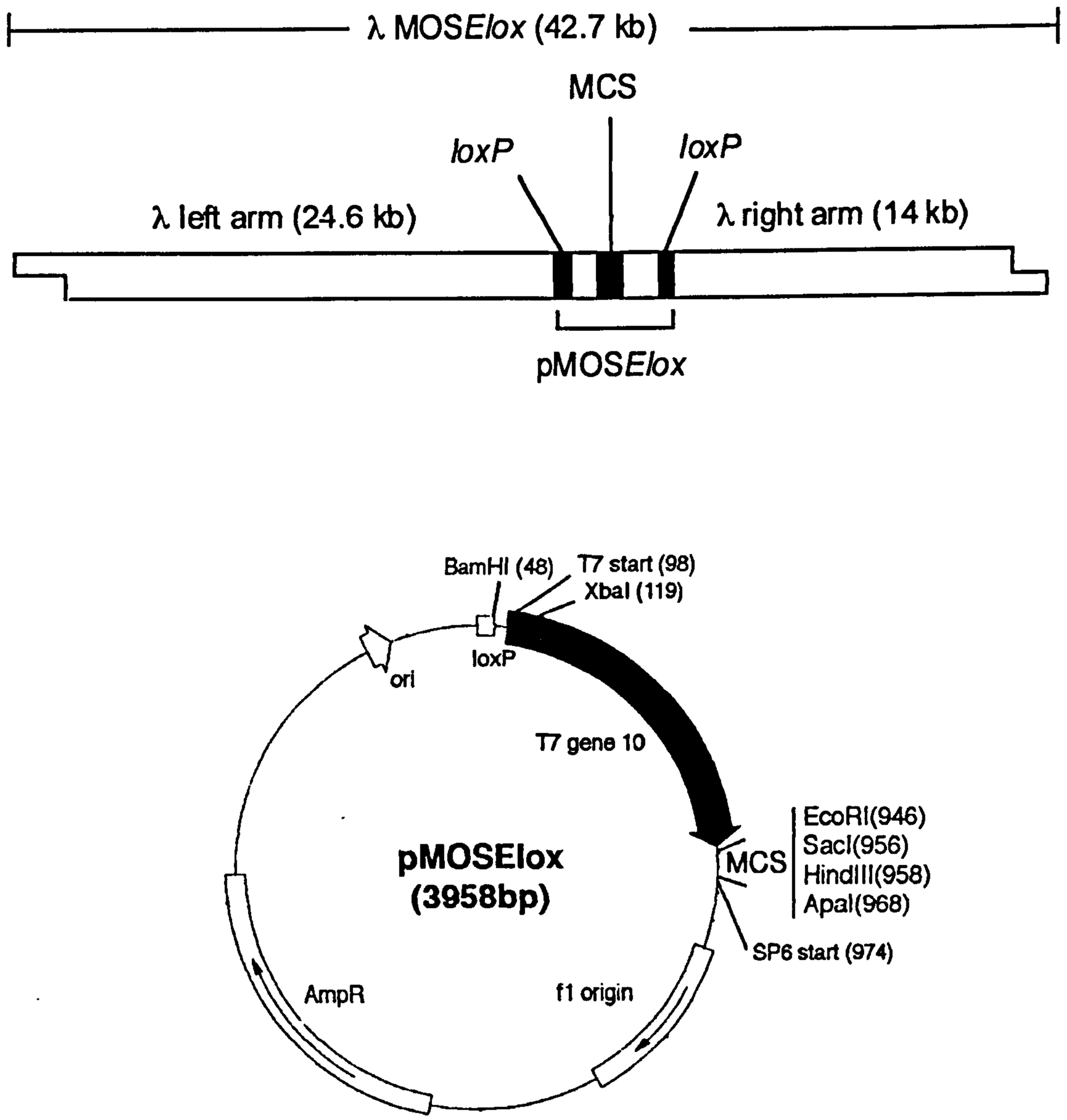


Figure 2.4. Map of λMOS*Elox* and pMOS*Elox*.
The MCS is located downstream from bacteriophage T7 transcription and translation signals and is followed by translation stop codons in all three reading frames, an SP6 promoter and a T7 transcription terminator. Some of the unique restriction sites are indicated.
See text for further details.

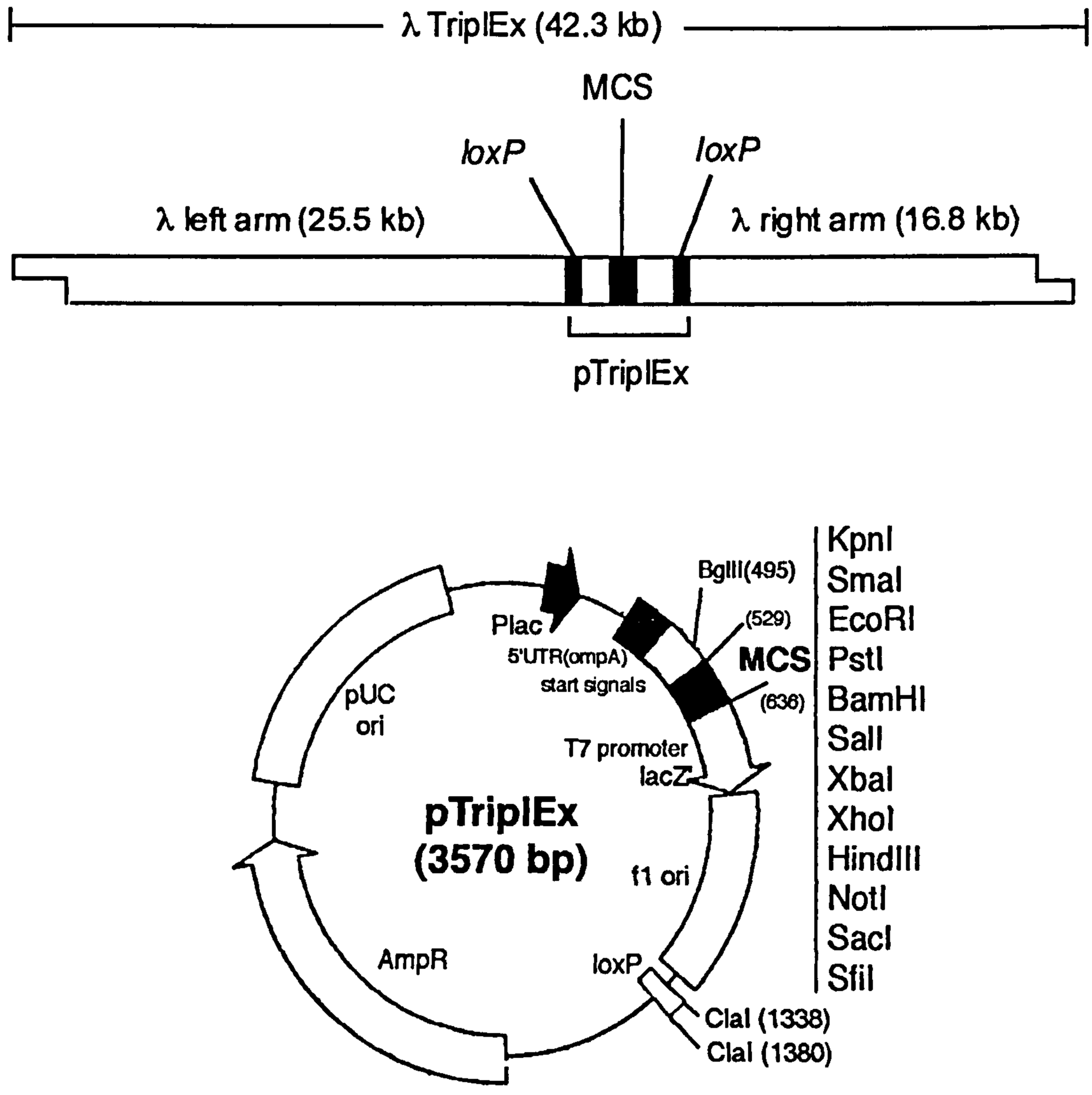


Figure 2.5. Map of λTriplEx and pTriplEx.

The *E. coli lac* promoter and operator provide regulated expression of inserts in *E. coli* hosts expressing the *lac* repressor (*lacI*^{q}). The MCS is embedded within the *lacZ* α-peptide allowing clones with inserts to be identified by blue/white screening. The 5' untranslated region (UTR) from the *E. coli omp A* gene stabilizes the mRNA. A T7 RNA polymerase promoter is located downstream of the MCS. Some of the unique restriction sites are indicated. See text for further details.

The ligation of EcoRI-"adapted" cDNA to EcoRI-digested, dephosphorylated λ vector arms was performed according to the manufacturer's protocol in the presence of T4 DNA ligase at 16°C overnight. While the concentration of vector arms was constant in reaction (at 1μg/10μl), different molar ratios of vector to cDNA were tried. The details of these ligation reactions are discussed together with the analysis of efficiency of cDNA library construction in the Results chapter.

2.8.6. *IN VITRO* PACKAGING OF LIGATION PRODUCTS

The whole of each 5μl or 10μl ligation reaction was packaged *in vitro* using Amersham's λ-DNA *in vitro* packaging cell extracts and protocol. The cell extracts provided are derived from two induced λ lysogens whose prophages carry different, but complementing, mutations in the genes required for assembly of mature phage particles. The recombinant λ DNA is packaged by these extracts into infectious phage particles which can then be introduced into bacterial host cells by infection. The collection of packaged phages thus obtained represents the unamplified cDNA library which can be stored as a normal phage stock in SM buffer (see Appendix I) for a few months at 4°C (total volume 500μl). The addition of a small amount of chloroform inhibits the growth of bacteria.

2.8.7. PLATING AND TITRATION OF λ PHAGE RECOMBINANTS. ESTIMATION OF RECOMBINATION EFFICIENCY

The following protocol was used to plate the λ phage-based cDNA libraries and to determine the titre of both unamplified and amplified cDNA libraries, as well as the background titre of vector alone.

Protocol

1. Prepare duplicate dilution series of each of the packaging reactions (or amplified cDNA libraries) in SM buffer (Appendix I).

As a general guideline, appropriate dilutions for an unamplified λ lysate are 1:5 to 1:20 and for amplified libraries 1:1,000 to 1: 10,000. For obtaining accurate phage titres it is important to use a fresh tip for each dilution.

2. Add an appropriate volume of each phage dilution to 100µl of *E. coli* phage plating cells (ER1647 for λMOS*Elox*-based phages and XL1-Blue for λ TriplEx-derived phages; see Table 2.2 for bacterial strains genotypes) prepared as described in 2.4.3. Allow the phage to adsorb at 37°C for 15 minutes.

3. Melt LB-top agar (see 2.3.1) and place in a 45°C water bath; convert it to M-top agar by adding maltose to a final concentration of 0.4%.

The maltose induces phage receptors on the bacterial cell surface.

4. To each plating mix add 4ml of liquid M-top agar at 45°C. Mix quickly and immediately pour on to a dry LB agar plate. After setting, incubate the plates inverted at 37°C for 12-16 hours.

5. Count the number of plaques (plaque-forming units, pfu) and calculate the titre of phage (pfu/ml):

$$\textbf{pfu/ml} = \frac{\text{number of plaques x dilution factor x } 10^3 \text{ µl/ml}}{\text{µl of diluted phage plated}}$$

Cloning efficiences are expressed either as pfu/µg λ vector arms or as pfu/µg insert DNA.

When using λTriplEx, the percentage of recombinant clones can be assessed by performing blue/white (*lacZ* α-complementation) screening in *E. coli* XL1-Blue. Follow the same procedure as described above, adding to the M-top agar 50µl of 0.1M IPTG and 100µl of 20mg/ml X-Gal stock solutions (Appendix I). Colour detection requires overnight incubation. The ratio of white (recombinant, since insertion of a cDNA clone disrupts the *lacZ* gene) to blue (non-recombinant) plaques gives an estimate of the recombination efficiency.

2.8.8. AMPLIFICATION AND STORAGE OF cDNA LIBRARIES

cDNA libraries in λ vectors need to be amplified in order to make a large, stable high-titre stock of the library. However, more than one round of amplification is not desirable, since slower growing clones may become underepresented. The number of plates required depends on how many independent clones are in the library to be amplified.

Protocol

1. Mix aliquots of the library suspension containing approximately 50,000 plaque-forming bacteriophages with 500µl of host cells prepared as described in 2.4.3.

2. Incubate in a 37°C water bath for 15 minutes.

3. Add to each tube 7 ml of M-top agar (see 2.8.7.) at 45 °C. Mix quickly and spread evenly on to 150mm plates of LB agar. Allow the top agar to harden at room temperature.

4. Incubate the plates at 37°C for 10-14 hours. Do not allow the plaques to get larger than 1-2mm.

5. Overlay the plates with 10ml of SM buffer. Allow the phages to diffuse into SM buffer overnight at 4°C and then for an hour at room temperature (preferably on a platform shaker at 30-40 strokes/min).

6. The pooled λ phage suspension represents the amplified library lysate. Lyse any bacterial cells by adding chloroform to 5% final concentration, mix well and incubate at room temperature for 10 minutes. Clear the lysate of cell debris by centrifugation at 7,000rpm for 10 minutes.

7. Recover the supernatant (amplified library) and add chloroform to a 0.3% final concentration. In this form the library can be stored at 4°C for a few months without a significant decrease of its titre. For long-term storage, add dimethylsulfoxide (DMSO) to a final concentration of 7% and store at -70°C.

2.8.9. *IN VIVO CRE*-MEDIATED SUBCLONING OF RECOMBINANT PLASMIDS

Conversion of recombinant phages to plasmid subclones was performed on individual positive (i.e., with informative-sized insert) plaques, following screening for inserts (accomplished by PCR with vector specific primers, as described in section 2.10 of this chapter). For this purpose, λMOS*Elox*- and λTriplex-based recombinant phages were transduced into *E. coli* non-expressing strains BM25.8 and BNN132 respectively, which provided the necessary *cre*-recombinase activity. The plasmid-bearing colonies were selected on LB-carbenicillin (50µg/ml) or LB-ampicillin (50µg/ml) agar plates.

Protocol

1. Pick a single, isolated colony of the host strain from the stock plate and use it to inoculate 10ml of LB supplemented with 10mM $MgSO_4$ and 0.2% maltose. Grow the culture until it reaches an OD_{600} of 0.5.

Strain BM25.8 requires about 5-6 hours at 37°C (with shaking at 200rpm) to reach the necessary cell density; strain BNN132 requires a further 1 hour incubation at 31°C to express the necessary cre-recombinase activity.

2. Mix an appropriate dilution of phage with 100µl of host cells.

Knowing the phage titre and assuming a 10-20% efficiency of the excision, calculate the dilution/volume of the phage eluate needed to obtain around 500 antibiotic resistant colonies (cfu)/plate.

3. Incubate the cell/phage mixture at 37°C (31°C for BNN132 strain) for 30 minutes, without shaking, to allow the phage to adsorb to the host cells.

An additional 1 hour incubation at 31°C with shaking (200rpm) increases the efficiency of excision for λTriplex/BNN132 system.

4. Spread the cell/phage mixture on the surface of LB agar plates supplemented with carbenicillin or ampicillin. Incubate the plates inverted at 37°C overnight.

When using carbenicillin, fewer "satellite" colonies are observed.

5. Count the number of colonies and calculate the efficiency of subcloning. Well-isolated colonies can be picked from this plates and use to prepare plasmid DNA (see 2.12.1).

2.9. HUMAN FOETAL RPE cDNA LIBRARY

An aliquot (100µl) of a human foetal RPE cDNA library used in this study was a gift from Dr. Dean Bok of UCLA, USA. It was derived from a one-round amplification of a commercially constructed cDNA library starting from mRNA isolated from cultured human foetal RPE cells. The cell culture had been initially set up with RPE cells from 17-22 weeks aborted human foetuses and maintained for 2 months until the confluent cells reached a dome-like appearance. The RPE cells from 12 culture flasks (150cm^2) were harvested and 90µg of $polyA^+$ mRNA was isolated from them. This represented the starting material for cDNA library construction.

The first strand cDNA synthesis was primed with an oligo(dT) linker-primer containing an internal XhoI site. EcoRI adaptors with the following sequence were used:

5′ AATTCGGCACGAG 3′

3′ GCCGTGCTC 5′

The cDNA library was directionally constructed in Uni-ZAP™ XR vector (Stratagene), using *Eco*RI and *Xho*I cloning sites upstream and, respectively downstream of the cDNA insert. Thus cDNA was inserted into the vector in a sense orientation with respect to the *lacZ* promoter.

The Uni-ZAP™ XR vector is suitable for screening with either DNA probes or antibody probes and allows *in vivo* excision of the pBluescript® phagemid (Figure 2.6), allowing the insert to be characterized in a plasmid system. The amino terminus of the *lacZ* gene provides α-complementation for blue/white colour selection of recombinant phagemids, while its promoter allows fusion protein expression with the β-galactosidase gene product.

The *E. coli* strain XL1-Blue MRF' (Table 2.2) used for amplification of library, blue/white colour selection of recombinants, screening, and *in vivo* excision of phagemids contains a F' episome (selectively maintained in the presence of tetracycline, due to *Tn10* gene) which serves several purposes. It contains a partial *lacZ* gene required for generation (by α-complementation) of a functional β-galactosidase, as well as the $lacI^q$ gene coding for the *lac* repressor, which blocks transcription from the *lacZ* promoter in the absence of inducer IPTG. Since some fusion proteins may be toxic to the *E. coli*, inhibition of their expression can potentially increase the representation of the library. The F' episome also contains the genes for expression of the bacterial F' pili, required for filamentous phage infection. This is important since the *in vivo* excision of a phagemid from a recombinant Uni-ZAP XR clone requires superinfection with a filamentous f1 helper phage. The nonsuppressing *E. coli* strain SOLR (Table 2.2) allows only the excised phagemid to replicate, and not the helper phage which contains an amber mutation.

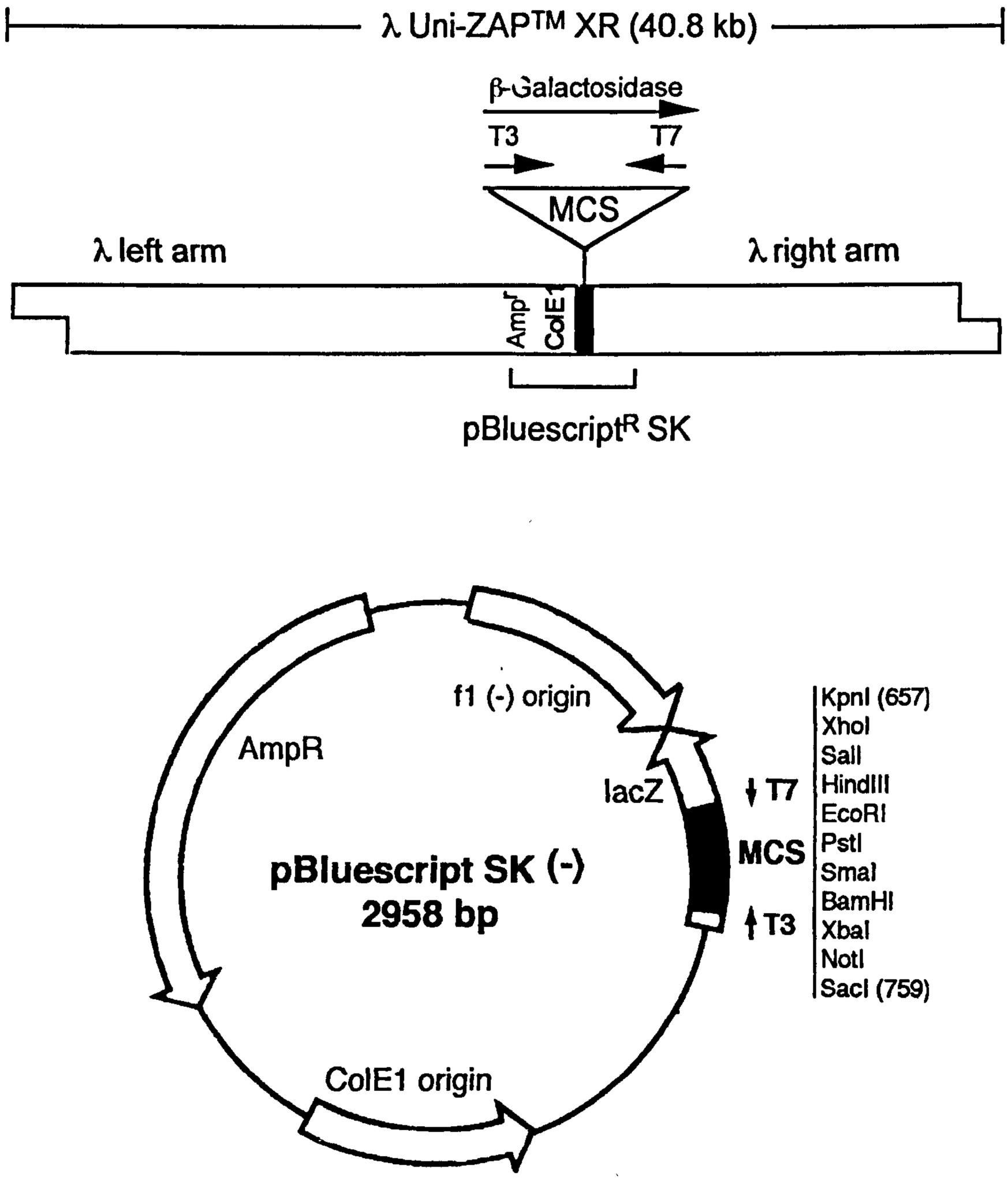


Figure 2.6. Map of the Uni-ZAP™ XR insertion vector and pBluescript®SK(-) phagemid.

The MCS (some of the unique cloning sites are indicated) is flanked by T3 and T7 RNA polymerase promoters and is orientated such that *lacZ* transcription proceeds from SacI to KpnI (this being indicated by the SK designation). The phagemid, which is derived from pUC19, has the ColE1 plasmid origin of replication and the filamentous phage f1 origin of replication ((-) indicates that the orientation of f1 origin is such that it allows recovery of the antisense strand of the *lacZ* gene in a host strain co-infected with helper phage).

Plating and titration of λ Uni-ZAP XR recombinants were performed using the *E. coli* strain XL1-Blue MRF', following the protocol described in section 2.8.7 and Stratagene's recommendations contained in the instruction manual of the ZAP-cDNA® Synthesis Kit. Recombinant pBluescript plasmids were *in vivo* excised according to manufacturer's protocol using ExAssist™interference-resistant helper phage with *E. coli* SOLR strain. Bacterial colonies of *E. coli* SOLR containing recombinant pBluescript-based plasmids were selected on LB-ampicillin agar plates (50µg/ml). The DNA screening protocol is described in section 2.11.

2.10. ANALYSIS OF RECOMBINANT CLONES BY PCR

The insert size of recombinant phages was evaluated by means of PCR using vector specific primers. This method was used to amplify insert sequences present in isolated phage clones and to assess the distribution of insert sizes on the whole cDNA library.

2.10.1. PRIMERS USED FOR AMPLIFICATION OF cDNA INSERTS

Primers used for amplification of cDNA inserts anneal to sites directly flanking the cloning sites in the respective vector. Their sequence, manufacturer and concentration of stock solution are listed in Table 2.6.

2.10.2. PCR AMPLIFICATION OF PHAGE LYSATES

PCR amplification reactions were set up using 5µl of individual phage plaques (eluted in SM buffer, overnight at 4°C) or 5µl of an appropriate dilution of whole cDNA library phage lysate. The total volume of each PCR reaction was 50µl and the final concentrations of the components in the reaction mix were: 0.1µM-0.2µM each primer (see Table 2.6), 200µM dNTPs (Pharmacia), 2mM Mg^{2+}, 1xGeneAmp® PCR buffer (Perkin-Elmer Ltd.), and 1.25U AmpliTaq® DNA polymerase (Perkin-Elmer Ltd.).

Table 2.6. Primers used for amplification of cDNA inserts.

Primer/ Sequence	**Vector specificity**	**Source/ Concentration of stock**
T7 gene 10 primer (19-mer) 5'TGAGGTTGTAGAAGTTCCG 3'	λMOSElox	Amersham (10μM) Custom made*(11μM)
T7 terminator primer (19-mer) 5'GCTAGTTATTGCTCAGCGG 3'	λMOSElox	Amersham (10μM) Custom made*(11μM)
5' Amplimer (26-mer) 5'CTCGGGAAGCGCGCCATTGTGTTGGT 3'	λTriplEx™	Clontech Lab., Inc. (20μM)
3' Amplimer (28-mer) 5'ATACGACTCACTATAGGGCGAATTGGCC 3'	λTriplEx™	Clontech Lab., Inc. (20μM)
T3 primer (20-mer) 5'AATTAACCCTCACTAAAGGG 3'	Uni-ZAP™/ pBluescript	Custom made* (10μM)
T7 primer (22-mer) 5'GTAATACGACTCACTATAGGGC 3'	Uni-ZAP™/ pBluescript	Custom made* (4μM)

* The custom made primers were synthesized by Perkin-Elmer Ltd.

The amplifications were performed in thin-walled reaction tubes, in a Perkin -Elmer GeneAmp PCR System 2400 cycler with the following program: 94°C for 1 minute, followed by 35 cycles of (94°C, 1 minute; 52°C-68°C, 1 minute; 72°C, 2 minutes), and a final extension at 72°C for 3 minutes. Control reactions were performed substituting the phage stock aliquot with DNA from respective vectors without any insert and with insert of known size (positive controls), and with double-distilled water (negative control). A third of each PCR reaction was analyzed on a 1% agarose gel containing ethidium bromide (according to the protocol described in section 2.12.3).

2.11. HYBRIDIZATION OF FILTERS OF cDNA LIBRARIES WITH cDNA PROBES

In order to determine the frequency of a cDNA clone in a whole cDNA library, filters of the cDNA library analyzed were screened with radioactive probes of the respective clone.

2.11.1. PREPARATION OF FILTERS

The following protocol was used to transfer recombinant phage plaques from plates prepared according to the protocol described in section 2.8.7 above to nitrocellulose filters (type HA, Millipore).

Protocol

1. Chill the plates for 1-2 hours at 4°C to harden the top agar.

2. Number the nitrocellulose filter and place it onto the plate, avoiding trapping air bubbles. Mark the filter in asymmetric locations for orientation by stabbing through it and into the agar with an 18-gauge needle. Leave the transfer to proceed for 2 minutes. If making duplicate filters, allow the second one to transfer for 4 minutes. Store the agar plates to 4°C for use after screening.

3. After lifting treat the membrane as follows: immerse in a 1.5M NaCl, 0.5M NaOH denaturation solution for 2 minutes; then, neutralize in 1.5M NaCl, 0.5M Tris-HCl (pH 8.0) for 5 minutes; finally, rinse in 0.2M Tris-HCl (pH 7.5), 2xSSC buffer for 30 seconds. Blot briefly on Whatman 3MM paper.

4. Fix the DNA to the filter by oven baking filters at 80°C for 2 hours.

2.11.2. HYBRIDIZATION AND WASHES OF FILTERS. SIGNAL DETECTION

The cDNA probes used for hybridization were prepared and radiolabelled with (α-^{32}P)dCTP as described in 2.7.3.c. The equipment used was that described in the same section.

Protocol

1. Make up the prehybridization solution (Table 2.7).

2. Prehybridize the filters at 42°C for a minimum of 2 hours in 3ml of prehybridization solution/filter.

Wet the filters completely in prehybridization solution prior to stacking and placing them in hybridization bottles. Avoid trapping air between the filters and the bottle wall.

Table 2.7. (Pre)hybridization solution for double-stranded probe (see Appendix I for components). For 20ml:

Solution	Volume (ml)	Final concentration
10xPIPES buffer	4	2xPIPES buffer
10% (w/v) SDS	1	0.5% (w/v) SDS
Formamide (deionized)	10	50% formamide
Distilled H_2O	4.8	
10mg/ml sonicated salmon sperm DNA (ssDNA)*	0.2	100µg/ml ssDNA

**Boil the salmon sperm DNA for 10 minutes before adding it to the prewarmed (50°C) preybridization solution.*

3. Decant the prehybridization solution and replace it with fresh prewarmed hybridization solution mixed with the denatured labelled probe.

4. Incubate the filters in the hybridization solution overnight at 42°C.

5. Wash the filters in 1xSSC buffer, 0.1% (w/v) SDS first at room temperature for 15 minutes, then at 50-65°C for 30 minutes. Increase the stringency of washes using 0.1xSSC, 0.1% (w/v) SDS for 15 minutes at 50-65°C. *Use about 10ml of wash solution per filter each time and do not allow the filters to dry between washes. If the filters are kept and stored moist, it is possible to wash them further to reduce background signal.*

6. After washing, place the filters between two sheets of plastic wrap, tape them in place in a cassette with intensifying screens. Expose to X-ray film (Fuji XR) overnight at -80°C. After developing the film, align it with the filters and the plates to identify the positive plaques.

2.12. PLASMID DNA

2.12.1. PREPARATION AND DETERMINATION OF CONCENTRATION OF PLASMID DNA

Plasmid DNA was isolated using various procedures, as a function of the downstream applications. The bulk of such preparations were directed towards manual and automatic sequencing and were performed using Qiagen Plasmid

and Kristal™ (Cambridge Molecular Technologies Ltd.) kits. For other purposes (such as restriction digestions, transformations, PCR amplifications, probe generation by random primed-labelling), double stranded plasmid DNA was also isolated with QIAprep Plasmid and HYBAID Recovery™ Plasmid Mini Prep kits.

Although all procedures are based on the modified alkaline lysis method (Birnboim & Doly, 1979), they differ mainly with respect to the material to which plasmid DNA is absorbed: Qiagen anion-exchange/Kristal binding resin for the methods designed to obtain sequencing-grade DNA, vs. different silica-gel membranes used in the other procedures. In the anion-exchange resin-based method, plasmid DNA is absorbed to the resin under appropriate low salt and pH conditions. RNA, proteins and low molecular weight impurities are removed by a medium salt wash, while plasmid DNA is eluted in a high salt buffer and concentrated and desalted by isopropanol precipitation. When using Kristal DNA Binding Resin, plasmid DNA is eluted directly into a low salt buffer ready for use, without the need for ethanol precipitation. In silica-based methods separation of plasmid DNA from other cellular components is based on coprecipitation, in high salt concentration, of denatured proteins, chromosomal DNA and insoluble complexes containing salt and detergent, while the shorter plasmid DNA renatures and remains in solution. Silica-gel membrane retains the DNA which is eluted, following a series of washes, in Tris-HCl buffer or water.

In order to avoid problems associated with plasmid DNA purified from *endA*+ (endonuclease A+) *E. coli* strains (Schoenfeld *et al.*, 1997), such as low yields, degradation during long term storage, resistance to restriction enzyme digestion, high backgrounds in automated fluorescent sequencing systems, when plasmids were first obtained in such a strain (for example, in the case of *in vivo* excision of pMOS*Elox* recombinant plasmids in *E.coli* strain BM25.8) they were transformed (see section 2.12.2) in a *endA1* (mutated) bacterial strain (such as DH5λα) prior to purifying plasmid DNA for sequencing and storage. The slower growing strain XL1-Blue (*endA1*), suitable for obtaining DNA of good quality which works well for sequencing, was used for growing pBluescript-based plasmids excised from the human foetal RPE cDNA library.

In all cases the manufacturer's protocol was followed, starting for mini-preparations with 1.5ml of bacterial culture, grown overnight in LB/Amp (100µg/ml). Plasmid DNA was resuspended or eluted in 50µl TE buffer (pH 7.5),

10mM Tris-HCl pH 7.6 or distilled H_2O. The concentration of plasmid DNA in solution was determined spectrophotometrically by measuring OD_{260} and using the conversion 1 OD_{260} unit of double-stranded DNA = 50µg/ml.

2.12.2. TRANSFORMATION OF PLASMID DNA INTO *E. COLI* DH5α

The protocol below is a modification of Hanahan's method (1983) and was adapted from the method used in Dr. Andy Hoyt's laboratory, The Johns Hopkins University, Baltimore. The expected frequency of transformation is 10^7-10^8 colony forming units (cfu)/µg pBR322. For composition of media see section 2.3.2.

Protocol

To make competent cells:

1. Grow an overnight culture of the strain to be transformed in LB at 37°C.

2. Dilute the overnight culture 1:100 in SOB. Incubate at 37°C until OD_{600}=0.4-0.5.

3. Put the culture on ice for 15 minutes and then harvest the cells by centrifugation at 2500rpm, 12 minutes at 4°C.

4. Resuspend the cells in 1/3 initial culture volume (step 2 above) of FSB and place on ice for a further 15 minutes.

5. Centrifuge as above and resuspend the bacterial cells in 1/10 initial culture volume of FSB.

6. Add dimethyl formamide (DMF) to a concentration of 3.5%, mix gently and leave on ice for 5 minutes. Add an additional aliquot of DMF to reach a final concentration of 7% and leave on ice for a further 5 minutes.

7. Aliquot 200µl competent cell suspension per sterile Eppendorf tube. Although it is best to use these aliquots immediately (see transformation protocol below), they can be frozen on dry ice and stored for a few weeks at -70°C (the competency decreases to 10^5-10^6 after a month of storage).

To transform competent cells:

1. Add plasmid DNA, in a volume as small as is practical, to an aliquot of competent cells. Incubate on ice for 30 minutes. Set up a control with no DNA.

2. Heat shock at 42°C for 90 seconds, then place on ice for 1 minute.

3. Add 1ml of SOC medium to each tube and incubate at 37°C for 1 hour.

4. Spin down the cells (10 seconds at maximum speed in a microcentrifuge), pour out the supernatant and resuspend the cells in the drop left.

5. Plate the cells on selective medium and incubate overnight at 37°C.

2.12.3. DIGESTION WITH RESTRICTION ENZYMES AND AGAROSE GEL ELECTROPHORESIS OF PLASMID DNA

Plasmid DNA was digested with restriction enzymes acording to manufacturer's instructions. Restriction fragments were analyzed by agarose (0.8-1.5%) gel electrophoresis in 1xTBE buffer (see Appendix I) containing 0.2µg/ml ethidium bromide. The voltage applied was of about 5 volts/cm gel. Appropriate size markers were run at the same time.

2.12.4. PURIFICATION OF DNA FRAGMENTS FROM AGAROSE GELS

Extraction of DNA from agarose TBE gels was performed using Kristal™ Gelex Resin (Cambridge Molecular Technologies Ltd.) according to manufacturer's instructions. The recovered DNA fragments were used for digestion with restriction enzymes and for labelling with ^{32}P by random priming (2.7.3.c).

2.13. SEQUENCING OF cDNA CLONES

The first cDNA clones from RPE cDNA libraries in this study were sequenced manually. The procedure used is outlined in the first part of this section. Since an automated sequencing facility became available in the Department, the majority of cDNA clones were sequenced in this way. The characteristics of the automated sequencing reactions are summarised in the second part of this section.

2.13.1. MANUAL SEQUENCING

In 1977 two methods for DNA sequencing were described: the chemical cleavage method (Maxam & Gilbert, 1977) and the enzymatic dideoxy chain-termination method (Sanger *et al.*, 1977). The latter one is the most routinely used at present, having been improved in terms of speed, ease and quality of sequence generation.

2.13.1.a. Sequencing reactions

The enzymatic method of DNA sequencing is based on the ability of a modified T7 DNA polymerase (Tabor & Richardson, 1987; Tabor & Richardson, 1995) to extend a primer, hybridized to the template to be sequenced, until a chain-terminating nucleotide is incorporated. For each sequence determination, four reactions are set up containing all four deoxyribonucleotide triphosphates (dNTPs) and a limiting amount of one of the dideoxyribonucleotide triphosphate (ddNTP). As the incorporation of a ddNTP stops the chain elongation, in each of the four reactions the growing oligonucleotide is terminated selectively at the respective nucleotide (A, G, C or T). The resulting fragments, each with a common origin given by the primer used to initiate the DNA synthesis, but ending at a different nucleotide, are separated according to size by denaturing polyacrylamide gel electrophoresis. By incorporating a labelled deoxyribonucleotide into the growing chain the extension products can be visualized on autoradiographic film.

In this study sequencing of double-stranded DNA (plasmid DNA prepared as described in section2.12.1) was performed with Deaza ^{T7}Sequencing™ mixes and T7 DNA Polymerase (Pharmacia Biotech), in the presence of (α-^{35}S)dATP (400Ci/mmol, ICN) as label. The Deaza ^{T7}Sequencing™ mixes are designed to overcome the appearance of "band compressions" which are caused by the secondary structures formed by G residues. This problem is reduced by incorporating a modified G residue, which has either a substitution of a carbon for a nitrogen atom at position 7 (as in 7-deaza guanine), or, besides such a substitution, also a removal of the exocyclic 3-amino group of guanine (as in inosine). T7 DNA polymerase can efficiently use the relevant dGTP analogues

(c^7dGTP and, respectively, c^7dITP). Following annealing of the primer to the template, the sequencing method recommended by the manufacturer of the sequencing mixes and enzyme involves two stages. A "labelling" reaction in which only deoxynucleotides at limiting concentrations are used, causing random chain termination. Then, a "termination" reaction in which high concentrations of deoxynucleotides are used with one dideoxynucleotide. In the end, reaction products are denatured and run on polyacrylamide gels. The sequencing reactions were performed with primers listed in Table 2.6 at molar ratios primer:template between 5:1 and 10:1.

2.13.1.b. Polyacrylamide sequencing gels

To resolve fragments differing in length by only one nucleotide, as the products of the sequencing reactions, high resolution, thin, denaturing polyacrylamide/urea gels (50cmx20cmx0.4mm) were used.

Protocol

1. Siliconize the plates before use with dimethyldichlorosilane solution (BDH) and assemble with 0.4mm spacers. Prepare the wells for loading the samples with 20 tooth combs.

2. To prepare one gel (final concentration 8M urea and 8% acrylamide in TBE), mix 40ml of 8% acrylamide/Urea/TBE stock solution (see Appendix I) with 30µl TEMED (N,N,N',N'-tetramethylethylenediamine, Sigma) and 400µl 10% ammonium persulphate (stock freshly prepared).

3. Pre-run the gel for 30 minutes at about 1500V, 30mA (LKB 2197 power supply, LKB Instruments). Denature the samples by boiling for 2 minutes. Load 2µl aliquots from each sequencing reaction per well.

4. Run the electrophoresis in 1xTBE buffer in the same conditions as the pre-run, for 2 hours for a short run and 4.5 hours for a long run.

5. After electrophoresis dry the gels on 3MM Whatman paper under vacuum at 80°C for 45 minutes. Expose the dried gels for 24-48 hours at room temperature using Fuji RX film.

2.13.2. AUTOMATIC SEQUENCING

Inasmuch as the DNA sequencing is a highly repetitive process, it is very suitable for automation. Smith *et al.* (1986) have first developed a strategy for the automated sequence analysis of DNA based on the use of four different fluorophores: fluorescein, NBD, tetramethylrhodamine and Texas Red. Currently, however, the automation is limited only to the data collection part of the process. Only recently Amersham Pharmacia Biotech/ Molecular Dynamics have announced the development of a new DNA sequencing system that allows automation by multicapillary parallel separation of sequencing reaction products, thus eliminating the use of conventional polyacrylamide gels, manual sample loading and sample lane tracking.

The currently used equipment is based on the use of a fluorescent label, rather than a radioactive one. The label can be attached either to the nucleotides that are used in the reaction or to the primer. If a different fluorescent label is used for each reaction, all four sets of reaction products can be electrophoresed in a single lane. The reaction products are detected while the polyacrylamide gel is running and data are collected on line by automated fluorescent detection instrumentation. The automatic sequencing protocols are based on thermal cycle amplification of the template DNA, which provides several advantages over conventional sequencing strategies. Among these are the following: reduced amount of template required, no need for alkaline denaturation/ethanol precipitation of the template, increased stringency of primer hybridization, reduced problems associated with secondary structure.

Developments in instrumentation, enzymology and dye chemistry have all contributed in the last few years to the refinement of various technologies in this area. For example, the most recent change in the dye chemistry used for automatic sequencing involves the introduction of a energy transfer dye system (Ju *et al.*, 1995). A donor dye absorbs light at the output wavelength of the laser used in fluorescent sequencing instruments and transfers its energy to an acceptor dye, thus increasing notably the signal strengths.

In this study two automated DNA sequencing technologies were used: Perkin-Elmer ABI-PRISM™ (UK) and MWG-Biotech LI-COR (Germany).

2.13.2.a. Dye-terminator chemistry of sequencing reactions for Perkin-Elmer ABI 373 and ABI PRISM 377 systems

Most of the automatic sequencing reactions in this study made use of the Perkin-Elmer ABI 373 DNA Sequencer in the Department. A smaller number of clones were sequenced on the ABI PRISM 377 DNA Sequencer, in the School of Tropical Medicine, Liverpool. In both cases, the method used was that of cycle sequencing with AmpliTaq DNA Polymerase, FS (Fluorescent Sequencing) and dye-labelled terminators.

AmpliTaq DNA Polymerase, FS, developed by Roche Molecular Systems/Perkin-Elmer, is a highly purified recombinant Taq DNA polymerase with a very low 5'-3' nuclease activity (important for the control of background noise), which allows an efficient incorporation of ddNTPs. The Ready Reaction Premix supplied by Perkin-Elmer contains dye-labelled dideoxynucleotides and deoxynucleotides (with dITP substituted for dGTP) in a concentration optimized to give a distribution of the signal between 10 and 700+ bases. Each of the four dideoxy terminators is tagged with a different fluorescent dye (Table 2.8). Thus, the growing chain is simultaneously terminated and labelled with the dye that corresponds to that base. Since each dye emits light at a different wavelength when excited by laser light, the products of all four reactions can be distinctively detected in a single gel lane. The detection is done one colour at a time as the fluorescently labelled DNA fragments migrate past the detector (ABI 373) or simultaneously for multiple fluorescent emissions by separating the individual wavelengths onto a CCD (charge coupled device) camera (ABI PRISM 377). The complete removal of the unincorporated fluorescent ddNTPs (achieved by ethanol purification) is critical for the accurate detection of sequence data.

Table 2.8. Rhodamine-based dyes used in ABI 373 sequencing system.

Dye	Emission λ	Reaction	Colour in electropherogram
R110	531nm	G	black
R6G	560nm	A	green
6-TAMRA	580nm	T	red
6-ROX	610nm	C	blue

About 1µg (roughly 0.4 pmol) of double-stranded template (plasmid DNA prepared as described in 2.12.1) was used per sequencing reaction. The template quality and quantity were assessed in all cases by electrophoresis alongside standards of known concentration. Each sequencing reaction contained between 4-5 pmol of the appropriate primer (Table 2.6). The manufacturer's protocols for cycle sequencing, purification of extension products by ethanol precipitation and preparation/loading the samples were followed. In the case of templates for which a clear sequence was not produced by a first attempt of sequencing, modifications of the following nature were made to the amplification protocol: DMSO to a final concentration of 5% was added in the sequencing mix; alterations of the annealing temperature (with 2-5°C) and/or of the annealing/ extension time (with 1-2 minutes) were tried. Electrophoresis was carried out in 6% acrylamide gel (SequaGel 6, National Diagnostics). The electrophoresis raw data was collected and interpreted by the base calling algorithms of the DNA Sequencing Analysis 1.2.1 Software. The output files containing the analyzed data and sequence were transferred to the sequencing application program Sequence Navigator (see section 2.14).

2.13.2.b. Sequencing reactions using labelled primers on LI-COR system

After a LI-COR DNA Sequencer 4200 IR^2 (MWG-Biotech, Germany) became available in the Department, some RPE cDNA clones were sequenced using this automated DNA image analysis system based on two-dye near-IR (NIR) fluorescence technology.

The methodology specific to this system is based on the use of two differently tagged primers for the amplification of the sequencing template in chain termination reactions. The NIR fluorescence dyes used are IRD41 (with emission at 807nm) and IRD700 (with emission at 705nm). The amplification is accomplished by a thermostable DNA polymerase in the presence of an optimized mix of all dNTPs and one ddNTP (four reactions are performed for each template). The extension products are detected in four sequencing lanes (A, G, C and T), following excitation of fluorophores by a dual laser (700nm and 800nm), with a dual infrared detector. This allows simultaneous readings of the

two different NIR emissions corresponding to the fluorescent dyes used for the end-labelling of primers. Two separate gel images are generated, each corresponding to the data collected for a different dye. Thus, the sequence of both strands can be determined at the same time.

Typically, 100ng DNA per kb template and 2 pmole of each dye-labelled primer (Table 2.9) were used per 20µl PCR reaction. Reactions were performed using Thermo Sequenase™ fluorescent labelled primer cycle sequencing kit with 7-deaza-dGTP (Amersham).

Table 2.9. Primers used for sequencing on LI-COR system.

Primer/Sequence	**Annealing temperature**
M13 universal, 18-mer 5'TGTAAAACGACGGCCAGT3'	57°C
M13 reverse, 18-mer 5'CAGGAAACAGCTATGACC3'	57°C

The primers were provided by MWG Biotech labelled with IRD-41 and IRD-700 dyes.

The protocols for PCR amplification and preparation of samples recommended by MWG-Biotech were followed. The electrophoresis was performed on RapidGel-6%-acrylamide gel (Amersham). Data collection and image analysis in real time were supported by Base ImagIR™ software.

2.14. BIOINFORMATICS ANALYSIS OF cDNA CLONES

2.14.1. PROCESSING OF SEQUENCE FILES

The sequence data collected from the automated sequencers was processed with Sequence Navigator™, DNA and protein sequence comparison software (Perkin-Elmer Applied Biosystems). The processing involved feature identification, editing and export of files in a format suitable for database searches. Some alignments and comparisons between a sequence and a reference or another sequence were also performed with this program.

The initial "cleaning up" of the sequence data made use of Factura - a separate program in the Sequence Navigator™ software package. Typically, multiple sequences were processed in a batch mode to identify vector sequences and ambiguity ranges. In all cases the sequences and the corresponding electropherograms were also individually examined for marking the identified features and for calling, if possible, the bases identified by the program as ambiguous. The confidence range was identified for each edited sequence file.

The sequence files thus created were imported into layouts of Sequence Navigator program, automatically removing the vector and ambiguity features identified in Factura. The main task performed with this program was export of the files into GCG™ file format, required for the subsequent database searches and comparisons.

2.14.2. DATABASE SEARCHES AND PROGRAMS FOR OTHER SEQUENCE ANALYSIS APPLICATIONS

Sequence comparisons against GenBank and EMBL databases were performed using FASTA and BLAST programs, either as stand-alone versions or as part of the GCG package. The following computing facilities were used: UK Human Genome Mapping Project (UK HGMP) Resource Centre, Hinxton, UK (http://www.hgmp.mrc.ac.uk/); University of Liverpool Computing Services (http://www.liv.ac.uk/) and SEQNET Daresbury Laboratory, Warrington, UK (http://www.seqnet.dl.ac.uk/).

The algorithm for the Basic Local Alignment Search Tool (BLAST) directly approximates alignments that optimize the local similarity (Altschul *et al.*, 1990). It performes quick searches, but with less stringency than FASTA. In particular, it is not good at finding matches that have insertions/deletions in them. A BLAST search can give a quick ckeck of the presence of a query sequence in the databases, but for obtaining the degree of similarity of the query sequence with other related sequences over its entire length the more sensitive program FASTA can be used (Pearson & Lipman, 1988). Since the latter option can be of particular relevance for comparisons of sequences with various degrees of similarity, very often in this study FASTA searches were preferred.

Following comparison with database entries, the assignments of ESTs were based on database annotations. Other applications of DNA sequence analysis and the relevant programs most used in this study are listed in Table 2.10.

Table 2.10. Other programs used for DNA sequence analysis applications.

Application	Program	Function description
Sequence editing	setkeys	redefines the keyboard
	seqed	sequence editor
Mapping	map	displays both strands of DNA with restriction map and protein translations
	mapplot	shows graphically all possible restriction sites in a sequence
	mapsort	finds and sorts restriction enzyme sites
Sequence comparison	compare	compares two protein or nucleic acid sequences
	gap	makes optimal alignment of two closely related sequences
	overlap	compares two DNA sequences to each other in both orientations
Database searching	fetch	gets sequences from the databases
	stringsearch and lookup	find sequences by searching their comments/data files for words
Primer design	prime	select oligonucleotide primers for a template DNA sequence
	primer[1]	
Sequence exchange	toprimer[2]	formats a GCG sequence into a PRIMER compatible file
	reformat	makes sequence files readable by GCG programs

1-Whitehead Institute program; 2-Extended GCG programs; all the others are GCG programs.

CHAPTER 3

RESULTS

RPE cDNA LIBRARIES CONSTRUCTION AND CHARACTERIZATION. ANALYSIS OF ESTs

The overall objective of this project was to reveal details of the transcriptional activity of RPE cells through a molecular cloning study. The rationale behind such an approach is supported by the fact that the specialized activities of these cells are only partially characterized at molecular level. Also, the molecular basis of alteration of their functional and phenotypic characteristics in various pathological conditions and in culture conditions mostly are not known and have been studied far more at the cellular level. Therefore, identification of genes active in the RPE, *in vivo* and *in vitro*, would help clarify its unique functions and may further our understanding of its characteristics in health and disease.

The assessment of the transcriptional activity profile of RPE cells required the availability of cDNA libraries of these cells. Therefore, the first part of the project was dedicated to construction of representative cDNA libraries of bovine RPE cells. Taking the view that the most informative set of data would be provided by the profile of genes expressed *in vivo* by RPE cells, a special emphasis was put on construction of a cDNA library of freshly dissected RPE cells. This proved to be the most difficult aspect of the project and after a few months of work it became clear that it would be a challenge of initially unsuspected proportions. The first section of this chapter describes the strategies for cDNA synthesis and cloning that I have used, alongside the control experiments employed to monitor the techniques involved by these procedures.

Since the cDNA libraries obtained were not of the desired quality in terms of number of recombinants with long inserts, and on the background

of restrictions imposed due to the BSE epidemic about a year after the start of the project, the strategy was shifted towards studying a good quality RPE cDNA library. This decision was enabled by the offer made by Dr. Dean Bok (UCLA, USA) of an aliquot of a commercially constructed cDNA library of cultured human foetal RPE cells.

Clones from all available RPE cDNA libraries were partially sequenced during the second part of the project. Most of the ESTs analyzed originated from the human foetal RPE cDNA library. Results of the ESTs analysis are presented in the second section of this chapter.

Finally, a group of clones, which constituted part of an interesting finding, were further analyzed and the relevant data are presented in the closing section of this chapter.

3.1. cDNA LIBRARIES OF RPE CELLS USED IN THIS STUDY

3.1.1. CONSTRUCTION OF BOVINE RPE cDNA LIBRARIES

3.1.1.1. Isolation of RPE cells for mRNA preparations

As described in chapter 1, the RPE cells form a monolayer of delicate cells juxtaposed between the retina and choroid. While their apical side is not anatomically joined to the neurosensitive retina (the microvilli of RPE cells surround photoreceptor OSs), their basal membranes are tightly cemented onto Bruch's membrane. The isolation procedure requires separation from the neurosensitive retina and detachment of RPE cells from the basal membrane. Since most of the RPE cells isolated from freshly dissected bovine eyes were used for mRNA preparations, special attention was given to optimizing conditions for their isolation, both qualitatively, i.e. pure and intact cells, and quantitatively, i.e. complete detachment of RPE cells from as large a surface as practical.

I found that the usual procedure for isolating bovine RPE cells using trypsin-EDTAx4Na (see 2.5.1), with an incubation time of 45 minutes at 37°C and subsequent washing of eye cup walls by pipetting up and down

the incubation solution, yields RPE cells that consistently set up good primary cultures. However, the number of cells thus isolated is in the range of 5-7x10^5 cells/bovine eye, representing less than 10% of the total number of RPE cells/ bovine eye. The total population was estimated at about 8x10^6 cells, knowing that there are at least 4x10^6 RPE cells/human eye (Forrester *et al.*, 1996), the surface area occupied by RPE in bovine eyes is about twice that in human eyes, and the size of RPE cells in the two species is similar.

The question that arose was how many RPE cells could efficiently be isolated per bovine eye. The efficiency, in this case, defines the maximum number of cells that can be isolated without contamination from fibroblasts or optic nerve constituent cells and without degradation which could subsequently impede the yield of mRNA isolation.

This problem was addressed by a series of experiments which monitored the isolation of bovine RPE cells from freshly dissected eye cups by trypsinization at 37°C for incubation times ranging between 45 minutes and 70 minutes. The optic nerve heads were isolated as described in section 2.5.1 and Figure 2.1. The solutions used for all experiments originated from the same batch preparation. Following the respective incubation time, the eye cup walls were washed by pipetting up and down portions of the incubation solution to help detach RPE cells that were loosened by trypsinization. The detachment of RPE cells was examined with the aid of histological staining (as described in 2.6.2.) of eye cups' cross-sections (Figure 3.1). The observations were correlated with the yields of RPE cells isolated in each case (Table 3.1).

The following observations were made. After an incubation time of 45 minutes the RPE cells had loose cell-cell connections but most of them appear still firmly attached on Bruch's membrane; this is why, under these conditions, only a small proportion of cells came off the monolayer. Following incubation for 50 minutes, most RPE cells were rounded up indicating isolation of each other but the attachment of the basement membrane still had not been sufficiently disrupted to allow majority of them to be washed away. Incubation times between 54 and 56 minutes resulted in appearance of gaps in the monolayer and in increase of the number of

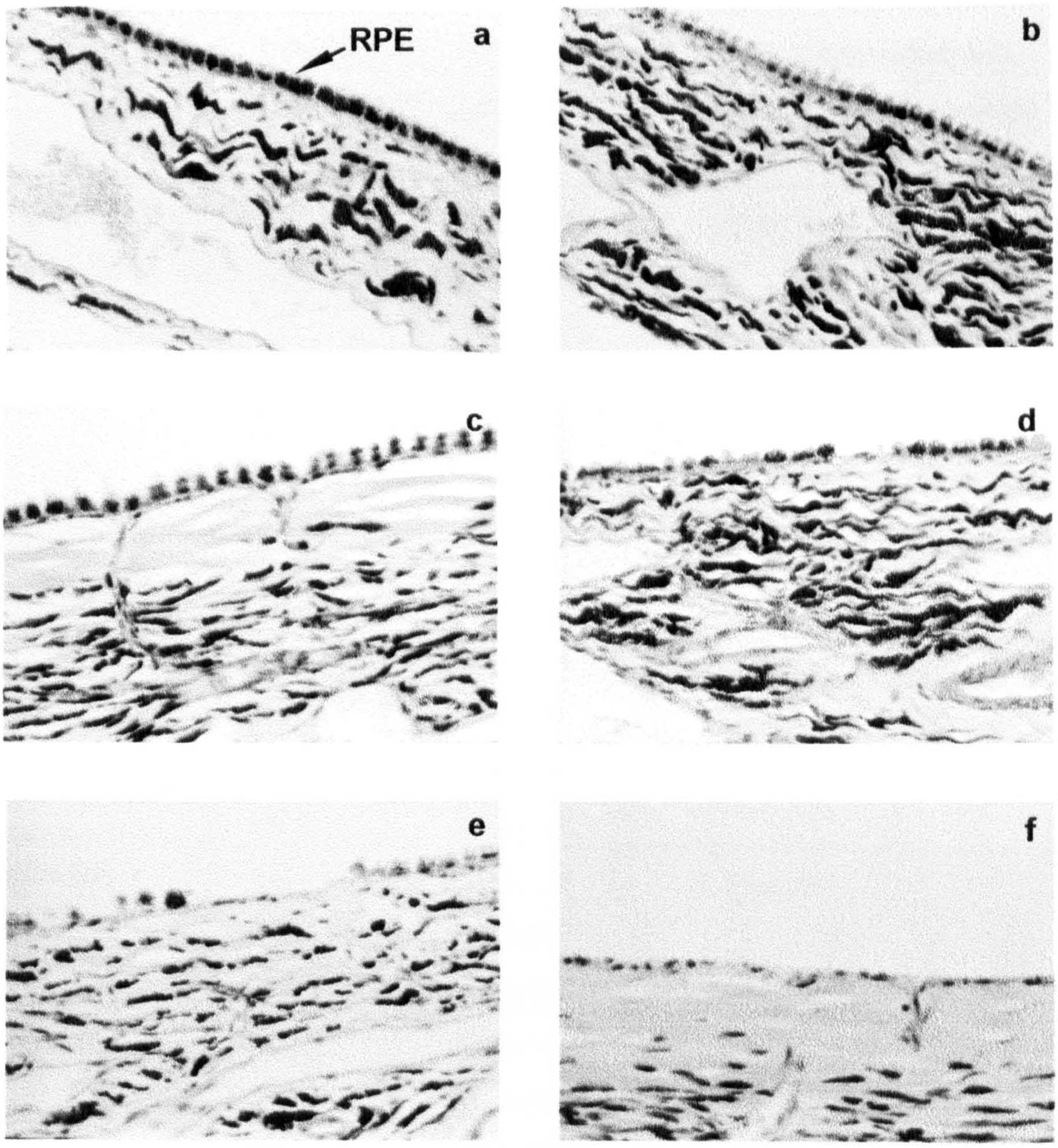


Figure 3.1. Histological staining of RPE/Bruch's membrane/choroid cross-sections examined by differential interference contrast microscopy illustrating stages of detachment of bovine RPE cells following trypsinization at 37°C (all x40).
(a) 45 minutes; (b) 50 minutes; (c) 54 minutes; (d) 56 minutes; (e) 60 minutes; (f) 70 minutes. See text for analysis of results.

Table 3.1. Yields of RPE cells obtained through trypsinization of freshly dissected bovine eye cups.

Incubation time	haemo-cytometer count (N)	RPE cells isolated/eye	Average number of RPE cells isolated/eye
45 minutes	26	$0.52x10^6$	**$0.54x10^6$**
	29	$0.58x10^6$	
	27	$0.54x10^6$	
50 minutes	35	$0.70x10^6$	**$0.69x10^6$**
	33	$0.66x10^6$	
	36	$0.72x10^6$	
54 minutes	42	$0.84x10^6$	**$0.87x10^6$**
	45	$0.90x10^6$	
	43	$0.86x10^6$	
56 minutes	55	$1.10x10^6$	**$1.08x10^6$**
	53	$1.06x10^6$	
	54	$1.08x10^6$	
60 minutes	50	$1.50x10^6$	**$1.56x10^{6*}$**
	52	$1.56x10^6$	
	54	$1.62x10^6$	
70 minutes	90	$2.70x10^6$	**$2.69x10^{6*}$**
	87	$2.61x10^6$	
	92	$2.76x10^6$	

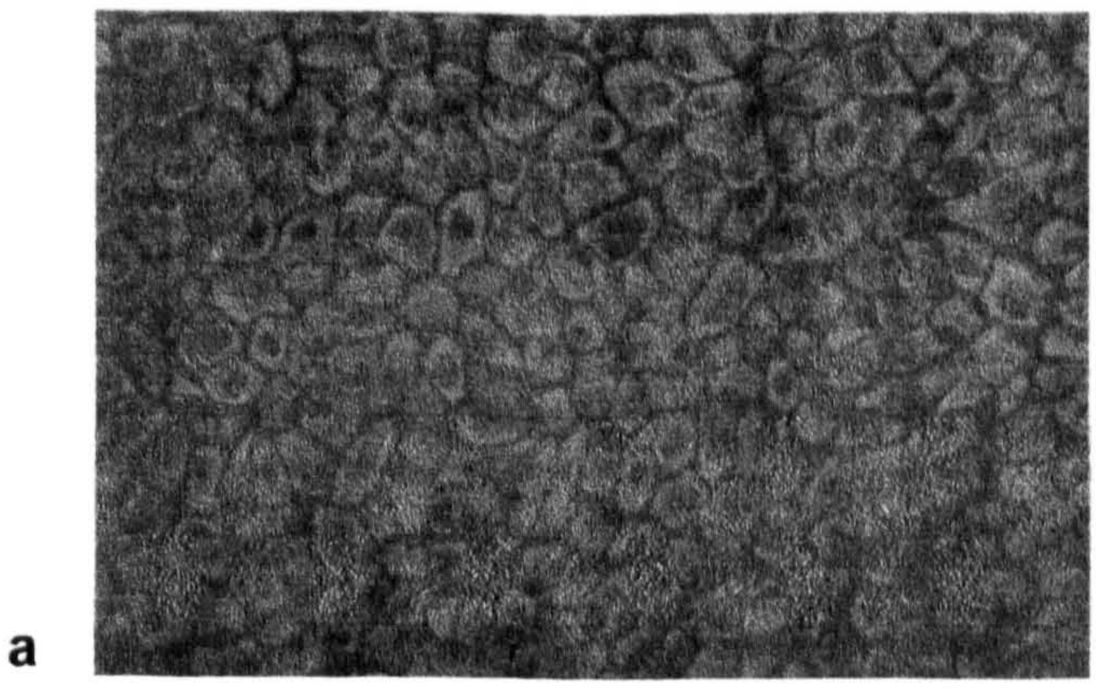
*The cell suspension was centrifuged, cells washed with PBS and resuspended in 2ml (3ml for 60 and 70 minutes time points) of culture media. Cells were then counted using a haemocytometer (N). Three determinations were made for each time point. RPE cells isolated/eye (found in 2ml, respectively 3ml suspension)=N x 10^4 x 2 (respectively, 3). The average presented is the arithmetic mean of the three determinations for each time point. * - time points when lysis of cells was evident.*

total cells isolated per eye. Although a further incubation of 4 minutes (60 minutes total) led to much wider gaps in the monolayer and an increased number of total RPE cells detached, some cells appeared to have undergone lysis (burst cells, of irregular shape were observed when counting the cells). Following trypsinization for 70 minutes, virtually all RPE cells could be washed away but an extensive lysis of cells also occured.

In the light of these results it appeared that the critical point for harvesting the RPE cells from the eye cups is at approximately 55 minutes, when the maximum possible number of cells could be isolated while avoiding lysis due to prolonged trypsinization. Usually, batches of 6-9 bovine eyes obtained from local abbatoirs (2-3 hours post-mortem) were handled at one time. The cells obtained were either used immediately for mRNA preparation or rapidly frozen in liquid nitrogen and stored at -70°C for later use.

Cultured bovine and human RPE cells were also used for mRNA preparations (used mostly for some of the Northern blotting experiments). The main characteristics of the bovine RPE cell cultures were as follows: after 7-8 days in primary culture colonies of pigmented, mostly polygonal cells were formed; within 2-3 weeks a monolayer of tightly packed cells was established (Figure 3.2.a); the mosaic pattern was subsequently lost in sub-culture, when cells appeared mostly flattened and unpigmented; the confluence was reached after 5-7 days post-passage. Sub-cultures of human RPE cells had similar morphological characteristics (see Methods' section 2.5.4). Cultured RPE cells up to fourth passage (bovine) and sixth passage (human) were used in this study.

The purity of RPE cells was assessed immunocytochemically (as described in 2.6.4), following culture of cells in 8-chambered cell culture slides (see 2.5.3). The presence of cytokeratins, intermediate filament components characteristic for RPE cells, was verified. Moderate to strong immunoreactivity confirmed that the cells were of RPE origin (Figure 3.2.b).



a

b

Figure 3.2. Cultured RPE cells.
(a) immunofluorescence of a primary culture of bovine RPE cells showing the mosaic-like pattern of cells after 16 days *in vitro* (vimentin-FITC monoclonal antibody; x100).
(b) light micrograph of preconfluent, flattened cells stained by immunoperoxidase technique (DAB staining) for cytokeratins illustrating the retinal pigment epithelial origin of cells (x600). Anti-vimentin and anti-GFAP stainings were negative (data not shown).

3.1.1.2. Poly A+ mRNA preparations from RPE cells and retina. Quality control of mRNA

Poly A+ mRNA was isolated directly from cells, without intermediate preparation of total RNA, using either Oligotex™ (Qiagen) or FastTrack® (Invitrogen) procedure. The RNase-free conditions were achieved by using a denaturing lysis buffer containing guanidinium isothiocyanate and a SDS-based buffer containing RNase/Protein Degrader (Invitrogen). mRNA was recovered from Oligotex particles and, respectively, from oligo(dT) cellulose in 5mM Tris-HCl pH7.5. The preparations using oligo(dT) cellulose required ethanol precipitation and concentration.

A summary of mRNA preparations from bovine RPE cells (freshly dissected and cultured) and bovine retina is presented in Table 3.2. The mRNA samples having $OD^{260}/OD^{280} \geq 1.8$ were regarded as pure and were used in downstream applications. The yield of mRNA obtained from both freshly dissected and cultured RPE cells indicated an average of 0.15 pg poly A+ mRNA/RPE cell.

The integrity of mRNA was assessed by Northern blotting. Figure 3.3 shows a Northern blot of mRNA directly isolated from freshly dissected bovine RPE cells hybridized with a 32P-labelled β-actin probe. The retina mRNA (available in larger quantities) was introduced as a positive control for the conditions of hybridization. A distinct band of hybridization of about 2kb is observed indicating an intact mRNA. This demonstrated the quality of mRNA to be good and suitable for cDNA synthesis.

Table 3.2. Summary of mRNA preparations from bovine RPE cells and retina.

No. and type of cells	P	Volume of prep (µl)	Spectrophotometrical measurements D	OD_{260}	OD_{280}	OD_{260}/OD_{280}	mRNA yield (µg)	mRNA/ cell (pg)
$2x10^6$ fresh RPE from 4 eyes	Q	40	nd				nd	nd
$3.5x10^6$ fresh RPE from 6 eyes	Q	40	5	0.045	0.025	1.80	0.36	0.10
$1.7X10^7$ fresh RPE from 19 eyes (2 preps)	F	20	50	0.070	0.038	1.84	2.80	0.16
$2x10^7$ fresh RPE from 21 eyes (2 preps)	F	25	50	0.050	0.027	1.85	2.50	0.12
$5.5x10^6$ cultured RPE cells	Q	40	20	0.032	0.018	1.78*	1.02	0.18
$3x10^7$ cultured RPE cells (3 preps)	F	20	100	0.055	0.030	1.83	4.4	0.14
110mg neuro-retina	F	20	100	0.080	0.043	1.86	6.4	nd

Procedures (P) used were: Oligotex™ Direct, Qiagen (Q); FastTrack®, Invitrogen (F). Where indicated, the cells available were divided in aliquots of up to $1x10^7$ cells/prep in order to avoid formation of a highly viscous lysate; the respective eluates obtained were then pooled and concentrated together.

nd - not determined.

** - the preparation with $OD^{260}/OD^{280} < 1.8$ was not used for cDNA synthesis.*

The cultured cells (up to 4th passage) were harvested when pre-confluent and confluent.

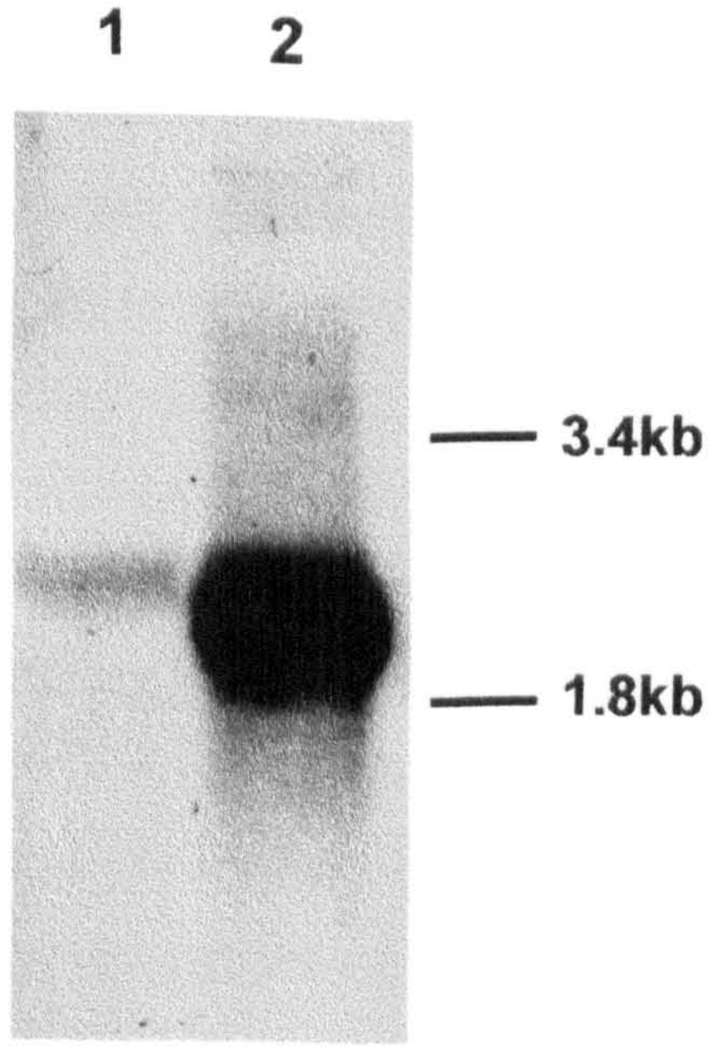


Figure 3.3. Northern blot of bovine RPE and retina poly A+ mRNA probed with a ^{32}P-labelled β-actin probe.
Lane 1 - 0.4μg of poly A+ mRNA from freshly dissected RPE cells
Lane 2 - 2μg of bovine retina poly A+ mRNA
The RNA samples were separated on a 1.3% formaldehyde/MOPS denaturing gel and then transferred into nylon Hybond-N membrane. The blot was probed with 40ng of a human β-actin probe labelled by random priming in the presence of (α^{32}P)dCTP to a specific activity of 1.83x10^9 dpm/μg. Hybridization and washes were carried out at 50°C. The signal was detected after 3 days exposure to X-ray film, with intensifying screen at -70°C.
Migration of yeast ribosomal RNA is indicated at the right.

3.1.1.3. Synthesis of bovine RPE cDNA

It was at this stage of the project when it became clear for the first time that the quantity of mRNA available for cDNA synthesis was limited. Also, a significant proportion of it had been used for quantitation and quality control of mRNA. I had two alternatives: either to keep accumulating RPE cells for a larger preparation of mRNA or to start synthesize cDNA from the available mRNA. Since already precious time had been lost, I chose the latter. I continued, though, to isolate and culture bovine RPE cells whenever the supply of bovine eyes was available. A few cDNA syntheses were thus made until the supply of material for dissection stopped due to the BSE epidemic.

The main characteristics of bovine RPE cDNA syntheses performed are summarized in this section while the cloning experiments of the resulting cDNA into λ-based vectors are described in the following section.

The method employed for cDNA synthesis involved three steps: (i) first strand cDNA synthesis primed with either random hexanucleotide primers $(dN)_6$ or oligo $(dT)_{25}$ primers; (ii) introduction of nicks in the mRNA strand with RNase H; (iii) synthesis of second strand cDNA in a nick translation reaction, followed by blunt-ending of double stranded cDNA. The reactions were carried out according to the instructions accompanying the cDNA Synthesis Module (Amersham) and TimeSaver® cDNA Synthesis Kit (Pharmacia). A standard cDNA synthesis (i.e., starting from 1μg mRNA) was performed in 20μl first strand and 100μl second strand synthesis reactions. Amounts between 0.5μg and 1.2μg mRNA were used per reaction. When less than 1μg mRNA was used as a template, the volumes of reaction components specified in the protocol accompanying the cDNA Synthesis Module (Amersham) were decreased proportionally; as recommended, the reverse transcriptase:mRNA and enzyme volume: reaction volume ratios were maintained the same.

The majority of first strand cDNA synthesis reactions were randomly primed. The assumption when making this choice was that cDNA clones thus generated would be more informative in obtaining coding sequences and in identifying genes. Also, as mentioned in the relevant section in

chapter 2, this strategy could overcome some of the problems associated with secondary structure by providing a series of alternative priming sites. Noteworthy is the fact that the amount of $(dN)_6$ primers recommended as optimum in the Amersham protocol (i.e., 7µM final concentration in the reaction mix) corresponds to conditions for medium-length cDNA synthesis according to Pharmacia's guidelines.

The cDNA synthesis was in all cases monitored by incorporating radioactive nucleotides during synthesis. Samples from first and second strand cDNA synthesis reactions were removed and analyzed for yield and product size distribution. The percentage of incorporation of radioactivity (%inc), needed for estimating the amount of cDNA synthesized, was calculated as the ratio incorporated counts (c_{inc})/total counts ($c_{inc}+c_{uninc}$) for first and second strand in separate labelling reactions and for first strand in sequencial labelling reaction; before calculating %inc for second strand in sequencial labelling, the amount of radioactivity incorporated into the first strand was subtracted, compensating also for the volume variation.

An RNA ladder provided by Amersham was used both as a test of the system components and as a size marker for gel analysis. The results of two cDNA syntheses, each starting from 0.5µg RNA ladder are given in Table 3.3. The percentage incorporation of radioactive label was 1.12%-1.20% in the first strand and 0.99%-1.08% in the second strand. From these figures, the amounts of each strand of cDNA synthesized were calculated as indicated in section 2.8.3.a. The calculations took into account the fact that there are 5 nmols of dCTP in a 10µl first strand reaction (the amount of labelled dCTP is negligible, i.e.- 1.65 pmols in 5µCi of 3000Ci/mmol specific activity) and that the amount of unlabelled dCTP does not change significantly in the second reaction mix. Between 78-84 ng and 69-77 ng of cDNA were synthesized in the first and, respectively, in the second strand reactions. Therefore, the total double stranded cDNA synthesized in the two control reactions described was about 160 ng and 148 ng. In the first strand reaction, in both cases, about 16% of the mRNA template was transcribed, and 88%-90% of the resulted first strand cDNA was subsequently transcribed into second strand. Since these yields were as expected according to Amersham's guidelines (around 10% RNA

transcribed into first strand cDNA and 90% first strand transcribed into second strand cDNA), I proceeded to cDNA synthesis from RPE mRNA. The products of cDNA synthesis from the RNA ladder were used as size markers in the analysis by gel electrophoresis of RPE cDNA (Figure 3.4):

The characteristics of bovine RPE cDNA syntheses are summarized in Table 3.4. Although the efficiency of RPE cDNA synthesis (described by the percentage of template transcribed in both reactions) was roughly similar to that achieved when using the RNA ladder, a few details are worthy of further note.

In all syntheses, the percentage of mRNA transcribed into first strand cDNA was lower than expected (as indicated in Amersham's literature, at least 20% of the mRNA population should be transcribed). The lowest yield of transcription was observed when the first strand cDNA synthesis was oligo(dT) primed (9.8%). The secondary structure of mRNA or a slight over-estimation of the mRNA amount could have contributed to this lower efficiency, more likely so since the copying into second strand was very efficient (96% template transcription).

When first strand cDNA synthesis was randomly primed, about 17% of mRNA was transcribed when using Amersham module , while with Pharmacia's kit the respective values were 13.5% and 10.3%. Notably, the primer concentration in the latter syntheses was 1μM which creates conditions for longer cDNA synthesis (compared with 7μM used in first syntheses). Generally, the yields of transcription into second strand cDNA were higher than 94% (the exception being the first synthesis, when the respective value was 90%).

The resulting double stranded cDNA was blunt-ended as described in Methods' section 2.8.1. The total amount of blunt-ended double stranded cDNA synthesized from mRNA of freshly isolated RPE cells was approximately 250 ng per synthesis when random hexanucleotide primers were used and 150 ng when oligo(dT) primers were used. Approximately 330 ng of cDNA of cultured RPE cells was obtained.

Table 3.3. Control cDNA syntheses using a RNA ladder monitored by incorporation of (α-^{32}P)dCTP into first and second strand synthesis reactions.

Amount of mRNA	**Primer**	**Conc. of primer**	**First strand cDNA synthesis**				**Second strand cDNA synthesis**			
			c_{inc}/ c_{uninc}	**% inc**	**cDNA synthesized**	**% template* transcribed**	**c_{inc}/ c_{uninc}**	**% inc**	**cDNA synthesized**	**% template* transcribed**
0.5 µg[1]	$(dT)_{25}$	0.4 µM	2143/ 175452	1.20	**84.0 ng**	**16.8**	10259/ 899070	1.08	**75.6 ng**	**90.0**
0.5 µg[2]	$(dT)_{25}$	0.4 µM	4858/ 428944	1.12	**78.4 ng**	**15.6**	3978/ 397878	0.99	**69.3 ng**	**88.4**

Reactions were performed according to the protocol of cDNA Synthesis Module (Amersham).

1 - sequential reaction labelling of cDNA synthesis with 5µCi of (α-^{32}P)dCTP in10µl first strand cDNA reaction and 25 µCi of the same label in 50µl second strand reaction;

2 - synthesis monitored by labelling the first and second strand reactions separately, with 10µCi of (α-^{32}P)dCTP per 10µl first strand reaction and, respectively, per 50µl second strand reaction;

The incorporated counts (c_{inc}), unincorporated counts (c_{uninc}) and percentage incorporation (%inc) were estimated in samples of 0.5µl and 2.5µl from first strand and, respectively, second strand cDNA synthesis reactions.

* - template referres to mRNA for first strand cDNA synthesis and, respectively, to first strand for second strand cDNA synthesis.

Table 3.4. Summary and analysis of cDNA syntheses of bovine RPE cells.

Amount of mRNA / Type of cells	Primer / Conc.	First strand cDNA synthesis				Second strand cDNA synthesis				ds cDNA
		c_{inc}/ c_{uninc}	% inc	cDNA synthesized	% template transcribed	c_{inc}/ c_{uninc}	% inc	cDNA synthesized	% template transcribed	
0.8 µg (freshly isolated RPE cells)[A, q]	$(dN)_6$ / 7µM	3394/ 338080	1.00	**140.0 ng**	**17.5**	13455/ 1396186	0.90	**126.0 ng**	**90.0**	**266 ng**
0.8 µg (freshly isolated RPE cells)[A, q]	$(dT)_{25}$ / 0.4µM	1963/ 346240	0.56	**78.4 ng**	**9.8**	7673/ 1321939	0.54	**75.6 ng**	**96.4**	**154 ng**
1.0 µg (cultured RPE cells) [A, q]	$(dN)_6$ / 7µM	3826/ 321962	1.17	**163.8 ng**	**16.3**	15032/ 1280671	1.10	**154.0 ng**	**94.0**	**338 ng**
1.0 µg (freshly isolated RPE cells) [P,S]	$(dN)_6$ / 1µM	5837/ 591429	0.97	**135.8 ng**	**13.5**	5803/ 619241	0.92	**128.8 ng**	**94.8**	**264 ng**
1.2 µg (cultured RPE cells) [P,S]	$(dN)_6$ / 1µM	5536/ 616230	0.89	**124.6 ng**	**10.3**	5212/ 614374	0.84	**117.6 ng**	**94.3**	**242 ng**

Syntheses were carried out with components from: **A** - cDNA Synthesis Module (Amersham); **P** - TimeSaver® cDNA Synthesis Kit (Pharmacia). Labelling was achieved in : **q** - sequential reactions (5µCi of (α-^{32}P)dCTP in first strand reaction and 25 µCi (α-^{32}P)dCTP in second strand reaction); **S** - separate reactions (10µCi (α-^{32}P)dCTP in each reaction). Samples of 0.5µl and 2.5µl from first strand and, respectively, second strand cDNA synthesis reactions were used for estimating incorporation of radioactivity.

The products of first and second strand labelled cDNA synthesis reactions were analysed by electrophoresis on a 1% alkaline agarose gel. The RPE cDNA appeared as a smear of heterogeneous molecular weight products from below 0.5 kb to approximately 4 kb, with the majority of synthesis products up to about 1.4 kb (Figure 3.4). No size fractionation was done at this stage in order to avoid loss of cDNA. Following phenol/ chloroform purification and ethanol precipitation, the blunt-ended double stranded cDNA was resuspended in 5-10µl TE buffer (pH 8.0) and subjected to adaptor ligation.

3.1.1.4. Cloning of bovine RPE cDNA into λ-based vectors

In order to be cloned into *Eco*RI-digested vector arms, the double stranded cDNA was *Eco*RI-"adapted", i.e. the blunt-ends of the cDNA were converted to *Eco*RI cohesive termini by ligation with adaptors. Each adaptor carries a blunt-end for ligation with cDNA and an *Eco*RI cohesive end. The *Eco*RI-adapted cDNA was next subjected to column purification (Pharmacia Sepharose® CL-4B spun columns) intended to size fractionate the cDNA and to eliminate the unreacted adaptors.

Ligation of adaptors to cDNA and, in particular, the subsequent purification/size fractionation of "adapted" cDNA represent steps of the cloning procedure that are diffficult to assay. Possible gel analysis and self-ligation test involve use of material that can not be recovered. However, the ligation of adaptors and size fractionation were monitored using a blunt-end control DNA supplied by Amersham (M13mp8 digested with *Hae*III). Adaptors were ligated to 1µg of blunt-end control DNA in a ligation reaction that made use of 250 pmoles of adaptor. As shown in Figure 3.5, the unreacted adaptors and the fragments smaller than 0.5kb were quantitatively removed by column purification. Also, the attachment of adaptors to the blunt-ends was confirmed by the increased size of fragments in lane 3 of the same figure compared with the corresponding fragments in lane 2.

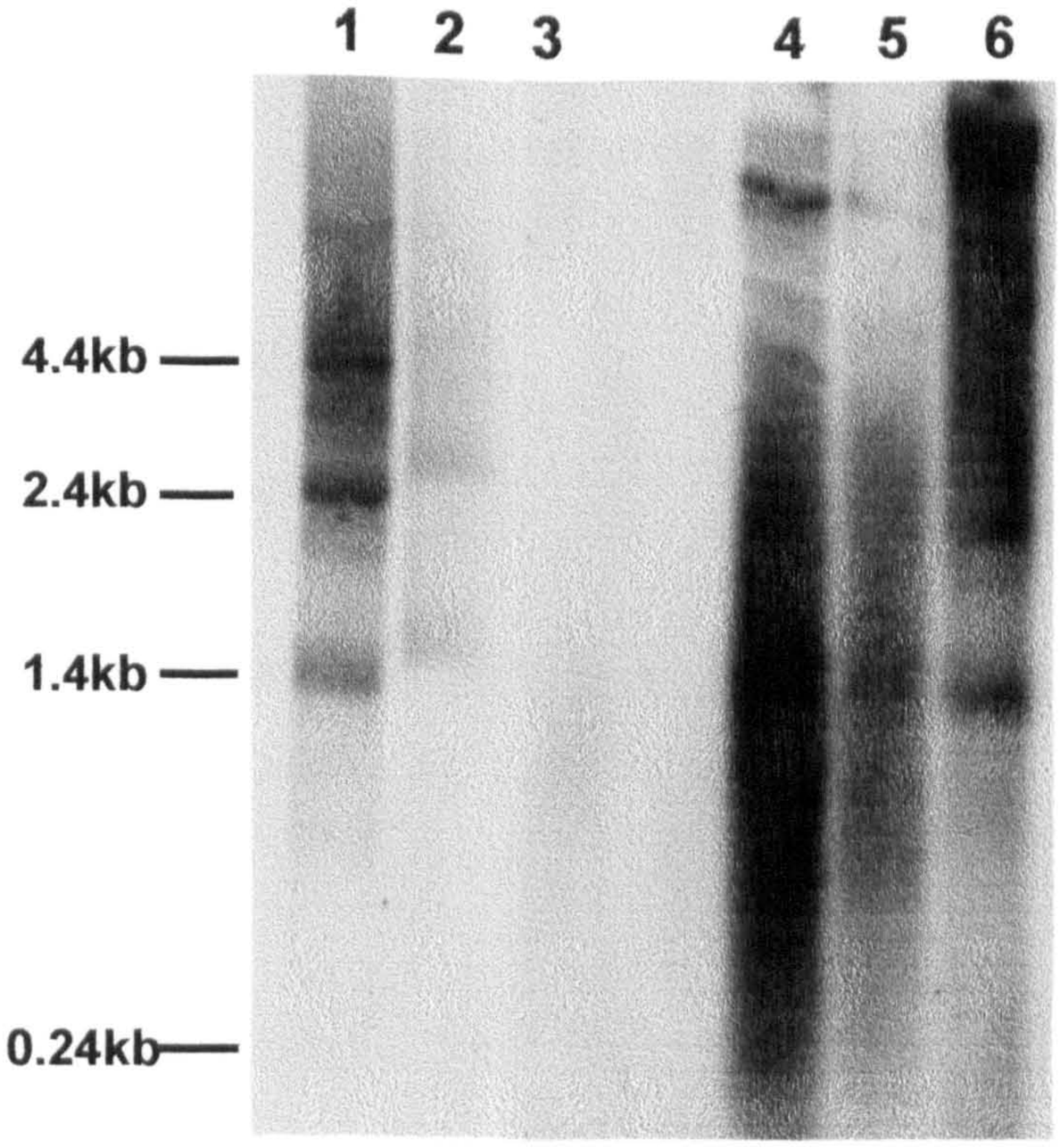


Figure 3.4. Alkaline agarose gel electrophoresis of cDNA synthesis from mRNA isolated from freshly dissected RPE cells.
Lanes 1, 2, 3 - first strand cDNA synthesis
Lanes 4, 5, 6, - second strand cDNA synthesis
Lanes 1 and 6 - control RNA ladder oligo$(dT)_{25}$ primed
Lanes 2 and 4 - RPE cDNA primed with random hexamers
Lanes 3 and 5 - RPE cDNA oligo$(dT)_{25}$ primed
The labelling was performed in sequencial reactions according to Amersham's protocol.
Overnight exposure at -70°C with intensifying screen.

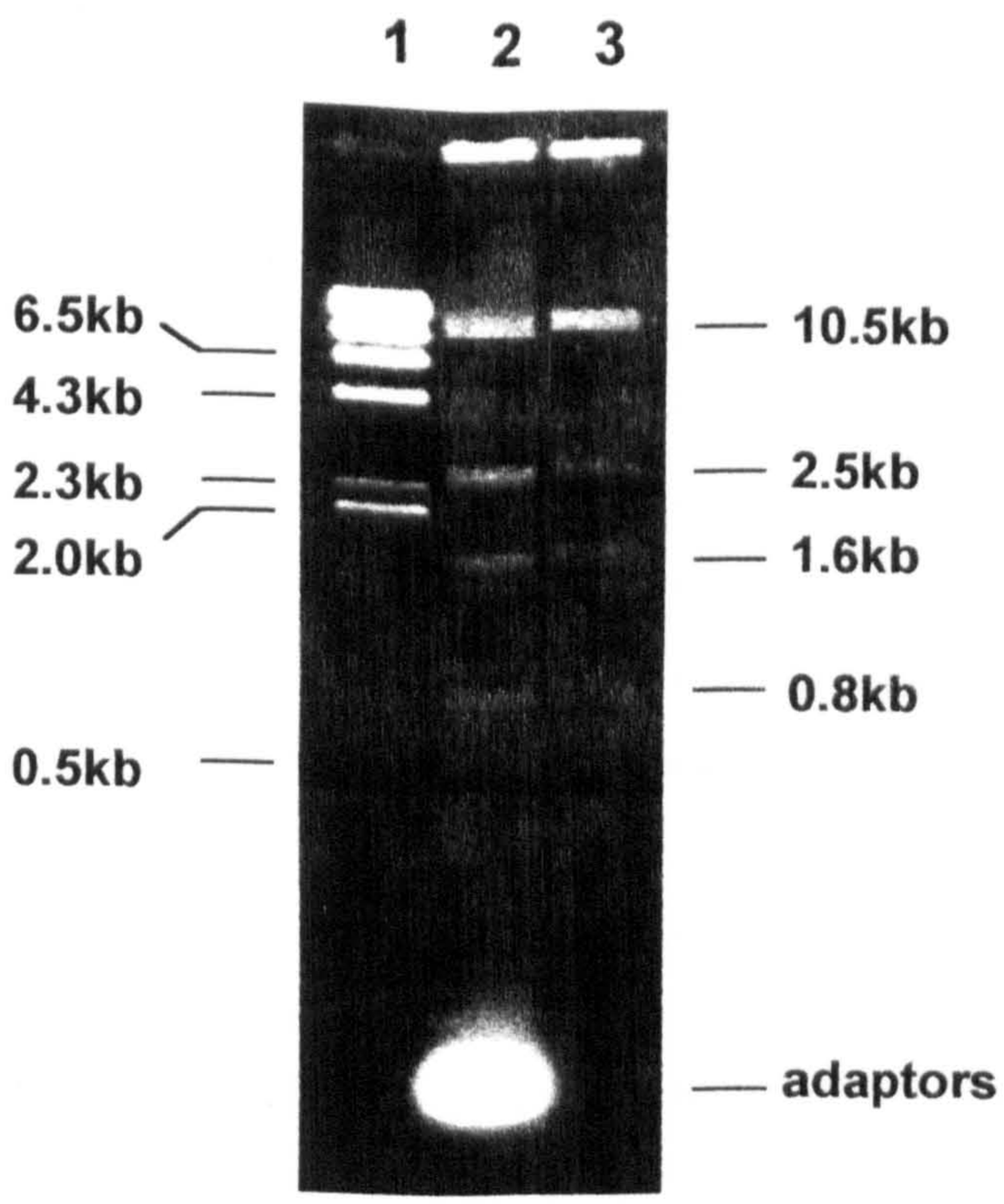


Figure 3.5. Gel analysis of adaptor ligation/spun column purification of blunt-end control DNA.
Lane 1 - λ /Hind III size marker (the corresponding fragment sizes are indicated on the left)
Lane 2 - sample removed from the adaptor ligation reaction of blunt-end control DNA, before incubation (approximately 0.4µg DNA)
Lane 3 - sample removed after column purification of "adapted" control DNA
The electrophoresis was performed in 1% agarose/TBE gel, in the presence of ethidium bromide.

In the case of bovine RPE cDNA, the adaptor ligation reaction was carried out with around 250-300 ng of cDNA (maximum amount available for cloning at any given time). The molar ratio adaptor:cDNA was maintained to about 80:1, as in the case of blunt-end control DNA described above. Following column fractionation, the 5'-hydroxyl groups on the *Eco*RI-adapted cDNA were converted to 5'-phosphate groups in a kinasing reaction in the presence of T4 polynucleotide kinase. After a further purification, by phenol extraction, the cDNA was finally ethanol precipitated.

The ligation of 5'-phosphorylated *Eco*RI-adapted cDNA to *Eco*RI-digested, dephosphorylated λMOS*Elox* (Amersham) or λTriplEx (Clontech) vector was carried out at 16°C overnight in the presence of T4 DNA ligase. Since the ratio of cDNA to vector in the ligation reaction is a critical factor in determining recombination efficiency in a cDNA library, two separate cDNA-containing ligation reactions were set up in parallel for each cDNA preparation: one with a quarter of the cDNA preparation and the other with the other three quarters of the same preparation. Between 0.5μg and 1μg of vector arms were used per ligation. Positive control ligations were set up with approximately 100 ng of *Eco*RI-adapted blunt-end control DNA and 100 ng of an *Eco*RI-ended control (3.7kb plasmid liniarized with *Eco*RI). A negative control ligation reaction contained vector arms alone (no insert). A separate λ phage *in vitro* packaging reaction was performed for each of the ligations using λ-DNA *in vitro* Packaging Module (Amersham) and obtaining 500μl recombinant phage stock in SM buffer (unamplified library). The titre of each library was determined by plating appropriate dilutions of the phage stock on *E. coli* strain ER1647 for λMOS*Elox*-based phages and on *E. coli* strain XL1-Blue for λTriplEx-based phages. In all cases, the highest titre was obtained for preparations which used the larger aliquot of cDNA in the ligation with vector arms. Table 3.5 summarizes the characteristics and titration results of these unamplified libraries.

The whole λ DNA gave a titre of $0.7\text{-}1\times10^{8}$ pfu/μg indicating that the λ DNA *in vitro* packaging reactions worked well. The titre produced by the ligation of *Eco*RI-digested, dephosphorylated λ arms was low (1×10^{5} pfu/μg), demonstrating a low background of uncut or non-dephosphorylated vector arms. The test ligations of λMOS*Elox* with *Eco*RI-adapted blunt-end

control DNA and with *Eco*RI-ended control DNA resulted in titres of $2.5x10^6$ pfu/µg vector arms and $5x10^6$ pfu/µg vector arms respectively, indicating that the ligation reactions were efficient. All ligations of λ vectors with adapted cDNA, with one exception, resulted in cloning efficiencies of $1x10^6$ - $2x10^6$ pfu/µg vector arms. Although lower than those given by the test ligations, these efficiencies indicated a high percent of recombinants in the cDNA libraries (over 90% in all cases).The exception was represented by the ligation of oligo(dT) primed cDNA which gave a titre of only $0.4x10^6$ pfu/µg vector arms.

For the cDNA libraries in λTriplEx vector (A2, B, and C1), the recombination efficiency was confirmed by blue/white (*lac Z* α-complementation) screening. IPTG/X-gal plates with about 500 plaques were examined after being incubated at 37°C for 18 hours. The recombination efficiency, given by the percentage of white (recombinant) plaques out of total number of plaques, was 97% for library A2, 89% for library B and 90% for library C1, thus indicating the presence of an insert in a high proportion of clones.

Table 3.5. Characteristics of bovine RPE unamplified cDNA libraries.

Library	cDNA origin	Primer	Adaptor	λDNA	Titre pfu/ml	total pfu*	pfu/µg λDNA	%rec
F-N	fresh RPE	N	A	λMOS*Elox*	$4x10^6$	$2x10^6$	$2x10^6$	95
F-T	fresh RPE	T	A	λMOS*Elox*	$0.8x10^6$	$0.4x10^6$	$0.4x10^6$	75
B	fresh RPE	N	P	λTriplEx	$1x10^6$	$0.5x10^6$	$1x10^6$	90
A2	cultured RPE	N	P	λTriplEx	$2x10^6$	$1x10^6$	$2x10^6$	95
C1	cultured RPE	N	A	λTriplEx	$1.5x10^6$	$0.7x10^6$	$1.5x10^6$	93
Test1	*Eco*RI-adapted blunt-end control	_	A	λMOS*Elox*	$5x10^6$	$2.5x10^6$	$2.5x10^6$	96
Test2	*Eco*RI-ended control	_	_	λMOS*Elox*	$1x10^7$	$5x10^6$	$5x10^6$	98
_	_	_	_	λMOS*Elox* arms	$2x10^5$	$1x10^5$	$1x10^5$	_
_	_	_	_	λMOS*Elox*	$1.5x10^8$	$0.7x10^8$	$0.7x10^8$	_
_	_	_	_	λTriplEx arms	$1x10^5$	$0.5x10^5$	$1x10^5$	_
_	_	_	_	λTriplEx	$1x10^8$	$0.5x10^8$	$1x10^8$	_

Primers used in the first strand cDNA synthesis: N-random hexanucleotide; T-anchored oligo(dT)$_{25}$. Adaptors: A-Amersham; P-Pharmacia (see Figure 2.3).
Pfu-plaque forming units.

Percentage recombinants (% rec) $= \frac{\text{Titre library-Background titre(λ arms only)}}{\text{Titre library}} X100$

3.1.2. INSERT-SIZE DISTRIBUTION IN THE BOVINE RPE cDNA LIBRARIES AND THE HUMAN FOETAL RPE cDNA LIBRARY

The next step in assessing the quality of cDNA libraries constructed was to carry out size analysis of the inserts. Firstly, the insert sizes were estimated by PCR using vector-specific primers flanking the cloning site (primers are described in Table 2.6). Between 20 and 30 randomly chosen clones were analyzed from each library. The results were rather dissapointing. In the majority of cases, fragments of up to only 300 bp were amplified indicating the presence of inserts of around 100 bp (the size of the vector fragment amplified by the corresponding pair of primers is 171 bp for λMOS*Elox* and 160 bp for λTriplEx).

Secondly, various experiments were designed to characterize by restriction the plasmids resulting from subcloning of λ clones (see Methods' section 2.8.9). The most informative digestions were the following: for λMOS*Elox* the combination of separate restrictions with *Kpn*I (two restriction sites in the vector and one site in the Amersham adaptors) and *Bam*HI (single cutting site in vector and site in the Amersham adaptors); for λTriplEx single restrictions with *Hind*III (single restriction site in vector) and with *Not*I (single site in Pharmacia adaptors). The restriction experiments showed that more than 70% of clones have small adapted inserts (i.e., shorter than 300 bp). Manual sequencing of some of these clones further revealed the presence of small cDNA inserts, adaptor-dimers and multimers in most of the cases.

On average, out of 10 randomly chosen plaques from each library, 2-3 yielded clones with inserts of over 350 bp. The PCR amplification products of two such clones with inserts of approximately 500 bp and 700 bp are included in the gel presented in the Figure 3.6. Clones identified as having informative-sized inserts (i.e., of at least 300 bp) were partially sequenced and the results are presented in section 3.2.

Although possible reasons for the dissapointing outcome of the cDNA cloning will be further disscussed in the final chapter of this thesis, it is worth mentioning here that two processing steps of cDNA cloning

procedure are thought to be possible stages at which problems arose: column purification/size fractionation and subsequent ethanol precipitation of cDNA. Losses of cDNA at these stages were most likely to occur, especially as less than 0.5µg cDNA was used, which is the amount recognized to be problematic for such processings. At the same time, due to the large excess of adaptors used in the previous ligation, a less efficient purification could have lead to a large proportion of the recovered material from the column being represented by adaptor multimers. Overall, the suspected main reason for the dissapointing result in construction of RPE cDNA libraries was the small amount of starting mRNA.

Given the small proportion of informative-sized inserts in the bovine RPE cDNA libraries constructed in-house, it became obvious that these are not suitable for the proposed study. A good quality cDNA library was preferable. The human foetal RPE cDNA library (gift from Dr.Dean Bok, UCLA, USA) appeared to be such a library, following assessment of percentage recombinants and average size of inserts.

The human foetal RPE cDNA had been commercially constructed from mRNA isolated from cultured RPE cells. The cDNA was directionally cloned into *Eco*RI/*Xho*I sites of Uni-ZAP™ XR vector (Stratagene) from which it can be subcloned into a pBluescript SK(-)-based phagemid. The working aliquot (designated library Z) for this study originated from one-round amplification of the original cDNA library. Upon receipt, cDNA library Z was subjected to physical characterization and its main features are presented below.

The library had a titre of 1.7×10^9 pfu/ml and 98% recombinants, as indicated by blue/white screening. The PCR amplification of inserts resulted in fragment sizes over 600 bp in 8 out of 10 plaques isolated at random. The size distribution of inserts in the whole library was evaluated by PCR amplification and compared with the insert-size distribution in two of the bovine RPE cDNA libraries constructed in-house. The results of this comparative analysis are presented in Figure 3.6.

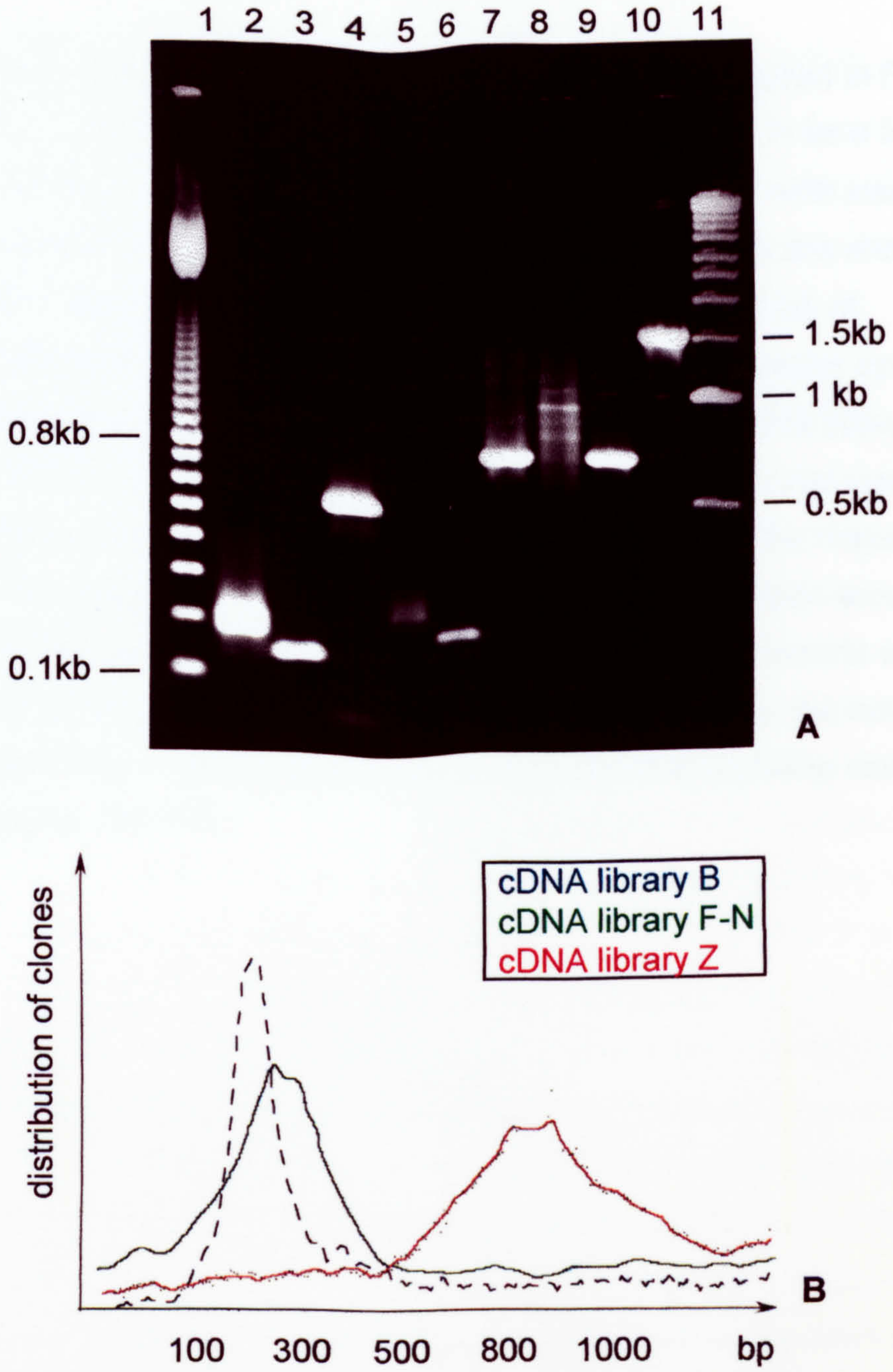


Figure 3.6. Evaluation of the insert-size distribution in different RPE cDNA libraries.

A. Agarose/TBE electrophoresis gel of PCR amplification products. PCR has been performed on diluted aliquots of phage lysate stocks of whole cDNA libraries and on some individual plaques using vector specific primers. lane1 - 100 bp ladder; lane 11 - kb ladder; lane 2 - library B (bovine fresh RPE cDNA in λTriplEx vector); lane 3 - λTriplEx vector - no insert; lane 4 - clone previously randomly isolated from library B; lane 5 - library F-N (bovine fresh RPE cDNA in λMOS*Elox* vector); lane 6 - λMOS*Elox* vector - no insert; lane 7 - clone randomly isolated from library F-N; lane 8 - library Z (human foetal cultured RPE cDNA in λUni-ZAP vector); lanes 9 and 10 - clones isolated from library Z.

B. Graphical representation (produced using NIH Image software) of distribution of clones as a function of insert-size in cDNA libraries B, F-N and Z (lanes 2, 5 and, respectively, 8 from above).

Both cDNA libraries of freshly dissected RPE cells analyzed in Figure 3.6 (library B in λTriplEx in lane2 and library F-N in λMOS*Elox* in lane 5) had the majority of inserts less than 300 bp long. Although clones with inserts of over 500 bp were isolated from these libraries (two of these are shown in the same figure, lanes 4 and 7, respectively), the fact that the bulk of amplification products were only 100-200 bp longer than the vector-specific amplification product (lanes 3 and 6, respectively) confirmed that these libraries contain a low percentage of clones with informative-sized inserts. Library Z, on the other hand, showed a distribution (lane 8) of the majority of its inserts in the range 500-1100bp, with a peak of the distribution around 850 bp. The PCR amplification products of two clones having inserts within these range are shown in lanes 9 and 10. Given these results, the decision was taken that library Z was the best source for randomly isolating clones for EST analysis in the RPE.

3.2. EXPRESSED SEQUENCE TAGS IN RPE

This section describes the results of the partial sequencing of clones from RPE cDNA libraries and the characteristics of the expressed sequence tags (ESTs) obtained.

3.2.1. GENERATION OF PARTIAL SEQUENCES OF RPE CLONES

3.2.1.1. Strategy for producing partial sequences

The approach for expression profiling in this study involved partial sequencing of randomly isolated cDNA clones, and aimed to form a collection of ESTs of the RPE. The assumption was that the transcriptional picture thus defined can contribute in two ways to our understanding of this tissue. Firstly, by identifying genes that are expressed in the RPE, it is reasonable to assume that the respective gene expression products (of previously known function or of novel function) are involved in an activity of this tissue. Secondly, the frequency of expression of various transcripts, measured by the recurrence of the same sequence in the transcriptional picture, can indicate the abundance of the corresponding transcripts in the mRNA population.

ESTs were generated by single sequencing reactions for each clone isolated. The partial sequence obtained was used for searching and comparison with database sequences. The clones subjected to sequencing were chosen at random but the general strategy for obtaining templates for sequencing was determined by the distribution of the insert-size in the cDNA libraries available and is described below.

Recombinant phage clones from bovine RPE cDNA libraries were first subjected to insert-size selection, due to the predominance of clones with short inserts in the libraries. Recombinant plaques were picked at random and eluted in SM buffer. The insert sequences present in the isolated clones were amplified by PCR using vector-specific primers (see Table 2.6). The PCR analysis enabled identification of phage clones

containing informative-sized inserts and estimation of the inserts size. Individual recombinant phages were then converted to plasmid subclones by *in vivo, cre*-recombinase-mediated excision accomplished by plating the phages on the *E. coli* non-expression strains BM25.8 (Amersham) and BNN132 (Clontech). Plasmid-bearing colonies were selected on LB-carbenicillin (50µg/ml) or LB-ampicillin (50µg/ml) agar plates and used for plasmid preparations. To avoid problems associated with plasmid DNA purified from *endA*+ (endonuclease A+) *E. coli* strains, such as BM25.8, recombinant pMOS*Elox* plasmids used for sequencing were purified after transformation of the respective clones in *E. coli* DH5α strain (*endA1*).

In the case of library Z (human foetal RPE cDNA library), which had a high percentage of recombinants containing cDNA inserts over 300 bp (the size considered to be informative for subsequent characterization of the respective cDNA), phagemids were obtained through mass excision from the phage library. To ensure statistical representation of the excised clones, the amount of amplified phage library (titre $1.7x10^9$ pfu/ml) subjected to mass excision was 6µl, i.e., approximately 10^7 pfu, representing ten times the primary library size (estimated in Dr. Bok's laboratory as 10^6 pfu). The excision was carried out using XI1-Blue MRF' *E. coli* cells at a 10:1 cells:phage ratio and ExAssist helper phage at a 1:1 helper phage:cells ratio. Aliquots of the stock containing the excised phagemids were used to infect SOLR *E. coli* cells in which the phagemids replicate. Plasmid-containing bacterial colonies were selected on LB-ampicillin (50µg/ml) agar plates. The titre of the excised phagemid library was $5x10^6$ cfu/ml. Individual, well isolated colonies were selected for plasmid preparations which were subsequently used directly for sequencing. Usually, no characterization of the size of cDNA insert was carried out prior to sequencing.

3.2.1.2. Template preparation

Since sequencing is very sensitive to template quality and quantity, special attention was given to plasmid preparation and quantitation. Double-stranded DNA templates for sequencing were produced using two

preparation methods: Qiagen Plasmid (mini preparations) and Kristal™-Cambridge Molecular Technologies Ltd.

Plasmid DNA obtained from minipreparations starting with 1.5ml overnight bacterial cultures was resuspended in 50µl 10mM Tris-HCl, pH 7.6. In order to determine the integrity of plasmid DNA and to estimate the concentration of preparations, 1-2µl aliquots were subjected to TBE/ agarose electrophoresis and quantified by comparison with standards. Figure 3.7.A represents a typical analysis of 15 plasmid preparations, pBluescript-based, originating from library Z. The concentration of DNA preparations was in the range 0.1µg-0.3µg/µl, indicating a plasmid DNA yield of 5µg-15µg/1.5ml culture. Panel B of Figure 3.7 shows the result of a double digestion *KpnI/SacI* of pBluescript-based plasmids. In each case, the double digestion resulted in a single vector-specific fragment of about 3kb and *KpnI/SacI/* insert fragment(s).

Appropriate aliquots of plasmid DNA preparations were used for sequencing. Template concentrations were not adjusted. In the case of low-yielding templates, depending on the total amount of plasmid DNA obtained, either an increased volume of the mini preparation was used for sequencing, or the respective templates were not sequenced.

Table 3.6 presents the summary of template preparations assessed by the number of successful sequencing reactions performed. Both DNA template preparation methods used produced double-stranded templates of comparable quality for automatic sequencing yielding a success rate of about 66%. Overall, considering both manual and automatic sequencing reactions performed, 67.7% of the reactions were successful in producing readable, useful sequences.

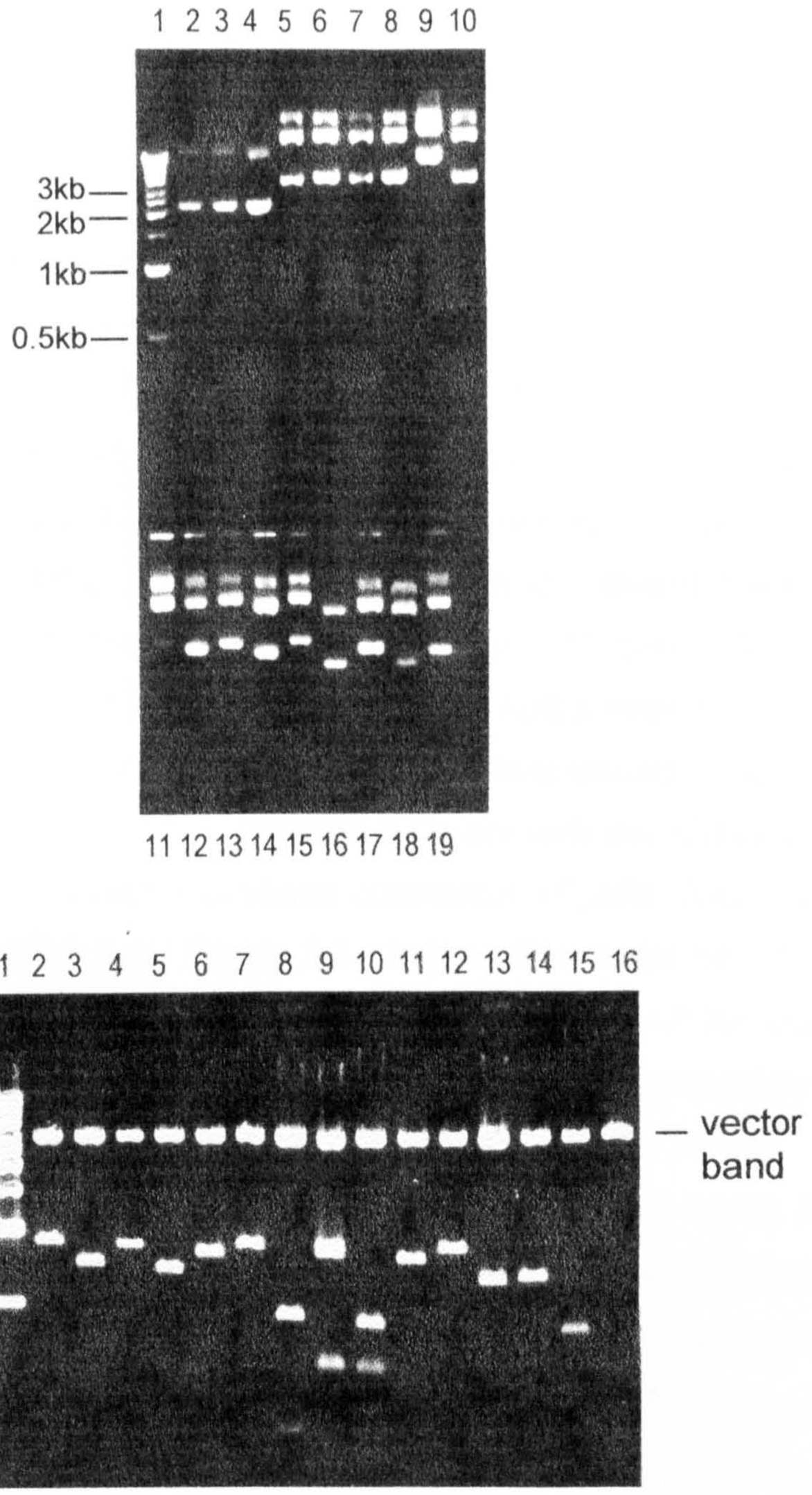


Figure 3.7. Analysis of plasmid DNA preparations.
A. Agarose gel TBE/ethidium bromide electrophoresis of undigested plasmid DNA samples (origin library Z). 1µl out of 50µl final volume was loaded in each well (lanes 5-19).
Standards (pUC18): 0.2µg in lane 2; 0.3µg in lane 3; 0.5µg in lane 4.
Size marker: kb ladder - lane 1.
B. Restriction analysis of pBluescript-based plasmids (see Figure 2.6). 0.5µg plasmid DNA was digested with *Kpn*I/*Sac*I and the resulting fragments were analysed by electrophoresis on 1% agarose/TBE gel containing ethidium bromide (lanes 2-15; pBluescript plasmid with no insert - lane 16).
Size marker: kb ladder - lane 1.

3.2.1.3. DNA sequencing and processing of sequence files

At start, sequencing was performed manually. However, shortly afterwards, a facility for automatic sequencing (Perkin-Elmer ABI-PRISM™) became available in the Department and thereafter this system was used for most of the templates. After carrying out trial sequencing reactions with primers listed in Table 2.6 (chosen according to the nature of vector), the best results in terms of length of unambigous sequence obtained were achieved with T3 primer (pBluescript-based plasmids), T7 gene 10 primer (pMOS*Elox* -derived plasmids) and 5'-amplimer (pTriplEx-based templates). These primers were therefore used for sequencing. The human foetal RPE cDNA library is a directional library with the cDNA inserts cloned into the Uni-ZAP XR vector in sense orientation (*Eco*RI - *Xho*I) with respect to the *lacZ* promoter (see Figure 2.6). In the pBluescript SK (-)-based phagemids which result by *in vivo* excision from Uni-ZAP XR clones, the *lacZ* transcription proceeds from *Sac*I to *Kpn*I. Therefore, sequencing recombinant double-stranded pBluescript phagemids with T3 primer produced the sequence of the plus (sense) strand of the cDNA insert. Also, sequencing from the 5' end of inserts was considered beneficial, being more likely to result in protein coding sequence and thus increasing the likelihood that database searches lead to assignment of putative identification.

A total of 300 sequencing reactions were performed (44 by manual and 256 by automatic procedure; see Table 3.6) on 272 individual clones (30 isolated from the bovine cDNA libraries and 242 from the human cDNA library). In order to improve the quality of the read sequence, 28 clones were subjected to sequencing a second time. 203 good, interpretable sequences were obtained (25 bovine, 178 human) and further analyzed.

Table 3.6. Template preparation summary for manual and automatic sequencing reactions.

Preparation method	No. sequencing reactions performed	No. successful	% successful
Manual sequencing†			
Qiagen	44	34	77.2
Automatic sequencing‡			
Qiagen	142	94	66.2
Kristal	114	75	65.8
Total	300	203	67.7

† - Manual sequencing reactions were performed on all clones isolated from bovine cDNA libraries and on a small number of clones from the human library.

‡ - Including 5 sequencing reactions for which manual sequencing has failed and 23 repeated automatic sequencing reactions.

Sequencing reactions were considered successful when yielded at least 150 nucleotides with fewer then 5% ambiguous base calls. The majority of sequencing reactions yielded much better data than this.

All bovine RPE partial sequences and approximately 5% of the total number of human RPE EST generated in this study were produced by manual sequencing. The sequencing reaction products were analysed by electrophoresis in 8% acrylamide/urea/TBE gels (see 2.13.1.b). The dried gels were exposed to X-ray film for 24-48 hours at room temperature. The sequences were read by direct examination of X-ray films. Reads of 200-250 nucleotides per clone were routinely obtained from one gel. In short-run sequencing gels (run for 2 hours at 30mA, 1500V) the sequence of vector, multiple cloning site and adaptor comprised up to 150 nucleotides (depending on the type of vector and primer used). Thus, the insert sequence obtained after a short-run electrophoresis had a length of 75-140 nucleotides. Further stretches of insert sequence up to 250 nucleotides were obtained by long-run (4 hours) electrophoresis of sequencing gels. Vector and adaptor sequences were removed from the final sequence.

The great majority of human RPE partial sequences were generated by automatic sequencing. The sequencing reaction products were analysed using ABI 373 and 377 DNA sequencers and version 1.2.1 data collection and base calling software (as described in 2.13.2.a). Since the extensive use of the automatic sequencing facility related to this sequencing project started shortly after the implementing of this facility in the Department, a series of test runs examined the influence of template preparation method, concentration of template and primer. The results of this optimization are incorporated in the description of the procedure (see Methods section 2.13.2.a). The output files contained between 450 and 700 nucleotides.

Usually, a portion of 100-150 bases at the end of run had a high percentage of ambigous base calls, due to peaks broadening and overlapping. Consequently, the 3' end of each sequence was cropped manually using Factura sofware (see 2.14.1). The vector sequences up to the cloning site and the overall clear sequence data were identified by submitting batch worksheets of 5-10 sequence files to automatic processing within Factura. The software settings were such that bases were removed from the beginning and the end of sequence until no more than one ambiguity remained out of 20 bases and the sequences with more than 10% ambiguities were rejected. Table 3.7 summarizes some of the features identified by Factura processing, compiled from the reports of batch analysis. Further processing of sequence files was done manually and included: adaptor sequence identification and removal; visual scanning of the whole sequence with particular attention to ambiguous bases; calling, when possible, of ambiguous bases by inspection of the electropherogram in the output file.

Table 3.7. Summary of sequence features identified by Factura.

Feature	Description
Vector range	up to 75 - 102 bases; for 78 clones, vector sequences were identified at the end of run
Ambiguity range (over 10%)	up to 45 - 57 bases and the last 150-200 of submitted sequence
Clear sequence length (within the confidence range)	140 - 698 bases
% ambiguities in the clear sequence	1% - 9%

3.2.2. ANALYSIS OF ESTs

3.2.2.1. Identification of ESTs

The analysis of sequence files resulted after the processing described in the previous section had two conceptually different parts: (1) automated analysis of clear, i.e.-insert only, sequencing data; and (2) evaluation of sequence comparison searches. The two parts differ significantly in the amount of human intervention involved. In the first part of the analysis which is the actual database searching, this intervention is restricted to choosing the algorithm of searching and to submitting the right format of files. The second part of analysis is achieved exclusively by individually inspecting and assessing the alignments.

All clear sequencing data obtained in this study were used for sequence comparisons against sequences in GenBank, EMBL and Swiss-Prot databases. The searches were performed in batch mode using FASTA and BLAST programs of the GCG package (computing facilities used are listed in section 2.14.2). Following analysis of the first sequences, it emerged that FASTA program is more suitable for the principal goal of this analysis, i.e.- identification and assignment of ESTs. Thus, the format of

FASTA output files enables inspection of various alignments along the whole sequence length. BLAST results, obtained otherwise more quickly and requiring less computational resources, reveal only the portion(s) of the sequence with the highest scores of matching. Therefore, while the BLAST approach proves very useful for assigning ESTs that have a high (i.e., generally over 80%) similarity with database entries, assessment of the rest of the ESTs is very difficult and usually requires a subsequent more sensitive comparison, of the FASTA type, supported by a different algorithm and yielding a different presentation of the result. Consequently, it appeared that performing FASTA searches from the beginning, for the number of sequences considered by this study, could be advantageous.

The assignment of ESTs was based on the annotation of database entries to which a match was found. Theoretically, the comparison searches can define three categories of results: (i) exact or good matches that allow an unambigous identification of ESTs; (ii) non-exact or partial match with database sequences; and (iii) low or uncertain levels of similarity with existing database entries. The fact that the search results of more than half of the analyzed ESTs fell into the latter two categories made the evaluation of sequence comparison searches more difficult and time-consuming than previously expected.

"Perfect" (exact) matches were represented by a minimum percentage match of 95% over the entire aligned region which was at least 100 bp long. To account for the relatively low accuracy expected of single-run sequencing, matches of over 80% identity along a minimum of 70% of the query sequence were considered good and allowed the straightforward assignment of respective ESTs. In several cases, ESTs matched two unrelated genes in the databases and hence were likely to be chimaeric. The identification of ESTs was given in connection with the coordinates of the sequence overlapping the assigning gene.

In all other cases (categories (ii) and (iii) of results above), in order to minimize false assignments, various searches were performed and evaluated (FASTA and/or BLAST searches against nucleotide and/or peptide databases). FASTA results allowed inspection of sequence similarity along the whole length of the analyzed sequence and multiple

alignments were assessed. When the degree of similarity with a database entry, over a relevant length (i.e., at least 100 bp, but usually 150 bp), was in the range 60%-80%, the corresponding EST was designated "similar to" the respective database sequence. The same designation was used, also, for matches with sequences of a different mammalian species. Percentage of match below 60% over the aligned sequence was regarded as an indication of no relevant similarity and the clones were "unassigned". Among the "unassigned" clones could be unique sequences of novel genes, not similar to other sequences in current databases.

Six RPE cDNA libraries (five bovine and one human) were used in this study to isolate cDNAs for identifying ESTs of RPE. Firstly, a small number of bovine ESTs were characterized from the in-house constructed cDNA libraries. The characteristics of these ESTs are presented in the next section. Once the human RPE cDNA library became available, due to its high percentage of recombinants with longer inserts (see 3.1.2), the emphasis was shifted towads isolating cDNA clones from this source. Therefore, the great majority of ESTs analyzed originated from the human library and the results of their analysis are presented in section 3.2.2.3.

3.2.2.2. Bovine RPE ESTs

Plaques from all five primary bovine RPE cDNA libraries constructed in-house (see Table 3.5) were randomly selected and subjected to PCR to determine those that have informative-sized inserts. The results of PCR and sequencing analysis are summarized in Table 3.8.

Table 3.8. Summary of PCR and sequencing analysis of clones randomly selected from primary bovine RPE cDNA libraries.

cDNA library	No. of plaques randomly selected	Templates sequenced (selected after PCR)	Total partial sequences obtained	rRNA ESTs	adaptor-multimer and polyA† clones	Total informative ESTs‡
F-N	22	7	6	—	2	4
B	20	5	3	1	1	1
F-T	30	9	8	1	3	4
A2	18	5	5	2	1	2
C1	18	4	3	—	2	1
TOTAL	**108**	**30**	**25**	**4**	**9**	**12**

† PolyA clones describe stretches of over 70 As only; two such clones were isolated from library F-T.

‡ Informative ESTs include sequences after removal of rRNA ESTs, adaptor-multimer and polyA clones.

From the total of 108 λMOS*Elox*- and λTriplEx-based recombinant plaques isolated, 30 were selected to be sequenced as they appeared to have an insert over 150bp long. The cDNA libraries sampled varied only slightly with respect to the ratio of clones containing informative-sized inserts to total number of clones randomly selected. This ratio was, roughly, between 1/4 and 1/3. All 30 phage clones selected were subjected to *in vivo* excision of plasmids which were then used as templates for manual single-run sequencing. Among the 25 partial sequences obtained,

four were from ribosomal genes and nine were represented by adaptor-multimers and, in the case of oligo(dT)-primed cDNA library F-T, long stretches of As flanked by adaptors. The remaining 12 sequences were regarded as informative, nuclear-coded ESTs, nine being of freshly dissected RPE cells (cDNA libraries F-N, B and F-T) and three originating from cultured RPE cells (cDNA libraries A2 and C1).

The results of sequence comparison searches of the informative bovine RPE ESTs against sequences in GenBank/EMBL databases are presented in Table 3.9. The ESTs were assigned to possible cellular role categories based on the classification used by Venter et al (Adams *et al.*, 1995). The indicated tissue distribution of homologous sequences is based on data from the same study cited above. Ten ESTs presented a relevant match (according to criteria described in 3.2.2.1.above) with an existent database sequence and were putatively identified. Out of these, seven ESTs (f1, f7, f.b3,T4, c1, c4 and c22) matched reported bovine gene sequences, with the percentage of identity over the aligned sequence ranging between 83.5% and 92.9%. Three other ESTs (f31.2, f31.6 and T2) were similar (percentage of identity between 78.3% and 92.3%) to clones from other species, namely mouse and human. This may be due to the fact that no corresponding bovine clones have been reported. Nine bovine ESTs were assigned to cellular role categories: 1-general metabolism; 3-gene/protein expression; 2-cell structure/motility; and 3-cell signalling/cell communication.

Table 3.9. Summary of results of sequence comparison of bovine ESTs with database sequences.

Clone	cDNA library of origin	Sequenced insert length (bases)	Match accession number	Putative identification	%ID (bp overlap)†	Tissue distribution‡
‡General metabolism:			**transport**			
c4	A2	92	GB_OM:BOVT1TRANS	bovine ADP/ATP translocase T1	87% (92)	Widely expressed
Gene/protein expression:			**protein synthesis: ribosomal proteins**			
f1	F-N	97	EM_OM:AF013214	Bos taurus acidic ribosomal phosphoprotein PO	86.3% (102)	Widely expressed
			protein synthesis: modification/targeting			
f31.6	F-N	117	EM_EST1:AA636556 EM_HUM1:HS417401	similar to a mouse cDNA clone and, respectively, to human trans-Golgi p230 protein	92.3% (117) 70% (113)	–
			protein synthesis: protein turnover			
T2	F-T	221	GB_PR:HUMPSC5	similar to human proteasome subunit HC5	78.3% (207)	Widely expressed
Cell structure/motility:			**general**			
f.b3	B	108	GB_OM:S77716S1	bovine laminin B1	92.9% (112)	Brain Uterus
			cytoskeletal			
c22	C1	194	GB_OM:BOVACT1	bovine cytoplasmic β-actin	90% (194)	Widely expressed
Cell signalling/communication:			**cell adhesion**			
f7	F-N	110	GB_OM:BOVPS	bovine P-selectin	83.5% (103)	Liver Testis
			intracellular transducers			
T4	F-T	209	GB_OM:BTAJ3123	Bos taurus FGF-2 binding protein	91% (178)	Brain Uterus
c1	A2	113	EM_OM:BTNA2	bovine transducin α-2	92.7% (110)	Retina
Unclassified:						
f31.2	F-N	78	EM_EST5:C80102	similar to 3'end EST Mus musculus blastocyst cDNA	85.5% (76)	–
T7	F-T	244	GB_PR2:HSAC02	unassigned (similar to human β-actin related pseudo-gene)	71.2% (153)	–
T8	F-T	229	EM_HTG:AC002978	unassigned (similar to human unassigned clone)	64.9% (211)	–

cDNA libraries F-N, B and F-T are of freshly dissected RPE cells and libraries A2 and C1 are of cultured RPE cells.
† bp overlap indicates the length of the best alignment; when the indicated overlap is longer than the query (EST) sequence, it includes the gap introduced by FASTA program for obtaining the best score.
‡ Cellular role categories and tissue distribution of the identified homologous/similar sequences are as described in (Adams, 1995 #180).

Two ESTs (T7 and T8) could not be assigned a putative identification based on the results of comparison searches, performed both at nucleotide and peptide level. However, possible open reading frames were identified for both of these sequences: T7 - starting at nucleotide 176 of the EST with no stop codon present for the rest of the sequence; T8 - containing the 3' end of a putative open reading frame ending at nucleotide 112 of the EST. In the case of all 12 ESTs described, the 3' end of the sequence read did not emerge into the vector sequence, indicating the presence of an insert longer than the determined EST. The presence of a photoreceptor OSs-specific clone (transducin α-2) may indicate a possible contamination. No clones matched known bacterial sequences indicating no bacterial contaminating sequences present.

3.2.2.3. Analysis of human RPE ESTs

Single-pass sequencing

288 antibiotic-resistant colonies were randomly selected following the mass excision of phagemids to which an aliquot of the human RPE cDNA library was subjected as described in section 3.2.1.1. above. Plasmid mini-preparations were carried out for all isolated clones. 46 preparations were considered unsuitable for sequencing, due to low yields and/or possible contamination with plasmid or genomic DNA.

A total of 242 human RPE clones were analyzed by single-pass sequencing. Of these, a set of 64 sequences did not meet the standards of sequence quality (% ambiguity over 10%) and were rejected from further analysis. The remaining 178 sequences were examined by the procedure described in section 3.2.1.3. The results of this analysis are shown in Table 3.10. Four sequences were identified as being pBluescript SK(-) vector.

Database analysis

The average length of the cDNA sequences used for comparison was about 350 bp. The sequence of each cDNA clone was compared with sequences in nucleotide and peptide sequence databases, using the programs FASTA and BLAST. Most sequence comparisons were carried

out at nucleotide level with sequences in Genbank database, using the FASTA algorithm at the UK HGMP-RC, Hinxton. Comparisons of six-frame conceptual translations of query sequences with the peptide sequences in the SwissProt database were performed using the program BLASTX. Results of database searches were inspected individually and the alignments evaluated. Sequences that, after an initial BLAST search, appeared less than 80% similar (over the length used for comparison) to known genes/ESTs were subjected to a repeated comparison using FASTA program. Putative identifications were given based on database annnotations of match sequences. The sequences of the clones with no database match were compared against one other to identify those encoding overlapping regions of the same cDNA. The accuracy of the single-pass DNA sequencing, as determined by comparison of the sequences of assigned human RPE cDNA clones from this study with known genes and ESTs in the databases, was estimated at approximately 96%.

Table 3.10. Overview of human RPE sequences analyzed in this study.

Total sequences analyzed	178
Length of ESTs	102 - 698 nt
ESTs < 200nt (%)	9 (5%)
ESTs with %N > 5% (%)	16 (9%)
Vector only	4
rRNA and mitochondrial DNA†	15
Repetitive DNA	3
Chimaeric clones	2
Total informative ESTs‡	154

nt - nucleotide; %N - % ambiguities.

† - Includes mitochondrial DNA for several tRNAs, rRNA and a loop attachment sequence.

‡ - After removal of vector only, rRNA and mitochondrial DNA†, repeats-containing, and chimaeric clones.

Three clones matched database entries containing Alu/dinucleotide repeats. Since Alu elements do not generally occur in protein-coding regions, but in intergenic/intronic sequences, 3'-untranslated regions or unspliced precursor RNAs, their frequency of appearance in an EST study is used as an indicative of the degree of potential genomic or unprocessed RNA contamination of the cDNA library. Hence, the human RPE cDNA library used in this study appears to have a low level of such contamination (3 out of 174 clones analyzed, i.e., 1.7%).

Identification and characterization of ESTs

The results of database searches for the 154 informative RPE ESTs are shown in tables 3.11- 3.14 organized as follows:

75 ESTs putatively identified as sequences of known/presumed function or localization (Table 3.11; summary of distribution into cellular role categories is presented in Table 3.12);

51 ESTs matching previously identified ESTs with no function assigned (Table 3.13); and

28 ESTs with no homology to known database sequences (Table 3.14).

Before proceeding to present these results, it is of importance to mention a few points. After evaluating the results of sequence comparison of the first batches of sequences obtained, it became clear that the attempt to elucidate the nature of sequences showing less than 60%-65% identity with database sequences would be of little benefit for this project. Such an endeavour could generate only a very limited, if any, amount of useful information which would, also, have a high degree of speculation being based mostly on mathematical means of deducing "relevant" biological information. The recognition of identifiable ESTs in the RPE was of far more importance. Hence the sequences less than 60% identical, over a region of at least 150 bp, with database sequences (human or of another species) were considered to have no useful homology. When the percentage identity with the best scored match in a search was in the range of 60%-70%, other characteristics of the alignment were considered, such as the gaps introduced in the alignment to produce that score and the length of the

longest uninterrupted string of identical nucleotides. In the listing of putative identifications (Table 3.11), the expression "similar to" was used to describe these matches for six ESTs.

Of the 154 informative ESTs, 75 (48.7%) matched the sequences of 62 genes encoding proteins of known or presumed function/localization recorded in the public databases (for list, see Table 3.11). Six ESTs were identified as derived from mitochondrial-encoded mRNAs. Eight transcripts (5.2%) were represented more than once. The most frequently isolated clones were those encoding transthyretin (serum thyroxine binding prealbumin) and those similar to kappa casein precursor, each of which was isolated 4 times. Three ESTs matched the sequence encoding cystatin C, a cysteine proteinase inhibitor. The five genes represented by two ESTs each encode a tyrosinase-related protein, ribosomal protein S14, ribosomal protein S9, a translationally controlled tumor protein and the mitochondrial cytochrome c oxidase subunit IV. The ESTs matching the same gene/EST but differing at the 5' end of the sequence were considered independent (not the result of library amplification) and were, therefore, all individually listed in Table 3.11.

The putatively identified ESTs were classified with respect to the biological roles of proteins encoded by the matched genes (Table 3.11). The six categories of biological roles used for the classification (metabolism, gene/protein expression, cell structure/motility, cell/organism defense, cell signalling/cell communication, and cell division) and their subcategories are as defined by Adams *et al.* (1995). The final category in the listing contains the "unclassified" ESTs, i.e. the ESTs matching sequences encoding proteins for which no definitive information regarding their cellular role was found. Such a classification should, obviously, be regarded as tentative, as it is constantly owing to the advancement of knowledge on cellular roles/functions of genes and proteins. Each putative identification is given alongside examples of the tissues from which the match cDNA/EST was previously isolated. As for the assignment to biological role categories, the examples of tissue distribution of ESTs were based on the information summarized by Adams et al. (Adams *et al.*, 1995) and on the information contained in the annotations of database entries (as in March 1998) and in searches of literature.

Table 3.11. Summary and classification of putatively identified human RPE ESTs.

The RPE ESTs are sorted according to cellular roles/functions of proteins encoded by the matched sequences.
The match accession number indicates the best match to a sequence in a public database. The percentage identity is given alongside with the length in base pairs of the nucleotide sequence match.
Tissue codes (for the examples of tissue distribution of the match sequence) are: a - breast/mammary gland; b - brain (including amygdala, cerebellum, hippocampus, pineal gland); c following any code indicates carcinoma/tumour; d - endothelial cells; e - embryo; f preceding any code indicates fetal; g - colon; h - heart; i - kidney; j - gall bladder; k - skin; l - lung; m - skeletal muscle; n - bone; o - ovary; p -pancreas; r - prostate gland; s - fibroblast; sm - smooth muscle; t - placenta; u - uterus; v - liver/spleen; w - white blood cells (including macrohages, T cells, monocytes); x - parathyroid; y - eye (including retina, RPE, cornea, ciliary body); z - testis.
HeLa and HL-60 - cell lines.

Clone/ EST	Length of EST (bases)	Match accession number	Putative identification	% ID (bp overlap)	Tissue distribution
Metabolism: amino acid					
K73	362	U09202	ornithine decarboxylase antizyme (Oaz)	88.7(265)	a, b, d, e, g, h, i, j, l, m, o, p, sm, t, v, w, y
K203	528	X51420	tyrosinase-related protein	92.5(516)	b, e, kc, y
L26	588	X51420	tyrosinase related protein	89.5(568)	b, e, kc, y
Metabolism: energy/TCA cycle					
K170	507	X69907	ATP synthase c subunit (P1 form), mitochondrial	92.3(489)	b, j
K40	466	M21575	cytochrome c oxidase COX subunit IV, mitochondrial	98(463)	b, e, g, p sm, t, v, w, y, z
L21	534	U90915	cytochrome c oxidase COX subunit IV, mitochondrial	90.3(527)	b, e, g, p, sm, t, v, w, y, z
K60	457	M16461	similar to NADH-cytochrome b5 reductase	66.6(350)	t
K34	461	M57911	NADH dehydrogenase subunit 2, mitochondrial	98.2(460)	b, m, p, w, z

Table 3.11. continued

Clone/ EST	Length of EST (bases)	Match accession number	Putative identification	%ID (bp overlap)	Tissue distribution
Metabolism: lipid					
K82	295	M16695	apolipoprotein D	87(294)	a, b, l, m, o, t, u, w, y
K38	472	D25283	galactocerebrosidase	91.9(472)	a, b, e, fh, j, l, n, t
K55	449	X54516	lipoprotein lipase	95.3(448)	b, n, fb, t, w
K174	483	L11702	similar to phospholipase D	76.7(467)	a, b, l, o, v,y
Metabolism: sugar/glycolysis					
K84	341	M31991	aldehyde dehydrogenase	81.3(342)	b, e, k, l, n, p, t, y
K161	314	J04469	creatine kinase, mitochondrial	90.2(314)	b
K67	309	L43581	iduronate-2-sulfate sulfatase (IDS)	88.9(279)	b, d, l, v, w
K36	365	M23725	M2-type pyruvate kinase	92.5(362)	a, b, d, e, h, j, k, l, n, y
Gene/protein expression: RNA synthesis					
L36	482	Z37166	BAT1-nuclear RNA helicase (DEAD family).	92(436)	a, b, d, e, g, h, m, o, t, v, w
K213	537	AF056087	secreted frizzled-related protein	85.7(440)	a, h, k, l, o
K47	463	L14599	cDNA encoding protein homologous to splicing factor PSF	96.3(463)	HeLa
Gene/protein expression: protein synthesis: post-translational modification					
K70	323	L11667	cyclophilin-40	96.9(322)	b, e, l, t, w, y
K135	232	M22382	matrix protein P1 (nuclear encoded), mitochondrial	92(232)	HL-60

Table 3.11. continued

Clone/ EST	Length of EST (bases)	Match accession number	Putative identification	%ID (bp overlap)	Tissue distribution
Gene/protein expression: protein synthesis: protein turnover					
K63	464	M11465	α-1 anti-trypsin	91.4(464)	e, u, v
K53	450	X87212	cathepsin C	96.4(450)	d, e, j, k, p, u, v, y
K22	364	X05607	cysteine proteinase inhibitor precursor cystatin C	93 (356)	a, b, d, e, g, i, o, p, r, t, v, y, z
K26	469	X05607	cysteine proteinase inhibitor precursor cystatin C	94 (465)	a, b, d, e, g, i, o, p, r, t, v, y, z
K210	447	X05607	cysteine proteinase inhibitor precursor cystatin C	97.6(423)	a, b, d, e, g, i, o, p, r, t, v, y, z
K136	400	S48220	similar to type I 5' iodothyronine deiodinase	82.3(276)	HeLa
K2	140	X95586	similar to MB1 gene; proteasome	78.4(134)	a, b, d, e, g
K6	150	D26512	membrane type matrix metallo-proteinase (MT-MMP)	99 (146)	b, u, y
K226	377	U02571	tissue inhibitor of metalloproteinase-3 precursor (TIMP-3)	93(377)	b, e, h, l, m, n, o, t, u, v, w, y, z
Gene/protein expression: protein synthesis:ribosomal proteins					
K90	365	U02032	ribosomal protein L23a	89.9(365)	r
K14	388	X73460	ribosomal protein L3	93.8(388)	b, d, g, h, m, n, v, y
K208	658	X73974	ribosomal protein L4	91(513)	a, b, e, g, h, l, m, o, t, u, y
K223	644	X69391	ribosomal protein L6	87.2(632)	z
K148	497	Z28407	ribosomal protein L8	96(497)	b, fl, g, h, k, l, sm, n, o, uc, y, z
K138	311	M13934	ribosomal protein S14	93(242)	a, b, d, e, g, h, i, j, k, l, m, n, o, p, t, u, v
K195	439	M13934	ribosomal protein S14	88(432)	a, b, d, e, g, h, i, j, k, l, m, n, o, p, t, u, v

Table 3.11. continued

Clone/ EST	Length of EST (bases)	Match accession number	Putative identification	%ID (bp overlap)	Tissue distribution
K165	432	X62691	ribosomal protein, homologous to yeast S24	87(431)	a, b, e, g, h, j, l, n, o, p, y
K191	698	U14970	ribosomal protein S5	97(698)	b, d, e, g, h, j, k, l, n, o, t, y
K62	461	M77233	ribosomal protein S7	98.8(242)	a, b, e, g, i, j, l, n, o, t, v
K18	425	U14971	ribosomal protein S9	91.6(416)	a, b, g, h, k, lc, m, n, oc, sn, t, v, z
K154	495	U14971	ribosomal protein S9	90(495)	a, b, g, h, k, lc, m, n, oc, sn, t, v, z
Cell structure/motility					
K193	498	X51521	ezrin	94(498)	HeLa
K39	450	X79683	laminin β2 chain	91.7(385)	b, v, t, z
K9	140	L19686	migration (glycosylation) inhibitory factor	97 (138)	y, w
Cell structure/motility: extracellular matrix					
K87	323	M32306	similar to epithelial glycoprotein	76.5(299)	b, k
K31	355	AA359338	EST similar to collagen, type X, alpha 1	98.5(224)	l, k
K225	538	J03040	SPARC/osteonectin	89.6(540)	a, b, d, e, g, h, i, j, m, n, r, t, v, y, z
Cell/organism defense					
K88	401	X14723	complement-associated protein SP-40	86.4(345)	b, d, h, l, o, p, v, y
K51	447	U12404	Csa-19(gene downregulated by cyclosporin A)	95.5(447)	a, b, d, e, g, h, j, l, n, o, t, v, y
Cell signalling/cell communication: cell adhesion					
K205	519	U79251	cDNA clone 23878, similar to opioid-binding cell adhesion molecule L34774	95(520)	b, y

Table 3.11. continued

Clone/ EST	Length of EST (bases)	Match accession number	Putative identification	% ID (bp overlap)	Tissue distribution
Cell signalling/cell communication: channels/transport proteins					
K35	368	M11714	transthyretin (serum thyroxine binding prealbumin)	93.5(336)	b, e, j, p, v, y
K108	349	M11714	transthyretin	93.9(348)	b, e, j, p, v, y
K147	483	M11714	similar to transthyretin	82.8(464)	b, e, j, p, v, y
K209	411	U19780	transthyretin precursor	96.3(402)	b, e, j, p, v, y
Cell signalling/cell communication: cytokines					
K13	415	L35263	Csaids (cytokine-suppressive anti-inflamatory drug) binding protein	94 (415)	w
K10	296	M76979	pigment epithelium differentiation factor	92 (288)	a, b, e, h, i, j, l, n, o, t, v, y
Cell signalling/cell communication: receptors/intracellular transducers					
L10	448	AA316189	EST similar to ras-related protein rap1B	82.4(408)	a, b, d, e, g, k, l, n, p, r, v, w
K66	420	U14910	RPE-retinal G protein-coupled receptor (rgr)	92.8(419)	y
Cell signalling/cell communication: protein modification					
L18	399	L36151	phosphatidylinositol 4-kinase	95(317)	b, h, t, v
Cell division/cell cycle					
K4	170	P11862	similar to growth arrest specific protein gas2	79.8 (164)	e, fl, fv, v

Table 3.11. continued

Clone/ EST	Length (bases)	Match accession number	Putative identification	%ID (bp overlap)	Tissue distribution
Unclassified					
K32	341	W94062	EST similar to SW:YA22_SCHPO Q09695 hypothetical 36.7 kd protein c2f7.02c in chromosome I	98(311)	fh
K41	449	J02984	insulinoma rig-analog DNA-binding protein	92.7(449)	p
K64	453	U22897	similar to nuclear domain 10 protein (ndp52)	61.6(258)	nc
K117	342	AA158218	EST similar to gb:M73628 kappa casein precursor	86(342)	d, e, fy, g, k, m, pc
K134	434	AA188076	cDNA similar to similar to gb:M73628 kappa casein precursor	87.2(434)	d, e, fy, g, k, m, pc
K149	376	AA854942	cDNA similar to gb:X16064 translationally controlled tumor protein	80.4(377)	h, n, xc
K222	481	AA081174	cDNA similar to gb:M73628 kappa casein precursor	86.5(483)	d, e, fy, g, k, m, pc
K230	344	D10653	cell surface glycoprotein	85.1(323)	w
L6	425	L08850	AD amyloid(synuclein)	95.5(421)	b
L7	439	X63679	TRAMP protein (translocating chain-associating membrane protein)	95.9(444)	HeLa
L12	370	N25956	EST similar to gb:J02931 tissue factor precursor	91.2(317)	h, k
L15	462	AA081169	EST similar to gb:M73628 kappa casein precursor	93(457)	d, e, fy, g, k, m, pc
L16	429	AA602366	EST similar to gb:X16064 translationally controlled tumor protein	94.9(430)	h, n, xc
L32	573	X76534	NMB (gene expressed in low-metastatic human melanoma cell lines)	96.7(564)	h, pc, s, t, u, xc

A feature resulted from the sorting of the ESTs according to their biological roles is that a significant fraction (more than a third) of the total number analyzed represent genes with involvement in gene/protein expression (see Table 3.12). Moreover, among the 26 ESTs in this category, 12 ESTs are for ribosomal proteins and 9 ESTs represent genes with roles in protein synthesis/protein turnover.

Table 3.12. Distribution of identified RPE ESTs into cellular role categories.

Cellular role category	No. of ESTs
Metabolism	16
Gene/protein expression	26,
with main subcategories:	out of which:
protein synthesis/turnover	9
protein synthesis/ribosomal proteins	12
Cell structure/motility	6
Cell/organism defense	2
Cell signalling/communication	10
Cell division	1
Unclassified	14

Fifty-one sequences (33.1% from the total informative ESTs) matched anonymous sequences, i.e. sequences reported by cDNA/ESTs and genomic sequencing projects with no identification of function (Table 3.13). The database ESTs matched by RPE ESTs were identified in cDNA libraries of, among others, brain, retina, cornea, embryo, breast, placenta, uterus, pancreas, liver, testis, fibroblasts. Match sequences from genomic sequencing studies usually originated from "sequencing in progress" or "unordered pieces" stages of respective projects.

Table 3.13. Human RPE ESTs similar to anonymous sequences reported from cDNA and genomic sequencing projects.

RPE ESTs characterized in this study are listed in numerical order and the type of project from which the match anonymous sequences were reported previously is indicated as follows: c - cDNA sequencing projects; g - genomic sequencing projects.
The information related to match sequences is based on database sequence annotations. For match sequences originated from genomic sequencing studies the identification of clone is given when indicated in the database annotation. Tissue distribution refers to the source of cDNA libraries from which the match ESTs were previously isolated. For tissue codes, see Table 3.11.

Clone/ EST	Length (bases)	Match accession number	Project type	%ID (bp overlap)	Tissue distribution
K1	172	U82668	g (clone RG030H15 from 7q31)	73 (178)	
K3	102	AC004019	g (clone193B12; HTGS phase 1)	82 (73)	
K5	103	W53035	c	82 (96)	g, fh, p
K11	246	Y10000	g (16p11)	82(91)	
K15	376	AA316757	c	97(376)	v
K16	453	AA487106	c	92(402)	l
K17	141	AA888961	c	72 (141)	b, t
K19	424	U28998	g (chromosome 12q13-15)	94 (375)	
K20	467	AA937045	c (IMAGE clone 1535165)	96(357)	
K28	128	AL021920	g (clone 163M9; HTGS phase 1)	93(125)	
K30	472	AA436265	c	97(276)	b
K33	456	AA031868	c	96(331)	u
K37	306	Z72004	g (cosmid J256K24, chromosome X)	78(254)	
K42	338	W465991	c	96.5(328)	f, h, oc, pc
K43	327	W27426	c	99(123)	e, fy
K56	409	AA581420	c	97.1(409)	b, w
K58	469	Z42652	c	98.7(279)	b
K101	378	AA086213	c	98(305)	m
K112	190	B58891	g (clone 2014E20)	94.1(186)	
K120	492	AA149563	c	93.3(492)	b, e, g, t, u
K121	395	AA157626	c	85.1(402)	e, fy, g, s
K124	320	HSPD03620	c	94(320)	m
K126	352	M85895	c	98.6(291)	b

Table 3.13. continued

Clone/ EST	Length (bases)	Match accession number	Project type	%ID (bp overlap)	Tissue distribution
K128	372	AL021026	g (clone 127D3; HTGS phase 1)	81.3(219)	
K131	415	U82672	g (chromosome X, clone QC15B1)	69.4(307)	
K132	271	AA252570	c	89(151)	ot, s
K137	204	D82106	c	84(200)	b, m
K141	429	AA171767	c	84.1(287)	h, oc
K145	330	R47884	c	90.4(332)	ac, l
K155	571	AA775795	c	89(562)	fv
K185	396	D80741	c	92(301)	b, l, oc, v
K194	309	U82207	g (chromosome 10, clone CITB_119_P_3; HTGS phase 3)	67.9(209)	
K198	426	AA723015	c	95(296)	b
K201	492	W07727	c	98.1(492)	l, s, xc
K202	423	AA811822	c	90(418)	k, xc
K215	472	AC002403	g (11p15.5)	90.3(258)	
K216	475	AA156180	c	95.6(473)	d, k, lc, u
K217	543	Z73417	g (X)	88(544)	
K224	474	AA653822	c	93.2(398)	a, b, h
K228	405	AA167138	c	83.7(406)	b, d, g, h, oc, p
L5	431	AC002395	g (HTGS phase 1)	63.1(358)	
L9	415	AC002399	g (16p11.2)	74.9(179)	
L17	329	AA633234	c	95.5(331)	b, fy, lc, p
L22	526	AC003035	g (HTGS phase 1)	84.3(498)	
L24	336	AA481404	c	99(336)	fb, y
L25	364	AA078369	c	92(286)	p, t
L29	344	W23228	c	94(313)	h, t, v, y
L30	624	Z93016	g (20q12-13.2)	94(625)	
L31	426	AA558022	c (IMAGE clone 1042025)	95(271)	
L33	447	AA704332	c	98(419)	b
L35	465	AA447512	c	97(446)	z

The sequences of 28 human RPE cDNA clones, representing 18.1% of the informative ESTs studied, were not related to any sequences in the public databases. These ESTs for which no putative identifications could be made are listed in table 3.14. Comparison of the sequences of these putatively novel cDNAs to one another indicated no relevant similarities between them.

Table 3.14 . RPE ESTs without database matches.

Clone/EST	Length (bases)	Clone/EST	Length (bases)
K21	394	K143	396
K24	467	K150	298
K27	339	K153	413
K46	389	K156	546
K48	454	K180	465
K61	436	K192	499
K69	319	K197	356
K72	317	K200	263
K92	385	K212	310
K107	228	K220	479
K109	443	K227	504
K116	336	L1	324
K118	306	L2	432
K142	376	L11	361

No investigation regarding the possible existence of open reading frames in the novel sequences was made since such an attempt would have been unlikely to give definitive identification of ORFs given the fact that the sequences under scrutiny were derived from partial one-pass sequencing and, most likely, were not of full-length clones. Hence, the significance of predicted translations from detected start codons or from assumed start codons upstream of the sequence investigated would have been speculative.

Genes for functions associated with protein turnover

Identification of clones matching genes encoding proteins of known or partially known function in other tissues indicated that the respective genes are also specifically expressed in the RPE. Moreover, given the random nature of clone selection used in this study, the repeated

occurrence of any sequence may indicate a moderate or high level of expression for that sequence which, therefore, is likely to have specific importance for the function of RPE. In this context, a prominent feature that resulted from the sorting of the RPE ESTs according to their possible cellular roles was that approximately a third of the total number of assigned ESTs represented genes with involvement in gene/protein expression, with a major subcategory concerned with protein turnover. In particular, the identification of three independent ESTs of the cysteine proteinase inhibitor cystatin C constituted an interesting finding. The cystatin C clones were subjected to further analysis whose results are presented in the following section.

3.3. INVESTIGATION OF RPE CYSTATIN C CLONES

3.3.1. FULL-LENGTH SEQUENCING OF RPE CYSTATIN C CLONES

Among the RPE cDNA sequences identified as matching genes involved in proteolytic activity and its regulation (Table 3.11), three clones, K22, K26 and K210, were found to match the database sequence of the cysteine proteinase inhibitor precursor cystatin C (accession number X05607; Abrahamson *et al.*, 1987). The sequences differ in length at the 5' end, indicating that the three clones are independent, i.e. they are the result of three distinct oligo dT-primed cDNA syntheses. This implies that they may reflect a potential moderate or high level of expression of cystatin C in the RPE. (For details see below.)

According to the database annotation and the reference cited above, the mRNA for precursor cystatin C is 771 bases long and features an open reading frame of 438 nucleotides between positions 75 and 512. A putative polyA signal appears at the position 755-760. The precursor cystatin C protein has a signal or leader peptide of 26 aminoacids encoded by the sequence in the region 75-152. The sequence of 120 aminoacids of mature cystatin C is encoded by the sequence between nucleotides 153

and 512. The above features were verified in this study (see figures 3.8 and 3.9).

All three clones of cystatin C identified in this study were fully sequenced. The sequencing of both strands was carried out using both of the automatic sequencing methodologies available in the Department. The primers used were as follows: T3 primer and T7 primer for sequencing reactions analyzed on the ABI 373; M13 reverse primer and M13 universal primer, both fluorescently labelled,for sequencing reactions performed according to Li-Cor methodology (see Methods sections 2.13.2. a and b). For all three clones, the sequence of the coding strand was obtained in sequencing reactions using T3 primer and M13 reverse primer, indicating the expected directional cloning. Sequencing was carried out at least twice for both strands of each clone.

Each clone has an insert of different length, thus showing that they are not merely the result of amplification of one initial clone. The sizes of inserts are as follows: 771 nucleotides for clone K22; 778 nucleotides for clone K26; and 421 nucleotides for clone K210. Figure 3.8 shows the schematic representation of the alignment of all cystatin C clones identified in this study with the databases sequence of mRNA coding for precursor cystatin C. All three clones extend with 3 nucleotides over the reported 3' end of the database sequence. Two clones (K22 and K26) contain the full-length precursor cystatin C coding region and differ only at their 5' end (K26 being 7 nucleotides longer). Compared with the reported cDNA for precursor cystatin C, clones K22 and K26 each likewise contain one mismatch (* in Figure 3.8); however, this is a one-base insertion in the 5' untranslated region (a G at position 63 of the database sequence) and therefore not affecting the encoded aminoacid sequence. The full sequence of the clone K26 (the longest of the 3 clones) and its conceptual translation product, precursor cystatin C, are presented in Figure 3.9, with the sequences coding for regions of functional importance for cystatin C inhibitory activity highlighted (see Discussion chapter). Clone K210 has an insert with about 55% of cystatin C mRNA. Clone K210 contains a variant in the 3' untranslated region (T instead of C at position 735) which is not

shared by either clones K22 and K26 nor by reported database cystatin C sequence.

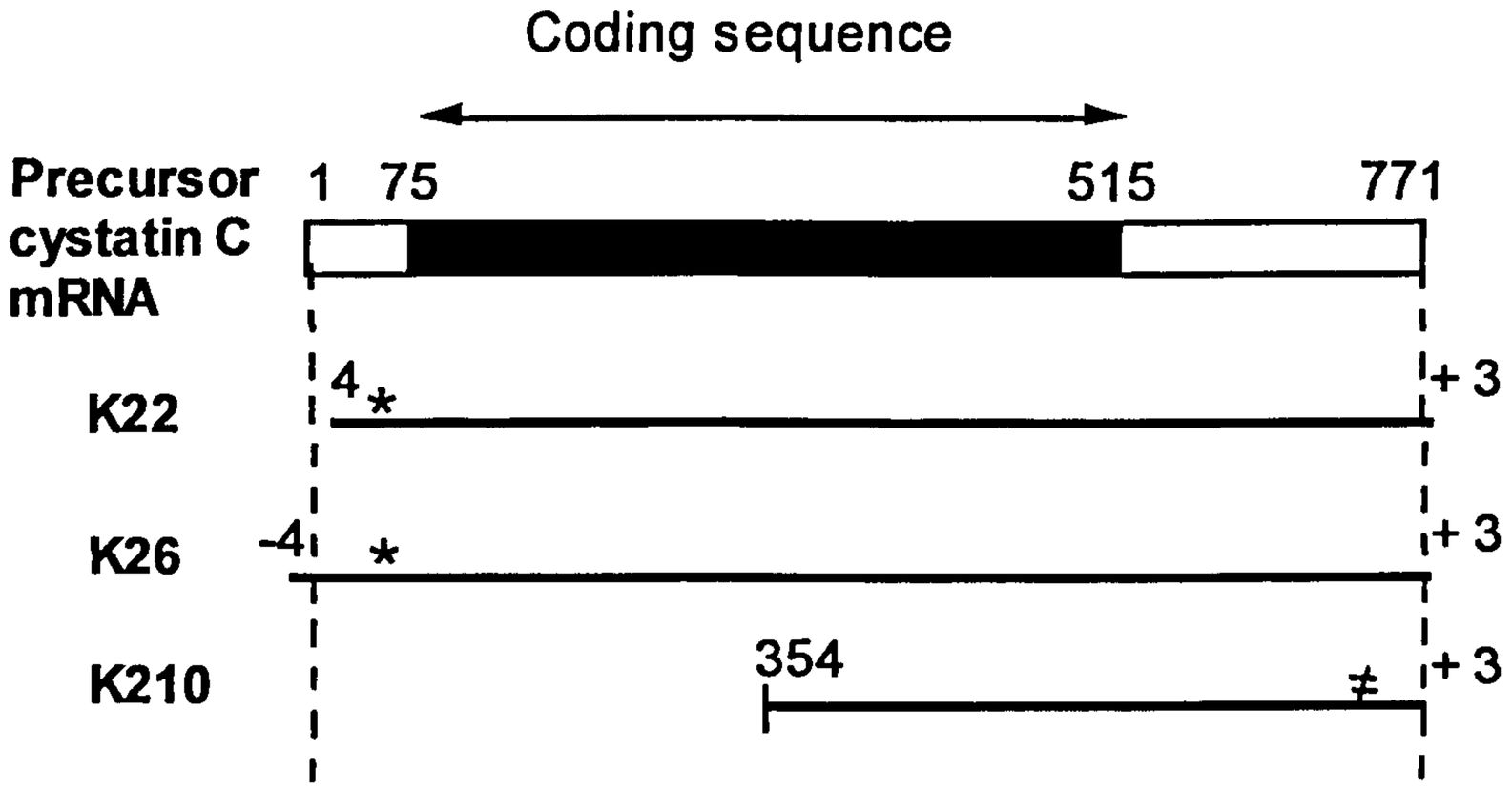


Figure 3.8. Schematic representation of the three clones coding for precursor cystatin C in RPE; *-indicates the mismatch; ≠ -substitution Numbering of the mRNA nucleotide sequence (accession number X05607) begins at first nucleotide and proceeds in the 5' to 3' direction.

3.3.2. ANALYSIS OF FREQUENCY OF CYSTATIN C CLONES IN THE HUMAN FOETAL RPE cDNA LIBRARY

The identification of three independent clones coding for cystatin C in the pool of 178 informative sequences analyzed from clones randomly chosen from the human foetal RPE cDNA library suggested that cystatin C clones may represent a significant proportion of the clones in the library investigated. In order to analyze the abundance of cystatin C clones in the cDNA library as a whole, blots representative for the entire library were hybridized with a radioactivelly-labelled probe specific for these clones. The design and preparation of cystatin C probes are described in section 3.3.2.1, while the conditions of the hybridization experiments are summarized in section 3.3.2.2.

-4 **GTAT**CGCAGCGGGTCCTCTCTATCTAGCTCCAGCCTCTCGCCTGCGCCCCACTCCCCGCGTCCCGC**G**TCCTAGCCGACC 74

75 ATG GCC GGG CCC CTG CGC GCC CCG CTG CTC CTG CTG GCC ATC CTG GCC GTG GCC CTG GCC 134

Met Ala Gly Pro Leu Arg Ala Pro Leu Leu Leu Leu Ala Ile Leu Ala Val Ala Leu Ala

135 GTG AGC CCC GCG GCC GGC TCC AGT CCC GGC AAG CCG CCG CGC CTG GTG GGA GGC CCC ATG 194

Val Ser Pro Ala Ala Gly Ser Ser Pro Gly Lys Pro Pro Arg Leu Val Gly Gly Pro Met

195 GAC GCC AGC GTG GAG GAG GAG GGT GTG CGG CGT GCA CTG GAC TTT GCC GTC GGC GAG TAC 254

Asp Ala Ser Val Glu Glu Glu Gly Val Arg Arg Ala Leu Asp Phe Ala Val Gly Glu Tyr

255 AAC AAA GCC AGC AAC GAC ATG TAC CAC AGC CGC GCG CTG CAG GTG GTG CGC GCC CGC AAG 314

Asn Lys Ala Ser Asn Asp Met Tyr His Ser Arg Ala Leu Gln Val Val Arg Ala Arg Lys

315 CAG ATC GTA GCT GGG GTG AAC TAC TTC TTG GAC GTG GAG CTG GGC CGA ACC ACG TGT ACC 374

Gln Ile Val Ala Gly Val Asn Tyr Phe Leu Asp Val Glu Leu Gly Arg Thr Thr Cys Thr

375 AAG ACC CAG CCC AAC TTG GAC AAC TGC CCC TTC CAT GAC CAG CCA CAT CTG AAA AGG AAA 434

Lys Thr Gln Pro Asn Leu Asp Asn Cys Pro Phe His Asp Gln Pro His Leu Lys Arg Lys

435 GCA TTC TGC TCT TTC CAG ATC TAC GCT GTG CCT TGG CAG GGC ACA ATG ACC TTG TCG AAA 494

Ala Phe Cys Ser Phe Gln Ile Tyr Ala Val Pro Trp Gln Gly Thr Met Thr Leu Ser Lys

495 **TCC ACC TGT CAG GAC GCC TAG** GGGTCTGTACCGGGCTGGCCTGTGCCTATCACCTCTTAT 554

Ser Thr Cys Gln Asp Ala End

555 GCACACCTCCCACCCCCTGTATTCCCACCCCTGGACTGGTGGCCCCTGCCTTGGGGAAGG 614

615 TCTCCCCATGTGCCTGCACCAGGAGACAGACAGAGAAGGCAGCAGGCGGCCTTTGTTGCT 674

675 CAGCAAGGGGCTCTGCCCTCCCTCCTTCCTTCTTGCTTCTCATAGCCCCGGTGTGCGGTG 734

735 CATACACCCCCACCTCCTGCAATAAAATAGTAGCATC**GGC 774**

Figure 3.9. Nucleotide sequence of the full-length cDNA for clone K26 coding for precursor cystatin C.
Numbering of the nucleotide sequence of the sense strand is in connection with the database mRNA sequence of precursor cystatin C (nucleotides 1 to 771; accession number X05607). The one-base (**G**) insertion detected in position 63 and the extended sequences revealed at 5' and 3' ends of the RPE clones are shown in bold.
The initiation consensus of the coding sequence and the polyadenylation signal are underlined. Nucleotides coding for Cys residues and fragments of sequence coding for other amino acids residues important for cystatin inhibitory function are highlighted.
In patients with hereditary cerebral haemorrhage with amyloidosis the boxed residue, Leu in cystatin, is replaced by Gln.

3.3.2.1. Cystatin C probe preparation

Three pairs of primers were used to amplify by PCR fragments of the cystatin C-encoding insert of clone K26. The primers were designed using the programs Prime (GCG package) and Primer (Whitehead Institute) and their characteristics are summarized in Table 3.15. Noteworthy is the fact that, due to the GC-rich sequence of cystatin C, a relatively small number of primer pairs could be designed for amplifying fragments longer than 250bp.

Table 3.15. Characteristics of PCR primer pairs specific for cystatin C clones.

	Pair P1/P2	Pair P3/P4	Pair P5/P6
forward primer	TCGGCGAGTACAACAAAG (244-261)	CAACGACATGTACCACAGCC (266-285)	GCGGGTCCTCTCTATCTAGC (5-24)
reverse primer	AAGAAGGAAGGAGGGAGGG (708-690)	AGCAAGAAGGAAGGAGGGAG (711-692)	TGGATTTCGACAAGGTCATT (498-479)
optimal annealing temperature	60.2°C	60.3°C	58.1°C
PCR product length	465bp	446bp	494bp

The sequence of primers is given in the 5'-3' orientation, followed, in brackets, by the coordinates of alignment with cystatin C sequence. Primers were custom-made by Perkin-Elmer Ltd. with the concentration of stock solutions in the range 4-16 pmol/μl.

The amplification of cystatin C-specific fragments was performed in 50μl PCR reactions, with 2.5ng of plasmid K26 and 12 pmoles of each primer per reaction (2mM Mg^{2+} and concentration of other components were as described in Methods' section 2.10.2). The temperature cycle program used consisted of 35 cycles of (94°C, 30"; 60°C, 1'; 72°C, 2') with a final extension

at 72°C for 7'. The amplification with the primer pair P5/P6 required a different concentration of Mg^{2+} (1mM) and an annealing temperature of 57°C. The results of the PCR amplifications are shown in Figure 3.10. Unique bands of the expected size were obtained in all three cases (though P5/P6 showed a smear of lower molecular weight material; this was not investigated further).

Following TBE/agarose electrophoresis, the bands corresponding to the PCR products resulted in the reactions P1/P2 and P3/P4 (reactions identified by the respective pair of primers used) were excized from gel. The DNA fragments were extracted with Kristal™ Gelex Resin (Cambridge Molecular Technologies Ltd.) in the presense of Na perchloride and resuspended in water at an approximate concentration of 25ng/μl.

DNA fragments resulting from the PCR amplification were subjected to restriction analysis in order to verify their identity. Restriction digestions were carried out, in separate reactions, with *Alu I* and *Bgl II* enzymes, whose restriction sites were expected to be present within the DNA fragments amplified. The digestions gave rise to fragments of the expected size (see Figure 3.11. a. and b.), thus providing reasonable proof that the DNA frgments amplified were the expected ones. In particular, the *Bgl II* restriction was considered informative since *Bgl II* is a 6bp cutter and, therefore, the probability of its site to appear by chance is 1 in 4^6 bp, i.e. 1 site in a DNA fragment of about 4kb.

Approximately 25ng of purified PCR product obtained as described above was labelled with (α-^{32}P)dCTP by random priming, according to the procedure described in the Methods section 2.7.3.c. The probe was separated from unincorporated nucleotides and used for hybridization of the whole cDNA library filters.

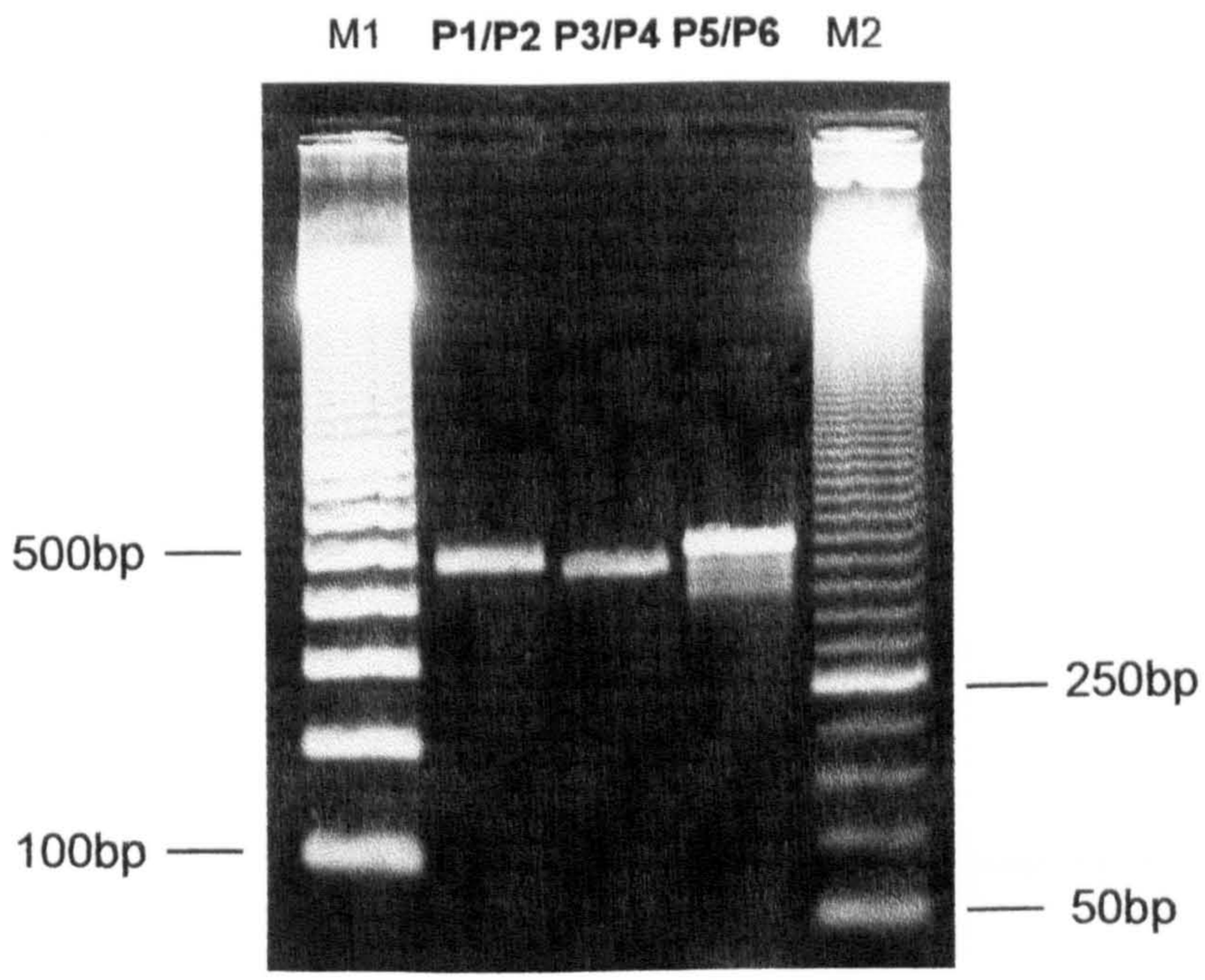


Figure 3.10. PCR amplification of cystatin C fragments.
Amplification reactions are identified by the pairs of primers used (conditions described in text). 15µl of each PCR reaction were subjected to electrophoresis in 1.5% agarose/ethidium bromide gel. Markers: M1 - 100bp ladder; M2 - 50bp ladder.

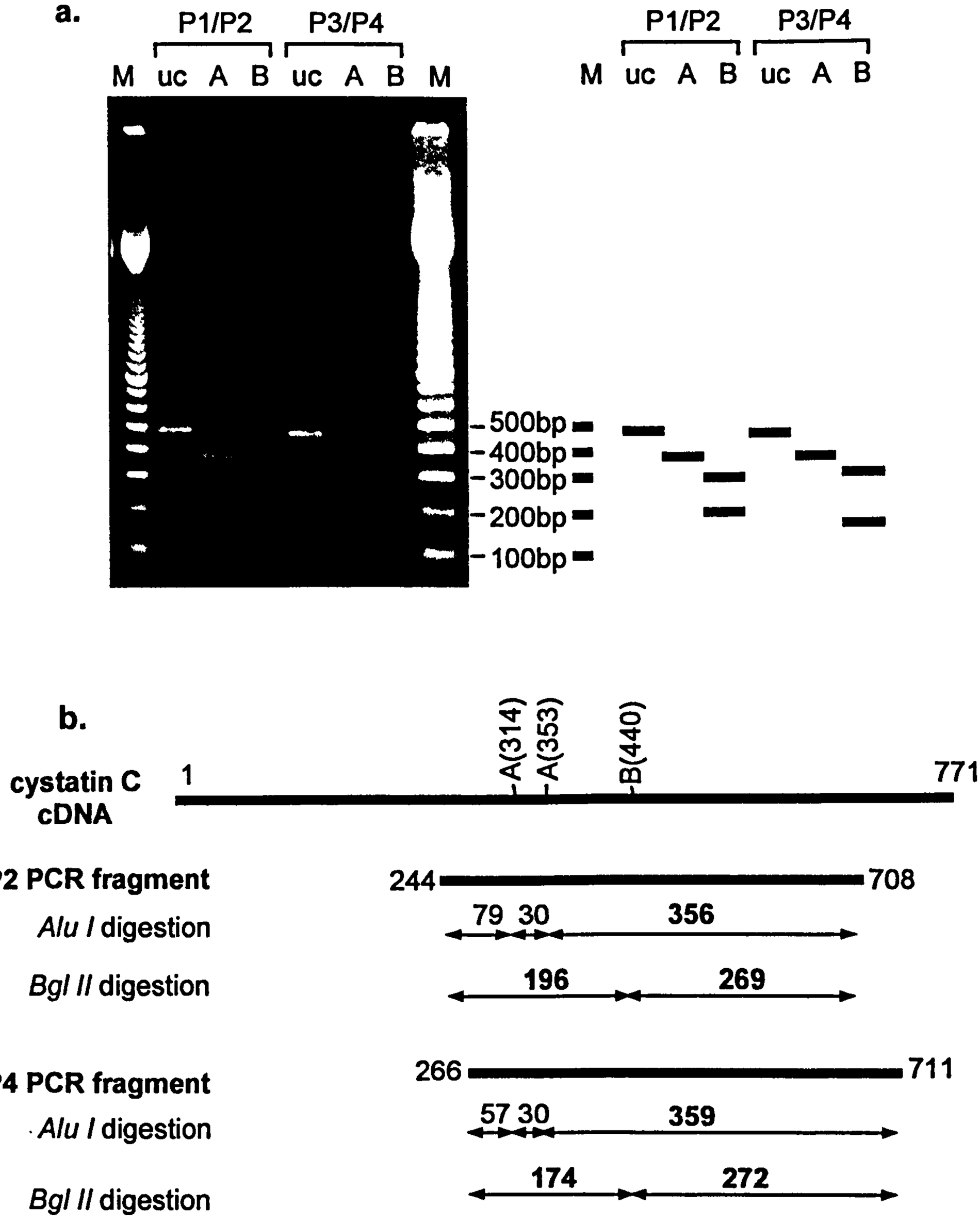


Figure 3.11. Restriction analysis of cystatin C fragments obtained by PCR amplification using P1/P2 and P3/P4 primer pairs.
a. Agarose gel electrophoresis of restriction digestion products.
uc - uncut; A - digestion with *Alu I*; B - digestion with *Bgl II*; M - 100bp ladder.
b. Diagram showing the fragments expected to arise from the restriction digestions.
Fragments seen on the gel are shown in bold typeface.

3.3.2.2. Hybridization of blots of the whole cDNA library with cystatin C probe

Diluted aliquots of the human foetal cDNA library were plated out as described in section 2.8.7. Serial dilutions were aimed at obtaining approximately 1500 plaques per plate. To ensure statistical representation of the library, an individual series of dilutions, starting with 1µl of the library phage stock, was made for each plate.

Plates were incubated at 37°C for 12-14 hours and then blotted onto nitrocellulose filters according to the protocol described in section 2.11.1. As a control for the consistency of hybridization, one plate was blotted onto triplicate filters. Control plates without plaques (bacteria that were not infected with phage) were also blotted and the resulting filters were used as negative controls for hybridization. After fixation and pre-hybridization (6 blots per hybridization bottle), blots were hybridized using 20ml hybridization solution (Table 2.7) and the whole probe resulted in a radioactive labelling reaction described above. A negative control filter was introduced into each batch of hybridization. The hybridization was carried out overnight at 42°C and the subsequent washes were performed until no radioactive signal could be detected on the control filters with the Geiger counter. The most stringent wash thus made was at 65°C in 0.1xSSC, 0.1% (w/v) SDS. Washed filters were exposed to X-ray film on which positive signals were identified after 4 hours exposure (overnight exposure produced very intense signals). Control filters presented no spots. The plaque corresponding to one of the positive spots was subjected to partial sequencing which confirmed the identity of the insert as cystatin C cDNA. The results of these blotting and hybridization experiments are summarized in Table 3.16 and represented in Figure 3.12.

Hybridization of filters of the human RPE cDNA library with a probe created from clone K26 indicated that the proportion of clones coding for cystatin C in this library is quite high (about 1.35%). This result is consistent with the identification of 3 independent clones of cystatin C among the 178 informative clones isolated at random from the analyzed cDNA library.

Table 3.16. Summary of hybridization of filters of whole human foetal RPE cDNA library with radioactive cystatin C probe.

Plate	Pfu/plate	Positive pfu	% positives
c1	1600	22	1.40
c2	1800	23	1.20
e1	1100	14	1.25
e2	1500	19	1.25
e3	1400	19	1.35
e4	1500	20	1.30
x1	2000	29	1.45
x2	1600	22	1.40
x3	1800	26	1.40
TOTAL	14300	194	1.35

The approximate number of plaque forming units (pfu) are indicated for each plate. Positive plaques were counted after 4 hours exposure of hybridized and washed filters to X-ray film. The % positive plaques are rounded up to the nearest decimal place. Plate x3 was blotted onto triplicate filters, which presented the same pattern of positive signals. Filters c1 and c2 are presented in Figure 3.12. No hybridization signals were detected on the control filters.

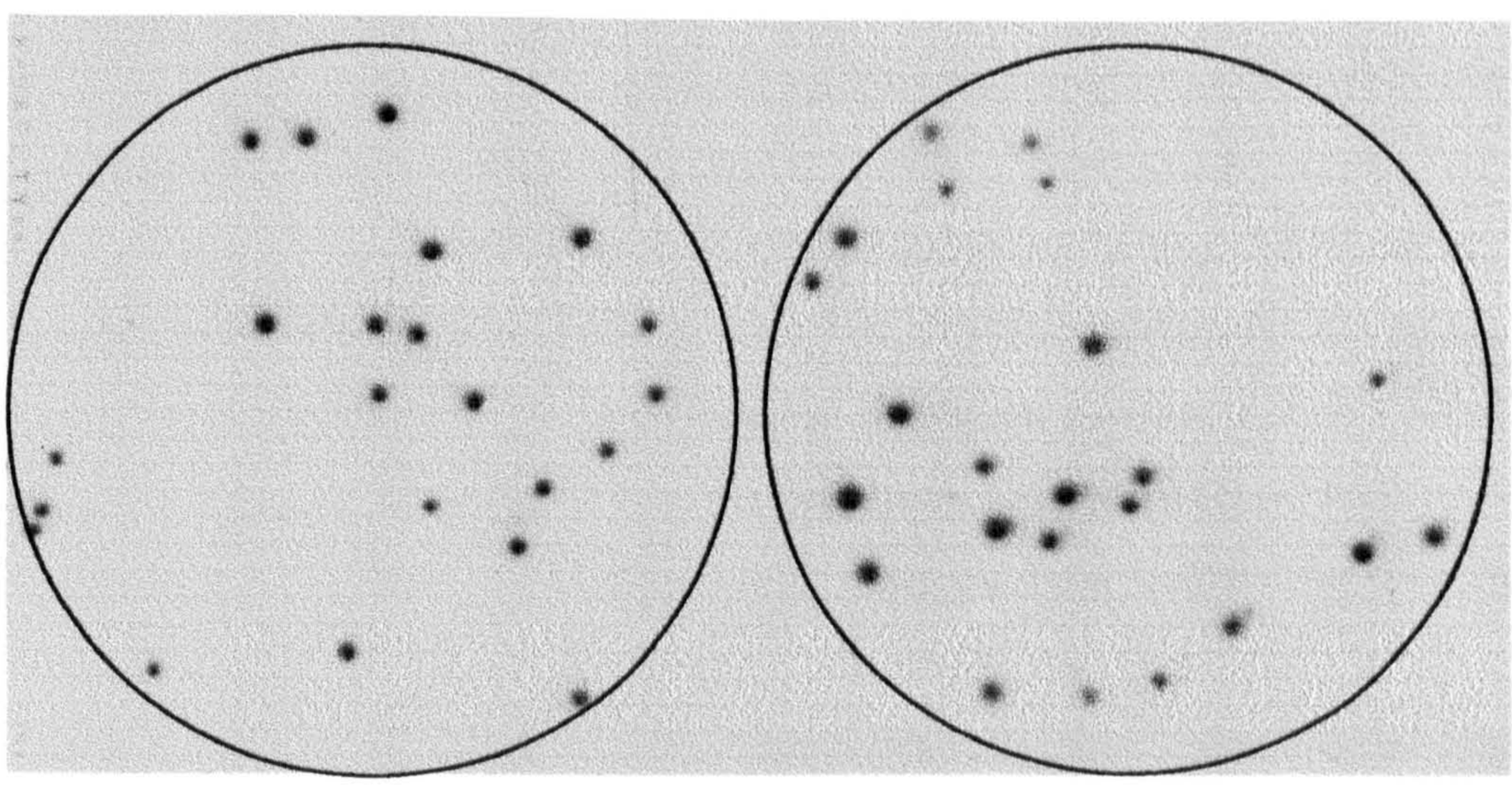


Figure 3.12. Blots of human foetal cDNA library hybridized with radioactive cystatin C probe.
Filters showed are c1 and c2 from the Table 3.16. Out of 1600, respectively 1800, plaque forming units, 22 and respectively 23, gave positive hybridization signals.
Hybridization controls (blots of plates with no plaques) presented no signals.
See text for description of hybridization and washing conditions.
4 hours exposure.

3.3.3. INVESTIGATION OF CYSTATIN C EXPRESSION BY NORTHERN TRANSFER HYBRIDIZATION

Northern blot analysis with a probe created from clone K26 was performed on mRNA isolated from cultured human RPE cells and on a collection of mRNAs from various tissues assembled commercially into a "human master blot". The aim of this analysis was to confirm the existence of mRNA specific for cystatin C in cultured RPE cells and to investigate the pattern of its expression in other tissues.

Approximately 1µg mRNA isolated from cultured human RPE cells was subjected to electrophoresis in a 1.5% MOPS/formaldehyde agarose gel (protocol and conditions as described in Methods section 2.7.3.a) alongside with 10µg total yeast RNA. The gel was blotted onto a Hybond-N nylon membrane, on which transferred RNA samples were fixed by baking at 80°C for 2 hours.

The Human RNA Master Blot™ (Clontech Lab., Inc.), used to assess cystatin C pattern of expression, was represented by a nylon membrane to which poly A+ RNAs from 50 human tissues were immobilized in separate dots, along with several controls. The mRNA samples on the master blot were normalized to the mRNA expression levels of eight different "housekeeping" genes. Hence, the hybridization signal obtained with a given probe should ensure a fairly accurate indication of the relative abundance of the specific mRNA in different tissues.

Both blots were hybridized to a K26-derived probe, prepared and labelled as described in the section 3.3.2.1 above. The hybridization was conducted at 42°C overnight and the subsequent washes were increasingly stringent, the final one being made with 0.1xSSC, 0.1% SDS at 65°C for the RPE blot and 50°C for the master blot (according to manufacturer's recommendations). Blots were exposed to X-ray film and a phosphoimager screen. The blot containing the RPE mRNA was stripped off and subsequently hybridized to a human β-actin probe as a control to ensure that the RNAs were of suitable quality. The hybridization to the actin probe was carried out overnight at 42°C and the washes were made at

45°C with 2xSSC, 1% SDS (conditions less stringent, due to species difference).

Hybridization of RPE mRNA with cystatin C probe gave rise to two close bands of the expected size (about 800bp) (Figure 3.13. A), probably due to the existence of two related transcripts. A very faint signal appeared on the lane of total yeast RNA (a band of a sligthly smaller mRNA). Both RPE and yeast RNA hybridized to β-actin probe (bands of relatively equal intensities, Figure 3.13.B). The two bands appear practically at the same position along the blot because the human β-actin probe, specific for the human mRNA species of 1138nt (GB_PR2:AB004047), hybridizes under the conditions of the experiment to yeast actin mRNA which is 1126 bases long (GB_PL:YSCACT).

The hybridization of human master blot to cystatin C probe revealed a wide expression of this transcript (Figure 3.14). The highest level of expression among the tissues examined was apparent in the salivary gland. All brain tissues represented in the blot, as well as the spinal cord also showed a considerable level of expression. Hybridization signals of similar intensity were detected for adrenal gland, lung and placenta. Among the foetal tissues sampled in the blot, the highest level of expression (moderate compared with adult tissues mentioned above) was observed in spleen. Obviously, due to the manner in which the master blot was constructed (dots of the whole mRNA population), all hybridization signals represent the pooled signal of all related transcripts from a particular tissue that hybridize to the probe. None of the negative controls (line H of the blot) hybridized to the cystatin C-specific probe.

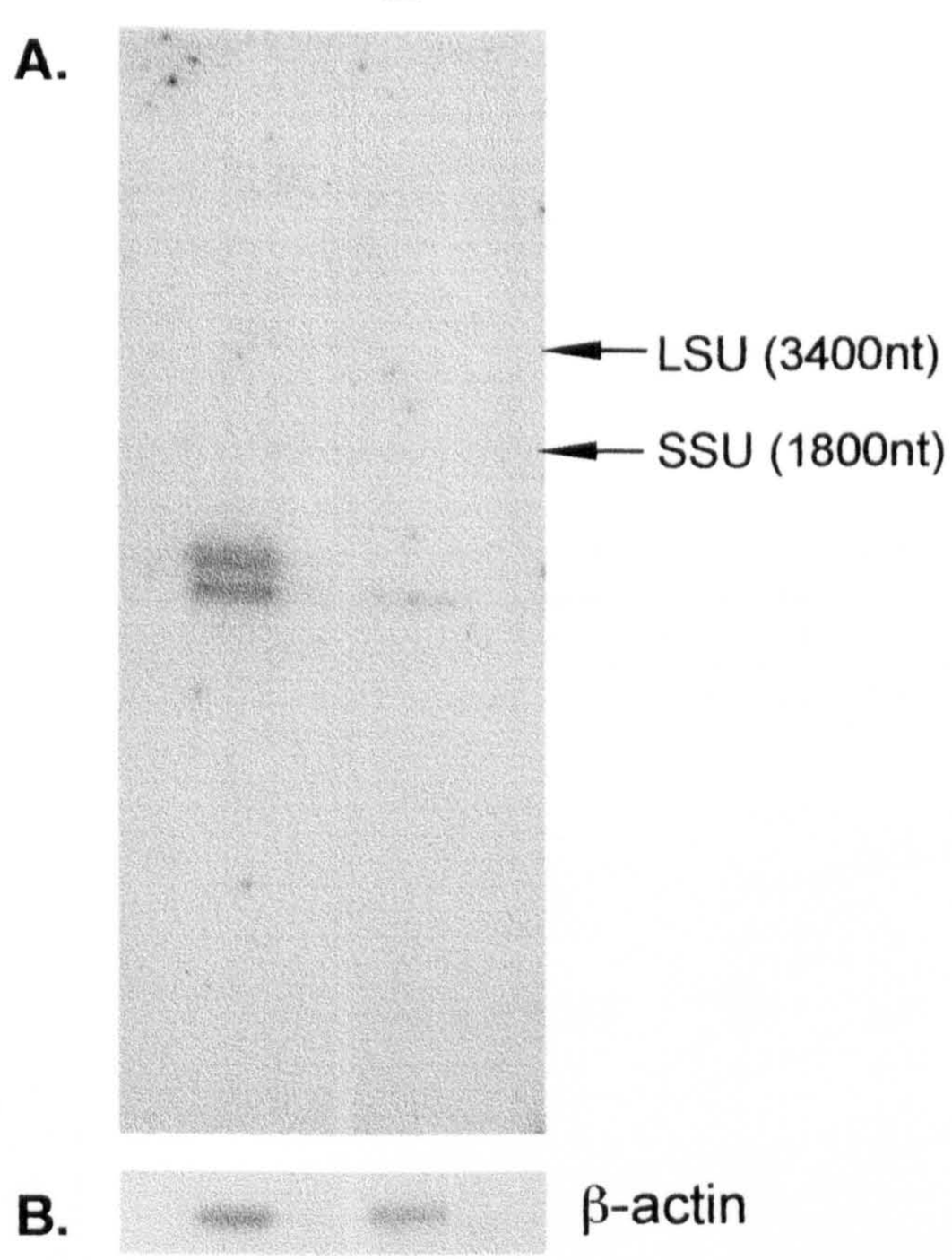


Figure 3.13. Northern blot hybridization of mRNA from cultured human RPE cells to cystatin C-specific probe.
A. Blot hybridized to ^{32}P-labelled cystatin C probe, derived from clone K26. 2 weeks exposure at -70°C with intensifying screen. Lane 1 - RPE mRNA (1µg); lane 2 - yeast total RNA (10µg). Position of yeast rRNAs was visualized by toluidine blue staining of total yeast RNA run on a separate lane on the same gel, cut prior to transfer.
B. Same blot hybridized to ^{32}P-labelled human β-actin probe. Overnight exposure to phosphoimager screen.

	1	2	3	4	5	6	7	8
A	whole brain	amyg-dala	caudate nucleus	cere-bellum	cerebral cortex	frontal lobe	hippo-campus	medulla oblongata
B	occipital lobe	putamen	subst. nigra	temporal lobe	thalamus	sub-thalamic nucleus	spinal cord	
C	heart	aorta	skeletal muscle	colon	bladder	uterus	prostate	stomach
D	testis	ovary	pancreas	pituitary gland	adrenal gland	thyroid gland	salivary gland	mam-mary gland
E	kidney	liver	small intestine	spleen	thymus	peri-pheral leukocyte	lymph node	bone marrow
F	appen-dix	lung	trachea	placenta				
G	foetal brain	foetal heart	foetal kidney	foetal liver	foetal spleen	foetal thymus	foetal lung	
H	yeast totalRNA 100ng	yeast tRNA 100ng	*E.coli* rRNA 100ng	*E.coli* DNA 100ng	Poly r(A) 100ng	human Cot1DNA 100ng	human DNA 100ng	human DNA 500ng

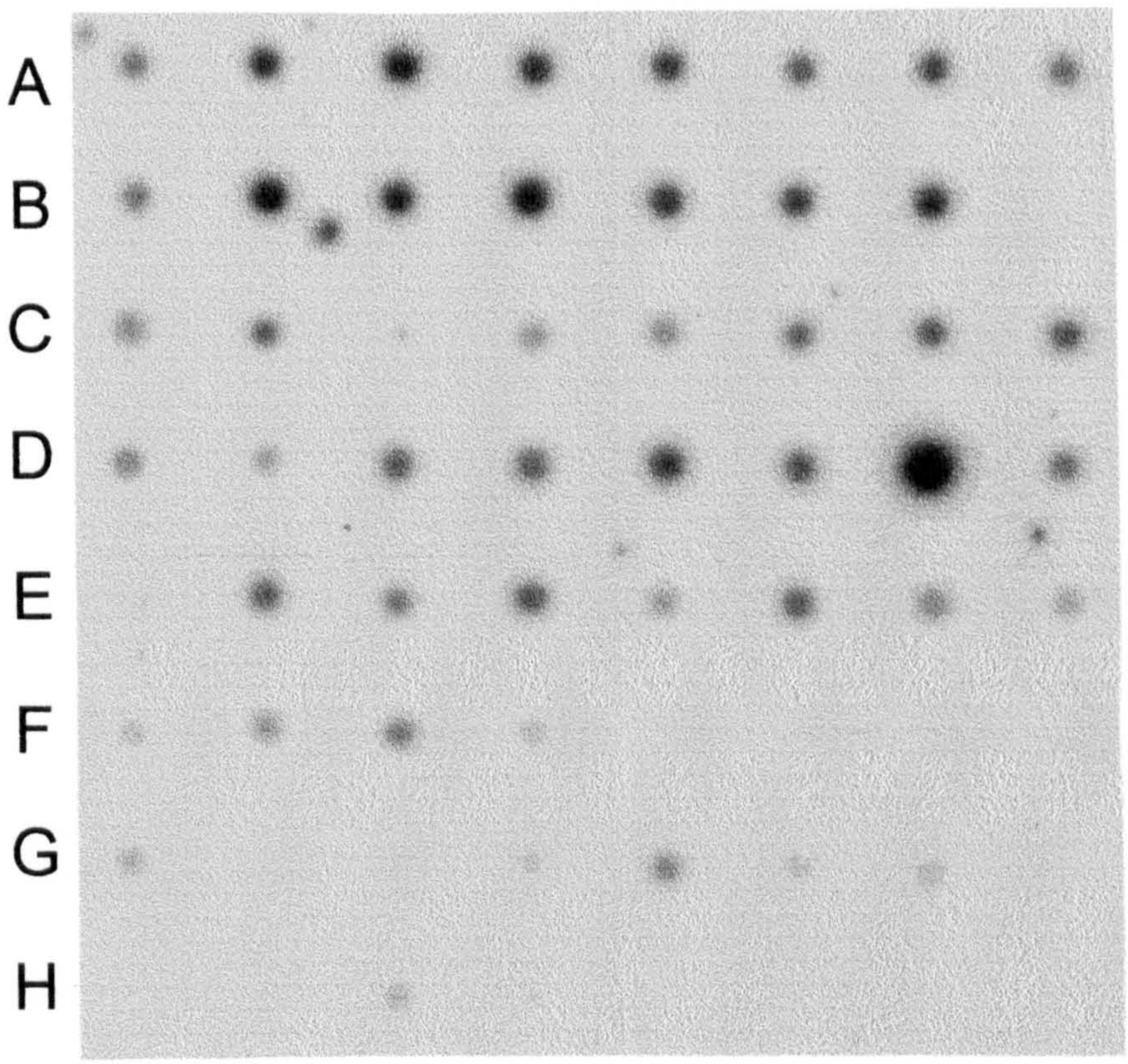


Figure 3.14. Northern blot analysis of cystatin C expression in various tissues.
The tissues represented in the Human RNA Master Blot^{TM} (Clontech) are indicated in the upper panel.
Lower panel - Blot hybridized to K26-derived cystatin C probe (^{32}P-labelled). 2 weeks exposure at -70°C with intensifying screen.

In summary, two experimental lines (ESTs analysis in the RPE and Northern blot analysis) provide evidence that cystatin C is expressed in the RPE of the eye. Moreover, the finding of cystatin C clones in a relatively high frequency in the cDNA library under analysis, as revealed first by the discovery of three independent cystatin C cDNA clones and then confirmed by the analysis of the frequency of this cDNA in the whole library, supports the hypothesis that cystatin C might have an important role in the RPE. The implications of this outcome and the possibilities for future investigations of cystatin C role(s) in the RPE are discussed in the final chapter of this thesis.

CHAPTER 4

DISCUSSION AND CONCLUSIONS

The study described in this thesis has achieved the objective set out at its initiation by revealing details of the transcriptional picture of RPE cells by random sampling of cDNA clones. The ESTs analysis led to identification of gene expression products which may be of specific importance for RPE cellular functions. However, this objective was not achieved by investigation of cDNA libraries of *freshly dissected* RPE cells, as was intended at the beginning of the project, but by the study of a cDNA library of *cultured* RPE cells. The decision to use the latter was determined by the unsatisfactory quality of cDNA libraries of freshly dissected bovine RPE cells constructed during the first part of the project.

Consequently, the discussions forming this chapter are structured in 3 main sections. The first section presents the conclusions from constructing of cDNA libraries of RPE cells and the possible ways of improving the outcome of such attempts. The second section discusses the characteristics of EST analyses carried out in this study, including implications of cystatin C clones in the RPE expression profile. Finally, possible future directions are considered.

4.1. CONSTRUCTION OF cDNA LIBRARIES OF RPE CELLS

Since relatively recently, synthesis of cDNA and construction of cDNA libraries has become common practice in molecular biology laboratories. Such libraries are used for a variety of projects and with a multitude of aims. A wealth of improving technologies support advances in this practice. Nevertheless, there are still a number of critical steps which are difficult to monitor and which affect the outcome.

Construction of a cDNA library involves between 13 and 20 steps, from isolation of cells to ligation of cDNA molecules with suitably modified ends to vector arms, including the purification and ethanol precipitation steps required at various stages during the procedure. Most of these steps have the potential of incurring loss of material. The highest such losses are likely to occur in the phenol purification/ethanol precipitation and in the column fractionation steps. Moreover, quantitation of material of interest is required in at least 3 stages (mRNA for first strand cDNA synthesis, ds cDNA for adaptor ligation, and "adapted" ds cDNA for ligation with vector arms) and is desirable or informative for at least some other stages of the procedure. Yet, the methods used for determination of yield and product size involve removal of material which can not be subsequently recovered, thus adding to the losses incurred.

Analysis of cDNA library construction performed in the present study points to three critical stages, which are individually discussed below:

(1). Isolation of RPE cells and mRNA preparation

(2). Reverse transcription of mRNA

(3). Size fractionation and precipitation of cDNA

Isolation of RPE cells and mRNA preparation

In the light of the facts presented above it appears obvious that an adequate quantity of starting material (i.e. cells and/or mRNA) is of major importance for a successful endeavour of constructing a cDNA library. As far as I am aware, at present the only cDNA library of freshly dissected bovine RPE cells of good quality (high percentage of recombinants with full-length DNA or long inserts) is that constructed in Dr. Tony Redmond's laboratory (National Eye Institute, NIH, USA). For the construction of this cDNA library, isolation of freshly dissected RPE cells, used for mRNA preparation, was performed on about 100 bovine eyes, all available during a short period of time (T.M. Redmond, personal communication). The resulting cDNA library was the source for the isolation of the cDNA for *RPE65*, a novel retinal pigment epithelium-specific microsomal protein (Hamel *et al.*, 1993) which subsequently generated much interest in the RPE/eye research community. Recently, the *RPE65* gene, encoding an

evolutionarily conserved 61-kDa protein, was the first RPE-specifically expressed gene to be implicated as a disease gene (mutations in *RPE65* associated with autosomal recessive childhood-onset severe retinal dystrophy were identified by (Gu *et al.*, 1997).

During the present project, a total of 50 bovine eyes were dissected in order to isolate RPE cells for mRNA preparation. In order to compensate for the limited availability of freshly isolated RPE cells and to fulfil the need for such starting material, special attention was given to the optimization of the isolation procedure for improving the yields of intact RPE cells obtained. Noteworthy is the fact that, for starting primary cultures, only a few viable cells are needed; hence, even when a relatively small proportion of RPE cells are isolated from freshly dissected eyes, there are sufficient cells that prove viable in culture, adhere and survive. Nevertheless, for obtaining more intact cells from each dissected eye, conditions of isolation should be optimized. Thus, it was observed that by increasing the incubation time of dissected bovine eye cups with trypsin solution by about 10 minutes (from 45 minutes, usually needed for harvesting material to set up primary cell cultures, to about 55 minutes) the yields of RPE cells detached doubled (see Table 3.1). Such altering of trypsinization conditions should, however, be very carefully monitored in order to avoid lysis of cells. The maximum number of RPE cells that could be isolated from one bovine eye, before signs of cell lysis occured, was about $1x10^6$ cells, which still represents less than a quarter of the total RPE cells in a bovine eye. Unfortunately, no exact conditions can be defined, as the optimum is very likely to be influenced by factors specific to each batch of eyes, such as age and time post-mortem, as well as the relative activities of different batches of trypsin. Working with a relatively small number of eyes at a a time (6 to 9) allows monitoring of dettaching/lysis of cells, but requires several successive experiments.

Messenger RNA preparations were made from up to $1x10^7$ freshly dissected bovine RPE cells (see Table 3.2). Similar numbers of cultured RPE cells (harvested at preconfluent and confluent stage of culture) were used for the same purpose. The yields were in the range 0.10-0.18 pg mRNA per cell, comparable with those obtained from other cell types by the

same procedure (examples in the information accompanying the Qiagen mRNA isolation kits). The mRNA obtained was of good quality, i.e. undegraded, as shown by Northern blotting (Figure 3.3), and not contaminated by protein, as indicated by spectrophotometric measurements. Although the direct isolation of mRNA from lysed cells reduced the number of manipulations, it may have been more economical to prepare total RNA first, which could have been used for tests assessing integrity/purity of preparation, avoiding losses of mRNA.

Reverse transcription of mRNA

Three attempts at constructing cDNA libraries of freshly dissected bovine RPE cells were made in this study, starting from approximately 1µg mRNA. Various stages of cDNA synthesis procedure were monitored and overall revealed that the synthesis did occur (having had incorporated radioactive label; see Table 3.4) and that the synthesis products were roughly of the expected size (Figure 3.4). Yet, the fact that the percentage of template transcribed in the first strand cDNA reaction was lower than expected (yields over 20% are indicated as routinely obtained with Amersham's cDNA synthesis module) may indicate that the quantity of mRNA used as template could have been slightly overestimated and that, consequently, insufficient template was used initially.

Under these circumstances, the outcome of the cDNA synthesis could have been improved by using a smaller amount of random hexanucleotide primers. Thus, according to Pharmacia's guidelines for use of TimeSaver® cDNA Synthesis kit, decreasing the ratio primers to template allows synthesis of medium and long cDNA to proceed more efficiently. Alternatively, synthesis of cDNA oligo-dT primed using total RNA could have been considered.

Size fractionation and precipitation of cDNA

Despite good percentage of recombinants yielded by all bovine libraries (Table 3.5), the actual size of inserts was dissapointingly small. Losses of cDNA in the column purification/size fractionation step are the most likely cause for this outcome. Losses could have also occured during

the subsequent precipitation of cDNA. These processing steps are most likely to incur loss of material when less than 0.5µg starting material is used.

The poor efficiency of column purification could have also yielded an over-representation of alternative reaction products of the ligation with adaptor reaction, such as adaptor dimers/multimers. In the case of Amersham adaptors, which contained 5'-hydroxyl group at their blunt ends, a likely explanation for the existence of such a large proportion of inserts containing adaptor multimers is that adaptors were partly retained by the column; such unseparated adaptors could ligate to each other, either through their blunt or *EcoRI*-cohesive ends, following the kinasing step. Also, the possibility that the cDNA was not neatly blunt-ended for the ligation of adaptors cannot be excluded and it could explain the subsequent failure of cloning, at least in part, the RPE cDNA. A self-ligation test to assess the efficiency of blunt-ending would have required another portion of the cDNA preparation.

4.2. BEGINNINGS OF THE TRANSCRIPTIONAL PICTURE OF RPE CELLS

4.2.1. ANALYSIS OF ESTs IN THE RPE

Efforts are underway in numerous laboratories to build transcriptional pictures of genes expressed in a particular tissue or cell. Sequencing underpins all these projects and relates them to general genome sequencing projects. Yet, the projects diverge considerably in the rationale of their experimental approaches with directions of study ranging widely from identification of specific or uniquely expressed genes (having subtractive cDNA libraries as the main tools) to identification of all expressed genes and semi-quantitative characterization of their expression. Some studies focus on identification of novel sequences expressed, others on characterization of disease genes.

The RPE is not among the tissues/cells for which a major systematic effort for identification and analysis of expressed gene species has yet been undertaken. Rather, knowledge in this field so far comes from studies that have focused directly on particular genes or activities. This explained the fact that all the RPE-related information contained in the "initial assessment of human gene diversity and expression patterns" generated by Venter's group (Adams *et al.*, 1995) was compiled from independent entries in the dbEST database. The only characterization of RPE ESTs previous to the study described in this thesis was based on the analysis of 108 3'-end sequences of randomly selected clones from a subtracted RPE cDNA library (Gieser & Swaroop, 1992). Of these ESTs, 36 were identified as representing known human genes or homologs of known vertebrate genes and 61 presented no homology to database sequences.

Parallels can also be drawn between characteristics of the study described in this thesis and attributes of the study of human fovea based on ESTs analysis (Bernstein *et al.*, 1996). The fovea is a region with the highest density of cone photoreceptors and of ganglion cells in the retina; it represents less than 5% of the total area of the retina. Difficulties in studying, especially at molecular level, this highly specialized region of the retina responsible for high visual acuity owe much to the difficulties in obtaining large amounts of fresh tissue for study. Bernstein et al. succeded in constructing an age-normalized fovea cDNA library in which enrichment for fovea-specific genes was proved by the estimated opsin to red/green cone pigment gene ratio (3.7 to 1, versus the corresonding ratio in the whole retina of 15 to 1 (Nathans *et al.*, 1986)). Single-pass sequencing was performed on 209 randomly isolated clones from this library. The analysis of the resulting ESTs revealed 121 clones (57.9%) matching sequences in GenBank; these included repetitive sequence motifs, rRNA sequences and mitochondrial expressed genes. The rest 42.1% of the clones (88 ESTs) presented no identical match to GenBank sequences and were therefore considered new; these included ESTs similar to database entries (minimum 55% sequence match over at least 70 bp).

4.2.1.1. Generation of RPE cDNA partial sequences

Considering the available resources, the target of this study was to identify 200 sequences expressed in the RPE. Altogether, 396 clones were randomly isolated from all RPE cDNA libraries available (108 bovine and 288 human) and 300 sequencing reactions performed, representing 272 individual clones. The success rate of sequencing reactions performed was 67% (203 good sequences obtained; see Table 3.6), comparing well with that obtained in major ESTs studies (52.4% to 71.5%; (Adams *et al.*, 1995)).

For defining the transcriptional picture of a tissue there are two main practical aspects to be considered: number of clones analyzed and degree of analysis. The distribution of efforts among these two aspects shapes decisively the outcome of the study. Within the analysis process itself, two main components can vary with regard to the time resources allocated: processing of the sequencing output files (the conclusions of which are presented below) and assessment of sequence comparison results (which is discussed in the following section).

In this study, for approximately the first half of the sequence files obtained, extensive processing was performed. The processing involved mostly manual calling of ambiguous bases by inspection of peaks of each sequence's electropherogram. Peaks that were not automatically called were usually preceded by a large peak or masked by dye blobs. Patterns of the following types were encountered:

- large "G" peak followed by a very small "T" or "C" peak;
- small "G" peak, frequently masked after a big "A";
- in a string of "G"s one of them, usually the third, showed a reduced signal;
- numerous weak "G"s in G-C rich regions;
- strings of "T"s followed by a very large peak of a different base obscuring adjacent peaks.

The accuracy of editing was tested by comparing the newly-called bases (always left in small letters in the sequence) with the corresponding bases in the best match sequence. Although the inspection and editing of

sequences obtained by automatic sequencing are important for improving the overall accuracy of base calling and quality of sequences, it was eventually realized that this extensive processing is not absolutely necessary for the immediate use of sequences obtained, i.e. comparison with database sequences. The software used routinely for these comparisons is sufficiently robust to tolerate some mismatches or ambiguities. Generally, good matches could still be easily identified and, therefore, especially when the interest lies in identification of sequences matching existent databases entries, a minimal intervention for the clarification of details of output sequence files is preferable.

4.2.1.2. Overview of the analysis of RPE ESTs

The partial sequencing strategy of randomly selected cDNA clones is straightforward in producing "raw" data of sequences expressed in a tissue or cell. Sequences of mitochondrial and ribosomal RNA are found in any primary collection of data thus generated. Some genomic contamination is also possible to be unveiled and potential contamination with prokaryotic or yeast sequences should be investigated. The real challenge, however, is to extract useful information from the analysis of what is left in the collection of sequences after elimination of identified contaminating sequences.

The RPE ESTs were categorized in 3 main groups, based on the evaluation of the results of comparison with sequences in the public databases: (1) ESTs identical or similar to genes coding for previously known or characterized proteins; (2) ESTs identical or similar to "anonymous" sequences; (3) ESTs matching no sequences in the databases. "Anonymous" sequences are sequences deposited in the public databases, generated by genomic or cDNA sequencing studies, whose gene expression products were not investigated.

The assessment of sequence comparison results is the most demanding part of the analysis of ESTs. After acquiring and analyzing data from the first approximately 100 ESTs, it seemed that the proportion of clones for which a clear identification is difficult to make is very high: about

half of the sequences were unaccountable with regard to the function encoded. A wide spectrum of situations was covered, including:

- various degrees of similarity to known sequences;
- similarity to anonymous or unassigned database sequences;
- no similarity with database entries.

The initial inclination was to emphasize the analysis of sequences with no clear identification. However, the assessments of these sequences were not of major importance for the final outcome of the study, most of the sequences so analyzed could still not be assigned and, therefore, not contributing to the identified details of the transcriptional picture.

Nevertheless, a conclusion important for the general strategy of an EST project emerged from the initial phase of the project described above. Since the effort to clarify the nature of sequences that pose problems in the first attempt at interpretation and/or assignment has very little chance of achieving its aim for a significant number of clones further analyzed, it is more beneficial to increase the number of clones analyzed. It is hard to escape from the conclusion that the unaccountable sequences are the "price that needs to be paid" on the way to building up the overall transcriptional picture.

Consequently, the rationale behind building a collection of sequences representing genes that are expressed in the RPE became the *identification* of gene expression products which are most likely to support the functions and activities of these cells (category 1 above). ESTs in both "anonymous" and unique sequences categories (categories 2, respectively), as well as those putatively identified as related or similar to known genes may represent genes coding for novel gene expression products. Yet, in order to draw any relevant conclusion regarding this matter, each one of the respective clones requires further investigation (full-length sequencing, isolation of full-length clones, investigation of open reading frames, etc) and any presumptions of function have to be tested experimentally. These investigations go beyond the initial scope of an EST study and reflect some of its limitations. Nevertheless, the limitations of an EST approach are fully counter-balanced by the value of positive identifications of ESTs which allow a transcriptional picture to be built.

Although all RPE cDNA libraries available for this study were sampled, the majority of ESTs analyzed were of clones isolated at random from the cDNA library of human foetal RPE cells, due to the substantially higher content of recombinants with informative-sized inserts from this source. Thus, of the total 203 sequences analyzed in this study, representing 25 bovine and 178 human RPE clones, a total of 166 informative RPE ESTs were defined. The remaining 37 sequences (13 bovine and 24 human) were eliminated from the subsequent analysis, being identified as representing ribosomal genes, repeats, chimaeric clones and adaptor multimers.

Analysis of bovine RPE ESTs

In the case of bovine RPE ESTs, the number of ESTs analyzed was insufficient for statistical analysis with respect to their distribution into possible cellular roles. However, a few remarks are of interest.

Most of the bovine ESTs that were putatively identified based on their identity/similarity with database sequences represent widely expressed genes (Table 3.9). The clones similar to P-selectin (83.5% over 103bp) and to *trans*-Golgi p230 protein (70% similarity with the human sequence over 113bp) may represent two exceptions. P-selectin is a cell adhesion molecule which has been shown to be expressed in platelets and endothelial cells (Strubel *et al.*, 1993). When exposed to the cell surface, upon endothelial cell activation, P-selectin promotes the binding of neutrophils and monocytes to the vessell wall at sites of tissue injury. p230 protein is a membrane protein associated with the cytosolic face of the *trans*-Golgi apparatus. Based on the analysis of its structural characteristics, it was suggested that p230 may play a role in vesicular transport from the *trans*-Golgi (Erlich *et al.*, 1996). The similarity of the two ESTs isolated from a bovine RPE cDNA library with the sequences encoding P-selectin and p230 proteins provides only a rough indication for the existence in the RPE of similar gene expression products, which may have specific importance for cell signalling/cell adhesion and, respectively, for secretory/trafficking activities of the RPE cells. Evidently, any conclusion regarding the nature and expression in the RPE of the genes represented

by these ESTs require further characterization. Until such investigations are carried out, even the possibility that these ESTs represent contaminants could not be excluded.

Three other bovine ESTs revealed by this study, grouped in the "unclassified" category, may represent new bovine ESTs. One of these ESTs showed a significant sequence similarity to a Mus musculus blastocyst cDNA, while the other two ESTs could not be assigned an identification following the sequence comparison searches.

Analysis of human RPE ESTs

Of the 178 informative sequences obtained of clones randomly isolated from the human RPE cDNA library, 154 informative ESTs were defined. 15 ribosomal, 2 chimaeric, 3 repetitive sequences and 4 vector-adaptor-only clones were excluded from the batch of sequences further analyzed. The evaluation of the 154 RPE ESTs surveyed in this study is summarized in Table 4.1.

Sequence comparison searches revealed significant homologies for 126 RPE ESTs (more than 60% identity over at least 150bp overlap). "Perfect" nucleotide sequence matches with database sequences were shown by 113 ESTs (categories 1/A/i and 1/B/i), i.e. by 73.4% of the total ESTs analyzed. However, a significant contribution to this proportion was made by sequences matching anonymous sequences previously reported by cDNA/genomic sequencing projects, with no function assigned. ESTs representing sequences with function or localization previously characterized at some extent amount to only 44.8% of the total ESTs analyzed (category 1/A/i above). This fact is a direct consequence of the high rate in which sequences have been deposited in public databases during the last few years and illustrates the current difference in pace between unveiling new sequence data and characterizing its function.

Thus, more than half of the ESTs analyzed represent genes whose expression products are not yet characterized, encompassing sequences identical to anonymous sequences in databases (category 1/B/i), sequences showing partial homology to known genes or ESTs (category

1/A/ii, respectively 1/B/ii) and sequences presenting no significant homology with other sequences in current databases (category 2).

The current survey of RPE ESTs revealed 41 new human ESTs (categories 1/A/ii, 1/B/ii and 2), representing potentially a maximum of 41 new genes. Of these, 13 ESTs (categories 1/A/ii, 1/B/ii) showed sequence similarities with database sequences (between 60% and 80% identity over entire length or at least 150 bp) and are, therefore, likely to represent related genes.

Table 4.1. Summary of sequence analysis of 154 informative ESTs of randomly isolated human RPE clones.

(1) 126 ESTs (81.8%) matched sequences in public databases
—
(A) 75 ESTs (48.7%) putatively identified as sequences of known/presumed function or localization
(i) 69 ESTs (44.8%) identical match
(ii) 6 ESTs (3.9%) similar to known genes
(B) 51 ESTs (33.1%) matched anonymous sequences with no function assigned
(i) 44 (28.6%) ESTs identical match
(ii) 7 (4.5%) similar to reported ESTs (less than 80% identity over their entire length)
(2) 28 ESTs (18.2%) presented no significant homology to database sequences

4.2.1.3. Classification of RPE ESTs with respect to their presumed cellular role(s)

The ESTs for which putative identifications were made (category 1/A) were assigned to cellular roles according to the known or putative involvement of the represented gene/protein in a cellular process (Table 3.11). ESTs representing genes for which the role was not known and could not be inferred were grouped in a category of "unclassified ESTs". Although other schemes of classification might be considered, grouping sequences that share similar functional characteristics and/or cellular roles (as defined in Adams *et al.*, 1995) appeared to best support the aim of this study.

Highly expressed activities of RPE cells are indicated by a high proportion of ESTs representing transcripts assigned to roles in metabolism and gene/protein expression: more than half of the total RPE ESTs assigned to cellular roles were associated with these two main categories (see Table 3.12). The expression of mitochondrial gene sequences (grouped in the subcategory metabolism: energy/TCA cycle) is consistent with a tissue with a high energy requirement.

A prominent feature that resulted from the sorting of the RPE ESTs according to their possible cellular roles, is that approximately a third of the total number of assigned ESTs represent genes with involvement in gene/protein expression. The high representation of transcripts involved in the protein synthesis machinery in the transcriptional picture revealed by this study is consistent with a high protein synthesis/turnover in the RPE cells. In particular, the strikingly high representation of clones coding for proteins involved in proteolytic activity and its regulation indicates that these proteins have important roles in supporting functions of the RPE.

In this study, 8 transcripts were represented by more than one EST each. Five of these were of genes assigned to metabolism and gene/protein expression categories, thus confirming the general trend in the distribution of RPE ESTs into functional role categories. One gene, that coding for precursor cystatin C, was represented by 3 independent ESTs. This rather unexpected result emerged as a main finding of this study and

stimulated a more in-depth investigation of these clones, which will be discussed in the following section of this chapter.

4.2.1.4. Tissue distribution of ESTs identified in the RPE

Of the 62 different putative identifications assigned to RPE ESTs, only 10 genes encode transcripts shown to be widely expressed, i.e. expressed in more than 20 tissues. The information compiled by Adams *et al.* (1995) regarding widely expressed genes in most major human organs and tissues was used for this assessment. The above-cited study found only 8 genes expressed in all 30 tissues sampled using at least 1000 ESTs. The fact that the present RPE study revealed ESTs of 3 of these most widely expressed genes (ribosomal proteins L3, S9 - 2 ESTs and S14 - 2 ESTs) is a good indication of the reliability of the transcriptional details unveiled. The majority of genes matched by RPE ESTs were previously shown to be present in less than 10 tissues. A total of 38 genes represented by RPE ESTs defined by the present study were not identified in the eye by previous EST studies. Although this is likely to be a direct consequence of the limits on the depth of sampling of any EST study, it illustrates, paradoxically, the significant potential of any such study for extending the profile of genes expressed for the tissue of interest.

The high degree of matches of RPE ESTs with genes expressed in brain (38 ESTs assigned to functional roles and 14 anonymous ESTs) is not surprising given the common neuroectodermic origin of these tissues, although it may also be a direct consequence of the fact that the brain tissues are the most extensively sampled material in EST studies.

4.2.2. LOOKING AT A PROMINENT COMPONENT OF THE TRANSCRIPTIONAL PICTURE OF RPE CELLS: CYSTATIN C

4.2.2.1. Characteristics of RPE cystatin C clones

An interesting feature of the EST analysis of the RPE cells was the high proportion of cDNAs assigned to cellular roles related to protein turnover. In particular, the cDNA data indicate that one of the most abundant transcripts in RPE is that of a specific cysteine proteinase inhibitor, cystatin C. Three independent ESTs of the precursor cystatin C were identified among 178 cDNA clones randomly selected from one cDNA library of RPE cells.

The full-length sequencing analysis of cystatin C-encoding RPE clones revealed that each clone has a slightly different insert (thus showing that they are not merely the result of amplification of one initial clone). Two of these clones (K22 and K26) contain the full-length precursor cystatin C cDNA (see Figure 3.8). Compared with the reported cDNA sequence for cystatin C (accession number X05607), the longest of the three RPE clones (K26) has a 4 bp longer 5'-untranslated region and a 3 bp longer 3'-untranslated region; the latter feature is also shared by clone K22. Both clones K22 and K26 contain one mismatch compared with the previously reported database sequence, but the encoded amino acid sequence is not affected. The cause of the appearance of a variant in the 3'-untranslated region which was detected for clone K210 is difficult to assess but could represent a sequence polymorphism in the population.

Noteworthy, the clarification of cystatin C sequence led to a conclusion of general interest for the analysis of data produced by automatic sequencing. Following full-length sequencing of the plus strand of all three clones, it appeared that there were 6 mismatches compared with the database sequence encoding precursor cystatin C, 4 of which appeared in the coding region. These mismatches were in codon degeneracy positions and thus not affecting the amino acid sequence encoded (3 codons appearing GAA instead of GAG and therefore still

coding for Glu and 1 codon being CAA instead of CAG, thus still coding for Gln). "A" and not "G" for each one of these positions, was clearly indicated in sequence output files (both in the electropherogram and automatic base calling). However, the sequence of the opposite strand revealed "C" in the corresponding positions, consistent with the reported database sequence. Visual inspection of the plus strand sequence confirmed a small but definite "G" peak concealed by a preceding big "A" peak. This clarification of sequence illustrates the fact that achieving a balance between speed of accumulation of data by automatic sequencing and the actual examination of sequence files bears a crucial importance for the accurate interpretation of data.

The proportion of clones coding for precursor cystatin C was shown by the plaque hybridization experiment (section 3.3.2.2) to be representative of the whole cDNA library of human foetal cultured RPE cells (1.35% of the clones of the library hybridized to a probe created from clone K26). The presence of cystatin C transcript in cultured human RPE cells was confirmed by Northern blot hybridization (Figure 3.13). The occurrence of two RNAs hybridizing to the cystatin C-specific probe may be due to the presence of splicing variants or to the existence of a related transcripts. Noteworthy is the fact that a similar pattern of hybridization has been previously reported for RNA of rat testes (see discussion in the following section).

Investigation of tissue distribution of cystatin C transcript (using a multiple tissue Northern blot) indicated a wide variety of tissues expressing cystatin C (see Figure 3.14). Practically all adult tissues examined presented a certain level of expression (various intensities) with the highest levels being apparent in salivary gland, brain tissues, adrenal gland, lung and placenta. Expression was barely noticeable in two tissues (skeletal muscle and kidney). The lack of expression in the kidney is particularly notable given the reports of other studies in which expression and function of cystatin C in the kidney were specifically investigated; the reason for this result was not determined and the possibility that the particular mRNA preparation spotted onto the blot was of poor quality cannot be excluded.

The role of cystatin C in the RPE has not been investigated up to this date. Previous studies of its expression in ocular tissues are quite controversial, as will be detailed further. Yet, based on the findings of the present study, it is likely that cystatin C is produced within the RPE and is functionally important. A brief description of current knowledge regarding cystatin C expression and roles follows to serve as an aid to further discussion.

4.2.2.2. Proteolysis and cystatin C

A wealth of physiological processes in all organisms require various proteolytic activities which are carried out by extracellular and intracellular proteases. Proteinases, which are enzymes capable of cleaving peptide bonds inside the polypeptide chain, are classified according to the nature of the catalytic group in the active centre and important for the enzymatic mechanism, as serine, cysteine, aspartic and metallo-proteinases (proteinases of unidentified catalytic mechanism are "unclassified"; Barrett, 1986).

Cysteine proteinases play a major role in cell protein turnover, being involved in the intracellular catabolism of peptides and proteins (Barrett & Kirschke, 1981), proteolytic processing of proenzymes (Taugner *et al.*, 1985) and prohormones (Marks *et al.*, 1986), extracellular breakdown of collagen (Etherington, 1980). They also may be involved in modulating the penetration of tissues by malignant cells (Sloane & Honn, 1984) and also by microorganisms (Barrett *et al.*, 1986b). Experimental evidence for a possible involvement of cysteine proteinases, alongside other lysosomal proteinases, in human brain disorders has been reviewed by Bernstein (1994). An important role in regulation of the physiological activity of cysteine proteinases is achieved by specific inhibitors which consequently are involved in a myriad of normal and pathological processes (for a review, see Roberts *et al.*, 1995).

Cystatin C belongs, alongside salivary cystatins S, SA, SN and cystatin D, to human group of Family 2 of the cystatin superfamily of protein inhibitors of cysteine proteinases (Barrett *et al.*, 1986a). Cystatin C shows

about 54% sequence homology with human salivary cystatins (Bobek & Levine, 1992). All family 2 cystatins (which includes the first detected and most studied cystatin, that from chicken egg-white (Anastasi *et al.*, 1983)) have two disulfide bridges, thus differing from Family 1 cystatins (also called stefins) comprising cystatin A and B which lack disulfide bridges and from kininogens (Family 3) which are more complex glycoproteins containing three copies of Family 2 cystatins and 9 disulphide bridges.

Cystatin C (encoded by the human gene CST3 (Abrahamson *et al.*, 1990)) is a small protein consisting of 120 amino acid residues in a single polypeptide chain which is synthesized with a signal leader peptide of 26 amino acid residues (Abrahamson *et al.*, 1987). Cystatin C is the tightest-binding inhibitor of lysosomal cysteine proteinases cathepsin B, H and L. Regions identified by functional studies as important for the inhibitory activity include the evolutionarily conserved glycine residue in the N-terminal region (Gly-11), the three residues preceding Gly-11 (Hall *et al.*, 1993) and, based on similarity with chicken cystatin, segments Gln-55_Gly-59 and Pro-105_Trp-106 (Lindahl *et al.*, 1992). Although disulphide bridges are generally characteristic of extracellular rather than intracellular proteins, cystatin C, as well as the other members of Family 2 cystatins, has been detected both inside and outside cells. Noteworthy though, most of the knowledge about cystatin C function relates to its extracellular activity.

Remarkably, cystatin C is present in all human biological fluids in concentrations implying a physiological role (Abrahamson *et al.*, 1986). As such, it is thought to be a major local regulator of the cysteine proteinase activity. In the cerebrospinal fluid (CSF) cystatin C is the dominating cysteine proteinase inhibitor (5.5 ratio concentration over that in blood plasma (Lofberg *et al.*, 1980)) and has, probably, a protective function for the cellular elements of the central nervous system. Cystatin C could be important in maintaining the integrity of cell surface proteins in the brain by suppressing undesirable proteinase activity, for example when proteinases are released from destroyed or dying cells. It is implicated in a hereditary form of cerebral hemorrhage (hereditary cystatin C amyloid angiopathy) characterized by the deposition of a variant of cystatin C (having one Leu residue substituted by a Gln) as amyloid in cerebral arteries walls,

associated with low concentration of cystatin C in CSF (Ghiso *et al.*, 1986). Cystatin C has also been identified in senile plaques of Alzheimer disease patients, suggesting an involvement in the pathogenesis of plaques possibly by influencing the processing cf the amyloid precursor (li et al., 1993).

Expression of cystatin C has been investigated and shown in a variety of tissues, such as salivary glands, brain, testes, seminal vesicles, kidney, liver, pancreas, intestine, stomach, lung, placenta (Abrahamson *et al.*, 1990; Turk & Bode, 1991; Twining, 1994). Notably, the brain tissue with the highest cystatin C mRNA level is the choroid plexus, according to a study in rat brain (Cole *et al.*, 1989), thus explaining the supply of cystatin C to the cerebrospinal fluid which is produced mainly by the choroid plexus.

Regulatory functions of cystatin C have been inferred for protection of kidney glomerular tissue (Baricos *et al.*, 1990) and for the control of proteolysis during spermatogenesis (Tsuruta *et al.*, 1993). Interestingly, the latter study described two sizes of transcripts of cystatin C in rat testes (of approximately 700 and 600 nucleotides), with the larger transcript seen also in the brain tissue and the shorter transcript expressed in the kidney. Similar observations were made for cystatin C transcripts in rat testes, i.e. two cystatin C mRNAs of approximately 700 bp and 500 bp in length were detected (Cole *et al.*, 1989). These observations suggest that cystatin C transcript may be subjected to tissue-specific splicing. The existence of two related genes also can not be excluded.

Secretion of cystatin C by human carcinoma cell lines (colon (Corticchiato *et al.*, 1992) and thyroid (Barka *et al.*, 1992)) and aberrant regulation of cystatin C associated with increased production of cysteine proteinases in the invasive phenotype of many metastatic cell lines (Chambers *et al.*, 1992) support the hypothesis that cystatin C plays a role in controlling cell invasiveness. Further support for this hypothesis is provided by the recent finding regarding spatial and temporal coordination of cystatin C expression with expression of cathepsins B and L during implantation and placentation of mouse embryo (Afonso *et al.*, 1997). Such a function may be part of a general role of modulating proteolytic degradation of extracellular matrix.

4.2.2.3. What role could cystatin C have in the RPE?

Cystatin C expression has not been investigated in ocular structures to any great extent and the results of the few studies published to date are contradictory. Cystatin immunoreactivity has been detected in monkey aqueous humour (Russell & Epstein, 1992). Given the similarities in composition between aqueous humour and CSF, and the existence of the blood-retinal and blood-CSF barrier in the eye and the central nervous system, respectively, it has been inferred that cystatin C is not an ultrafiltrate of blood plasma. Its site of synthesis in chick eye has been suggested to be the epithelium of the ciliary body (Colella, 1996). However, in the rat eye (Barka & Van der Noen, 1994), the expression of cystatin C was evident only in the sclera and retina and was found to be absent from the ciliary body. The retinal activity was associated with glia and rods by *in situ* hybridization, but the authors pointed out themselves that "it is difficult to establish the exact cellular location of the hybridization signal". It was suggested that a possible explanation for the differences of the findings in the two studies may be species variation (Colella, 1996); by analogy with the brain, where it was suggested that expression of cystatin C first evolved in the choroid plexus and additional sites appeared later in the evolution of mammalian species (Tu *et al.*, 1992), an evolutionary shift of the site of cystatin C synthesis from the ciliary epithelium to other cells in the eye could be speculated.

Neither study investigated specifically cystatin C expression in the RPE. To my knowledge, the only report indicating expression of cystatin C in the RPE, previous to the present study, is that of Gieser & Swaroop (1992) who identified two ESTs representing cystatin C in a subtracted cDNA library of cultured RPE cells.

Based on the findings of the present study, it is likely that cystatin C is, at least *in vitro*, functionally important for the regulation of RPE-driven proteolytic events. Cystatin C activity seems to be associated in particular with transporting epithelia and the RPE is very much a transporting epithelium (Bok, 1993; Joseph & Miller, 1991; Quin & Miller, 1992). It is not known what the target destination of cystatin C might be and whether its roles are carried out inside the RPE cells and/or elsewhere in ocular

structures. Three plausible possibilities regarding the functional importance of cystatin C expressed by the RPE cells are discussed here.

(1) Cystatin C could be involved in modulating cysteine proteinase activity around the rods and cones in the interphotoreceptor matrix (IPM). The photoreceptor cell is a delicate terminally differentiated cell placed in an environment, i.e. IPM, known to modify by transport activities and synthesis of matrix components. The presence of lysosomal enzymes in the IPM is due, at least in part, to secretion by the RPE (Adler, 1989). Modulation of enzymes likely to cause damage to photoreceptors would agree with the known role of the RPE as "nursemaid of the photoreceptors". Moreover, cystatin C is a component of ocular biological fluids which provide a nutritive or protective role such as tears (Barka *et al.*, 1991) and aqueous humour; hence, its presence in the nutritive environment of IPM would not be surprising.

(2) RPE cells have a highly developed phagolysosomal system involved in degradation of shed photoreceptors outer segments. Cysteine proteinases of the cathepsin B type are bound in lysosomes (Tsung *et al.*, 1984) and could have a role, either directly in the digestion process or in the processing/activation of other lysosomal enzymes involved in this function ((Rakoczy *et al.*, 1994); see section 1.2.1.4). Noteworthy is the fact that cathepsin B has been demonstrated to activate serine proteinases by inactivation of normal protective serine proteinase inhibitors (Johnson & Travis, 1977). If cystatin C is involved in regulation of lysosomal enzyme activity concerned with degradation of shed photoreceptor membranes then it becomes a key player in the biology of photoreceptor turnover.

(3) The RPE lies on a basement membrane which in turn lies on the complex collagenous layer of Bruch's membrane. These tissues undergo systematic remodelling throughout life and changes in Bruch's membrane are central to pathological conditions, such as age-related macular degeneration (see section 1.2.2.5). Since cysteine proteinases are involved in collagen degradation (see section 4.2.2.2 above), a posible role for cystatin C in regulation of turnover and remodelling of Bruch's membrane cannot be excluded.

4.3. FUTURE DIRECTIONS

Two main experimental lines will be pursued to further the study of cystatin C expression in the RPE: endogenous expression, and expression of cystatin C - Green Fluorescent Protein (GFP) fusion proteins.

(1) Endogenous expression of cystatin C. Two questions will be addressed:

a) is there cytological evidence for cystatin C synthesis in the RPE *in vivo* ? and

b) can cystatin C protein be detected in the RPE?

Concerning synthesis of cystatin C mRNA, *in situ* hybridization (ISH) experiments will be designed to localize this mRNA inside RPE and neuroretina cells. Non-radioactive ISH will be carried out on tissue specimens, using digoxigenin-labelled RNA probes. Complementary RNA probes will be generated by transcription of cystatin C cDNA insert of clone K26, taking advantage of the directionality of cloning in pBluescript SK(-) vector.

At the protein level, cystatin C will be detected using a specific antibody by immunofluorescence. Experiments will be carried out on sections of retina/RPE and on cultured RPE cells, in pre-confluent, confluent and post-confluent culture conditions. These investigations will utilise a commercially available rabbit anti-human cystatin C polyclonal antibody (with specificity for both cystatin C and its precursor) and a fluorescently-labelled goat anti-rabbit IgG as secondary antibody. The distribution of the protein in relation to cell polarity can be informative with regard to the possible role performed by cystatin C in the RPE.

(2) Expression of cystatin C - GFP fusion proteins in transfected RPE cells. In order to examine the way in which the RPE cells handle the protein once it is synthesized, a genetic reporter system based on GFP will be exploited. Cystatin C cDNA inserts, including and excluding the fragment coding for the signal peptide, provided by one of the clones K22, K26 which contain the full-length cystatin C cDNA, will be fused in frame to C-terminus of tag protein. In these constructs the stop codon of cystatin C cDNA will end

the translation of fusion protein. The fluorescent tag on cystatin C protein will allow *in vivo* localization of the fusion protein. These observations will help formulate the possible roles of cystatin C in ocular structures.

Epilogue

More RPE ESTs certainly will be characterized in the future, so helping to build up a list of genetic and biochemical components that contribute to a functioning RPE cell. As the investigation described in this thesis showed, it is conceivable that the analysis of the transcriptional picture thus revealed will unravel genes of biological importance for the functions carried out by RPE cells and of potential medical interest. Cystatin C may be one of them.

REFERENCES

Abrahamson, M., Barrett, A., Salvegen, G. & Grubb, A. (1986). Isolation of six cysteine proteinase inhibitors from human urine. Their physicochemical and enzyme kinetic properties and concentrations in biological fluids. *J Biol Chem* **261**, 11282-11289.

Abrahamson, M., Grubb, A., Olafsson, I. & Lundwall, A. (1987). Molecular cloning and sequence analysis of cDNA coding for the human cysteine proteinase inhibitor cystatin C. *FEBS Lett* **216**(2), 229-233.

Abrahamson, M., Olafsson, I., Palsottir, A., Ulvback, M., Lundwall, A., Jensson, O. & Grubb, A. (1990). Structure and expression of the human cystatin C gene. *Biochem J* **268**, 287-294.

Adams, M., Dubnick, M., Kerlavage, A., Moreno, R., Kelley, J., Utterback, T., Nagle, J., Fields, C. & Venter, J. (1992). Sequence identification of 2,375 human brain genes. *Nature* **355**, 632-634.

Adams, M., Kelley, J., Gocayne, J., Dubnick, M., Polymeropoulos, M., Xiao, H., Merril, C., Wu, A., Olde, B., Moreno, R., Kerlavage, A., McCombie, W. & Venter, J. (1991). Complementary DNA Sequencing: Expressed Sequence Tags and Human Genome Project. *Science* **252**, 1651-1656.

Adams, M., Kerlavage, A., Fields, C. & Venter, J. (1993). 3,400 new expressed sequence tags identify diversity of transcripts in human brain. *Nature Genet* **4**, 256-267.

Adams, M., Kerlavage, A., Fleischmann, R., Fuldner, R., Bult, C., Lee, N. & Kirkness, E. (1995). Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. *Nature* **377 Suppl**, 3-17.

Adler, A. (1989). Selective presence of acid hydrolases in the interphotoreceptor matrix. *Exp Eye Res* **49**, 1067-1077.

Adler, A. & Kluznik, K. (1982). Proteins and glycoproteins of the bovine interphotoreceptor matrix: Composition and fractionation. *Exp Eye Res* **34**, 423-434.

Adler, A. & Martin, K. (1983). Lysosomal enzymes in the interphotoreceptor matrix: acid protease. *Curr Eye Res* **2**, 359-365.

Affara, N., Bentley, E., Davey, P., Pelmear, A. & Jones, M. (1994). The identification of novel gene-sequences of the human adult testis. *Genomics* **22**, 205-210.

Afonso, S., Romagnano, L. & Babiarz, B. (1997). The expression and function of cystatin C and cathepsin B and cathepsin L during mouse embryo implantation and placentation. *Development* **124**, 3415-3425.

Agarwal, N., Swaroop, A., Zheng, K., Liou, J.-d., O'Rourke, K., Graves, K., Gieser, L., Del Monte, M. & Yang-Feng, T. (1995). Expression and chromosomal localization of cDNA clones from an enriched human retinal pigment epithelial (RPE) cell line library: identification of two RPE-specific genes. *Cytogenet Cell Genet* **69**, 71-74.

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. & Watson, J. (1989). RNA Synthesis and RNA Processing. In *Molecular Biology of the Cell* 2nd edit. (Adams, R., ed.), pp. 523-550. Garland Publishing, Inc., New York.

Altschul, S., Gish, W., Miller, W., Myers, E. & Lipman, D. (1990). Basic Local Alignment Search Tool. *J Mol Biol* **215**, 403-410.

Anastasi, A., Brown, M., Kembhavi, A., Nicklin, M., Sayers, C., Sunter, D. & Barrett, A. (1983). Cystatin, a protein inhibitor of cysteine proteinases. *Biochem J* **211**, 129-138.

Anderson, D. & Fisher, S. (1975). Disc shedding in the rod-like and cone-like photoreceptors of tree squirrels. *Science* **187**, 953-955.

Aparicio, S., Morrison, A., Gould, A., Gilthorpe, J., Chaudhuri, C., Rigby, P., Krumlauf, R. & Brenner, S. (1995). Detecting conserved regulatory elements with the model genome of the japanese puffer fish, Fugu rubripes. *Proc Natl Acad Sci USA* **92**, 1684-1688.

Baricos, W., Cortez, S., Le, Q., Zhou, Y., Dicario, R., O'Connor, S. & Shah, S. (1990). Glomerular basement membrane degradation by endogenous cysteine proteinases in isolated rat glomeruli. *Kidney Int* **38**, 395.

Barka, T., Asbell, P., van der Noen, H. & Prasad, A. (1991). Cystatins in human tear fluid. *Curr Eye Res* **10**, 25-34.

Barka, T. & Van der Noen, H. (1994). Expression of the cysteine proteinase inhibitor cystatin C mRNA in rat eye. *Anat Rec* **239**, 343-348.

Barka, T., Van der Noen, H. & Patil, S. (1992). Cysteine proteinase inhibitor in cultured human medullary thyroid carcinoma cells. *Lab Invest* **66**, 691-700.

Barrett, A. (1986). An introduction to the proteinases. In *Proteinase inhibitors* (Barrett, A. & Salvesen, G., eds.), pp. 3-54. Elsevier Science Publishers B.V., Amsterdam.

Barrett, A., Fritz, H., Grubb, A., Isemura, S., Jarvinen, M., Katunuma, N., Machleidt, W., Muller-Esterl, W., Sasaki, M. & Turk, V. (1986a). Nomenclature and classification of the proteins homologous with the cysteine proteinase inhibitor chicken cystatin. *Biochem J* **236**, 312.

Barrett, A. & Kirschke, H. (1981). Cathepsin B, cathepsin H and cathepsin L. *Meth Enzymol* **80**, 535-561.

Barrett, A., Rawlings, N., Davies, M., Machleidt, G., Salvesen, G. & Turk, V. (1986b). Cysteine proteinase inhibitors of the cystatin superfamily. In *Proteinase inhibitors* (Barrett, A. & Salvesen, G., eds.), pp. 515-569. Elsevier Science Publishers B.V., Amsterdam.

Barry, R., Canada, F. & Rando, R. (1989). Solubilization and partial purification of retinyl ester synthetase and retinoid isomerase from bovine ocular pigment epithelium. *J Biol Chem* **264**, 9231-9238.

Basinger, S., Hoffman, R. & Matthes, M. (1976). Photoreceptor shedding is initiated by light in the frog retina. *Science* **194**, 1074-1076.

Basu, P. K., Sarkar, P., Menon, I., Carre, F. & Persad, S. (1983). Bovine retinal pigment epithelial cell cultured in vitro:growth characteristics, morphology, chromosomes, phagocytosis ability, tyrosine activity and effect of freezing. *Exp. Eye Res.* **36**, 671-683.

Baxendale, S., Abdulla, S., Elgar, G., Buck, D., Berks, M., Micklem, G., Durbin, R., Bates, G., Brenner, S., Beck, S. & Lehrach, H. (1995). Comparative sequence-analysis of the human and pufferfish Huntingtons-disease genes. *Nature Genetics* **10**, 67-76.

Berman, E. (1971). Acid hydrolases of the retinal pigment epithelium. *Invest Ophthalmol* **10**, 64-68.

Berman, E. (1979). Biochemistry of the Retinal Pigment Epithelium. In *The Retinal Pigment Epithelium* (Zinn, K. M. & Marmor, M. F., eds.), pp. 83-102. Harvard University Press, Cambridge, Mass.

Bernstein, H. (1994). The many faces of lysosomal proteinases (cathepsins) in human neuropathology. A histochemical perspective. *Eur J Histochem* **38**, 189-192.

Bernstein, P., Law, W. & Rando, R. (1987). Isomerization of all-*trans* retinoids to 11-*cis* retinoids in vitro. *Proc Nat Acad Sci USA* **84**, 1849-1853.

Bernstein, S., Borst, D., ME, N. & Wong, P. (1996). Characterization of a human fovea cDNA library and regional differential gene expression in the human retina. *Genomics* **32**, 301-308.

Besharse, J. (1984). Photoreceptor disk shedding and phagocytosis-mechanism and regulation. *J Cell Biol* **99**, A115.

Besharse, J. & Dunis, D. (1982). Rod photoreceptor disc shedding in vitro: inhibition by cytochalasins and activation of colchicine. In *The Structure of the Eye* (Hollyfield, J., ed.), pp. 85-96. Elsevier Biomedical, New York.

Besharse, J. & Iuvone, P. (1983). Circadian clock in Xenopus eye controlling retinal serotonin N-acetyltransferase. *Nature* **305**, 133-135.

Birnboim, H. & Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl Acids Res* **7**, 1513-1522.

Bobek, L. & Levine, M. (1992). Cystatins - inhibitors of cysteine proteinases. *Critical Reviews in Oral Biology and Medicine* **3**, 307-332.

Boguski, M., Lowe, T. & Tolstoshev, C. (1993). dbEST - database for "expressed sequence tags". *Nature Genetics* **4**, 332-333.

Bok, D. (1982). Autoradiographic studies on the polarity of plasma membrane receptors in retinal pigment epithelial cells. In *The Structure of the Eye* (Hollyfield, J., ed.), pp. 247-256. Elsevier North Holland, New York.

Bok, D. (1985). Retinal photoreceptor-pigment epithelium interactions. *Invest Ophthal Vis Sci* **26**, 1659-1694.

Bok, D. (1988). Retinal photoreceptor disc shedding and pigment epithelium phagocytosis. In *Retinal Diseses:Medical and Surgical Management* (Ryan, S., Ogden, T. & Schachat, A., eds.), Vol. 1, pp. 69-81. C.V. Mosby Co., St.Louis.

Bok, D. (1990). Processing and transport of retinoids by the retinal pigment epithelium. *Eye* **4**, 326-332.

Bok, D. (1993). The retinal pigment epithelium: a versatile partner in vision. *J Cell Sci* **Suppl 17**, 189-195.

Bok, D. & Hall, M. (1971). The role of the pigment epithelium in the etiology of inherited retinal dystrophy in the rat. *J Cell Biol* **49**, 664-682.

Bok, D., O'Day, W. & Rodriguez-Boulan, E. (1992). Polarized budding of vesicular stomatitis and influenza virus from cultured human and bovine retinal pigment epithelium. *Exp Eye Res* **55**, 853-860.

Bok, D. & Young, R. (1979). Phagocytic Properties of the Retinal Pigment Epithelium. In *The Retinal Pigment Epithelium* (Zinn, K. M. & Marmor, M. F., eds.), pp. 148-174. Harvard University Press, Cambridge, Mass.

Bosch, E., Horwitz, J. & Bok, D. (1993). Phagocytosis of outer segments by retinal pigment epithelium:phagosome-lysosome interaction. *J Histochem Cytochem* **41**, 253-263.

Botcherby, M. (1997). Of mice and men. *Genome News* **22**, 21-23.

Boulton, M., Moriarty, P., Jarvis-Evans, J. & Marcyniuk, B. (1994). Regional variation and age-related changes of lysosomal enzymes in the human retinal pigment epithelium. *Br J Ohthalmol* **78**, 125-129.

Brenner, S. (1990). The human genome - the nature of the enterprise. *Ciba Foundation Symp* **149**, 6-17.

Brenner, S., Elgar, G., Sandford, R., Macrae, A., Venkatesh, B. & Aparicio, S. (1993). Characterization of the Pufferfish (Fugu) genome as a compact model vertebrate genome. *Nature* **366**, 265-268.

Burke, J. & Twining, S. (1988). Regional comparisons of cathepsin D activity in bovine retinal pigment epithelium. *Invest Ophthalmol Vis Sci* **29**, 1789-1793.

Buse, E., Eichmann, T., de Groot, H. & Leker, A. (1993). Differentiation of the mammalian retinal pigment epithelium in vitro: influence of presumptive retinal neuroepithelium and head mesenchyme. *Anat Embryol* **187**, 259-268.

Cahill, G. & Besharse, J. (1992). Light-sensitive melatonin synthesis by Xenopus photoreceptors after destruction of the inner retina. *Vis Neurosci* **8**, 487-490.

Caldwell, R. & McLaughlin, B. (1983). Permeability of retinal pigment epithelial cell junctions in the dystrophic rat retina. *Exp Eye Res* **36**, 415-427.

Campochiaro, P. & Glaser, B. (1985). Endothelial cells release a chemoattractant for retinal pigment epithelial cells in vitro. *Arch Ophthalmol* **103**, 1876-1880.

Campochiaro, P., Jerdan, J. & Glaser, B. (1986). The extracellular-matrix of human retinal-pigment epithelial-cells in vivo and its synthesis in vitro. *Invest Ophthalmol Vis Sci* **27**, 1615-1621.

Chaitin, M. & Hall, M. (1983a). Defective ingestion of rod outer segments by cultured rat pigment epithelial cells. *Invest Ophthalmol Vis Sci* **24**, 812-831.

Chaitin, M. & Hall, M. (1983b). The distribution of actin in cultured normal and dystrophic rat pigment epithelial cells during the phagocytosis of rod outer segments. *Invest Ophthalmol Vis Sci* **24**, 821-831.

Chambers, A., Colella, R., Denhardt, D. & Wilson, S. (1992). Increased expression of cathepsins L and B and decreased activity of their inhibitors in metastatic, ras-transformed NIH 3T3 cells. *Molecular Carcinogenesis* **5**, 238-245.

Chen, C.-C. & Heller, J. (1977). Uptake of retinol and retinoic acid from serum retinol-binding protein by retinal pigment epithelial cells. *J Biol Chem* **252**, 5216-5221.

Cohen, L. H. & Noell, W. K. (1965). Relationship between visual function and metabolism. In *Biochemistry of the Retina* (Graymore, C. N., ed.), pp. 36-49. Acad. Press Inc., London.

Cole, T., Dickson, P., Esnard, F., Averill, S., Risbridger, G., Gauthier, F. & Schreiber, G. (1989). The cDNA structure and expression analysis of the genes for the cystein eproteinase inhibitor cystatin C and for beta2-microglobulin in rat brain. *Eur J Biochem* **186**, 35-42.

Colella, R. (1996). Cystatin C mRNA is expressed by the ciliary epithelium of the chick eye. *J Histochem Cytochem* **44**, 77-79.

Colley, N., Clark, V. & Hall, M. (1987). Surface modification of retinal pigment epithelium cells: effects on phagocytosis and glycoprotein composition. *Exp Eye Res* **44**, 377-392.

Collins, J. & Hohn, B. (1978). Cosmids: A type of plasmid gene-cloning vector that is packageable in vitro in bacteriophage lambda heads. *Proc. Natl. Acad. Sci.* **75**, 4242-4246.

Corticchiato, O., Cajot, J., Abrahamson, M., Chan, S., Keppler, D. & Sordat, B. (1992). Cystatin C and cathepsin B in human colon carcinoma: expression by cell lines and matrix degradation. *Int J Cancer* **52**, 645-652.

Coulombre, A. J. (1979). Roles of the Retinal Pigment Epithelium in the Development of Ocular Tissues. In *The Retinal Pigment Epithelium* (Zinn, K. M. & Marmor, M. F., eds.), pp. 53-57. Harvard University Press, Cambridge, Mass.

Crouch, R., Goletz, P., Yu, S. & Redmond, T. (1997). A possible role for RPE65 in retinoid processing. *Invest Ophthalmol Vis Sci* **38**, S304.
Cunha-Vaz, J. (1979). The blood-ocular barriers. *Surv Ophthalmol* **23**, 279-296.

Dowling, J. & Sidman, R. (1962). Inherited retinal distrophy in the rat. *J Cell Biol* **49**, 664-682.

Edwards, R. (1982). Culture of mammalian retinal pigment epithelium and neural retina. *Methods Enzymol* **81**, 39-42.

Eldred, G. (1995). Lipofuscin fluorophore inhibits lysosomal protein degradation and may cause early stages of macular degeneration. *Gerontology* **41**(Suppl 2), 15-28.

Elledge, S. J., Mulligan, J. T., Ramer, S. W., Spottswood, M. & Davis, R. W. (1991). YES:A multifunctional cDNA expression vector for the isolation of genes by complementation of yeast and Escherichia coli mutations. *Proc Natl Acad Sci USA* **88**, 1731-1735.

Erlich, R., Gleeson, P., Campbell, P., Dietzsch, E. & Toh, B.-H. (1996). Molecular charectreization of *trans*-Golgi p230. *J Biol Chem* **271**, 8328-8337.

Etherington, D. (1980). *Ciba Foundation Symposium.*
Feeney-Burns, L. & Eldred, G. (1985). The fate of the phagosome; Conversion to "age pigment" and impact in human retinal pigment epithelium. *Trans Ophthalmol Soc UK* **103**, 416-421.

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

Feinberg, A. & Vogelstein, B. (1984). "A technique for radiolabeling DNA restriction endonuclease fragments to high specific acyivity". Addendum. *Anal Biochem* **137**, 266-267.

Finnemann, S., Marmorstein, A., Neill, J. & Rodriguez-Boulan, E. (1997). Identification of the retinal pigment epithelium protein RET-PE2 as CE-9/OX-47, a member of the immunoglobulin superfamily. *Invest Ophthalmol Vis Sci* **38**, 2366-2374.

Fliesler, S., Cole, G. & Adler, A. (1990). Neural cell adhesion molecule (NCAM) in adult vertebrate retinas: Tissue localization and evidence against its role in retina-pigment epithelium adhesion. *Exp Eye Res* **50**, 475-482.

Forrester, J., Dick, A., McMenamin, P. & Lee, W. (1996). *The Eye. Basic Sciences in Practice*, WB Saunders Co Ltd, London.

Ghiso, J., Pons-Estel, B. & Frangione, B. (1986). Hereditary cerebral amyloid angiopathy: the amyloid fibrils contain a protein which is a variant of cystatin C, an inhibitor of lysosomal cysteine proteases. *Biochem Biophys Res Commun* **136**, 548-554.

Gieser, L. & Swaroop, A. (1992). Expressed sequence tags and chromosomal localization of cDNA clones from a subtracted retinal pigment epithelium library. *Genomics* **13**, 873-876.

Glaser, B., Campochiaro, P., Davis, J. & Sato, M. (1985). Retinal pigment epithelial cells release an inhibitor of neovascularization. *Arch Ophthalmol* **103**, 1870-1875.

Grierson, I., Hiscott, P., Hogg, P., Robey, H., Mazure, A. & Larkin, G. (1994). Development, repair and regeneration of the retinal pigment epithelium. *Eye* **8**, 255-262.

Grierson, I., Mazure, A., Hogg, P., Hiscott, P., Sheridan, C. & Wong, D. (1996). Non-vascular vitreoretinopathy:the cells and the cellular basis of contraction. *Eye* **10**, 671-684.

Grisanti, S. & Guidry, C. (1995). Transdifferentiation of retinal pigment epithelial cells from epithelial to mesenchymal phenotype. *Invest Ophthalmol Vis Sci* **36**, 391-405.

Gu, S., Thompson, D., Srikumari, C., Lorenz, B., Finckh, U., Nicoletti, A., Murthy, K., Rathmann, M., Kumaramanickavel, G., Dento, M. & Gal, A. (1997). Mutations in the RPE65 gene cause autosomal recessive childhood-onset severe retinal dystrophy. *Nature Genet* **17**, 194-197.

Gubler, U. & Hoffman, B. (1983). A simple and very efficient method for generating cDNA libraries. *Gene* **25**, 263-269.

Gundersen, D., Orlowski, J. & Rodriguez-Boulan, E. (1991). Apical polarity of Na,K-ATPase in retinal pigment epithelium is linked to a reversal of the ankyrin-fodrin submembrane cytoskeleton. *J Cell Biol* **112**, 863-872.

Gundersen, D., Powell, S. & Rodriguez-Boulan, E. (1993). Apical polarization of N-CAM in retinal pigment epithelium is dependent on contact with the neural retina. *J Cell Biol* **121**, 335-343.

Hackett, S., Conway, B. & Campochiaro, P. (1989). Subretinal fluid stimulation of retinal epithelial cell migration and proliferation is dependent on certain features of the detachment or its treatment. *Arch Ophthalmol* **107**, 391-394.

Hall, A., dalboge, H., Grubb, A. & Abrahamson, M. (1993). Importance of the evolutionarily conserved glycine residue in the N-terminal region of human cystatin C (Gly-11) for cysteine endopeptidase inhibition. *Biochem J* **291**, 123-129.

Hall, M. & Abrams, T. (1987). Kinetic studies of rod outer segment binding and ingestion by cultured rat RPE cells. *Exp Eye Res* **45**, 907-922.

Hall, M. & Abrams, T. (1991). The phagocytosis of ROS by RPE is not inhibited by mannose-containing ligands. *Exp Eye Res* **53**, 167-170.

Hall, M. O., Bok, D. & Bacharach, A. D. E. (1969). Biosynthesis and assembly of the rod outer segment membrane system. Formation and fate of visual pigment in the frog retina. *J Mol Biol* **45**, 397-406.

Hamel, C., Tsilou, E., Harris, E., Pfeffer, B., Hooks, J., Detrick, B. & Redmond, T. (1993a). Developmentally regulated microsomal protein-specific for the pigment-epithelium of the vertebrate retina. *J Neurosci Res* **34**, 414-425.

Hamel, C., Tsilou, E., Pfeffer, B., Hooks, J., Detrick, B. & Redmond, T. (1993b). Molecular-cloning and expression of RPE65, a novel retinal-pigment epithelium-specific microsomal protein that is posttranscriptionally regulated in-vitro. *J Biol Chem* **268**, 15751-15757.

Hanahan, D. (1983). Studies on transformation of Escherichia coli with plasmids. *J Mol Biol* **166**, 557-580.

Hardison, R., Oeltjen, J. & Miller, W. (1997). Long human-mouse sequence alignments reveal novel regulatory elements: A reason to sequence the mouse genome. *Genome Res* **7**, 959-966.

Hargrave, P. A., McDowell, J. H., Curtis, D. R., Wang, J. K., Juszczak, E., Fong, S. L., Rao, J. K. M. & Argos, P. (1983). The stucture of rhodopsin. *Biophys Struct Mech* **9**, 235-244.

Hayasaka, S. (1983). Lysosomal enzymes in ocular tissues and diseases. *Surv Ophthalmol* **27**, 245-258.

Hayasaka, S., Hara, S. & Mizuno, K. (1975a). Distribution and some properties of cathepsin D in the retinal pigment epithelium. *Exp Eye Res* **21**, 307-313.

Hayasaka, S., Hara, S. & Mizuno, K. (1975b). Partial purification and properties of cathepsin D in the retinal pigment epithelium. *Invest Ophthalmol* **14**, 617-620.

Hayasaka, S., Hara, S. & Mizuno, K. (1976). Lysosomal enzymes in subretinal fluid. *Graefes Arch Klin Exp Ophthalmol* **200**, 13-20.

Hayasaka, S. & Shiono, T. (1982). Alpha-fucosidase, alpha-mannosidase and beta-N-acetyl-glucosaminidase of the bovine retinal pigment epithelium. *Exp Eye Res* **34**, 565-569.

Hewitt, A. & Adler, R. (1994). The retinal pigment epithelium and interphotoreceptor matrix: Structure and specialized functions. In *Retina Basic Science and Inherited Retinal Disease* 2nd edit. (Ryan, S., TE, O. & Schachat, A., eds.), Vol. 1. Mosby-Year Book, Inc., St. Louis.

Hewitt, A., Krawczewicz, L. & Newsome, D. (1984). Changes in Bruch's membrane proteoglycans associated with age and retinitis pigmentosa. *Invest Ophthalmol Vis Sci* **Suppl. 25**, 57.

Hewitt, A., Nakazawa, K. & Newsome, D. (1989). Analysis of newly synthesized Bruch's membrane proteoglycans. *Invest Ophtahlmol Vis Sci* **30**, 478-486.

Hillier, L., Lennon, G., Becker, M., Bonaldo, F., Chiapeli, B., Chissoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M., Tan, F., Thierry-Meg, J., Trevaskis, E., Underwood, K., Wohldman, P., Waterston, R., Wilson, R. & Marra, M. (1996). Generation and analysis of 280,000 human expressed sequence tags. *Genome Res* **6**, 807-828.

Hiscott, P. & Grierson, I. (1991). Subretinal membranes of proliferative vitreoretinopathy. *Br J Ophthalmol* **75**, 53.

Hiscott, P., Grierson, I. & McLeod, D. (1984). Retinal pigment epithelial cells in epiretinal membranes: An immunohistochemical study. *Br J Ophthalmol* **68**, 708-715.

Hiscott, P., Larkin, G., Robey, H., Orr, G. & Grierson, I. (1992a). Thrombospondin as a component of the extracellular matrix of epiretinal membranes: comparisons with cellular fibronectin. *Eye* **6**, 566-569.

Hiscott, P., Waller, H., Grierson, I., Butler, M., Scott, D., Gregor, Z. & Morino, I. (1992b). Fibronectin synthesis in subretinal membranes of proliferative vitreoretinopathy. *Br J Ophthalmol* **76**, 486-490.

Ho, M.-T., Massey, J., Pownall, H., Anderson, R. & Hollyfield, J. (1989). Mechanisms of vitamin A movement between rod outer segments, interphotoreceptor retinoid-binding protein, and liposomes. *J Biol Chem* **264**, 928-935.

Hoar, R. M. (1982). Embryology of the eye. *Environm. Health Perspect.* **44**, 31-34.

Hoffman, S., Friedlander, D., Chuong, C. & Grumet, M. (1986). Differential contributions of Ng-CAM and N-CAM to cell adhesion in different neural regions. *J Cell Biol* **103**, 145-158.

Hogan, M. J., Wood, I. & Steinberg, R. H. (1974). Phagocytosis by pigment epithelium of human retinal cones. *Nature* **252**, 305-307.

Hogan, M. J., Alvarado, J. A. & Weddell, J. E. (1971). *Histology of the Human Eye. An Atlas and Textbook*, WB Saunders Co, London.

Houlgatte, R., Mariage-Samson, R., Duprat, S., Tessier, A., Bentollia, S., Lamy, B. & Auffray, C. (1995). The Genexpress Index: A resource for gene discovery and the genic map of the human genome. *Genome Res* **5**, 272-304.

Ii, K., Ito, H., Kominami, E. & Hirano, A. (1993). Abnormal distribution of cathepsin proteinases and endogenous inhibitors (cystatins) in the hippocampus of patients with Alzheimer's disease, Parkinsonism-dementia complex on Guam, senile dementia and in the aged. *Virchows Archiv* **423**, 185-194.

Irons, M. (1987). Cytochemical localization of Mn++-dependent pyrimidine 5'-nucleotidase activity in isolated rod outer segments. *Exp Eye Res* **44**, 780.

Johnson, D. & Travis, J. (1977). Inactivation of human alpha1-proteinase inhibitor by thiol proteinases. *Biochem J* **163**, 639-641.

Joseph, D. & Miller, S. (1991). Apical and basal membrane ion transport mechanisms in bovine retinal pigment epithelium. *J Physiol* **435**, 439-463.

Ju, J., Kheterpal, I., Scherer, J., Ruan, C., Fuller, C., Glazer, A. & Mathies, R. (1995). Design and synthesis of fluorescence energy-transfer dye-labeled primers and their application for DNA sequencing and analysis. *Anal Biochem* **231**, 131-140.

Junqueira, L. C., Carneiro, J. & Kelley, R. O. (1992). Vision: the photoreceptor system. In *Basic Histology* 7th edit., pp. 470-496. Appleton & Lange, Norwalk, Connecticut.

Kahn, A., Wilcox, A. & Polymeropoulos, M. (1992). Single pass sequencing and physical and genetic mapping of human brain cDNAs. *Nature Genet* **2**, 180-185.

Kanai, M., Raz, A. & Goodman, D. (1968). Retinol binding protein: the transport protein for vitamin A in human plasma. *J Clin Invest* **47**, 2025-2044.

Kato, K. (1990). A collection of cDNA clones with specific expression patterns in mouse brain. *Eur J Neurosci* **2**, 704-711.

Katz, M. (1989). Incomplete proteolysis may contribute to lipofuscin accumulation in the retinal pigment epithelium. *Adv Exp Med Biol* **266**, 109-116.

Katz, M. & Shanker, M. (1989). Development of lipofuscin-like fluorescence in the retinal pigment epithelium in response to protease inhibitor treatment. *Mech Ageing Dev* **49**, 23-40.

Kennedy, C., Rakoczy, P. & Constable, I. (1995). Lipofuscin of the retinal pigment epithelium: a review. *Eye* **9**, 763-771.

Kirchhof, B., Kirchhof, E., Ryan, S. & Sorgente, N. (1989). Vitreous modulation of migration and proliferation of retinal pigment epithelial cells in vitro. *Invest Ophthalmol Vis Sci* **30**, 1951-1957.

Kornberg, A. (1992). DNA polymerase I of *E. coli* in DNA replication. In *DNA replication* 2nd edit. (Kornberg, A. & Baker, T., eds.), pp. 101-164. E. Freeman and Co., San Francisco.

Korte, G., Repucci, V. & Henkind, P. (1984). RPE destruction causes choriocapillary atrophy. *Invest Ophthalmol Vis Sci* **25**, 1135-1145.

Kruse, T., Reiber, H. & Neuhoff, V. (1985). Amino-acid transport across the human-blood CSF barrier - an evaluation graph for amino-acid concentrations in cerebrospinal-fluid. *J Neurol Sci* **70**, 129-138.

Lai, C.-M., DiGrandi, S., Baines, M., Rakoczy, P. & Constable, I. (1997). *ARVO, Ft Lauderdale, USA*.

LaVail, M. M. (1976). Rod outer segment disk shedding in the rat retina. Relationship to cyclic lighting. *Science* **194**, 1071-1074.

Lennon, G., Auffray, C., Polymeropoulos, M. & Soares, M. (1996). The I.M.A.G.E. Consortium: An integrated molecular analysis of genomes and their expression. *Genomics* **33**, 151-152.

Li, L. & Turner, J. (1988). Inherited retinal dystrophy in the RCS rat: prevention of photoreceptor degeneration by pigment epithelial cell transplantation. *Exp Eye Res* **47**, 911-916.

Liew, C., Hwang, D., Fung, Y., Laurenssen, C., Cukerman, E., Tsui, S. & Lee, C. (1994). A catalog of genes in the cardiovascular-system as identified by expressed sequence tags. *Proc Natl Acad Sci USA* **91**, 10645-10649.

Lindahl, P., Abrahamson, M. & Bjork, I. (1992). Interaction of recombinant human cystatin C with the cysteine proteinases papain and actinidin. *Biochem J* **281**, 49-55.

Lofberg, H., Grubb, A., Sveger, T. & Olsson, J. (1980). The cerebrospinal fluid and plasma concentrations of gamma trace and beta2-microglobulin at various ages and in neurological disorders. *J Neurol* **223**, 159-168.

Marks, N., Berg, M. & Benuck, M. (1986). Preferential action of rat brain cathepsin B as a peptidyl dipeptidase converting pro-opioid oligopeptides. *Arch Biochem Biophys* **249**, 489-499.

Marmor, M. (1990). Control of subretinal fluid: experimental and clinical studies. *Eye* **4**, 340-344.

Marshall, J. (1970). Acid phosphatase activity in the retinal pigment epithelium. *Vis Res* **10**, 821-824.

Martini, B., Pandey, R., Ogden, T. & Ryan, S. (1992). Cultures of human retinal pigment epithelium. Modulation of extracellular matrix. *Invest Ophthalmol Vis Sci* **33**, 516-521.

Matsumoto, B. & Besharse, J. (1985). Light and temperature-modulated staining of the rod outer segment disc tips with lucifer yellow. *Invest Ophthalmol Vis Sci* **26**, 628-635.

Maxam, A. & Gilbert, W. (1977). A new method for sequencing DNA. *Proc Natl Acad Sci USA* **74**, 560-564.

Mayerson, P. & Hall, M. (1986). Rat retinal pigment epithelial cells show specificity of phagocytosis in vitro. *J Cell Biol* **103**, 299-308.

Mazure, A. & Grierson, I. (1992). In vitro studies of the contractility of cell types involved in proliferative vitreoretinopathy. *Invest Ophthalmol Vis Sci* **33**, 3407-3416.

McCombie, W., Adams, M., Kelley, J., Fitzgerald, M., Utterback, T., Khan, M., Dubnick, A., Kerlavage, J., Venter, J. & Fields, C. (1992). Caenorhabditis elegans expressed sequence tags identify gene families and potential disease gene homologues. *Nature Genet* **1**, 124-131.

McKechnie, N. M., Boulton, M., Robey, H. L., Savage, F. J. & Grierson, I. (1988). The cytoskeletal elements of human retinal pigment epithelium: in vitro and in vivo. *J Cell Sci* **91**, 303-312.

McNaughton, P. (1990). Light response of vertebrate photoreceptors. *Physiol Rev* **70**, 847-883.

Miller, J. H. (1972). *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Miller, S., Hughes, B. & Machen, T. (1982). Fluid transport across retinal pigment epithelium is inhibited by cyclic AMP. *Proc Natl Acad Sci USA* **79**, 2111-2115.

Miller, S. & Steinberg, R. (1976). Transport of taurine, L-methionine and 3-O- methyl-D-glucose across frog retinal pigment epithelium. *Exp Eye Res* **23**, 177-189.

Miller, S. & Steinberg, R. (1982). Potassium transport across the frog retinal pigment epithelium. *J Membr Biol* **67**, 199-209.

Milner, R. & Sutcliffe, J. (1983). Gene expression in rat brain. *Nucleic Acids Res* **11**, 5497-5520.

Mund, M. L. & Rodrigues, M. M. (1979). Embryology of the Human Retinal Pigment Epithelium. In *The Retinal Pigment Epithelium* (Zinn, K. M. & Marmor, M. F., eds.), pp. 45-52. Harvard University Press, Cambridge, Mass.

Nathans, J. & Hogness, D. S. (1984). Isolation and nucleotide sequence of the gene encoding human rhodopsin. *Proc Natl Acad Sci USA* **81**, 4851-4855.

Nathans, J., Thomas, D. & Hogness, D. S. (1986). Molecular genetics of human color vision: the genes encoding blue, green, and red pigments. *Science* **232**, 193-202.

Newman, T. (1994). Genes galore: A summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDna clones. *Plant Physiol* **106**, 1241-1255.

Nishida, K., Adachi, W., Shimizu-Matsumoto, A., Kinoshita, S., Mizuno, K., Matsubara, K. & Okubo, K. (1996). A gene expression profile of human corneal epithelium and the isolation of human keratin 12 cDNA. *Invest Ophthalmol Vis Sci* **37**, 1800-1809.

O'Day, W. & Young, R. (1978). Rhythmic daily shedding of outer-segment membranes by visual cells in the goldfish. *J Cell Biol* **76**, 593-604.

Okajima, T., Pepperberg, D., Ripps, H., Wiggert, B. & Chader, G. (1989). Interphotoreceptor retinoid-binding protein: Role in delivery of retinol to the pigment epithelium. *Exp Eye Res* **49**, 629-644.

Okubo, K., Hori, N., Matoba, R., Niiyama, T., Fukushima, A., Kojima, Y. & Matsubara, K. (1992). Large scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression. *Nature Genet* **2**, 173-179.

Okubo, K., Itoh, K., Fukushima, A., Yoshii, J. & Matsubara, K. (1995). Monitoring cell physiology by expression profiles and discovering cell type-specific genes by compiled expression profiles. *Genomics* **30**, 178-186.

Olson, M., Hood, L., Cantor, C. & Botstein, D. (1989). A common language for physical mapping of the human genome. *Science* **245**, 1434-1435.

Ostwald, T. & Steinberg, R. (1980). Localization of frog retinal pigment epithelium Na,K-ATPase. *Exp Eye Res* **31**, 351-360.

Ottonello, S., Petrucco, S. & Maraini, G. (1987). Vitamin A uptake from retinol-binding protein in a cell-free system from pigment epithelium cells of bovine retina. Retinol transfer from plasma retinol-binding protein to cytoplasmic retinol-binding protein with retinyl-ester formation as an intermediate step. *J Biol Chem* **262**, 3975-3981.

Pautler, E. & Tengerdy, C. (1986). Transport of acidic amino-acids by the bovine pigment-epithelium. *Exp Eye Res* **43**, 207-214.

Pearson, W. & Lipman, D. (1988). Improved tools for biological sequence comparison. *Proc NAtl Acad Sci USA* **85**, 2444-2448.

Philp, N. & Nachmias, V. (1981). Phagocytosis by retinal pigment epithelium explants in culture. *Exp Eye Res* **33**, 47-53.

Pino, R., Essner, E. & Pino, L. (1982). Location and chemincal composition of anionic sites in Bruch's membrane of the rat. *J Histochem Cytochem* **30**, 245-252.

Politi, L., Lee, L., Wiggert, B., Chader, G. & Adler, R. (1989). Synthesis and secretion of interphotoreceptor retinoid-binding protein (IRBP) by isolated normal and rd mouse retinal photoreceptor neurons in culture. *J Cell Physiol* **141**, 682-690.

Prunte, C. & Kain, H. (1995). Enzymatic digestion increases permeability of the outer blood-retinal barrier for high-molecular-weight substances. *Graefes Arch Clin Exp Ophthalmol* **233**, 101-111.

Quinn, R. & Miller, S. (1992). Ion transport mechanisms in native human retinal pigment epithelium. *Invest Ophthalmol Vis Sci* **33**, 3513-3527.

Rakoczy, P., Baines, M., Kennedy, C. & Constable, I. (1996). Correlation between autofluorescent debris accumulation and the presence of partially processed forms of cathepsin D in cultured retinal piment epithelial cells challenged with rod outer segments. *Exp Eye Res* **63**, 159-167.

Rakoczy, P., Lai, C., Baines, M., Di Grandi, S., Fitton, J. & Constable, I. (1997a). Modulation of cathepsin D activity in retinal pigment epithelial cells. *Biochem J* **324**, 935-940.

Rakoczy, P., Lai, M., Baines, M. & Constable, I. (1997b). *ARVO, Ft. Lauderdale, USA*.

Rakoczy, P., Mann, K., Cavaney, D., Robertson, T., Papadimitreou, J. & Constable, I. (1994). Detection and possible functions of a cysteine protease involved in digestion of rod outer segments by retinal pigment epithelial cells. *Invest Ophthalmol Vis Sci* **35**, 4100-4108.

Raymond, S. M. & Jackson, I. J. (1995). The retinal pigmented epithelium is required for development and maintenace of the mouse neural retina. *Curr Biol* **5**, 1286-1295.

Redmond, T., Yu, S. & Pfeifer, K. (1997). An RPE65 knockout: Targeted disruption of the mouse RPE65. *Invest Ophthalmol Vis Sci* **38**, S699.

Regan, C., de Grip, W., Daemen, F. & Bonting, S. (1980). Degradadtion of rhodopsin by a lysosomal fraction of retinal pigment epithelium: biochemical aspects of the visual process. *Exp Eye Res* **30**, 183-191.

Rizzolo, L. & Heiges, M. (1991). The polarity of the retinal pigment epithelium is developmentally regulated. *Exp Eye Res* **53**, 549-553.

Roberts, R., Mathialagan, N., Duffy, J. & Smith, G. (1995). Regulation and regulatory role of proteinase inhibitors. *Critical Reviews in Eukaryiotic Gene Expression* **5**, 385-436.

Russell, P. & Epstein, D. (1992). Protein analysis of monkey aqueous humor. *Curr Eye Res* **11**, 1239-1246.

Saari, J. & Bredberg, L. (1982). Enzymatic reduction of 11-*cis* retinol bound to cellular retinal-binding protein. *Biochem Biophys Acta* **715**, 266-272.

Saari, J., Bredberg, L. & GG, G. (1982). Identification of the endogenous retinoids associated with three cellular retinoid-binding proteins from bovine retina and retinal pigment epithelium. *J Biol Chem* **257**, 13329-13333.

Sambrook, J., Fritsch, E. & Maniatis, T. (1989). *Molecular Cloning. A Laboratory Manual.* 2nd edit, 1. 3 vols, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sameshima, M., Uehara, E. & Ohba, N. (1987). Specialization of the interphotoreceptor matrices around cone and rod photoreceptor cells in the monkey retina, as revealed by lectin cytochemistry. *Exp Eye Res* **45**, 751-762.

Sanger, F., Nicklen, S. & Coulson, A. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* **74**, 5463-5467.

Sarks, S. (1980). Drusen and their relationship to senile macular degeneration. *Aust J Ophthalmol* **8**, 117-130.

Sasaki, T., Song, J., KogaBan, Y., Matsui, E., Fang, F., Higo, H., Nagasaki, H., Hori, M., Miya, M., Murayama-Kayano, E., Takiguchi, T., Takasuga, A., Niki, T., Ishimaru, K., Ikeda, H., Yamamoto, Y., Mukai, Y., Ohta, I., Miyadera, N., Havukkala, I. & Minobe, Y. (1994). Toward cataloguing all rice genes: Large-scale sequencing of randomly chosen rice cDNAs from a callus cDNA library. *Plant J* **6**, 615-624.

Schoenfeld, T., Mendez, J., Storts, D., Portman, E., Patterson, B., Frederiksen, J. & Smith, C. (1997). Effects of bacterial strains carrying the *endA1* genotype on DNA quality isolated with Wizard plasmid purification systems. *Promega Notes* **53**, 12-20.

Sellner, P. (1986). The movement of organic solutes between the retina and pigment epithelium. *Exp Eye Res* **43**, 631-639.

Seyfried-Williams, R. & McLaughlin, B. (1984). Acid phosphatase localization in normal and dystrophic retinal pigment epithelium. *J Neurocytol* **13**, 201-214.

Shaper, N., Meurer, J., Joziasse, D., Chou, T., Smith, E., Schnaar, R. & Shaper, J. (1997). The chicken genome contains two functional nonallelic beta 1,4-galactosyltransferase genes - Chromosomal assignment to syntenic regions tracks fate of the two gene lineages in the human genome. *J Biol Chem* **272**, 31389-31399.

Shimizu-Matsumoto, A., Adach, W., Mizuno, K., J, I., Nishida, K., Kinoshita, S., Matsubara, K. & K, O. (1997). An expression profile of genes in human retina and isolation of a complemetary DNA for a novel rod photoreceptor protein. *Invets Ophthalmol Vis Sci* **38**, 2576-2585.

Shiono, T., Hayasaka, S. & Mizuno, K. (1983). Effect on temperature and pH on release of enzymes from lysosomes of the bovine retinal pigment epithelium in vitro. *Exp Eye Res* **36**, 871-876.

Sloane, B. & Honn, K. (1984). Cysteine proteinases and metastasis. *Cancer Metastasis Rev* **3**, 249-263.

Smith, L., Sanders, J., Kaiser, R., Hughes, P., Dodd, C., Connell, C., Heiner, C., Kent, S. & Hood, L. (1986). Fluorescence detection in automated DNA sequence analysis. *Nature* **321**, 674-678.

Smith-Thomas, L., Richardson, P., Parsons, M., Rennie, I., Benson, M. & MacNeil, S. (1996). Additive effects of extra cellular matrix proteins and platelet derived mitogens on human retinal pigment epithelial cell proliferation and contraction. *Curr Eye Res* **15**, 739-748.

Spitznas, M. & Hogan, M. (1970). Outer segments of photoreceptors and the retinal pigment epithelium. *Arch Ophthalmol* **84**, 810-819.

Steinberg, R. & Miller, S. (1979). Transport and Membrane Properties of the Retinal Pigment Epithelium. In *The Retinal Pigment Epithelium* (Zinn, K. & Marmor, M., eds.), pp. 205-225. Harvard University Press, Cambridge, Mass.

Stramm, L., Haskins, M. & Aguirre, G. (1989). Retinal pigment epithelial glycosaminoglycan metabolism: Intracellular vs. extracellular pathways. In vitro studies in normal and diseased cells. *Invest Ophthamol Vis Sci* **30**, 2118-2131.

Strubel, N., Nguyen, M., Kansas, G., Tedder, T. & Bischoff, J. (1993). Isolation and characterization of a bovine cDNA encoding a functional homolog of human P-selectin. *Biochem Biophys Res Comm* **192**, 338-344.

Sudo, K., Chinen, K. & Nakamura, Y. (1994). 2058 expressed sequence tags (ESTs) from a human fetal lung cDNA libarary. *Genomics* **24**, 276-279.

Sutcliffe, J., Milner, R., Shinnick, T. & Bloom, F. (1983). Identifying the protein products of brain specific genes with antibodies to chemically synthesized peptides. *Cell* **33**, 671-682.

Swaroop, A., Xu, J., Agarwal, N. & Weissman, S. (1991). A simple and efficient cDNA libary subtraction procedure: isolation of human retina-specific cDNA clones. *Nucleic Acids Res* **19**, 1954.

Tabor, S. & Richardson, C. (1987). DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc Natl Acad Sci USA* **84**, 4767-4771.

Tabor, S. & Richardson, C. (1995). A single residue in DNA polymerases of the *Escherichia coli* DNA polymerase I family is critical for distinguishing between deoxyribonucleotides and dideoxyribonucleotides. *Proc Natl Acad Sci USA* **92**, 6339-6343.

Takeda, J., Yano, H., Eng, S., Zeng, Y. & Bell, G. (1993). A molecular inventory of human pancreatic islets: sequence analysis of 1,000 cDNA clones. *Human Molec Genet* **2**, 1793-1798.

Tarnowski, B. & McLaughlin, B. (1988). Phagocytic interactions of sialated glycoproteins, sugar, and lectin-coated beads with rat retinal pigment epithelium. *Curr Eye Res* **6**, 1079-1089.

Taugner, R., Buhrle, C., Nobiling, R. & Kirschke, H. (1985). Coexistence of renin and cathepsin B in epithelioid cell secretory granules. *Histochemistry* **88**, 102-108.

Tsung, P., Lombardini, J. & Holly, F. (1984). Intracellular distribution and properties of cathepsin B in the rat retina and its inhibition by a cytosolic fraction. *Exp Eye Res* **38**, 73-79.

Tsuruta, J., O'Brien, D. & Griswold, M. (1993). Sertoli cell and germ cell cystatin C: stage-dependent expression of two distinct mesenger ribonucleic acid transcrits in rat testes. *Biol Reprod* **49**, 1045-1054.

Tu, G.-F., Aldred, A., Southwell, B. & Schreiber, G. (1992). Strong conservation of the expression of cystatin C gene in choroid plexus. *Am J Physiol* **263**, R195.

Turk, V. & Bode, W. (1991). The cystatins: protein inhibitors of cysteine proteinases. *FEBS Lett* **285**, 213-219.

Twining, S. (1994). Regulation of proteolytic activity in tissues. *Current Rev Biochem Mol Biol* **29**, 315-383.

Weinberg, R. A. (1985). The molecules of life. *Scientific American* **253** (4), 34-43.

Wiedemann, P. & Weller, M. (1988). The pathophysiology of proliferative vitreoretinopathy. *Acta Ophtahlmol* **189 (suppl)**, 3-15.

Wilcox, D. (1988). Vectorial accumulation of cathepsin D in retinal pigmented epithelium: effects of age. *Invest Ophthalmol Vis Sci* **29**, 1205-1212.

Williams, D. & Fisher, S. (1987). Prevention of rod disk shedding by detachment from the retinal pigment epithelium. *Invest Ophthalmol Vis Sci* **28**, 184-187.

Womack, J. & Kata, S. (1995). Bovine genome mapping - evolutionary inference and the power of comparative genomics. *Curr Opinion Gen & Development* **5**, 725-733.

Wood, W. I., Gitschier, J., Lasky, L. A. & Lawn, R. M. (1985). Base composition-independent hybridization in tetramethylammonium chloride: A method for oligonucleotide screening of highly complex gene libraries. *Proc Natl Acad Sci USA* **82**, 1585-1588.

Woodcock, D. M., Crowther, P. J., Doherty, J., Jefferson, S., Decruz, E., Noyer-Weidner, M., Smith, S. S., Michael, M. Z. & Graham, M. W. (1989). Quantitative evaluation of Escherichia coli host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucleic Acids Res.* **17**, 3469-3478.

Young, R. (1971). Shedding of discs from rod outer segments in the rhesus monkey. *J Ultrastruct Res* **34**, 190-203.

Young, R. (1977a). The daily rhythm of shedding and degradation of cone outer segment membranes in the lizard retina. *J Ultrastruct* **61**, 172-185.

Young, R. (1977b). The daily rhythm of shedding and degradation of rod and cone outer segment membranes in the chick retina. *Invest Ophthalmol Vis Sci* **17**, 105-116.

Young, R. (1987). Pathophysiology of age-related macular degeneration. *Surv Ophthalmol* **31**, 291-306.

Young, R. W. (1967). The renewal of photoreceptor cell outer segments. *J Cell Biol* **33**, 61-72.

Young, R. W. & Bok, D. (1969). Participation of the retinal pigment epithelium in the rod outer segmet renewal process. *J Cell Biol* **42**, 392-403.

Young, R. W. & Droz, B. (1968). The renewal of protein in retinal rods and cones. *J Cell Biol* **39**, 169-184.

Zauberman, H. (1979). Adhesive Forces between the Retinal Pigment Epithelium and Sensory Retina. In *The Retinal Pigment Epithelium* (Zinn, K. M. & Marmor, M. F., eds.), pp. 192-204. Harvard University Press, Cambridge, Mass.

Zimmerman, W., Godchaux, W. & Belkin, M. (1983). The relative proportions of lysosomal enzyme activities in bovine retinal pigment epithelium. *Exp Eye Res* **36**, 151-158.

Zinn, K. M. & Benjamin-Henkind, J. V. (1979). Anatomy of the Human Retinal Pigment Epithelium. In *The Retinal Pigment Epithelium* (Zinn, K. M. & Marmor, M. F., eds.), pp. 3-31. Harvard University Press, Cambridge, Mass.

APPENDIX I

2.2.2. PREPARATION OF REAGENTS

8% Acrylamide/Urea/TBE stock solution

For 200ml:

20ml 10xTBE

53ml 30% (w/v) acrylamide/bis-acrylamide

(ratio 37.5:1) (Severn Biotech Ltd.)

100g urea (AnalaR, BDH)

30ml distilled H_2O

Alkaline electrophoresis buffer (10x)

500mM NaOH

10mM Na_2EDTA

Alkaline loading buffer (5x)

50% sucrose

0.4% bromophenol blue

1x alkaline electrophoresis buffer

100x Denhardt's solution

2% (w/v) BSA (bovine serum albumin)

2% (w/v) Ficoll™

2% (w/v) PVP (polyvinylpyrrolidone)

Formaldehyde gel loading buffer

900µl formamide (deionized)

180µl 10x MOPS buffer

120µl 37% formaldehyde

1mg xylene cyanol

1mg bromophenol blue

IPTG (isopropyl β-D-thiogalactopyranoside) stock solution

0.1M (23mg/ml) in H_2O

Filter sterilize and store at 4°C.

X-Gal (5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside)

20mg/ml in dimethylformamide (DMF)

Store at -20°C.

10x MOPS buffer

200mM 3-(N-morpholino)propane-
sulphonic acid (MOPS)

50mM CH_3COONa

10mM EDTA

Adjust the pH to 6.5-7.0 with NaOH

Phenol/chloroform (1:1 v/v)

Equilibrate the phenol with TE buffer of pH 7.8.

Add a volume of chloroform equal to that of the
TE saturated phenol.

Mix and allow the two phases to separate.

Use the lower phenol/chloroform phase.

Store in dark bottles at 4°C.

10xPIPES buffer

4M NaCl

0.1M 1,4-piperazine-diethanesulphonic acid (PIPES)

pH 6.5

Salmon sperm DNA (10mg/ml)

(non-homologous DNA, for blocking non-specific hybridization)

Shear salmon sperm DNA (Sigma) by passing it through an 18-gauge needle at least 5 times. Then boil it for 10 minutes and quickly chill on ice. Store at -20°C.

SM buffer

0.1M NaCl

0.01M $MgSO_4$

0.03M Tris base

0.01% Gelatin

Adjust the pH to 7.5 with HCl.

Autoclave and store at 4°C.

20x SSC

3M NaCl

0.3M Na_3citrate

Adjust the pH to 7.0 with 5M NaOH

10xTBE buffer

0.89M Tris base

0.89M boric acid

0.025M EDTA

pH8.3

TE buffer

10mM TRIS-HCl (pH 7.5 - 8.0)

1mM EDTA

pH 7.5 - 8.0